A DEFENSE OF EQUILIBRIUM REASONING IN ECONOMICS

by

Jennifer Soyun Jhun

BA in Philosophy and Economics, Northwestern University, 2008

Submitted to the Graduate Faculty of

The Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2016

UNIVERSITY OF PITTSBURGH

KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Jennifer Soyun Jhun

It was defended on

May 27, 2016

and approved by

Robert Batterman, Professor of Philosophy, University of Pittsburgh

Sheldon Smith, Professor of Philosophy, University of California, Los Angeles

Dissertation Advisor: Mark Wilson, Distinguished Professor of Philosophy, University of Pittsburgh

Dissertation Advisor: James Woodward, Distinguished Professor of History and Philosophy of Science, University of Pittsburgh

A DEFENSE OF EQUILIBRIUM REASONING IN ECONOMICS

Jennifer Jhun, PhD

University of Pittsburgh, 2016

A DEFENSE OF EQUILIBRIUM REASONING IN ECONOMICS

Jennifer Soyun Jhun, PhD

University of Pittsburgh, 2016

Critics both within and outside of philosophy have challenged economics wholesale as unscientific. In particular, economics seems unable to predict future events because it relies on assumptions like equilibrium conditions, which stipulate that the economy tends to stay in its current state absent external forces. The popular background view that gives rise to this criticism is that the job of science is to uncover laws of nature, by appeal to which we can determine (usually deductively) the future behavior of a dynamical system as it evolves. I argue that lawlike statements in economics have a very different role than this: they provide a means of understanding in terms of how efficient a particular system is. This account is more faithful to the history and the practice of economics. Perhaps surprisingly, it also accounts better for the explanatory power of some laws of physics.

By reinterpreting *ceteris paribus* assumptions, and equilibrium assumptions more generally, as tools both for articulating constraints on a system as well as for identifying opportunities for a

particular system, I am able to take into consideration the ways in which both engineers and policy makers aim to design or test such complex systems for *stability*. Macroscopic properties such as stability cannot be reduced to details about the individual atoms that make up bulk material or the individual agents that make up the economy. Yet, the behaviors of the micro-constituents and of the macroscopic aggregate are related. In order to address this lacuna between the micro and the macro, I explore the possibility of exploiting newer methods in the material sciences, namely *multi-scale modeling*, that have been useful in talking about these interesting and desirable macroscopic properties. These methods use not one but multiple models at different temporal and spatial scales to describe a system, without prioritizing any particular one. Given the substantial methodological and formal analogies between thermodynamics and economics, the success of the multi-scale framework in the former suggests it will be similarly useful for the latter.

TABLE OF CONTENTS

PREFACE IX

1.0 INTRODUCTION: SOME PERCEIVED PROBLEMS OF ECONOMICS 1

2.0 WHAT'S THE POINT OF *CETERIS PARIBUS?* (OR: HOW TO UNDERSTAND SUPPLY AND DEMAND CURVES) 8

2.1 CARNOT CYCLES AND SUPPLY-DEMAND CURVES 13

2.2 A LESSON FROM THE HISTORY OF ECONOMIC THOUGHT: ALFRED MARSHALL ON *CETERIS PARIBUS* PROVISOS 21

2.3 SCIENTIFIC UNDERSTANDING IN TERMS OF EFFICIENCY 28

2.4 AGAINST ORTHODOXY: REJECTING THE SEMANTIC COMPLETER ACCOUNT 33

2.5 CONCLUDING REMARKS 43

3.0 DYNAMIC AND DISEQUILIBRIUM BEHAVIOR 45

3.1 PRELIMINARIES 45

3.2 DIAGNOSTIC REMARKS 47

3.3 A SECOND LESSON FROM THERMODYNAMICS: FALSE EQUILIBRIA AND FINANCIAL CRISES 54

3.4 THE 2007-2010 FINANCIAL CRISIS 59

3.5 CONCLUDING REMARKS 65

4.0 MODELING STABILITY ... 67

4.1 STABILITY, EQUILIBRIA, AND AGGREGATION ... 68

4.2 OBSTACLES TO EQUILIBRIUM BEHAVIOR ... 73

4.3 CORRECTIVE MEASURES AND CONTROL THEORY ... 81

4.4 MEANWHILE, IN ECONOMICS ... 83

4.4.1 DSGE Models ... 83

4.4.2 Heterogeneity and Meso-Structure ... 88

4.4.3 Policy Design as an Optimal Control Problem ... 94

5.0 CONCLUDING REMARKS: WHAT DOES THE LUCAS CRITIQUE CRITIQUE? ... 99

5.1 ECONOMICS VS. ECONOMETRICS, STRUCTURAL VS. REDUCED-FORM ... 102

5.2 MULTISCALE MODELS TO THE RESCUE? ... 108

BIBLIOGRAPHY ... 110

LIST OF FIGURES

Figure 1. Supply and demand equilibrium. 10

Figure 2. Post-supply shock. 11

Figure 3. A Carnot Engine. 15

Figure 4. A Pressure-Volume Diagram for a Carnot Engine. 16

Figure 5. Deadweight loss. 20

Figure 6. A Perfect Crystal Lattice. 50

Figure 7. A Crystal Lattice with Dislocations. 50

Figure 8. Before and After: A Crystal Lattice Subjected to an External Force. 51

PREFACE

I am especially thankful to the following individuals.

For their patient, encouraging, and insightful mentorship: Robert Batterman, Marciano Siniscalchi, Sheldon Smith, Mark Wilson, and James Woodward.

For minding the state of my wellbeing: Agnes Bolinska, Kathleen Cook, Kathryn Lindeman, Tyke Nunez, Jasmin Özel, Meghan Page, Patricia Palacios, Benjamin Schulz, and Kathryn Tabb.

For comic relief: Ray Jhun.

For reasons to look forward: Rade Radjenovich.

1.0 INTRODUCTION: SOME PERCEIVED PROBLEMS OF ECONOMICS

A common criticism (even from economists) is that economics is not useful – nor, even more surprisingly, does it even aim to be. The culprit is an over-reliance on and misuse of "idealization": too much formalization without justification has made economics a field that seems unconcerned with giving a true or accurate representation of the real world. Most of economic analysis is centered on the *equilibrium* concept. The state of equilibrium is generally described as the state from which the market players have no incentive to deviate, or from which the economy has to tendency to change once reached given that there are no outside influences, the state towards which the economy will gravitate towards...All are descriptions of the state where "economic forces are in balance."

The equilibrium concept in economics is tied to other concepts, some them normative, e.g. *rationality* and *efficiency*. For example, in competitive markets, should a supplier or a buyer stray from the equilibrium state, they could do *better* by returning to equilibrium. In the case of competitive markets, this means that one couldn't do better without making someone else worse off. The equilibrium problem is thus a problem of optimization under constraints. What it is better ("rational") to do is to behave in the relevant optimizing fashion. The rational economic state is the state of equilibrium where resources are allocated efficiently.

It shouldn't surprise us that what happens in the world fails to align with our theoretical calculations. To criticize economics simply on that basis is to react to an oversimplified picture of economics, and in fact of science in general. I argue that this critical view only results from a commitment to a narrow conception of what *scientific laws* and should look like and function. This usually stems from an implicit preference for the Hempelian deductive-nomological picture. In this vein, explanations look like arguments, the conclusion of which is a description of the target phenomena. Amongst the premises, there is at least one "law of nature." According to this picture, the explanatory problem we face boils down to a dynamic initial value problem. Once we specify these initial (and boundary) conditions, then with the help of these laws we can deduce the behavior of the system over time. This view of what counts as scientific enquiry muddies what it is that economic methodology is actually doing by misinterpreting the purpose of idealizations used in model construction and ignoring the significant of both their historical and current development.

These prejudices often misconstrue how economists actually motivate and use idealizing strategies when it comes to modeling. Chapter 2 considers economics' use of *ceteris paribus* – "all other things equal" – assumptions, which occur explicitly and are systematically deployed in statements considered to be laws of economics. The textbook supply-demand model is a clear demonstration of how simplifying *ceteris paribus* assumptions work. When supply of a product equals demand at a particular quantity and price, we say that the market is at equilibrium. An external force – such as a natural disaster or trade policy – might disrupt this equilibrium state, shifting either or both the demand and supply of the market. The market should then shift to a new equilibrium quantity and price. We can summarize this line of reasoning with one of the laws of supply and demand, e.g. "A downward shift in supply will, *ceteris paribus,* lead to an increase in price." In order to characterize the equilibrium state of a single market, we deploy *ceteris paribus* assumptions that allow us to isolate the particular market we're interested in from economic effects outside of that market, even though this is rarely ever the case.[1]

Complaints about the legitimacy of such laws and the over-idealization involved to formulated them are nearly always targeting equilibrium analysis. We know, of course, that external forces will affect markets other than the one we are primarily interested in, that there are feedback effects, and so on. But *ceteris paribus* clauses allow us to temporarily ignore these other complicating effects so that we may characterize the causal relationship between, say, the price of a certain commodity and shift in supply that has been brought about by an external disruptive event. However, there is a popular conception that *ceteris paribus* clauses are actually meant to flag the incomplete nature of economic laws. This interpretation maintains that the *ceteris paribus* clause is a placeholder for a list of other conditions that are assumed to hold. That is, "If A, then *ceteris paribus*, B" is an abbreviation for: "If A and (blank) and (blank)...then B." The *ceteris paribus* clause is a placeholder for this list of other conditions that are assumed to hold so that properly filled in, the sentence turns out to be a true description.

This, however, is the incorrect way to understand the role of *ceteris paribus* clauses. Once we clear away this naive conception, we clear away a secondary misconception as well: that the "de-idealizing" process simply consists in incorporating the missing information we initially omitted. But furthermore, this reveals that criticisms of equilibrium theory are based on a misconception of what we expect scientific laws to do. This misconception, it turns out, grounds not only many philosophers' appraisal of the status of the social sciences but the

---

[1] My focus will thus be mainly on the competitive partial equilibrium concept, where the equilibrium state is reached when the supply of commodities equals consumers' demands for those goods. Partial equilibrium analysis focuses on a single market as opposed to multiple interrelated segments of the economy. This contrasts with general equilibrium analysis, whose gaze is the whole of the economy.

physical sciences, too. For example, thermodynamics – a paradigm case of a successful branch of science as any – deals with everyday materials. Yet, positing quasi-static processes induced by an external controlled force driving a system from one equilibrium state to another is a fundamental feature of thermodynamics. This contrasts with the idea of a theory as issuing dynamical pictures that can track, via derivation, the evolution of a system as it unfolds autonomously.

I argue instead that the invocation of *ceteris paribus* is a methodology for "uncovering dominant behaviors" (Wilson forthcoming). In fact, as an examination of Alfred Marshall's (1890) work will show, the use of *ceteris paribus* has been motivated this way ever since the beginning of modern economic thought. Now, what I already take as obvious, and that I am not arguing for, is that adequate analysis implicitly needs consideration of characteristic relaxation times given external forces. No equilibrium model tells us what the relaxation time of a process will be. Rather, uncovering such dominant behaviors, e.g. salient causal relationships at particular scales of the system, are aimed at setting benchmarks for the *efficient performance* of systems against which we compare actual performance. Central to this is the notion of *manipulation*, and Jim Woodward's (2003) interventionist account of causation finds solid ground here.

Having clarified the methodological role of *ceteris paribus* (and in general, equilibrium) idealizations in economics, in Chapter 3, I consider the question of how we do deal with the dynamical behavior of economic systems that are *not* at equilibrium. Now that we have justified the usage of equilibrium reasoning, how do we discharge those idealizing *ceteris paribus* assumptions, if not by filling in all the details we initially omitted? How can we take theory and put it to use in real life cases, such as understanding and predicting phenomena like financial crashes?

One source of difficulty when it comes to modeling these extra complications is the bottom-up nature of most microeconomic reasoning. That is, economic models are often built from the case of the individual upwards to groups. In order to get some information about consumers as a whole, it (mistakenly) simply mashes together the information specified about each individual. Distinct, but often held in tandem with the bottom-up attitude, is the reductionist attitude. This is the stance that a more fundamental or legitimate science has laws that are made in terms of the most individual constituents so that the aggregate behavior is implied (usually deductively). These two attitudes limit the ways we can build more expansive economic models and downplays the usefulness and explanatory power of the models already available, including very basic partial equilibrium models as well as more complicated ones such as the dynamic stochastic general equilibrium model that I discuss later in Chapter 4. The usual naive thought is that once we have accumulated all the details at a micro-level, we will be able to infer or deduce information about what goes on at the macro level.

The strategies we found in the simpler thermodynamic cases are useful explanatory tools in more complex scenarios too (e.g. solids rather than ideal gases). Dismissing such modeling strategies as over-idealizing and thus inapplicable is to inappropriately divorce the strategy from the corrective architecture in which it resides. By combining Pierre Duhem's (1910) notion of false equilibrium – a system that is technically not at equilibrium may nonetheless be in some sense stable – with tools from contemporary engineering and material sciences, I suggest that the natural next step is to employ *multi-scale modeling* methods. It's often the case that we wish to assess and induce a higher-order property like stability, but that is information about which we cannot glean from a microscale perspective. For example, the steel we use in constructing buildings will be *elastic*, which is a macro-scale property that arises from the interactions between many heterogeneous micro-constituents which are heterogeneous and may include defects. It is only by examining goings-on at the *meso*-scale, neither the micro- nor the macro-scale, that we can explain this property and bring it about in manufacturing. By way of an example – the recent 2007/2008 subprime mortgage crisis – I show how multi-scale models provide an attractive framework for exploring actual economic behavior while still exploiting equilibrium methods. In fact, in order to arrive at a satisfactory diagnosis of the crisis we have to consider the crisis from multiple scales.

Chapter 4 extends this discussion further, addressing questions regarding how to understand financial stability, the dynamics of crises, and the corrective measures that economists take to ensure stability and keep crises from spiraling out of control. A consideration of well-known theoretical difficulties involving *aggregation* (using the transition from individual consumer to aggregate consumer demand as an example) and *market incompleteness* (where insuring oneself against all cases of risk is impossible) motivates further the multi-scale framework. These are the difficulties that even contemporary models such as dynamic stochastic general equilibrium (DSGE) models struggle with: systems with interacting heterogeneous agents and market incompleteness will exhibit non-linear behavior and give rise to novel features appearing at higher order scales. I propose that corrective measures – institutions such as fiscal or monetary policy – should be interpreted as meso-scale interventions that regulate the adaptive behavior of the economy. Furthering the engineering analogy, we can think of economic policymaking as an exercise in constructing *control theory*. Familiar examples include cruise controls and float valves in toilets; these are tools that help us concentrate the behavior of a system within a desirable range.

Chapter 5 closes with a discussion of the *Lucas Critique*. This critique, which sits at the intersection between economic theory and econometric practice, maintains that we cannot adequately assess policy effects if the very structure of the economy changes with the implementation of policy. Models such as the DSGE can be said to “pass” or “fail” the Critique. I suspect that most these assessments rely on a misunderstanding of the Lucas Critique, and that Lucas’s insight can be aligned with our own modeling concerns.

2.0 WHAT'S THE POINT OF *CETERIS PARIBUS?* (OR: HOW TO UNDERSTAND SUPPLY AND DEMAND CURVES)

> [E]conomists developed a nearly irresistible predilection for deductive reasoning. As a matter of fact, many entered the field after specialising in pure or applied mathematics. Page after page of professional economic journals are filled with mathematical formulas leading the reader from sets of more or less plausible but entirely arbitrary assumptions to precisely stated but irrelevant theoretical conclusions. (Leontief 1982, 107)

From the perspective of philosophy of science, investigative procedures in modern economics often appear quite unsatisfactory. As a result, it is tempting to distinguish the laws of economics from the laws of physics in the following way. The laws of physics are universal and exceptionless; they are exemplars of scientific achievement that other disciplines should aspire to. Not only do the laws of economics admit of exceptions, they apply only under extremely unrealistic assumptions and appeal to "equilibrium conditions" rarely witnessed in real life. We hedge economic generalizations with the proviso that *all other things are equal*, or *ceteris paribus,* e.g. "an increase in the price of a commodity X leads, c*eteris paribus*, to a decrease of the quantity demanded." Perhaps one day this generalization can be replaced with an explicit, comprehensive law of consumer demand that is exceptionless. In a barb directed at Eugene Fama and Richard Shiller at a 2013 roundtable of fellow Nobel laureates, chemist Martin Karplus dismissed the science of economics – "if one wants to call it a science" – as an unsatisfactory tool when it came to making sense of the actual behavior of actual markets. Physicist

Richard Feynman (1981) once commented: "Social science is an example of a science which is not a science…They follow the forms…but they don't get any laws…They haven't got anywhere yet…Maybe someday they will." So it is common to interpret *ceteris paribus* laws, as Earman and Roberts (1999) do, as at best signaling "an embryonic theory on its way to being developed to the point where it makes definite claims about the world" (466).

In this essay I argue there is a class of statements hedged by *ceteris paribus* clauses that have a robust explanatory role of their own.[2] This role diverges from what philosophers usually expect of laws of nature, and is not borne out by more sophisticated, mature formulations of the lawlike statement in question. In order to elucidate this alternative role of *ceteris paribus* clauses, I draw upon the way it is used in particular in economics, where it is invoked regularly and in a specific way. Unlike many philosophers, I do not assume that the primary role of laws is to deductively imply a description of the behavior of a system over time. So, the role of *ceteris paribus* laws is not primarily to hedge would-be predictive statements of this kind. Rather, they help us understand a dynamic system in terms of how efficient it is by giving us the tools to explore what the relevant causal relationships that characterize a particular problem of interest. And this is not unique to economics. The theoretical discovery of Carnot's cycle for heat engines, which I will discuss later in this paper in more detail, is a landmark in the history of physics and chemistry. The Carnot cycle traces out the optimal performance of a heat engine that does work and then returns to its original state. While it is impossible to get a steam engine to follow the perfectly controlled path of the cycle, an engineer's appreciation of how well an actual

[2] Unlike some other approaches, my account is not general. The philosophical literature on *ceteris paribus* laws typically looks for an account that applies to all generalizations that have exceptions. I am not concerned about giving an account of *ceteris paribus* laws that would include statements such as "*ceteris paribus*, birds fly" alongside "*ceteris paribus*, aspirin relieves headaches."

engine is performing compared to what is optimally possible is central to her understanding of the machine's workings.

Basic economic constructs are also hypothetical. One might wonder whether they are merely convenient "fairly tales" as game theorist Ariel Rubinstein (2006) once remarked. I think they are not, and no more so than constructs such as the Carnot engine. Consider the classic so-called law of demand: *ceteris paribus,* the price of a commodity and the quantity demanded are inversely related (the blue line in Figure 1). These *ceteris paribus* assumptions allow us to characterize *equilibrium* states for limited markets Figure 1 and talk about the dynamic behavior of markets in terms of *comparative statics* Figure 2.

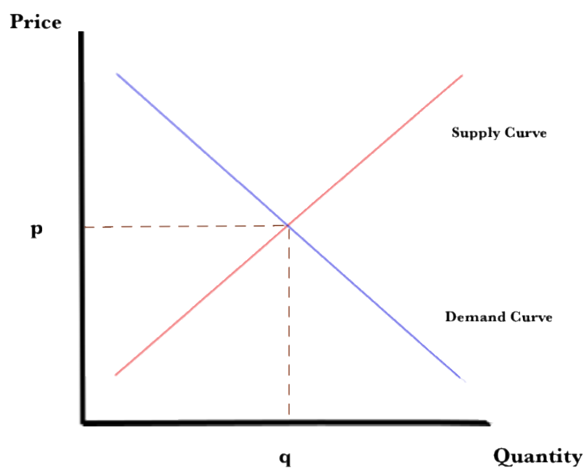

Figure 1. Supply and demand equilibrium.

That we can draw either curve at all in a manner as above requires stringent assumptions. For example, consider the downward sloping consumer demand curve (in blue). This line represents the inverse relationship between price and quantity of a good. In order to establish this relationship, we ignore changes in the population, in consumers' preferences and incomes, and in the prices of other goods that would otherwise prevent us from drawing such a nice linear curve. The equilibrium price and quantity here are **p** and **q**.

Now, suppose that some outside influence affects the supply curve in red, pushing it to where the dotted red line is. The new equilibrium price and quantity are at **p*** and **q***. The price is now higher and the new quantity is lower. If our analysis compared the new equilibrium state with the old one, then we have investigated the pure causal effect of a supply shock, i.e. we have undertaken an exercise in *comparative statics*, illustrated below.

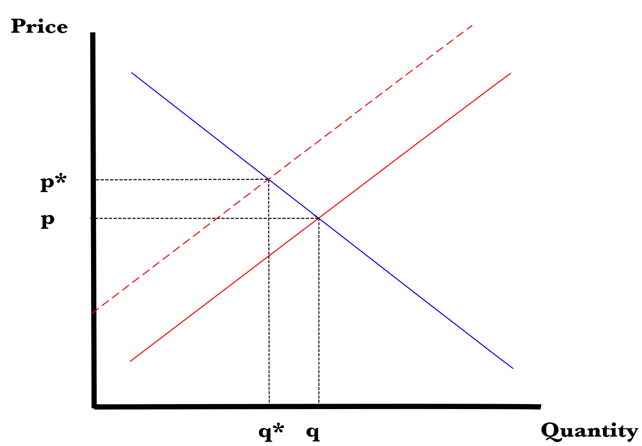

Figure 2. Post-supply shock.

But the surrounding economy is constantly in flux, and a change in price or demand of a single commodity will affect and be affected by other goings-on. Because it's unclear what bearing it has on the real world, and so the equilibrium method has met widespread criticism from methodologists of science. Economics "hasn't helped us improve our predictive abilities," which "suggests it is still far from being a science, and may never be" (Rosenberg and Curtain 2013).

This paper is a reaction against the dismissive attitude, which is grounded in the assumption that the target of analysis must be treated as an autonomously evolving system. There are other ways we can approach phenomena.[3] **Section II** establishes that the use of pure supply and demand curves is analogous to an engineer's use of the pathway drawn by the Carnot cycle. They supply the theoretical base upon which we can do empirical work. I show in **Section III** that the original motivation of *ceteris paribus* reasoning introduced in Alfred Marshall's 1890 *Principles of Economics* aligns with this interpretation and characterizes how we still approach economic problems. The modern worry, as former Treasury Secretary Timothy Geithner (2014) puts it, is what to do since "[f]inancial crises cannot be reliably predicted, so they cannot be reliably prevented." Ultimately, economists wish to understand how a fluctuating economy behaves so that through suitable controlled interventions, its functioning can be improved. Despite the unpredictability of crises, "governments can do a lot to reduce the damage of financial disasters…they can reduce the system's vulnerability of risk, as well as the risk that crises will spiral out of control" (388).

Interpreting *ceteris paribus* clauses as standing in for information we haven't uncovered yet (but which, if we had, we would have *laws*) would be trying to cram the kind of understanding valued by the engineer or the economist into the narrow confines of "predicting the temporal behavior of an autonomously evolving system". But this strikes me as a philosophical mistake. The substantial similarities between economic and thermodynamic reasoning suggest that *ceteris paribus* laws are used to (1) articulate the causal constraints on a

---

[3] For another account of explanation that, while distinct from mine, is similarly motivated, see Batterman (2001). There, he focuses on asymptotic reasoning that involves sacrificing detail and precision. In addition, Batterman and Rice (2014) suggest a minimal models conception of explanation that could also supplement ours.

system and (2) flag loci for possible interventions and manipulations. **Section IV** articulates an account of understanding that accommodates these roles – understanding in terms of *efficient performance* – that builds on recent work by James Woodward (2003) on interventionist theories of causation. I argue in **Section V** that the distinction between economics and other sciences cannot be made on the basis of whether the statements of that science use *ceteris paribus* assumptions. In fact, one common interpretation of the *ceteris paribus* proviso, the *semantic completer interpretation* that underlies views like that of Earman and Roberts, commits the very philosophical mistake that I avoid.

2.1 CARNOT CYCLES AND SUPPLY-DEMAND CURVES

Consider the following remarkable claim from Einstein (1949):

> (Thermodynamics) is the only physical theory…of which I am convinced, that within the framework of applicability of its basic concepts will never be overthrown.

What makes thermodynamics so impressive?[4] Like economics, thermodynamics studies systems where we do not (and often need not) know the exact configuration of its individual constituents. Measuring the temperature of a heated pot of water does not require our knowing all the atomic details of the water, the pot, etc. Despite not being considered a fundamental physical theory,

[4] One reason why this is so surprising is because by the time Einstein said this, both statistical and quantum mechanics (which most assume to be more fundamental) were already established subfields of physics.

thermodynamics is still successfully employed today, notably in engineering contexts.[5]

In the 19th century, Rudolph Clausius attempted to explain what we now call entropy by appealing to the Carnot cycle I described in the introduction. The Carnot cycle consists of four *reversible* processes – processes that could in theory be "reversed" to their initial states with no trace that a change had taken place at all. This implies that at any given time, the system is in thermodynamic equilibrium with its surroundings, which means that it moves infinitely slowly.[6] Imagine a piston cylinder, with ideal gas inside, as depicted below [Illustration C]. At the first stage of the cycle (1) we submerge it in a heat reservoir like an infinite vat of water. The gas expands, pushing the piston up and leading to a decrease in pressure. That is, the engine *does* work. The heat reservoir and the engine remain the same temperature while volume increases. Then, (2) we remove the piston, insulate it so that we lose no heat during transit, and lift the piston as the gas continues to expand. Next, (3) we place the piston cylinder in a cooler reservoir. Then we compress the piston again. Temperature remains the same while pressure increases and volume decreases. The piston cylinder releases heat. Work is done *on* the engine. In the last step (4) we further compress the piston after removing the cylinder and insulating it in order to raise the temperature, returning it to the state we started in before putting it back in the heat bath. (1) and (3) are *isothermal* processes involving no temperature change. (2) and (4) are *adiabatic* processes that involve no heat transfer.

---

[5] There is often no need to appeal to a micro-theory like statistical mechanics when thermodynamics is sufficient and more appropriate in many applied contexts. It's often claimed that statistical mechanics is more fundamental; however, arguments that purport to establish this are problematic, often relying on the fact that the equivalence holds only in the thermodynamic limit as volume goes to infinity. And then there's the further issue as to whether statistical mechanics is reducible to quantum theory. I am not concerned with either proposed reduction here.

[6] This we call *quasi-static*, and while it is actually a distinct concept from reversibility it is a necessary condition for it.

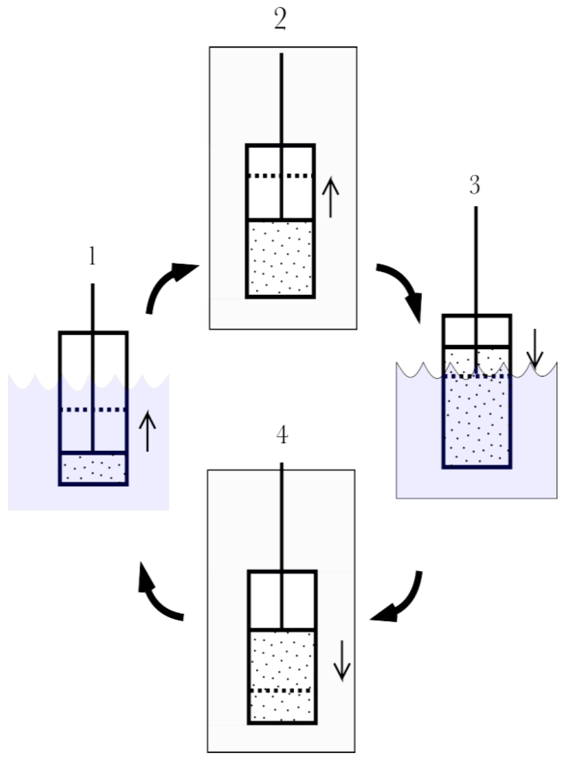

Figure 3. A Carnot Engine.

In a way, the posited stages of a Carnot cycle are all physically absurd and scarcely predictive of any behavior we might ever witness. Gases are not ideal. There are no infinite heat baths. There is no way to perfectly insulate anything. Paradoxically, things are at equilibrium, and yet this system is at the same time dynamic.[7] Why undertake this bit of reasoning at all? Indeed, we can get some empirical information out of this exercise. One consequence is that this Carnot engine represents the upper efficiency of any heat engine, which takes in heat and converts it into mechanical work. This maximal efficiency limit depends solely on the temperature differential between heat reservoirs. No real engine will do better than the Carnot

[7] As the economist Keynes (1924) noted, what good is a model if "in the long run, we're all dead"? After all, "the long run is a misleading guide to current affairs...Economists set themselves too easy, too useless a task if in tempestuous seasons they can only tell us that when the storm is past the ocean is flat again" (80- 82).

engine because real engines produce waste heat due to friction.

Nothing *a priori* tells us what the correct description of a physical system looks like. But by positing a hypothetical external controller that can *in principle* adjust one variable to see how some other variable will vary (holding all the others constant), we can articulate which pairs of variables are causally tied to one another.[8] Each pair consists of one extensive variable and one intensive variable – one variable that represents a force and the other the displacement.[9] Volume and pressure are one such pairing [Figure 4].

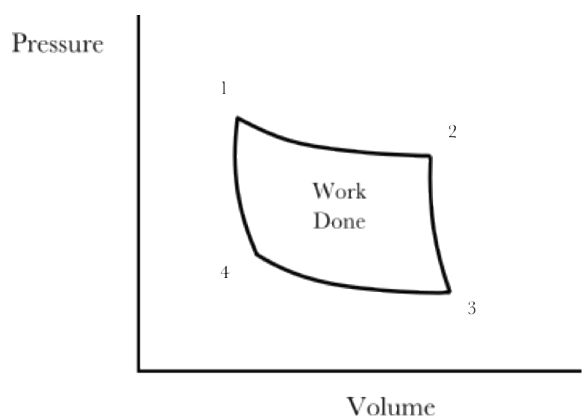

Figure 4. A Pressure-Volume Diagram for a Carnot Engine.

We need such conjugate pairs because they provide the vocabulary to talk about systems and articulate, for instance, the outlines of the Carnot cycle. Toggle the volume in *just such* a way,

[8] These are called *conjugate pairs*.

[9] The displaced quantity is extensive (scales with the size of the system), while the force that does not scale in this matter is intensive. An *extensive* parameter, like Internal Energy (U), Entropy (S), Volume (V), or Number of Particles (N), is a parameter that scales with the system, i.e. U(*a*S,*a*V,*a*N)=*a*U(S,V,N). An *intensive* parameter, like temperature (T), pressure (P), or chemical potential (μ), is a parameter that does not scale with the system (at equilibrium). In other words, T(*a*S,*a*V,*a*N)=T(S,V,N).

and we thereby adjust the pressure by *this* much. The real, empirical behavior of heat engines will play out in less perfect and more inefficient ways, but within these contours.[10] The interior of the figure above represents the most about of work that *any* heat engine (given the temperature differentials of the heat reservoirs) will ever be able to do. They represent important limits and guidelines on what can and does occur in the real world, and so remain crucial in most manufacturing and engineering contexts.

What's crucial to bear in mind here is that the Carnot cycle is one manifestation of the Second Law of Thermodynamics, so the idealizations employed undergird some of the most fundamental Laws of Thermodynamics. Letting S be entropy:

$$S(\text{initial}) \leq S(\text{final})$$

Entropy itself is not a directly measurable quantity. We do, however, talk about *changes* in entropy, standardly defined as follows:

$$dS = dQ/T$$

where Q is heat and T temperature. The equation only holds if we assume that the transitions are reversible. Positing these processes allows us to articulate the entropy concept – now we can express *inefficiency* in terms of entropy. Real engines involve an *increase* in entropy while the Carnot engine does not. So even if it is not in general possible to manufacture processes that are actually at equilibrium, (1) the equilibrium benchmark still provides a useful tool, e.g. in the study of transport phenomena, which are always irreversible and in their limiting cases are approaching equilibria, and (2) it turns out this kind of

[10] For more complicated systems, the number and kinds of conjugate pairs we need for adequate modeling will depend on what kind of work the system is doing. There is also no *a priori* guarantee that all the conjugate pairs will fit nicely into a state equation describing the system as a whole.

hypothetical equilibrium reasoning undergirds the theoretical foundations of sciences such as thermodynamics.

These observations provide a robust analogy to equilibrium analysis in economics. Despite their static nature, those thermodynamic snapshots allow us to investigate dynamical properties. In Figure 3 of the Carnot engine, I represented the engine as moving through various processes by drawing in arrows. These arrows essentially represent a succession of variable driven equilibrium states. Demand and supply curves are exact analogs of these arrows. When economists characterize equilibrium states in order to employ the method of comparative statics, they compare two equilibria to get a handle on how some feature - such as the price of a product – responds to changes in the market (Figure 2). Tracing out a demand or supply curve itself invokes a hypothetical external controller toggling one variable to see how only one other variable will vary in response – e.g. a rise in a commodity's price means that consumers are willing to buy less of it. We envision such processes as perfectly smooth adjustments from one equilibrium state to another. In so doing, we ignore how long it would take to reach the new equilibrium, the fluctuations that route might include, the effects that shift has on outside markets, and the effects that those outside markets have on our target market. In this way we can ensure that the change in target variable is due solely to the external force we subject it to.[11]

This external controller is explicitly invoked in partial equilibrium exercises when we

---

[11] Theoretical (and perhaps actual) exceptions to the law of demand are Giffen and Veblen goods, since their demand increases with increases in price and vice versa, though for different reasons. Giffen goods typically refer to staple commodities – e.g. bread or potatoes – as opposed to Veblen goods, usually luxury goods. In the case of the former, an increases in the price of bread will lead to an increase in its consumption since consumers' income will be diverted towards it and away from other luxury goods (such as meat). For the latter, the high price of, say, the newest technological gadget, makes it all the more appealing to consumers. Diagnosing a good as a Giffen or Veblen one relies, nonetheless, on comparative exercises like the one here.

investigate market disturbances. In Figure 2 above, the supply (say for widgets) moves downward due to a shock.[12] This could be due to external factors, such as the main widget factory burning to the ground. Alternatively, I myself may affect the supply curve by downsizing my factory. Both of these events would move the equilibrium price upward. In the meantime, I have put aside various inconveniences to my analysis. For instance, I assume that the new equilibrium state will be reached in fairly short order, and that the local widget market that I am considering is sufficiently independent of other gadget markets in the vicinity. In reality, the price will take time to adjust. and it is quite possible that the new actual price will be lower than the calculated equilibrium price. That is, the economy also exhibits behavior that we might think of as due to "frictions" that impede its functioning.[13]

Selling the new, lower quantity of product at a higher price is *efficient* in that at this new equilibrium state, nobody can do better without making someone worse off. If I were to sell or price my commodities somewhere other than at equilibrium, I would incur what's called *deadweight loss*, i.e. social welfare could be improved [Figure 5]. Suppose that instead of the equilibrium price and quantity, I decide to sell at the price and quantity indicated at the black dot (a government mandated price ceiling could effect this). The yellow shaded area represents the loss in efficiency, sometimes called allocative inefficiency, which is the cost to society due to our failing to meet equilibrium conditions. Here we have a shortage of widgets. All benefits from exchange would have been realized at the equilibrium state; we fail to do so here. If I am trying to alter the supply curve myself, the new quantity and price

---

[12] Strictly speaking, shocks (in economic jargon) are unpredictable events. For the purposes of my argument, however, I refer to exogenous impacts generally on a system as shocks.

[13] Perhaps unfortunately, neoclassical equilibrium theory often diagnoses any "non-equilibrium behavior as due to frictions, though it goes by other names: rigidity, stickiness, and so on.

represent the maximal extent to which I can affect the economy by those means alone.[14]

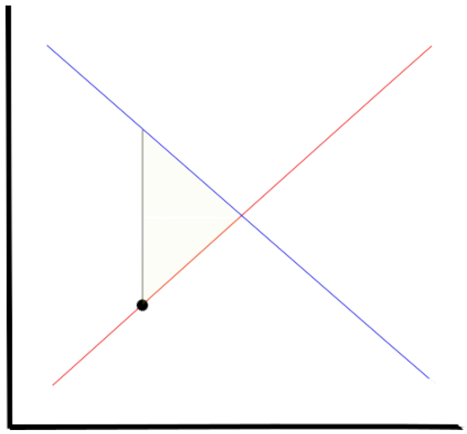

Figure 5. Deadweight loss.

Consider the temperature-entropy diagram [Figure 4] again. Notice that if we are anywhere in the interior of the quadrilateral region traced out by the legs of Carnot cycle, we are not doing as well as we could otherwise be. Similar to deadweight loss, this is a loss in efficiency, and represents the amount by which I could be doing better (i.e. do more work). Because these lines set the benchmark for what would be efficient performance, the supply/demand curves represent the limits that we try to approximate as best as we can.

---

[14] This is called Pareto efficiency. It should be no surprise that the formal manifestation of efficiency, the Pareto frontier, is also used in engineering.

2.2 A LESSON FROM THE HISTORY OF ECONOMIC THOUGHT: ALFRED MARSHALL ON *CETERIS PARIBUS* PROVISOS

I suggest that lawlike generalizations in economics such as "an increase in price leads to a decrease in supply" serve as strategic anchors around which we do empirical work in the way that thermodynamic equations do. So *ceteris paribus* assumptions in economics do much the same thing that positing reversible processes do in thermodynamics. In fact, economic use of *ceteris paribus* has been motivated exactly this way for some time.

While the use of *ceteris paribus* assumptions has been around for centuries, its usage in economics is usually associated with Alfred Marshall and his well-known partial equilibrium method.[15]

> The element of time is a chief cause of those difficulties in economic investigations which make it necessary for man with his limited powers to go step by step; breaking up a complex question, studying one bit at a time, and at last combining his partial solutions into a more or less complete solution of the whole riddle. In breaking it up, he segregates those disturbing causes, whose wanderings happen to be inconvenient, for the time in a pound called Ceteris paribus. **The study of some group of tendencies is isolated by the assumption other things being equal: the existence of other tendencies is not denied, but their disturbing effect is neglected for a time.** The more the issue is thus narrowed, the more exactly can it be handled: but also the less closely does it correspond to real life. Each exact and firm handling of a narrow issue, however, helps towards treating broader issues… (V.V.10, emphasis mine)

His exploitation of the *ceteris paribus* pound, where inconveniences to our analysis are kenneled away for the time being, is quite clearly motivated in part by issues of tractability. But he means something more than just making a difficult problem manageable. For instance, he acknowledges that the effects of causes may unfold at different rates. Thus, depending on what the current

---

[15] It is still introduced to economics students in textbooks in this way.

problem is, some causes may be less relevant than others. Changes may occur so rapidly that their effects negate one another within a short period of time, e.g. rapid fluctuations in weather. In addition to factors that cancel one another out, we also omit consideration of other interfering factors such as ones that simply take too long to be of present interest. The following example of the fish market illustrates the practical questions one considers when making economic decisions, e.g. what kind of employment to seek for the next year or two:

> We now impound fluctuations due to the weather in ceteris paribus, and neglect them provisionally: they are so quick that they speedily obliterate one another, and are therefore not important for problems of this class. And for the opposite reason we neglect variations…too slow to produce much effect in the year or two during which the scarcity of meat lasts…we give our full attention to such influences as the inducements which good fishing wages will offer to sailors to stay in their fishing homes for a year or two, instead of applying for work on a ship. (V.V.4)

Furthermore, Marshall was well aware of the risks posed by holding on to *ceteris paribus* assumptions for too long; even seemingly negligible fluctuations could accumulate over time (V.V.20). We cannot invoke such assumptions indiscriminately or for an indefinite period of time. Impounding has to be done carefully and with good reasons. Something that seems negligible in the short run, for example, may not be so negligible in the long run.[16]

What Marshall was interested these "normal conditions," which are deliberately ambiguous, specifiable only with the help of *ceteris paribus* restrictions.[17] And certainly he recognized that the analysis was incomplete. He certainly intended, though never completed, a

---

[16] Despite Marshall's explicit disclaimer, partial equilibrium analysis has borne the brunt of some heavy criticism targeted at its use of idealization and resultant unrealistic quality. Perhaps the most notable is from Piero Sraffa in his 1925 and 1926 papers that target these *ceteris paribus* assumptions, which motivated some to transition to general equilibrium models.

[17] On the other hand, Marshall was not interested in what we usually call perfect competition – he uses the phrase free competition, and perfect market (the latter which, at best, means perfect information).

follow-up second volume to the *Principles* on dynamical behavior. Despite this, Marshall did have substantive if not terribly precise assumptions about the role of scientific laws and what they looked like in their most desirable form.

> Those physical sciences…are not all of them strictly speaking "exact sciences." But they all aim at exactness. That is, they all aim at precipitating the result of a multitude of observations into provisional statements, which are sufficiently definite to be brought under test by other observations of nature. These statements, when first put forth, seldom claim a high authority. But after they have been tested by many independent observations, and especially after they have been applied successfully in the prediction of coming events, or of the results of new experiments, they graduate as *laws*. A science progresses by increasing the number and exactness of its laws; by submitting them to tests of ever increasing severity; and by enlarging their scope till a single broad law contains and supersedes a number of narrower laws, which have been shown to be special instances of it. (I.III.4)

Even a law of physics, such as the law of gravitation, was "a statement of tendencies" (I.III.7). So in principle, economics and physics were doing the same thing; physics was just better at it in terms of exactness. That is, the statements of physics were less probabilistic than those of economics, which was akin to the "science of the tides". He offers the Nautical Almanac as a successful case of exact physics:

> It is a very exact statement - so exact that mathematicians can calculate a Nautical Almanac, which will show the moments at which each satellite of Jupiter will hide itself behind Jupiter. They make this calculation for many years beforehand; and navigators take it to sea, and use it in finding out where they are. Now there are no economic tendencies which act as steadily and can be measured as exactly as gravitation can: and consequently there are no laws of economics which can be compared for precision with the law of gravitation. (I.III.8)

Economic phenomena fail to be as exact in part because they have some inherent degree of uncertainty, and for this reason measurement is difficult. Yet, despite being cognizant of the limitations of economic theorizing, what Marshall pictured as scientific improvement remains somewhat unclear. It is not obvious what the larger organizational scheme, in which "meta-laws" supersede the special instances, looks like. This is made more difficult by the fact that he does not say much about what it is for a law to be narrower in scope or to be a special instance of something else, nor what this "contains and supersedes" relation looks like.

Perhaps, in the end, he had a prototype of the deductive-nomological scheme in mind. But closer reflection on his remarks on the future of economics indicates a subtler view than that of a simplistic "deductive" view of science. In his 1920 *Industry and Trade*, Marshall disavows this view of science as artificial and only applicable within very controlled environments.

> Absolute certainty is possible only in regard to (1) particular individual facts; and (2) deductions by strict reasoning from axiomatic premisses, such as those of pure mathematics. Even sciences, which deal with concrete facts and conditions as definite and immutable as those of physics appear to be, cannot claim certainty over the whole of their area. In biological sciences the area over which certainty extends is relatively very small; and in the social sciences it is less than in those which deal with the lower forms of life. (43)

Both economics and physics state tendencies, and can claim only some range of certainty over the "whole of their area." While scientific progress clearly meant improved predictive power, on the other hand Marshall's vision of the completed economic science was not an obvious picture of dynamical laws of the kind we might expect: formulae that spell out the trajectory of a system once we fill in initial (and boundary) conditions.[18]

[18] Indeed, his vision of the improved economic science was fully dynamic:

> The unsatisfactory character of these results is partly due to the imperfections of our analytical methods, and may conceivably be much diminished in a later age by the gradual improvement of our scientific machinery. We should have made a great advance if we could represent the normal demand price and supply price as functions both of the amount normally produced and of the time at which that amount became normal. [Appendix H.11]

Elsewhere, he imagines:

Just as important as what might happen in the future relative to the present was the handling of contemporaneous causes all within the same time period. The complexity of the subject matter, Marshall thought, would ultimately require the discipline to move away from the methods of mechanics, which gave the appearance that economics could only provide static analyses of phenomena, and take its cue from the biological sciences.

> The Mecca of the economist lies in economic biology rather than in economic dynamics. But biological conceptions are more complex than those of mechanics; a volume on Foundations must therefore give a relatively large place to mechanical analogies; and frequent use is made of the term "equilibrium," which suggests something of statical analogy. This fact, combined with the predominant attention paid in the present volume to the normal conditions of life in the modern age, has suggested the notion that its central idea is "statical," rather than "dynamical." But in fact it is concerned throughout with the forces that cause movement: and its key-note is that of dynamics, rather than statics (Preface 19).

Unfortunately, the biological examples he offered, nor the extent to which they were more than merely suggestive, are not always clear. Some of these analogies addressed the dynamic behavior of markets themselves, e.g. comparing the growth of industry to a forest composed of many trees (individual firms).[19] [20]

> …a surface, of which the three dimensions represented amount, price, and time respectively. If we had marked on each curve the point corresponding to that amount which, so far as can be foreseen, seems likely to be the normal amount for the year to which that curve related, then these points would form a curve on the surface, and that curve would be a fairly true long-period normal supply curve for a commodity obeying the law of increasing return. [Footnote 84]

[19] Marshall had observed as an empirical matter that the development of industries tended to come with increasing returns: outputs increased in proportionally larger amounts than the amount by which the inputs were increased. This would be an advantage for large producers, making it more likely for a few large firms to monopolize the market, contrary to the assumptions of a competitive market. Marshall appealed to the imagery of the trees of a forest to explain this away: a developing industry consists of a number of firms in different stages of the life cycle. The trees-forest analogy was also used to allude to the specialization of labor that reinforced the

However apt (or not) those analogies might be, we will instead focus on another role for biology that Marshall saw, one that was different than as a metaphor for economic phenomena. He also used it as a metaphor for economic science. Though never completed, Marshall intended to write a follow-up volume to the *Principles* on dynamic behavior that would flesh out economic science. He had imagined that the complexity of economics would require finding inspiration in biology, rather than classical mechanics (Preface 19). Not only did biological examples supply analogues for economic phenomena, but also provided metaphors for economic science itself.

> The modern economic organism is vertebrate; and the science which deals with it should not be invertebrate. It should have that delicacy and sensitiveness of touch which are required for enabling it to adapt itself closely to the real phenomena of the world; but none the less must it have a **firm backbone of careful reasoning and analysis.** (Appendix B 46, emphasis mine)

If statics is always really a part of dynamics, *ceteris paribus* reasoning can not be eliminated from the completed science. Properly conceived, then, *ceteris paribus* assumptions are fundamental to economic reasoning, and though an economic statement "is not descriptive, nor does it deal constructively with real problems," it serves as "the theoretical backbone of our knowledge of the causes which govern value..."(V.I.4).[21] [22]

---

development of the economy as a whole, i.e. to the division of labor of systems that form increasingly specialized interacting segments that interact.

[20] In addition, he had in mind the division of function in the human brain and the innovation/automation distinction as biological examples. But by 1920 and the rise of joint-stock companies, he seems less sure of the biological metaphor, conceding that the "representative firm" was no longer the smaller, privately owned firm of the earlier century that would die out when its owners did.

[21] The motto preserved in *Industry and Trade* is still "the many in the one and the one in the many," which appears in Appendix C of *Principles*. This phrasing appears as well after his mention of the tactic/strategy analogy in Appendix C. In 1898 he offers up the example of rent as progress towards this that makes it a bit clearer: rent as at one many things, and yet many things that fall under the same heading.

Like the deployment of reversible processes in thermodynamic reasoning, *ceteris paribus* assumptions in economic reasoning help establish sets of relevant variables that are in principle dependent on one another. We draw out those causal relations by Articulating these variables establishes at the same time the scope of the problem at hand. In thermodynamics, we might establish the maximal efficiency limit of an engine. In economics, we might assess the extent of the effect of an economic perturbation (an influx of artists who insist on being pesceterians moving to the area) on a particular market (the fishing industry).[23] As Marshall notes, it is clear that the idealizations utilized by both economics and thermodynamics have a purpose that is not

---

> For we are learning that what is commonly called the rent of land is really a very complex thing made up of many elements, some of which differ more widely from one another than it, as a whole, differs from profits, or than some elements of it differ from wages. And as the obverse of this movement, those elements in rent, in profits, and in wages, which are similar to one another, are being drawn together, and the particular laws which govern them are being subsumed under more general laws common to all.

[22] This idea of the backbone of science functions the same way as Wilson's (forthcoming) symbolic "Clausius submanifold." The manifold is the metaphorical skeleton representing the physically impossible processes that real processes approximate.

> …we will collect together all of the possible temporal developments that a target system of a fixed construction might display, each of which we picture as a ribbon carrying the system forward in time...The entire collection…constitutes our primary manifold...In addition, winding as supportive skeleton throughout the whole wad of trajectories is a closed surface (or 'submanifold') that I shall call the 'Clausius submanifold'…Inscribed within this submanifold are various lines that, following standard practice, we shall call 'reversible paths,' although they actually do not represent any real form of temporal development in themselves...However, many of the off-the-submanifold ribbons (which do symbolize genuine temporal processes) wind fairly closely to the paths on the submanifold…(19-20)

[23] Perhaps intended to be comical, this was Marshall's own example in the *Principles* (V.V.18).

strictly descriptive or predictive. Instead, the idealizations employed have a demarcating role. They set limits on what is empirically possible and specify the relevant variables (namely causal ones) for a given problem of interest. In so doing, they also establish the limits of analysis. In their proper contexts, *ceteris paribus* statements *were* exact. For Marshall, he considered the law of supply and demand as general, though in any form we find them they are always hedged by ceteris paribus clauses.

2.3 SCIENTIFIC UNDERSTANDING IN TERMS OF EFFICIENCY

Ceteris paribus assumptions flag a methodology that we use to formulate clean answers to causal questions by positing an external (theoretical) controller manipulating one variable at a time. And while not explicit, the fundamental laws of thermodynamics are grounded on such a methodology. The causal relationships expressed by *ceteris paribus* laws draw the boundaries of what is empirically possible. Inextricably tied to this constraining role are opportunities for system control, and therefore what I call understanding in terms of efficiency.[24]

In applied contexts, our aim is to identify causal relationships apt for manipulation and the extent to which we can manipulate them in order to achieve desired effects. Notice that even in theoretical contexts, such as formulating the laws of supply and demand or the laws of thermodynamics, we nonetheless formulate them in terms of variable manipulation. Borrowing

[24] This allows us to understand economic reasoning in the form of these equilibrium analyses as attempting to function as a kind of (optimal) *control theory*. Given the operationalization of economics during the Cold War period (for discussion of post-war economic thought, see Erickson et al (2013) and Düppe and Weintraub (2014)), there seems to be some historical basis to this possibility. For an explicit interpretation of economics as a control theory, see Dorfman (1969).

some terminology from recent work by Woodward (2003), *ceteris paribus* assumptions provide baseline conditions in order to assess how *invariant* the causal relationships between different variables are. Woodward defines invariance thusly: "When a relationship is invariant under at least some interventions…it is potentially usable in the sense that…*if* an intervention on X were to occur, this would be a way of manipulating or controlling the value of Y" (16). The problem an economist or an engineer faces is figuring out which X's and which Y's matter, and to what extent they do. Assessing these relationships requires theoretical interventions that involve changing the value of a single variable and tracking the way it affects some other variable. What the relevant variables are depends on what the problem of interest is, and so they will also depend on how isolated that economic or thermodynamic system is from its surroundings. This assessment involves determining which factors may be safely neglected, e.g. those small daily weather fluctuations on the fish market. Manufacturing engines does not require our knowing all positions and velocities of the interacting gas particles contained inside. So *ceteris paribus* assumptions are a methodology for establishing the maximal invariance of relevant variables with respect to each other, while at the same time establishing what those variables are.

The knee-jerk reaction is to suspect something paradoxical is going on here: interventions in the real world are always what we might call "ham-fisted" or "fat-handed" – an intervention will induce an effect along more than one causal path. After all, real interventions involve more than *one* thing changing, and often several causal processes will mutually interact. Woodward's interventions are what one might call "surgical" (Pearl 2000) or "immaculate" (Bogen 2004) – the effects of any other variables are ignored or fixed during the hypothetical manipulation.[25] So

[25] Bogen (2002) is more concerned with the counterfactual semantics of Woodward's account, which Woodward is not explicitly committed to nor especially concerned with (see Woodward 2004 for this response).

how do we make sense of what actually happens in terms of what would happen?[26] My answer is that for many cases in physics and in economics formulating what *would* happen must be prior to the question of what will happen. Furthermore, doing so is not a stepping-stone to answering the question of what will happen; rather, it illuminates ways in which we could in principle *be doing better*. That is, we now know something about how efficiently a system is performing.[27] Marshall himself noted more than a century earlier that the problem faced was that "the various elements govern one another *mutually*, and not *successively* in a long chain of causation" (V.XIV). To disentangle them effectively at all we needed to invoke *ceteris paribus* assumptions. And this is, in fact, the very strategy used to construct a Carnot cycle. Like the Carnot cycle, our supply and demand curves draws the boundaries of what counts as *most efficient*. We can compare how we are actually doing against this baseline, as I did when evaluating the deadweight loss of setting my widget prices lower than they should be (see Figure 5 above).

The strategy embodied by *ceteris paribus* methodology is ubiquitous in other sciences. Woodward (2002) offers the example of a simple Hebbian learning rule expressing the relationship between change in the synaptic weight of a neuron to its pre- and post-synaptic activity.[28][29] While the Hebbian yields causal information, Woodward notes that we do not have to spell out all the conditions that hold in order to see that the Hebbian holds. While the Hebbian

---

[26] Machamer (2004) follows a similar line, emphasizing that articulating a mechanism is to tell us how actually a thing will work as opposed to how it would work.
[27] This might look like I am presupposing what it is to do better or what it is to be "efficient"; to assuage these worries, notice that although I use the term "efficiency" all this means is that we are optimizing some variable of interest.
[28] The Hebbian was initially intended to explain associative learning.
[29] Letting η be a constant, x_j and y_k pre-and post-synaptic signals, and $\Delta w_{kj}(t)$ the change in the synaptic weight of neuron k,

$$\Delta w_{kj}(t) = \eta y_k(t) x_j(t).$$

yields causal information, Woodward notes that we do not have to spell out all the conditions that hold in order to see that the Hebbian holds.[30] However, when a particular neuron conforms to this equation, we still know something about the conditions of its applicability without knowing the full range of those conditions. This information is important in the context of experimentation. In neuroscience and neural net modeling, the above equation specifies the change in synaptic weight between two neurons in proportion to their simultaneous activation. Roughly, this means that the synapses between neurons that happen to fire together will be strengthened.[31] So we have a method that can help us detect hubs of brain activity and also induce them, e.g. by subjecting a neuron to repeated inputs.

Given causal constraints articulated in terms of manipulable causal relations, we then know along what dimensions our system is most effective and to what extent we fail to achieve them (and so at the same time, what maneuvers are available to us for improvement). The examples we have been concerned with include the following: increase in the price of a good, strengthen the synapses of neurons in a particular region of the brain, or construct an engine with given heat reservoir temperatures. A pure decrease in supply pushes the new equilibrium price higher but at lower quantities sold. This might be doable in very controlled or isolated economies, but to a lesser degree or perhaps not at all in more integrated ones. We can control (at least sometimes) for neural plasticity.[32] While we cannot build Carnot engines, we can do fairly

[30] Other notable exceptions include volume learning, and the rule itself can be shown to be unstable for any neuron model since it does not place any upper limit on the strength of neuron connections, i.e. the magnitude of synaptic weight.

[31] That is, "when an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it" (Hebb 1949, 62).

[32] In neuroscience, one relevant inquiry concerns the neural efficiency hypothesis, which suggests that someone that is more intelligent "makes use" of fewer neurons because things are wired together "more efficiently" when performing cognitive tasks.

well. We may try to improve the efficiency of engines by creating greater temperature differentials or by adjusting the inner workings of the engine to reduce losses due to friction.

Idealizing methods that ignore disturbing factors are, thus, effective ways of dealing with complex, dynamic systems. Equilibrium exercises have clear, contemporary implications in applied economics, and in many cases are successful. More than a hundred years since Marshall, equilibrium analysis is still a dominant strategy in economics, in particular when it comes to matters of policymaking. In reaction to financial crises, notably since the 1990s and especially since the 2007 crash, the Federal Reserve conducts an annual review of the banking sector. One thing it checks for is whether the banking sector can absorb large shocks and yet continue operating normally – these are so-called *stress tests*. If the banking sector is able to withstand economic shocks, disastrous effects on individuals (and other regions of the economy at large) can be curtailed. The analysis is clearly counterfactual – *if* a crisis were to occur, then we want the economy to be structured in a way as to withstand it.[33] Our aim is not predictive in the strict sense that we should be able to deduce something's behavior at any point in the future. Rather, economic policy is often concerned with the question of how to best curtail the effects of financial crises in the case that one should occur, independently of whether or not a crisis does occur. Therefore, it may not give a precise answer as to how events will proceed afterward.

What we have is a methodology to establish some *a priori* expectation on how driving forces relate to the responses elicited. These may not tell us, however, about what will *actually* happen at any particular future moment. A Carnot engine doesn't tell us much about how actual engines work, and a partial equilibrium analysis won't tell us exactly how the actual fish market

[33] "…[S]tress tests estimate the exposure, the possible losses and the general reaction of the financial system to a specific event, but not the probability of the event occurring" (Committee on the Global Financial System, 2000).

will behave. But we do know the best our particular engine could ever do, or how much of a market change is possible given a shock to the economy. Given these benchmarks, we try to approximate the best case possible, whether it involves smoothly functioning engines or resilient markets. The methodology of using *ceteris paribus* yields understanding in terms of diagnostic causal relationships between salient variables; we use those variable relationships to determine and evaluate how efficient a system is and could be.

2.4 AGAINST ORTHODOXY: REJECTING THE SEMANTIC COMPLETER ACCOUNT

In the previous section, I proposed an interpretation of *ceteris paribus* assumptions that accommodates its original intended usage and still persists today in practice. *Ceteris paribus* methods serve to articulate relevant causal relationships, and so aid us in our understanding a system in terms of how efficiently it can perform. As a consequence, what I have said runs completely counter to the point of view pursued by a large group of prominent philosophers.[34] Indeed, this view even affects some of Marshall's prose.[35]

---

[34] In what follows, I consider a handful explicitly, but there is a long line of literature in both philosophy and economics (sometimes overlapping with the debate on reductionism, often on the subject of whether or not science qualifies as a science) that I take to represent the view I am opposing here, e.g. Nagel (1961), Hempel (1988), Cartwright at some times (1995), Hausman (1992), to some extent Klein (1947), Rosenberg (1992), and Strevens (2012), amongst many others. The view I am opposed to takes it to be the *default* interpretation of *ceteris paribus* assumptions that something unfinished or incomplete is going on, whereas my point is that there is a fairly well-developed methodological role.

[35] Despite Marshall's commitment to an untypical picture of science, some of his language suggests it was contaminated by elements of the orthodox view. Hempelian language like "subsumed by more general laws" connotes this kind of picture. Therein, I believe, lies the

The upshot of the standard picture of science as seeking and issuing immutable laws of nature is an over-privileging of predictive power as a theoretical virtue. This manifests as one popular understanding of *ceteris paribus* laws, the *semantic completer interpretation.* According to this view, "If A, then *ceteris paribus*, B" is an abbreviation for: "If A and… and…. then B." The *ceteris paribus* clause is a placeholder for this list of other conditions that are assumed to hold. But given the similarities with thermodynamics detailed in Section II, looking at a demand curve as representing a deficient species of law would be just as odd as, say, considering the idealized temperature-entropy curves for the Carnot engine as representing a deficient law of thermodynamics. As I have been arguing, the purpose of these idealizations is to articulate variable relationships for diagnostic purposes.[36]

Fodor (1991) and Schiffer's (1991) seminal debate over whether or not social science statements could count as laws was grounded in a common assumption, despite their differing verdicts. This assumption, which has been inherited by contemporary views, identifies *ceteris paribus* statements as problematic because they admit of exceptions. To turn them into robust laws, we would have to undertake the following project. A state such as "desiring cake" can be

---

tension in his enquiry. Despite preserving in *Industry and Trade* from *Principles* the motto of "the many in the one and the one in the many," the language of subsuming appears again:

> For we are learning that what is commonly called the rent of land is really a very complex thing made up of many elements, some of which differ more widely from one another than it, as a whole, differs from profits, or than some elements of it differ from wages. And as the obverse of this movement, those elements in rent, in profits, and in wages, which are similar to one another, are being drawn together, and the particular laws which govern them are being subsumed under more general laws common to all.

[36] This doesn't imply that completer accounts are *never* appropriate or useful, though it does seem to get the primary motivation for why such assumptions are used undergirding the methodology of some scientific endeavors incorrect.

realized in any number of ways at the neurological level. A lower-level state that instantiates this higher-order one might take the form of "neurons *a* an*d b* light up and *c* doesn't." We spell out what it would be to desire cake in terms of neurons lighting up, and in addition, we have to rule out other goings-on that may interfere and render the statement false. I may have overriding desires that would prevent my pursuing cake, such as a conflicting desire to make it to a department meeting on time. For Fodor, who maintains that there are such things as *ceteris paribus* laws, this means that there are circumstances which in addition to the realizer of "desiring cake" – the comprehensive description of such circumstances is called the "completer" – that will imply my pursuit of cake. As long as realizers have completers, the statement counts as a law. So suitably filled out, "Desiring cake leads to pursuit of cake" is a law according to this view.

The relevant items of interest in the special sciences involve higher-order predicates and are instantiated at the micro-level in more ways than one. Whatever being hungry looks like (say it's realized when a-neurons and b-neurons light up, but not c-neurons), even if *that* particular instance did not terminate it acquiring cake, it's all right if enough other statements where "someone is hungry" is replaced with "a and b, but not c light up" end up being true. In sum, for Fodor, "ceteris paribus, if I'm hungry, I'll get cake" is true just in case for every realizer of "being hungry" there's a completer C for which "realizer of being hungry + C" implies that I seek cake, *or* if there isn't anything that serves as an adequate completer for that particular realizer of being hungry so that it implies cake consumption, that realizer can figure in other laws in the network (such as (2) or (3)).[37] Thus, even if one particular realization of a predicate doesn't have a

[37] A completer is any additional information that in addition to the realizer state, would imply the consequent of the law in question. This information may be: "there are no disturbing external

completer, we may still be able to formulate laws using it.

So for Fodor, it's against a given *network* of statements that we evaluate whether a special science statement can qualify as a law. Multiple-realizability plus a network conception of laws grounds a ceteris paribus generalization. There's still more work that has to be done to make this suggestion feasible; Fodor’s account will allow too many things will count as ceteris paribus generalizations/laws. For instance, “ceteris paribus, if someone is hungry, she'll pursue cake” is a candidate for lawhood. It seems then that we are unable to distinguish between real generalizations and sham ones, since just about anything can be made to count as a ceteris paribus law.[38]

Unfortunately, there isn’t much by way of saving the semantic completer account. A project that involves hunting for completers will allow nearly any generalization to count as a law. If all that’s needed is filling out the completers, we can read “If A, then *ceteris paribus,* B” as “If A, then either B (or not B).” This renders such statements trivial instead of informative. Given how fruitless the completer endeavor seems, Earman and Roberts (1999) claim that “that there is no distinctively philosophical problem about *ceteris paribus*, but there is a scientific problem: what is needed is not finer logic chopping but better science” (460). Contra Cartwright (1983), “typical theories from fundamental physics are such that *if* they were true, there would be

---

causes.” So, in order to avoid charges of triviality, someone who advocates for a semantic analysis may insist on a further, distinct qualification that *ceteris paribus* claims may only be invoked if (and only if) some appropriate completer exists.

[38] Completer accounts manifest in several ways, but in general the ceteris paribus clause serves as a semantic buffer against the accusation that such a statement is either false or empirically inert. A structurally similar approach, which is not semantic, is a “completer" account in terms of epistemic conditions rather than truth-conditions. Pietrosky and Rey’s (1995) *ceteris paribus* laws have explanatory value and supposedly guarantee confirmable reasons why some event did not come about even if the relevant antecedent holds by appealing to theories independent of the *ceteris paribus* law's theory. For them, a *ceteris paribus* law's status as to whether it's vacuous or not is distinct from the matter of whether it's true.

precise proviso-free laws (446)." So, the statements of the special sciences do not count as laws because ceteris paribus statements simply need more development, not more filler.

> … Einstein's gravitational field law asserts – without equivocation, qualification, proviso, *ceteris paribus* clause – that the Ricci curvature tensor of spacetime is proportional to the total stress-energy tensor for matter-energy; the relativistic version of Maxwell's laws of electromagnetism for charge-free flat spacetime asserts – without qualification or proviso – that the curl of the **E** field is proportional to the partial time derivative of the **B** field, etc. (1999: 436)

Not only does physics contain *real* laws, but it has always been the aim of physics to do so: "We also claim that the history of physics and the current practice of physics reveal that it is the goal of physicists to find such strict, proviso free laws." According to them, the presence of *ceteris paribus* indicates a statement that serves as a landmark in a scientific landscape that is yet in progress. Like Earman and Roberts, Smith (2002) agrees it is only at the level of fundamental physics that we have laws that hold without exception; *ceteris paribus* clauses have no role there.

When economists assert the law of supply demand, they also assume that there are no outside interferences. Earman and Roberts presumably have no difficulty with the laws of economics if they are restricted to the lazy interpretation of *ceteris paribus*. But the reason that we do this is not because we are too lazy to fill them out. What about the other, presumably more robust senses of *ceteris paribus*? Here are their requirements up close. Formally put, the two conditions for *ceteris paribus* to hold in the non-lazy, weak sense is that:

> (i) There should be no condition ψ which can be stated in the language of X, which may not occur "naturally" but which can be realized using the techniques of X, and which defeats. By "defeats" we mean that when ψ obtains, is not even approximately true for X.

(ii) there be conditions θ such that when θ obtains, is true or approximately true for X. (461)[39]

What they mean by "naturally", for instance, is unclear. And in any case, Earman and Roberts require too much. The first condition alone simply rules out defeaters, even ones that can come about via manipulation! And consider that what's often the point of economic analysis is not only to get at what causal relationships are in play between variables of interest, but also to manipulate them. There's no reason to suppose, of course, that there is a clean delineation between natural and unnatural occurrences. This kind of definition flatly rules out a field like economics as something that issues laws. The alternative, for Earman and Roberts, is that we're either in the process of getting to exceptionless laws of this sort – which is unlikely – or that we are merely stating statistical relationships between well-defined variables. But often in economic analysis what we are trying to do is delineate variables, and it is not *given* that we have well-defined ones.

All parties we have discussed are far too restrictive on what they allow social science statements to do. Because these conceptions misunderstand what kind of tool *ceteris paribus* is in economics, they get the distinction between the social science laws and the fundamental science laws wrong. The debates we've looked at tend to assume that there is a clear distinction between things that are more fundamental or less fundamental, usually cashed out in terms of higher and lower orders (the latter being more fundamental). But this isn't the best way for us to

[39] There is a third condition to add to these if we wish to talk about their *strong* interpretation of *ceteris paribus* laws: "the condition θ of (ii) obtains in "most" of the intended applications of φ in X." (463).

orient ourselves. Economic progress doesn't typically consist in in making economic generalizations (and thus converting them into laws) by filling them out so that they yield universal, exceptionless statements at the lowest, most fundamental level.

But, one might suggest that economics is not one of these sciences that is in its early stages. To account for social science statements that seem fairly well developed, Earman et al. (2002) propose that "[t]here are plenty of important things that indicative-mood sentences can do other than state propositions or facts" (290). Perhaps, they suggest, they state statistical relationships between well-defined variables. This isn't satisfactory for two reasons. First, this is inadequate since our goal is to make sense of *causal* relations, not mere correlations.[40] Second, in economic analysis – and in thermodynamics – what we are trying to do is delineate variables, and it is not *given* that we have well-defined ones.[41]

---

[40] And this is not strictly speaking the only thing they say. They specify that: "In population H, a variable *P* is positively statistically correlated with variable *S* across all sub-populations that are homogeneous with respect to the variables $V_1 \ldots V_n$" (1999: 467). The articulation of the relationship between these variables is what I take *ceteris paribus* analysis to be trying to determine in the first place.

[41] Earman et al (2002) might suggest that I am missing the point. After all, a diagnosis of "why it has *looked* to people as if there is a problem of CP laws in the vicinity" is that:

> What makes it easy to miss the distinction between a *theory* consisting of a set of non-hedged laws and an *application of a theory* that might be hedged (though, again, in an easily stateable way) is that differential equations of evolution type – like the one we imagined deriving above – and their consequences are often thought of as *laws*. If one takes them to be laws, one expects them to be *part* of the theory in question and, thus, it looks like the theory contains hedged laws. But differential equations of evolution type are not laws; rather, they represent Hempel's applications of a theory to a specific case. They are derived using (unhedged) laws along with non-nomic modelling assumptions that fit (often only approximately) the specific case one is modelling. Because they depend on such non-nomic assumptions, they are not laws. For example, because Kepler's "law" that planets travel in ellipses is derived from laws together with the assumption that there are only two bodies in the universe, it is not a law in spite of the normal nomenclature. Lange's example of the "law" of heat expansion of metals is

Other philosophers have taken a different strategy that allows for the statements of the social sciences to be invariant under some *range* of counterfactual scenarios, rather than requiring them to be exceptionless simpliciter. This would both avoid the demand to provide completers as well as leave room for *ceteris paribus* to perform a role that is substantially lawlike. But they often still mistakenly assume a distinction between "more" and "less" fundamental laws depending on how stable or invariant they are, i.e. how more or less broadly they hold. So even more permissive accounts of lawhood, such as Mitchell's (2003) may not go far enough in capturing the primary role of *ceteris paribus* laws in economics. Michell's account emphasizes that the distinction between laws of the social sciences and the laws of physics or logic is not a matter of a difference in kind (where one kind counts as law and the others don't), but rather they differ in degree (in terms of regularity):

> At one end of the continuum are those regularities, the conditions of which are stable over all time and space. At the other end are the so-called accidental generalizations. And in the vast middle is where most scientific generalizations are found. It is my view that to reserve the title of "law" for just one extreme end is to do disservice to science by collapsing all the interesting variations within

---

> derived from a differential equation under the assumption that there are no boundary stresses, but that is a non-nomic boundary condition. (285-286)

We all agree that laws need not take the form of dynamical claims that spell out the evolution of a system. But then I am unsure how these remarks reconcile with conditions (i)-(iii) in the case for economics, as I view economic statements and thermodynamic statements as doing much the same thing (and according to them, thermodynamics *does* satisfy these clauses). Furthermore, even if their account is in part compatible with my reading of *ceteris paribus* laws, the question of whether or not something counts as a real law distracts from the primary methodological role of *ceteris paribus* assumptions in sciences where they are routinely used, as in economics.

Furthermore, given that they *do* consider phenomenological thermodynamics as providing the most plausible candidates for what they consider to be genuine *ceteris paribus* law, and that this is because thermodynamics is (presumably) reducible to statistical mechanics – and yet, demanding conditions that seem to rule out statements of the special sciences as laws almost right away – indicates an underlying predisposition to treat certain kinds of statements as laws.

science into one category: nonlaws. (138)

This spectrum is meant to capture the gamut of scientific statements from the law of conservation of energy (fairly close to our notion of an ideal law), Galileo’s law of free fall (further away from the ideal), Mendel’s law of inheritance (yet even further), all the way to accidental generalizations on the opposite end (138-140). But I do not think the difference between the laws of economics and those of physics (or anything else) is one of degree either – i.e. that the laws of one are “more invariant” than the other. I think this all puts the cart before the horse. Economic statements hedged by *ceteris paribus* clauses, such as the laws of supply and demand, are not invariant (or otherwise) just as they are. They can be used to analyze any part of the economy over any amount of time, depending on what the relevant economic problem is. If anything is invariant (or not), it is particular claims about particular markets. They do not make sense if we haven’t already carved out some particular part of the economy over some period of time. And the scope of what we are interested in, insofar as the economy is concerned, is often dictated by questions of economic policymaking.

Ultimately, I am not concerned with whether or not the statements of economics or other social sciences are laws, but rather with the assumptions that give rise to this question in the first place.[42] The background assumptions that underlie these debates are rather orthogonal to the concerns that motivate the *ceteris paribus* methodology; completer accounts are simply a philosopher's way of giving voice to predictivist assumptions about understanding in science. My account emphasizes the purely methodological role *ceteris paribus* assumptions play and the surprising amount of work it actually does in practice. The *ceteris paribus* rider is not an

[42] I also remain fairly neutral (I hope) with respect to the kinds of aims one ought to have in the background (“unification” for instance, being one, and “predictive accuracy” in some narrow sense being another).

abbreviation for a list of other persisting conditions. Nor does it indicate that the statement is a special kind of scientific generalization that is entirely different from other things considered to be laws. They are tools for the identification of invariant causal relationships, and the scope within which they are invariant, so that we can engineer efficient systems and processes. As we have seen there is no significant difference in methodology between the idealizations invoked by *ceteris paribus* assumptions in economics and those sometimes deployed in other sciences such as thermodynamics. While this does not imply that there is *no* difference between economics and the other sciences, it does show that any such difference cannot be justified by pointing at the presence of the *ceteris paribus* rider.

Nor is the scientist's goal to turn *ceteris paribus* generalizations *into* laws. One might think undertaking that kind of project will increase one's scientific understanding. But that individual is already committed to the orthodox conception of dynamical laws that I have been reacting against. There are substantial reasons to think that it is not a feasible project. Formal results in economics such as the 1950's Sonneschein-Mantel-Debreu theorem, which states that traditional assumptions on rational agents do not imply any interesting properties about the aggregate population (which, if we are economists, we are very interested in), are strong indicators that the-more-details-the-better approach is doomed from the start.

Completer accounts miss out entirely on this diagnostic kind of understanding in terms of efficiency that is crucial in applications. Instead they try to cram valuable scientific information into a form in which it does not fit. Like my account, a completer account is motivated in part by the fact that what we want to know is the range of circumstances in which a particular generalization holds. But they take *ceteris paribus* laws to be approximations of – and thus in some way deficient forms of – more general laws. On my account, *ceteris paribus* laws are not deficient forms of anything, though *real life systems* may be deficient, i.e. inefficient, when contrasted with them.

2.5 CONCLUDING REMARKS

Discussions in philosophy of science frequently take *ceteris paribus* assumptions to signal that the lawlike statement they figure in is not yet complete. This is orthogonal to the way *ceteris paribus* assumptions are often actually employed in the discipline in which they appear explicitly the most often (economics), and obscures the fact that similar strategies are employed in other sciences such as physics. *Ceteris paribus* riders flag that an evaluative methodology is in play, articulating constraining relationships drawn by pairs of variables. These relationships also make explicit the kinds of interventions we can make sense of, whether we actually effect such interventions or not.

One might initially think the obvious way to improve predictive accuracy is to treat an idealized model as a useful base case. The received view has it that science progresses by making its way to a more comprehensive and realistic picture by adding those extra details that we missed the first time around. This is not what scientists are actually doing much of the time in practice. Nor is this the way we should measure the degree to which something succeeds or fails to be a genuine science. The distinction between the laws of thermodynamics and the laws of economics is not a difference in kind since they both involve idealization involving *ceteris paribus* assumptions, and for the same reasons. They are also not different by degree from one another, because the *ceteris paribus* rider flags a background methodology aimed at constraint and control in both disciplines.

Instead of envisioning science as primarily describing autonomously evolving systems, philosophers of science should approach dynamical systems, including economic and thermodynamic ones, as the kinds of things we can theoretically intervene on. This shift in perspective has dramatic effects, and yet still preserves a sense in which we talk about the *dynamical* properties of systems without having to give up our static idealizing strategies. We move away from describing and predicting the behavior of autonomously evolving systems to evaluating the behavior of systems in terms of the ways they can be manipulated and improved. This recognizes idealizations via *ceteris paribus* hedging as crucial causal tools that are usefully deployed in sciences quite generally.

3.0 DYNAMIC AND DISEQUILIBRIUM BEHAVIOR

3.1 PRELIMINARIES

Even our most sophisticated "formulas, or models, are only pale reflections of the real world, and sometimes they can be woefully misleading" (Freedman 2011, 76). The criticism runs deep – all the way down to the fundamentals of economics, which still ground current practice. For instance, economist Steve Keen (2015) chastises the Fed for its failure to foresee the global financial crisis of the late 2000s, and "the fact that [t]he Fed was caught completely unawares by the crisis should have led to some recognition that maybe, just maybe, its model of the economy was at fault," targeting in particular the use of *equilibrium* models. At any given time, things are *not* at equilibrium. Given that this is unrepresentative of actual economic behavior, one might think there must be something wrong with economics.

As I demonstrated in Chapter 2, there are significant methodological parallels between equilibrium reasoning in well-established subfields of physics and in economics. In particular, equilibrium methods in both thermodynamics and in economics aim at identifying salient scale-sensitive causal variables. This chapter aims to further clarify the proper role of such idealizing strategies in applied context, where – in both thermodynamics and in economics – we face the task of modeling and understanding complex, dynamic, heterogeneous systems.

Section III locates the failure to appreciate the appropriate role of equilibrium reasoning in a combination of two attitudes that dominate philosophy of science. The first is the inclination to treat every system as one that evolves autonomously. Most philosophical assessments of scientific theory as a result overemphasize the virtue of predictive power, conceiving of models as representations of dynamical systems that generate descriptions of future states given initial and boundary conditions. The second is a tendency to favor "bottom-up" model building – the idea that we can model complex systems by building up from the details of their micro-constituents.

However, in both economics and physics, the relationship between micro-and macro- scale models is less than straightforward. So how do we talk about complex, heterogeneous, dynamic systems given our available methods? Equilibrium methods can still help. To reorient our perspective on equilibrium strategies, in **Section IV** I draw inspiration from Pierre Duhem's *false equilibrium* concept in order to introduce what I call "scale-linked equilibrium" states. These are states where systems may macroscopically be stable, yet their micro-constituents are not at *true* equilibrium. Systems in such states are ubiquitous and can be reliably brought about. For instance, amorphous solids like glass are slowly, but ever so surely, relaxing towards a final goo-like state. So while the system as a whole may be at equilibrium for our intents and purposes (e.g. on a short enough time scale), but it may contain defects and structural features that prevent the occupants of lower-order micro-scales from immediately settling into their own respective equilibrium states. Dealing with non-equilibrium, non-ideal behavior will therefore look different from what most philosophers might expect. It requires a framework to accommodate such interesting, and important, structural details, yet we need not discard our traditional modeling tools. The *multi-scale* framework, already put to use in other fields such engineering and biology, is such an apparatus in which we can usefully embed equilibrium methods.

3.2 DIAGNOSTIC REMARKS

There is a tempting analogy that I think it is a mistake to make. Suppose we're interested in the behavior of an ordinary object moving through space such as a block sliding down an incline. All of these objects are conglomerates of smaller things – at bottom, particles. One might think that the *most* fundamental scientific theory should have the resources to describe those smallest constituents. For a system that consists of a bunch of particles, then, all we need is the position and velocity of each particle (plus the laws of mechanics). So analogously, for complicated economic problems, one might think that we just need to add in more details, such as individual details for our heterogeneous agents, in addition to our usual laws of supply and demand.

The idea is that once we have accumulated all the details at a micro-level, we will be able to infer or deduce information about what goes on at the macro-level. Perhaps, then, the goal of science is to build a theory about the smallest constituents of a system such that we can deductively infer what kinds of higher-level phenomena will occur given all the information about those micro-constituents. Two distinct, but commonly held, attitudes give rise to this. The first is the assumption that the scientist is after laws of nature that describe the behavior of a system as it evolves autonomously. But given the centrality of and similarities between quasi-static reasoning in thermodynamics and comparative static reasoning in economics, it seems rather to be the case that the scientist must see the system as one that she can theoretically control.

The second source of modeling difficulties is the bottom-up nature of most microeconomic reasoning. That is, economic models are often built from the case of the individual upwards to groups. A bottom-up methodology aims to achieve a picture of the whole economy first by gleaning information about its constituents and then somehow putting all that information together. It's common, though unwarranted, to assume that such a strategy is possible in the physical sciences, i.e. though we do not need to specify the individual constituents of a body in order to determine its behavior, in principle we could.[43] [44]

In actuality the bottom-up attitude will actually give patently wrong answers when it comes to most of the questions we ask about the real world. Goings-on at higher order scales and lower order ones are in some way coordinated, though this relationship cannot be captured by deductive relations. To take a simple example:

> …consider a rubber band, which is composed of a tangle of long polymer chains built from carbon, hydrogen and allied atoms. From the point of view of orthodox statistical mechanics, these molecular chains possess a "frozen order" that is hard to diagnose but keeps the material from satisfying the standard conditions on equilibrium. Indeed, considered on a very long time scale, such chains will break down into a simple soup of unlinked molecules. (Wilson forthcoming)

Rubber bands are metastable viscous liquids. Wait for long enough (several hundred years), and a rubber band will eventually degrade into a molecule soup. It doesn't do so right away because the molecular chains have structural properties so that they degrade over a long time-scale – so long that we may neglect it for practical purposes – the "frozen order" that Wilson speaks of. If

[43] Often accompanying this attitude is the tendency towards reductionism, a Nagelian artifact that has somehow managed to persist. This is especially evident when a philosopher claims that thermodynamics reduces to statistical mechanics, which is strictly speaking false except in the special case of the thermodynamic limit as volume approaches infinity. This is the stance that a more fundamental or legitimate science has laws that are made in terms of the most individual constituents or at least can be somehow "reduced" to them, usually deductively.

[44] See Butterfield and Bouatta (2011) for this view, and Batterman (2013) for an argument against.

we really were to consider the rubber band at true equilibrium, we would be treating it as if it were soupy already, and "we will obtain truly rotten estimates of the material's properties…"[45]

Another example of an everyday kind of problem we might worry about is how to manufacture materials in order to construct sturdy buildings that can sustain some wear and tear. Consider a steel beam, which has a crystalline structure at the microscopic level. However:

> If we engaged in a purely bottom-up lattice view about steel, paying attention only to the structure for the pure crystal lattice, then we would get completely wrong estimates for its total energy, for its average density, and for its elastic properties. The relevant Hamiltonians require terms that simply do not appear at the smallest scales. (Batterman 2013, 268)

Knowing all the details about the microstructure of the steel beam won't help us at all. Looking at just the macro-scale won't do it either; a steel beam looks fairly homogeneous from that perspective, too. Even in this fairly mundane case, it turns out that structural features crucial to

---

[45] Wilson (forthcoming) elaborates on the rubber-band example:

> Instead, physicists have directly probed dislocation behavior at a higher "mesoscopic" level and can supply quite sharp empirical rules for how ably these structures directly store strain energy in terms of their two-dimensional bending (dislocations must be approached as string-like "objects" at a higher scale resolution for these geometrical claims to make sense). We can then assume that our dislocation "objects" will quickly assume configurations where their own stored energy becomes minimized at their own characteristic scale, without releasing much of this residual energy to the molecules themselves. In other words, an unattached dislocation string will minimize its energy by straightening itself out and a chain connected end-to-end will seek a circular shape (thereby trapping a significant amount of stored energy). In most real life materials, dislocations appear as a profuse and knotted snarl of interweaving lines, so calculating their collective minimal states of strain energy can prove quite challenging. But if we can do this, we can allow each scale level to minimize strain energy in its own way, dependent upon the ambient temperature. Exploiting our independent empirical knowledge of several different scale-focused behaviors in this combined manner, we can augur the material's final macrolevel behavior far more accurately than if we had tried to work in either a wholly bottom up (molecular modeling) or top down (Cauchy-Green-Stokes scaling) mode. (33)

steel's ability to withstand shocks are located neither at the micro- nor the macro-scale, but at meso-scales in between.10 Giving rise to these features are dislocations, structural imperfections, "non-equilibrium defects" that contain trapped energy even though the overall structure of the material they appear in may be at equilibrium.

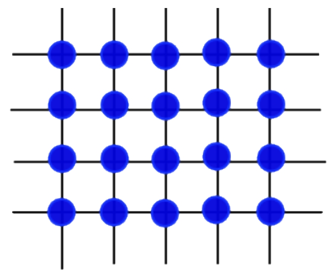

Figure 6. A Perfect Crystal Lattice.

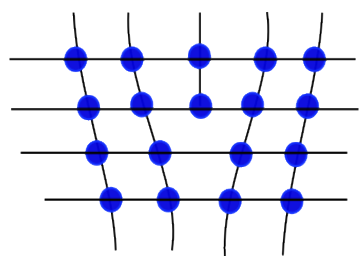

Figure 7. A Crystal Lattice with Dislocations.

They form in a material during deformation or via solidification (when a material crystallizes). At the microscopic level they appear as small irregularities, but they also tend to occur in groups, appearing as chains (grain boundaries) that become apparent at higher scales. When external pressure is applied, the dislocations move through the structure they reside in instead of dispersing. Shear occurs in a material when molecular bonds gradually break and reform along these irregularities, finally resulting in a slightly deformed state.

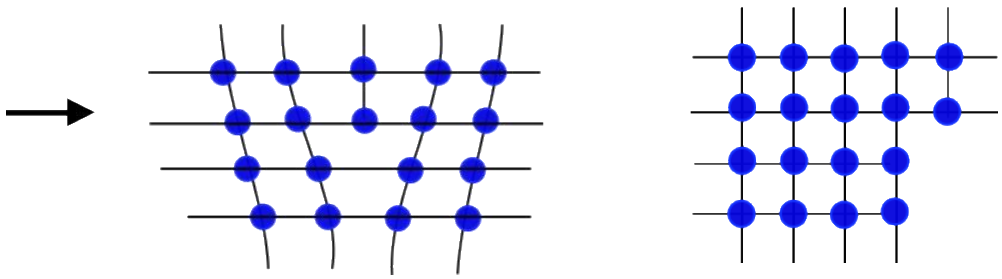

Figure 8. Before and After: A Crystal Lattice Subjected to an External Force.

This is certainly preferable to brittle behavior, which a perfect crystalline structure would exhibit. This property we are interested in, elasticity, is nonetheless epistemically inaccessible at the molecular level. Bulk materials will come to rest at a different equilibrium than would be predicted had we simply used bottom-up calculations on the information gleaned from the molecular crystalline lattice structure alone. What counts as equilibrium at one scale may not be at all what counts as equilibrium at another scale.[46]

This is certainly preferable to brittle behavior, which a perfect crystalline structure would exhibit. This property we are interested in, elasticity, is nonetheless epistemically inaccessible at the molecular level. Bulk materials will come to rest at a different equilibrium than would be predicted had we simply used bottom-up calculations on the information gleaned from the

[46] Both disciplines deal with systems with many interacting, heterogeneous parts, but the details of the individual constituents are often not the main target of analysis, nor is it necessary in many cases. These variable relationships we use to describe a system is exactly the identification of what Wilson (forthcoming) calls dominant behaviors

> We want to know "when does behavior X become strongly dominant when a mass of Y's collectively interact?" but mathematics stupidly replies, "I do not understand your term 'dominant'; in my eyes every scrap of behavior is equally important." So we must often trick her into supplying the answers we seek.

molecular crystalline lattice structure alone. What counts as equilibrium at one scale may not be at all what counts as equilibrium at another scale. The characterization of the equilibrium state itself does not say anything about the dynamical behavior of an actual thermodynamic system for the same reasons a comparative statics exercise fails to tell us what exactly will occur to a market given a disturbance. All that an equilibrium analysis elucidates are in-principle relationships between variables of interest, and those variables are indexed by which scale we have specified. This also implies that what counts as dominant behavior – behavior that we think of as salient, like elasticity – is sensitive to scale too.

But information across scales may be impossible to integrate into an axiomatic, deductive system. If one asks, "What do all the individual particles of a stable structure look like?" there's a good chance that the question is not a productive one to ask in the first place, even if we ignore problems of tractability. The same difficulty arises when one asks: "what do all the individual actors in a stable economy look like?" Even the most elementary concepts in microeconomics fail to yield a straightforward transition from individuals to aggregates. For example, the law of consumer demand tells us that an increase in the price of a commodity will (*ceteris paribus*) lead to a decrease in the demand of that commodity. Intuitively, the more something costs, the less willing I am to buy it. One might try to investigate the applicability of the law to a group of individuals do this by aggregating all the individual demand curves of a population together. But significant restrictions must be in place for a smooth transition from individuals to the aggregate, and the end result is not a useful one even when so. The 1950's Sonneschein-Mantel-Debreu (SMD) theorem is the formalization of this difficulty.[47] It states that the rationality assumptions

[47] For the aggregate function to depend only on societal income and the price (in addition other usual restrictions, e.g. that demand curves all be independent) the preferences of the individual

that apply to individual demand functions do not carry over to the aggregate, so that familiar utility maximization assumptions on individuals do not lead to any testable, interesting properties of the population as a whole.[48] That is, there is no general result on what the aggregate demand function looks like.[49] If the micro-details cannot imply that the aggregate has a desirable property like stability, then what does? This suggests two possibilities: that we require non-standard assumptions on individual agents, or that more (meso- and) macroscopic properties should be introduced from the get-go.[50] [51]

Multi-scale modeling is a strategy that combines both without being ad hoc. Those features of thermodynamic thinking are often masked by an excessive focus on cases that involve only simple (often ideal) gases (in what are called "standard" conditions). For other more complicated objects, however, such as solids that we encounter all the time, we require a framework that functions as a bit of corrective architecture that buttresses strategies like positing quasi-static processes (and equilibrium states), so that they count as useful modes of reasoning.

---

need to be identical and quasi-homothetic. Then there's the further problem of whether equilibria can be unique. In a two-agent, two-commodity exchange economy, if agents view commodities either as perfect substitutes or as perfect complements, there is no unique equilibrium. Rather there is a continuum of equilibria.

[48] The only things that do seem to carry over is homogeneity of degree zero, continuity, and obeying Walras' Law (excess market demand/supply is zero) and boundary behavior when prices approach zero.

[49] Arrow's Impossibility Theorem drives this point even further home. It demonstrates how difficult this can be even under some basic assumptions. Under some intuitive constraints on preferences to ensure fairness, no social preference function for a group of individuals can be built from the preferences of those individuals (at last in an exchange economy).

[50] And indeed this is something that comes up in contemporary debates now over the issue of "microfoundations" – the endeavor that tries to build a picture of the macro-economy in micro-economic terms.

[51] Grandmont (1992) and Hildenbrand (1994) are two well-known attempts that recognize the role of meso-scale details, in particular population heterogeneity, in achieving stable equilibrium. See also see Blundell and Stoker (2005), Lewbel (1994), Rizvi (1997) and Kirman (1992).

The multi-scale framework can function as this corrective architecture; equilibrium methods are a way to coherently articulate salient scale-dependent information without prioritizing any scale over another. We can think of meso-scale structures as acting a bit like "filters" for, e.g. policy effects, as they reach the microscale (individuals). For example, in their investigation of Latin American policy, Rome and Caracas (2001) suggest that economy analysis often neglects such crucial meso-structures:

> A consensus has emerged in the academic literature, and among national policymakers and multilateral agencies, that structural adjustment particularly during the 1980s, was not structural enough, as it failed to focus on some critical links between policy measures, intended objectives, and their actual outcomes. In Latin America, by focusing almost exclusively on price reforms ("getting prices right") and macro-economic variables, policy reforms tended to ignore important structural and institutional traits that have hindered the achievement of the alleged goals of economic growth and increased human welfare. The negative social effects of the reforms, for instance, have only recently been recognised by policymakers and started to be addressed through the design of safety nets for the newly unemployed and the newly poor.

3.3 A SECOND LESSON FROM THERMODYNAMICS: FALSE EQUILIBRIA AND FINANCIAL CRISES

The behavior of systems that are complex with interacting parts behave differently from what is theoretically conjectured due to the presence of dislocations and grain boundaries in the system. As mentioned, we do not expect our thermodynamic calculations to hold in the empirical world.

These micro- and meso-imperfections give rise to interesting behaviors at the macroscale.[52] One might think the obvious way to improve the model's predictive accuracy and explanatory power is to treat idealized models as useful base cases, from which we could build up theories incorporating those features initially neglected like microscopic defects and omitted frictional forces. That is, we just make our way to a more comprehensive and realistic theory by adding those extra details. But this is to give in to the two received attitudes I rejected in the previous section.

Dislocations may seem like theoretical inconveniences, but in practical engineering contexts they are both useful and necessary. Manufacturing materials requires us to take into account the behavior of objects at different scales, but does not presuppose some privileged scale as fundamental. As I noted earlier, the presence of dislocations (and grain boundaries) is what allows materials to bear stress and strain. We could import this lesson into economics: not all the behavior we are interested in is at the level of individual agents nor derivable from it. The strategies we use to intervene on and stabilize our economy – namely via fiscal and monetary policy – require our understanding scale-relative behaviors and frictions in the way that an engineer understands and successfully manipulates bulk materials.

Standard methods in classical physics such as Hamiltonian and Lagrangian methods of describing systems and their time evolution are unable to satisfactorily handle ubiquitous, real

---

[52] As a disclaimer that not all physical phenomena that we usually associate as lagging behavior, off- equilibrium behavior, and the like, need be due to be frictional forces strictly speaking. For example, one might expect lags to occur when two bodies exert force against one another without sliding (think normal force). This derives from the gravitational force multiplied by the mass of the object itself. These may not be frictional forces, but different kinds of interactions. Thus we ought to be cautious in drawing any substantive analogies in economics to one of "friction" in physics.

life phenomena such as friction.[53] In thermodynamic cases, friction between interacting constituents (amongst, undoubtedly, many other things) prevents systems from reaching calculated equilibrium values. Yet, accounting for friction is what allows us to make sense of a certain class of stable states. Let's instead consider for a moment Pierre Duhem's work and his false equilibrium concept.[54] Duhem distinguished two different kinds of such states.[55] The first kind includes systems that are out of equilibrium and unstable *per se*, but seem to remain stable until massively disturbed, while the second refers to systems moving very slowly towards equilibrium. Cases of the first sort include surfused (undercooled) liquids and supersaturated liquids. Oxygen and hydrogen will stably coexist even though they may not at equilibrium, until ignited by a spark. On the other hand, substances such as diamond or petroleum may remain stable, though out of equilibrium, for long periods of time and seemingly even when subject to fairly large perturbations. Similarly, amorphous polymers like glass are slowly, but ever so surely, relaxing towards a final goo-like state.

Duhem himself did not anticipate multi-scale modeling in its contemporary form, and though he certainly thought equilibrium reasoning to have something like the role I've been advocating here, his notion of "false equilibrium" does not distinguish between scales. But we can, and so further refining Duhem's notion, these systems are at what I will call "scale-linked equilibrium" states, which many real world phenomena inhabit. Systems that are in scale-linked

---

[53] To spell all these out in a Hamiltonian or Lagrangian would be unwieldy; we would have to describe the interactions between every two particles, or alternatively use a catch-all dissipation function. Various technical difficulties arise in classical Lagrangian and Hamiltonian approaches to physics. The Lagrangian approach is not even an autonomous method, but is actually a hybrid that exploits Newtonian formalism.

[54] See Duhem (1893).

[55] Which he termed "apparent" or "real" (1902, 436).

equilibrium states may be at equilibrium at the macro-scale, but structural features prevent individual constituents of those systems from settling into their respective equilibrium states.[56]

When we manufacture steel beams for construction, what we are really aiming for is a material that resides in a scale-linked equilibrium state. Despite being non-equilibrium features, dislocations are to an extent desirable defects in steel beams; we can still reliably handle them and even bring them about. We purposefully inject such frozen order into the structure of materials at the meso-scale so that shocks will disperse differently throughout the micro-scale. In metals, we can increase their ability to bend by heating them to very high temperatures – a process called "annealing," by which we manipulate the grain structure – and then "freezing" that structure in place by quenching them in cold water. The resultant material is both manipulable and stable, appearing at least at the macro-scale to remain as it is for long periods of time and displays predictable behavior under some range of external shocks.[57]

There are analogous phenomena in economics that all (somewhat unfortunately) are classified as examples of "friction".[58] Sometimes it goes by other names: rigidity, stickiness, and so on. Frictions, in economics, are what cause the economy from exhibiting the clean, theoretical trajectories way our models depict. But such obstructions can help stop financial contagion by slowing it down; most of this has to occur at the institutional level of banks. Banks can and do fail; we would like to prevent failure on a massive scale due to spillover effects. A bank fails due

---

[56] Though Duhem's primary examples are chemical, he also offered a mechanical analogy (1896). Consider the box sliding on the incline again. In a frictionless classical mechanical picture, this would not be at equilibrium. But in the real world, under a certain threshold for the incline of the plane, the actual equilibrium state is different from the theoretical equilibrium point precisely because of the effects of friction. The box will stay where it is until the incline reaches some critical angle.

[57] In engineering this is what they call "metastable" rather than just merely stable.

[58] The economic literature tends to call everything that can be considered to be non-equilibrium behavior as due to friction.

to insolvency: it is unable to hold up its end to its depositors and creditors. An asset is more or less liquid if it can be easily exchanged on the market without affecting its price too much. In the case of a bank panic (and everybody rushes out to withdraw their money), money becomes increasingly illiquid as a bank becomes less and less able to give out the necessary funds to its depositors. But this kind of failure can sometimes be truncated and absorbed, so that cascading effects do not reach individual customers. In 2008, Washington Mutual failed – the "largest bank failure in US History" (Sidel, Enrich, and Fitzpatrick 2008). However, JP Morgan Chase stepped in to buy all of its banking operations in exchange for a payment to the Federal Deposit Insurance Corporation (FDIC, which insures depositor's bank accounts), thus absorbing the shock. All Washington Mutual customers then effectively became JP Morgan Chase members, but to individual customers the change was only nominal.[59] In fact, like in the materials sciences, it is now common practice now to conduct *stress tests* on the economy. When the US Federal Reserve conducts stress tests, one thing it checks for is whether the banking sector can absorb large shocks and yet continue operating normally. If the banking sector is able to withstand economic shocks, disastrous effects on individuals (and other regions of the economy at large) can be curtailed.

[59] Of course, there were some losers in the transaction – WashMu stock and debt holders were left with nothing.

3.4 THE 2007-2010 FINANCIAL CRISIS

Only in extreme cases does looking at a sample of a system at the atomic scale reveal anything noteworthy about the macroscopic properties of that system.[60] While what properties do transfer may be a relevant question, often we are interested in properties that we do not even have the appropriate theoretical vocabulary to talk about at the atomic level. The example of the steel beam and some of the banking structure of the economy serves to show that one crucial property we are interested in is *stability*. Stability for materials is straightforward enough for our purposes – the ability to withstand stress without buckling. In economics, we'll take stability to mean a system resides in a stable equilibrium state – if perturbed, it will return to its original position. After all, it's not just *any* equilibrium state that we are interested in, because some equilibrium states are unstable.[61] If an economy is in an unstable equilibrium state, a disturbance may cause it to move to a new, very different equilibrium state.[62] On the other hand, if we're in the vicinity of a stable equilibrium state, we'll tend towards equilibrium.[63]

---

[60] In fact, there is extra work that needs to be done in order to demarcate a sample as representative, with an entire subfield of applied mathematics (homogenization theory) dedicate to this.

[61] A conceptual point relating to the concept of efficiency: the main results coming out of general equilibrium are the welfare theorems, which relate (Pareto) efficiency and equilibria. The first tells us that every equilibrium state is an efficient state – changes can't be made from that state without making someone worse off. The second one is more complicated, requiring the convexity assumptions mentioned here. Not every efficient state is an equilibrium state – after resources are allocated, wealth must be shuffled around within the agent population in order to support that allocation arrangement as an equilibrium state.

[62] For some, the heavy use of equilibrium was reason to be pessimistic. Consider Hicks: "Now the reason for this sterility of the Walrasian system is largely, I believe, that he did not go on to work out the laws of change for his system of General Equilibrium" (1939, 61).

[63] Note that we do not want to be in such a stable state that the economy cannot grow any further. Part of the economist's job is therefore to help nudge the economy along a growth-path.

The intractability of a bottom-up model of economic stability means that we should adopt a different tactic to model such features. Elasticity in materials is due to the presence of meso-scale structure that we can bring about via manufacturing. Meso-scale structures like institutions – whether they manifest physically (banks) or not (legal measures) serve the same function in economic systems. In this section I consider the subprime mortgage crisis of 2007 to demonstrate how appealing to the thermodynamics of solids and multi-scale methods can provide insight into the economy.[64] In particular, the upshot is that meso-scale structures like institutions may serve as sources of frozen order in economic systems.

Suppose I wish to buy a house and take out a mortgage loan. I go to my local bank, which decides whether or not to let me buy a house depending on their expectation that I will pay back the amount (plus interest) reliably over a fixed amount of time, as compared to using it as a rental property that earns income over that same period of time. We'd expect the bank to refuse me a mortgage if the former is not expected to be at least as profitable as the latter. If I receive a mortgage loan, I agree to pay the principal of the loan back at some interest rate. However, I may default on the loan or may pay off the loan early – meaning that the bank might not receive the entire intended loan amount plus interest bank. In order to keep loaning out money, this bank would have to replenish its funds. That bank will sell this mortgage to a bigger bank in return for a fixed amount of money, ridding itself of the risk that comes with the mortgage loan itself and enabling itself to make further loans.

Instead of refusing mortgages to homebuyers who were not expected to repay, a bank could simply unload the risk of that loan to another bank. Since housing prices were expected to

---

[64] Further discussion can be found in Stiglitz (2010a) and Johnson and Kwak (2010).

keep increasing, the worst-case scenario meant that a bank could just repossess a house. Loans traversed an upward path to bigger and bigger banks in this way. During the housing crisis of the late 2000s, investors with more money available to invest than things to invest in turned to mortgage backed securities.[65] Individual mortgages were bought, grouped together, and sold again and again until they ended up at a large bank on Wall Street. There they were bundled together and sold finally to investors. As borrowers paid off their loans, these large banks received the income from mortgage payments, and investors bought into shares of that mortgage income. These shares are mortgage-backed securities, and demand for them exploded. Mortgages were handed out to borrowers with shaky income, histories of debt, and bad credit.[66]

Moving the risk upstream means that the individual bank selling the loan rids itself of the risk involved, but it also creates incentives to lie in order for mortgage bonds to continue being generated in order to meet demand. Leading up to the crisis, those loans (so-called “toxic loans”) targeting high-risk borrowers often involved fabricating crucial information such as how much income they earned. At the level of individual mortgages, however, everything looks safe since the house serves as collateral even if the borrower defaults.

Mortgage loans were chopped up and bundled together in various ways as they made their way upwards. Individually risky loans found themselves in collateralized debt obligations (CDOs) with high ratings from credit ratings agencies using outdated housing data. From all appearances, these risk-laden packages of mortgage loans looked on the surface like good bets. A CDO is itself divided into “tranches,” which vary in their risk profiles and are ranked

---

[65] The other option would have been to invest in US treasury bonds, but investors wanted high-return low-risk investments and Greenspan wasn’t keen on raising the federal interest rate that that time.

[66] For more details, see Davidson and Blumberg (2008).

accordingly: senior, mezzanine, and equity (unranked). This ranking determines who gets interest and principal payments first and bears the brunt of default losses. The more junior tranches are affected by losses first, for instance, while senior tranches holding most of the debt actually tend to retain their capital value unless there is a widespread crisis. An investor picks which tranche to buy into depending on how they assess risk and reward – higher risk and higher reward seeking investors would buy into the junior tranches, while low risk but low reward seeing investors bought into the senior ones.[67]

The intuition is that the bundling is meant to be a buffer against loss by distributing risk, both default and prepayment risk.[68] A mortgage might come along with an adjustable rate, so that one pays a low interest rate the first few years of the loan but a high one afterwards. The borrower is likely to pay off the loan those first few years, but less likely those later ones. If I prefer low returns with low risk, rather than high returns with high risk, in that case, I prefer the tranche that contains that less risky bit of the loan.

Tranches became incredibly complex. Financial intermediaries such as hedge banks often mediated buyers and financial products. Consequently, most individuals probably did not know what they were buying into. Risky assets were being bought and sold by individuals that had incomplete knowledge of how risky they were. Mortgage loans, tranches in CDOs, etc. were also generally treated as behaving independently of one another, so correlations were often not modeled explicitly if properly at all. As one economist put it, one major problem with the models being used was that "they usually do not incorporate the financial sector. This is all the most

---

[67] See Hamilton (2009) for a summary of tranches.
[68] For a discussion of risk distribution and absorption, see DiMartino and Duca (2007).

surprising since they were used by Central Banks and advisers to policy makers in monetary policy" (Garcia 2011). Measuring credit correlation was in any case, a terribly difficult and tedious enterprise. Even the information that would have been necessary for it was often times limited in availability (corporate balance sheets are only published quarterly, for instance). [69]

Eventually, borrowers began to default as the price of their homes exceeded their incomes earned (which were not rising). Homes flooded the market, and as the supply of houses went up their value went down. Panic, insurance policies backing credit defaults couldn't make good on their guarantees when homeowners starting defaulting on their loans, and individuals betting against the subprime mortgage market all contributed to the forces exerting pressure on the market. Mortgage bundling, which was meant to spread out the risk of defective loans, merely masked those defects; if the riskiness of the loans being bundled were independent of one another (which is unlikely), such overall risk would be reduced. So, despite highly rated tranches of mortgage-backed securities, these interrelated risks stayed trapped in these packages, affecting their behavior in unanticipated ways. It was not particular individual loans, but rather the

---

[69] Part of the risk measure of a tranche package involves correlations amongst mortgages. JP Morgan offered the following metaphor to describe the effects of correlation trading – a financial strategy used by exploiting correlations between financial instruments. If two stocks are correlated, they tend to behave similarly. However, how strong that correlation is varies over time, and a change in correlation may then be followed by a return to the mean trend. The investor can then exploit changes in correlations in order to make a profit. Consider a blindfolded cat walking through a room that has mousetraps scattered about. If a cat could express her preferences, she'd probably prefer to talk through a room where the mousetraps tend to cluster (the trap locations are highly correlated). This corresponds to a lower-rated tranche. On the other hand, if the cat has nine lives, as fictional cats do, she'd prefer a room with the mousetrap locations to have low correlations – while she's more likely to step on any particular trap, she's certainly less likely to step on nine in a row. (JP Morgan, as quoted in Moore 2004).

organization of the mortgage market that contributed to the crisis. Ultimately, the Federal Reserve had to bail out these big banks – by injecting liquidity into the system.

In this case (but not generally) securitization - in particular *competitive* securitization – was one of the prime causes of the mortgage crisis, giving rise to informational asymmetries and principal agent problems.[70] After all, not all agents in the mortgage market had incentives to be honest with one another about the product (e.g. bad mortgage bundles).[71] Securitization involves pooling together debt, and often the pooling will disguise a bad debt. While bad debt might be buffered by being mixed with less risky loans, higher-level parties such as investors buying those securities are privy to less information than the original institution from which a loan originated, and thus they inherit the loan-specific risk embedded in the security. In order to combat this, a third party financial institution, which may be the government, is needed to guarantee those securities.[72]

Dislocations allow materials to bend without breaking (to some extent). What we do look for are structural features that dissipate energy (or in the case of economics, risk) in interesting

[70] Securitization is a process by which assets, like debts, get pooled together, and third parties can buy the cash flows resulting from interest and principal payments.

[71] What's a principle agent problem? In its barest form, it only involves two parties: a principal and an agent, where the agent can affect or act on behalf of the latter. However, he may have his own incentives and act in his best interests rather than the principal's. Thus there is a moral hazard problem, requiring additional incentive structure. Principal agent problems involve more asymmetries, such as an authority asymmetry, preference asymmetry, and so on. See also Ashcraft and Schuermann (2008) and Lang and Jagtiani (2010) for discussion.

[72] When there is less competition, securitizers have buyer power. But when securitizers lose that power and the originators of loans have supplier power instead, securitizers had incentives to participate in loose underwriting practices – the process of evaluating the liability of a borrower - in the race to issue as many loans as possible. In the US economy, government sponsored enterprises (GSEs) is the largest of these institutions that sets the underwriting standards. Stable economies in Europe have a high market concentration of securitizers – that is, a small number of firms control the industry overall – and fairly rigid underwriting practices. That means these institutions were less able to offer things like subprime mortgages.

ways. We can describe bubble bursts as the point at which the economy finally buckles after bending. Government policy is one good place to explore how we try to prolong the bending period, or how we try to ameliorate the economy in anticipation of external forces, and so on. The goals of exercises such as modern bank stress tests, after all, are to identify and remedy weakness, enhance resiliency to these large shocks, and prevent spread. Systems that contain such interesting structural features will behave differently from idealized cases, such as the box on the frictionless inclined plane or the container of ideal gas. In this way we depart from the assumption that good science is predictive science, since such analyses do not tell us what *will* happen but at best what *would* happen in specific circumstances.

3.5 CONCLUDING REMARKS

I have argued that equilibrium methods in economics are rationalizable, and importantly useful, idealizing strategies that have analogous uses in other sciences where they are not subjected to criticism (in particular in thermodynamics). Characterizing equilibrium states helps us characterize the system of interest by helping us delineate causal relationship between salient, scale-dependent variables. Philosophers of science have underappreciated equilibrium reasoning because they often assume that a final scientific theory should contain laws of nature that describe the most micro-scale entities of a system. Then, according to this picture, from those micro-scale details we can then determine the behavior of an aggregate over time.

On the other hand, this misrepresents actual scientific practice – both physical and social – because scientists often do not privilege the most micro-scale, nor do they aim to reduce analysis to talk of just one unique scale. The equilibrium method is most at home in a *multi*-scale framework, wherein we can still gain an understanding about the *dynamics* of systems even though we may use *static* methods with the aim to design systems that are in scale-linked equilibrium. This framework is thus especially useful for investigating whether a system has desirable properties like stability, making it apt for analyses of financial crises.

4.0 MODELING STABILITY

"Stability" can characterize a number of different things in economics. Sometimes equilibrium states are said to be stable; absent disturbances, they have no tendency to change. When we achieve *monetary* stability, inflation is being sustained at low rate without too many fluctuations. *Output* stability means output is growing at a constant rate. The *exchange rate* should be stable, since we would like to be confident in the values of currencies so as to facilitate trade. In this chapter we focus on *financial* stability – quite broadly, when our financial institutions (e.g. banks, and those entities that enable and facilitate the flow of funds) are functioning the way they ought to be.

> Broadly, financial stability can be thought of in terms of the financial system's ability: (a) to facilitate both an efficient allocation of economic resources—both spatially and especially intertemporally—and the effectiveness of other economic processes (such as wealth accumulation, economic growth, and ultimately social prosperity); (b) to assess, price, allocate, and manage financial risks; and (c) to maintain its ability to perform these key functions—even when affected by external shocks or by a build up of imbalances—primarily through self-corrective mechanisms. (Schinasi 2004, 8)

Stability in this sense is connected with the familiar notion of efficient performance, as the World Bank (2016) notes:

> Most [definitions] have in common that financial stability is about the absence of system-wide episodes in which the financial system fails to function (crises). It is also about resilience of financial systems to stress.
>
> A stable financial system is capable of efficiently allocating resources, assessing and managing financial risks, maintaining employment levels close to the economy's natural rate, and eliminating relative price movements of real or

> financial assets that will affect monetary stability or employment levels. A financial system is in a range of stability when it dissipates financial imbalances that arise endogenously or as a result of significant adverse and unforeseen events. In stability, the system will absorb the shocks primarily via self-corrective mechanisms, preventing adverse events from having a disruptive effect on the real economy or on other financial systems. Financial stability is paramount for economic growth, as most transactions in the real economy are made through the financial system.
>
> The true value of financial stability is best illustrated in its absence, in periods of financial instability. During these periods, banks are reluctant to finance profitable projects, asset prices deviate excessively from their intrinsic values, and payments may not arrive on time. Major instability can lead to bank runs, hyperinflation, or a stock market crash. It can severely shake confidence in the financial and economic system.

I have argued thus far that equilibrium methods are tools for outlining economic problems that primarily concern the efficient performance of a system in question, setting up the foundations for identifying relevant (and potentially manipulable) causal relationships. This chapter elaborates further what I suggested in Sections 3.4 and 3.5: that questions of economic stability should be understood within a larger, multi-scale framework, within which equilibrium methods are embedded. In this context, policymaking can be understood as constructing appropriate meso-scale mechanisms designed to keep the economy's behavior from fluctuating too wildly. We can (and economists do, at least implicitly) think of these mechanisms in the way that engineers design *control systems*, which we encounter everywhere, such as cruise control systems, autopilot programs, thermostats, and float and flush valves in toilet tanks.

4.1 STABILITY, EQUILIBRIA, AND AGGREGATION

How does *stability* figure into formal economic models? The minimal formal assumptions that equilibrium models make to imply just *existence* of an equilibrium state do not alone guarantee a *unique* equilibrium state. But we will not be concerned with global uniqueness, which requires stronger assumptions.[73] Multiple equilibria are more likely and realistic: "Under very general (but not universal) conditions, competitive general equilibria do exist" though "they are globally unique, however, only in quite special cases, although they are generally locally unique" (Fisher 1983, 33). Formally, an equilibrium price vector is locally unique if there isn't another one arbitrarily close to it; this means that equilibria only need to be "far away enough" from one another. We could still even use comparative static methods once we delineate the vicinity of a local vector to see how shocks might affect the market.[74]

Furthermore, we are not interested in just any equilibrium states, but rather *stable equilibrium* states.[75] If an economy is in an unstable equilibrium state, a disturbance may cause it to move to a new, very different equilibrium state.[76] On the other hand, if we're in the vicinity of a stable equilibrium state, we'll tend back towards equilibrium.[77] So stability is really a dynamical property quite different from uniqueness or existence; establishing the latter two implies nothing about stability.

---

[73] For instance, we have to rule out wealth effects – changes in demand due to the changes in values of goods. These occur when changes in a consumer's wealth affects their consumption patterns, because people tend to spend more money when they think that they are richer. Some kinds of commodities will therefore rise in price.

[74] There is a serious limitation to this, however. This only works in cases of no more than three variables. See also Hands and Mirowski (1998) and Kehoe et al (1989).

[75] A conceptual point relating to the concept of efficiency: the main results coming out of general equilibrium are the welfare theorems, which relate (Pareto) efficiency and equilibria. The first tells us that every equilibrium state is an efficient state – changes can't be made from that state without making someone worse off. The second one is more complicated, requiring the convexity assumptions mentioned here. Not every efficient state is an equilibrium state – after resources are allocated, wealth must be shuffled around within the agent population in order to support that allocation arrangement as an equilibrium state.

[76] For more remarks on dynamical stability, see Hicks (1946). He also states elsewhere that "Now the reason for this sterility of the Walrasian system is largely, I believe, that he did not go on to work out the laws of change for his system of General Equilibrium" (1939: 61).

[77] Though, we do not want to be in such a stable state that the economy cannot grow any further.

Additionally, stability will require us to look not just at one particular commodity like we did in the simplest partial equilibrium models, but larger, more comprehensive (one might call them *general equilibrium*) contexts. Investigating these dynamic properties will inevitably require looking the interactions between scales.[78] But asking this in the form of the question, "what do all the individual actors in an economy look like in a stable economy?" seems to lead nowhere. As mentioned earlier in Chapter 3, the Sonneschein-Mantel-Debreu (SMD) theorem is testament to this: under standard assumptions on agents, no interesting macroscopic properties are guaranteed to emerge from the population. That is, they don't give rise to any particular equilibrium vector, meaning that properties at the micro scale do not transfer to the macro scale. For instance, we could assign all agents a downward sloping demand curve, and yet we may not see that behavior in the aggregate.[79]

Perhaps we should work the other way around. Instead of trying to ascertain what micro-details give rise to what macro-details, we could try investigating some other scale. Notably, Grandmont (1992) and Hildebrand (1994) cite *population heterogeneity* as contributing to achieving stable equilibrium. Ackerman (2001) describes these as unlike general equilibrium methods in that they "attempt to deduce regularities in aggregate economic behavior from the *dispersion* of individual characteristics" (emphasis mine) rather than impose any restrictions at the level of individuals.[80] Hildenbrand (1994) in particular explicitly recognizes such aggregative

[78] In fact, since the 1990's some have taken the SMD theorem to motivate arguments against the microfoundations project, as for instance Rizvi (1994) and Kirman (1992) do.
[79] Kirman (1992) demonstrated that the SMD result can result for a population of almost identical individuals (the result depends on the way consumer preferences and their initial endowments are specified).
[80] See also Lewbel (1994) Rezvi (1997) and Kirman (1998).

properties as in some important sense *emergent*, rather than resulting from a straightforward transmission of properties from the micro to the macro scales.

> I believe that the relevant question is not to ask which properties of the individual demand behavior are *preserved* by going from individual to market demand but rather to analyze which new properties are *created* by the aggregation procedure.

Such distributional approaches impose constraints from the macro level, assuming only minimal conditions at the individual level. However, according to Ackerman (2002), both are unsatisfactory:

> First, it has not yet succeeded. The assumption of a smooth distribution of consumer characteristics seems to help, but has not entirely freed the proof of market stability from arbitrary restrictions on individual preferences or aggregate demand function.
>
> Second, even if the statistical approach were to succeed in explaining past and present market stability, it would remain vulnerable to future changes in preferences. Suppose that it is eventually demonstrated that the empirically observed dispersion of consumer preferences is sufficient to ensure stability in a general equilibrium model. This finding might not be reliable for the future since, in the real world, fads and fashions episodically reorganize and homogenize individual preferences. That is, coordinated preference changes involving the media, fashions, celebrities, brand names, and advertising could, in the future, reduce the dispersion of consumer preferences to a level that no longer guaranteed stability. (123-124)

Of course, the current distribution used for one's model will fail if those properties change. We are, at most, only committed to local predictive power. To know when stability arises for particular distributions of consumer preferences counts as something for us, even if it only tells us some conditions under which stability of a certain kind prevails.

But while Ackerman's second complaint should not worry us, the first one still should. Let's take a closer look at what these authors mean by "behavioral heterogeneity" and by "stability". Say every individual has a demand function f which assigns to each price vector p a bundle of commodities $f(p)$. Everybody has a distinct demand function, but say they all have the same amount of income. The mean aggregate demand at some price vector is:

$$f(p, w) = \int_F \mathrm{f}(p, w)\, d\tau$$

where τ is some Borel probability measure over the (infinite dimensional) demand function space F. We might take “behavioral heterogeneity” to mean that this probability measure has broad support (i.e. not concentrated on a subset of the demand function set). Hildebrand (1998) notices that because an infinite dimensional space is too unwieldy, Grandmont selects a finite parameter space for convenience. Call this finite-dimensional parameter space C, and let T be a mapping of C in F. A probability measure on C that induces τ on F is the image measure under mapping T.

Certain ways of picking the mapping T will yield a mean demand function that behaves very nicely. But it also means that the well-behaved nature of the function might not be, as Quah (2007) points out “related to heterogeneity (in some meaningful sense) of the distribution of demand functions,” but rather to the particular mapping chosen. For Grandmont, the result is that agents are different but not really all *that* different from one another. In this case, stability specifically means (expenditure) insensitivity to price changes. It’s questionable how this particular result could be practically helpful for us, e.g. in terms of policy-making.[81] It has merely snuck in the assumption at the ground level that households will generally behave the same way.[82]

---

[81] Hildenbrand’s 1983 results, though they jibe fairly well with empirical results, also require some narrow assumptions. He accepts minimal rationality conditions, mostly out of convenience, as well as that consumers had a common preference relation, that income distribution was independent of prices, and that the distribution of individual expenditures is decreasing.

[82] On the other hand, Giraud and Maret (2001), imply that the constancy of expenditure is a truly aggregative result. Unlike Grandmont, they employ a non-parametric method, and claim that the insensitivity of the budget share is due to a balancing out effect at the macro level rather than insensitivity at the micro level. That is, budget share expenditures remain constant because if one household decides to spend more, some other one will spend less, rather than it being the case that all households tend to behave in one way.

The aggregation difficulties I've described demonstrate just how hard it is to coordinate models at the micro- and more macro-scales in order to account for stability. A multi-scale model, however, could accommodate different levels of scale analysis without requiring one as derivable from the other by way of deduction, and may be a more suitable framework for analyzing stability. Aggregation is only straightforward in the strictest of scenarios, with extremely homogeneous agents, perfect information, and so on.[83] Interesting aggregate behaviors arise precisely because these assumptions at the individual level fail, no matter how useful they are in very limited cases. In economics, it's standard to assume a *representative agent* – an individual entity in the model that acts as a proxy for all individual actors. In a typical model only one such actor would be specified. Within a multi-scale framework we can admit different kinds of representative agents, each of which might represent the dominant behavior of its respective scale.

4.2 OBSTACLES TO EQUILIBRIUM BEHAVIOR

In material cases, dislocations prevented systems from reaching calculated equilibrium values. But taking into account defects allows us to make sense of a certain class of stable states – those regions of false equilibrium that many real world phenomena inhabit.

[83] And even in such cases not much of interest gets inherited by the aggregate. For every individual in the economy, we can construct an excess demand function. The aggregate excess-demand function only inherits the following traits from individual functions: continuity, homogeneity of degree zero, values of excess market demands are zero (Walras' Law), and a boundary condition that specifies that as prices go to zero demand increases.

In economics, something that gets called a "friction" is any impediment to the flows of capital, labor, services, etc. *Frictional unemployment* specifically refers to the systematic and normal unemployment that results from workers searching for or transitioning between jobs, and is by an large considered to be due to voluntary behavior.[84] More specifically, it is the unemployment that arises from the time delay that results from individuals moving between jobs. These are sometimes called search or matching frictions, and arise because labor and laborers are both heterogeneous groups. Unemployment does arise for other reasons, the other two cases being cyclical (due to business cycles) and structural (new technology, for instance, which would mean the demand for old technology declines). During recessions, people tend to hold on to the jobs they already have (so there's less frictional unemployment but likely more cyclical unemployment).

Economists also call *transaction costs* "frictions".[85] Trading on the stock market incurs a commission fee per trade, and the amount that I pay for a bond is its wholesale price plus what's called the bond markup (the difference between the price paid for by a broker and the price that I pay for it). Other such costs are harder to spot, such as bid-ask spreads and market-moving costs. The former refers to the difference between the highest price a buyer is willing to pay and the lowest price the seller is willing to accept. The latter refers to a market's reacting in anticipation of an agent (typically an institution rather than a particular individual) who is anxious to rapidly sell or buy a large block of securities (even before the trade is completed), driving the price

[84] This distinguishes such normal unemployment levels from cyclical unemployment (due for instance, by a cycling economy that may dip into a recession) or structural unemployment (where there is a supply/demand mismatch).

[85] See Coase (1937) for a precursory definition, though he did not use the term itself. He indicated that these costs were simply those associated with engaging in the price mechanism – the cost of partaking in the market.

lower or higher than it would have been otherwise and thereby paying a bit more for purchases or earning a bit less when selling. All these things impede perfect transactions. [86]

Other economic frictions are more obviously analogous to physical frictions. All economic processes are non-isolated dynamical processes, and any variable that does not react immediately to changes, even behaves as if *resistant* to changes, can be said to be *sticky*. Price stickiness, for instance, occurs often and in fairly mundane circumstances. One might find that firms are generally unwilling to change the prices of their goods even in cases where prices ought to move upward from a theoretical standpoint. For example, a firm may be unwilling to change the price of its product because then it must re-advertise its product at those new, higher prices – reaching the new equilibrium price will disruptively incur firm adjustment prices. And furthermore, if one firm predicts that the other firms won't raise their prices, that same firm has an incentive to stick with its current prices for fear of losing the market share. If I am one of several widget-sellers in Pittsburgh and the demand for widgets goes up, adjusting my prices would require me to change all my signs and advertisements (so-called menu costs). If my business is a small one, I might decide that this is not worth the trouble. But even if I could cover these costs without difficulty, my customers may change loyalties to competitors offering widgets at lower prices (and that effect, too, is limited by customer loyalty). In more recent literature, *information* has also been recognized as something that can exhibit stickiness.[87] At any given time, agents are probably considering information that is out of date, though less so today due to advances in technology.

---

[86] Transaction costs can arise *because of* incomplete markets, defined to be markets where the number of Arrow-Debreu securities is fewer than the number of states of nature.

[87] Such frictions interfere with efficient decision making in a larger variety of practical contexts, such as in choosing health insurance (Handel et al 2015).

Part of the reason economists failed to predict the last financial crisis is simply because their models were incomplete: models failed to account for the finance sector, and even the ones that did were quite far off the mark.[88] The economy is usually treated as having two dimensions: a "real" dimension and a "financial" dimension. The real economy deals with the exchange of actual goods and services, while the financial economy includes all those institutions that facilitate those exchanges. But the financial sector is not a ghost sector that stands in between exchanging parties, which was its intended purpose:

> Thanks to lax supervision and regulation, a significant part of the banking sector was no longer focused on its traditional role of serving the real economy – by mobilizing savings and channeling them in a cost-effective fashion to the most productive investment opportunities, and by managing associated risks. Instead, banks sought to develop an independent and separate status that would deliver lots of dollars...
>
> Even the common labeling of the industry itself changed – from "financial services" to just "finance." (El-Erian 2016, 34).

Many contemporary financial instruments had no representation of the financial sector at all.[89] But how bad a crisis gets can depend on whether it involves the financial sector; as Jorda et al (2013) point out that recessions that involve financial crises tend to be more prolonged with larger impacts than regular business cycle fluctuations.

And even when we do take into account the financial sector, the details get thorny. Often an economist will simply "graft additional so-called financial 'frictions' on otherwise fully well behaved equilibrium macroeconomic models, built on real-business-cycle foundations and augmented with nominal rigidities" (Borio 2012).[90] Not accounting for frictions in the right way

---

[88] Tovar (2009) blamed the failure on its reliance on the Ricardian equivalence: government financing, either via taxes or by debt, was to be treated the same way.
[89] See Landmann (2014).
[90] "Many of the recent DSGE models with financial frictions abstract from bank balance sheets altogether by modeling all lending as direct" (Benes et al 2014, 4).

will prevent us from understanding the contribution of nonlinearities to the cyclical behavior of the economy (Boissay et al 2013). Benes et al (2014) claim the more recent dynamic stochastic general equilibrium models

> …are not very useful for evaluating macro- prudential policy tradeoffs under such conditions, first because by construction they do not capture the effects of nonlinearities, and second because they ignore the special role played by banks in contributing to vulnerabilities and nonlinearities.

Relationships between aggregate sectors are not merely relationships between "bigger versions" of the constituents of those aggregates, e.g. individual households or firms, because institutions such as the financial sector get in the way.[91] Institutions and policies, which we understand as meso-scale structure, act as "filters" that policy effects pass through from the macroscopic level to the microscopic level.[92]

In very simple (ideal) materials, we can move from the micro- to the macro-scale of a system that has crystalline structure via averaging. But when dislocations multiply due to stress, the correlations between them increase, and to get information about the macroscale we have to

---

[91] Tobin (1972)

[92] And this goes for economics in general, not just the financial sector, and is indeed an international issue. In their investigation of Latin American policy, Rome and Caracas (2001) suggest that economy analysis often neglects such crucial structures in the same way that we have thus far:

> A consensus has emerged in the academic literature, and among national policymakers and multilateral agencies, that structural adjustment particularly during the 1980s, was not structural enough, as it failed to focus on some critical links between policy measures, intended objectives, and their actual outcomes. In Latin America, by focusing almost exclusively on price reforms ("getting prices right") and macro-economic variables, policy reforms tended to ignore important structural and institutional traits that have hindered the achievement of the alleged goals of economic growth and increased human welfare. The negative social effects of the reforms, for instance, have only recently been recognised by policymakers and started to be addressed through the design of safety nets for the newly unemployed and the newly poor.

do a bit more work at the mesoscale, e.g. by representative volume methods. Such methods allow us to coarse-grain a system by delineating chunks in the material. Each chunk is the smallest unit that can capture more macro-behavior.

We can easily move between the macro and micro scales via the representative agent if markets are *complete*; the problem is that markets usually aren't. In a complete market, there is a market for every good (in other words, there is an equilibrium price for every commodity at every possible state of the world). Complete markets are an idealization.[93] The actual future state of the world is uncertain from the perspective of the present, so we don't know what future states of the world will look like. Suppose, for simplicity, we only assume the following two kinds of commodities. One is a risk-free bond (which will give me some guaranteed amount no matter how the future turns out to be – think a fixed interest rate).[94] The other is another kind of asset, called a security, which allows us to insure ourselves against future events, which pays off say 1 dollar for every possible state, and nothing otherwise.[95] To insure myself against all possible future uncertainty, I would buy a security for every one of these future possible states.[96] To rule out arbitrage in the equilibrium state (with complete markets), the price of the bond would have to equal the price of the securities I pay for. This set of securities is a *portfolio* that acts as if I've

---

[93] What I describe here are tools from the *Arrow-Debreu* model of general equilibrium, which (like partial equilibrium models) we think about the market in terms of the satisfaction of wants. See Arrow (1951), Debreu (1951), and Arrow and Debreu (1954). One might conceive them as the formalization of Walrasian general equilibrium, though it's not clear whether Walras would find it satisfactory.

[94] So a risk free bond that pays X after a year at a rate of return Y would cost me X/(1+ .0Y) now.

[95] The securities need not have any dollar worth – it could be 1 unit of output, if we don't want to consider *money* at all, but instead focus on real exchange.

[96] The price now of such a security (for a particular state), would be the price divided by 1+r with r being my discount rate.

bought a risk-free bond. Given the market is complete and I can buy a security for any possible state of the world, I'm guaranteed a payoff of 1 dollar after the state of the world is revealed.

Brown et al (1996) claim that incorporating market incompleteness within the standard general equilibrium framework is a promising strategy:

> The general equilibrium with incomplete markets (GEI) model follows the standard Arrow-Debreu model in most of its methodological assumptions: agents optimize based on rational expectations, there is perfect competition, and markets clear. But the GEI model is much richer in its ability to describe and analyze important economic phenomena, such as financial innovation, the relevance of corporate control and financial policy, and the potential positive role for government intervention.

In a complete market, we are able to exploit the representative agent apparatus because the ratios of the marginal utilities of these commodities are the same for everyone, and individual behavior will track changes in the aggregate (i.e. individual consumption is proportional to aggregate consumption). In a complete market model we can prove the existence of the equilibrium state in terms of this representative agent. But this is not so when you have incomplete markets.

An incomplete market network doesn't absorb shock the same way a complete market can. The better connected a bank is (and thus is able to interact more easily with a number of other banks), for instance, the better it can mitigate idiosyncratic losses particular to its capital, though at the same time interconnections might make the network prone to *contagion.* Allen and Gale (2000) contrast "complete" with "incomplete" bank networks, where the former involves banks in symmetrical relationships with one another and the latter involves banks in such relationships only with neighbors.[97] They find that financial contagion is less damaging in complete networks for the very reason that risk is spread out among many banks (given that the

[97] There is a third kind not mentioned here, which involves incomplete networks with entirely separate banking networks.

topology of the system is homogeneous), though Nier et al (2007) show this also that connections may make an individual in the system more susceptible to contagion.[98] In the case of financial crises, what we want to do is curtail contagion effects, i.e. we have to be able to isolate risk or alter the way it disperses throughout a system.[99]

So we can understand financial stability by how prone a system is to *systemic risk* – the risk that an entire system might collapse due to how susceptible it is to contagion effects. We need to be able to model richer non-homogeneous incomplete structures and include friction-like forces without resorting to the bottom-up reductionist modeling strategy that I have rejected.[100] Looking to how physicists deal with the thermodynamics of solids may provide a good guide as to how to proceed. Recall that understanding the elasticity of a bulk material like steel requires analysis at different scale levels. Some crucial structural features only appear at meso-scale levels which we cannot assume from knowing the details of the lower level scales. This is true as well of a rubber band. Young's modulus, the ratio of stress to strain on a material for this particular kind of deformation (other kinds of deformations correspond to other kinds of moduli, e.g. the shear modulus) describes the stiffness of the material in question. But to describe this kind of property, we have to take into account these meso-scale dislocations.[101] Analogously,

---

[98] There are different results on whether integrated networks of banks amplify or absorb shocks.
[99] For a discussion on how to define contagion, see Kaminsky and Reinhart (1998) and Gregorio and Valdez (2001)
[100] For similar reasons, we should also be wary of top-down approaches.
[101] Strategies that try to work from the bottom up (or even continuum top-down methods) would get things wrong:

> Such a methodology tacitly presumes that, in fairly short order, temporal jiggling will break up the dislocation structures and release their higher scale energy into the lower scale lattice. Because this doesn't happen with dislocations, molecular scale equilibrium calculations often give poor results (Wilson, forthcoming).

economic models that do not account for the fact that higher-order structures such as institutions interact with the lower-scale (heterogeneous) individuals in non-trivial ways will give bad predictions as to how the economy will behave.[102] [103]

4.3 CORRECTIVE MEASURES AND CONTROL THEORY

Like the physical and materials sciences, economics is also an enquiry in how to best construct systems with particular desirable features. Sub-disciplines such as *market design* aim to ameliorate the problem of incomplete markets by constructing a market so that the relevant transactions to take place (satisfying some desirable properties, like fairness). Another such sub-field of economics that is also motivated by interventionist purposes is that of (computational) *mechanism design,* a branch of game theory, where often the goal is to create incentives in the market to bring about better results that are currently possible. For example, one could try to design incentive structure that encourages market participants to reveal private information. These kinds of maneuvers could be interpreted as attempts to make incomplete markets *more* incomplete.

In this section I suggest that economic theory, including the usage of equilibrium idealizations, can be interpreted as functioning as a *control theory*, which attempts to determine the dynamic behavior of systems that take inputs and produce outputs (and are subject to

[102] See Peters (2004) for a discussion of frictions and the mesoscale in the context of physics.
[103] There's no obvious observable distinction between being in a world that is currently changing equilibrium states and one in which behavior is simply fluctuating around equilibrium points, though one is evolving due to external forces and the other due to internal frictions working themselves out.

feedback effects). These control projects are typically aimed at the meso-scale, with a view to achieve stability at other scales (e.g. the banking sector can absorb shocks and prevent them from fully reaching individuals, which means more macroeconomic stability).

Control theory, the analysis of control systems, originated as a field of applied mathematics and was considered to be a fairly autonomous discipline in the 1960s.[104] Control systems manage the dynamical behavior of other systems – for example, cruise control systems help manage the behavior of the automobile by smoothing out fluctuations around a target speed. The classic example of a control system is the centrifugal governor, the precursor to today's cruise control systems. Originally found in steam engines, the centrifugal governor helped maintain a vehicle's speed.[105] Driving an automobile without a governor would require us to manually adjust how much fuel was injected into the engine every single time road conditions changed. Every adjustment, however, comes along with a risk of overcompensation.[106] This was precisely the problem encountered by steam engine operators during the industrial revolution. When the load on the engine increases, it will begin to slow down, but an injection of fuel would cause it to swiftly increase speed. One might let in too much fuel, and the engine would oscillate as a result. The system in which the governor is embedded then reliably behaves within some pre-specified window, even when external circumstances change.[107] While we no longer use steam engines, the basic principles of centrifugal governors can be found in today's cruise

---

[104] Control theory is the study and design of control systems, where by system we actually mean differential equations (so they may not characterize a physical system).

[105] Feedback is an informational notion: part of the information about a system's output figures into its input. Alternatively, it describes the coupled dynamics of two systems that influence each other.

[106] Not all hysteretic behaviors are desirable.

[107] Auyang (1998) points out that we can also talk about patterns of regularities by classifying the control parameter values into different categories according to what responses they elicit.

control system, which keep the velocity of a car constant despite encountering disturbances, such as air resistance and changes in slope of terrain. These machines can self-regulate their behavior by comparing actual outputs with target outputs and adjusting their behavior accordingly.

Economic institutional measures act like governor-like mechanisms at the meso-scale that keep an economy from spinning out of control in response to external conditions. Their job is to maintain the stability of the economy, often by intervening in a structural way. The automatic stabilizer is a policy that acts as a negative feedback loop. This is government policy that kicks in without direct intervention from policymakers, usually in the form of taxes (e.g. income tax) or transfer systems (e.g. unemployment insurance) to offset changes in gross domestic product (GDP). A progressive income tax will affect different people in different ways, either raising their tax rate or lowering it relative to income earned, such that higher earners pay a higher amount and lower earners pay a lower amount of taxes. The financial sector that we discussed in Chapter 3 (in particular Section 3.4) serves as something like meso-scale structure divvying up the economy (say, to prevent the spread of a shock to the economy) in the way lines of dislocations will demarcate grains of steel from one another.

4.4 MEANWHILE, IN ECONOMICS

4.4.1 DSGE Models

Consider one particular contemporary economic model, the *dynamic stochastic general equilibrium* (DSGE) model, and some of the ways that our worries bear upon it. The development of the contemporary DSGE model (and financial modeling post-crisis in general) embodies these very methodological struggles that we have discussed so far, which are quite general to *complex adaptive systems*, a system composed of many interacting parts that will react and adapt to its environment.[108] The DSGE was intended to allow flexibility in modeling, including meso-scale features such as institutional structures (which may be specified exogenously, and include budget constraints and monetary policy, or may be endogenous to the economy).[109] Given that it is a dynamic model, it explicitly recognizes the intertemporal nature of the behavior it models. It is also what we call a "microfounded" model, aiming to explain economic behavior in terms of individuals that are optimizing, rational agents (i.e., representative agents).[110] [111]

The DSGE model has its origins in a tension between Classical economic thought and (New-)Keynesian economic thought regarding abnormal market behavior. The former, which are associated with the theory *real business cycles*, sees such behavior as rational reactions to exogenous shocks. So periodic expansion and contractions are actually the economy reacting rationally. The older DSGE models had no role for monetary policy at all, and seemed to aim are pure description by aiming at calibration rather than estimation.[112] Keynesians, however, diagnose such behavior as a market failure of some kind – namely what I've called *frictions*.

---

[108] At least, it is used by most US institutions. European institutions use the cointegrated vector autoregression (CVAR) approach, which I do not address here

[109] For instance, the Christiano et al (2011) DSGE model for evaluating monetary policy includes the caveat the aggregate output can't be straightforwardly implied by aggregate input and technology, since aggregate output would depend on how resources were distributed across firms.

[110] The New Keynesian version consists of an aggregate demand, aggregate supply, and a Taylor rule. Its real business cycle counterpart considers capital, labor, consumption and technology, with its setup involving an expected utility maximizing household.

[111] These two schools substantially differing only in how to treat aggregate *supply* in particular.

[112] More seriously, though, if the structure parameters of the model are wrong than no amount of calibration would help except by sheer coincidence.

Despite that these two schools seem to disagree on what the actual economic facts are, since the 1990s and 2000s macroeconomics has reached some kind of consensus, the “new neoclassical synthesis” (Duarte 2012). The DSGE hybrid model is thus also called the “New Consensus Model”.

There are three key ingredients in a standard DSGE model: a supply block, demand block, and a monetary policy equation. The monetary policy equation takes inputs from the supply and demand blocks. This equation describes how central banks set nominal interest rates, raising it (e.g. when inflation occurs) and lowering it (during economic lulls) as necessary. Lowering the interest rate encourages consumption, though too much subsequent demand risks inflation. This might then require raising the interest rate. Because each block is treated as a representative agent, and thus its behavior is modeled explicitly, we say that this model is microfounded.

But the DSGE model was heavily criticized in particular after the Global Financial Crisis due to its *failure* to predict the crisis.[113] There is no consensus as to how helpful the DSGE model is insofar as it is treated as a forecasting tool, and there are a plethora of reasons why people think it fails. Sometimes we get assumptions external to the DSGE model itself wrong. For instance, simulations typically apply shocks that fall in a Gaussian or normal distribution when instead economic time series data seem rather to be characterized by fat tails, meaning that more extreme events happen more often that is often assumed in models (Ascari et al 2008).[114]

---

[113] Of course, professional opinion on its failures are subtler – for example, a post-Keynesian may be dissatisfied with how thee model treats endogenous money in particular.
[114] Newer, state-of-the-art models, i.e. versions of the Smets and Wouters (2007) model, are used by Central Banks and fare better along some dimensions.

But other problems seem to be endogenous to the DSGE model, especially due to the heterogeneous nature of the economy.[115] A standard DSGE model assumes complete markets.[116] But as I've pointed out, incomplete – or missing – markets, make the picture more realistic (and more difficult), and indeed one could cite incomplete markets as contributing to the crisis.[117] In incomplete financial markets, risk may get distributed inefficiently because some banks bear more risk than others and may not be able to unload it under stress.[118] [119] The policymaker wants to guard against systemic risk, so institutions like central banks and regulatory measures themselves taken to buffer the market against shocks can be interpreted as ameliorating the shortfalls of the incomplete market. One way it might do is is by trying to "complete" the market by absorbing shocks in the appropriate kind of way, namely by providing liquidity. Other banks may be unwilling to interact with another one if it is bearing a particularly heavy load of risk by declining to provide funds. Even systematic regulatory checks on the financial system can be interpreted as governor-like, in that they regularly check the international economy for weakness and providing recommendations as to how to improve.

---

[115] See Sbordone (2010)

[116] Those that do not assume complete markets are, however, present in the literature and have been gaining traction since 2005.

[117] See for rejections of the assumption of complete markets Attanasio and Davis (1996), Benhabib et al (2013), Fisher and Johnson (2006) and Jappelli and Pistaferri (2006).

[118] In economic modeling, incomplete markets can be modeled either exogenously or endogenously. The first route specifies the details of what securities agents can and do trade. The standard incomplete markets model involves consumers being able to only purchase one kind of (non-contingent) bond with borrowing constraints. This means that the incompleteness is imposed exogenously. On the other hand, we can model incompleteness as arising endogenously; that is, incompleteness can be derived from more fundamental frictions rather than imposed from the get-go. This typically seem to arise when agents possess private information or there are transaction costs involved. See Kocherlakota (1996).

[119] Buiter (2009) criticizes DSGE use for the assumption of complete markets, bemoaning the "The unfortunate uselessness of most 'state of the art' academic monetary economics."

For example, The IMF began instituting regular stress tests after the Asian crisis in the 1990s. In the US, for instance, the Fed runs the Comprehensive Capital Analysis and Review, or the CCAR, which is conducted annually. The CCAR tests the robustness of institutions, and the Fed might suggest revising stock distribution plans in light of the test results. The Fed surveys large complex bank holding companies (BHCs) with more than 50 billion in assets and the methods they use to assess their own capital needs (and even methodology documentation such as model validation). In 2014, the Fed assessed (in a second kind of test) 30 BHCs, emphasizing qualitative improvements rather than just establishing that a BHC was sufficient with enough quantitative capital. They were asked to have forward-looking capital planning plans that anticipate stresses of differing degrees, including not just a baseline case but severe (and extremely severe) cases. The Federal Reserve expects these BHCs to have sufficient capital to withstand a highly stressful operating environment and be able to continue operations.

The CCAR tests have both a qualitative and a quantitative component, extending further than the standard BHC tests, which tests for adequate capital and delivers pass/fail marks. The CCAR additionally test for what banks will do with their capital for shareholders (e.g. pay out dividends or buy back stocks from shareholders). The Fed requests that banks submit *capital action plans*, looking to see whether the bank could function not just under stress."[120] Each bank had to be able to handle its own idiosyncratic risks. The Fed examines the capital adequacy of the bank, its internal assessment processes, and its individual plans. Afterwards, banks will publish their results on how distribute capital over the next four quarters (the message to the public meant to be a reassuring one). If, however, a bank is too weak or there are calculation

---

[120]One external commentator (Bible 2013) further recommended these be "robust forward-looking capital planning process that accounts for their unique risks."

errors spotted, it may have its dividends capped. In 2013, Federal Reserve Governor Daniel Tarullo stated:

> … the Federal Reserve's review of capital plans provides a regular, structured, and comparative way to promote and assess the capacity of large bank holding companies to understand and manage their capital positions, with particular emphasis on risk-measurement practices…The financial crisis showed not only that regulators needed to increase capital requirements and conduct regular stress tests, but also that firms need strong internal processes to evaluate their own capital needs based on their individual risks and circumstances.

For instance, a bank may try to insure itself against shocks by hoarding capital; this is, however, bad news for other banks that may be interconnected to it. So a measure of just an individual firm's susceptibility to risk could be informative to the policymaker. *Macroprudential* measures, aimed at the overall welfare of an economic system, as well as *microprudential* ones, aimed at individual institutions like banks, are necessary for the functioning of the economy. A comprehensive analysis requires examining both kinds of measures in order to account for the interrelationships between institutions such as banks.[121] [122] And it is these relationship that matter when "the main cause of systemic collapse is endogenous risk" (Brunnermeier et al 2009, 59), and so those relationships are also the targets for policy.

4.4.2 Heterogeneity and Meso-Structure

How exactly we can measure and model systemic risk is still work in progress. Newer modeling approaches, in particular network theory, has offered some tools to help capture structure detail

[121] Of course, some may advocate that whenever a bubble is detected that we ought not *pop* it – this is the Jackson Hole view.
[122] This is sometimes referred to as *granularity* (Bremus and Buch (forthcoming) and Gagliardini and Gourieroux (2015)) as a response to the nonlinear behavior of financial systems.

that may be missed by representative agent type methods that may over-homogenize a sector. Acemoglu et al (2012) recognize that "We show that, as the economy becomes more disaggregated, the rate at which aggregate volatility decays is determined by the structure of the network capturing such linkages," indicating a clear multi-scale interpretation. As Gabaix (2011) points out, "any economic fluctuations are not, primitively, due to small diffuse shocks that directly affect every firm. Instead, many economic fluctuations are attributable to the incompressible 'grains' of economic activity, the large firms…idiosyncratic shocks to large firms have the potential to generate nontrivial aggregate shocks that affect GDP, and via general equilibrium, all firms." Firms experience shocks differently, but these idiosyncratic shocks will affect aggregate behavior.

Despite the bottom-up appearance of network models, they have allowed us to gain some understanding of important higher-order features (called *topological* features). That a bank is larger than another might be interpreted as being more interconnected than others. This interconnectedness only appears at higher scales and that could be understood as a meso-scale structure (i.e. as a "risk hub").[123] Some works explicitly identify a *clustering coefficient* (e.g. Fagiolo 2007), but such coefficients rely on a bottom-up construction of the network.[124] [125] Other

---

[123] Indeed, matching markets (and measures to create matching markets) are one interventionist method to try to improve the market performance namely by providing a conduit for information exchange.

[124] The clustering coefficient is a measure of how likely a node's neighbors are in a group, e.g. connected to one another.

[125] This may be unsatisfactory, as the data regarding bilateral relationships between banks is not readily available. Network theory itself and its extensions into other sciences, notably biology, are still in development: "there is no procedure that enables one to simultaneously assess whether a network is organized in a hierarchical fashion and to identify the different levels in the hierarchy in an unsupervised way" (Sales-Pardo et al 2007). Second, some of the more developed multi-scale models are actually *multi-layer models* where a "node" inhabits several

developments in statistics also look quite promising, in particular multilevel and random/fixed effects models that treat data as explicitly hierarchical.[126] Ahelegbey and Giudi (2014) propose a hierarchical Bayesian graphical model, whose purpose is to discern *what* the upper-level structure (connections between countries' banking sectors) is given a network of individual banks. In the case of the financial crisis, epicenters of these hubs are where artificial measures such as bailouts or capital controls are effected. We would target the node in a network with the highest eigenvalue centrality – its measure of importance in a network. Individuating the representative constituents at this higher level would also at the same time represent obstacles where frictional effects (i.e. regions where liquidity transfer would be problematic) would originate.

The trajectory of the development of financial modeling is motivated by many of the same reasons multiscale models are motivated in physical sciences. Additional recent work ranges from examining whether multi-scale structures exist in financial time-series (see Ferreira et al 2006), to more explicitly importing statistical mechanical tools such as hierarchical modeling and mean-field approximation approaches (Aoki 1994). Given that we have explored parallels with classical thermodynamics, it might seem natural that progress in economic modeling should start taking inspiration from statistical mechanics.

Thus far, there is some promising work in financial time series modeling, especially since the behavior of financial systems seem to be unpredictable precisely due to their nature as nonlinear complex systems. For example, one commonly accepted stylized fact is that stock

---

different kinds of communities, yet our multiscale goal involves the inhabitants of scales to be in some sense distinct kinds of entities.

[126] Though they may be subject to the aggregation problem that we are familiar with as the "ecological fallacy" (Gelman 2001).

market returns exhibit power law behavior. Given general similarities between financial and physical systems – both are systems that involve many interacting heterogeneous constituents that exhibit nonlinear behaviors – we might be tempted to draw a further analogy. Perhaps, as Sornette (2003) and Gabaix et al (2008) suggest, financial systems exhibit critical behavior in the way physical systems do, so we might be able to regain predictability by thinking of salient phenomena such as crises as akin to phase transitions like the spontaneous magnetization of ferromagnets below a critical temperature. This project would usefully intersect with the more recent literature adapting network theory to economic contexts, since both can then characterize interesting behavior that arises endogenously through coordinated panic behavior (in the case of statistical mechanics, this corresponds to correlations becoming infinite near critical points). In particular, it may be useful to employ such methods in order to constrain the plausible micro-theories that we may be interested in formulating. Johansen et al (2008) and Sornette (2013) suggest that the financial data behaves a particular way akin to critical phase transitions near the times of crashes in a way that narrows down the possible network structures of agents that might give rise to such behavior. And using models at multiple scales, in conjunction with empirical data, will help insulate the enquiry from another difficulty: financial time series may simply *appear* as if they are following power law behaviors. But if the modeling strategy requires active negotiation between candidate macro- and micro-/meso- theories, then the enquiry seems much less speculative than eye-balling some data.

One could ultimately worry that we would *not* be justified to borrowing methods from physics, because the kinds of systems under consideration are fundamentally different. For example, one might argue that economic systems and physical systems are simply too different for the most salient parallels to actually be exploitable, and that attempts to do so are really a manifestation of physics envy. Lo and Mueller (2010), suspect that the kind of uncertainty present in physical systems is simply different in kind from the kind present in economic systems. Lo and Mueller claim "economic logic goes awry when we forget that human behavior is not nearly as stable and predictable as physical phenomena."

The reason for this is because of the kind of uncertainty inherent in economic systems. *Uncertainty* in economics is a technical term that contrasts with *risk*. When I flip a coin (and say, bet a dollar on heads), I face a scenario that is risky – I don't know which outcome will actually result, but I do know what the possible outcomes are and the probability distribution over those outcomes. However, much of my everyday life consists of *uncertain* decisions – decisions where I know even less than in the risky scenario. Sometimes, with enough data or with some mathematical trickery, I can reduce the uncertainty to risk. Many economic scenarios of interest contain uncertainty that cannot be reduced to risk – Lo and Mueller refer to these kinds of risk as:

> [S]ituations in which there is a limit to what we can deduce about the underlying phenomena generating the data. Examples include data-generating processes that exhibit: (1) stochastic or time-varying parameters that vary too frequently to be estimated accurately; (2) nonlinearities too complex to be captured by existing models, techniques, and datasets; (3) non-stationarities and non-ergodicities that render useless the Law of Large Numbers, Central Limit Theorem, and other methods of statistical inference and approximation; and (4) the dependence on relevant but unknown and unknowable conditioning information.

Changes in rules and other behavior irregularities as well as the possibility that we might need multiple models to capture the data-generating processes might give rise to this kind of uncertainty. Other sources of uncertainty may be impossible to anticipate. In the face of such radical uncertainty, which strains the validity of our models, we need:

> …to develop a deeper under- standing of what is going on and to build a better model, which may be very challenging but well worth the effort if possible … There are no simple panaceas; as with any complex system, careful analysis and

> creativity are required to understand the potential failures of a given model and the implications for a portfolio's profits and losses.

We need to determine *which* features that contribute to radical uncertainty can be consolidated into a manageable risk. But "the other response is to admit ignorance and protect the portfolio by limiting the damage that the model could potentially do," which they describe as falling within the ken of *risk management*.[127]

According to these authors, the assessment of economics is that the field has not made the same kind of scientific progress that physics has, but that's because the systems that it investigates are much *more* complex. But some of the problems that Lo and Mueller pinpoint – nonlinear behavior, for example – is also typical of some physical systems (weather is one obvious example). Some of the methods in complex systems analysis are adapted to fields such as biology which do face such difficulties. And furthermore, their recommendation – "to delineate the domain of validity of each model" – is typically how economic methodology has been motivated anyway, even in its most fundamental form (i.e. even the equilibrium concept). In particular, simpler cases of risk are used strategically to manage more complicated cases of uncertainty.[128] [129]

---

[127] And there are corresponding technical maneuvers to make when it comes to formal modeling:

> …stop-loss policies that incorporate the risk of ruin or franchise risk; limits on capital commitments and positions (to prevent "outlier" signals from making unusually large or concentrated bets); statistical regime-switching models that can capture changes in the profitability of the strategy; and Bayesian decision rules to trade off "Type I and Type II errors.

[128] Another criticism, similar in spirit, is Bookstaber (2010), who pinpoints the reactive and adaptive nature of agents as preventing a robust parallel to physical systems. But while he, too, recognizes that economic systems are interactive, it's that these interactions are distinctively *intentional* that they fail to be captured with current quantitative methods. But his suggestion is that "A better analogy than physics or biology is a military one." This is puzzling, because interactive, intentional, strategic behavior has been studied for some time as game theory. And,

4.4.3 Policy Design as an Optimal Control Problem

Notice that many of the problematic cases in economics, including those that involve modeling feedback, are problematic because they are not isolated systems. Because the policymaker is interested in system regulation, it may be plausible to think of policy problems as *control problems* in the way that engineers do in mechanical, engineering, and now software and computing contexts. In particular, given the nature of equilibrium reasoning, we ought to be thinking of problems as *optimal control problems*. Optimal control theory is a more advanced method for solving dynamic optimization problems that can be adapted to various disciplines other than physics. When we tackle an optimal control problem, we manipulate the dynamical behavior of a system via another control system in a way that *optimizes* some feature of that behavior over time – mainly via some kind of cost minimization. For instance, when we want to

---

in particular, newer evolutionary methods are attempts that try to elaborate how it is that interesting behaviors (retaliatory, cooperative, etc.) might arise.

[129] What about agent-based computational economics (ACE), which are sometimes touted as alternatives to the DSGE model (and on top of that, immune to the Lucas Critique)?[129] These specify the rules that agents obey over time, and we let the system evolve on their own after specifying some initial configuration of agents. While there has been recognition that there is interaction from the macro to the micro, since aggregate features will influence individual behavior, such structures should not necessarily be posited a priori. But ACE methods only involve the specification of initial (and boundary) conditions with no further intervention from the experimenter, it is hard to see what kind of story is actually being told about feedback, so in itself it does not address Bookstaber's concerns. The hope is that a multi-scale model could successfully reap the benefits of ACE, which endeavors to describe emergent macrobehavior from interactions between heterogeneous agents, without having to commit to "bottom-up culture-dish approach to the study of economic systems" (Tesfasion 2002, 1).

launch a satellite into orbit, we would like to do so in a way that gets the job done in a way that maximizes fuel efficiency and minimizes travel time.[130]

The interpretation of economic theory as optimal control theory is not new. It has been applied extensively to the problem of economic growth, and in particular with respect to optimal savings and consumption, with Ramsey (1928) being one of the first to explicitly employ the calculus of variations in order to calculate the optimal rate of savings. Dorfman (1969) claimed outright that "[optimal] control theory is formally identical with capital theory, and that its main insights can be attained by strictly economical reasoning."

Some well-known criticisms of the plausibility of control theory as a macroeconomic method tend to oversimplify matters. For instance, Kydland and Prescott (1997) worried that optimal control theory treated variables, which were in reality not invariant, as invariant:

> [O]ptimal control theory is an appropriate planning device for situations in which current outcomes and the movement of the system's state depend only upon current and past policy decisions and upon the current state. But, we argue, this is unlikely to be the case for dynamic economic systems. Current decisions of economic agents depend in part upon their expectations of future policy actions. Only if these expectations were invariant to the future policy plan selected would optimal control theory be appropriate in situations in which the structure is well understood, agents will surely surmise the way policy will be selected in the future. Changes in the social objective function reflected in, say, a change of administration do have an immediate effect upon agents' expectations of future policies and affect their current decisions. (474)

Their result is that employing optimal control theory would, at best, lead to consistent solutions that are sub-optimal. This is especially so in cases of moral hazard: we would rather not find ourselves in the scenario where we would have to subsidize a bank's bailout, but refusing to do

---

[130] As an aside, there are newer applications of optimal control theory in biology (Lenhart and Workman 2007) and it also has precedents in medicine (e.g. Inoue et al 1976).

so when a bank does happen to fail might lead to even more disastrous consequences. So a time-consistent policy might achieve the policymaker's objective (and be better for society overall) in the long run, but might make people unhappy in the short-run (e.g. bailing out a failing bank in the midst of a crisis). It's only when policies are credible that there is no information asymmetry between the policymaker and the public that discretionary policy from a central bank will be effective. The alternative, they suggest, is to stick in general to monetary *rules* where a credibility problem won't arise because the regulatory conventions are already pre-established. If, however, the policymaker had knowledge that other agents did not, and the agents were restricted to very simple behaviors (i.e. simply adopting the policy that was recommended): "There is *no* way control theory can be made applicable to economic planning when expectations are rational" even when there is a "well-defined and agreed-upon, fixed social objective function" (473). That is, after the announcement of a policy recommendation, we are in a *new* economic setup and must start the Sisyphean analysis over again.

Despite this, in 1980 they suggest that "In spite of the time inconsistency problem, we do not think the determination of the optimal policy and the resulting return [from taxes] is without interest. The optimal return is a benchmark with which to compare the return of the time consistent policy under a particular set of institutional constraints" (80). This is entirely consistent with understanding of equilibrium reasoning that I proposed in Chapter 2 and the way in which I envision multi-scale modeling in the economic context. In fact, they further suggest that "In environments where policy rules can be changed only after an extended period of public deliberation, the time consistent rule may be nearly optimal" (90). This can be interpreted in the following way: a state that involves the dissemination of information can, like a stressed material, need to undergo some (possibly characteristic) "relaxation time" before ultimately landing at a new equilibrium state (or, approximately so). The implicit assumption is that the policymaker's problem is one that spans more than one temporal scale, and if so her challenge is not only to manufacture an effective policy but to find out at which scale is appropriate for its implementation.

More recently, especially during the last decade, the Fed has taken moves to use optimal control methods as a tool in formulating monetary policy. In her 2012 address at the Haas, Janet Yellen explicitly invokes optimal control methods in describing how the Fed might achieve long term target goals of low/stable inflation and unemployment and at the same time handle short-term fluctuations: "A high priority for the Chairman was to further clarify the FOMC's interpretation of the long-term objectives implied by its dual mandate to promote maximum employment and stable prices." This project involves looking at the problem from two time scales, since there are "two tasks for monetary policymakers: first, setting policy to pursue the longer-run objectives; and second, adjusting policy in response to shocks to minimize shorter-run fluctuations around those objectives."

In the past, many central banks often only targeted one (inflation), whereas Yellen suggests that we must incorporate both. This means that we must keep the following in mind:

> The first is that, if the FOMC is doing its best to minimize deviations from its objectives, then, over long periods, both unemployment and inflation will be about equally likely to fall on either side of those objectives. To put it simply, if 2 percent inflation is the Committee's goal, 2 percent cannot be viewed as a ceiling for inflation because that would result in deviations that are more frequently below 2 percent than above and thus not properly balanced with the goal of maximum employment. Instead, to balance the chances that inflation will sometimes deviate a bit above and a bit below the goal, 2 percent must be treated as a central tendency around which inflation fluctuates. The same holds true for fluctuations of unemployment around its longer-run normal rate.
>
> The second property, which to me is the essence of the balanced approach, is that reducing the deviation of one variable from its objective must at times involve allowing the other variable to move away from its objective. In particular, reducing inflation may sometimes require a monetary tightening that will lead to a

> temporary rise in unemployment. And a policy that reduces unemployment may, at times, result in inflation that could temporarily rise above its target.

Indeed, modeling policy quite explicitly as a control system has been anticipated for some time. Aoki and Nikolov (2006) suggest incorporating an *integral term* in a policymaking equation, akin to what one finds in equations for proportional-integral-derivative (PID) controller.[131]

> We argue that the optimal price level targeting rule performs best because it has elements of integral control, in the sense that the policy reacts to the integral of past deviations of inflation from its target value. Integral control terms can reduce the propagation of policy mistakes in two ways. First, when mistakes are persistent (due to, for example, persistent mismeasurement of the natural interest rate), the integral terms can reduce the impact effects of the policy errors. Second, history dependence in optimal policy rules generate endogenous persistence in equilibrium dynamics that is independent of persistence in exogenous disturbances, as emphasised in Woodford (1999). This endogenous persistence may work as a propagation mechanism of policy mistakes for certain policy rules. Policy rules that have elements of integral control terms can reduce this endogenous propagation.

An example of *proportional* control would be the toilet valve or the centrifugal governor. Departures from some target speed would elicit a proportional response from the control system to decrease it. *Proportional-derivative* control systems incorporate a derivative, also known as a *damping*, term. This term can be thought of as reducing the jitteriness of the behavior as system tries to adjust to a new target. The *integral* term reduces steady-state error – the buildup of errors from past behavior.

[131] We could take the elementary governor discussed in 4.3 and add in a bit more because it turns out that even the governor that did not adjust speed until a certain threshold was reached may still a bumpy ride: it was simply just putting bands around target speeds meant that when we did throttle the engine, though it happened less frequently the more “lurching” there would be. In order to combat this, we could add an integral term that would smooth out those lurches.

5.0 CONCLUDING REMARKS: WHAT DOES THE LUCAS CRITIQUE CRITIQUE?

> This essay has been devoted to an exposition and elaboration of a single syllogism: given that the structure of an econometric model consists of optimal decision rules of economic agents, and that optimal decision rules vary systematically with changes in the structure of series relevant to the decision maker, it follows that any change in policy will systematically alter the structure of econometric models.
>
> (Lucas 1976, 41)

I have argued that philosophers and economists are often mistaken about the methodological goals of economic idealization, especially the use of the equilibrium concept. Even its historical development points at a goal that is different from *tout court* prediction, which is consistent with the goal of even today's economic models such as the ones that central banks use to assess and implement monetary policy, despite their shortcomings. Clarifying the role of the equilibrium concept in economics also helps clarify the way it is employed in other sciences, suggesting that the dominant view in philosophy of science is overly restrictive in what qualifies as science (and scientific progress).

I want to conclude with a few remarks on how our investigation bears upon the *Lucas Critique*, a historical landmark in the development of contemporary econometric modeling. Econometric forecasting *is* concerned with prediction. But how to achieve it? In 1976, Robert

Lucas argued that data-based descriptive large-scale macroeconomic models are not suitable for evaluating the effects of a new policy, since "any change in policy will systematically alter the structure of econometric models" (1976, 41). Forecasting models could not be expected to be forward looking if the method by which they established variables were backward looking! If the underlying structure of the economic system changed whenever people incorporated their knowledge about upcoming policy implementation, then one's model essentially relies on obsolete presumptions and would give inaccurate results.

So much for trying to compare the potential effects of different policies. In order for such an exercise to be useful, we would have to assume the stability of the equation describing the dynamical behavior of the economy. This means that "agents' views about the behavior of shocks to the system are invariant under changes in the true behavior of these shocks" (Lucas XX) If we consider a state equation describing the economy that has the following functional form:

$$y_{t+1} = F(y_t, x_t, \theta_t, \varepsilon_t)$$

y_t is a vector of state variables, x_t is a set of exogenous forcing variables, θ is a fixed parameter vector, and ε representing (i.i.d.) random shocks. A policy will specify various current and future values for $\{x_t\}$. The problem is that (F, θ) – which may be difficult to recover from the data anyway – may vary systematically with changes in $\{x_t\}$.

> According to the theory of economic policy, one then simulates the system under alterative policies (theoretically or numerically) and compares outcomes by some criterion. For such comparisons to have any meaning, it is essential that the structure [i.e. the equation characterizing the dynamic behavior of the economy] not vary systematically with the choice of [those variables characterizing the "policy"].
>
> Everything we know about dynamic economic theory indicates that this presumption is unjustified.

The parameters that encode agents' expectations may depend on what policy regime is in place. So those expectations (and agents' optimal policy rules) may change when the regime changes, too. One way Lucas's proposed solution has been understood was a push to seek "deep" parameters invariant to policy changes. He noted that this was generally a problem for long-term forecasting, more often than (but not excluding) short-term forecasting:[132]

> For short-term forecasting, these arguments have long been anticipated in practice, and models with good (and improvable) tracking properties have been obtained by permitting and measuring "drift" in tile parameter vector 0. Under adaptive models which rationalize these tracking procedures, however, long-run policy simulations are acknowledged to have infinite variance, which leaves open the question of quantitative policy evaluation.

Lucas recommended that we would have to concentrate our efforts on finding the additional parameters that we would incorporate in the model that would track the economy's behavioral adaptations to changes in the system in addition to the parameter that described those changes.[133] In modern jargon, what we need for our policy models are "deep" parameters that are policy-invariant.

Perhaps instead we ought to model the policy as a function of the state of the system:

$$x_t = G(y_t, \lambda, \mu_t)$$

Where G is known, λ is a fixed parameter vector, and a μ_t a vector of random disturbances. The modified state equation for the rest of the economy is then:

$$y_{t+1} = F(y_t, x_t, \theta(\lambda), \varepsilon_t)$$

---

[132] Note, however, the even Lucas was aware of aberrant cases even in the short term. Sometimes changes happen quickly enough that the functional form being used to model the economy must be changed right away. Furthermore, and somewhat strangely, Lucas also briefly suggests that even long-term features that we other omit for convenience should be included as factors *in* short-term projects (e.g. in the short-term objective function that's being maximized).

[133] There are, quite possibly, other options that Lucas did not consider apart from these two. He does remark that this is related to the rules-vs.-policy debate that we saw Kydland and Prescott address.

θ then captures behavioral changes with changes in λ.

> In a model of this sort, a *policy* is viewed as a change in the parameters λ, or in the function generating the values of policy variables at particular times. A change in policy (in λ) affects the behavior of the system in two way: first by altering the time series behavior of x_t; second by leading to modification of the behavioral parameters $\theta(\lambda)$ governing the rest of the system. Evidently, the way this latter modification can be expected to occur depends crucially on the way the policy change is carried out. If the policy change occurs by a sequence of decisions following no discussed or pre-announced pattern, it will become known to agents only gradually, and then perhaps largely as higher variance of "noise." In this case, the movement to a new $\theta(\lambda)$, if it occurs in a stable way at all, will be unsystematic, and econometrically unpredictable. If, on the other hand, policy changes occur as fully discussed and understood changes in *rules*, there is some hope that the resulting structural changes can be forecast on the basis of estimation from past data of $\theta(\lambda)$.

Given his remarks, I want to investigate briefly two purported implications of his remarks that would have an impact on economic methodology. First, I want to suggest that one distinction – reduced-form as opposed to structural modeling – is accompanied by the same ambiguities that effect the distinction between partial and more general (or micro as opposed to more macro) models. The second is whether or not Lucas's remarks, strictly speaking, imply that adequate models require microfoundations (I suspect that this is not necessarily the case).

5.1 ECONOMICS VS. ECONOMETRICS, STRUCTURAL VS. REDUCED-FORM

The so-called Lucas Critique has actually been around for some time (Lucas's elaboration of it just happened to be more dramatic). Even prior to Lucas's papers in the 1970s, econometrics had distinguished between reduced-form and structural parameters (and it is still sometimes debated which kind of approach we ought to prefer and when).[134] [135] They tend to be presented as contrasting to one another: the latter being more a-theoretical than the former, since structural models derive from an underlying theory. Reduced form equations for reduced-form parameters represent endogenous parameters solely in terms of exogenous ones.[136] One need not have an underlying economic theory to formulated something in reduced-from – we merely observe the statistical relationship between variables.[137] Sometimes "reduced-form" methods are also called "experimental," and might involve a good bit of data-fitting.

> A notion that can be sensibly associated with the term *structural* is of an equation that stands alone, that makes sense by itself, and that has a certain autonomy. In a market system, for example, one can consider the demand equation singly, considering the optimizing behavior of a consumer. This will lead to an equation giving quantity demanded as a function of price and income. Similarly, the supply equation has a life of its own…These equations are structural in that each can be studied on its own. Adding the equilibrium condition leads to a complete system with two variables determined in the model: price and quantity. The *reduced form* consists of an equation for each of these variables. It is difficult to generate economic insight by considering either of these equations separately. (Christensen and Kiefer 2009, 6)

The Lucas Critique is often interpreted as a push towards *structural* models, which can deliver useful counterfactual analyses, and therefore can be informative insofar as we may be interested in speculating beyond the data. They may not, however, cleanly separate endogenous variables

---

[134] The distinction between reduced-form and structural parameter initially concerned observability.

[135] And also note that Lucas was responding to the models of his time in particular; today, we tend to treat any macroeconomic model as needing to pass the Lucas critique.

[136] E.g. Logic and Probit models.

[137] Sometimes structural parameters can be derived from reduced-form ones.

from exogenous variables in their formulation (for example, one can have a structural equation with endogenous variables on both the left and right hand sides).

> A key distinction in econometrics is between essentially descriptive models and data summaries at various levels of statistical sophistication and models that go beyond mere associations and attempt to estimate causal parameters. The classic definitions of causality in econometrics derive from the Cowles Commission simultaneous equations model that draw sharp distinctions between exogenous and endogenous variables, and between structural and reduced form parameters. Although reduced form models are useful for some purposes, knowledge of structural or causal parameters is essential for policy analysis. (Cameron and Trivedi 2000)

These structural models are helpful for dealing with real-life phenomena that might complicate our modeling efforts, such as feedback effects. Though Timmins and Schlenker (2009) here are referring to environmental economic contexts, the lesson could be generalized to any complex system.

> Structural models are able to address explicitly those feedbacks. This can have a number of advantages. For example, many of the covariates in these models will be endogenous…The structure of the model may suggest instrumental variables strategies to deal with those endogenous variables and recover consistent estimates, where otherwise valid instruments may be hard to find. (353)

On the other hand, reduced-form methods may be more useful in other contexts.

> Counter-factual policy simulations that include out-of-sample predictions by definition require structural parameters that model how people readjust in response to the policy…Reduced-form studies, on the other hand, are better suited to identifying key high-level response parameters as opposed to primitives in a structural model. (352)

Reduced-form modeling includes analyses of *treatment effects*: in a quasi-experimental setup we might investigate how manipulating *one* thing will affect how that (local) system unfolds. Structural equations establish other kinds of interrelationships – for example, there may be interesting relationships between endogenous variables which would be washed out if one insists on reducing the structural equations into reduced-form.

We can always derive reduced-form equations from structural ones, though not the other way around. Most of the time that we encounter structural equations (note that these simultaneous equations) in the form of:

$$\mathbf{y}A + \mathbf{x}\mathbf{B} = \boldsymbol{\varepsilon}$$

Where **y** is a 1xN vector of endogenous variables, **x** is a 1 x M vector of predetermined variables, and **ε** is a 1 x N vector of disturbances. A and B are MxN arrays of coefficients. The reduced form for this equation would then be:

$$\mathbf{y} = \mathbf{x}\Pi + \mathbf{z}$$

Where $\Pi = -BB^{-1}$ and $z = \boldsymbol{\varepsilon}\ B^{-1}$. A concrete textbook example (from Hoover 2012) is the relationship between money (m) and GDP (Y). Suppose they're interrelated (money depends on GDP and vice versa)

(1) $m=\alpha y+\varepsilon_m$

(2) $y=\beta m+\varepsilon_y$,

However,

> If we have a model like equations (1) and (2), including knowledge of α and β and of the statistical properties of ε_m and ε_y, then answering counterfactual questions (probabilistically) would be easy. The problem in the Cowles Commission view is that we do not know the values of the parameters or the properties of the errors. The real problem is the inferential one: given data on y and m, can we infer the unknown values? As the problem is set out, the answer is clearly, "no." (93)

We need a bit more information to figure out causal information of the kind that reduced-form (and its narrower form in claims that rely on *ceteris paribus* idealizations) if there is more observable information; that is, some observable information that we can treat as exogenous to these endogenous ones. For instance, interest rate (r) and price (p) A bit of technical legwork yields the following as reduced-form correspondents:

(3) $m=\Pi_1p+\Pi_2r+E_m$

(4) $y=\Gamma_1p+\Gamma_2r+E_y$,

Π_1, Π_2, Γ_1 and Γ_2 are the estimated coefficients and E_m and E_y are regression residuals.

Is there is some sense in which the distinction resembles that between general and partial equilibrium methods? The latter might correspond to the reduced-form modelling methods just summarized here. This is because the causal claims of partial equilibrium model are analogous to reduced-form statements. In fact, the kinds of causal claims that I have been considering as paradigmatic are reduced-form in a much narrower way – typically, there is only one exogenous variable and one endogenous variable.

Given that reduced-form methods give us more obvious causal information, why prefer structural form analysis at all? Well, the counterfactual work would be easy if we actually knew all the values of the coefficients and the error terms in the original structural equations of m and y. But we only in general have (at least initially) those observables m and y. Estimating the other terms of interest requires us to posit a direction of causation, which a regression (in and of itself) will not select.

> On the one hand, the structural approach is particularly useful in the presence of feedback loops, or when rational agents can reoptimize in response to large policy shifts. Feedback loops pose difficult endogeneity problems that generally make them more difficult to estimate using reduced-form methods. Counter-factual policy simulations that include out-of-sample pre-dictions by definition require structural parameters that model how people readjust in response to the policy (Timmins and Schlenker 2009, 352).

Despite that they supply (subtly) different kinds of information to the economist, both are often useful together:

> For example, reduced-form studies can be used to estimate high-level response parameters of interest, e.g., the effect of environmental amenities (clean air, good weather, etc) on humans (e.g., infant health) or the ecosystem (e.g., crop yields) that are then used in structural models to simulate policy responses by economic agents.

I noted, in various ways throughout this dissertation, that it was not straightforward how to move from more partial to more general models; I suspect that for similar reasons reconciliation efforts between reduced-form and structural methods will be difficult. Rather than that we need to look exclusively for structural equations, however, I take it that the proper moral of the Lucas Critique (and subsequent developments in econometrics) is actually that we ought to use both, and that to get a useful enquiry that provides counterfactual analysis requires careful coordination between both. It is thus not unsurprising that the caricature of these two kinds of analyses often sketches one as theoretic and the other as empirical.

So, what does this mean for the microfoundations project? Many have taken the Lucas Critique to imply that what we need is to explicit model consumer expectations, tastes, and technology. Defenders of DSGE models, and others that pursue microfounded models, claim that such approaches avoid the Lucas Critique. But it's unclear whether they actually succeed in doing so. Furthermore, the Lucas Critique does not necessarily necessarily imply microfoundations if the proper lesson is the one I have briefly summarized in Section 5.1.

Rather, Lucas seems to suggest that any model that incorporates rational-agent assumptions would be susceptible to the Critique if the rational agent is misspecified. This is precisely the way in which DSGE models can and do fail – "a model which is not correctly specified (e.g. because it uses the wrong specification for price stickiness, or because it does not take into account agent heterogeneity) will also display parameter instability issues in response to policy shocks, and may provide inaccurate policy advice" (Hurtado 2007, 7).

Lucas's further criticism is not, as some people seem to suggest (and if not, the language certainly implies it), that we need to pursue a more reductionist project: "The failure of those models to ground their supply and demand curves in the decisions of rational, forward-looking consumers and producers was a fatal flaw" (Ohaian et al 2009, 2240). Rather, I think that the economic tools that we considered as methodological guides to modeling – these *ceteris paribus* (and by extension equilibrium) methods – are motivated in the same way that Lucas motivates his methodological critique.

5.2 MULTISCALE MODELS TO THE RESCUE?

A speculative note: would multi-scale modeling would allow the economic modeler to sidestep the Lucas critique, or at least the kinds of difficulties that motivate microfoundational and structural approaches to economic modeling? In particular, can a multi-scale model account for *feedback effects*? I think it's promising for some of the cases the Lucas Critique is concerned with, at least in their caricature form. A policy at a higher scale filters down to the lower scale, altering features there (e.g. individual agents' expectations). These then give rise to unanticipated behaviors at the macroscale because the original predictive enquiry did not take into account the micro-sensitivity to the macro-policy. That is, using Lucas's notation, $\theta(\lambda)$ is a parameter that the macro-model will have to consult a more micro-scale for.

This may not generally be the case. The textbook example introduced earlier in this chapter, the interrelationship between money and GDP, and their associated "endogenous" counterparts (interest and price) can be clearly interpreted as residing on different scales. Furthermore, even if they did, the difficulty in economics is that we do not yet have any establish method for bridging scales into a coherent multi-scale model, whereas we do in physics. For example, when we manipulate a rubber band we see it reach its macro-level quasi-equilibrium fairly quickly, though at the micro-scale things will iron out over a longer time scale. In order to describe what happens at the macroscopic scale, we often still need to consult details at the microscopic scale. Describing an inhomogeneous elastic string, for example, requires constructing a wave equation – a continuum equation – that includes in it micro-terms. We now have newer methods in the material sciences, such as renormalization group and homogenization methods, which allow for some mobility between scales. An additional problem is that we still do not have a good grasp of the micro-theory in economics, whereas again in the case of physics we do (at least much of the time).

To my knowledge, such strategies do not currently have successful and broadly accepted formal analogues in economics. However, in the case of assessing and mitigating systemic risk, the policymaker's project quite clearly requires coordinating assessments of overall systemic risk with how that risk is distributed amongst the constituents of that system. The challenge for the policymaker-qua-engineer, for example, is to design a structure so that the banks are connected *enough* but not *too* connected. Furthermore, how that risk gets distributed post shock may also depend on factors that are internal to the financial system itself, meaning that a shock to the economy may also have *endogenous* causes. As a first glance, the multi-scale framework that I have sketched seems like it may be a good investigative apparatus for considering several seemingly distinct models together in order to deal with the kinds of problems that concerned Lucas and still concerns economists today.

BIBLIOGRAPHY

Ackerman, F. (2001) Still Dead After All These Years: Interpreting the Failure of General Equilibrium Theory. *Journal of Economic Methodology* 9(2): 119-139.

Ahelegbey, D.F. and P. Giudi (2014) Bayesian Selection of Systemic Risk Networks. *Advances in Econometrics* 34: 117 - 153

Allen, F. and and D. Gale (2000) Financial Contagion. *The Journal of Political Economy* 108(1): 1-33

Aoki, M. (2004) *Modeling aggregate behavior and fluctuations in economics. Cambridge Books*.

Arrow, K.F. and G. Debreu (1954) Existence of an Equilibrium for a Competitive Economy. *Econometrica* 22(3): 265-290.

Arrow, K.F. (1951) *Social Choice and Individual Values*, Yale University Press.

Ascari, G., Fagiolo, G., and Roventini, A. (2015). Fat-tail distributions and business-cycle models. *Macroeconomic Dynamics*, *19*(02), 465-476.

Ashcraft, A. and T. Schuermann (2008) Understanding the Securitization of Subprime Mortgage Credit. *Federal Reserve Bank of New York Staff Reports*, 318.

Attanasio, O., and Davis, S. J. (1994). *Relative wage movements and the distribution of consumption.* NBER Working Paper No. w4771.

Auyang, S. (1998) *Foundations of Complex-system Theories in Economics, Evolutionary Biology, and Statistical Physics.* Cambridge University Press,

Batterman, R. (2001) *The Devil in the Details*: *Asymptotic Reasoning in Explanation, Reduction, and Emergence*. Oxford University Press.

Batterman, R. and C. Rice (2014) Minimal Model Explanations. *Philosophy of Science* 81(3): 349– 376.

Batterman, R. (2013) The Tyranny of Scales. In Robert Batterman (ed.), *The Oxford Handbook of Philosophy of Physics*, Oxford: Oxford University Press.

Benes, J. M., M. Kumhof, and D. Laxton (2014) Financial Crises in DSGE Models: A Prototype Model. IMF Working Paper

Benhabib, J., Bisin, A., and Zhu, S. (2014). The wealth distribution in Bewley models with investment risk. NBER Working Paper No. w20157.

Bernstein, D.S. (2007) Ivory ghost. *IEEE Control Systems Magazine* 27: 16–17.

Bible, D. (2013) Letter to Robert deV. Frierson: Regarding the Notice of Proposed Rulemaking on Capital Assessment and Stress Testing Information Collection OMB Control Number: 7100-0341. *Manuscript.*

Blancard, O. and L.H. Summers (1986) Hysteresis and the European Unemployment Problem. *NBER Macroeconomics Annual* 1, 15-78.

Board of Governors of the Federal Reserve (2013) 2013 Banking and Consumer Regulatory Policy [Press Release]. Retrieved from: https://www.federalreserve.gov/newsevents/press/bcreg/20130314a.htm

Bogen, J. (2002) Analyzing Causality: The opposite of counterfactual is factual. *Manuscript.*

Blumberg, D. and A. Davidson (2008) The Giant Pool of Money. Radio Broadcast. *This American Life*, Episode 35.

Blundell, R. and T. Stoker (2005) Heterogeneity and Aggregation. *Journal of Economic Literature* 43: 347-391.

Boissay, F., F. Collard, and F. Smets. (2013). Booms and systemic banking crises. *ECB Working Paper No. 1514*

Bookstaber (2010) Here's Why Econophysics Will Never Work, Though Traders May Dream, *Business Insider.* Retrieved from: http://www.businessinsider.com/heres-why-econophysics-will-never-work-though-traders-may-dream-2010-8

Borio, C. (2014). The financial cycle and macroeconomics: What have we learnt? *Journal of Banking and Finance*, *45*, 182-198.

Bremus, F., and Buch, C. M. (forthcoming). Granularity in Banking and Growth: Does Financial Openness Matter? *Journal of Banking & Finance*.

Brown, D. J., DeMarzo, P. M., and Eaves, B. C. (1996). Computing equilibria when asset markets are incomplete. *Econometrica: Journal of the Econometric Society*, 1-27.

Brunnermeier, M. K., Crockett, A., Goodhart, C. A., Persaud, A., & Shin, H. S. (2009). *The fundamental principles of financial regulation* (Vol. 11). ICMB, International Center for Monetary and Banking Studies.

Butterfield, J. and N. Bouatta (2011) Emergence and reduction combined in phase transitions. *Manuscript.*

Callen, H.B. (1960) *Thermodynamics*. New York: John Wiley and Sons.

Cartwright, N. (1995). "Ceteris Paribus" Laws and Socio-Economic Machines. *The Monist*, *78*(3), 276-294.

Chatelain, J. B., and Ralf, K. (2012). The failure of financial macroeconomics and what to do about it. *The Manchester School 80*(s1), 21-53.

Christiano, L.J., M.Trabandt, and K. Walentin, (2011) DSGE Models for Monetary Policy Analysis. In B.M.. Friedman, and M.Woodford (eds.): *Handbook of Monetary Economics* (3A), The Netherlands: North-Holland. 285-367.

Coase, Ronald (1937). "The Nature of the Firm." *Economica* 4(16) 386–405.

Committee on the Global Financial System (2000) "Stress testing by large financial institutions: current practice and aggregation issues." Basel: Bank for International Settlements. Retrieved from: www.bis.org/pub/cgfs14.htm

Debreu, G. (1951). The coefficient of resource utilization. *Econometrica: Journal of the Econometric Society*, 273-292.

De Vroey, M., and Duarte, P. G. (2013). In search of lost time: The neoclassical synthesis. *The BE journal of macroeconomics*, *13*(1), 965-995.

DiMartino, D. and J. Duca. (2007) The Rise and Fall of Subprime Mortgages. *Federal Reserve Bank of Dallas Economic Letter* 2(11) 1-8.

Dimian, M. and P, Andrei. (2014) *Noise-driven phenomena in hysteretic systems*. Springer.

Dorfman, R. (1969) An Economic Interpretation of Optimal Control Theory. *The American Economic Review* 58(5): 817–831.

Duhem, P. (1902). *Thermodynamique et Chimie.* Paris: Hermann.

Duhem, P. (1896) Sur les Déformations Permanentes et l'hystérésis. In *Sur les Déformations Permanentes et l'hystérésis.* Bruxelles: Hayez, Imprimeur de l'Académie Royale de Belgique.

Duhem, P. (1893). *Introduction à la Mécanique Chimique*. Paris: Carré.

Düppe, T. and E.R Weintraub (2014) Siting the New Economic Science: The Activity Analysis Conference of June 1949. *Science in Context* 27(3): 453–483.

Earman, J., and J. Roberts (1999) "Ceteris Paribus, There is no Problem of Provisos. *Synthese* 118: 439–478.

Earman, J., J. Roberts, and S. Smith (2002) Ceteris Paribus Lost. *Synthese* 57(3): 281-301.

Einstein, A. (1949) Autobiographical notes. In P. Schilp (ed.), *Albert Einstein: Philosopher-scientist: The Library of Living Philosophers Vol. 7*, 1–95. LaSalle: Open Court.

Erickson, P. J. Klein, L. Daston, R. Lemov, T. Sturm, and M. Gordin (2013) *How Reason Almost Lost Its Mind: The Strange Career of Cold War Rationality*. Chicago: University of Chicago Press.

Ewing, J.A. (1885) Experimental research in magnetism. *Philosophical Transactions of the Royal Society of London* 172(2)

Fagiolo, G. (2007). Clustering in complex directed networks. *Physical Review E*, *76*(2), 026107.

Ferreira, M.A.R., D. Higdon, H. Lee, and M. West (2006) Multi-scale and hidden resolution time series models. *Bayesian Analysis* 1(4): 947-967.

Feynman, R. (1981) The Pleasure of Finding Things Out. BBC Horizon.

Fisher, F. M. (1983) *Disequilibrium Foundations of Equilibrium Economics*. Cambridge: Cambridge University Press.

Fisher, F.M. (1976) The Stability of General Equilibrium: Results and Problems. In M. J. Artis and A. R. Nobay (eds.) *Essays in Economic Analysis: Proceedings of the Association of University Teachers of Economics Annual Conference*. Cambridge: Cambridge University Press.

Fisher, I. (1933). The debt-deflation theory of great depressions. *Econometrica: Journal of the Econometric Society*, 337-357.

Fisher, J. and D. Johnson (2006) Consumption mobility in the United States: Evidence from two panel data sets. *B.E. Journal of Economic Analysis and Policy* 6(1).

Fodor, J. (1991) You Can Fool Some People All of the Time, Everything Else Being Equal; Hedged Laws and Psychological Explanations. *Mind* 100: 19–34.

Fodor, J. (1987) *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge: MIT Press.

Fodor, J. (1974) Special Sciences, or the Disunity of Science as a Working Hypothesis. *Synthese* 28: 97-115.

Freedman, D. (2011) A Formula for Economic Calamity. *Scientific American* 305(5):76-9.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, *79*(3), 733-772.

Gagliardini, P., and Gouriéroux, C. (2013). Granularity adjustment for risk measures: Systematic vs unsystematic risks. *International Journal of Approximate Reasoning*, *54*(6), 717-747.

Gallegati, M., and Kirman, A. (2012). Reconstructing economics: Agent based models and complexity. *Complexity Economics*, *1*(1), 5-31.

Garcia, N. (2011) DSGE Macroeconomic Models: A Critique. *Economie Appliquée* 1: Jan.-March.

Gatti, D. D., Gaffeo, E., and Gallegati, M. (2010). Complex agent-based macroeconomics: a manifesto for a new paradigm. *Journal of Economic Interaction and Coordination*, *5*(2), 111-135.

Gelman. A.. Park. D.K.. Ansolabehere. S.. Price. P.N.. Minnite. L.C. (2001). Models. assumptions and model checking in ecological regressions. *Journal of the Royal Statist. Soc. A.* 164. 101–118

Giraud, G., and Maret, I. (2001). Behavioral Heterogeneity in Large Economies. *Personal manuscript. Footnotes*, *1*.

Grandmont, J.-M. (1992) "Transformations of the commodity space, behavioral heterogeneity, and the aggregation problem." *Journal of Economic Theory* 57: 1-35.

De Gregorio, J., and R.O. Valdes (2001). Crisis transmission: Evidence from the debt, tequila, and Asian flu crises. In S. Claessens and K. Forbes (Eds.). *International financial contagion*. Springer Science & Business Media

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

Hamilton, R. B. (2009) *The Road to America's Economic Meltdown.* Bloomington: Authorhouse.

Handel, B.R., J. Kolstad, and J. Spinnewijn, (2015) Information Frictions and Adverse Selection: Policy Interventions in Health Insurance Markets. NBER Working Paper No. 21759

Hausman, D. (1992) *The Inexact and Separate Science of Economics*. Cambridge: Cambridge University Press.

Hempel, C. G. (1988). Provisos: A philosophical problem concerning the inferential function of scientific laws. In A. Grünbaum & W. Salmon (eds.), *The Limits of Deductivism*. University of California Press, Berkeley, 190-36

Hurtado, S. (2013) DSGE Models and the Lucas Critique. Bank of Spain Working Paper 13XX.

Mirowski, P., and Hands, D. W. (1998). A paradox of budgets: the postwar stabilization of American neoclassical demand theory. *History of Political Economy*, 260.

Hebb, D. (1949) *The Organization of Behavior.* New York: Wiley and Sons.

Hicks, J. (1939) *Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory.* Oxford: Clarendon Press.

Hildenbrand, W. (1994) *Market Demand: Theory and Empirical Evidence*. Princeton University Press.

Hoover, K. (2012) Economic Theory and Causal Inference. In U. Mäki (ed.): *Handbook of the Philosophy of Economics*. San Diego: North Holland, 89-113.

Inoue, M., F. Kajiya, H. Inada, A. Kitabatake, M. Hori, S. Fukui, H. Abe, S. Takasugi, and T. Furukuwa (1976) Optimal control of medical treatment: Adaptive control of blood glucose level in diabetic coma. *Computers and Biomeidcal Research* 9(3): 217-228

Jappelli, T., and Pistaferri, L. (2006). Intertemporal choice and consumption mobility. *Journal of the European Economic Association*, *4*(1), 75-115.

Johnson, S. and J. Kwak (2010) *13 Bankers: The Wall Street Takeover and the Next Financial Meltdown.* Pantheon Books.

Jordà, Ò., Schularick, M., and Taylor, A. M. (2013). When credit bites back. *Journal of Money, Credit and Banking*, *45*(s2), 3-28.

Kaminsky, G., Lizondo, S., and Reinhart, C. M. (1998). Leading indicators of currency crises. *Staff Papers-International Monetary Fund*, 1-48.

Keen, S. (2015) The Fed Has Not Learnt From the Crisis. *Forbes*. Available online at: http://www.forbes.com/sites/stevekeen/2015/03/28/the-fed-has-not-learnt-from-the-crisis/

Kehoe, T. J., Levine, D. K., Mas-Colell, A., and Zame, W. R. (1989). Determinacy of equilibrium in large-scale economies. *Journal of Mathematical Economics*, *18*(3), 231-262.

Keynes, J.M. (1923) *A Tract on Monetary Reform.* London: Macmillan.

Kirman, A. P. (1998). *Economies with interacting agents* (pp. 17-51). Springer Berlin Heidelberg.

Kirman, A. (1992) Whom or What Does the Representative Individual Represent? *Journal of Economic Perspectives,* 6(2): 117–36. *History of Political Economy* 38: 228-245.

Klein, L.F. (1947) *The Keynesian Revolution*. New York: Macmillan.

Kocherlakota, N. R. (1996). The equity premium: It's still a puzzle. *Journal of Economic literature*, *34*(1), 42-71.

Krugman (1980). Scale economies, product differentiation, and the pattern of trade. *The American Economic Review* 70(5): 950-959.

Landmann, O. (2014). *Short-run macro after the crisis: The end of the'new'neoclassical synthesis?* Discussion Paper Series No. 27, University of Freiburg, Department of International Economic Policy.

Lang, W. and J. Jagtiani (2010) The Role of Credit Risk Management and Corporate Governance. *Atlantic Economic Journal* 38(2): 123-144.

Lenhart, S., and Workman, J. T. (2007). *Optimal control applied to biological models*. CRC Press.

Leontief, W. (1982) Letter: Academic Economics. *Science* 127: 104–7.

Lewbel, A. (1994) Aggregation and Simple Dynamics. *American Economic Review*, 84(4): 905–18.

Lo, A. W., and Mueller, M. T. (2010). WARNING: Physics envy may be hazardous to your wealth! Manuscript.

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

Marshall, A. (1890) *Principles of Economics*. Macmillan: London.

Machamer, P. (2004) Activities and Causation: The Metaphysics and Epistemology of Mechanisms. *International Studies in the Philosophy of Science* 18: 27–39.

Meyers, R. A. (Ed.). (2010). *Complex systems in finance and econometrics*. Springer Science and Business Media.

Mitchell, S. (2003) *Biological Complexity and Integrative Pluralism.* Cambridge University Press.

Moore, P. (2004). *The ABC of exotic structured credit.* London: Incisive Financial Publishing Ltd.

Nagel, E. (1961). *The structure of science: Problems in the logic of scientific explanation.* Harcourt, Bruce, and World Publishing.

Nier, E., Yang, J., Yorulmazer, T., and Alentorn, A. (2007). Network models and financial stability. *Journal of Economic Dynamics and Control*, *31*(6), 2033-2060.

Pearl, J. (2000) *Causality: Models, Reasoning, and Inference.* Cambridge University Press.

Pietroski P. and G. Rey (1995) When Other Things aren't Equal: Saving Ceteris Paribus Laws from Vacuity. *British Journal for the Philosophy of Science* 46: 81–110.

Peters, R. (2004) Friction at the Mesoscale. *Contemporary Physics* 45(6): 475-490

Quah, J. (2007) Demand is heterogeneous in Grandmont's model. Working Paper, Nuffield College, Oxford, No. 2001-W12.

Ramsey, F. P. (1928). A mathematical theory of saving. *The economic journal 38*(152), 543-559.

Rizvi, S. (1997) The Sonnenschein-Mantel-Debreu Results after Thirty Years. *History of Political Economy* 38: 228-245.

Rizvi, S. (1994) The microfoundations project in general equilibrium theory. Cambridge Journal of Economics 18: 357-377.

Rosenberg, A. and T. Curtain (2013, Aug. 25) What is Economics Good For? *The New York Times:* SR9.

Rosenberg, A. (1992). *Economics-Mathematical Politics or Science of Diminishing Returns?* University of Chicago Press.

Rubinstein, A. (2006) Dilemmas of an Economic Theorist. Manuscript.

Schiffer S. (1991). Ceteris Paribus Laws. *Mind* 100: 1–17.

Kirman, A. (1992) Whom or What Does the Representative Individual Represent? *Journal of Economic Perspectives,* 6(2): 117–36.

Sachs, J. (1986) “High Unemployment in Europe. Diagnosis and Policy Implications.” NBER Working Paper No. 1830.

Sales-Pardo, M., Guimera, R., Moreira, A. A., and Amaral, L. A. N. (2007). Extracting the hierarchical organization of complex systems. *Proceedings of the National Academy of Sciences*, *104*(39), 15224-15229.

Scarf, H. (1960). Some examples of global instability of the competitive equilibrium. *International Economic Review*, *1*(3), 157-172.

Schurz, G. (2001) Pietroski and Rey on *Ceteris Paribus* Laws. *British Journal for Philosophy of Science* 52, 359–370.

Schurz, G. (2002) Ceteris Paribus Laws: Classification and Deconstruction. *Erkenntis* 57(3): 351–372.

Schinasi, G. (2004) Defining Financial Stability. IMF Working Paper.

Sidel, R. D. Enrich, and D. Fitzpatrick (Sept. 26 2008) WaMu is Seized, Sold Off to J.P. Morgan, in Largest Failure in U.S. Banking History. *Wall Street Journal.*

Sraffa, P. (1998) [1925] On the relations between cost and quantity produced. In L. Pasinetti (ed.) *Italian Economic Papers*. Oxford University Press.

Sraffa, P. (1926) The Laws of Returns Under Competitive Conditions. *Economic Journal* 36: 535– 550.

Stiglitz, J. (2010a). *Freefall: America, Free Markets and the Sinking of the World Economy.* W.W. Norton and Co.

Stiglitz (2010b) Foreclosure Moratorium, Government Stimulus Needed to Revive US Economy. Interview by Amy Goodman with *Democracy Now*.

Smets, F., and R. Wouters (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. National Bank of Belgium Working Paper.

Smith, S. (2002) Violated Laws, *Ceteris Paribus* Clauses, and Capacities. *Synthese* 130(2): 235–264.

Strevens, M. (2012). Ceteris paribus hedges: Causal voodoo that works. *The Journal of Philosophy*, *109*(11), 652-675.

Timmins, C. and W. Schelnker (2009) Reduced-Form Versus Structural Modeling in Environmental and Resource Economics. *Annual Review of Resource Economics* 1: 351-380

Tobin, J. (1972) Inflation and Unemployment. *American Economic Review* 62(1/2): 1-18.

Tovar, C. E. (2009) DSGE models and central banks. *Economics: The Open-Access, Open-Assessment E-Journal*, *3*.

Walras, L. (1874) *Elements of Pure Economics, or the Theory of Social Wealth*, transl. W. Jaff, 1954, London: George Allen and Unwin Ltd.watts

Wilson, M. (forthcoming) *Physics Avoidance and Other Essays*. Manuscript.

Wilson, M. (2008) *Wandering Significance: An Essay on Conceptual Behavior*. Oxford University Press.

Woodward, J. (2004) Counterfactuals and Causal Explanation. *International Studies in the Philosophy of Science* 18(1): 41–72.

Woodward, J. (2003) *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press.

Woodward, J. (2002) There is no such thing as a *ceteris paribus law*. *Erkenntis* 57(3): 303–328.

World Bank. (2016) Financial Stability. Retrieved from: www.worldbank.org/en/publication/gfdr/background/financial-stability