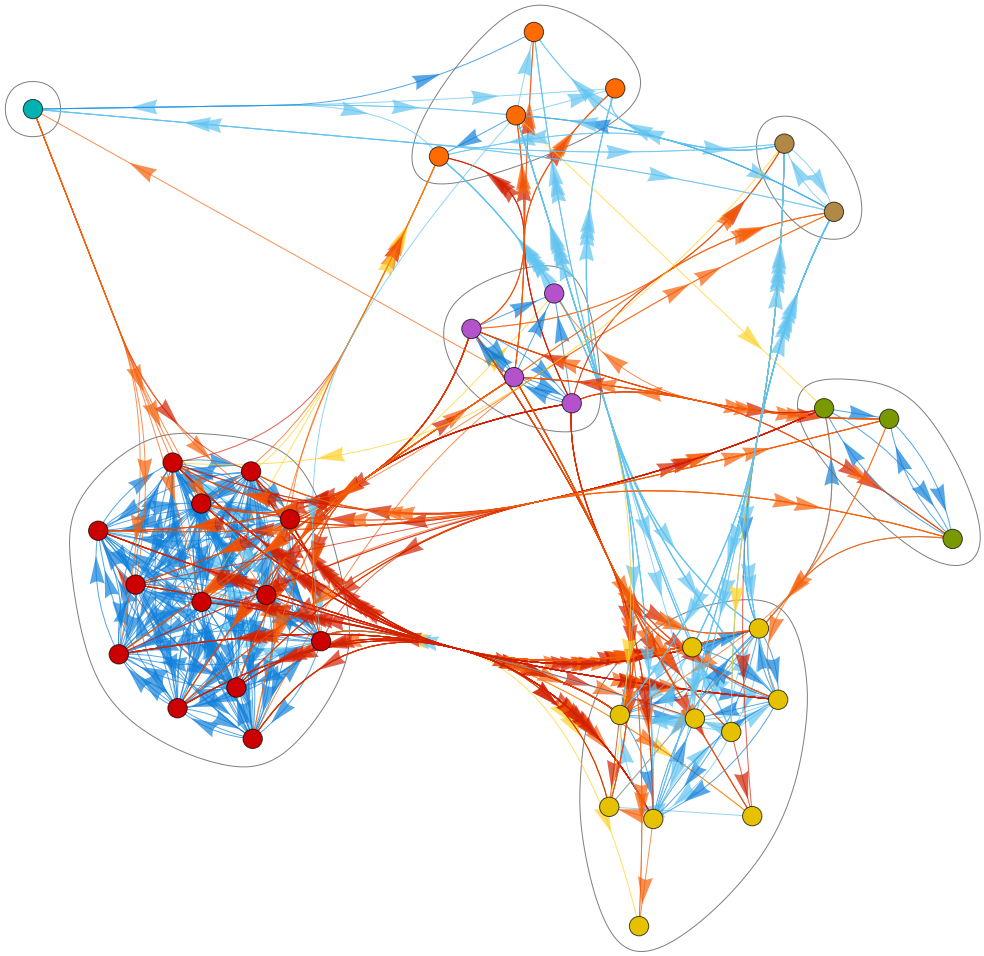
FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

FACULTY OF SCIENCES

Live Agent-Based Models



GHENT UNIVERSITY

Kevin Hoefman

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

FACULTY OF SCIENCES

Live Agent-Based Models

Supervisor: Prof. Dr. Koen Schoors
Co-supervisor: Prof. Dr. Jan Ryckebusch

Dissertation submitted in fulfillment
of the requirements for the degree of:
Doctor of Economics and of Science: Physics

Ghent University
Faculty of Economics and Business Administration:
Department of Economics
Faculty of Sciences:
Department of Physics and Astronomy

Academic year 2019-2020

The figure on the cover depicts the Eve Online political coalitions network from Chapter 8.

GHENT UNIVERSITY

Kevin Hoefman

To Kristel

May one and one always be eleven

Acknowledgements

When I embarked on this grand PhD adventure, I was hoping it would change some things in my life. Little did I know it would change me. Looking back at the roller-coaster experience that is writing a dissertation, while the ride wasn't always easy (and neither it should), I can only feel grateful for the entire experience.

It is only fitting to start by thanking my promotor, Koen Schoors, without whom this PhD would not be. The man with a thousand ideas, half of them crazy and most of them brilliant, when Koen considered computer games as testing grounds for economic and social theories, he saw potential where others saw risk. His enthusiasm translated into a funding proposal which, in the words of Jan, "the jury could only accept". And accept it they did, providing for an entire team of LAB-M researchers. This dissertation is only one of its tangible results.

Second, I want to thanks my co-promotor, Jan Ryckebusch, whose Social Physics research motivated the multidisciplinarity of this dissertation. Jan's enthusiasm for exploring the correspondences between sub-atomic physics and human social behavior is inspiring, and I can totally see why many of Jan's students consider him their favorite teacher. I look forward to our next collaboration, on refining Maxwell's demon concept into a heuristic explanation for the exponential price behavior observed in Eve Online's virtual market.

Co-author and coach, Aaron Bramson was instrumental in bringing the academic publications of this dissertation to realisation. Aaron was also the first to recognize the implications of being able to separate social from functional value in Eve Online's market data. We didn't get to publish in Science after all, but we did well.

Research does not happen in a vacuum, and I want to thank the professors, staff, and fellow PhD students at the department of economics, and of the LAB-M group, for being their amicable and inspirational selves. In particular I would like to thank Vicky and Tom, with whom I had many a conversation on many a different topic, often to the bewilderment of our other dear colleagues, trying to get some work done on the other side of the office wall.

To family and friends, and most of all my parents: thank you for everything.

And finally, I want to thank Kristel, my amazing wife, wine lady and champagne ambassador, without whom I would have starved to death five years ago.

Ghent, May 2020
Kevin Hoefman

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 A Crisis of Models

In August 2009, complexity economists Doyne Farmer and Duncan Foley argued that policy makers were unable to lead the world out of the financial crisis of 2008, because even the best types of available macroeconomic models[1] failed to provide useful guidance [Farmer and Foley, 2009]. As a consequence, they wrote, "the leaders of the world are flying the economy by the seat of their pants". Farmer and Foley suggested agent-based models (ABMs) as a way forward, yet recognized:

> Agent-based models are not a panacea. The major challenge lies in specifying how the agents behave, and in particular, in choosing the rules they use to make decisions. ... To make agent-based modelling useful we must proceed systematically, avoiding arbitrary assumptions, carefully grounding and testing each piece of the model against reality and introducing additional complexity only when it is needed. [Farmer and Foley, 2009, 686]

ABMs are a radical departure from classic macroeconomic modeling. Instead of describing aggregate behavior of populations of (typically) a single representative economic agent, ABMs model the behavior of *individual* agents. Macro-behavior is the emergent result of the interaction between these agents. Herein lies the crux of agent-based modeling: if the behavior of the agents, and/or the complex interaction between them, is misspecified, the emergent outcome of the ABM will be disconnected from reality and useless for real-world policy-making.

The methodology of Live Agent-Based Models (LAB-M), which we introduce in this thesis, attempts to address some of the challenges of *microfounding* macroeconomic models in general, and agent-based models in particular.

[1] Econometric and DSGE models, see chapter 2.

1.2 Outline

A decade has passed since Farmer and Foley's call to arms. Yet we find the state of macroeconomics largely as the authors described it in 2009. Econometric models aren't reliable for long-term forecasting because "they are empirical statistical models that are fit to past data. These successfully forecast a few quarters ahead as long as things stay more or less the same, but fail in the face of great change". DSGE models, in turn, "assume a perfect world, and by their very nature rule out crises of the type we are experiencing now" [Farmer and Foley, 2009, 685]. But agent-based models still haven't replaced the old models, leading to the uncomfortable realisation that economic policy makers are still flying blind today. We take a more in depth look at macroeconomic models in chapter 2.

Our understanding of how people make decisions has changed considerably during the last 10 years. Whereas in the 20th century, the model of *rational utility-maximizing agents* reigned supreme, since the 21st century an alternative system of decision-making based on *heuristics* has received mainstream recognition, to the point where in 2011, some researchers claimed an "emerging science of heuristics". With our understanding of human decision-making still rapidly evolving, it is no wonder that macroeconomic agent-based models, which rely on an accurate description of agent behavior, have not made much progress. We look at rationality and human decision-making in greater detail in chapter 3.

Economic networks are complex systems. Economic researchers have increasingly relied on techniques from social physics, like network theory, to adequately model the economic interaction between agents. Macroeconomic agent-based models need to specify the complex interaction between economic agents correctly to obtain realistic emergent behavior. We look at complexity science in chapter 4.

Virtual worlds allow researchers to study real people in a simulated yet natural environment, where they don't feel observed. Because virtual worlds are generated on a computer, all occurring activity can be accurately recorded. The resulting datasets are vastly more comprehensive, granular and reliable than data collected from surveys or real-world observations. This presents a novel way to answer research questions pertaining to economics and social physics, as well as other academic fields of research. We call these environments Live Agent-Based Models, reflecting on their hydrid nature, between field experiments and the artificial data-generating processes of agent-based models, with real people as the agents providing the underlying micro-behavior. We look at virtual worlds and live agent-based models in more detail in chapter 5.

To illustrate our thesis of Live Agent-Based Models as useful testing grounds for economic and social theories, we present some of our empirical contributions based on virtual worlds data. In chapter 6, we document the software implementation of the methodology used in chapter 7 and chapter 8. In chapter 7, we apply

three different datasets and a technique called Temporal Webs, to study the propagation of information through networks. In chapter 8, we study the dynamic evolution of diplomatic relations between Eve Online alliances, by testing variations and extensions of Structural Balance Theory. And in chapter 9, we study the relationship between price and quality of goods in a virtual market, using Hedonic Pricing Theory as a reference.

We summarize our work and offer an outlook for future research in chapter 10.

References

[Farmer and Foley, 2009] Farmer, J. and Foley, D. (2009). The economy needs agent-based modeling. *Nature*, 460:685–6.

Part I

Theory

Chapter 2

Models and the Economy

2.1 Models and Reality

That more realistic models produce better predictions seems intuitively reasonable. But there are sound arguments against this logic. In his 1953 essay *The Methodology of Positive Economics*[1], Milton Friedman argued that the relationship between the significance of a theory and the realism of its assumptions is, if anything, inverse. The purpose of a theory, Friedman stated, is to explain much by little: to abstract the crucial elements from the mass of complex and detailed circumstances, and permit valid predictions on the basis of them alone. Therefore,

> the relevant question to ask about the "assumptions" of a theory is not whether they are descriptively "realistic," for they never are, but whether they are sufficiently good approximations for the purpose in hand. And this question can be answered only by seeing whether the theory works, which means whether it yields sufficiently accurate predictions. [Friedman, 1953, 153].

Paul Samuelson famously critiqued Friedman's argument as a logical fallacy, calling it "the F-twist" [Samuelson, 1963, 232]. Herbert Simon referred to "Friedman's principle of unreality", and vehemently rejected it as deliberately misleading: "The expressed purpose of Friedman's principle of unreality is to save classical theory in the face of the patent invalidity of [its underlying assumptions]" [Simon, 1963, 230]. Simon called for a theory based on solid, observable foundations:

> Let us make the observations necessary to discover and test true propositions ... Then let us construct a new market theory on these firmer foundations. [Simon, 1963, 230]

[1] According to Daniel Hausman "the most influential work on economic methodology of [the 20th] century" [Hausman, 2007, 145]

Whether economic models should be based on realistic propositions remains a matter of debate today. Philosopher of science Daniel Hausman finds that opinion has shifted towards realism: about two-thirds of the articles in the February, 2018 American Economic Review were based on empirical studies, whereas twenty-five years earlier, only about one-eighth of the first issue of the 1993 American Economic Review relied on any empirical studies [Hausman, 2018]. Yet despite the trend, the debate is far from settled. In their influential essay *The Case for Mindless Economics*, Faruk Gul and Wolfgang Pesandorfer argued that the findings of behavioral economics are irrelevant to economics, and deny that the realism of the assumptions of economic models matters [Gul and Pesendorfer, 2008]. In contrast, British philosopher and economist Tony Lawson argued in his 2017 *What is wrong with modern economics, and why does it stay wrong?* that modern economics is "mostly simply irrelevant", because

> its formulations, in the main, are patently and repeatedly unrealistic, and so able to provide little or no explanatory insight or understanding of the world in which we live. [Lawson, 2017, 26]

2.2 Macro behavior and Microfoundations

A similar but distinct question is whether the laws of aggregate macro-behavior should be derived from individual behavior. The idea goes back at least to John Stuart Mill (1806 – 1873), who wrote in his 1843 *A System of Logic* that:

> The laws of the phenomena of society are, and can be, nothing but the laws of the actions and passions of human beings united together in the social state. ... Human beings in society have no properties but those which are derived from, and may be resolved into, the laws of the nature of the individual man. [Mill, 1843, 879]

The chief opponent of this methodological individualism was the pre-eminent French sociologist Émile Durckheim, who in 1927 argued that social phenomena are irreducible to individual psychology [Homans, 1987, 67]. Durckheim's doctrine used to be almost universal in sociology. It also underpins the first generation of macroeconomic models. As we shall see, the evolution of macroeconomic modeling shows a growing trend towards foundation in micro behavior.

2.3 A Short History of Macroeconomic Models

2.3.1 Econometric Models

2.3.1.1 The Cowles Commission

After the stock market crash of 1929, Alfred Cowles, an investment counselor from Colorado Springs, realized he did not understand the workings of the economy [Christ, 1994]. After being put in touch with Irving Fisher, then president of the fledgling Econometric Society, Cowles decided to finance a new economics organization, as well as a journal called *Econometrica*, still one of the most respected economic journals today [The Econometric Society, 2019].

The Cowles Commission was founded in 1932, and Econometrica began publishing in 1933. Many of the researchers involved with the Cowles Commission, such as Ragnar Frisch who originated the term "econometrics" [Fisher, 1941, 187], Trygve Haavelmo and Tjalling Koopmans, "the two chief originators of Cowles' theoretical econometric work" [Christ, 1994, 31], Kenneth Arrow and Gérard Debreu, who laid the foundation of general equilibrium theory [Arrow and Debreu, 1954], and Herbert Simon, who originated Behavioral Economics (see chapter 3), would go on to receive Nobel Prizes in economics.

The objective of the Cowles Commission was to develop an all-inclusive model of the economy, that could be used for forecasting and policy-making:

> The Cowles program was intended to combine economic theory, statistical methods, and observed data to construct and estimate a system of simultaneous equations that could describe the workings of the economy. The aim was to learn from such a system of equations how economic policy could improve the performance of the economy [Christ, 1994, 31].

Irving Fisher had applied the system-of-equations approach to economic equilibrium in 1891 to earn Yale's first PhD in political economy[1] [Dimand, 2019, 3]. The foundation of what is now called General Equilibrium analysis, the system-of-equations technique is intuitively attractive: given a set of mathematical functions that describe the rules of the economy (eg. a function for unemployment, a function for market pricing, etc.), solving this system of equations tells us how the economy works "in equilibrium". Whether the system of equations is indeed solvable (known as the identification problem) was a major contribution of the Cowles Commission. Today, the basic solution is taught in high school algebra: a system of equations has a unique solution if the number of independent variables equals

[1] As Fisher discovered a few weeks before submitting his thesis, French economist Leon Walras had already published the technique in 1874, which is why we now refer to Walrasian equilibrium.

the number of equations[1]. This was the holy grail of econometric modeling: to formulate a set of equations succinctly describing the economy, based on a number of economic variables equal to the number of equations. Solving this system would reveal the equilibrium at which the economy operates[2].

While the Cowles Commission developed many techniques fundamental to modern econometrics (for an overview see [Christ, 1994]), the ultimate goal of the Cowles program, an all-inclusive model of the economy, proved elusive. Looking back in her 1990 *The history of econometric ideas*, author Mary Morgan observed that:

> By the 1950s the founding ideal of econometrics, the union of mathematical and statistical economics into a truly synthetic economics, had collapsed. [Morgan et al., 1990, 264]

Liu [Liu, 1960] (and later Sims [Sims, 1980]) pointed out that due to the complexity of the real economy, the number of variables affecting the economy is so large, and their simultaneous interaction so pervasive, that the system of equations cannot be identified: there are simply too many variables compared to the number of equations. And in the mid 1970s, the econometric approach came under all out attack.

2.3.1.2 The Lucas Critique

In his 1976 *Econometric Policy Evaluation: A Critique* [Lucas Jr, 1976], Robert Lucas Jr. famously attacked the use of econometric models in long-term policy-making as dangerous and fundamentally flawed. Before the 1970s, econometric models had revealed a historical relationship between inflation and unemployment: when unemployment went up, inflation went down, and vice versa[3]. Following the interventionist philosophy of John Maynard Keynes, and facing economic stagnation in the early 1970s, policy makers tried to exploit this relationship by deliberately creating inflation, in the hope that the economy would pick up as a result. But the observed historical relationship between inflation and unemployment broke down, and instead of the anticipated recovery, the mid to late 1970s experienced both economic stagnation *and* high levels of inflation.

Lucas blamed the undesired outcome on a lack of *microfoundations* in econometric models. Since an economy consists of individual agents who adapt to economic policy, Lucas argued, changes in economic policy induce agents to change

[1] Although Walras relied on the basic solution, a complete solution for identification is more complex. The first general proof of equilibrium existence was provided by Arrow and Debreu in 1954 [Arrow and Debreu, 1954].

[2] This is the static interpretation of equilibrium which was prevalent at the time; for dynamic stochastic equilibrium see DSGE models, below.

[3] The so-called Phillips curve.

their behavior. As a result, (past) observed behavior cannot be relied on for policy making: since changes in policy change the structure of the macro-model itself, via the adaptive behavior of individual agents:

> 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 Jr, 1976, 41]

Lucas concluded that econometric models can have good short-term forecasting properties, yet for "longer term forecasting and policy simulations ... will lead to large, unpredictable errors" [Lucas Jr, 1976, 26].

2.3.2 DSGE Models

Econometric models, Lucas argued, are essentially *backward-looking* - they estimate the structure of the economy from past data, and project it into the future, not accounting for individual responses to changes in policy. Lucas called for a dynamic equilibrium model that could explain business cycles and could account for the economy's response to economic shocks, such as disruptions in the oil supply, without making Keynesian assumptions.

As an early example of such a model, Lucas proposed Thomas Sargent's 1976 classical macroeconomic model for the United States [Sargent, 1976]. These models evolved into the modern Dynamic Stochastic General Equilibrium (DSGE) models [Sargent, 1987], with Bayesian DSGE models being the current state of the art [Smets and Wouters, 2007].

2.3.2.1 Dynamic and Stochastic

DSGE models are somewhat of a misnomer, as what sets them apart from econometric models isn't covered by the name. The term "Dynamic" refers to changes in time, which DSGE modelers implemented into their equations by adopting differential equations techniques from engineering. "Stochastic" variables, accounting for random shocks (see General Equilibrium, below), had already been developed by the Cowles Commission.

Since modern econometric (VAR) modelers also use these techniques, this makes for a rather confusing situation in terminology, as modern econometric models are also Dynamic, Stochastic, General Equilibrium models. Econometric and DSGE models share a common core: advanced (Dynamic and Stochastic) mathematics, founded on the (General Equilibrium) system-of-equations framework. DSGE models did not replace econometric models, as Lucas might have

expected. The modern DSGE model is an econometric model, with *microfoundations*: forward-looking dynamics based on the theory of Rational Expectations (see below).

2.3.2.2 General Equilibrium

Before the 1930s, systems-of-equations modelers thought of equilibrium as static: the equations in the model described the workings of the economy, and solving the system revealed the ideal state of the economy:

> When Keynes wrote [his General Theory], the terms equilibrium and classical carried certain positive and normative connotations which seemed to rule out either modifier being applied to business cycle theory. The term equilibrium was thought to refer to a system at rest, and some used both equilibrium and classical interchangeably with ideal. Thus an economy in classical equilibrium would be both unchanging and unimprovable by policy interventions. [Lucas Jr and Sargent, 1979, 201]

The idea of classic equilibrium had been rejected by the experience of the Great Depression: it was clear to the members of the Cowles Commission that a perpetually stable economy could not be the end of the story. To account for deviations from equilibrium, the Cowles Commission developed techniques to incorporate uncertainty into the model. Economic variables were augmented with a deviation factor representing unknown errors due to incorrect measurement, uncertainty of expectations, or any other factor that would cause the variable to deviate from the "correct" value. These so-called "stochastic" variables required the development of new mathematical techniques and insights, pioneered by Kenneth Arrow and Gérard Debreu. Stochastic variables changed the meaning of model equilibrium: instead of predicting a static, ideal state, the "equilibrium" outcome of economic behavior under uncertainty could now predict states which we would generally consider "out of equilibrium":

> In recent years, the meaning of the term *equilibrium* has changed so dramatically that a theorist of the 1930s would not recognize it. This development, which stemmed mainly from work by K. J. Arrow (1964) and G. Debreu (1959), implies that simply to look at any economic time series and conclude that it is a disequilibrium phenomenon is a meaningless observation. [Lucas Jr and Sargent, 1979, 201]

In recent literature, a distinction between Walrasian "equilibrium" and Schumpeterian "out-of-equilibrium" economics is sometimes made [Hanusch and Pyka,

2007]. But it is worth noting that, even though DSGE models are equilibrium models, this does not mean they predict single, unique equilibria in the old, static sense. Equilibrium models with stochastic variables allow for whole ranges of possible "equilibrium" outcomes[1]. DSGE modelers can (and typically do) make additional assumptions to narrow down the possible outcomes, in order to obtain useful predictions. Crucially, this exposes DSGE models to the same vulnerabilities that agent-based models (see below) are sensitive to: DSGE models with unrealistic assumptions about individual behavior are not microfounded in any meaningful sense.

2.3.2.3 Microfoundations

> "The basic logic in strict microfoundations and in General Equilibrium Theory, generally speaking, is to try to derive macroeconomic properties from assumptions on economic agents considered individually" [Rizvi et al., 1994, 357].

What separates DSGE models from econometric models is the focus on microfoundations. The most enduring effect of the Lucas critique, microfoundations are macro-model choices consistent with microeconomic theory, particularly regarding future expectations.

Economists have generally agreed that future expectations are important determinants of current behavior. In his 1978 "Rational expectations and the dynamic structure of macroeconomic models: A critical review", Robert Schiller wrote:

> It is easy to produce [examples] of macroeconomic behavioral relations which depend essentially on public expectations. In fact, if one looks at one of the major macroeconomic models one is impressed that most of the essential behavioral relations are based on assumptions about how expectations are formed. Public expectations of future inflation rates, interest rates, rental rates, income, and particular components thereof influence behavior current in a fundamental way. [Shiller, 1978, 2]

Lucas explained his criticism of econometric models in more detail in his 1979 paper with Thomas Sargent [Lucas Jr and Sargent, 1979]: all models make assumptions about underlying behavior, but the assumptions made by Keynesian economists regarding agent expectations are unrealistic and inconsistent with behavior of individual agents as described in microeconomics:

[1] The famous Sonnenschein [Sonnenschein, 1973], Mantel [Mantel, 1974] and Debreu [Debreu, 1974] results; for a more in depth treatment see [Rizvi et al., 1994]

> The casual treatment of expectations is not a peripheral problem in [Keynesian] models, for the role of expectations is pervasive in them and exerts a massive influence on their dynamic properties (a point Keynes himself insisted on). The failure of existing models to derive restrictions on expectations from any first principles grounded in economic theory is a symptom of a deeper and more general failure to derive behavioral relationships from any consistently posed dynamic optimization problems. [Lucas Jr and Sargent, 1979, 299]

Lucas argued that macroeconomics needed *foundations in microeconomic theory* [Lucas Jr and Sargent, 1979, 298]. Lucas wasn't the first to call for more consistency between macro and micro; in 1967, Kenneth Arrow had referred to the relationship between microeconomics and macroeconomics (or the lack thereof) as a 'major scandal' [Arrow, 1967, 734].

It must be noted that the Lucas Critique also had an important political dimension. Keynesian economists take their moniker from economist John Maynard Keynes, who argued in his 1936 *The General Theory of Employment, Interest and Money* [Keynes, 1936] that government intervention in the economy can be beneficial to society. In contrast, classical economists argue that economic agents will rationally act in their own interest, and no intervention is needed or indeed helpful. Lucas argued that Keynesian economists ignored classical microeconomics in the choices underlying their models:

> A main theme of [Keynesian] theoretical work since the General Theory has been the attempt to use microeconomic theory based on the classical postulate that agents act in their own interests to suggest a list of [model restrictions]. ... Modern probabilistic microeconomic theory almost never implies either the exclusion restrictions suggested by Keynes or those imposed by macroeconometric models. [Lucas Jr and Sargent, 1979, 298-99]

To solve what is essentially a problem of accounting for the future, Lucas adopted the theory of Rational Expectations (also see chapter 3), which he had explored in 1974 to explain involuntary unemployment, together with Edward Prescott. Rational Expectations imbues agents with a limited form of perfect foresight. Lucas conceded that the validity of this solution depended on psychological rather than economic theory:

> [W]e have provided no account as to how workers arrive at the state of perfect knowledge of the probability distributions relevant to their decision problem. Ultimately, this is a question for psychological rather than economic theory, so we do not apologize for framing it here in ad hoc "adaptive" terms. [Lucas Jr and Prescott, 1974, 204]

In retrospect, Lucas' conclusion that Keynesianism was dead, captured by the title of his 1979 paper "After Keynesian Macroeconomics", was premature; Keynesianism is very much alive today. The call for proper microfoundations was, however, hugely influential, probably not in the least due to the virulent way in which Lucas attacked the predominantly Keynesian macroeconomic profession. Lucas was so successful that a decade later, James Tobin remarked on the microfoundations counter-movement of the 1980s that:

> This [microfoundations] counter-revolution has swept the profession until now it is scarcely an exaggeration to say that no paper that does not employ the 'microfoundations' methodology can get published in a major professional journal, that no research proposal that is suspect of violating its precepts can survive peer review, that no newly minted Ph.D. who can't show that his hypothesized behavioral relations are properly derived can get a good academic job. [Tobin, 1986, 350]

2.3.2.4 Representative Agents

For Lucas, *foundation in microeconomic theory* meant that macroeconomic models had to account for the rationality of agents regarding future expectations. But individualistic behavior is difficult to incorporate in equilibrium models, which represent the economy by a single set of mathematical functions. The mainstream DSGE practice is to represent individual agents as a single homogeneous RARE (Representative Agent with Rational Expectations) agent: a hyper-rational actor whose behavior captures all relevant micro-level dynamics. As a result, DSGE models cannot truly account for macroeconomic features that depend on heterogeneous differences between agents, such as borrowing behavior. It did not take long before this approach was severely criticized:

> the 'representative' agent deserves a decent burial, as an approach to economic analysis that is not only primitive, but fundamentally erroneous. [Kirman, 1992, 119]

2.3.3 Agent-Based Models

In his 1986 *Rationality of Self and Others in an Economic System*, Kenneth Arrow questioned the soundness of the DSGE interpretation of the General Equilibrium model which he had helped found with his proof of equilibrium existence in 1954. Arrow pointed out that rationality, by itself, is a weak assumption: even assuming perfectly rational agents at the micro level, DSGE models predict a whole spectrum of (different) macroeconomic outcomes. To restrict the possible outcomes in order to make meaningful predictions, modelers have to make additional, often questionable assumptions:

> the [DSGE] researcher is tempted into some strong assumptions. In particular, the homogeneity assumption seems to me to be especially dangerous. It denies the fundamental assumption of the economy, that it is built on gains from trading arising from individual differences. Further, it takes attention away from a very important aspect of the economy, namely, the effects of the distribution of income and of other individual characteristics on the workings of the economy. [Arrow, 1986, 390]

In his concluding remarks, Arrow accepted "the insight of Herbert Simon on the importance of recognizing that rationality is bounded" [Arrow, 1986, 398] (see chapter 3), and called for "a more consistent assumption of computability in the formulation of economic hypotheses" [Arrow, 1986, 398].

To spearhead the development of new modeling techniques, Arrow teamed up with physicist Philip Anderson, chairing the now-famous 1987 *Evolutionary Paths of the Global Economy* conference at the fledgling Santa Fé Institute [Arthur, 2010]. The objective of the conference was to bring together economists and natural scientists, to introduce the natural scientists to the basic relevant work in economic analysis, and economists to the basic techniques for dealing with complex dynamical systems developed in the natural sciences. The conference was a success, and the Santa Fé Institute based its first research program on its results. The proceedings of the conference were published in the 1988 *The Economy as an Evolving Complex System* [Anderson et al., 1988].

The efforts at the Santa Fé Institute led, via Artifical Adaptive Agents [Holland and Miller, 1991], Evolutionary Economics [Lane, 1993] and Agent-Based Computational Economics [Tesfatsion, 2002], to macroeconomic Agent-Based Modeling [Arthur, 2010, 152].

2.3.3.1 Building blocks

Agent-Based Models (ABMs) are a radical departure from the way General Equilibrium models are constructed. Instead of a system of mathematical functions, ABMs use object-oriented programming techniques [Tesfatsion, 2002, 14-15] to construct complex adaptive systems based on atomic building blocks called *agents* [Holland and Miller, 1991, Wooldridge and Jennings, 1995]: programmed entities representing the individual actors of the economy. Macro-behavior is the emergent result of the interaction between these agents.

Giorgio Fagiolo identifies the main ingredients of economic ABMs [Fagiolo et al., 2007]:

- *a bottom-up perspective* Agents are the individual building blocks of the model; the aggregate macro-properties of the model emerge from the individual behavior of agents and the interaction between them.

- *heterogeneity* Agents can be given different properties, meaning ABMs can simulate the result of interactions between agents with heterogeneous preferences and behaviors.
- *bounded rationality* Because agent behavior is programmed, it can be made to deviate from the rationality assumptions that characterize traditional quantitative economic models.
- *networked direct interactions* The connections between the agents can be made to contain structures, such as subgroups of agents or local networks.

2.3.3.2 Strengths and Weaknesses

In a sense, macroeconomic ABMs are a generalization of General Equilibrium models; ABMs programmed with the assumptions of General Equilibrium models should come to identical conclusions:

> Agent-based, non-steady-state economics is also a generalization of equilibrium economics. Out-of-equilibrium systems may converge to or display patterns that are consistent — that call for no further adjustments. If so, standard equilibrium behavior becomes a special case [of agent-based] behavior. [Arthur, 2006, 4]

But because ABMs are simulated on a computer instead of analytically solved in a system of equations, the potential variation in the behavior of agents, and the interaction between them, is unrestricted. This is a strength, but also a potential weakness, as the outcome of an ABM is the emergent result of the choices made by the programmer for the behavior of the agents. If the underlying micro-behavior is incorrectly specified, the whole ABM will be "wrong". But one should be careful not to overstate this argument: General Equilibrium results are also critically dependent on the assumptions of the modeler. The only real difference is that the DSGE modeler imposes their assumptions in more constrained, mathematical form.

2.3.4 The Great Recession and Beyond

Following the 1980s, the "new neoclassical synthesis" grounded on the DSGE model became the new dominant macroeconomic model for policy analysis. But in the aftermath of the financial crisis of 2008, criticism was scathing and widespread. In his 2011 *Rethinking Macroeconomics: What Failed, and How to Repair it*, Joseph Stiglitz pointed out that:

> Prediction is the test of a scientific theory. But when subject to the most important test — the one whose results we really cared about —

> the standard macroeconomic models failed miserably. [Stiglitz, 2011, 592]

Yet despite the many calls for a re-evaluation of policy analysis, and while macroeconomic ABMs such as [Deissenberg et al., 2008] have been proposed, an overview of the literature on economic forecasting shows that macroeconomic Agent-Based Models have so far failed to gain traction. Princeton University Press' 2016 *Economic Forecasting* [Timmermann and Elliott, 2016], for example, doesn't even mention agent-based models at all.

2.4 The Microfoundations Delusion

> DSGE models are very poor in forecasting, but so are all other approaches. [Edge and Gurkaynak, 2011]

Given the known shortcomings of DSGE models[1], why have agent-based models so far not taken a more prominent role in economic analysis? We argue that DSGE models and agent-based models suffer from different forms of *microfoundations delusion* (borrowing the phrase from [King, 2012]). Simply put, neither DSGE models nor macroeconomic ABMs are microfounded in any meaningful sense. This is especially problematic for agent-based models, whose methodology is explicitly based on the assumption that individual behavior can explain macro results.

The case for a proper understanding of human decision-making is as relevant today as when Herbert Simon made it in the 1960s (see Models and Reality, above). But realistic decision-making is only half the story. Interdependence, interaction, networks and trust are now understood to be vital elements in economic behavior. In his 2011 *Complex economics: individual and collective rationality*, Kirman pointed out that:

> it is the interaction between individuals that is at the heart of the explanation of many macroeconomic phenomena and [...] we will never move towards an understanding of such phenomena without considering this interaction as a central feature of the economy [Kirman, 2011, 7]

Macroeconomic agent-based models must get both the decision-making behavior of individuals, and the complex interaction between them, right. This is a daunting task. We take a deeper look at Complexity in chapter 4.

The case against microfounded DSGE models is straightforward: empirical evidence from the last 30 years overwhelmingly rejects the model of the rational

[1] For a recent comparison between DSGE models and ABMs, see [Fagiolo and Roventini, 2016].

economic agent. In addition, DSGE models often make assumptions that are in direct conflict with micro evidence (see Arthur, above; more recently Korinek: "DSGE models [...] frequently impose a number of restrictions that are in direct conflict with micro evidence" [Korinek, 2017, 8]).

DSGE models may still be founded in (current) microeconomic theory. But microeconomic theory itself, when based on perfectly rational, forward looking agents, is unfounded in reality. We take a deeper look at human decision-making, and rationality in particular, in the next chapter.

References

[Anderson et al., 1988] Anderson, P., Arrow, K., and Pines, D. (2018(1988)). *The Economy as an Evolving Complex System*, volume 5 of *Santa Fe Institute studies in the sciences of complexity*. Westview Press, CRC.

[Arrow, 1967] Arrow, K. J. (1967). Samuelson collected. *Journal of Political Economy*, 75(5):730–737.

[Arrow, 1986] Arrow, K. J. (1986). Rationality of self and others in an economic system. *Journal of Business*, pages S385–S399.

[Arrow and Debreu, 1954] Arrow, K. J. and Debreu, G. (1954). Existence of an equilibrium for a competitive economy. *Econometrica: Journal of the Econometric Society*, pages 265–290.

[Arthur, 2006] Arthur, W. B. (2006). Out-of-equilibrium economics and agent-based modeling. *Handbook of computational economics*, 2:1551–1564.

[Arthur, 2010] Arthur, W. B. (2010). Complexity, the santa fe approach, and non-equilibrium economics. *History of Economic Ideas*, pages 149–166.

[Christ, 1994] Christ, C. F. (1994). The cowles commission's contributions to econometrics at chicago, 1939-1955. *Journal of Economic Literature*, 32(1):30.

[Debreu, 1974] Debreu, G. (1974). Excess demand functions. *Journal of mathematical economics*, 1(1):15–21.

[Deissenberg et al., 2008] Deissenberg, C., Van Der Hoog, S., and Dawid, H. (2008). Eurace: A massively parallel agent-based model of the european economy. *Applied Mathematics and Computation*, 204(2):541–552.

[Dimand, 2019] Dimand, R. W. (2019). Léon walras, irving fisher and the cowles approach to general equilibrium analysis. In *10th conference of the International Walras Association, University of Lausanne*, pages 13–14.

[Edge and Gurkaynak, 2011] Edge, R. M. and Gurkaynak, R. S. (2011). How useful are estimated dsge model forecasts? *Available at SSRN 1810075*.

[Fagiolo et al., 2007] Fagiolo, G., Moneta, A., and Windrum, P. (2007). A critical guide to empirical validation of agent-based models in economics: Methodologies, procedures, and open problems. *Computational Economics*, 30(3):195–226.

[Fagiolo and Roventini, 2016] Fagiolo, G. and Roventini, A. (2016). Macroeconomic policy in dsge and agent-based models redux: New developments and challenges ahead. *LEM Working Paper Series*, (2016/17).

[Fisher, 1941] Fisher, I. (1941). Mathematical method in the social sciences. *Econometrica: Journal of the Econometric Society*, pages 185–197.

[Friedman, 1953] Friedman, M. (1953). The methodology of positive economics. Reprinted in *The philosophy of economics: An Anthology* (2007), pages 145–179.

[Gul and Pesendorfer, 2008] Gul, F. and Pesendorfer, W. (2008). The case for mindless economics. *The foundations of positive and normative economics: A handbook*, 1:3–42.

[Hanusch and Pyka, 2007] Hanusch, H. and Pyka, A. (2007). *Elgar companion to neo-Schumpeterian economics*. Edward Elgar Publishing.

[Hausman, 2007] Hausman, D. M. (2007). *The philosophy of economics: an anthology*. Cambridge University Press, 3 edition.

[Hausman, 2018] Hausman, D. M. (2018). Philosophy of economics. *The Stanford Encyclopedia of Philosophy (Fall 2018 Edition)*. https://plato.stanford.edu/archives/fall2018/entries/economics/.

[Holland and Miller, 1991] Holland, J. H. and Miller, J. H. (1991). Artificial adaptive agents in economic theory. *The American Economic Review*, 81(2):365–370.

[Homans, 1987] Homans, G. (1987). *Behaviourism and After*.

[Keynes, 1936] Keynes, J. M. (2018(1936)). *The general theory of employment, interest, and money*. Springer.

[King, 2012] King, J. E. (2012). *The microfoundations delusion: metaphor and dogma in the history of macroeconomics*. Edward Elgar Publishing.

[Kirman, 2011] Kirman, A. (2011). *Complex economics: individual and collective rationality*. Routledge.

[Kirman, 1992] Kirman, A. P. (1992). Whom or what does the representative individual represent? *Journal of economic perspectives*, 6(2):117–136.

[Korinek, 2017] Korinek, A. (2017). Thoughts on dsge macroeconomics: Matching the moment, but missing the point?

[Lane, 1993] Lane, D. A. (1993). Artificial worlds and economics, part i. *Journal of evolutionary economics*, 3(2):89–107.

[Lawson, 2017] Lawson, T. (2017). What is wrong with modern economics, and why does it stay wrong? *Journal of Australian Political Economy, The*, (80):26–42.

[Liu, 1960] Liu, T.-C. (1960). Underidentification, structural estimation, and forecasting. *Econometrica*, 28(4):855.

[Lucas Jr, 1976] Lucas Jr, R. E. (1976). Econometric policy evaluation: A critique. In *Carnegie-Rochester conference series on public policy*, volume 1, pages 19–46. North-Holland.

[Lucas Jr and Prescott, 1974] Lucas Jr, R. E. and Prescott, E. C. (1974). Equilibrium search and unemployment. *Journal of Economic theory*, 7(2):188–209.

[Lucas Jr and Sargent, 1979] Lucas Jr, R. E. and Sargent, T. J. (1979). After keynesian macroeconomics. *Federal Reserve Bank of Minneapolis Quarterly Review*, 3(2):295–319.

[Mantel, 1974] Mantel, R. R. (1974). On the characterization of aggregate excess demand. *Journal of economic theory*, 7(3):348–353.

[Mill, 1843] Mill, J. S. (1974(1843)). A system of logic, in robson, j.m. (ed.). *The Collected Works of John Stuart Mill*, 8.

[Morgan et al., 1990] Morgan, M. S. et al. (1990). *The history of econometric ideas*. Cambridge University Press.

[Rizvi et al., 1994] Rizvi, S. A. T. et al. (1994). The microfoundations project in general equilibrium theory. *Cambridge Journal of Economics*, 18(4):357–77.

[Samuelson, 1963] Samuelson, P. (1963). Problems of methodology - discussion. *The American Economic Review*, 53(2):231–236.

[Sargent, 1976] Sargent, T. J. (1976). A classical macroeconometric model for the united states. *Journal of Political Economy*, 84(2):207–237.

[Sargent, 1987] Sargent, T. J. (2009(1987)). *Dynamic macroeconomic theory*. Harvard University Press.

[Shiller, 1978] Shiller, R. J. (1978). Rational expectations and the dynamic structure of macroeconomic models: A critical review. *Journal of monetary economics*, 4(1):1–44.

[Simon, 1963] Simon, H. A. (1963). Problems of methodology - discussion. *The American Economic Review*, 53(2):229–231.

[Sims, 1980] Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, pages 1–48.

[Smets and Wouters, 2007] Smets, F. and Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American economic review*, 97(3):586–606.

[Sonnenschein, 1973] Sonnenschein, H. (1973). Do walras' identity and continuity characterize the class of community excess demand functions? *Journal of economic theory*, 6(4):345–354.

[Stiglitz, 2011] Stiglitz, J. E. (2011). Rethinking macroeconomics: What failed, and how to repair it. *Journal of the European Economic Association*, 9(4):591–645.

[Tesfatsion, 2002] Tesfatsion, L. (2002). Agent-based computational economics: Growing economies from the bottom up. *Artificial life*, 8(1):55–82.

[The Econometric Society, 2019] The Econometric Society (2019). Econometrica. https://onlinelibrary.wiley.com/journal/14680262.

[Timmermann and Elliott, 2016] Timmermann, A. and Elliott, G. (2016). *Economic Forecasting*. Princeton University Press.

[Tobin, 1986] Tobin, J. (1986). The future of keynesian economics. *Eastern Economic Journal*, 12(4):347–356.

[Wooldridge and Jennings, 1995] Wooldridge, M. and Jennings, N. R. (1995). Intelligent agents: Theory and practice. *The knowledge engineering review*, 10(2):115–152.

Chapter 3

Models of Man

3.1 The Rational Utility Maximization Hypothesis

The field of economics has a long history of modeling human decision-making by assuming rationality [McCormick, 1997], succinctly described by Stigler as:

> There are three characteristics of a rational consumer:
> 1. His tastes are consistent.
> 2. His cost calculations are correct.
> 3. He makes those decisions that maximize utility. [Stigler, 1987, 52]

According to Joseph Schumpeter, the idea that individuals make decisions by maximizing utility goes back at least as far as François Quesnay (1694-1774):

> [Quesnay] manifestly thought that if every individual strives to realize maximum satisfaction, then all individuals will "of course" achieve maximum satisfaction. The fact that one of the best brains of our science could have been content with such an obvious non sequitur is indeed food for thought: low standards of rigor and sloppiness of thinking have been worse enemies of scientific economics than has been political bias. [Schumpeter, 1954, 233]

Interestingly, the early classical economic thinkers did not consider the economy as consisting of rational utility-calculating agents. Adam Smith, for example, in his 1759 *The Theory of Moral Sentiments* writes that a person's

> own passions are very apt to mislead him; sometimes to drive him and sometimes to seduce him to violate all the rules which he himself, in all his sober and cool hours, approves of. [Smith, 1759]

It is only with the gradual introduction of mathematics into economic thinking, which Irving Fisher contends started with Jevons in 1871 [Fisher, 1892], that assumptions of utility maximization enter the economic discourse. The appeal of maximization is clear: it can easily be expressed in mathematical form. If utility is expressed as a function, then maximizing utility is a matter of differentiating the utility function and setting the result equal to zero (the full requirement is that the utility function is twice differentiable). Especially in the age before computers, mathematical solvability was an important advantage, even if it meant a significant simplification of reality.

3.2 Bounded Rationality

In a series of seminal publications [Simon, 1955, Simon, 1956, Simon, 1957], Herbert A. Simon was the first to provide a systematic exploration of the limits to rationality. Simon was awarded the Alfred Nobel Memorial Prize in Economic Sciences in 1978 for "his pioneering research into the decision-making process within economic organizations" [The Royal Swedish Academy of Sciences, 1978]. He later reflected:

> The big problem for me, from the very beginning of my work with Clarence Ridley, was to reconcile the way that decisions were actually made in organizations with the way that the economists pretended that they were made. [Golembiewski, 1988]

Simon worked on this problem throughout his academic career. In *Administrative Behavior*, published in 1947 and based on his doctoral dissertation, Simon explained the problems of the rational utility maximization hypothesis:

> The implication might be drawn from this discussion that any rational choice between alternatives involves a complete description of the possibility consequential on each alternative and a comparison of these possibilities. We would have to know in every single respect how the world would be changed by our behaving one way instead of another, and we would have to follow the consequences of behavior through unlimited stretches of time, unlimited reaches of space, and unlimited sets of descriptive variables. Under such conditions even an approach to rationality in real behavior would be inconceivable. [Simon, 1947, 38]

Simon observed that instead "behavior is determined by the irrational and non-rational elements that bound the area of rationality" [Simon, 1947, 241]. This was the earliest statement of what would become Bounded Rationality, formalized by Simon in his 1957 *Models of Man: Social and Rational* [Simon, 1957].

To be clear, Simon did not invent the concept of limits to rationality. The semantic field of Bounded Rationality has a long history (see Table 3.1) which parallels the use of the rational utility maximization hypothesis in economics. People have always realized that perfect rationality is a stretch.

Keyword Heuristic											**Total**
Limited Intelligence	1840										73
Finite Intelligence		1880									40
Incomplete Rationality			1922								7
Limited Rationality				1945							105
Administrative Rationality					1947						48
Approximate Rationality						1948					9
Bounded Rationality							**1957**				**626**
Finite Rationality								1972			17
Constrained Rationality									1978		7
Restricted Rationality										1983	1
Bounded Intelligence											0

Table 3.1: The emergence of the semantic field of Bounded Rationality, with term employed, year of earliest occurrence, and total occurrences in the JSTOR database [Klaes et al., 2005]

What made Simon *incontournable* was the fact that his 1955-1957 publications treated the concept of limited rationality in mathematical form, thereby turning the expression into a technical term requiring a model: "The broader aim, however, in constructing these definitions of 'approximate' rationality is to provide some materials for the construction of a theory of the behavior of a human individual or of groups of individuals who are making decisions in an organizational context" [Simon, 1955, 114]. Simon had appealed to economists in their own language. Behavioral Economics was born.

3.2.1 Satisficing

Simon distinguished between reaching an *optimal* solution, which requires full understanding of the problem, and the feasibility of reaching a *satisfactory* solution by comparing results to a minimum pay-off. He pointed out that the process of reaching a decision is drastically simplified if the criterion of success is determined by the solution being merely acceptable [Simon, 1955, 108]. Simon called this process "satisficing" in contrast to "maximizing". In his 1956 *Rational Choice and the Structure of the Environment*, he observed:

> Evidently, organisms adapt well enough to "satisfice"; they do not, in general, "optimize." [Simon, 1956, 129]

Simon concluded that maximization assumes individual behavior that is unnecessarily complicated for coping with the environment, and that economic theory should be more closely in agreement with how humans actually behave in the real world:

> In particular, no "utility function" needs to be postulated for the organism, nor does it require any elaborate procedure for calculating marginal rates of substitution among different wants.
>
> The analysis set forth here casts serious doubt on the usefulness of current economic and statistical theories of rational behavior as bases for explaining the characteristics of human and other organismic rationality. It suggests an alternative approach to the description of rational behavior that is more closely related to psychological theories of perception and cognition, and that is in closer agreement with the facts of behavior as observed in laboratory and field. [Simon, 1956, 138]

This last paragraph illustrates an epistemological distinction between Simon and many others: Simon believed that in order to model economic macro-behavior, one should start from the reality of human behavior. This approach would later become the premise of agent-based modeling.

3.2.2 Heuristics

Heuristics are 'rules of thumb' which the human mind employs to solve complex problems by generating an approximating solution [Romanycia and Pelletier, 1985]. Simon first explored the concept of heuristics in his 1956 paper, to refer to simple strategies that generate satisfactory results in complex environments. He employed as an illustration a simple organism that can observe its environment and move through it, and which needs a strategy to find food in such a way that it will not starve. Simon submitted that the rational way for the organism to behave is to explore the surface at random looking for food, which it eats when it finds it. If this is satisfactory, the organism will spend the rest of the time resting before resuming its search for a new meal.

Simon pointed out that this behavior is not particularly remarkable, except in comparison to the behavior postulated by economists, which involves the organism calculating the optimal food gathering pattern before making a decision, a course which would require full knowledge of the environment, as well as high cognitive ability to compare all available options.

3.2.3 Computational Complexity

To understand the role of heuristics in human decision-making, it is useful to look at the study of computational complexity. If human beings solve decision prob-

lems analytically as suggested by the rational utility maximization hypothesis, then the computational complexity of decision problems is relevant to human decision-making: whereas simple problems could be solved analytically, more complex problems defying analytical solutions would require heuristic workarounds[1].

Computational Complexity Theory, or simply Complexity Theory - not to be confused with Complexity Science, see chapter 4 - has been a central discipline of theoretical computer science since the mid 1960s. It provides a tool for the classification of the complexity of feasible problems [Du and Ko, 2014].

The still-unresolved question of $P \stackrel{?}{=} NP$ lies at the center of complexity theory. In terms of stylized facts it is relatively easy to understand: let *P* be the set of problems that can be solved analytically by a modern computer[2]. Let *NP* be the set of problems with no known analytical solution[3]. If P = NP, then these still-unsolved problems are analytically solvable, but we haven't found the correct algorithms yet. If the answer is no, then these problems by their very nature defy analytical solution and, important for our purpose, cannot be analytically solved by a human brain[4]. It is generally considered likely that $P \neq NP$, but this is still unproven. It is one of the seven 'millennium problems' for which the Clay Mathematics Institute offers a 1$ million reward. [CMI, 2019]

There are many known real-world problems, sometimes surprisingly common ones, that (for now) defy analytical solution. The optimal solution to these problems can only be found by simulating every possibility and comparing the results. This is where heuristics come in. Within the class of problems that for now cannot be solved analytically, we distinguish two categories[5]:

- NP: a possible solution to these problems can be *verified*, but finding the *optimal* solution requires a brute-force evaluation of all possible options. An example of such a problem is the code for opening a safe: evaluating a given code against the correct solution is easy, but solving the problem without knowing the correct code requires brute force experimenting until a working (satisfying) combination is found.

[1] Computational Complexity deals with the question of whether problems have an *algorithmic* solution. In economics we can use the two terms interchangeably, since modern analytical models are implemented on a computer: problems that are algorithmically solvable in computer science terms are analytically solvable in mathematical terms.

[2] P stands for Polynomial: a deterministic Turing machine, like a modern computer, can solve these problems in polynomial time.

[3] NP stands for Non-deterministic Polynomial: a Non-deterministic Turing machine could solve these problems in polynomial time. But these machines do not (yet) exist. Normal computers can only "solve" these problems by testing all the possible combinations and comparing the results.

[4] Unless the human brain is a non-deterministic Turing machine, which for now it is understood not to be.

[5] Complexity Theory recognizes a third category, that of NP-complete problems, but this term isn't useful for our analysis.

- NP-hard: these problems are *at least* as hard as NP. Like NP, these problems cannot be solved analytically, and in addition, verifying possible solutions also requires a full exploration of the problem space. Such is in the case in the Traveling Salesman Problem [Karp, 1972], where the validity of a suggested optimal route can't be confirmed without checking all other routes and comparing the results.

For "mere" NP problems, where possible solutions are verifiable without having to solve the entire problem, "satisficing" offers an obvious optimization to maximizing. Maximizing utility through analysis is not even possible in these situations, since an analytic solution to these problems (as far as we know) does not exist. But coming up with "satisfactory" solutions (typically through learning) and verifying their validity is a feasible strategy.

3.2.4 Artificial Intelligence

Simon's early insights into human decision-making were instrumental to the emerging field of Artificial Intelligence. The 2016 *Artificial Intelligence: A Modern Approach* describes how, at the two-month workshop at Dartmouth in the summer of 1956 where the term Artificial Intelligence found its first official usage, Simon and his co-author Allen Newell were well ahead of the curve:

> Two researchers from Carnegie Tech, Allen Newell and Herbert Simon, rather stole the show. Although the others had ideas and in some cases programs for particular applications such as checkers, Newell and Simon already had a reasoning program, the Logic Theorist (LT) ...
>
> Soon after the workshop, the program was able to prove most of the theorems in Chapter 2 of Russell and Whitehead's Principia Mathematica. Russell was reportedly delighted when Simon showed him that the program had come up with a proof for one theorem that was shorter than the one in Principia. [Russell and Norvig, 2016, 17]

In 1958, Simon and Newell published *Heuristic Problem Solving: The Next Advance in Operations Research*, in which they explained how new insights in human cognitive ability could be used to make computers think:

> [W]e now have the elements of a theory of heuristic (as contrasted with algorithmic) problem solving; and we can use this theory both to understand human heuristic processes and to simulate such processes with digital computers. [Simon and Newell, 1958, 6]

Simon and Newell continued their collaboration in the 1950s and 1960s, and integrated their ideas in their 1972 magnum opus Human Problem Solving [Newell and Simon, 1972], which treated a human as a processor of information and offered a general theory of human problem solving. For their pioneering work, in 1975 Simon and Newell were awarded the Turing Award, informally considered the Nobel Prize of Computer Science. Simon and his colleagues were credited with the creation of the first AI program:

> One of the most important outcomes of this approach to computer science was Simon's development — and strong advocacy — of heuristic programming. ... Simon and his colleagues Allen Newell and J.C. Shaw employed this notion of heuristic problem-solving in the first successful AI program, the Logic Theorist (LT) of 1955-56. [ACM, 1975]

Heuristic techniques based on Simon's approach are still used in the field of artificial intelligence today. Home energy management controllers for renewable energy [Rahim et al., 2016], drone delivery routing [Gahm et al., 2019], and machine scheduling [Kitjacharoenchai et al., 2019] are some examples.

3.3 Bounded Rationality Unbound

It is ironic that, while Simon's insights in human decision-making were embraced by computer scientists to develop *artificial* intelligence, within the field of Economics, Simon's ideas would soon be subverted, or downright rejected.

Two major counter-movements against Bounded Rationality became mainstream in the 1970s: the 'Chicago school' of economic thought based on the theory of Rational Expectations advocated by Lucas and Sargent [Sent, 1997], and the so-called 'heuristics and biases literature of the 1970s' pioneered by Tversky and Kahneman which, although seemingly sympathetic to Simon's work, framed heuristics as the source of systematic errors against rationality [Sent, 2005].

3.3.1 Rational Expectations

In 1961, John Muth published *Rational Expectations and the Theory of Price Movements*. In it, he directly challenged Simon and his concept of Bounded Rationality, instead arguing that economic models do not assume *enough* rationality:

> It is sometimes argued that the assumption of rationality in economics leads to theories inconsistent with, or inadequate to explain, observed phenomena, especially changes over time (e.g., Simon [29]). Our

> hypothesis is based on exactly the opposite point of view: that dynamic economic models do not assume enough rationality. [Muth, 1961, 316]

Muth's central thesis was that people's expectations are essentially the same as the predictions of economic theory. Though individuals may not have perfect information, Muth argued, efficient markets integrate the available information in such a way that population expectations reflect the structure of the entire system - through markets, people's population expectations are always correct.

Muth made an important claim to illustrate his argument: under rational expectations, publicly-made predictions will have no substantial effect on the operation of the economic system; only predictions based on inside information can have this effect [Muth, 1961, 318]. The theory of Rational Expectations and the Efficient Market Hypothesis (that prices reflect all available information, and only inside information is useful for generating above-market returns) are not identical, but it is clear that the two concepts are closely related.

The theory of Rational Expectations was hugely successful in terms of publications. In his 1983 *A Rational Expectations Approach to Macroeconometrics*, Frederic Mishkin claimed "a veritable explosion in the number of empirical studies that either use or can use" the theory, and provided an exhaustive list of 44 publications to convince the reader of the productivity of the theory [Mishkin, 1983, 2-3]. Rational Expectations remains one of the fundamental underlying concepts of modern DSGE models today [Stiglitz, 2018].

3.3.2 Prospect Theory

In the early seventies, cognitive psychologists Amos Tversky and Daniel Kahneman published a series of papers exploring the role of heuristics in the formation of subjective probabilities [Kahneman and Tversky, 1972, Tversky and Kahneman, 1973, Tversky and Kahneman, 1974]. Their findings entered the economic discourse with *Prospect Theory: An Analysis of Decision under Risk*, published in *Econometrica* in 1979 [Kahneman and Tversky, 1979]. Surprisingly, nowhere in these papers do the authors define the term 'heuristic', or refer to previous publications explaining the term. Neither do they cite Simon's previous work on heuristics and Bounded Rationality, or even mention Simon at all (Kahneman would rectify this in 2003). Nonetheless, the similarities are clear:

> Less obvious, however, is the fact that the deviations of subjective from objective probability seem reliable, systematic, and difficult to eliminate. Apparently, **people replace the laws of chance by heuristics,** which sometimes yield reasonable estimates and quite often do not. [Kahneman and Tversky, 1972, 431]

In the introduction of their 1973 paper *Availability: A Heuristic for Judging Frequency and Probability*, the authors write:

> We propose that when faced with the difficult task of judging probability or frequency, people employ a limited number of heuristics which reduce these judgments to simpler ones. [Tversky and Kahneman, 1973, 209]

And in the introduction of the 1974 *Science* paper *Judgment under Uncertainty: Heuristics and Biases*, we find:

> This article shows that people rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations. [Tversky and Kahneman, 1974, 1124]

In *Prospect Theory*, the 1979 mathematical formulation of their previous publications, Kahneman and Tversky made no mention of Simon's Bounded Rationality, instead appearing to discover the concept independently. The abstract opens with the sentence:

> This paper presents a critique of expected utility theory as a descriptive model of decision-making under risk, and develops an alternative model, called prospect theory. [Kahneman and Tversky, 1979, 263]

Once Prospect Theory took off, Tversky and Kahneman became the new fathers of Behavioral Economics. Simon was largely forgotten. In 2004, Brian Loasby observed: "All that most economists know about Herbert Simon is that he wrote about bounded rationality and organizational behavior" [Simon et al., 2004, 259].

Crucially, Tversky and Kahneman's treatment of heuristics was very different from Simon's. Whereas to Simon, heuristics were useful tools for dealing with complex problems, Tversky and Kahneman treated heuristics as the source of systematic errors of judgement:

> This article has been concerned with cognitive biases that stem from the reliance on judgmental heuristics. ... several of the severe errors of judgment reported earlier occurred despite the fact that subjects were encouraged to be accurate and were rewarded for the correct answers. ... The reliance on heuristics and the prevalence of biases are not restricted to laymen. Experienced researchers are also prone to the same biases - when they think intuitively. [Tversky and Kahneman, 1974, 1130]

In other words, heuristics were to blame for *deviations from* rational behavior. If only people were more rational, these "severe errors of judgment" might not occur. Rationality was restored to normative behavior in economics. Heuristic reasoning was deviant, and had to be "meliorated" [Stanovich, 1999]. In her 2005 *Simplifying Herbert Simon*, Esther-Mirjam Sent summarized:

> Kahneman and Tversky's insights were opposite those of Simon in the sense that they started from the rationality assumption that has characterized mainstream economics and next analyzed departures from this yardstick, rather than developing an alternative one. This explains the difference between old behavioral economics, which draws on Simon's work, and new behavioral economics, which takes off from the research of Kahneman and Tversky. [Sent, 2005, 230]

3.4 Two Minds in One Brain

During the eighties and nineties, new empirical research revealed a profound disconnect between academic theory and real world observations. The concept of efficient markets came under increasingly heavy fire [Minsky, 1986, Grossman and Stiglitz, 1980], and empirical experiments showed that expectations of lay people [Jonung and Laidler, 1988] and experts alike [Brown and Maital, 1981] didn't conform to the predictions of Rational Expectations theory.

3.4.1 Dual-System Models

The main theme of the 1970s and early 1980s had been demonstrating that actual, descriptive human behavior diverged from normative, rational behavior (for an overview see [Kahneman et al., 1982]). Trying to make sense of the gap between descriptive and normative behavior, then, was the subject of often heated debate during the 1980s and 1990s [Cohen, 1981, Kahneman and Tversky, 1983, Gigerenzer, 1996].

During this period, so-called two-process theories of reasoning came to prominence [Evans, 1984, Epstein, 1994]. In his 1994 *Integration of the Cognitive and the Psychodynamic Unconscious*, Epstein noted "convergence of the views on two processing systems, one automatic and intuitive and the other abstract and analytical" [Epstein, 1994, 714]. He proposed a global theory of personality, which:

> distinguishes an experiential system that is intimately associated with affect ..., and that operates according to a set of inferential rules that differ from those of a relatively affect-free, abstract, analytical, rational system. The two systems are assumed to operate in parallel and to interact with each other" [Epstein, 1994, 713].

In their seminal 2000 *Individual differences in reasoning: implications for the rationality debate?*, Stanovich and West reviewed the large literature of empirical studies on heuristics and biases, and demonstrated that a subset of these results could only be explained by a dual-system rationality based on two modes of cognition, which they labeled System 1 and System 2 [Stanovich and West, 2000, 658].

Repeating some of the experiments of Tversky and Kahneman [Stanovich and West, 1999], Stanovich and West had demonstrated that in some experiments, the "correct" answer by normative, rational logic tended to be given by participants with *lower* cognitive ability, against what would be expected. Also, in some experiments, participants on average did not change their opinion even after they were given the correct answer, clearly believing that an alternative answer was more rational than the normative answer. For example, after participating in a first administration of the famous Prisoner's Dilemma [Stanovich and West, 1999, 366-369], participants were given the (presumably correct) normative answer, along with a non-normative alternative, and allowed to play the game a second time. There was no significant difference in the proportions of subjects who changed their answer from the non-normative to the normative answer or vice versa, indicating that the divergence in answers was due to something other than computational limitations due to cognitive ability:

> (Normative answer) "No matter what the other person does, I am better off competing. If the other person cooperates and I compete, I get $25 rather than $20. If the other person competes and I compete, I get $10 rather than $5. Competing is always the better strategy for me, so it is the better choice."
>
> (Alternative answer) "The rational thing for both of us to do is to both cooperate and get $20 rather than to both compete and get only $10. The other player probably realizes this, too, just like I do. Therefore, I should cooperate so that we both end up with $20, rather than $10."

3.5 The Return of Bounded Rationality

In August 2001 (published in 2002), Daniel Kahneman and co-author Shane Frederick accepted the validity of dual-process theories, adopting the System 1 / System 2 terminology of Stanovich and West. However, Kahneman and Frederick still argued for the negative heuristics and biases understanding, proposing that the role of System 2 (the rational system) was to correct System 1 (the intuitive system) when it makes mistakes, as illustrated by the chapter title: "Making Biases Disappear: a Task for System 2" [Kahneman and Frederick, 2002, 14]. Regardless, this publication was a milestone in the field of Behavioral Economics for its acceptance of dual-process theory:

> The ancient idea that cognitive processes can be partitioned into two main families – traditionally called intuition and reason – is now widely embraced under the general label of dual-process theories (Chaiken and Trope, 1999; Hammond, 1996; Sloman, 1996, Chapter 22 this volume). Dual-process models come in many flavors, but all distinguish cognitive operations that are quick and associative from others that are slow and governed by rules (Gilbert, 1999). We adopt the generic labels 'System 1' and 'System 2' from Stanovich and West (Chapter 24). [Kahneman and Frederick, 2002, 3]

Then, in his 2003 *A Perspective on Judgment and Choice: Mapping Bounded Rationality*, Kahneman finally lifted intuition from out of the shadow of rationality. Reflecting on his 2002 Nobel Prize, Kahneman unambiguously placed his earlier work with Amos Tversky in the territory of Herbert Simon's bounded rationality:

> The work cited by the Nobel committee was done jointly with the late Amos Tversky (1937-1996) during a long and unusually close collaboration. Together, we explored a territory that Herbert A. Simon had defined and named - the psychology of bounded rationality (Simon, 1955, 1979). [Kahneman, 2003, 697]

Employing a 1973 example from Simon and co-author Chase, Kahneman recognized that the intuitive mind can develop high expertise on its own (through learning), rather than having to be corrected by the rational mind:

> In the examples discussed so far, intuition was associated with poor performance, but intuitive thinking can also be powerful and accurate. High skill is acquired by prolonged practice, and the performance of skills is rapid and effortless. The proverbial master chess player who walks past a game and declares, "White mates in three," without slowing is performing intuitively (Simon & Chase, 1973), as is the experienced nurse who detects subtle signs of impending heart failure (Gawande, 2002; Klein, 1998). Klein (2003, chapter 4) has argued that skilled decision makers often do better when they trust their intuitions than when they engage in detailed analysis. [Kahneman, 2003, 699]

Daniel Kahneman went on to champion dual-system thinking, introducing it to the wider audience through his seminal *Thinking, Fast and Slow* [Kahneman, 2011].

3.6 The Emerging Science of Heuristics

The 21st century has revolutionized our understanding of heuristics. Recent studies are demonstrating its usefulness in human decision-making [Todd and Gigerenzer, 2003, Goldstein and Gigerenzer, 2008]. Heuristics are making their way into scientific fields like social psychology [Strack and Deutsch, 2015] and neuroscience [Spunt, 2015]. Gigerenzer and Gaissmaier in their 2011 *Heuristic decision-making* even referred to an "emerging science of heuristics", giving a modern interpretation of heuristics as "efficient cognitive processes, conscious or unconscious, that ignore part of the information." [Gigerenzer and Gaissmaier, 2011, 451]

Similarly, dual system models have been applied in scientific fields such as social psychology [Cushman, 2013, Strack and Deutsch, 2015], consumer psychology [Samson and Voyer, 2012], artificial intelligence [Lieto et al., 2015], and design science [Kannengiesser and Gero, 2019]. Not everyone agrees that dual-system models are an accurate representation of human cognition [Rustichini, 2008, Mugg, 2016]. But this a healthy state of ongoing research, which is where we are today.

How heuristics work, how they fit into dual-system models, and how the heuristic and the rational mind interact (or whether they really are distinct systems at all) is still to be determined by further research. With the workings of the human brain still up in the air, it is no wonder that economic models based on these insights are still in their infancy.

References

[ACM, 1975] ACM (1975). The turing award. https://amturing.acm.org/award_winners/simon_1031467.cfm.

[Brown and Maital, 1981] Brown, B. W. and Maital, S. (1981). What do economists know? An empirical study of experts' expectations. *Econometrica: Journal of the Econometric Society*, pages 491–504.

[CMI, 2019] CMI (2019). Millennium problems. http://www.claymath.org/millennium-problems/p-vs-np-problem.

[Cohen, 1981] Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, 4(3):317–331.

[Cushman, 2013] Cushman, F. (2013). Action, outcome, and value: A dual-system framework for morality. *Personality and social psychology review*, 17(3):273–292.

[Du and Ko, 2014] Du, D.-Z. and Ko, K.-I. (2014). *Theory of computational complexity*. Hoboken, New Jersey : John Wiley & Sons.

[Epstein, 1994] Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American psychologist*, 49(8):709.

[Evans, 1984] Evans, J. S. B. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology*, 75(4):451–468.

[Fisher, 1892] Fisher, I. (1892). *Mathematical investigations in the theory of value and prices, and appreciation and interest*.

[Gahm et al., 2019] Gahm, C., Kanet, J. J., and Tuma, A. (2019). On the flexibility of a decision theory-based heuristic for single machine scheduling. *Computers & Operations Research*, 101:103–115.

[Gigerenzer, 1996] Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to kahneman and tversky.

[Gigerenzer and Gaissmaier, 2011] Gigerenzer, G. and Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, 62:451–482.

[Goldstein and Gigerenzer, 2008] Goldstein, D. G. and Gigerenzer, G. (2008). The recognition heuristic and the less-is-more effect. *Handbook of experimental economics results*, 1:987–992.

[Golembiewski, 1988] Golembiewski, R. T. (1988). Nobel laureate simon'looks back': A low-frequency mode. *Public Administration Quarterly*, 12(3):275.

[Grossman and Stiglitz, 1980] Grossman, S. J. and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American economic review*, 70(3):393–408.

[Jonung and Laidler, 1988] Jonung, L. and Laidler, D. (1988). Are perceptions of inflation rational? some evidence for swedens. *The American Economic Review*, 78(5):1080–1087.

[Kahneman, 2003] Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *American psychologist*, 58(9):697.

[Kahneman, 2011] Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

[Kahneman and Frederick, 2002] Kahneman, D. and Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The psychology of intuitive judgment*, 49:81.

[Kahneman et al., 1982] Kahneman, D., Slovic, S. P., Slovic, P., and Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge university press.

[Kahneman and Tversky, 1972] Kahneman, D. and Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive psychology*, 3(3):430–454.

[Kahneman and Tversky, 1979] Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263–291.

[Kahneman and Tversky, 1983] Kahneman, D. and Tversky, A. (1983). Can irrationality be intelligently discussed? *Behavioral and Brain Sciences*, 6(3):509–510.

[Kannengiesser and Gero, 2019] Kannengiesser, U. and Gero, J. S. (2019). Design thinking, fast and slow: A framework for kahneman's dual-system theory in design. *Design Science*, 5.

[Karp, 1972] Karp, R. M. (1972). Reducibility among combinatorial problems. In *Complexity of computer computations*, pages 85–103. Springer.

[Kitjacharoenchai et al., 2019] Kitjacharoenchai, P., Ventresca, M., Moshref-Javadi, M., Lee, S., Tanchoco, J. M., and Brunese, P. A. (2019). Multiple traveling salesman problem with drones: Mathematical model and heuristic approach. *Computers & Industrial Engineering*, 129:14–30.

[Klaes et al., 2005] Klaes, M., Sent, E.-M., et al. (2005). A conceptual history of the emergence of bounded rationality. *History of political economy*, 37(1):27–59.

[Lieto et al., 2015] Lieto, A., Radicioni, D. P., and Rho, V. (2015). A commonsense conceptual categorization system integrating heterogeneous proxytypes and the dual process of reasoning. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*.

[McCormick, 1997] McCormick, K. (1997). An essay on the origin of the rational utility maximization hypothesis and a suggested modification. *Eastern Economic Journal*, 23(1):17–30.

[Minsky, 1986] Minsky, H. P. (1986). Stabilizing an unstable economy.

[Mishkin, 1983] Mishkin, F. (1983). *A rational expectations approach to macroeconometrics*. U. of Chicago Press for the National Bureau of Economic Research.

[Mugg, 2016] Mugg, J. (2016). The dual-process turn: How recent defenses of dual-process theories of reasoning fail. *Philosophical Psychology*, 29(2):300–309.

[Muth, 1961] Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica: Journal of the Econometric Society*, pages 315–335.

[Newell and Simon, 1972] Newell, A. and Simon, H. A. (1972). *Human problem solving*. Prentice-Hall.

[Rahim et al., 2016] Rahim, S., Javaid, N., Ahmad, A., Khan, S. A., Khan, Z. A., Alrajeh, N., and Qasim, U. (2016). Exploiting heuristic algorithms to efficiently utilize energy management controllers with renewable energy sources. *Energy and Buildings*, 129:452–470.

[Romanycia and Pelletier, 1985] Romanycia, M. H. and Pelletier, F. J. (1985). What is a heuristic? *Computational Intelligence*, 1(1):47–58.

[Russell and Norvig, 2016] Russell, S. J. and Norvig, P. (2016). *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited,.

[Rustichini, 2008] Rustichini, A. (2008). Dual or unitary system? two alternative models of decision making. *Cognitive, Affective, & Behavioral Neuroscience*, 8(4):355–362.

[Samson and Voyer, 2012] Samson, A. and Voyer, B. G. (2012). Two minds, three ways: dual system and dual process models in consumer psychology. *AMS review*, 2(2-4):48–71.

[Schumpeter, 1954] Schumpeter, J. A. (1954). *History of Economic Analysis*. New York: Oxford University Press.

[Sent, 1997] Sent, E.-M. (1997). Sargent versus simon: bounded rationality unbound. *Cambridge Journal of Economics*, 21(3):323–338.

[Sent, 2005] Sent, E.-M. (2005). Simplifying herbert simon. *History of political economy*, 37(2):227–232.

[Simon, 1947] Simon, H. A. (1947). *Administrative behavior: A study of decision-making processes in administrative organization*. Macmillan.

[Simon, 1955] Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, 69(1):99–118.

[Simon, 1956] Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological review*, 63(2):129–138.

[Simon, 1957] Simon, H. A. (1957). Models of man; social and rational.

[Simon et al., 2004] Simon, H. A. et al. (2004). *Models of a man: Essays in memory of Herbert A. Simon*. MIT Press.

[Simon and Newell, 1958] Simon, H. A. and Newell, A. (1958). Heuristic problem solving: The next advance in operations research. *Operations research*, 6(1):1–10.

[Smith, 1759] Smith, A. (1759). *The Theory of Moral Sentiments*.

[Spunt, 2015] Spunt, R. (2015). Dual-process theories in social cognitive neuroscience.

[Stanovich, 1999] Stanovich, K. E. (1999). *Who is rational?: Studies of individual differences in reasoning*. Psychology Press.

[Stanovich and West, 1999] Stanovich, K. E. and West, R. F. (1999). Discrepancies between normative and descriptive models of decision making and the understanding/acceptance principle. *Cognitive psychology*, 38(3):349–385.

[Stanovich and West, 2000] Stanovich, K. E. and West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and brain sciences*, 23(5):645–665.

[Stigler, 1987] Stigler, G. (1987). *The Theory of Price*. New York: Macmillan Publishing Company.

[Stiglitz, 2018] Stiglitz, J. E. (2018). Where modern macroeconomics went wrong. *Oxford Review of Economic Policy*, 34(1-2):70–106.

[Strack and Deutsch, 2015] Strack, F. and Deutsch, R. (2015). The duality of everyday life: Dual-process and dual system models in social psychology. *APA handbook of personality and social psychology*, 1:891–927.

[The Royal Swedish Academy of Sciences, 1978] The Royal Swedish Academy of Sciences (1978). Alfred nobel memorial prize in economic sciences. https://www.nobelprize.org/prizes/economic-sciences/1978/press-release/.

[Todd and Gigerenzer, 2003] Todd, P. M. and Gigerenzer, G. (2003). Bounding rationality to the world. *Journal of Economic Psychology*, 24(2):143–165.

[Tversky and Kahneman, 1973] Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, 5(2):207–232.

[Tversky and Kahneman, 1974] Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131.

Chapter 4

Complexity Science

4.1 Simple Solutions for Complex Problems

The human brain handles complex problems in a variety of ways, often involving reduction to simpler, typically linear alternatives [Dawes and Corrigan, 1974]. Simplification strategies can be overall beneficial: in their 1999 *Simple Heuristics That Make Us Smart*, cognitive psychologists Gerd Gigerenzer and Peter Todd argued for a "less-is-more" effect, where less knowledge results in more accurate predictions [Gigerenzer and Todd, 1999, 45–50]. People are also capable of complex, non-linear problem solving, but need to be made aware of the advantages of doing so; absent this, people tend to rely on linear solutions even in non-linear situations [Hammond and Summers, 1965]. Improper linear deduction can lead to mistakes of intuition, and has been observed in students of mathematics [Van Dooren et al., 2008] and economics [De Bock et al., 2014]. An early example is found in Plato's dialog *Meno*, where Meno's pupil, when asked by Socrates to double the area of a square, confidently "solves" the problem by doubling the sides of the square, incorrectly applying linear logic to a quadratic situation.

Distributions are also useful tools for simplification, by reducing information about large groups to only a few key numbers, called *moments*. In the simplest distribution, where everyone in the group is identical, information can be reduced to a single *average*: the first moment. In a normal or Gaussian distribution, group information can be summarized by the average, combined with a second moment called the *variance*, which fully captures how information varies around the average. Partly due to their simplicity, normal distributions are ubiquitous in scientific models. Economic shocks, too, are typically modeled as normally distributed, even though it is well understood this does not hold in the real economy [Batiz-Zuk et al., 2015]. As a consequence, economic models in general, and DSGE models in

particular, are known to underestimate the effects of extreme shocks [Ascari et al., 2015].

Our ability to reduce complex problems to simple solutions may be an evolutionary asset, but there are situations where properly accounting for non-linearity is crucial. The failure of Lehman Brothers in 2008 (one out of approximately 7000 banks [FRED, 2020]), for example, proved to have a hugely disproportionate effect on the financial system as a whole.

4.2 What is Complexity Science?

Complexity Science, then, is the study of non-trivial relationships and distributions. It is not a single theory, but a field of interest characterized by the realisation that the whole is more than the sum of its parts, first identified by Herbert Simon in his 1962 *The Architecture of Complexity*:

> Roughly, by a complex system I mean one made up of a large number of parts that interact in a nonsimple way. In such systems, the whole is more than the sum of the parts, not in an ultimate, metaphysical sense, but in the important pragmatic sense that, given the properties of the parts and the laws of their interaction, it is not a trivial matter to infer the properties of the whole. [Simon, 1962, 468]

Complexity Science has been characterized as a 'grey' science for the 'stuff in between' [Richardson et al., 2000]. Albert-László Barabási, one of the foremost complexity researchers of this day, observed in 2007 that

> a complete theory of complexity does not yet exist. When trying to characterize complex systems, the available tools fail for various reasons. [Barabási, 2007, 33]

Despite the lack of an all-compassing definition, Arthur Holt identified three broad interpretations of complexity [Holt et al., 2011, 360-362]: (1) the *general interpretation* of complexity offered by Simon, above, (2) *computational complexity* based on the information theory of Shannon & Weaver [Shannon, 1948], and (3) *dynamic complexity* based on heterogeneous interacting agents, which is most widely used in economics, in the form of agent-based models.

4.3 Complexity Economics

Complexity economics originated in 1987 from the efforts of physicist Philip Anderson and economist Kenneth Arrow to apply complexity science to economics at the Santa Fé Institute [Arthur, 2014]. Arrow critiqued the appropriateness of the

equilibrium approach to macroeconomic modeling, and called for a new methodology (also see previous chapter). In 1997, Brian Arthur, the first director of the interdisciplinary Economics Program at Santa Fé, outlined the six fundamentals of this new approach [Arthur, 1997, 3–4]:

- Dispersed interaction among heterogeneous agents
- No global controller in the economy
- Cross-cutting hierarchies with tangled interactions
- Continual adaptation and learning by evolving agents
- Perpetual novelty
- Out-of-equilibrium dynamics with no presumption of optimality

An early example of this new approach was the articifical stock market developed by the Santa Fé Institute in 1989 [Arthur et al., 1996], which was able to recreate some of the more interesting dynamics also observed in the real world. However, this was still a far cry from a valid alternative for conventional economic modeling. In his influential 1997 *The End of Science: Facing the Limits of Knowledge in the Twilight of the Scientific Age*, science journalist John Horgan ridiculed complexity as part of a series of failed fads which he labeled "chaoplexology":

> So far chaoplexologists have created some potent metaphors: the butterfly effect, fractals, artificial life, the edge of chaos, self-organized criticality. But they have not told us anything about the world that is both concrete and truly surprising, either in a negative or in a positive sense. [Horgan, 1997, 266]

Mathematical economist Barkley Rosser Jr. surveyed the contributions of complexity economics in his 1999 *On the Complexities of Complex Economic Dynamics*. He concluded that while Horgan's views were extreme, and complexity economics had made useful contributions by undermining conventional thinking, in particular the validity of rational expectations, Horgan's skepticism concerning the ultimate value of complexity was generally justified:

> [Horgan] must be granted that it is hard to identify a concrete and surprising discovery, rather than "mere metaphor," that has arisen due to the emergence of complexity analysis. Rather, complexity theory has shifted the perspective of many economists towards thinking that what was viewed as anomalous or unusual may actually be the usual and expected. [Rosser, 1999, 187]

4.4 A Renewed Sense of Urgency

The realisation that conventional models had been unable to predict the financial crisis of 2008 despite being "forward-looking" created a new impetus for complexity-based economic research. Advocates of conventional models tried to refute the perception that conventional models had "failed", but while Lucas may have been right that DSGE modelers never claimed their models could predict crises, his argument only confirmed Doyne's critique that economic policy makers were flying blind:

> The charge is that the Fed's FRB/US forecasting model failed to predict the events of September 2008. Yet the simulations were not presented as assurance that no crisis would occur, but as a forecast of what could be expected conditional on a crisis not occurring. [Lucas Jr., 2009]

Two aspects of complexity in particular had been underestimated during the buildup to the crisis. First, the structure of the banking network, and the centrality of a small number of systemically important banks, insurers and reinsurers like AIG [McDonald and Paulson, 2015], played a major role in the accumulation of fragility. When the crisis broke out, it spread through the banking network in the form of *contagion* [Acemoğlu et al., 2016]: banks that were connected to other, impaired banks through lending relationships quickly found themselves cut off from funding. Second, although it was already understood at the time that financial shocks are not normally distributed [Ghose and Kroner, 1995], DSGE models implement shocks as normally-distributed, leading these models to underestimate the probability of a serious downturn, which in turn led policy makers to trust in diversification of risk, a policy which would have worked under normally distributed linear conditions, but was quickly scrapped after the crisis:

> Indeed, policy discourse based on assumptions underlying DSGE models had a kind of incoherence: before a crisis, the conventional wisdom called for diversification—as much as possible, e.g. through securitization and financial linkages/risk sharing. After the onset of a crisis, discourse turned to contagion. [Stiglitz, 2018, 80]

4.4.1 Network Science

Network science is a subfield of Complexity Science, which aims to understand the influence of networks on the actors they connect. By representing these actors as nodes and their mutual interactions as edges, the researcher tries to find order in what seems a miriad of equally unimportant connections. Studying the properties of the network as distinct from the characteristics of the actors it connects, network

science has proven its value in a range of academic disciplines like biology [Assenov et al., 2008], epidemiology [Barabási et al., 2011], neuroscience [Sporns et al., 2005] and engineering [Pagani and Aiello, 2013].

Within the field of economics, network science is used to understand how people, firms and countries interact, for example through trade [Chaney, 2014] or the extension of credit [Minoiu and Reyes, 2013]. Since the highly interconnected interbank credit market acted to facilitate the meltdown of financial markets during the crisis of 2008, the study of interbank credit networks has received serious attention, both in order to understand the nature of the problem [Markose et al., 2012] and as a tool for policy evaluation [Perillo and Battiston, 2018].

Temporal networks [Bramson et al., 2017], one of the most recent innovations within network science, are used to understand cascades: how information in general, and bank defaults in particular [Roukny et al., 2013], propagate through time via the interconnectivity of the network. In chapter 7 and chapter 8 we study the propagation of information and of political relationships, respectively.

4.4.2 Non-normal Distributions

In their seminal 1999 paper *Emergence of scaling in random networks*, network scientist Albert-László Barabási and physicist Réka Albert linked a particular network structure, characteristic for naturally growing networks such as the internet, to a power-law distribution [Mantegna and Stanley, 1995]. In contrast to normally distributed data, where most of the information is clustered around an average and large deviations are rare, power-law distributions characterize a diverse range of natural and man-made phenomena that do not demonstrate "normal" behavior: the populations of cities, the intensities of earthquakes, and the sizes of power outages are some examples of phenomena demonstrating "scale-free" behavior. These phenomena cannot be characterized as simply as other measurements, a sign of complex underlying processes that merit further study [Clauset et al., 2009].

One characteristic of scale-free phenomena are so-called "fat tails": a statistical property indicating that exceptional situations occur more often than normal. In the wake of the 2008 financial crisis, researchers began testing financial networks for scale-freeness. Although there is general agreement on a fat-tailedness in interbank networks [Fricke and Lux, 2015], and similar observations have been made about other sources of economic data [Fagiolo et al., 2008], debate about the nature of complex distributions is ongoing [Vandermarliere et al., 2015].

4.5 Complex Macroeconomic Agent-Based Models

Going through the literature of the past few years one cannot help but wonder if time has stood still. The same debate that erupted in 2009 following the financial

crisis between proponents and detractors of mainstream models is still ongoing, almost verbatim. No progress seems to have been made by either side. In his 2018 *Complex Agent-based Models*, Mauro Gallegati, professor of economics at the Marche Polytechnic University, pointed out that while "the dominant economic theory does not contemplate the possibility of a large crisis", "this notion is now being questioned and a new perspective is emerging" [Gallegati, 2018, vii]. In contrast, in his 2018 *Is something really wrong with macroeconomics?*, Ricardo Reis, professor of economics at the London School of Economics, laments that "it is worrying to see the practice of rigorously stating logic in precise mathematical terms as a flaw instead of a virtue." [Reis, 2018, 133].

Despite the renewed drive to develop alternatives for the mainstream approach, which is universally recognized as being flawed, a complex macroeconomic agent-based model does not yet exist. The problem seems to be that making a complex model is in itself complex. Non-linearity implies that small deviations in the underlying conditions can have large effects on the outcome (a phenomenon colloquially known as the Butterfly Effect). In terms of economic agent-based models, this means that small mistakes in how agents are modeled, and in the way they interact, can make a complex model completely wrong.

By itself, the challenge of getting an ABM exactly right due to the complex nature of the model itself is no argument for sticking with linear DSGE models, which can't model complexity altogether. But it does seem fair that the economics profession sticks with what it knows, until the alternative is shown to work better. In order to get to that point we must better understand how people make decisions, and how they interact. In the next chapter, we propose a novel methodology for studying these properties and interactions.

References

[Acemoğlu et al., 2016] Acemoğlu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2016). Networks, Shocks, and Systemic Risk. In *The Oxford Handbook of the Economics of Networks*. Oxford, Oxford University Press.

[Arthur, 1997] Arthur, W. B. (1997). *The economy as an evolving complex system II*. Addison-Wesley.

[Arthur, 2014] Arthur, W. B. (2014). Economic complexity: A different way to look at the economy. https://medium.com/sfi-30-foundations-frontiers/economic-complexity-a-different-way-to-look-at-the-economy-eae5fa2341cd.

[Arthur et al., 1996] Arthur, W. B., Holland, J. H., LeBaron, B., Palmer, R., and Tayler, P. (1996). Asset pricing under endogenous expectations in an artificial stock market. *The economy as an evolving complex system II*, 27.

[Ascari et al., 2015] Ascari, G., Fagiolo, G., and Roventini, A. (2015). Fat-tail distributions and business-cycle models. *Macroeconomic Dynamics*, 19(2):465–476.

[Assenov et al., 2008] Assenov, Y., Ramírez, F., Schelhorn, S.-E., Lengauer, T., and Albrecht, M. (2008). Computing topological parameters of biological networks. *Bioinformatics*, 24(2):282–284.

[Barabási, 2007] Barabási, A.-L. (2007). The architecture of complexity. *IEEE Control Systems Magazine*, 27(4):33–42.

[Barabási et al., 2011] Barabási, A.-L., Gulbahce, N., and Loscalzo, J. (2011). Network medicine: a network-based approach to human disease. *Nature reviews genetics*, 12(1):56–68.

[Batiz-Zuk et al., 2015] Batiz-Zuk, E., Christodoulakis, G., and Poon, S.-H. (2015). Credit contagion in the presence of non-normal shocks. *International Review of Financial Analysis*, 37:129–139.

[Bramson et al., 2017] Bramson, A., Hoefman, K., van den Heuvel, M., Vandermarliere, B., and Schoors, K. (2017). Measuring propagation with temporal webs. In *Temporal network epidemiology*, pages 57–104. Springer.

[Chaney, 2014] Chaney, T. (2014). The network structure of international trade. *American Economic Review*, 104(11):3600–3634.

[Clauset et al., 2009] Clauset, A., Shalizi, C. R., and Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review*, 51(4):661–703.

[Dawes and Corrigan, 1974] Dawes, R. M. and Corrigan, B. (1974). Linear models in decision making. *Psychological bulletin*, 81(2):95.

[De Bock et al., 2014] De Bock, D., Van Reeth, D., Minne, J., and Van Dooren, W. (2014). Students' overreliance on linearity in economic thinking: An exploratory study at the tertiary level. *International Review of Economics Education*, 16:111–121.

[Fagiolo et al., 2008] Fagiolo, G., Napoletano, M., and Roventini, A. (2008). Are output growth-rate distributions fat-tailed? some evidence from oecd countries. *Journal of Applied Econometrics*, 23(5):639–669.

[FRED, 2020] FRED (2020). Commercial Banks in the U.S. Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org/series/USNUM.

[Fricke and Lux, 2015] Fricke, D. and Lux, T. (2015). On the distribution of links in the interbank network: evidence from the e-mid overnight money market. *Empirical Economics*, 49(4):1463–1495.

[Gallegati, 2018] Gallegati, M. (2018). *Complex Agent-Based Models*. Springer.

[Ghose and Kroner, 1995] Ghose, D. and Kroner, K. F. (1995). The relationship between garch and symmetric stable processes: Finding the source of fat tails in financial data. *Journal of Empirical Finance*, 2(3):225–251.

[Gigerenzer and Todd, 1999] Gigerenzer, G. and Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford University Press, USA.

[Hammond and Summers, 1965] Hammond, K. R. and Summers, D. A. (1965). Cognitive dependence on linear and nonlinear cues. *Psychological Review*, 72(3):215.

[Holt et al., 2011] Holt, R. P., Rosser Jr, J. B., and Colander, D. (2011). The complexity era in economics. *Review of Political Economy*, 23(3):357–369.

[Horgan, 1997] Horgan, J. (1997). *The End of Science: Facing the Limits of Knowledge in the Twilight of the Scientific Age*. Paperback ed. New York: Broadway Books.

[Lucas Jr., 2009] Lucas Jr., R. (2009). In defence of the dismal science. https://www.economist.com/finance-and-economics/2009/08/06/in-defence-of-the-dismal-science.

[Mantegna and Stanley, 1995] Mantegna, R. N. and Stanley, H. E. (1995). Scaling behaviour in the dynamics of an economic index. *Nature*, 376(6535):46.

[Markose et al., 2012] Markose, S., Giansante, S., and Shaghaghi, A. R. (2012). 'too interconnected to fail' financial network of us cds market: Topological fragility and systemic risk. *Journal of Economic Behavior & Organization*, 83(3):627–646.

[McDonald and Paulson, 2015] McDonald, R. and Paulson, A. (2015). Aig in hindsight. *Journal of Economic Perspectives*, 29(2):81–106.

[Minoiu and Reyes, 2013] Minoiu, C. and Reyes, J. A. (2013). A network analysis of global banking: 1978–2010. *Journal of Financial Stability*, 9(2):168–184.

[Pagani and Aiello, 2013] Pagani, G. A. and Aiello, M. (2013). The power grid as a complex network: a survey. *Physica A: Statistical Mechanics and its Applications*, 392(11):2688–2700.

[Perillo and Battiston, 2018] Perillo, C. and Battiston, S. (2018). A multiplex financial network approach to policy evaluation: the case of euro area quantitative easing. *Applied network science*, 3(1):49.

[Reis, 2018] Reis, R. (2018). Is something really wrong with macroeconomics? *Oxford Review of Economic Policy*, 34(1-2):132–155.

[Richardson et al., 2000] Richardson, K. A., Cilliers, P., and Lissack, M. (2000). Complexity science: A 'grey'science for the 'stuff in between'. In *Proceedings of the First International Conference on Systems Thinking in Management*, pages 532–7.

[Rosser, 1999] Rosser, J. B. (1999). On the complexities of complex economic dynamics. *Journal of economic Perspectives*, 13(4):169–192.

[Roukny et al., 2013] Roukny, T., Bersini, H., Pirotte, H., Caldarelli, G., and Battiston, S. (2013). Default cascades in complex networks: Topology and systemic risk. *Scientific reports*, 3:2759.

[Shannon, 1948] Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423.

[Simon, 1962] Simon, H. A. (1991(1962)). The architecture of complexity. In *Facets of systems science*, pages 457–476. Springer.

[Sporns et al., 2005] Sporns, O., Tononi, G., and Kötter, R. (2005). The human connectome: a structural description of the human brain. *PLoS computational biology*, 1(4).

[Stiglitz, 2018] Stiglitz, J. E. (2018). Where modern macroeconomics went wrong. *Oxford Review of Economic Policy*, 34(1-2):70–106.

[Van Dooren et al., 2008] Van Dooren, W., De Bock, D., Janssens, D., and Verschaffel, L. (2008). The linear imperative: An inventory and conceptual analysis of students' overuse of linearity. *Journal for Research in Mathematics Education*, pages 311–342.

[Vandermarliere et al., 2015] Vandermarliere, B., Karas, A., Ryckebusch, J., and Schoors, K. (2015). Beyond the power law: Uncovering stylized facts in interbank networks. *Physica A: Statistical Mechanics and its Applications*, 428:443–457.

Chapter 5

Live Agent-Based Models

Live Agent-Based Modeling (LAB-M) is a novel methodology for studying the behavior and interactions of people in a world of bounded rationality and imperfect self-awareness. LAB-M combines the benefits of natural field experiments with the extensive data generation of agent-based models. This combination allows us to study participants in a natural environment, and because the data is collected rather than self-reported or observed, we can distinguish between how participants actually behave and how they think they behave. In chapter 9, for example, we find that while most people believe price-quality should be a linear relationship, in reality they pay exponential increases in price for linear increases in quality.

5.1 Field Experiments in the Economic Sciences

Experimental research, one of the key methods of scientific discovery, is notoriously difficult to conduct in the field of economics. Paul Samuelson and William Nordhaus describe the problem in the 12th edition of *Economics* [Samuelson and Nordhaus, 1985] as follows:

> The economic world is extremely complicated. There are millions of people and firms, thousands of prices and industries. One possible way of figuring out economic laws in such a setting is by controlled experiments. A controlled experiment takes place when everything else but the item under investigation is held constant. Thus a scientist trying to determine whether saccharine causes cancer in rats will hold "other things equal" and only vary the amount of saccharine. Same air, same light, same type of rat.
>
> Economists have no such luxury when testing economic laws. They cannot perform the controlled experiments of chemists or biologists

> because they cannot easily control other important factors. Like astronomers or meteorologists, they generally must be content largely to observe.

Experimental methods in economics, pioneered by Vernon Smith who in 2002 shared the Nobel Prize in economics for "having established laboratory experiments as a tool in empirical economic analysis, especially in the study of alternative market mechanisms" [The Royal Swedish Academy of Sciences, 2002], have gained more of a foothold in the last two decades. In his 2011 *Why Economists Should Conduct Field Experiments and 14 Tips for Pulling One Off*, experimental economist John A. List addresses the main question of experimental research: whether the results can be generalized to non-laboratory settings, due to either the setting of the experiment (lab vs. field experiment) or the awareness of the participant that they are part of a study (framed vs. natural experiment). List considers four types of experiments in the spectrum of field experimentation in economics [List, 2011, 4-7]:

- *lab experiments*: Experiments held in a small, controlled setting. Participants are usually students who are aware of the nature of the experiment.
- *"artefactual" field experiments*: These mimic lab experiments, except that they use "nonstandard"subjects: participants who are not students, but drawn from the market of interest.
- *framed field experiments*: To avoid the possibility that the lab setting influences the results, the experiment is conducted in the natural environment of the subject rather than in a laboratory setting.
- *natural field experiments*: These occur in the environment where the subjects are naturally undertaking certain tasks, and the subjects *do not know* that they are participants in an experiment.

Natural field experiments combine the most attractive elements of the experimental method and naturally occurring data: randomization and realism, allowing the researcher to avoid biases common to the other experimental approaches, such as self-selection:

> generally the subjects who choose to participate in the experiment are those who expect to gain the most (perhaps because they believe they are likely to get good results from the treatment). As a result, the estimated causal effect from these other experimental types, while valid, might not generalize to the target population of interest — which of course includes the subpopulation (often the majority) that did not volunteer for the experiment when offered the opportunity. [List, 2011, 6]

5.2 Empirical Validation of Agent-Based Models

In his 2007 *A Critical Guide to Empirical Validation of Agent-Based Models in Economics*, Giorgio Fagiolo addresses the difficulty of empirically validating economic agent-based models against real world data: examining the extent to which the output traces generated by a particular model approximates reality, typically described by one or more 'stylized facts' drawn from empirical research. Fagiolo treats both the real world and the model as *data generating processes*:

> [Real world] features are usually described in terms of causal relations and it is usually assumed that some causal mechanism (deterministic or stochastic) has generated the data. We call this causal mechanism the 'real-world data generating process' (rwDGP). A model approximates portions of the rwDGP by means of a 'model data generating process' (mDGP). ... The extent to which the mDGP is a good representation of the rwDGP is evaluated by comparing the simulated outputs of the mDGP with the real-world observations of the rwDGP. [Fagiolo et al., 2007, 200]

Empirical validation of a model, then, is the process of comparing the results of the *mDGP* to the *rwDGP*. A generalized procedure is illustrated in Figure 5.1.

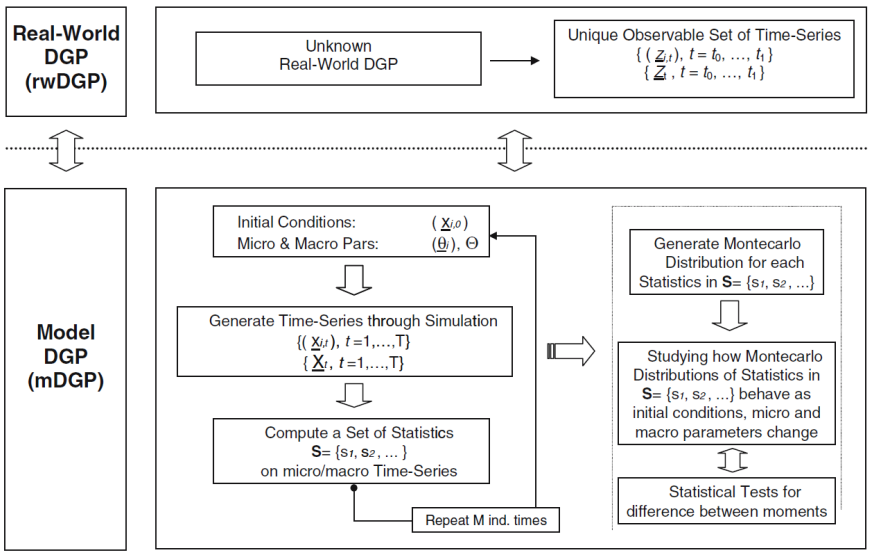


Figure 5.1: A procedure for studying the output of an agent-based model, by comparing the model data generating process (mDGP) to real world data (rwDGP) [Fagiolo et al., 2007, 203]

Fagiolo identifies three approaches to empirical calibration, of which two (the Indirect Calibration approach and the Werker-Brenner approach) are suitable to macro economic modeling, and considers a number of methodological and operational issues associated with empirical calibration [Fagiolo et al., 2007, 206–213]:

1. Assessing fitness amongst a class of models does not automatically help identify a true (but unknown) underlying model.
2. Empirically calibrating models encourages the modeler to focus on variables and parameters that are readily calibrated, and for which data already exists.
3. Effective calibration requires a wealth of high quality data, whereas real-world economic data tends to be biased and incomplete because of (1) costs of organizing and collecting raw data, (2) bias in the collection process and (3) issues related to the nature of the phenomenon being observed, for example rarely occurring phenomena.
4. Correctly identifying the nature of the relationship between the model *mDGP* and the real-world *rwDGP*: is the rwDGP ergodic? If so, what are the initial values of the relevant micro and macro variables?
5. If the *mDGP* and *rwDGP* are sufficiently stationary, can we correctly synchronize them by correctly setting t_0?
6. Are the observed micro and macro parameters time dependent? This influences the structural relationship of the model.
7. Do the predictions of the model take into account data that lies outside the current regime, in the way that real economic agents' expectations project current data into the future? (the Lucas Critique)

In his conclusion, Fagiolo summarizes the "core problem" of real-world economic datasets:

> This points us to a final core problem; the availability, quality and bias of available data sets. Empirically-based modeling depends on high quality data sets. Unfortunately, the data sets that exist are invariably pre-selected. Not all potential records are retained; some are fortuitously bequeathed by the past but others are not captured. The data sets that do exist are invariably biased. ... As econometricians know only too well, it may simply be the case that data that would have assisted in a particular discussion has simply not been collected. [Fagiolo et al., 2007, 223]

In practice, the large gap between model *mDGP* and real-world *rwDGP*, due to the conditions described above, so far appears difficult to reconcile.

5.3 Virtual World Data Generating Process

Advances in computer technology during the eighties and early nineties enabled the development of agent-based modeling. During this period, researchers built the artificial worlds of evolutionary economics [Kwaśnicki, 1999]. These simulations were considered promising, because they could reproduce real-world behavior that traditional models could not. But as we saw earlier, simulation and reality remain difficult to reconcile.

In the last two decades, further progress in computational power has enabled the creation of virtual worlds populated by large communities of *real* people. Contrary to simulation, the phenomena observed in these virtual worlds are causally driven by human behavior. And because virtual worlds exist on a computer, conditions within the world are precisely quantified. Player actions and responses can be registered, generating a wealth of data far beyond what is possible in a laboratory experiment.

The resulting virtual world Data Generating Process, or *vwDGP*, constitutes a bridge between the agent-based model *mDGP* and the real-world *rwDGP*, opening up new possibilities for scientific study of economic behavior and interaction.

5.3.1 Virtual Worlds

Virtual worlds are computer-based environments offering unique opportunities for testing economic and social theories. Virtual worlds have a number of attractive properties for research:

- A computer keeps track of the state of everything that populates the world, resulting in very detailed datasets
- Properties of everything existing in the world are precisely defined
- Data is available about the entire population of participants
- All participants can be given access to all relevant details, meaning participants can be studied under conditions of (almost) perfect information
- Participants tend to have more homogeneous preferences, because virtual worlds are invariably simpler than the real world

Virtual worlds offer rich environments, where the possible interactions between agents are programmed into the environment. For example, if the makers of a virtual world want their players to be able to communicate with each other inside the environment, they must explicitly enable this by writing the program code that handles the commmunication. If the intent is to allow players to trade virtual items between one another, the makers of the game could design and implement a

marketplace. However, because the agents inside the world are real, living players, the underlying "agent behaviour" is real. As a result, a virtual world acts as a "Live Agent-Based Model": capturing perfectly the emergent dynamic behaviour of groups of people based on a combination of real underlying dynamics and motivations, integrated by the deliberately designed interactions enabled by the makers of the virtual world.

Because these activities are connected in a complex interplay of dependencies, players inside the game are motivated in similar ways as those real-world actors whose behavior we wish to study. Players invest time and real-world money to acquire virtual resources in the game world. These resources are then used to enhance the players' capabilities, influence and/or status. When players lose these resources in conflict or through mistake, the time, effort and money they used to acquire them are permanently lost. This feature creates genuine scarcity and risk-aversion that fuels realistic economic behaviors [BBCNews, 2014, Goh, 2018, Hoefman et al., 2019].

The constraints imposed by a virtual world are also a blessing. Although the scale of these games can be quite large, with the most successful of these sporting millions of participants (in fact larger than some countries), managing and supervising these worlds is still possible, because all world behavior is specified via program code which can be altered and reviewed. These worlds mimic complex behavior, for example by allowing alliances of players to tax their members, while the system is still invariably simpler than the encyclopedic tax codes of the real world. Whereas many of these worlds organize factions of participants with political relations, these standings between factions tend to be captured by a single number, rather than a complicated mixture of treaties, international laws, and historical conventions. All the necessary ingredients for complex socio-politico-economic dynamics exist, but they are all simplified. A virtual world is like a computer simulation, except with humans instead of algorithms as the driving force.

5.3.1.1 Massively Multiplayer Online Games

Within the classification of virtual worlds exists the category of so-called Massively Multiplayer Online Games (MMOGs). Whether a virtual world is a computer game can be difficult to determine. Second Life [Anstadt et al., 2013], a classic example of a virtual world that does not advertise itself as a computer game, allows its participants to build creations of their own, including games which others can play. In contrast, Eve Online (see below) is advertised as a game, but the virtual world does not force its players into any specific game-behavior. These so-called Sandbox Games [Squire, 2008], which offer players a world of possibilities without coercing them into scripted behavior, offer the best opportunities for research. Since players aren't pushed into pre-determined behavior (such as defeating the bad guy), observed behavior is emergent. But because of the fun

element typical of games, players are drawn to participate in this world, which can be studied without participants realizing they are part of a study.

5.3.2 Desired Properties

Although many computer games are too simple to provide support for such emergent behavior, some large-scale virtual worlds embody dynamics that are representative of the behavior we would like to study in the real world. The challenge is to find a virtual world that is:

- Large enough to support complex interaction between participants
- Contains dynamics (for example market behavior) of the type the researcher wants to study
- Behavior is not the result of a pre-selection due to participant demographics, or built in incentives.

If we find such a world, we have an almost perfect environment for testing economic and social theories.

5.3.3 Eve Online

Eve Online (EVE) is an open-ended, Massively Multiplayer Online Game set in a science fiction universe created by Icelandic company CCP Games in 2003. More than 500,000 players compete for resources and territory while engaging in a variety of professions and activities including mining, manufacturing, trading, piracy, exploration, and combat, both versus the environment and against other players. The game provides its players with a virtual world and the tools to explore, and players have the freedom to choose what, when, and how to approach the available content, including the purchase and sale of goods. Players are free to join corporations, which can form into alliances, which are in turn part of implicit coalitions in a multi-tiered hierarchy of political arrangements. Alliance membership determines friends and foes, which regulates conflicts on a scale ranging from single-player economic sabotage to prolonged territorial war involving thousands of players.

EVE's game environment is broken into 7930 solar systems of which 3524 are conquerable by the players. These conquerable systems are where the alliances can hold sovereignty over a system and ownership of its stations. Ownership allows players to control access to stations, collect taxes on transactions occurring inside the system, and extract the resources available in the system. The game contains 12,709 distinct items that can be bought and sold between players including ships, ship modules, minerals, ammunition, blueprints, and many more. Available items

and their properties (which players can view at any time at zero cost) are decided by CCP and only rarely adjusted. Market prices, on the other hand, are endogenously determined by the market behavior of players via a double auction system that matches buy orders with sell orders. Players can buy and sell anywhere in the virtual universe but, for reasons of efficiency, market activity tends to cluster in hubs. Two thirds of all market transactions are conducted in a single central trading hub. Transaction data is collected on an individual basis, generating an extremely detailed economic dataset collected throughout the period of the game's existence, illustrated in Figure 5.2.

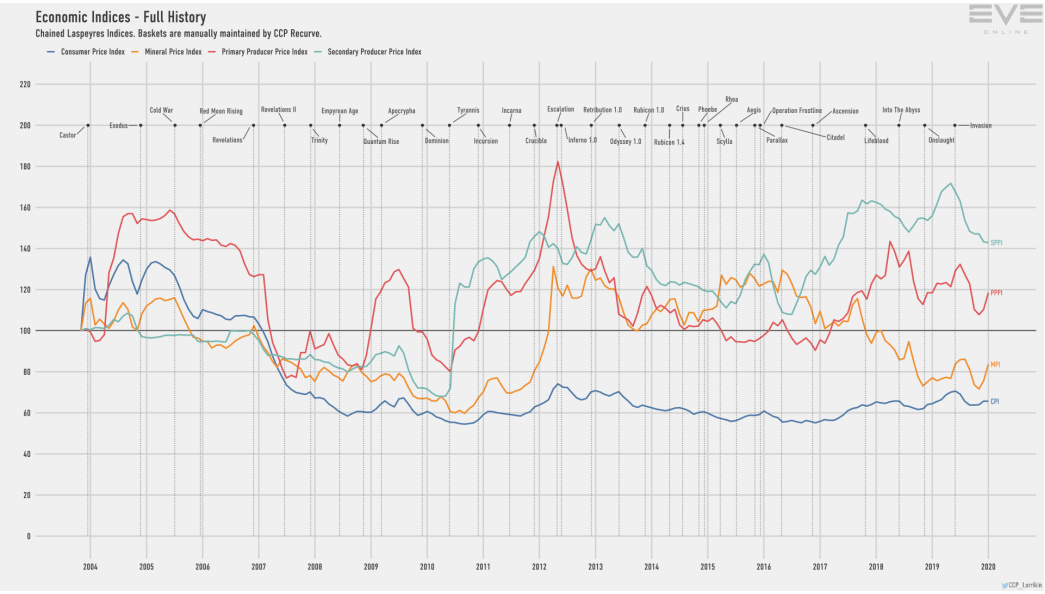


Figure 5.2: CCP Games has been collecting detailed economic time series data within Eve Online since 2003 [CCP Games, 2020]

EVE's economic realism is driven by a strict destruction policy: when a player loses a ship to the activity of in-game agents (so-called Non-Player Characters of NPCs) or because of the hostile actions of other players, the ship is permanently destroyed. Replacing the ship requires building it anew, either by mining the resources and producing the ship directly, or (more often) by trading with a player who specializes in these activities. This cycle of creation and destruction means that players on average are constantly replacing their economic goods, which not only drives market behavior, but also leads to derived activities like transportation and protection.

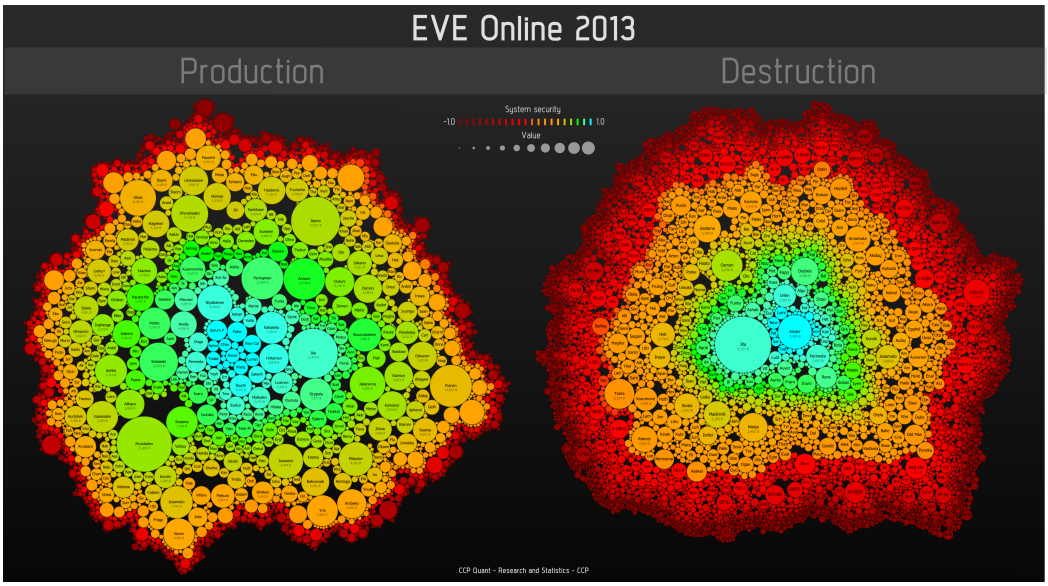


Figure 5.3: Production and destruction per solar system, colored by security status of the system for the year 2013. Production is predominantly situated in high security systems, while destruction happens mainly in low security areas. This dynamic generates secondary economic activities, such as transportation. [CCP Games, 2014]

5.3.4 Other Candidates

Other virtual worlds that have been the subject of academic research are Everquest [Castronova, 2004], Second Life [Kaplan and Haenlein, 2009], and Pardus [Thurner et al., 2012].

5.4 Applications

5.4.1 Empirical Validation of Agent-Based Models

The exercise of empirical validation of agent-based models, by comparing the model *mDGP* to the real-world *rwDGP*, is marred by a number of difficulties that remain unsolved. It would be interesting to investigate whether a virtual world *vwDGP* could bridge the distance between the two data generating processes: as the model *mDGP* is easier to map onto the virtual world *vwDGP* because of the increased availability of reliable data of the *vwDGP*, and the real-world *rwDGP* is easier to relate to the virtual world *vwDGP* because both are causally driven by the behavior of real people.

5.4.2 Model Ergodicity and LAB-M Bootstrapping

Non-ergodic systems are a particularly difficult class of processes to model. Simply put, ergodicity describes the tendency of a system to return to a stable state. Non-ergodic systems have no such tendency, meaning (stochastic) changes during the process permanently influence the future behavior of the system. The macroeconomy is understood to be a non-ergodic system, which makes it particularly difficult to model. In particular, correctly identifying the right starting conditions is crucial to the outcome of a non-ergodic model. A possible solution worth investigating would be to bootstrap macroeconomic agent-based models using data from a virtual world *vwDGP*.

5.4.3 Directly tapping the *vwDGP*

Virtual world *vwDGPs* contain data about human behavior that is vastly more fine-grained and complete than available real-world data. This opens up new avenues for testing economic, social, and political theories that have so far been difficult to test in real-world conditions, as well as potentially developing entirely new theories, when the new data reveals previously unobserved dynamics. Macroeconomic agent-based models in particular depend on an accurate understanding of human decision-making, as well as a correct model of the complex interaction between people. The data generated by a virtual world *vwDGP* offer novel ways of gaining insight into these dynamics.

In the chapters that follow, we put to the test the exploration of a virtual world *vwDGP* to study complex interaction and behavior. In chapter 8, we use Eve Online alliance data to study the dynamics of political relations in the context of Structural Balance Theory. In chapter 7, we compare virtual world Eve Online data to real-world Twitter and Russian banking data, to study the dynamics of network propagation using Temporal Networks. And in chapter 9, we use Eve Online pricing data to study market behavior using Hedonic Pricing Theory.

References

[Anstadt et al., 2013] Anstadt, S. P., Bradley, S., Burnette, A., and Medley, L. L. (2013). Virtual worlds: Relationship between real life and experience in second life. *The International Review of Research in Open and Distributed Learning*, 14(4).

[BBCNews, 2014] BBCNews (2014). Eve online virtual war 'costs $300,000' in damage.

[Castronova, 2004] Castronova, E. (2004). The price of bodies: A hedonic pricing model of avatar attributes in a synthetic world. *Kyklos*, 57(2):173–196.

[CCP Games, 2014] CCP Games (2014). Insights into 2013 production and destruction. https://www.eveonline.com/article/insights-into-2013-production-and-destruction.

[CCP Games, 2020] CCP Games (2020). Monthly Economic Report, January 2020. https://www.eveonline.com/article/q624rx/monthly-economic-report-january-2020.

[Fagiolo et al., 2007] Fagiolo, G., Moneta, A., and Windrum, P. (2007). A critical guide to empirical validation of agent-based models in economics: Methodologies, procedures, and open problems. *Computational Economics*, 30(3):195–226.

[Goh, 2018] Goh, A. (2018). How to build a robust game economy.

[Hoefman et al., 2019] Hoefman, K., Bramson, A., Schoors, K., and Ryckebusch, J. (2019). The impact of functional and social value on the price of goods. *PLoS One*, 13(11):e0207075.

[Kaplan and Haenlein, 2009] Kaplan, A. M. and Haenlein, M. (2009). The fairyland of second life: Virtual social worlds and how to use them. *Business horizons*, 52(6):563–572.

[Kwaśnicki, 1999] Kwaśnicki, W. (1999). Evolutionary economics and simulation. In *Computational techniques for modelling learning in economics*, pages 3–44. Springer.

[List, 2011] List, J. A. (2011). Why economists should conduct field experiments and 14 tips for pulling one off. *Journal of Economic perspectives*, 25(3):3–16.

[Samuelson and Nordhaus, 1985] Samuelson, P. A. and Nordhaus, W. D. (1985). *Economics*. 12th ed. New York: McGraw-Hill.

[Squire, 2008] Squire, K. (2008). Open-ended video games: A model for developing learning for the interactive age. *The ecology of games: Connecting youth, games, and learning*, pages 167–198.

[The Royal Swedish Academy of Sciences, 2002] The Royal Swedish Academy of Sciences (2002). Alfred nobel memorial prize in economic sciences. https://www.nobelprize.org/prizes/economic-sciences/2002/smith/facts/.

[Thurner et al., 2012] Thurner, S., Szell, M., and Sinatra, R. (2012). Emergence of good conduct, scaling and zipf laws in human behavioral sequences in an online world. *PLOS ONE*, 7(1).

Part II

Methodology

Chapter 6

Creating a Temporal Network Propagation Engine using C++

6.1 Introduction

In chapter 7, we study the dynamics of information propagation through networks via a technique called temporal webs, using three datasets: alliance relationships in Eve Online, interbank lending in Russia, and emotional affect in the Twitter network. Of these, the Eve Online network was by far the largest; in fact so large that we were forced to write our own analysis software. We then expanded on this framework for chapter 8, by allowing for Structural Balance Theory with weighted edges, to study coalition changes within the Eve Online network of alliances through time.

A network follows simple rules, but these do lead to a few counter-intuitive challenges that need to be accounted for. A normal network consists of N *nodes*, connected via *edges* to potentially N-1 other nodes. Edges can be *directed* or *undirected*. Undirected edges are symmetric, meaning that if person A is friends with person B, then person B is also friends with person A, and the information about the relationship needs to be stored only once. In a network with directed edges, the relationship can be asymmetric, so the states of the relationship between A and B, and vice versa, need to be stored separately.

A network of N nodes has a potential of N * (N-1) directed edges. As a result, the edge size of a network increases quadratically with the number of nodes. This is an important property for the implementation of a temporal network (or any network for that matter): it is mainly the edges connecting the nodes that present the bottleneck in terms of computer memory, which in turn poses an upper limit on network size.

6.2 The Eve Online Temporal Triad Network

In our analysis of the Eve Online alliance network, we are interested in the (in)stability of triads of alliances. From an underlying alliance network of N nodes, we generate a triad network consisting of N_T unique *triad nodes*, to analyse how changes in the stability within triplets of alliances propagate through the triad network as underlying alliance relationships change through time:

$$N_T = \frac{N!}{3! * (N-3)!} = \frac{N * (N-1) * (N-2)}{6}$$

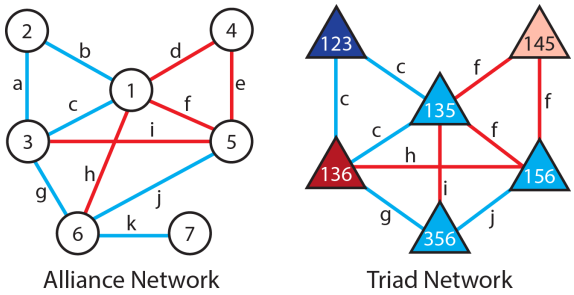


Figure 6.1: Left: an alliance network. Right: the derived triad network. Combinations of alliances (left) form the triad nodes (right), and the edges between alliances (left) form the edges that connect the triads (right). Blue lines indicate friendly, red lines are hostile relationships.

A triad uniquely combines three alliance pairs. Changes in triad states are driven by changes in the states of the underlying pairs of alliances (for example, two allies turning enemies). As a result, triad changes (only) propagate between triads that share an alliance pair. For example, the triads formed by alliances {1, 2, 3} and {1, 2, 4} share the {1, 2} alliance pair. If the relationship between alliances 1 and 2 changes, both triads are affected. In network terms, this means edges in the triad network are replications of the edges (= pairs of alliances) in the underlying alliance network. This is illustrated in Figure 6.1, where the c edge connecting alliances 1 and 3 in the alliance network is replicated 3 times to form the c edges between triads {1, 2, 3}, {1, 3, 5} and {1, 3, 6} in the triad network.

To create a triad network generated from an alliance network of N nodes, each edge in the underlying alliance network is replicated (N-2) times, connecting (N-2) triads in the triad network. As a result, each triad node is connected to (N-3) other triads for each alliance pair, for a total of 3 * (N-3) other triads. In a triad network generated from an alliance network of N nodes, the number of edges M_T is:

$$M_T = 3 * (N-3) * N_T$$

So far, we are still lacking a time element. The triad network at this point could be used to investigate triad properties at one point in time, but to analyse propagation *through* time, we would need to replicate the triad network T times, once for each day of observation in our dataset, and connect the triad nodes between time steps. The resulting structure, called a *temporal network*, is connected through both network space and time. Applied to our triad network approach, the resulting *temporal triad network* consists of N_{TT} temporal triad nodes:

$$N_{TT} = N_T * T$$

For each time step within the temporal network there is a triad network, with nodes interconnected as described earlier. In addition, at each time step except the last, each triad node is also connected with the equivalent of itself at the next time step. This gives us a total number of temporal triad edges M_{TT} in our temporal triad network:

$$M_{TT} = M_T * T + N_T * (T - 1)$$

Our Eve Online alliance dataset covers 607 unique alliances over a period of 437 days. As a result, each time slice of the temporal triad network contains 37 million triad nodes, and fully constructed, the Eve Online temporal triad network contains 16.2 billion nodes and 29.4 trillion edges.

6.3 Network Representation

To save a network to file, and/or to construct a network in computer memory for analysis, the network information needs to be structured in an organised format. There are two broad techniques for representing a network: via adjacency matrix, or via a connection list.

6.3.1 Adjacency Matrix

An adjacency matrix is a two-dimensional, NxN matrix, where the nodes of the network are both the rows and the column entries. The cells inside the matrix contain information about each matching {row, column} pair, these being the edges of the network. For example, for a simple network where nodes are either connected or not connected, cell values of 0 could stand for unconnected, with values of 1 indicating the connected {row, column} pairs.

Adjacency matrices are an elegant and flexible way of representing network data. In addition, they allow network operations to be calculated in mathematical form using matrix algebra. Adjacency matrices are often used, however, because

adjacency matrices are two-dimensional constructs, their size increases quadratically with the number of nodes. For large and/or sparse networks, storing a value for every possible edge can be wasteful.

To represent the Eve Online alliance temporal triad network, which contains 16.2 billion nodes, an adjacency matrix is clearly not the right solution. With 37 million temporal nodes per time step, a single time slice of the temporal network would have a corresponding adjacency matrix of 1.369 quadrillion cells.

6.3.2 Connection List

A connection list is an enumeration of all nodes in the network with at least 1 connection, followed by the nodes to which it is connected. For example, the connection list of a network with node 1 connected to nodes 3, 5, and 12 would contain an entry reading "1 3 5 12".

Connection lists are more resource-efficient than adjacency matrices, as only information about existing edges is stored. The disadvantage is that storing more detailed information than simply whether nodes are connected is more complex with a connection list.

6.3.3 Networks and Object-oriented Programming

The conceptual structure of networks - organised in nodes and edges - naturally lends itself to the object-oriented programming paradigm, in which data is organised in modular entities called objects, which refer to other objects via memory addresses (implemented through *pointers* or *references*, depending on which programming language).

By treating the nodes of a network as the objects of an object-oriented program, and storing inside each object a connection list of the memory addresses of the other objects to which the node connects, we can generate networks in computer memory in a way that is straightforward to analyse, while at the same time, being vastly more efficient resource-wise than by using adjacency matrices.

6.4 Technical Considerations

To store any kind of data, a computer uses RAM memory. Network data is simply the collection of all relevant information about the nodes that form the network, and the edges between them. How large a network can be analysed by a computer is determined by the amount of available computer memory[1].

[1] Some big data techniques exploit other computer resources, like the hard disk, as extra storage, but this comes with a considerable hit to execution time.

6.4.1 Computer Memory

From an application software perspective, computer memory is simply a long list of bytes. One byte can store a number between 0 and 255; bytes are clustered to store larger numbers, or more complex information like floating point values. Computer memory comes in a magnitude of tens of billions of bytes. A computer with 64 gigabytes (GB) of memory, for example, has 64 billion bytes of RAM for storing information. The more RAM available on a computer, the larger the networks it can analyse.

6.4.2 32 vs 64 bit processing

Reading from and writing to computer memory happens via memory addressing. Objects refer to other objects via memory addresses. Given the prevalence of connections in network analysis, memory addresses are an important implementation consideration.

A memory address is a 32 bit or a 64 bit number, containing the position of another variable in memory. The choice of 32 or 64 bit is important, because it determines whether referencing to another variable takes 4 bytes (32 bit) or 8 bytes (64 bit) of memory.

There is a trade-off: a 32 bit number has an upper limit of around 4.3 billion, meaning a 32 bit program can only use up to 4 gigabytes of RAM (in practice it is less, around 3.5 GB). In contrast, the upper limit of a 64 bit number is 2^{64}, meaning a 64-bit memory address can directly access approximately 16 exabytes of memory, well beyond the maximum range of computers today. However, the memory address itself takes up more space: 8 bytes instead of 4.

6.4.3 Programming Languages

Two broad categories of programming languages exist: *script languages*, which hide some of the complexity of computer programming for the benefit of ease of use, and *compiled languages*, which give the programmer more control over efficient use of computer resources, at the cost of a higher learning curve and longer development time.

Both language families are equally important. For smaller networks, script languages such as Python are the better alternative: all other things equal, in programming shorter is always better than longer; readable is always better than cryptic; publishable in a wide circle of researchers is always better than restricted to a specialised audience. But when computer resources are a limiting factor, a compiled language can determine whether an analysis can be run on a computer at all.

For our specific situation, C++ offers itself as the best candidate, as it is a compiled, object-oriented language known for speed and efficient use of resources.

6.5 Program Design

The full Eve Online alliance temporal triad network contains 16.2 billion nodes and 29.4 trillion edges. At 8 bytes of memory per edge, and ignoring node size, merely connecting the network in computer memory would require 235 terabytes of computer memory. This is well beyond the 64 gigabytes of RAM available for desktop computers at this time of writing.

6.5.1 First Optimization: State Machine design pattern

If we model propagation through the temporal network as an iteration of propagations between two time steps (t → t+1 ; t+1 → t+2 ; ...) then, instead of making an object for every triad at every time step, we only need to implement two time slices of the temporal network. Taking this logic one step further, we need to implement only a single time slice, with node objects capable of storing two node states at consecutive time steps. One way to accomplish this is by using a State Machine design pattern[1].

6.5.1.1 Alliance Data and State objects

For each of the 607 alliances in our dataset, we implement two objects: a Data object, which stores the alliance data for the entire dataset, and a State object, which manages (only) the previous and current states of an alliance at a given time step:

- Whether the alliance existed
- The membership count of the alliance
- The number of solar systems the alliance holds sovereignty over

In addition, the alliance State object has functionality for determining its relationship with other alliances at a certain time, as well as the distance between its own territory and that of other alliances, by retrieving this information from the Data object.

6.5.1.2 Optimization Results

By reducing T time slices of temporal triad nodes to a single time slice of temporal nodes storing two triad states at once, the size of the temporal triad network is reduced to 37 million nodes, albeit storing double data, and 67 billion edges connecting the nodes. The edges remain the biggest contributor to memory use; connecting this network at 8 bytes per edge requires 536 gigabytes of RAM.

[1] Design patterns are tried and tested solutions for regularly occurring software design problems.

6.5.2 Second Optimization: Triad Network Topology

To generate the triad network from a network of N alliances, the edges of the underlying alliance network are replicated N-2 times. If we can correctly link up all the triad nodes while avoiding this edge replication, we can make another significant optimization.

We do this by implementing the edges of the underlying alliance network as (unique) objects. The triad nodes then connect to these Edge objects via their memory addresses, and vice versa. The resulting {Triad ↔ Edge ↔ Triad} object model preserves the conceptual structure of the triad network, while avoiding costly edge replication.

6.5.2.1 Edge State objects

We already have unique Alliance state objects, one for every alliance in the dataset. By creating an Edge state object for each pair of Alliance state objects, and having the Edge objects refer to the Alliance objects via their memory addresses, we obtain a list of unique Edge state objects, whose states are determined by the states of their underlying alliance pairs.

Each Edge object manages the previous and current states of the edge they represent at a given time step:

- Directed and undirected edge relationships (friendly, neutral, hostile)
- Geographical distance between the two alliances

The Edge objects also store a list of the memory addresses of all Triad objects they interconnect. This way, these memory addresses need to be stored only once.

6.5.2.2 Triad State objects

Triads are triplets of alliances. Their states are uniquely defined by the states of the triplets of Edges, connecting between these alliances. This means we can create Triad state objects as a triplet of memory addresses of Edge state objects, one for each unique combination of three Edge state objects.

Each Triad object manages the previous and current state of the Triad they represent at a given time step:

- The state of the Triad according to Structural Balance Theory (weakly balanced, strongly balanced, ...)
- Weights by membership count, sovereignty count, distance calculated by average location and closest distance, and the average between these weight metrics. These are used in the triad analysis in chapter 8.

6.5.2.3 Optimization Results

In this design, each Triad only stores three memory address. The new design still has 37 million Triad node objects, but the amount of edges in the triad network now equals the amount of edges in the underlying alliance network. Note that in this design, memory addresses no longer correspond with the edges of the triad network, instead acting as the glue interconnecting the Triad and Edge objects.

This optimized temporal triad network takes up (only) 3.58 gigabytes of RAM, distributed almost 50/50 between node states and connections data:

- State data: 1.80 GB
  - 607 Alliance state objects: 48,560 bytes.
  - 183,921 Edge state objects: 16,185,048 bytes.
  - 37,090,735 Triad state objects: 1,780,355,280 bytes.
- Triad ↔ Edge connections, at 8 bytes per memory address: 1.78 GB
  - 183,921 Edge objects, connecting to 605 Triad objects per edge: 890,177,640 bytes.
  - 37,090,735 Triad objects, connecting to 3 Edge objects per triad: 890,177,640 bytes.

Note that the above only includes network data. The Eve Online alliance standings data is stored separately, by the Alliance Data objects, which the Alliance State objects access to obtain their states. This separation between network and underlying data will become relevant in the next section, where we replicate the network a number of times to speed up the analysis.

6.6 Multi-threaded Network Analysis

Modern computer processors (CPUs) are composed of several "cores", effectively small, single processors. Each core is able to execute a single process, or two if the core is hyperthreaded, in parallel with the other cores. Effectively, multi-core processors act as multiple computers, housed in a single casing. Parallel processing is especially relevant for supercomputers, which achieve their "super" speed by featuring hundreds of cores, executing hundreds of analyses at the same time. The most recent generation of desktop processors, as of this writing, features 10 cores, each able to run 2 processes or *threads*, for a total of 20 parallel processes[1].

Parallel processing can increase the speed of a single analysis, *if* the analysis can be split up. Geographically splitting the network into independent subnetworks would complicate our analysis, but we can subdivide the time period:

[1] https://www.intel.com/content/www/us/en/products/processors/core/i9-processors.html

instead of analysing 437 days in a single, long analysis, we can assign a shorter period to each available core, and integrate the subresults into a single outcome after the cores have completed their analysis.

Since cores run independently from one another, each core would need to have its own instantiation of the temporal triad network. This points to an interesting trade-off between RAM memory and execution speed: by replicating the triad network a number of times, up to the limitation of our total computer memory, we can accelerate the analysis a number of times, up to the limitation of available cores.

For our analysis, we used a desktop computer capable of running 8 parallel processes, with 32 GB of available RAM. The analysis took 10 hours to complete.

6.7 Conclusions

A temporal network is a particular kind of network, with specific implementation challenges. In theory, temporal networks are any type of network, replicated and interconnected a number of times across time. By implementing a State Machine design pattern, it is possible to avoid this replication, which for large networks can make the difference whether the temporal network can be analysed at all.

For our triad network, it was possible to further reduce memory usage by implementing a second optimization, specific to the topology of our network. The applicability of this solution depends on the specific network. We would advise looking for opportunities for optimization within the network topology where possible.

Finally, temporal networks lend themselves to multi-threaded network analysis which, even on a desktop computer, offers significant opportunities for increasing analysis speed.

6.8 Appendix

Following here are the relevant parts of the C++ classes that define the state objects described above.

```
class AllianceState final
{
public:
    // constructor(s), destructor, and other member functions

    // state functions
    void GenerateState(const Date& date);
    double CalculateDistance(const Date& date, const AllianceState* otherPtr) const;
    double GetStanding(const Date& date, const AllianceState& otherRef) const;
```

```
private:
    // connections data
    AllianceData* m_MyDataPtr = nullptr;

    // Alliance state data
    bool m_Exists, m_ExistsPrevious;
    Coordinate m_Coordinate, m_PreviousCoordinate;
    int m_MemberCount, m_PreviousMemberCount;
    int m_SovereigntyCount, m_PreviousSovereigntyCount;
};

class Edge final
{
public:
    // constructor(s), destructor, and other member functions

    // state functions
    void GenerateState(const Date& date);
    double CalculateDistance(const Date& date) const;

private:
    // connections data
    AllianceState *m_FirstPtr, *m_SecondPtr;
    std::vector<Triad*> m_VecConnectedTriadPtrs;

    // Edge state data
    Standing m_Standing12, m_Standing12Previous;
    Standing m_Standing21, m_Standing21Previous;
    Standing m_UndirectedStanding, m_UndirectedStandingPrevious;
};

class Triad final
{
public:
    // constructor(s), destructor, and other member functions

    // state function
    void GenerateState();

private:
    // connections data
    std::array<Edge*, 3> m_ArrEdgePtrs;

    // Triad state data
    TriadState m_State, m_StatePrevious;
    std::array<double, 5> m_ArrWeights;
};
```

Part III

Empirical Contributions

Chapter 7

Measuring Propagation with Temporal Webs

7.0 Preface

The following paper was published in the 2017 *Temporal Network Epidemiology*, edited by Naoki Masuda and Petter Holme [Masuda and Holme, 2017]. It fits in this dissertation as a contribution to Complexity research, itself in the domain of Social Physics, by using data from a virtual world (Eve Online alliance relationships data), to further develop our understanding of the methodology of Temporal Networks.

My contributions to this paper were providing the data on Eve Online alliance relationships, writing the analysis software for tracking propagation through temporal webs (see chapter 6), and the empirical analysis of propagation through the temporal networks, generated from the datasets used in this chapter: Twitter affect data, Russian interbank data, and Eve Online alliance relationships data.

7.1 Abstract

We present a form of temporal network called a "temporal web" that connects nodes across time into a single temporally extended acyclic directed graph as a way to capture contingent behaviors. This representation is especially useful for uncovering and measuring social influence. We first present the general temporal web technique and then use it to analyze three empirical datasets: political relationships in the game EVE Online, interbank loans of the Russian banking system, and Twitter posts regarding the H1N1 vaccine. For each dataset we provide a detailed breakdown of the contingent behaviors using an approach we call temporal influence abduction. We then construct a temporal web for each one and describe

the patterns of propagation found. Based on these patterns of propagation we infer more general properties of influence and the impact of certain types of behaviors in each system.

7.2 Introduction

Tracking and measuring the propagation of diseases, ideas, etc. across a population is an academic problem of great interest as well as a practical problem with important implications. There are many approaches using system dynamics and/or network contagion – and the use of various kinds of temporal networks is the latest advancement in this effort. Even within temporal networks there is significant variation among approaches, as exemplified by the papers in this volume [Masuda and Holme, 2017]. Multiple demands of analysis have lead researchers towards temporal networks [Holme and Saramäki, 2012, Holme, 2015]; two major pulls are preserving the time ordering of interactions [Lerman et al., 2010, Pfitzner et al., 2013, Kim and Anderson, 2012, Grindrod and Higham, 2013, Rocha and Blondel, 2013] and constructing centrality measures for dynamic networks [Mantzaris and Higham, 2013, Nicosia et al., 2013, Bramson and Vandermarliere, 2015, Xu and Wang, 2017].

Contributors to the field of temporal networks are still sorting out best practices and identifying which construction is most appropriate for which questions, and towards that end we demonstrate the use of a version of temporal networks and related measures called "temporal webs" to capture and analyze a variety of problems. Temporal webs are distinguished by their use of cross-temporal interaction and/or inheritance links. So, rather than being a sequence of network time slices connected by node membership, they are always monolithic graphs of the interaction structures across time. Similar, or perhaps even identical, structures have presumably gone by other names; we are not trying to make a serious nomenclature stake, just to identify the set of constructs we are addressing here. A pure temporal web has only cross-temporal edges to create a single acyclic directed graph. This construction has certain advantages in communication networks for which transmission and reception may take several time steps. It also embodies some specific advantages for analysis through the availability of approaches that work on large, sparse directed adjacency matrices.

In what follows we first describe the specifics in building one of the temporal web-style networks and some analysis approaches for them. The focus here is on how to approach problems, and especially the question of influence, by thinking about dynamic interactions as a temporal web. We then go on to demonstrate this approach by applying it to a variety of propagation problems on empirical networks. In this paper we use temporal webs to describe cascades of animosity on political alliance networks in the game EVE Online, understand risk propagation

on the Russian interbank loan network, and the diffusion of emotional affect in the Twitter social network. Although we interpret the results in light of each subject matter, our focus here is on the methodology rather than the substantive issues, so no domain knowledge into these subjects is expected or required.

Temporal web thinking is particularly useful for identifying the most influential nodes in a network — one of the key pursuits in network epidemiology and network theory more generally [Pastor-Satorras and Vespignani, 2002, Kempe et al., 2005, Colizza et al., 2007, Kimura et al., 2010, Kitsak et al., 2010, Yu et al., 2010, Lü et al., 2011, Chen et al., 2012, Dekker, 2013, Sikic et al., 2013, Lawyer, 2015, Bramson and Vandermarliere, 2016, Malliaros et al., 2016, Xu and Wang, 2017]. One can think of influence in two ways, one is to identify the role that an individual plays in a particular chain of events, and the other is to identify its potential role in all possible chains of events. In the first case it is specific to a sequence of observations and in the second it is a dispositional property (like being fragile or brave) that can only be assessed hypothetically (e.g., how much propagation would it cause under thus and such a scenario). In this paper we address the former concept of influence by examining the propagation of node characteristics in empirical temporal networks to assess the influence in that particular chain of events.

7.3 What Are Temporal Webs?

The general description of a temporal network is a mathematical structure that captures interactions across time. In the case of *dynamic networks* the key feature is that the conduits of interactions (i.e., the network edges) change across time so that the graph structure itself must be time-indexed. One can simulate dynamics across these networks, but the network properties of the nodes (e.g., how many friends one has) change over time and complicate the analysis. In contrast, a *time-layered network* typically has a static potential interaction structure and each layer reveals the interactions which occurred during that time step (e.g., how many friends one actually talked to). The difference between the two is fuzzy: if one captures a corporation's email sent each day as a temporal network it can be considered as either (1) a changing collection of interaction patterns or (2) a base set of cooperating colleagues and their de facto communications. This ambiguity is due largely to the primary strength of networks: the edges can represent any relationship among any objects, loosely defined. One is free to interpret the time-varying connections as changes in the link structure or activity across a link structure.

Here we describe a version of temporal networks that emphasizes the activity aspect, and especially transmission activity across time. Instead of capturing the interactions occurring within a time slice for each slice of time, a *temporal web* traces a link across time from the cause to effect. Although this may represent

several different scenarios, it is most natural to think of it in terms of simulations with simultaneous updating in which an agent at time t has an effect on other agents at time $t+1$. Each agent (whatever a node represents) has certain information (its state) at time t and that dictates its behavior, including its interaction behavior, at that time. Then each agent may change its state based on its behavior and input from other agents to reach a new state at $t+1$. Thought of in this way it is obvious that for this situation the best representation is to have nodes at t interacting with (affecting) nodes at $t+1$ instead of nodes at t. As mentioned earlier, one distinguishing feature of this approach is that the result is always a single acyclic digraph rather than connected layers of time-slice networks.

Naturally there is no limitation to discrete, integer, or uniform time increments; although that case is the simplest to represent.[1] It is possible, for example, that a message sent on Monday is read on Tuesday by some people, but on Wednesday or Friday by others. It is simple to incorporate links that connect a node at t to other nodes at $t+s$ ($s>0$) for possibly heterogeneous s as long as the s times are discrete time steps. This is so because these connections would fit naturally into a temporal web adjacency matrix filling in spots outside the superdiagonal blocks of the t to $t+1$ connections. Allowing for continuous time dynamics is also possible, but it requires a switch in representation and an accompanying change in algorithms to analyze propagation dynamics, so we leave that out of the current work.

Pure temporal webs are a representation choice that is only appropriate for some purposes and datasets. Temporal webs can obviously be generated from temporally layered symmetric graphs by changing a link between nodes A and B at time t to a pair of directed links from $A(t)$ to $B(t+1)$ and from $B(t)$ to $A(t+1)$. But such a conversion naturally carries with it the assumptions and interpretation of temporal webs, which may or may not be appropriate for a particular temporally layered graph. Specifically, temporal webs (as interpreted here) are best used for actions such as the spread of a disease or communication of an idea and less appropriate for other relationships if those are not connections that happen across time (like co-location, club membership, or being connected by a road). In many cases actions are instantaneous in their effects, and a chosen time resolution (e.g., daily updates) masks the true behavioral pattern by obscuring the temporal ordering of events. Some of our methods presented below are designed to address exactly these issues by creating temporal webs from data so that they make the best of the actual influence relationships across time, even if that isn't a pure temporal network, and performing analyses using the general temporal web thinking as appropriate.

[1] There is also no strict limitation that the interactions be instantaneous – they may be spans of time in which the agents are related – but that is a further extension beyond what we cover here. See [Viard and Latapy, 2014, Wehmuth et al., 2015] for more information on link-stream graphs.

7.4 Analyzing Temporal Webs

Because pure temporal webs generate a single acyclic digraph from the entire system behavior over time, there are many off-the-shelf directed network measures that can be immediately applied. However, although these measures will return a value, the interpretation of that number is often twisted or impossible. They fail to actually inform our understanding of the system's dynamics. By analogy, I can find the average value of a list of phone numbers, but that doesn't mean the result is a viable phone number or meaningful in any other way.

As a specific example, the diameter of a temporal web is always equal to the time duration T (or infinity when some node at $t = 0$ has no time-order preserving path to some node at $t = T$). This is so because for any propagation there must be at least T steps along any path when the duration of the run is T to reach the final step. At each step all links connect from t to $t + 1$, meaning that along any path from $A(0)$ to $B(T)$ it can only make one network step per time step. So, even though diameter is a generally useful measure for assessing connectivity (even in acyclic directed social networks), it is not useful for temporal webs due to their specific construction.

Other measures, such as betweenness centrality and clustering coefficients are similarly uninformative for temporal webs for reasons discussed in [Bramson and Vandermarliere, 2015]. There has been considerable attention paid to centrality measures on temporally layered networks [Braha and Bar-Yam, 2006, Holme and Saramäki, 2012, Kim and Anderson, 2012, Rocha and Masuda, 2014] as well as temporal clustering [Cui et al., 2013, Nicosia et al., 2013]. The merits of these measures for those structures are still being evaluated, but their value after adapting them for temporal webs is much more suspect. Certainly there is a desire for measures of centrality, clustering, communities and other network features on temporal webs, but they mean something slightly different than on other temporal networks and must be differently formulated. We do not put forward or test such temporal network measures in this work, but rather propose a way of thinking about transtemporal interaction data to ground the development of such measures.

7.4.1 Temporal Knockout

In order to benchmark the success of network measures in capturing all-things-considered influence, a measure called *temporal knockout score* (TKO) was presented in [Bramson and Vandermarliere, 2015] and refined in [Bramson and Vandermarliere, 2016]. At the limit of an infinite number of possible disease scenarios, this measure considers every possible disease transmission trajectory among a population, and then for each of those trajectories it tests how much the disease morbidity changes when each agent at each time is removed from the population. Because each agent at each time is represented as a node in a temporal web, these

"agent-time" nodes constitute the unit of measure for infections. The agent-times (temporal nodes) that consistently have the greatest effect on the disease magnitude have the largest TKO score. Because this measure exhaustively covers all possible disease trajectories, it fully captures potential influence. By narrowing the set of test scenarios it can provide a more focused contingency analysis, such as to particular initial agents or particular interaction structures, and measure context-dependent influence.

Temporal knockout does not strictly depend on the temporal web structure, but the temporal magnitude measure of disease morbidity that it uses does depend on transtemporal records of the agent states. For example, in a hybrid temporal web in which the effects of interactions within a time-step are represented as a change in state during that period, then the interaction edges are not transtemporal; however, as long as there is inheritance of states across time it is possible to calculate the magnitude of the spread and therefore TKO. TKO does have its drawbacks however. Because it analyzes the removal of every agent at every time and reruns the dynamics in that counterfactual scenario it is extremely computationally intensive. Furthermore, the approach is intrinsically mechanistic. After the removal of an agent-time one must rerun the dynamics from that point forward, and doing so requires using some mechanism to change the agents' states. The mechanism can be as simple as a probability of infection, but the resulting measure of influence is dependent on the mechanism used. Is the case of the empirical temporal webs used in this analysis, we do not know the mechanism that produced the observed propagation, so we must either infer a mechanism and use that to determine TKO, or we can apply an alternative technique that still captures the essence of contingent marginal change in propagation magnitude.

7.4.2 Temporal Out Component Paths and Refinements

As a way to partially ameliorate the prohibitively long computational times of TKO analyses, [Bramson and Vandermarliere, 2015] explored the use of various proxy measures on temporal webs of simulated disease spread. One class of such measures utilizes variations on the temporal out component of a node; i.e., all the nodes reachable from a given agent at a given time. One branch of variations included different ways of weighting the future: the number of reachable temporal nodes, the total length of all paths to those nodes, or the out component paths divided by the number of infections at that time step. Another branch of variations looked at ways to adjust the out component measures based on the history up to that point: weighing the contribution of each node in the out component by its in-degree, its in-component size, or the number of redundant paths leading to it. That paper also introduced a measure called *nexus centrality* as the sum of the lengths of all the paths running through that node as a way to capture bottlenecks.

Previously these measures were applied to simulated disease dynamics on a variety of network topologies to determine how well they matched TKO identifications. The results presented in [Bramson and Vandermarliere, 2015] indicate that the temporal out component paths (TOCP) measures was highly correlated with TKO over 1000 simulations in both SEIR and SEIS dynamics. Although TOCP techniques consistently and drastically out-performed other network measures calculated on the base interaction network on average, the TKO values for temporal nodes is highly variable from run to run – sometimes even having negative values. Negative values occur when removing an agent at a certain time results in more overall infection. Though perhaps unintuitive, this results from the time varying network connections: the neighbors infected early on spread to only a few of their neighbors, but when infected later through another pathway they infect many more.

In addition to being much faster to calculate than TKO, the use of out component paths does not require a mechanism to evaluate influence. These measures can be calculated on the kinds of empirical networks examined below to provide an indication of influence. Although we won't have the TKO measure to benchmark against, it is still possible to assign a score to each temporal node for its potential impact (the full network) or actual impact (just through nodes that are "infected") based on the temporal out components. In this chapter we analyze behavior by types of nodes rather than particular nodes and thus use aggregated measures of out component paths to assess patterns in propagation potential.

Another advantage of these out component based proxy measures for TKO is the ease of converting them into measures for quantitative (rather than categorical) propagation and to continuous time flows (rather than discrete time steps). Instead of counting the number of infected agent-time nodes in a temporal out component one can calculate the product of the time and quantity for each reachable agent-stream. This is essentially the area under the curve for the property across time starting after their interaction with the focal agent-time. The details depend on the nature of that property; for example, the calculation could use the value of the property at the time of interaction as a baseline value and thus determine whether the interaction has had a net positive or negative effect on the property. These considerations are taken up in other work that explores continuous time temporal webs, but here we continue using temporally coarse-grained data.

7.4.3 Temporal Influence Abduction

When analyzing influence on empirical temporal webs, i.e., ones for which there is a particular thing that already happened, we need to rework the idea of temporal knockout to fit the application. We still would like to know how much of the propagation each agent at each time is uniquely responsible for, but the data

provides what is essentially the result of one run of a simulation. In order to assess influence (rather than just the luck of how things turned out), we need to compare the observed system behavior to counterfactual behavior. As already mentioned, one way to do this is to model a generative mechanism that recreates the observed behavior and use it to populate hypothetical system behavior for comparison. Many machine learning techniques (such as Bayesian networks and Markov models) use this approach, and something like this could be developed that is specifically adapted to temporal webs. Although here we take a more direct statistical approach to describe consistent patterns in the dynamics, the behavioral patterns we uncover could very well be used for such a generative model in the future.

The general approach described here, called *temporal influence abduction* (TIA), uncovers the particular impact each agent has on the others conditional on its state (or its changing state). This is abductive because it infers the likely influence pattern as the best supported explanation for the available evidence, without proposing a mechanism by which the influence occurs. Take disease spreading as a familiar example. We can examine several relationships among nodes and their disease states to uncover whether they are playing a key role in spreading the disease: How many of one's neighbors become infected after interacting with an infected focal agent? How different is this proportion from the proportion when the focal agent is uninfected? How consistently do the each of disease states have this difference in effect across time? How similar are all the agents in these effects? Given these ideas about how influence would effect the actual spread, we can generate hypotheses about who is the most influential agent and quantitatively test them for being the best explanation.

A more concrete example makes this clearer. For categorical data, conditional infection rate can be assessed by collecting for each agent the number of (or proportion of infectable) neighbors that become infected each time step that the focal agent is infected. That is one distribution of values for each agent and across agents. Then do the same for when the agent is not infected. Those distributions can be compared using a Kolmogorov-Smirnov two-sample test to determine whether they are significantly different; the z-score acts as a measure of their being different patterns. Of course one must also confirm that the distribution mean is higher when agents are infected than when uninfected because we are generating hypotheses about disease propagation. To check the consistency of the effect we can look at the dispersion of the two distributions; if the distribution is tight then it is more likely to be a systematic feature of propagation rather than a spurious connection. Finally, for a particular agent to be a key player in the spread of disease, an agent would also need to have consistently higher scores on all these measures than other agents.

Temporal influence abduction can also be performed on quantitative (rather

than categorical) agent properties. Rather than a disease state of susceptible or infectious, one might be interested in how much confidence each person has in some idea and how that level of confidence spreads. In this case we can measure how much the confidence level of neighbors changes towards the confidence of a focal agent (total delta absolute distance) multiplied by the confidence of that focal agent (because a lack of confidence does not spread). With these values calculated at each time step we can also derive a distribution of values, but without categories we must discern patterns in the distribution endogenously. The details of such an analysis will vary from dataset to dataset depending on what the property is; for example, a lack of confidence should be down-weighted because a lack of something cannot propagate, but other properties (even other epistemic properties) may be just as contagious at every level. In what follows we provide hybrid TIA analyses of three datasets via a coarse-graining of the value ranges with the plan to follow up each one with a more in-depth continuous quantitative analysis in future work.

7.4.4 Measuring Empirical Propagation of Observed Properties

Measuring the influence of agents and times on the spread of hypothetical diseases on an abstracted social network is useful for many theoretical reasons. But capturing and understanding the observed spread of a property across an empirical network requires a distinct set of measures and methods. For one, rather than being in one or another disease state, the nodes in these systems may have multiple relevant properties that each take a range of values. Additionally, rather than probabilities of transmission, what we have is particular actual transmissions. Although we know which transmissions/interactions actually take place, we often don't have the potential interaction structure nor its implicit properties such as edge density or degree distribution. Furthermore, although we can trace the temporal node properties across the temporal web and provide measures of influence in terms of increased and decreased marginal spread, we cannot (as is usual in empirical science) bridge the correlation-causation divide nor eliminate exogenous influences.

Despite these limitations, the temporal web techniques described above can shed some new light on observed propagation dynamics. Below we examine three temporal network datasets with propagation, demonstrate the kind of conditional effect analysis required for temporal influence abduction, examine the temporal out components of types of actors in each system, and tie these results back into substantive insights into each system.

7.5 Propagation of Frustration on Political Relation Networks

Our first dataset is a temporal web extension of Structural Balance Theory (SBT) [Heider, 1946, Cartwright and Harary, 1956, Harary, 1959] that makes the propagation dynamics of frustration explicit. SBT utilizes networks of positive and negative relationships among some sets of agents (typically people or countries) and provides a characterization of when such a signed social network is balanced or how frustrated it is. The implication is that frustrated relationships are more likely to change than balanced ones, and so the theory is implicitly one about the driver of signed social network dynamics. There exist analyses of how much the aggregate frustration changes over time according to generated dynamics [Hummon and Doreian, 2003, Antal et al., 2006] or empirical networks [Leskovec et al., 2010, Szell et al., 2010, DuBois et al., 2011], but no approaches that directly track the spread of frustration from triad to triad.

According to strict SBT a triad of nodes is frustrated whenever there is an odd number of negative links among them.[1] Thus if a triad of agents A, B, and C is frustrated then exactly one or exactly all three of its edges are negative. When any single edge (say $A \bullet\!\!-\!\!\bullet B$) changes valence, the triad becomes balanced – regardless of whether that edge was originally positive or negative. Now consider that agents A and B were also in a triad with agent D, and that the $\triangle_{ABD}$ triad was balanced. When the edge $A \bullet\!\!-\!\!\bullet B$ changes valence to balance the triad $\triangle_{ABC}$, this *necessarily* causes the neighboring triad $\triangle_{ABD}$ to become unbalanced (that is, frustrated[2]). In fact, all triads using the $A \bullet\!\!-\!\!\bullet B$ edge will flip between balanced and frustrated. In some cases these local changes can balance a large number of frustrated triads, increasing overall stability. This change can also create a large number of newly frustrated triads. These newly frustrated will likely then change other edge valences to become balanced, which serves to push the frustration starting with $\triangle_{ABC}$ through the social network. The methods presented in this section aim to track this frustration propagation using a temporal web of the network of triads.

7.5.1 Constructing a Network of Triads

In our general temporal web approach, the nodes carry the properties that are propagated, so to apply this technique to structural balance theory we first convert the social network with signed edges to one in which the nodes are frustrated or bal-

[1] Originally it was formulated as an odd number of links across all paths, but the triad version has become dominant [Abell, 1968], cf. [Facchetti et al., 2011].

[2] For our analysis below we distinguish among strongly and weakly balanced and frustrated. Thus for us being unbalanced implies being one of the two kinds of frustrated, but that is not the only terminology. 'Balanced' can also refer to what we call 'strongly balanced', and hence our 'weakly balanced' would be 'unbalanced' but not frustrated.

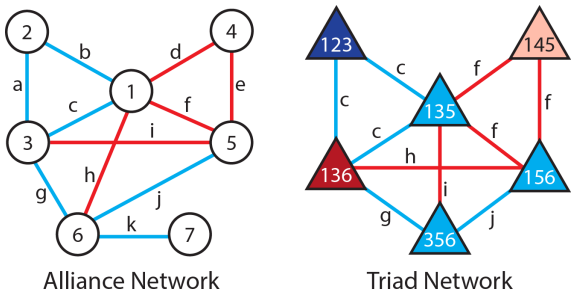


Figure 7.1: An example of converting a signed social network into a triadic network with node properties corresponding to the frustration property of each triad in the alliance network. Dark blue is strongly balanced, light blue is weakly balanced; likewise dark red and pink are strongly and weakly frustrated, respectively. The frustrated triad $\triangle_{136}$ can become balanced by changing any one of the edges c, g, or h. Note that although changing c will balance $\triangle_{136}$, that edge is also used by the triads $\triangle_{123}$ and $\triangle_{135}$ and so such a change will force both them to become frustrated. If $\triangle_{135}$ then rebalances itself by changing edge f this will propagate by balancing $\triangle_{145}$ and making $\triangle_{156}$ unbalanced. The triad network makes clear the level of frustration and the consequences of particular edge flips.

anced. Our way of doing this is to construct a *triadic node* for each connected triplet of agents in the network that holds a property for the state of the triad (e.g., how many positive and negative edges or whether it is frustrated). Triadic nodes are connected by an edge whenever the triads they represent have an edge in common, but the triadic graph edge representing $A \bullet\!\!-\!\!\bullet B$ in the original graph will occur as many times in the triadic graph as that edge is shared by triads.[1] Although the number of triadic nodes depends on the structure of the network, and especially the edge density and clustering coefficient, because edges can be reused in many different triads, there are typically many times as many triadic nodes as there are agents in the social network. An example of converting a signed social network into such a triadic network appears in Figure 7.1

In nearly all SBT analyses the network structure is static and only the sign of the links changes. For applications to empirical data, however, the approach must also accommodate dynamic networks to reflect both new or lost edges among existing agents as well as agents entering and leaving the system. Just as a single edge sign flip can alter the frustration of a large number of triads that include that edge, the addition/deletion of a single new edge may sharply increase/decrease the number of triads; however, these changes cannot change the frustration of existing triads. In addition to frustration propagation we must therefore also consider ex-

[1] Triads can only share one or zero edges. If two triads share two edges, then they both also include the nodes that are the source and target of those edges. The minimum number of source and target nodes for two edges (in a non-reflexive and non-multigraph network) is three. If two triads share three nodes, then they are identical triads.

ogenous events in both the changing of edge valences as well as edge creation and deletion. An aggregate approach to measuring system-wide frustration is inadequate for that task, so we focus on marginal changes in frustration in the temporal out component of state-changing triadic nodes conditional on each type of state change.

If the network is balanced, any cascade must be initiated by an exogenous change. Even in an unbalanced network, exogenous events can flip edge valences. However, as we will soon see, empirical political relationships are more tolerant to long-term frustration than SBT implies they should. Regardless of the spark, a single edge sign change will reveal itself as a change in state for all triadic nodes that use that edge. Then when any triadic node switches state – whether from frustrated to balanced or vice versa – all the neighboring triads must also flip their state unless: (1) an even number of changes occurs via multiple changing neighbors, or (2) a link is removed thus dissipating the triad.

7.5.2 Description of the Dataset

Here we apply our social frustration tracking technique to data of political standings from Eve Online (EVE), a massively multiplayer online game. Players in the game can create and run corporations, which have between zero and more than 10,000 player members. The leader (or an elected group of directors) of a corporation has complete power to determine corporate policies and represent the corporation. Corporations can (and usually do) form official collectives known as alliances. The number and sizes of alliances change over time, but most alliances have a few thousand character members (some are empty and the largest alliances include more than 17,000 members spread out over hundreds of corporations). The leadership of one of the member corporations will control the alliance and has the sole authority to set the lists of enemies, competitors, non-aggressors, and allies. These standings are set on a scale from -10 (enemy) to 10 (friend); setting it to zero means actively neutral (which is in effect like being an enemy), but they can also be removed completely. These alliance standings need not be reciprocal, though they almost always are. In the game, the icons for ships of an enemy alliance will show up as red on the screen, alerting players that they are both potential predators and potential prey. Friendly alliance ships appear green, while the icons for ships in alliances with unset standings appear white (as do asteroids, space bases, and other unaligned game objects). Attacking members of friendly alliances is possible, but such players are typically punished by their own alliance by paying a fine or being kicked out (potentially losing access to their personal assets).

The game environment is broken into 7930 solar systems of which 3524 are conquerable by the players. These conquerable systems are where the alliances can hold sovereignty over a system and ownership of its stations. Ownership allows

them to decide who is allowed to place and retrieve resources from its stations, collect taxes on transactions occurring in that system, and have access to the natural resources in that system. There are many alliances that do not hold sovereignty over any systems, and therefore the political standings of these alliances only serve to establish conflict permissions. Here we perform an analysis of all alliances with more than 200 player members (which includes all sovereign alliances and many large alliances holding no territory).[1]

As already mentioned, the size and number of alliances change over time depending on player actions. We analyze the alliance standings data from 2/4/2015 to 4/17/2016 and within this time frame there are 606 unique alliances with 200 player members or more at some point during our time frame. On average there were 328 large alliances in play each day, which means there is a lot of turn-over among the alliance with 200+ members. Each alliance can set its standing to every other alliance, but in practice the standings matrix is quite sparse. The alliance matrix also reveals a block structure showing coalitions of alliances: an unofficial collection of alliances that agree to support each other without any mechanism in the game to associate them. Because the standings relationships are nearly perfectly reciprocated, we construct the triad network from a symmetric version of the standings matrix by setting the links to be negative if either direction is negative and positive if both directions are positive.[2]

We primarily make use of a looser variation of SBT for determining frustration in signed social networks that seems more appropriate for EVE. Part of the fun of a multiplayer game like EVE is attacking other players, so the stable outcome of globally friendly relationships is not expected here. Furthermore, in (game and real-world) politics we expect there to be long-standing mutual animosities among more than just two coalitions [Davis, 1967]. These coalitions may form temporary truces with each other to team up against a mutual and otherwise unbeatable enemy, but then go back to being enemies when the immediate threat is over. This is true for world politics [Axelrod and Bennett, 1993] and we should expect the same for EVE. Thus in the looser version of SBT we consider the triple negative triads as being weakly frustrated (essentially non-frustrated but non-balanced) compared to strongly frustrated single negative triads [Davis, 1967, Szell et al., 2010]. We can also consider triple positive triads as being more strongly balanced than the weaker "mutual enemy" single positive triads; however, in the current analysis we count them both equally as balanced. As you can see in Figure 7.2, the triple neg-

[1] In future work we will compare these results to the subset of sovereign alliances. Because the standings of sovereign alliances have a larger effect on the players, we expect there to be a stricter adherence to structural balance, but here we are primarily interested in whether the temporal web approach can capture and detect frustration propagation.

[2] A more nuanced approach involving continuous-time analyses of frustration dynamics also weights the edges by both the value set and considering both directions. Presenting such an analysis requires details into EVE and into structural balance considerations not pertinent to our goal here of demonstrating the use of temporal webs to track frustration propagation.

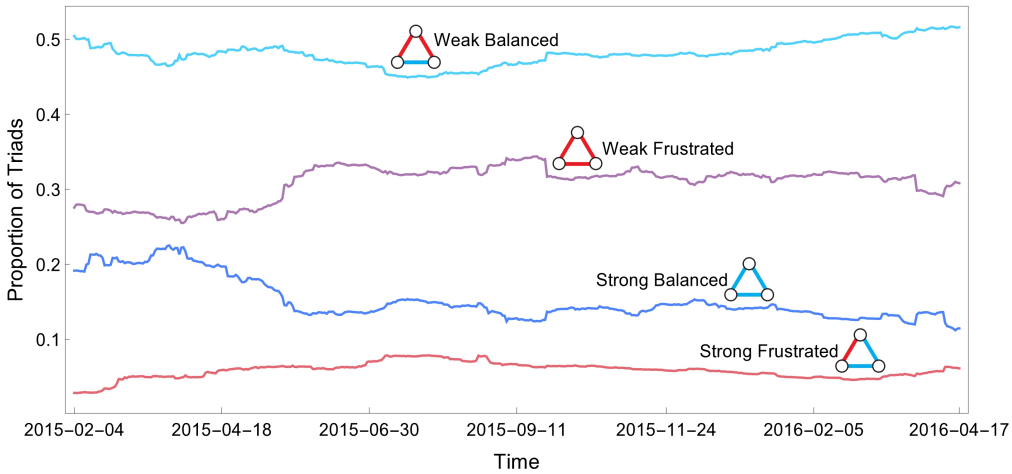


Figure 7.2: Time series showing the proportions of each state of triad each day. Although the number of triads more than tripled during this period the relative proportions are strongly persistent. Roughly half are balanced "mutual enemy" relations while a third are weakly frustrated triple negative triads. Their consistent presence bolsters the claim that within EVE these triads do not actually cause political frustration. Triple positive triads, which are the norm within coalitions, are overall underrepresented. On average across time less than six percent of the triads are truly frustrated, lending evidence to a loose structural balance mechanism in this context that is consistent with anecdotal accounts within the game.

ative weakly frustrated triads make up roughly a third of the triads across time. Within coalitions everybody is friendly with each other and they all share the same enemies, but nearly everybody in coalition X is aggressive to everybody in both coalitions Y and Z.

7.5.3 Temporal Web Analysis Results

The set of alliances having at least 200 members went from 311 at the beginning, peaked at 349 in the middle of the dataset, and dropped to 312 at the end, but the number of triads steadily (though nonmonotonically) grew from 50,491 to 170,296. The disproportionate increase in the number of triads comes from an increase in edge density in the alliance network. So the network here is dynamic both in the sense that nodes come in and drop out of the system and that the edges among them also form, dissolve, and change sign. This poses no difficulties for a temporal web analysis because each agent at each period is a distinct temporal node object. They are only connected to future selves through a chain of cross-temporal links. And because our temporal nodes are the triads, this opens the possibility to identify cases when ending a relationship that is part of a frustrated

triad is used instead of balancing it.

In order to determine whether alliances use link elimination as a strategy for frustration elimination, we look at the difference in rates of triad dissolution for triads of each of the four kinds of frustration. Because relationship changes may be motivated by a variety of game considerations besides political frustration, there will be a base rate of dissolution. The hypothesis is that dissolution counts as a strategy only if the rate for strongly frustrated triads is significantly greater than the base rate. Although the overall number of triads increases over time, there are still 116,799 cases of standings links being removed. Removing a standings link from the alliance network removes all the triadic nodes and triadic links that utilize it, and doing so cannot affect the state of the remaining triads. Conversely, we can look at how many frustrated triads are generated when (1) a new alliance enters the considered set or (2) a new link is generated between existing alliances.

Scenario				
Triadic Node Creation	0.185	0.0799	0.491	0.244
Triadic Node Persistence	0.146	0.0584	0.484	0.312
Triadic Node Dissolution	0.228	0.123	0.443	0.206

Table 7.1: Summary of results showing the proportion of the triad states when they are created, that exist in the system over time, and when they are removed. This only includes adding and removing triads through link creation and destruction; i.e., excluding nodes entering/leaving. Approximately 8% of newly formed triads are strongly frustrated. The type proportions at creation closely match the overall number in the system with a bit less strong frustration (∼5.8%) and a bit more weak frustration. The proportions of eliminated triad types show a greater focus on strongly balanced and strongly frustrated triads (∼12%), and a tendency against dissolving weakly frustrated triads.

From the results in Table 7.1 we can see that around 8% of new triads formed through link creation inject strong frustration into the system. This proportion is not much higher than the proportion of strongly frustrated triads across time, meaning both that on average the frustration introduced is partially resolved (possibly through the removal of the links that created the strongly frustrated triads) and that this process is potentially responsible for injecting all of the observed frustration. However, recall that if this is the case then it would also mean that frustration rarely propagates and is instead sequestered to the original frustrated triad until it is dissolved. To determine to what degree frustration propagates we need to look both at state transition behavior and neighbor behavior.

Table 7.2 shows the relative rates of triad state transformations to reveal biases in how the different triads states change their states. Most notably we see how stable the triads' states are: on average more than 99% of all triad states stay the same from day to day. Strongly frustrated triads are the least stable, but only by

Triad state at time t	Triad state at time $t+1$				
→ ?	0.993	0.00251	0.000624	0.00000171	0.00415
→ ?	0.00346	0.987	0.00408	0.000119	0.00561
→ ?	0.0000480	0.000372	0.996	0.00125	0.00243
→ ?	0.000000729	0.00000627	0.00129	0.997	0.00175

Table 7.2: Summary of results of the proportions of triad state changes including though the deletion of edges (but not nodes). Blue edges are positive, red edges are negative, and gray triads are nonexistent. The large proportions along the diagonal indicate that alliance standings are highly stable and therefore from day to day all triad states are likely to persist. Notably the strongly frustrated triads are the least stable while the weakly frustrated (triple negative) are the most stable by tiny margins.

small margin. Because the political standings, and therefore triad states, can be long lasting even when there is frustration, we also look at the next state of the triad conditional on some change occurring in Table 7.3.

Triad state at time t	Triad state at time $t+1$				
→ ?		0.344	0.0856	0.000235	0.570
→ ?	0.261		0.307	0.00895	0.423
→ ?	0.0117	0.0907		0.305	0.592
→ ?	0.000240	0.00206	0.423		0.575

Table 7.3: Summary of results of the proportions of conditional triad state changes including through the creation and deletion of edges (but not nodes). Blue edges are positive, red edges are negative, and gray triads are nonexistent. Although it is possible for triads to change multiple edge valences in one step because of our use of daily data, we see here that most changes are one edge flip away. However, in the first row we can see that nearly 8.5% of changes from strongly balanced triads go to weakly balanced within a day. This is seven times more frequent than the next largest 2-step change (from weakly balanced straight to strongly balanced). Similarly, weakly balanced triads are more than three times more likely to become weakly frustrated than strongly frustrated. This evidence supports a claim of general avoidance of strong frustration.

Table 7.3 shows the same counts as Table 7.2 but renormalized without including the static entries. From this we can more clearly see the changes that occur when they actually occur. The first observation is the large proportions of triads that are dissolved through link removal; except for strongly frustrated triads they are more likely to be removed than changed. The second observation is that when a triad does change state it is proportionally more often to be a single valence flip

away. Note that because of our daily time resolution, it is possible for many edges to change to occur in one step, but what we see is that this is a rare occurrence. This is consistent with the high level of standing persistence seen in Table 7.2.

Looking in more detail at the differences among triads states, we can see that when single negative (strongly frustrated) triads change they are split similarly between becoming triple positive or weakly balanced triads and the least likely to dissolve. In contrast, the single positive (weakly balanced) triads are three times more likely to switch to weakly frustrated than to strongly frustrated and are the most likely to be dissolved. This may indicate a conscious aversion to making changes that generate strongly frustrated triads, so much so that it is better to dissolve the relationship than to make it frustrated. All of this helps us understand the dynamics of frustration changes in the dataset, but does not yet address the question of propagation.

We can look more directly at the propagation effects in Table 7.4. The cases in which non-existent triads stay non-existent (bottom row) act as baseline rates for the transitions of neighbors because the neighbors are not reacting to anything.[1]

In the row showing proportions of strongly frustrated triads becoming strongly balanced ( → ) the 14.3% of neighbors also changing from strongly frustrated to strongly balanced (red highlight in Table 7.4) is much higher than the baseline (0.02%). These are mostly neighbors sharing the negative edge, so when it becomes positive all triads using it gain an extra positive edge. This change also balances neighboring weakly frustrated triads by converting one negative link to a positive link (13.2% of neighbors). You can also see in that row that 5.34% of neighbors change from weakly balanced to strongly frustrated (green highlight in Table 7.4). While this number is small, it is still a considerable number of new strongly frustrated triads and much larger than the base line rate. This is one side of the propagation phenomenon we are looking for — the side of balancing by changing from negative to positive.

Because frustrated triads can also become (weakly) balanced by changing one of the positive links to a negative, there is another side of the propagation story. Specifically, in the row of strongly frustrated changing to weakly balanced ( → ) we see that 7.6% of the neighboring triads go from strongly balanced to strongly frustrated blue highlight in Table 7.4), while 10.36% change from strongly frustrated to weakly balanced (yellow highlight in Table 7.4). So for both directions of edge changes, more neighboring triads become balanced than become frustrated. Together this indicates that whether it is an explicit decision or an implicit pressure, it is the case that changes in standings tend to reduce the overall

[1] One can include all static focal node states, but the number of cases of triads staying non-existent overwhelms the effect of including the others so that the largest differences is one in ten thousand. Furthermore, a comparison of the existent static to the non-existent static cases shows that the neighbor transition proportions are similar to within one in a thousand, bolstering our claim that these static cases can act as a baseline.

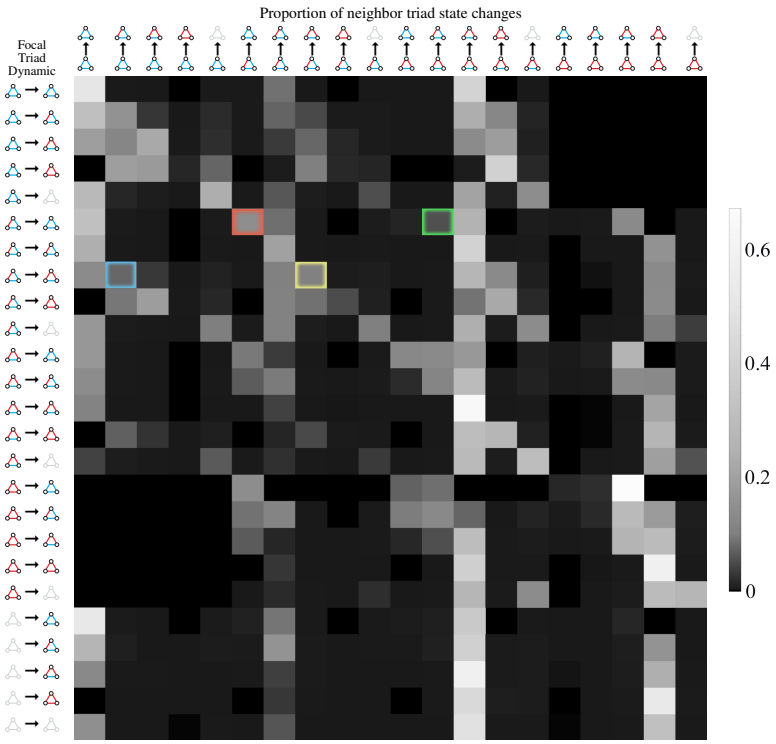


Table 7.4: Summary of results of the triad states changes among neighbors, contingent on each type of focal node state change (including creation through new links and no change). Specifically, for each focal triad's change from t to $t+1$, the table shows the proportion of each type of triad state changes of those triads that share an edge in the standings data during that same period (i.e. neighbors in the network of triadic nodes). The construction of the triad network requires that when an edge in the standings network changes state, all triads containing (and connected by) that edge also change state. Boxed cells are discussed in more detail in the text.

(strong) frustration of the system. This tentative conclusion stands as evidence in defense of loose structural balance.

Looking at the behaviors of individual triads is probably not insightful, so we focused on characterizing the behavior/influence of triad types. The proportions of conditional reactions in Table 7.4 combined with the proportion of behaviors in Tables 7.1 and 7.2 go far in capturing the behavior of the system. These state and rate proportions provide the necessary information to generate simulations with similar behaviors and/or perform a Bayesian-style analysis of the alliances' behavior. However, as stated in Section 7.4.3, we are taking a more direct descriptive approach in this work.

Collectively the results show that, based on immediate neighbors, frustration

does propagate to some degree, but it loses momentum with each step as more stabilize than become frustrated. These averages across all system dynamics are used to detect global tendencies, but below we use the out components of balancing triads to track the trajectories of each frustration cascade to see if these conditional neighbor changes chain together. Before moving on to that topic, we make some other conclusions directly related to structural balance theory.

Edge Elimination Strategy. We can discern from the results so far the degree to which edge removal is used by the alliances as an (implicit or explicit) strategy to reduce political frustration. This can be seen both from strongly frustrated triads transitioning into nonexistent ones, and also the relative frequency of transitioning into a nonexistent state instead of a frustrated one. Table 7.2 shows that strongly frustrated triads have the highest rate (by a small margin) of being dissolved (0.00561), indicating that eliminating the triads makes up some of the difference of their lower stability. Table 7.3 shows that conditional on being in the strongly frustrated state, the triads are relatively more likely to switch to a neighboring state (i.e., change one edge valence) than to eliminate the frustrated triad through edge deletion. This result is counter to the hypothesis of an explicit edge elimination strategy.

Now recall that when a triad changes state to balance itself, that edge valence change will unbalance any other balanced triads also using that edge. Table 7.3 shows that these balanced triads tend to dissolve rather than become strongly frustrated, lending evidence to a claim that edge elimination may count as a strategy for frustration *avoidance* rather than frustration removal. One way to think of this is that when a strongly frustrated triad is forced to change it is usually to one form of balanced triad, so there is little pressure to dissolve it unless players were consciously considering the aftershocks it would create. The observed change pattern thus provides evidence that the players commanding alliances in EVE are largely obeying the tenets of structural balance implicitly but not purposefully.

Temporal Influence Abduction. The results described above comprise the basis of the temporal influence abduction analysis described in section 7.4.3. Rather than assessing whether any particular alliance, or any particular triad, is responsible for propagating frustration, we have investigated the behaviors of each type of triad state. We are interested in whether a frustrated triad is more likely than not to cause other triads to become frustrated in an effort to balance itself. To do this we need to look at the neighbors of changing triads to see what effects they have, and specifically we are interested in triads that change from (strongly) frustrated triads to balanced triads. Table 7.4 shows the proportions of all changes in neighbors immediately effected by changes in (or lack thereof) of each agent at each period. Although standings among persistent alliances are rather stable (indicated by the low variance in proportions for the rows with unchanging focal frustration), when there is a change the nature of structural balance forces a change in the frustra-

tion/stability of some neighboring triads. So this reflects less the behavior of the players and more the rules of structural balance. However, we can confirm that certain transitions are more likely than others to shift frustrations to other triads.

A strongly frustrated triad has one negative edge and it can become balanced by either changing that negative one to a positive one (26.1%) or by changing one of the positive edges to another negative one (30.7%). When changing to a positive one, that would propagate by converting weakly balanced neighboring triads into strongly frustrated ones, but this doesn't happen often (9.1%). Strongly balanced triple positive triads convert to strongly frustrated triads 24.4% of the time, thus they are the main vehicle for frustration propagation. In fact, from Table 7.3 you can see that the strongly balanced triads are 3.8 times more likely to become strongly frustrated, and even then 57% of the time they do change, they are dissolved instead.

Triple positive triads are common among the alliances within a coalition, and when a pair of such alliances have a falling out all of the other alliances in the coalition are still friends with both. This causes an upswell in frustration: 62% of neighbors stay the same (mostly because they weren't using the edge that went negative) but 15% become strongly frustrated. When a triple positive instead skips over the strongly frustrated step and directly becomes weakly balanced (adding two negative), 12% of its coalition partner's triads become strongly frustrated, although 22% also skip the frustration step and pick sides within a day. From this it doesn't seem that balancing triads is the main source of frustration propagation, but we can look from another perspective to make sure.

These probabilities so far have been conditional on the focal triad dynamic, but we get a slightly different view when we normalize by all the events that could directly create a frustrated neighbor (i.e., proportions down the column of all focal triad transitions that result in strongly frustrated neighbors). Frustrated focal agents spreading frustration when they become (either kind of) balanced only make up 13.6% of the frustration injection into the system. 27.3% comes from triple positives gaining a negative edge just mentioned above, 21.5% comes from weakly balanced triads becoming weakly frustrated, and 9.4% comes from when weakly balanced triads become frustrated. All other transitions induce 5% or less of the frustration into the system. From this data we see that changes to balance the network do happen, but they tend to remove frustration rather than spread it. Most frustration injection comes from pairs of alliances in a coalition becoming enemies; this destabilizes both their internal coalition triads and their mutual enemy triads.

Temporal Out Components. The above analysis looks at immediate changes to understand how well the micro-dynamics conform to structural balance, but propagation is a macro-dynamic that can only be discerned by looking at longer-term patterns. This is where we pull in the temporal out component. We want to

understand how far the effects of correcting for frustration ripple through the network. There is a trigger event (the reduction in frustration of a triad) that causes immediate flips in the frustration state of some neighboring triads. If any neighboring triads become frustrated from these flips, then that counts as part of the magnitude of the trigger node. We follow all these frustrated triads across time (because they need not react immediately) to see if they balance themselves and push the frustration to their neighbors (potentially back to the original trigger). We count both the number of triads that are frustrated (cumulative cases) and the number of frustrated temporal nodes (magnitude) that could be the result of that trigger.

Of course there are also link deletions that cut off frustration without propagation and multiple trajectories in which changes from triggers effect the same downstream temporal nodes. For deletions we do not need any adjustment; whether it is a conscious strategy of the alliance leaders or an implicit reaction to their situation, deleting an edge involved in a frustrated triad simply removes the frustration without propagating it. For frustrated triads in overlapping out components (i.e., causal overdetermination) we have three choices: count it for both because either is sufficient (common to cumulative cases conditional on the initial agent), weight the contribution by the amount of inflow redundancy (like in [Bramson and Vandermarliere, 2015]), or count it for neither because neither *uniquely* contributes to the spread (like TKO in [Bramson and Vandermarliere, 2016]). Really, we can parameterize the weighting to run the spectrum from counting none or all of the redundant inflow, but conceptually there is a difference here compared to disease propagation: one edge flip frustrates a balanced triad, but a second one balances it again. Thus redundant paths of frustration propagation combine in structural balance in a way that is very different from most other systems in which it merely accumulates. Because of this a frustrated triad would not be doubly caused by frustration coming in from two paths, the two sources of frustration would instead combine to balance the triad. So we can ignore redundancy here and simply look at the number of strongly frustrated triads in the frustrated subgraph of the temporal out component of each trigger event.

As shown in Figure 7.3, each cascade event is triggered by a frustrated triad becoming balanced and pushing that frustration to neighboring triads. The duration of a cascade event is the maximal path length between the trigger and all the nodes in its temporal out component (i.e., the longest branch). The cumulative cases measure counts the number of distinct triads that become frustrated as a result of the trigger event, and only counts itself if it again becomes frustrated. Magnitude counts the distinct temporal nodes in the temporal out component that are "infected" by the trigger event.

You can see in Figure 7.4 the duration (length) and magnitude (color) of each cascade in our time period. If multiple triads trigger the exact same changes in

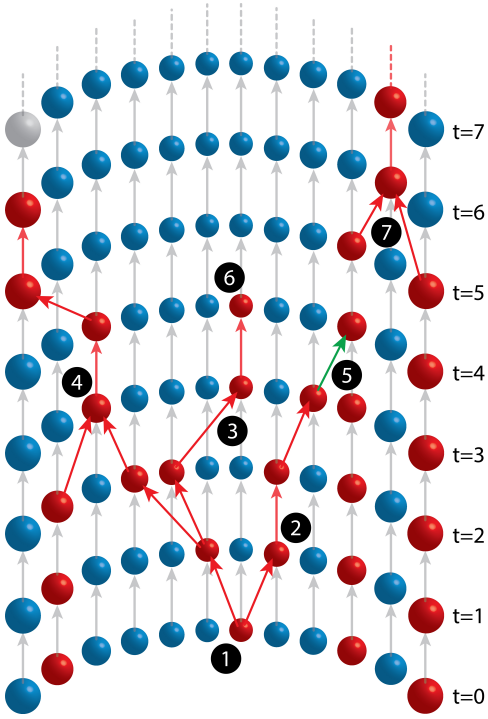


Figure 7.3: This diagram represents an idealized version of the triad network being analyzed to demonstrate key points in the analysis (non-active triad links and weakly frustration details are left off for clarity). (1) A trigger event occurs whenever a triad converts from strongly frustrated to either kind of balanced triad and that change is not part of a chain of events originating earlier. (2) The frustration may propagate to neighboring triads if they become strongly frustrated while the triggering nodes becomes balanced. (3) It is possible that the frustration will get pushed back onto the originating node. (4) In some cases multiple node balancing events may lead into a new frustrated triad; in such cases both are considered responsible and the propagation calculations are considered independently. (5) If the neighboring node is already frustrated then this does not count as spreading and path ends there. (6) In some cases a triad will balance without spreading frustration at all, which also marks the end of the path. (7) Multiple balancing events can initiate the same series of spread events, in which case they are all considered triggers. Focusing on the trigger event at (1), we can calculate the cumulative cases (8 distinct triads including itself due to reinfection), the longevity (6 time steps for the longest path), and the magnitude (12 frustrated temporal nodes in the frustrated subgraph of its temporal out component).

neighboring triads, then that is recorded as a single cascade. The blocks seen in the figure are the result of separate, but related, cascades in which initially different subsets of triads were affected by the initial trigger events (because of slightly different sets of neighboring triads) that typically converge on one chain of propagation causing them to end at the same time.

The duration of the infections is 45% correlated with the magnitude, and the cumulative cases measure is 91.9% correlated. The maximum value for magnitude equals cumulative cases times duration, and indeed that combination is 94.9% correlated with magnitude. In general, the correlation with cumulative cases diminishes as the amount of reinfection increases, thus indicating that in this dataset the reinfection rate is low. We already know that persistence of state is high for triads; however, with frustration lasting as long it does here, there would be ample time for reinfection and we do not see it.

Because the cumulative cases measure correlates much better with magnitude than duration, we can infer that the breadth of the infection makes up the greatest volume of the cascade size despite the apparent long durations. In this case the maximal path length is much longer than the average path length across all the branches. It could be that weighting the triads' frustration by their importance will reveal that the propagation pushes the frustration to corners of the systems where it doesn't interfere with players' activities (more below).

There are 439 days in the dataset, but because triggers and propagations require changes in triad and neighboring triad states, there are two fewer time steps for this analysis. For there to be several events with magnitudes reaching past 300,000 temporal nodes, many hundreds or thousands of distinct triads must be involved. Recall that the dense connectivity within a coalition of just a couple dozen alliances is already thousands of triads, and most edges are used by many triads. When one alliance in such a coalition changes its political position towards another, all of those triads become frustrated. Over a few days the other alliances align to either eject the defector or create a schism in the coalition. Some of the frustration may never be ameliorated, but most of it is. Earlier we sought the source of frustration in the system, and it does indeed seem that these events can inject frustration and spread it to other triads, many of which do not become balanced even a year later.

7.5.4 Conclusions on Structural Balance Propagation

We have presented a detailed analysis of the conditional behaviors of alliance standings changes based on the frustration of the triads formed according to an appropriately loose version of structural balance theory. Although only the temporal out component segment explicitly uses the temporal web formulation to measure influence properties of types of triads, the other behavioral measures provides necessary background understanding. Temporal influence abduction contrasts the

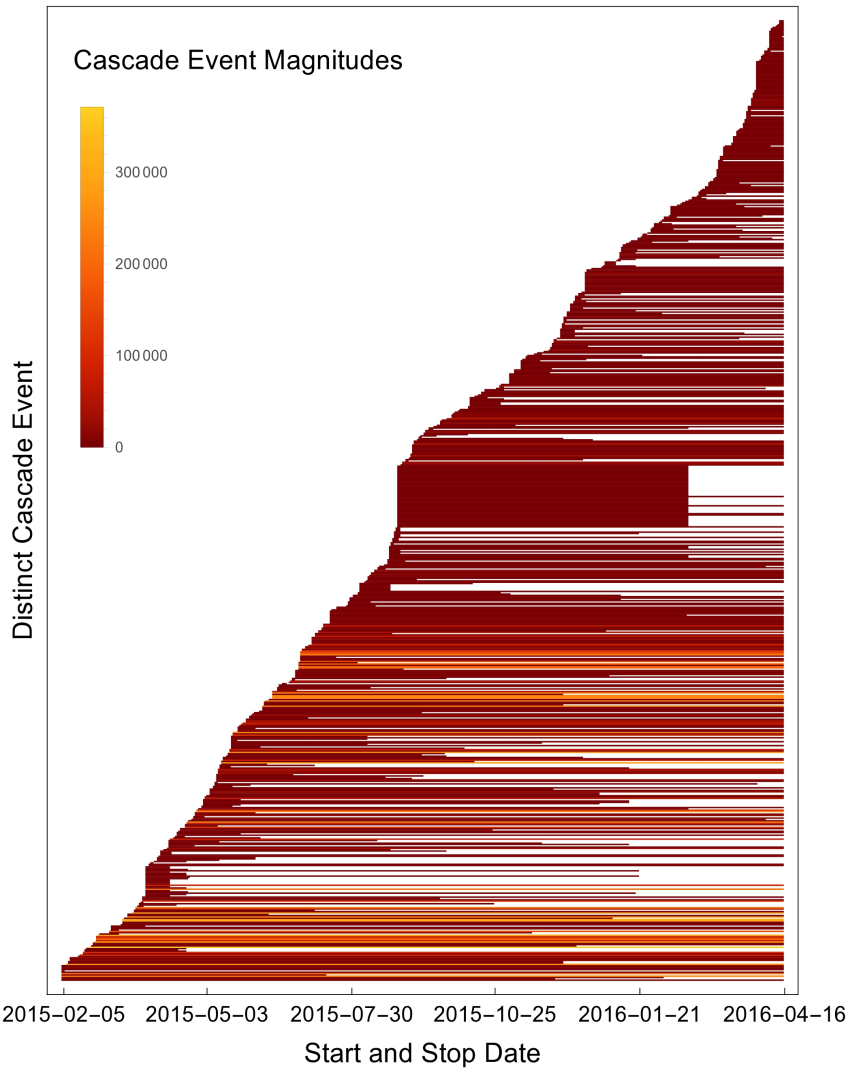


Figure 7.4: Each cascade event is represented by a horizontal line from its trigger day to its end day. Colors represent the eventual magnitude of that event. As mentioned earlier, there is a baseline level of frustration in the system and 98.7% of strongly frustrated triads stay that way; this can be seen from the long durations of the cascades even when the magnitude (and hence the number of cumulative cases) is low.

kinds of results derived from the temporally extruded representation from what can be done using aggregate behavior. By utilizing the contingent behavior proportions we can construct a generative model of triad changes and compare the temporal web results. Such tests are necessary to bridge the knowledge gap between the kinds of analyses we are accustomed to and those enabled by temporal networks.

Among the things we learned is that propagation in structural balance has two components: the first is mechanistic changes in triads that share an edge when that edge changes sign, and the second is how players respond to those changes. Because our empirical dataset came from a virtual online world, we do not have problems of dirty or missing data, but it is still necessary to separate signal from noise. In this analysis we defined our trigger events via an exogenous shock of frustrated triads becoming balanced without being an adaptation to the states of neighboring triads. We then traced the shock through neighbors of the trigger balancing themselves and frustrating their neighbors, and so forth through time and across the network. That is not the only possible trigger event, and not the kind of event that accounts for most frustration injection into the system. By changing the analysis to other triggers we will find different propagation patterns within the same dataset. Although the propagation results are clear, there are competing explanations that can also be explored using the technique demonstrated above.

For now we can tentatively conclude that frustration does in fact propagate through the network, and players do respond to frustration in the system in a way that is consistent with structural balance theory. However, our analysis also revealed that this adherence to SBT is likely to be implicit rather than through the conscious decisions of players to balance their standings. Furthermore, the short temporal path lengths and the patterns of contingent dynamics indicate that players more strongly work to avoid frustration than they do ameliorate it once it arrives. In a game based on conflict it is not strange to think that some frustration would persist, but the strong role of cooperation and coordination across alliances suffices to pressure players into forming stable and coherent political relationships.

7.5.5 Other Considerations for Future Work

As a further refinement we consider a measure of frustration weighted in several ways. As already mentioned, we are extending this analysis to the subset of sovereign alliances to determine whether they more strongly conform to the tenets of structural balance. We can do this as a separate category, and we can also do this by weighting alliances using various functions of the number of systems they hold sovereignty over. The reasoning is that the larger the alliance's territory, the more pressure it would feel from its players to have a consistent political policy. For similar reasons of importance, it is also natural to include all alliances and weight

them by their membership.

Another obvious weight is the distance between the alliances. The game universe of EVE is large, and in some cases the territories of aggressive alliances are far away from each other. In such cases there may be little push to adjust the standings because they have little effect on the actions of the players. The *effective frustration* among alliances captures this feature by down-weighting the frustration by the distance between the centroids of their sovereign territory. We are interested in the difference in the amount of frustration, the speed at which it is ameliorated, and any differences in the specifics of the dynamics. Through these and other weightings we can explore whether frustration propagates in such a way that it also minimizes effective frustrations, such as trading frustration with larger sovereign alliances with frustration with smaller landless alliances. That is, even if the number of frustrated triads increases, the total level of frustration may still be decreasing.

Signed social network analysis beyond structural balance has gaining increased interest [Doreian and Mrvar, 2009, Costantini and Perugini, 2014, Ciotti et al., 2015], but it is still not yet a well-developed field. Although we pursue some measures of signed temporal networks in other forthcoming work, here we focused on the propagation of frustration encoded in nodes representing triads, an approach that itself does not require specifically signed network measures. The richness of the EVE dataset will open up more opportunities to develop and refine augmented temporal network measures for multi-graphs as well because alliances have many different relations to each other beyond their political standings.

A major practical consideration here, and one shared by temporal network research more generally, is the time resolution of events. Most of the players who are in charge of large alliances log into the game nearly every day and update their standings quickly in response to events. In this analysis we do coarse-graining by using values at the time of daily server maintenance, but a perturbation may arise, propagate, and dissolve within a day and therefore not be detected. In future work we explore the use of a continuous time version that incorporates each standings change as it happens.

Finally, here we analyzed a temporal network of the political triads, and showed how frustration propagates across the political network via changes in shared edges. Another way to do the conversion, which we apply elsewhere, is to keep the alliances as the nodes and count the number/proportion of frustrated triads that each alliance is a member of. In this case, rather than tracking which relationships are influential in spreading frustration through the system, this alliance-centric version determines which agents are responsible for pushing frustration to other agents. In order to do this, however, one must have a measure of influence that can handle quantitative changes (rather than just categorical changes), which is reserved for future work.

7.6 Propagation of Risk Factors on Interbank Networks

Our second dataset will showcase the use of the temporal web technique in order to study the propagation of systemic risk in interbank loan networks. During the recent financial crisis, the interbank lending market proved to be one of the most important channels of financial risk contagion and hence of systemic risk. The malfunctioning of this overly interconnected market caused a liquidity drought across financial markets with consequences reverberating throughout the entire economy. Since then, research into interbank markets has proliferated. The dominant subject is to uncover the topology of interbank markets, to understand how they function, and how they could catalyze a systemic meltdown. There are many approaches to understanding the problem and recommending solutions, but the most promising ones are grounded in network science and/or agent-based modeling [Haldane, 2009, Buchanan, 2009, Georg, 2013, Hüser, 2015].

In this chapter we approach these matters by analyzing interbank lending as a temporal web. This technique will not only enable us to quantify the extent of cascades of bank troubles but it will also enable us to have a better understanding of the micro dynamics of the spreading process. We will use temporal influence abduction and out components to address questions like: Does lending to an unhealthy bank tend to negatively effect the lender? Do unhealthy banks only interact with other unhealthy banks or not? Are there lending patterns specific to at-risk banks? Are there lending practices that act as firewalls to contain a risk cascade? This section starts with a description of the dataset, then moves on to a discussion of the uncovered micro dynamics and ends with the insights garnered from the temporal web.

7.6.1 Description of Dataset

We use a dataset on the Russian interbank market provided by the private financial information agency Mobile for the period from August 1998 to October 2004. This dataset has two parts; the first part contains monthly bank balances for most Russian banks (described in detail in [Karas and Schoors, 2005]). From this first part we take two variables: total bank assets and capital. The second part of the dataset contains monthly reports "On Interbank Loans and Deposits" and represents a register of all interbank loans issued in the Russian market. For each loan we know its size, interest rate, issuer, receiver, reporting date, and maturity date.[1]

[1] In this chapter, we restrict ourselves to short-term loans, defined as loans with a maturity of less than a week. These account for more than 80% of the transactions both in terms of the number and of the volume. The reasons for this restriction is because the data provide information about the repayment date and not about the issuance date of the loans. This makes it hard to infer the exact duration of the connection between two banks for the long-term loans.

More details on the dataset can be found in [Karas and Schoors, 2010] and [Vandermarliere et al., 2015].

Because the balance sheet data is only reported monthly, we aggregate the loans to be on the same time scale. One way to encode this interbank lending scenario as a temporal web is that if bank A lent money to bank B in a given month t there is a directed link from bank B at month t to bank A at month $t+1$. The link expresses the fact that bank A has an exposure to bank B in month t and hence, for example, a deterioration in the state of bank B at month t might trigger a deterioration in the state of bank A in month $t+1$. For the time being we only work with unweighted links (instead of weighted by the amount of the loan) to keep the analysis simple.[1]

Figure 7.5 shows the time series of the number of banks active on the interbank market (solid green line) and the number of issued loans (dotted gray line). During the 1998-2002 period the interbank network experiences growth: we observe a steady and comparable increase in the number of active banks and of issued loans. Starting around 2002 the interbank network gradually matures as revealed by the number of active banks flattening out while the number of loans per bank continues to grow. Note, however, the strong variation in the number of issued loans from the second half of 2003 onwards.

As is indicated in Figure 7.5, our sample period includes two crises: one in August 1998 and one in the summer of 2004. Both crises resulted in a partial meltdown of the Russian interbank market. They coincide with the edges of the sample period and are clearly marked by a reduction in the number of active banks and issued loans. The first crisis got triggered on August 17, 1998 when Russia abandoned its exchange rate regime, defaulted on its domestic public debt and declared a moratorium on all private foreign liabilities. The second crisis was ignited by an investigation of banks accused of money laundering and sponsorship of terrorism. This gave rise to a wave of distrust among banks and a consequent liquidity drought.

7.6.2 Temporal Web Analysis Results

Before we are able to track the propagation of risk in interbank networks, we first need a measure of a bank's riskiness. A commonly used measure to express risk is the equity ratio of a bank: the total equity held by a bank divided by its total assets. The higher this ratio, the more a bank will be able to cushion shocks to its balance sheet. A continuous value-based analysis could look at any changes in the banks'

[1] Because the loans we are using have a maturity of less than one week, it may not be the case that A still has an exposure to B in month $t+1$ if the loan was initiated in t. However, we do not know the conditions of that repayment. The idea is that lending to a bank that is at risk of default does not immediately elevate the lender's risk; failing to be paid back or having a high-risk asset on the books however may elevate a lender's risk, and that shows up in the next balance sheet.

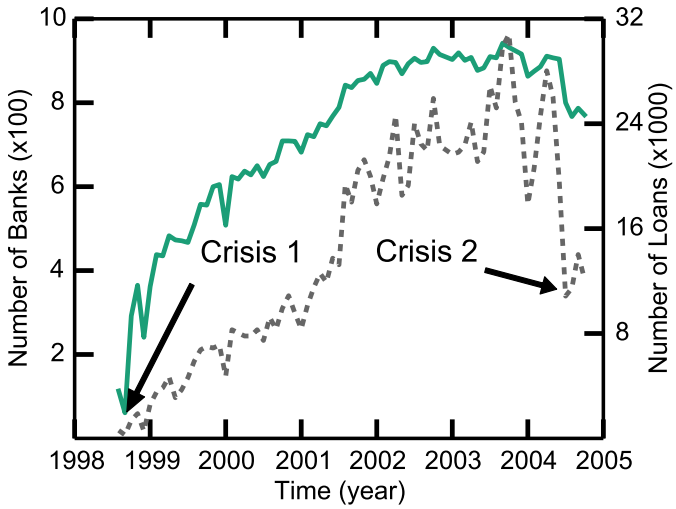


Figure 7.5: The time dependence of the number of active banks (solid green line) and the number of interbank loans (dashed line) in the Russian interbank network between 1998 and 2005. Data are aggregated over a month. The arrows indicate the start of the two "crises." (Figure reused from [Vandermarliere et al., 2015] with permission.)

equity ratios, but for the current analysis we apply a coarse-graining of the banks' equity ratios into the five categories defined in Table 7.5. During most of the period under analysis the Russian banking system had an equity ratio requirement of 7%. Banks above that are considered healthy, but many banks are well above that (up to 100%) and their behavior is distinct from merely healthy banks. The final column in Table 7.5 shows the proportion of banks that are in the corresponding equity ratio category across time. Zombie banks are banks with negative equity, meaning they are past bankrupt, but which still have a license to operate and hence still appear in the dataset. Although the percent of Zombie banks is small, their effect on system risk may not be. Dead banks are the ones that have closed (or not yet opened), and so have neither an equity ratio nor a temporal node.

Having categorized banks by our measure of financial riskiness, we can now look at the aggregate changes in their risk categories. Table 7.6 shows the proportions of month on month bank state transformations. We find a high level of consistency from period to period (76.3 to 95.2%), with Healthy+ banks as the most stable and Stressed banks the least. Unsurprisingly, Zombie banks have the highest proportions of closures (7.8%), but more than 1.4% of Stressed banks also close without going through the Zombie state first. Also note that a large number (but small proportion) of Healthy and Healthy+ banks also close directly without becoming Stressed or Zombie banks.

Bank State	Equity Ratio	Percent of Banks
Healthy+ (+)	$14\% \leq E$	0.743
Healthy (H)	$7 \leq E < 14\%$	0.191
Stressed (S)	$0 \leq E < 7\%$	0.0507
Zombie (Z)	$E < 0\%$	0.0149
Dead (D)	NA	NA

Table 7.5: For the current analysis we apply a coarse-graining of the data into these five categories by their equity ratio.

Bank state at time t	Bank state at time $t+1$				
	Healthy+	Healthy	Stressed	Zombie	Dead
Healthy+	0.952	0.0419	0.00222	0.000422	0.00376
Healthy	0.145	0.802	0.0482	0.00154	0.00387
Stressed	0.0465	0.155	0.763	0.0221	0.0130
Zombie	0.0273	0.0143	0.0463	0.834	0.0777

Table 7.6: Summary of results of the proportions of bank state changes from t to $t+1$ compiled across the entire dataset.

Other results gleaned from Table 7.6 include that somewhat surprisingly 8.8% of the banks that are Zombies at t actually recover into positive equity ratios the next month. The Stressed banks improve the following month in 20.2% of cases across time and only 3.5% become Zombies or close. Banks may stay Stressed for a while, but when they transition it is usually for the better. With the exception of Zombie banks becoming Healthy+, we can see that banks are much more likely to change incrementally through the adjacent categories rather than jump. These figures provide a foundation for understanding risky level changes in the system and the base transformation rates needed to create a simulation with similar aggregate behavior.

Temporal Influence Abduction. The next step in the analysis is to investigate whether the interbank lending network actually is a channel of risk propagation by looking at the conditional changes of lenders. Lending money exposes the lender to the situation of the borrower because a (sudden) drop in the health of the borrower might mean that the money will never be returned. This in turn influences the balance sheet of the lender and could result in a drop of the health of the lender too. When a bank lends to an unhealthy bank and then itself becomes less healthy, we can interpret this as the financial risk spreading from the borrower to the lender. Table 7.7 gives an exhaustive overview of the state changes of lenders from t to

$t+1$ conditional on the state of the borrower at t. The results are compiled across the entire dataset and reflect the micro dynamics of possible risk propagation.

Borrowing bank state at time t	Lending bank state changes from t to $t+1$									
	+→+	+→H	+→S	+→Z	+→D	H→+	H→H	H→S	H→Z	H→D
Healthy+	0.589	0.0352	0.000279	0.000010	0.000508	0.0334	0.276	0.0103	0.000034	0.000191
Healthy	0.504	0.0338	0.000287	0.000009	0.000479	0.0330	0.337	0.0125	0.000009	0.000218
Stressed	0.472	0.0323	0.000359	0	0.000431	0.0333	0.324	0.0153	0.000180	0.000359
Zombie	0.310	0.0464	0	0	0	0.0333	0.199	0.0232	0	0

Borrowing bank state at time t	Lending bank state changes from t to $t+1$									
	S→+	S→H	S→S	S→Z	S→D	Z→+	Z→H	Z→S	Z→Z	Z→D
Healthy+	0.00182	0.00950	0.0423	0.000034	0.000117	0.000015	0.000020	0.000064	0.000953	0
Healthy	0.00167	0.0123	0.0633	0.000070	0.000113	0.000052	0.000009	0.000148	0.00147	0
Stressed	0.00187	0.0157	0.0977	0.000647	0.000144	0.000144	0.000108	0.000324	0.00528	0
Zombie	0.00706	0.0151	0.270	0.0101	0	0	0	0.00907	0.0766	0

Table 7.7: Summary of results showing the proportions of lending bank state changes from t to $t+1$ (columns) conditional on borrower types (rows) compiled across the entire dataset. This shows the conditional effects of riskiness of the borrower on the stability of the lender. Note that the proportions are normalized across the rows, and the rows are split between the top and bottom subtables.

The boxed entries in Table 7.7 highlight the cases in which the state of the lending bank dropped after lending to a risky borrower. Comparing values down those columns it is clear that lending to a Stressed or Zombie bank has a comparatively higher rate of decreasing the risk status of lending banks, but comparing across rows shows that these are still low proportions. So this finding corroborates the idea that risk can propagate via the interbank lending network. However, we can also see that Stressed and Zombie banks recover proportionately more after borrowing from other Stressed and Zombie banks. This may have more to do with who is willing to lend to these banks than the effect of the loans on equity ratios.

To further test the risk spread hypothesis we can focus on those entries in which the banks' risk status got worse. These figures are summarized in Table 7.8. The first data column shows the proportions of lenders having their equity ratios drop at least one category (including death). Banks lending to Zombie banks are the most likely to see a reduction in stability. Some of those stability changes are from Healthy+ to Healthy, or from Healthy+ to Dead (although presumably they likely died for other reasons). So the second data column shows the proportions of lenders having their equity ratio dropped into the Stressed or Zombie categories, or from the Stressed to Zombie category. This also reveals a slightly larger proportion of banks becoming stressed contingent upon lending to Zombie. From these combined results we tentatively conclude that based on immediate neighbors, financial risk does spread to some degree across the interbank lending network.

Temporal Out Components. Temporal influence abduction indicates that

Borrower Type	Got worse	Below 7%
Healthy+	0.0467	0.0106
Healthy	0.0475	0.0129
Stressed	0.0497	0.0165
Zombie	0.0796	0.0333

Table 7.8: This table focuses on the type of borrowing bank for cases in which the lender becomes more risky. The first data column shows the proportions of lenders having their equity ratios drop at least one category (including death). Banks lending to Zombie banks are the most likely to see a reduction in stability. However also note that because there are so many more Healthy+ and Healthy banks than Stressed and Zombie banks, most decreases in equity ratios happen after borrowing from Healthy+ and Healthy banks. The effect revealed here is that, conditional upon lending to an at-risk bank, a lending bank is marginally more likely to become at risk itself.

financial risk can spread from a borrowing bank to a lending bank, its "neighbors" in the temporal web of the interbank market. In turn, these neighbors can also affect their neighbors in a cascade of risk propagation. Hence an initial bad loan of a bank can ripple through the bank system to increase overall system risk. In this section, our goal is to see whether the conditional neighbor changes seen above can indeed be chained together into risk-propagating cascades.

First it is important to understand the dynamics of bank lending, bank risk, and how they translate into a temporal web. Figure 7.6 shows an idealized scenario for a temporal web of an interbank loan network. In this case a trigger event is when a non-risky bank borrows from a risky (Stressed or Zombie) bank and then becomes risky itself. The infecting bank is not itself infected because it was already risky. The infected bank is also infectious as long as it stays in one of the risky states; any bank lending to it and becoming risky also becomes infected and infectious and added to the cascade originating from the trigger. A chain ends whenever an infected bank recovers into a healthy state or closes.

In this scenario the cumulative cases measure captures the number of unique banks involved in a cascade (excluding the trigger unless it becomes healthy and then reinfected by lending to a risky bank in the cascade that it started). In the widest chain there were 38 unique banks involved, but on average 3.59 banks are involved per cascade with most involving just one or two banks. The longest chain of infections lasted 40 months, but the average was 6.89 months (remembering that our data was in monthly time increments).

The cascade patterns for this temporal web are also distinct from the structural balance data in several key ways as can be seen in Figure 7.7. There are a few small blocks of related cascades, but most are distinct events with separate subsets of banks involved. The most striking thing about these results is how regular they

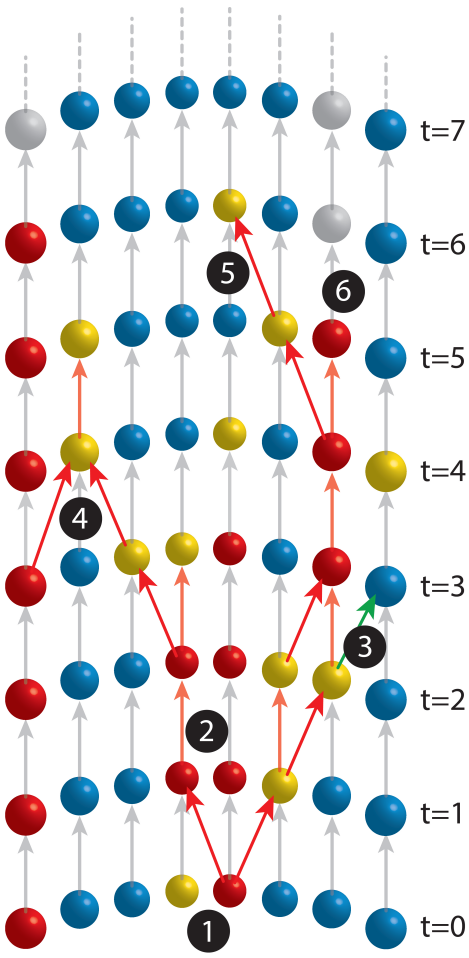


Figure 7.6: This diagram represents an idealized version of the interbank loan temporal web to demonstrate the events that trigger and propagate financial risk through the system. (1) A trigger event occurs when a risky bank (with equity ratio below 7%) borrows money from another bank and that bank's equity status becomes worse. (2) The risk injected by that trigger is not counted when maintained by the triggering borrower, but it accumulated through paths of contagion. (3) If a lending bank does not become riskier in the next period it is not counted as part of the cascade, even if it becomes riskier later without additional inputs. (4) A bank may lend to multiple risky banks and turn riskier; this counts in the magnitude of the contagion for each borrower. (5) The trigger node maintaining risk does not add to magnitude, but recovering and then becoming risky again does add to the magnitude if it occurs in the bank's own temporal out component. In this case the focal bank produces 6 cumulative cases (including itself through reinfection), a duration of 6 periods, and a magnitude of 14 temporal nodes.

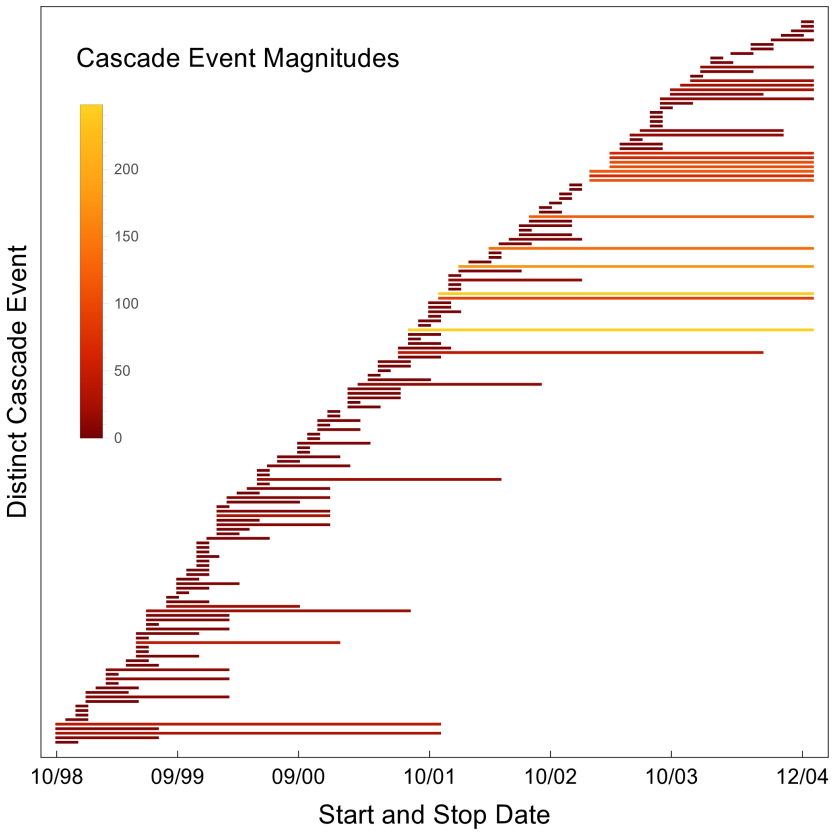


Figure 7.7: This figure shows each of the 160 cascades from their start date to their end date color coded by cascade magnitude. Unlike the EVE structural balance analysis nearly all the cascades finish within the data time frame because most stressed and Zombie banks either recover or die over several months.

are. There is no burstiness in cascade creation and long cascades occur throughout the period of analysis. Considering the number of banks and interbank loans in the dataset, the cascade results indicate that they are actually rare events.

Magnitude is the number of infected bank-months (temporal nodes) and so incorporates how long those banks stay in a risky state after infection (and also includes reinfection). For this dataset magnitude and cumulative cases are 96.4% correlated, and magnitude is 97.7% correlated with cumulative cases times max duration. Duration itself is 82.2% correlated with magnitude. These high correlations indicate that the more banks that are infected the more time some bank is infected because the set of infected banks is staggered. It also implies that the reinfection rate is low. There are a few cascades that last a long time (38 months)

involving only a few banks, but most larger cascades are ones that are both longer and involve more banks.

7.6.3 Conclusions on Propagation of Risk Factors in Interbank Networks

As many other studies have found before us (see [Hüser, 2015] for an excellent review), our temporal web analysis lets us tentatively conclude that interbank lending is a channel of systemic risk. Next to highlighting the micro dynamics, we were able to expose chains of risk propagating throughout the interbank system. We acknowledge the exploratory nature of this analysis and offer this as a demonstration of how to apply temporal webs to a field of increasing importance.

Across all cascades a total of 574 banks were involved with a total magnitude of 2646 temporal nodes. As we saw in Figure 7.7 most cascades were short lived and involved few banks, and only a few were large-scale events. In some cases individual banks become infected and stay stressed for more than a year, but the longest cascades involve multiple chains in which one bank passes off the infection to another bank that carries it forward. Perhaps counterintuitively, the largest chains originate in and last through the healthiest part of the bank system history. Although not many new cascades start during the second crisis, it is the case that several cascades last until that period, perhaps indicating that it is a build-up effect rather than a spontaneous reaction to a particular trigger.

Of course we are not claiming that the infecting loan was the only, or the primary, source of the increased bank risk among infected banks. Specifically we do not account for the size of the loans in this demonstration. By weighting the edges by the size of the loans we could generate a size-weighted magnitude score that more accurately reflects the spread of financial risk. Such an analysis is a natural pairing to a move to a continuous values of bank risk instead of categories. Furthermore the equity ratio is not the only measure of riskiness, so that is another avenue of expansion. Even with these limitations of the current analysis, there is a clear propagation pattern in the interbank loan network. Risk does seem to spread from risky banks to healthy banks through bad loans at least some of the time. The harm is usually quickly ameliorated, but in some cases (probably in concordance with exogenous factors) the bad loan does in fact trigger a cascade of reduced financial stability. These cascades are, we argue, best captured and measured with temporal webs because they make explicit the chains of events that constitute such cascades.

7.7 Propagation of Emotional Affect on Twitter Networks

Micro-blogging is an easy way for people to share information and opinions about a variety of subjects. Twitter is currently the largest micro-blogging platform on which millions of people voice their opinions and share information daily using messages (tweets) within the 140 character limit. The content of these tweets creates a massive stream of data with possible uses in subjects such as public opinion [Bollen et al., 2011, Pagolu et al., 2016, Bermingham and Smeaton, 2011, Jahanbakhsh and Moon, 2014, Lampos et al., 2013], information dissemination [Serrano and Iglesias, 2016, Riquelme and Gonzalez-Cantergiani, 2016, Zhao et al., 2013, De Domenico et al., 2013], and even event tracking [Lampos et al., 2010, Guille and Favre, 2015, Kumar et al., 2014].

To utilize these messages to measure opinion or intention it is essential to know the sentiment of the tweets; this is the first step in transforming tweets into useful data. We will not go into detail about the many techniques that exist for this sentiment extraction as this is not our focus, but instead refer to the large body of work on automated sentiment analysis [Liu, 2015, Saif et al., 2013]. Most of these techniques use some sort of natural language recognition, either by the use of dictionaries or trained neural networks, to map words, word combinations, and symbols onto emotions.

The role of sentiment on the probability of online material going viral has gathered much attention in the last years, especially for market research purposes. Various social media platforms play an important role in this spreading of content and every platform has its own characteristics that are used to different ends. By analyzing the sharing of articles in ”The New York Times” Berger and Milkman found that content virality is positively correlated to its positivity and emotions linked to activation of arousal (awe, anxiety and anger) and negatively correlated to sadness [Berger and Milkman, 2012]. However, their dataset might contain only a limited type of content and not be representative of all platforms, especially considering these results seem to be in contradiction with the folk theory that negative news sells.

Hansen et al. looked into this conflict in [Hansen et al., 2011] using retweets in twitter, which is part social media and part news and information dissemination. They found that there was a difference between news content and non-news content. News content's virality was connected to negative sentiment content, while for non-news content positive sentiments supports virality. Heimbach extended these results and found that the connection between positivity and virality was non-linear and thus interacts in a more complicated way as first suggested by classic theories [Heimbach and Hinz, 2016].

Previous analyses have only looked at the conditional behavior of content

based on sentiment to make generalization about the effects of sentiment on spreadability. This is similar to aspects of temporal influence abduction, but by using temporal webs, we can follow and study its longer-term propagation through the system and find the most important spreaders (and bottlenecks) of positive (and negative) sentiment. Sentiment is of importance in various subjects because it is an indicator of intention towards acting [Bermingham and Smeaton, 2011, Lampos, 2012, Salathé and Khandelwal, 2011, Salathé et al., 2013] – people that have a positive sentiment towards vaccination or a presidential candidate will more likely get vaccinated or vote for this candidate. Understanding the dynamics behind sentiment propagation can thus contribute to the precision of predictions about future behavior.

To uncover features of the propagation of a property like sentiment through a network, it is necessary to include chains through time. This is especially so in the case of social networks wherein the time-order of communication plays an essential role in social contagion. There has been some work on Twitter data using dynamical network techniques and temporal analysis [Taxidou and Fischer, 2014, Moro, 2012, Kwak et al., 2010], but to our knowledge there has not yet been explicitly temporal network analyses research conducted on it (perhaps for reasons of computational limitations). Below we present a temporal web analysis of one Twitter dataset that includes the network of followers, the tweets sent, and the encoded sentiment of those tweets.

7.7.1 Description of the Dataset

The dataset used here was shared with us by Prof. Marcel Salathé who used it in [Salathé and Khandelwal, 2011, Salathé et al., 2013]. The dataset contains 411,720 tweets between August 2009 and January 2010 from users based in the United states which contain keywords pertaining to vaccination.[1] The 4,793,160 follower-followee connections between the 101,852 authors of these tweets were also gathered. By combining these sources we have a social network in which the nodes are people tweeting about vaccines and the directed edges are tweets that go from the author to all his or her followers.

In his papers Salathé et al. investigated the characteristics and dynamics of health behavior sentiments towards the then novel influenza A(H1N1) vaccine (the time window of the data coincides with the fall wave of the swine flu pandemic). To obtain the sentiment of the collected tweets they utilized an ensemble method combining a naive Bayes and a maximum entropy classifier. The training set was generated by students that all together rated 88,237 tweets and 47,143

[1] This includes retweets in both the old and the current form. The old form of retweet was just copy-pasting the other person's text and manually adding RT. The current form of retweets was rolled out for everyone on Nov 19, 2009. The keywords included are: vaccine, vaccinated, vaccinate, vaccinating, immunized, immunize, immunization, and immunizing.

unique tweets. They were presented with four options for every tweet: "positive", "negative", "neutral", and "irrelevant". Irrelevant tweets were those that were not about the influenza vaccine and the others reflect the emotional valence of the sentiment. They report an accuracy of 84.29% for this ensemble classifier, and while the quality of the sentiment detection effects our results we simply take their results as given to demonstrate the temporal web approach on this kind of dataset. We refer to [Salathé and Khandelwal, 2011] for the details of their sentiment encoding methodology.

It is important to mention that the potential contact network is taken to be static. If at any moment within the dataset user B is a follower of user A, there is a corresponding edge between them in the constructed social network. Changes in followers and followees are thus not taken into account. Furthermore, we do not currently have a measure of the rate of change in followers to judge how much this feature of the data may be affecting our results. The concern is that our method expects the sentiment of tweets by A to effect the sentiment of tweets of A's followers. If we find that this is not true for some B, we don't know whether that is because B is unaffected or because B wasn't actually a follower for much of the dataset. While acknowledging this weakness in the current dataset we can still assess the qualitative tendency of sender sentiment to effect followers (even if the specific values of our propagation measures are off by some unknown quantity).

In what follows we present an analysis of the data using temporal influence abduction to uncover the sentiment of tweets conditional on the tweets they see from the people they follow. Then, using temporal webs we investigate the propagation of this sentiment within the network of vaccine-related tweets. With this technique, we are able to find longer patterns in the dynamics of sentiment propagation; we can move from "Does sentiment affects immediate virality?" to "Is it really infectious?".

7.7.2 Constructing the temporal web

Our first move is to cull the data by only including tweets related to the topic of vaccines and thus exclude the tweets tagged as "irrelevant". Because we are leaving a continuous time version of the temporal web technique for future work, we have to choose a coarse-graining of the data to construct a temporal web from the data. To do this we aggregate the tweets by day so that each time step represents all persons active in the network and their tweets within a certain day.

Although there are some alternatives to analyzing the spread of sentiment in this dataset, our goal here is to determine whether a person who is exposed to a lot of negative tweets increases that person's negativity, and whether exposure to positive messages increases the positivity of tweets made. To do this we need to establish the mood of a person based on their tweets. There are three sentiments

that can be expressed in a tweet: "positive", "negative", and "neutral". The mood of a person is therefore uniquely represented by two numbers:

$$\text{Mood Positivity} = \frac{\text{number of positive tweets sent}}{\text{total number of tweets sent}} \tag{7.1}$$

$$\text{Mood Negativity} = \frac{\text{number of negative tweets sent}}{\text{total number of tweets sent}} \tag{7.2}$$

One could calculate the ratio of these two to get a single index, but that would hide the proportion of tweets with sentiment compared to all tweets.[1] Furthermore, for the purpose of comparing whether negativity or positivity is more contagious, the separate scores are more useful.

Using these equations we calculate a mood for each person at each time step based on their activity. If we look at time step t and a person has not tweeted up to that point, that person is coded as not having a mood (a null value that is different from being neutral). A person who does not tweet in time step t, but has tweeted in a previous time step, is assigned the mood from the previous time step. We call this *temporal inheritance* and reflects a (sometimes explicit, sometimes implicit) connection from an agent to its immediate future self. This convention is common in disease and epistemic temporal networks to reflect that the agent's state stays the same unless acted upon, but is less clearly appropriate for moods of people on Twitter.

The next step is to construct the temporal web from the flow of information in the system. We use the follower network and connect each followee at time t to all of his/her followers at time $t+1$. The edges are cross-temporal to reflect the effect of the tweets on the calculated moods of the receivers even though the tweets themselves may have been seen immediately. To answer the question of sentiment spread, we calculate the positivity and negativity of a person's feed, which is based on the collection of all tweets a person is exposed to from all the people followed. This aggregation again leads to two numbers that characterize the mood of a person's feed:

$$\text{Feed Positivity} = \frac{\text{number of positive tweets exposed to}}{\text{total number of tweets exposed to}} \tag{7.3}$$

$$\text{Feed Negativity} = \frac{\text{number of negative exposed to}}{\text{total number of tweets exposed to}} \tag{7.4}$$

We generated histograms of the distributions of each person's mood and (rather surprisingly) both the positivity and negativity are distributed almost entirely into

[1] If both positivity and negativity have 5 tweets out of 10, then the sentiment ratio is 1:1. If they both have 5 out of 1000 then again the ratio is 1:1. However, in the former case they both have medium sentiment levels and in the later case they both have low sentiment levels. In other work we provide a combined sentiment measurement, but the added complication is beyond the scope of this work.

three narrow spikes: one near zero, one near the center, and one at the extreme. We repeated this for the feed sentiment values and found a high concentration on and near zero (though more dispersed than for mood), a narrow spike near the center, and a narrow spike on the extreme. Thus both mood and feed can be cleanly broken down into distinct categories according to the scheme in Table 7.9.

Category	Range
low	$0 \leq m \leq \frac{1}{3}$
medium	$\frac{1}{3} < m < \frac{2}{3}$
high	$\frac{2}{3} \leq m \leq 1$

Table 7.9: Categorization schemes for both mood and feed positivity and negativity based on the proportion of tweets having emotional affect for that day.

In Table 7.10 we show the distribution of the mood and feed negativity and positivity among the three categories of Table 7.9. We thus see that the spikes near zero make the low category by far the largest of the three. That feed negativity has a higher value for the medium category and lower value for the high category compared to mood negativity indicates that most people, even if they have a highly negative mood, do not only follow other people with a highly negative mood, but will also follow people that either send out purely informative tweets (low negativity) or that are more positive on average.

Category	Mood negativity	Feed negativity	Mood positivity	Feed positivity
low	0.903	0.933	0.853	0.955
medium	0.0106	0.0324	0.0101	0.0201
high	0.0868	0.0347	0.137	0.0251

Table 7.10: The distributions of both moods and feeds for each category separated into positives and negatives. Most tweets lack sentiment, but there is a notably higher level of highly positive tweet writers despite a lack of highly positive tweet readers. For each group the value for the medium category is the lowest, but especially for the moods.

7.7.3 Temporal Web Analysis Results

The first step is to look at the transitions in mood from time step t to time step $t+1$ by type, shown in Table 7.11. Here we separately analyze the positive and negative moods since there might be an asymmetry in the propagation. We see that there is high stability in moods, with the low level being the most stable. We also see that for both positive and negative moods there is nearly an order of magnitude

difference between the probability of going from one extreme to the other than going to an adjacent (medium) level. For medium we see it is much more likely to transition to low than it is to high.

All States

a) Negative Sentiment

Mood at time t	Mood at $t+1$ low	medium	high
low	0.999	0.000208	0.000674
medium	0.0198	0.976	0.00373
high	0.00511	0.000281	0.995

b) Positive Sentiment

Mood at time t	Mood at $t+1$ low	medium	high
low	0.999	0.000236	0.000823
medium	0.0191	0.978	0.00294
high	0.00283	0.000108	0.997

Conditional On Transition

c) Negative Sentiment

Mood at time t	Mood at $t+1$ low	medium	high
low		0.236	0.764
medium	0.842		0.158
high	0.948	0.0521	

d) Positive Sentiment

Mood at time t	Mood at $t+1$ low	medium	high
low		0.223	0.777
medium	0.869		0.134
high	0.963	0.0368	

Table 7.11: Summary of results of the proportions of mood changes from t to $t+1$ compiled across the entire dataset. Tables (a) and (b) show all states for positivity and negativity respectively. Although all three levels are extremely stable, medium levels are the least stable and low levels are the most stable. Subtables (c) and (d) renormalize the data conditional on there being a change to highlight what those changes are.

The conditional state changes in Table 7.11 highlight both the similarity of positive and negative mood changes and the pattern of changing directly between high and low instead of through medium. Low sentiment is three times more likely to transition to high than medium levels, and high sentiment is 24 times (positive sentiment) or 15 times (negative sentiment) more likely to transition into low than medium levels. Mediums levels are roughly six times more likely to drop than to increase. Despite these lower levels, the medium level sentiment holders could play a key role in propagation if they are typically low level sentiment people with temporary spikes of higher sentimentality. To discern whether or not this is the case, we need to look at the conditional changes in sentiment level.

Temporal Influence Abduction. To start looking at the effect of the exposed positivity and negativity on the mood of a person, we present the mood change proportions per feed sentiment level they get exposed to in Table 7.12. As before, we see that most people stay in the low sentiment category (regardless of the feed level) and the next largest group stays in the high category. This raises the idea that there are two dominant types of Twitter users: broadcasters of information and opinion pushers. Their behavior is changed little by the inputs they receive.

The feed sentiment level does not have a large influence on these trends on an aggregate level but threads of propagation might be buried in the noise.

a) Negative Sentiment

Feed negativity at time t	Change in mood from t to $t+1$								
	l→l	l→m	l→h	m→l	m→m	m→h	h→l	h→m	h→h
low	0.870	0.000181	0.000575	0.000188	0.00929	0.0000351	0.000600	0.0000324	0.119
medium	0.840	0.000180	0.00125	0.000361	0.0142	0.0000790	0.00143	0.0000677	0.142
high	0.829	0.000127	0.00101	0.000216	0.0125	0.0000636	0.000903	0.000127	0.156

b) Positive Sentiment

Feed positivity at time t	Change in mood from t to $t+1$								
	l→l	l→m	l→h	m→l	m→m	m→h	h→l	h→m	h→h
low	0.793	0.000189	0.000649	0.000180	0.00913	0.0000268	0.000553	0.0000208	0.196
medium	0.849	0.0000968	0.00126	0.000221	0.0116	0.0000692	0.000775	0.0000692	0.137
high	0.792	0.0000873	0.000661	0.0000624	0.0111	0.0000749	0.000574	0.0000250	0.195

Table 7.12: Summary of results of the proportions of mood changes from t to $t+1$ compiled across the entire dataset contingent on the mood of the feed. The extreme levels and the non-changing states again dominate. A majority of people stay in the low level, followed by the high level independent of the feed mood.

By looking deeper at the contingent behaviors we can uncover a few other details that hint at the presence of influence propagation. For example, for the positive posts, people with a medium level are much more likely to switch to a high sentiment level given a high feed sentiment level compared to other feed levels. In the negative sentiment case, people with a low sentiment mood are twice as likely to switch to a high level after receiving a high-level feed than they are to stay low or go to medium. Negative medium level people are also marginally most likely to switch to a high level given a high feed level. Because most people stay in the same category, these small marginal differences may indicate the influence of the feed on one's sentiment level, but it is difficult to perform meaningful statistical tests to check for the significance of these small differences due to the structure of the data across people, their networks, and time.

Temporal Out Component Analysis. As before, to really see whether a property spreads across the network, one must look at the network across time and chain together the changes of interest. Given the large size of the network (101,852 nodes and 4,793,160 edges at each daily time step) the small conditional change probabilities might still constitute cascades of influence across the temporal web. The first step is to refine our definition of what it means for a person to be influenced by another in this context.

We use the following criteria to identify sentiment propagation. If person B is a follower of person A, and person A has an upwards change in sentiment from $t-1$ to t, then we say there is propagation along the edge from A to B if person

B shows a change in sentiment in the same direction as A from t to $t+1$. One can think of the temporal web as a transformation of the data into a structure in which the temporal nodes are changes in sentiment levels with a property value of 'increased', 'static', or 'decreased'. Temporal node α at layer L connects to temporal node β at layer $L+1$ if and only if β follows α and α tweets something on both times $t-1$ and t. If A's proportion of sentiment-having tweets is higher at t than $t-1$, then the α node at L has the property 'increased' and is considered infectious. If person B sends a larger number of emotional tweets in $t+1$ compared to t then the temporal node β has the 'increased' property at layer $L+1$ and we say that propagation occurs from A to B between t and $t+1$.

Unlike the previous two analyses, in the Twitter case there is a separation between being infectious and being infected. Nodes are only infectious on the time steps after an increase in their proportion of sentiment-laden tweets. Agents only become infected when they perceive such an increase, and then increase their sentiment level themselves in response. At that point they are both infected and infectious. If they remain in the elevated state then they remain *infected*, but because their sentiment level is constant they are *not infectious*. Because we have three levels of sentiment, it is possible for a person to go from low to medium, and then from medium to high. Both increases will make them infectious, but after reaching the high state there is nowhere to go except to stay high or drop. Once the level decreases from the elevated state the person is considered recovered and is no longer infected. Recall that states are inherited across times without tweets, this allows us to identify changes from past moods to judge increases and decreases, but their mood can only change if they actively tweet something that day.

We can perform separate analysis for the subset of nodes that are infectious and the ones that are infected. This distinction produces some interesting insights into the flow of sentiment on the Twitter temporal web, but for now we abbreviate these analysis to present only the core results. The main difference is that when all infected nodes are included, the magnitude and duration last as long as somebody stays in a heightened state, and this sometimes lasts until the end of the dataset. By contrast, when only the infectious nodes are included there is a strict limit on the length of time a particular cascade can dwell in a node (one or two days). The difference is just one of bookkeeping and measurement; whether to perform an analysis of the infectiousness of sentiment or of the social impact of sentiment spread.

Here we will focus on whether sentiment propagates on the Twitter network and, if it does, whether positivity or negativity propagates better? We found 539 distinct cascades in the positive sentiment temporal web and 653 distinct cascades in the negative sentiment temporal web. As you can see in Figure 7.8 most of these cascades were small, infecting fewer than a dozen people. In the case of the negativity spread (right scatter plot) there is a large conspicuous series of cascades of

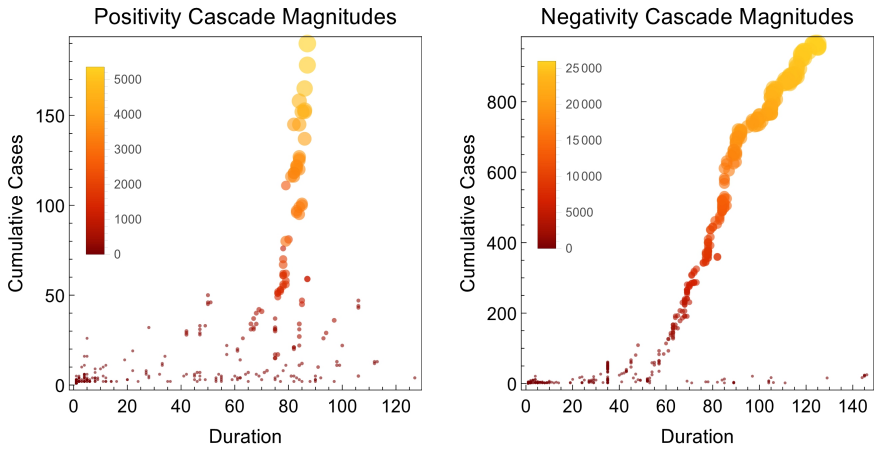


Figure 7.8: Scatter plots for propagation of positivity (left) and negativity (right) in which each disk represents a distinct cascade. The disk's location indicates its duration and cumulative cases, while the color and size indicate magnitude. Note the difference in scaling, especially the y-axis.

increasing magnitude, cumulative cases, and duration. Starting on September 15, 2009 there is a single chain of sentiment spread that on each day infects between one and dozens of people and lasts until November 29, 2009. Although there are three slightly different sets of initially infected agents that form three different triggers (hence distinct cascades), the paths leading from those events quickly converge. As the chain progresses through time there are dozens of other trigger events along the way that converge into this one main stream. It is like many small tributaries feeding a main river. So although each independent trigger counts as a distinct cascade, nearly all the nodes involved in the cascades are shared.

Looking at the timelines for the infectious cascades this is even more conspicuous because all those cascades start at different times but end on the same day. If we count the overlap differently, therefore, we would get very different quantitative results for magnitude and cumulative cases. We revisit this point below. The convergence of chains of propagation is not uncommon, in fact it happens to a lesser scale four other times in the negativity cascades and there are five small cases of convergence in the positivity cascades, but we didn't expect to see any chains this long or broad from this dataset. To discover such a long chain, and one that has so many independent triggers, means that not only does large scale propagation of sentiment occur on the Twitter networks, large scale propagation is inevitable (at least on this subset of Twitter).

Now we address the question of whether positivity or negativity is more viral. The scatter plots in Figure 7.8 share some features, but also have some clear differences. The summary statistics of the positivity and negativity cascades presented

in Table 7.13 demonstrate clearly that negativity is more successful in spreading in this Twitter data. The cascades of negativity are more frequent, larger, and longer lasting. This is in line with what Salathé et al. found in their papers, namely that negative sentiment spreads much more easily than positive. However, the bulk of the magnitude of negativity is contained in that one huge chain, whereas the chains in positivity are more diverse. If we change the way we account for out component overlap it could drastically effect this result; we revisit this issue in the next section.

	Positivity Cascades	Negativity Cascades
Number of cascade events	539	653
Mean Duration	25.2	46.6
Max Duration	127	146
Mean Cumulative Cases	14.5	225.7
Max Cumulative Cases	190	965
Mean Magnitude	340.3	5620.1
Max Magnitude	5355	25993

Table 7.13: Comparison of the cascade features of positive and negative sentiment propagation. A Kolmogorov-Smirnov two-sample test reveals that these distributions are different with an extremely high significance, which is already clear from these summary statistics.

Although infectiousness has a very limited time frame, specifically the period after a mood change, the magnitude is calculated throughout the infected period. If people undergo long-term changes in their posting sentiment propensities then the durations and magnitudes will be greater, and magnitude will correlate less with cumulative cases. What we find is that for positivity, magnitude and cumulative cases are 98.2% correlated, while in negativity they are 99.9% correlated. Durations are 59.6% and 82.3% correlated with magnitude for positivity and negativity respectively. This means that spreading to new people dominates the magnitude measure and changes are short-lived. It also means that cumulative cases times duration is actually slightly less correlated with magnitude than cumulative cases alone. This finding is consistent with the earlier claim that people are generally either spreaders of information or of opinion; they may change their style based on a particular issue, but quickly return to their base rate.

7.7.4 Conclusions on Propagation of Sentiment on a Twitter Network

Here we showed that propagation of the type similar to sentiment propagation, in which person A changes his/her mind from time $t-1$ to t, causing person B to change his/her mind from time t to $t+1$ (which is a different type of propagation than the last two types), can easily be tracked and quantified within the presented temporal web framework without losing information in aggregation. Although the aggregated change proportions showed little evidence of spread, the temporal web analysis shows that cascades of change actually play a large role in who changes sentiment and when.

The particular chain patterns found in this dataset are quite interesting and unexpected. The single long chain in the negativity propagation brings to the forefront questions about how to count overlapping out components. As mentioned above, in this work we are treating them as distinct cascades, multiply counting each infected temporal node for each independent event that could have infected them. This measures, for each trigger, its potential to spread sentiment; i.e., what sentiment it would have spread even if all other cascade events were eliminated. However, this is a poor measure of the marginal effect on sentiment spread because most of infections are infecting the already infected. To capture marginal effect one needs to identify each cascade's unique influence by identifying temporal nodes that would not have been infected without that event. The intermediate approach of weighting each path from each trigger event node potentially best captures the contribution of each temporal node in each event, but the appropriate weighting may depend on the application and the distribution of events.

The reason this is important ties back to the principal use of temporal webs to measure influence. The next step in such an analysis is to compare the score each trigger node achieves according to various (temporal and static) network measures to the magnitude it generates. Alternate ways of measuring impact on the temporal web will produce different comparisons and highlight different network features. This then feeds back into the role that capturing temporally extruded cascades plays in identifying useful network measures of influence.

Moving forward with the Twitter analysis, a natural addition is to combine the two sides of the sentiment to also analyze the interplay between positive and negative sentiment. For example, here we determined whether having a more positive feed leads to more positive tweets, and we can extend this to include having fewer negative tweets as well. The current dataset on the avian flu is rather limited in its scope and has features that we believe are particular to this topic. By applying this technique to a variety of Twitter subgraphs on a variety of unrelated topics we can make broader claims about general features of propagation on this social network platform.

7.8 Summary

This chapter provided a general description of measuring influence on dynamic social networks using a temporal web structure of the interactions across time. We also proposed a way to derive the rules that people use via a technique called temporal influence abduction. This approach uses aggregated contingent behaviors to uncover the kinds of information one would need to build a simulation that recreates data like the single observed one. In future work we will use these contingent behavior rules to generate multiple alternate counterfactual histories for each dataset and compare the propagation patterns of the real and simulated temporal web. One motivation to do this is to determine the role of luck (i.e., a convergence of probabilistically rare events) in the appearance of large-scale cascades and/or to determine to what degree some amount of propagation is inevitable in dynamic networked systems.

For each of the datasets analyzed here cumulative cases (and cumulative cases times duration) were very highly correlated with magnitude. This is not generally true based on previous work with disease simulations, and so reflects a substantive feature of spread on these networks. The improved accuracy of the magnitude measure is one of the proposed benefits of the temporal web approach over running dynamics on static networks or analyzing flattened networks. In the three cases presented above, this benefit is limited because the easier and more traditional measures act as reasonable proxies for the temporal web measure. This is generally the case when either cumulative cases or the duration is the main driving factor of total spread size, and less so when both factors play a significant role. There are many techniques useful for studying propagation, studies like this help us home in on the features of network contagion for which the temporal web approach provides unique insights.

Although work on this method is still fresh, it is gaining momentum. Demonstrations of how to use temporal networks (beyond just temporal webs) and the benefits of such applications are critical to them gaining a wider acceptance. Our sincere hope is that this detailed analysis of three applications of temporal webs to empirical data provides such an example. These datasets are already large, and because each agent at each period is encoded as a temporal node, analyzing them requires especially efficient algorithms to run. Although this may act as a barrier to entry for some researchers, we are eager to collaborate with interested parties in extending our applications of temporal webs.

References

[Abell, 1968] Abell, P. (1968). Structural balance in dynamic structures. *Sociology*, 2(3):333–352.

[Antal et al., 2006] Antal, T., Krapivsky, P. L., and Redner, S. (2006). Social balance on networks: The dynamics of friendship and enmity. *Physica D: Nonlinear Phenomena*, 224(1):130–136.

[Axelrod and Bennett, 1993] Axelrod, R. and Bennett, D. S. (1993). Landscape theory of aggregation. *British Journal of Political Science*, 23(02):211–233.

[Berger and Milkman, 2012] Berger, J. and Milkman, K. L. (2012). What makes online content viral? *Journal of Marketing Research*, 49(2):192–205.

[Bermingham and Smeaton, 2011] Bermingham, A. and Smeaton, A. F. (2011). On using twitter to monitor political sentiment and predict election results. In *Sentiment Analysis where AI meets Psychology (SAAIP) Workshop at the International Joint Conference for Natural Language Processing (IJCNLP)*.

[Bollen et al., 2011] Bollen, J., Mao, H., and Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1):1–8.

[Braha and Bar-Yam, 2006] Braha, D. and Bar-Yam, Y. (2006). From centrality to temporary fame: Dynamic centrality in complex networks. *Complexity*, 12(2):59–63.

[Bramson and Vandermarliere, 2015] Bramson, A. and Vandermarliere, B. (2015). Dynamical properties of interaction data. *Journal of Complex Networks*, 4(1):87–114.

[Bramson and Vandermarliere, 2016] Bramson, A. and Vandermarliere, B. (2016). Benchmarking measures of network influence. *Scientific Reports*, 6:34052.

[Buchanan, 2009] Buchanan, M. (2009). Meltdown modelling. *Nature (London)*, 460(7256):680–682.

[Cartwright and Harary, 1956] Cartwright, D. and Harary, F. (1956). Structural balance: A generalization of heider's theory. *Psychological Review*, 63(5):277–293.

[Chen et al., 2012] Chen, D., Lü, L., Shang, M.-S., Zhang, Y.-C., and Zhou, T. (2012). Identifying influential nodes in complex networks. *Physica A: Statistical mechanics and its applications*, 391(4):1777–1787.

[Ciotti et al., 2015] Ciotti, V., Bianconi, G., Capocci, A., Colaiori, F., and Panzarasa, P. (2015). Degree correlations in signed social networks. *Physica A*, 422:25–39.

[Colizza et al., 2007] Colizza, V., Barrat, A., Barthelemy, M., Valleron, A.-J., and Vespignani, A. (2007). Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. *PLoS Medicine*, 4(1):e13.

[Costantini and Perugini, 2014] Costantini, G. and Perugini, M. (2014). Generalization of clustering coefficients to signed correlation networks. *PLoS ONE*, 9(2):e88669.

[Cui et al., 2013] Cui, J., Zhang, Y.-Q., and Li, X. (2013). On the clustering coefficients of temporal networks and epidemic dynamics. In *2013 IEEE International Symposium on Circuits and Systems (ISCAS2013)*, pages 2299–2302. IEEE.

[Davis, 1967] Davis, J. A. (1967). Clustering and structural balance in graphs. *Human Relations*, 20:181–187.

[De Domenico et al., 2013] De Domenico, M., Lima, A., Mougel, P., and Musolesi, M. (2013). The anatomy of a scientific rumor. *Scientific Reports*, 3:2980.

[Dekker, 2013] Dekker, A. H. (2013). Network centrality and super-spreaders in infectious disease epidemiology. In *20th International Congress on Modelling and Simulation (MODSIM2013)*.

[Doreian and Mrvar, 2009] Doreian, P. and Mrvar, A. (2009). Partitioning signed social networks. *Social Networks*, 31(1):1–11.

[DuBois et al., 2011] DuBois, T., Golbeck, J., and Srinivasan, A. (2011). Predicting trust and distrust in social networks. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom)*, pages 418–424.

[Facchetti et al., 2011] Facchetti, G., Iacono, G., and Altafini, C. (2011). Computing global structural balance in large-scale signed social networks. *PNAS*, 108(52):20953–20958.

[Georg, 2013] Georg, C.-P. (2013). The effect of the interbank network structure on contagion and common shocks. *Journal of Banking & Finance*, 37(7):2216–2228.

[Grindrod and Higham, 2013] Grindrod, P. and Higham, D. J. (2013). A matrix iteration for dynamic network summaries. *SIAM Review*, 55(1):118–128.

[Guille and Favre, 2015] Guille, A. and Favre, C. (2015). Event detection, tracking, and visualization in twitter: a mention-anomaly-based approach. *Social Network Analysis and Mining*, 5(1):1–18.

[Haldane, 2009] Haldane, A. (2009). Rethinking the financial network. Speech delivered at the Financial Student Association, Amsterdam.

[Hansen et al., 2011] Hansen, L. K., Arvidsson, A., Nielsen, F. A., Colleoni, E., and Etter, M. (2011). Good friends, bad news. affect and virality in twitter. In *Future Information Technology. Communications in Computer and Information Science*, volume 185, pages 34–43. Springer.

[Harary, 1959] Harary, F. (1959). On the measurement of structural balance. *Behavioral Science*, 4(4):306–323.

[Heider, 1946] Heider, F. (1946). Attitudes and cognitive organization. *Journal of Psychology*, 21:107–122.

[Heimbach and Hinz, 2016] Heimbach, I. and Hinz, O. (2016). The impact of content sentiment and emotionality on content virality. *International Journal of Research in Marketing*, 33(3):695–701.

[Holme, 2015] Holme, P. (2015). Modern temporal network theory: a colloquium. *The European Physical Journal B*, 88(9):1–30.

[Holme and Saramäki, 2012] Holme, P. and Saramäki, J. (2012). Temporal networks. *Physics Reports*, 519(3):97–125.

[Hummon and Doreian, 2003] Hummon, N. P. and Doreian, P. (2003). Some dynamics of social balance processes: Bringing heider back into balance theory. *Social Networks*, 25(1):17–49.

[Hüser, 2015] Hüser, A.-C. (2015). Too interconnected to fail: A survey of the interbank networks literature. Technical report.

[Jahanbakhsh and Moon, 2014] Jahanbakhsh, K. and Moon, Y. (2014). The predictive power of social media: On the predictability of u.s. presidential elections using twitter. In *arXiv preprint arXiv:1407.0622*.

[Karas and Schoors, 2005] Karas, A. and Schoors, K. (2005). Heracles or sisyphus? finding, cleaning and reconstructing a database of russian banks. Working paper 327, Ugent.

[Karas and Schoors, 2010] Karas, A. and Schoors, K. (2010). A guide to russian banks data. *SSRN: http://ssrn.com/paper-1658468*.

[Kempe et al., 2005] Kempe, D., Kleinberg, J., and Tardos, É. (2005). Influential nodes in a diffusion model for social networks. In *Automata, languages and programming*, pages 1127–1138. Springer.

[Kim and Anderson, 2012] Kim, H. and Anderson, R. (2012). Temporal node centrality in complex networks. *Physical Review E*, 85(2):026107.

[Kimura et al., 2010] Kimura, M., Saito, K., Nakano, R., and Motoda, H. (2010). Extracting influential nodes on a social network for information diffusion. *Data Mining and Knowledge Discovery*, 20(1):70–97.

[Kitsak et al., 2010] Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E., and Makse, H. A. (2010). Identification of influential spreaders in complex networks. *Nature Physics*, 6:888–893.

[Kumar et al., 2014] Kumar, S., Liu, H., Mehta, S., and Subramaniam, L. V. (2014). From tweets to events: Exploring a scalable solution for twitter streams. *arXiv preprint arXiv:1405.1392*.

[Kwak et al., 2010] Kwak, H., Lee, C., Park, H., and Moon, S. (2010). What is twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web*, pages 591–600. ACM.

[Lampos, 2012] Lampos, V. (2012). On voting intentions inference from twitter content: a case study on uk 2010 general election. *Computing Research Repository (CoRR) arXiv:1204.0423*.

[Lampos et al., 2010] Lampos, V., De Bie, T., and Cristianini, N. (2010). Flu detector - tracking epidemics on twitter. In *ECML PKDD*, pages 599–602. Springer.

[Lampos et al., 2013] Lampos, V., Lansdall-Welfare, T., Araya, R., and Cristianini, N. (2013). Analysing mood patterns in the united kingdom through twitter content. *Computing Research Repository (CoRR) arXiv:1304.5507*.

[Lawyer, 2015] Lawyer, G. (2015). Understanding the influence of all nodes in a network. *Scientific reports*, 5:8665.

[Lerman et al., 2010] Lerman, K., Ghosh, R., and Kang, J. H. (2010). Centrality metric for dynamic networks. In *Proceedings of the Eighth Workshop on Mining and Learning with Graphs*, pages 70–77. ACM.

[Leskovec et al., 2010] Leskovec, J., Huttenlocher, D., and Kleinberg, J. (2010). Signed networks and in social and media. In *CHI 2010: Machine Learning and Web Interactions April 10-15, 2010, Atlanta, GA, USA*.

[Liu, 2015] Liu, B. (2015). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press.

[Lü et al., 2011] Lü, L., Zhang, Y.-C., Yeung, C. H., and Zhou, T. (2011). Leaders in social networks, the delicious case. *PLoS ONE*, 6(6):e21202.

[Malliaros et al., 2016] Malliaros, F. D., Rossi, M.-E. G., and Vazirgiannis, M. (2016). Locating influential nodes in complex networks. *Scientific Reports*, 6:19307.

[Mantzaris and Higham, 2013] Mantzaris, A. V. and Higham, D. J. (2013). *Temporal Networks*, chapter Dynamic communicability predicts infectiousness, pages 283–294. Springer.

[Masuda and Holme, 2017] Masuda, N. and Holme, P. (2017). *Temporal network epidemiology*. Springer.

[Moro, 2012] Moro, E. (2012). Temporal network of information diffusion in twitter.

[Nicosia et al., 2013] Nicosia, V., Tang, J., Mascolo, C., Musolesi, M., Russo, G., and Latora, V. (2013). *Temporal Networks*, chapter Graph Metrics for Temporal Networks, pages 15–40. Springer.

[Pagolu et al., 2016] Pagolu, V. S., Challa, K. N. R., Panda, G., and Majhi, B. (2016). Sentiment analysis and of twitter and data for and predicting stock and market movements. In *International conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*.

[Pastor-Satorras and Vespignani, 2002] Pastor-Satorras, R. and Vespignani, A. (2002). Immunization of complex networks. *Physical Review E*, 65:036104.

[Pfitzner et al., 2013] Pfitzner, R., Scholtes, I., Garas, A., Tessone, C. J., and Schweitzerk, F. (2013). Betweenness preference: Quantifying correlations in the topological dynamics of temporal networks. *Physical Review Letters*, 110:198701.

[Riquelme and Gonzalez-Cantergiani, 2016] Riquelme, F. and Gonzalez-Cantergiani, P. (2016). Measuring user influence on twitter: A survey. *Information Processing & Management*, 52(5):949–975.

[Rocha and Blondel, 2013] Rocha, L. E. and Blondel, V. D. (2013). Flow motifs reveal limitations of the static framework to represent human interactions. *Physical Review E*, 87(4):042814.

[Rocha and Masuda, 2014] Rocha, L. E. and Masuda, N. (2014). Random walk centrality for temporal networks. *New Journal of Physics*, 16(6):063023.

[Saif et al., 2013] Saif, H., Fernández, M., He, Y., and Alani, H. (2013). Evaluation datasets for twitter sentiment analysis: a survey and a new dataset, the sts-gold. In *1st Interantional Workshop on Emotion and Sentiment in Social and Expressive Media: Approaches and Perspectives from AI (ESSEM 2013)*.

[Salathé and Khandelwal, 2011] Salathé, M. and Khandelwal, S. (2011). Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control. *PLoS Computational Biology*, 7(10):e1002199.

[Salathé et al., 2013] Salathé, M., Vu, D. Q., Khandelwal, S., and Hunter, D. R. (2013). The dynamics of health behavior sentiments on a large online social network. *EPJ Data Science*, 2(1):1–12.

[Serrano and Iglesias, 2016] Serrano, E. and Iglesias, C. A. (2016). Validating viral marketing strategies in twitter via agent-based social simulation. *Expert Systems with Applications*, 50:140–150.

[Sikic et al., 2013] Sikic, M., Lancic, A., A., Antulov-Fantulin, N., and Stefancic, H. (2013). Epidemic centrality – is there an underestimated epidemic impact of network peripheral nodes? *The European Physical Journal B*, 86(10):1–13.

[Szell et al., 2010] Szell, M., Lambiotte, R., and Thurner, S. (2010). Multirelational organization of large-scale social networks in an online world. *PNAS*, 107(31):13636–13641.

[Taxidou and Fischer, 2014] Taxidou, I. and Fischer, P. M. (2014). Online analysis of information diffusion in twitter. In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14 Companion, pages 1313–1318, New York, NY, USA. ACM.

[Vandermarliere et al., 2015] Vandermarliere, B., Karas, A., Ryckebusch, J., and Schoors, K. (2015). Beyond the power law: Uncovering stylized facts in interbank networks. *Physica A*, 428:443–457.

[Viard and Latapy, 2014] Viard, J. and Latapy, M. (2014). Identifying roles in an ip network with temporal and structural density. In *2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 801–806. IEEE.

[Wehmuth et al., 2015] Wehmuth, K., Ziviani, A., and Fleury, E. (2015). A unifying model for representing time-varying graphs. In *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pages 1–10. IEEE.

[Xu and Wang, 2017] Xu, S. and Wang, P. (2017). Identifying important nodes by adaptive leaderrank. *Physica A*, 469:654–664.

[Yu et al., 2010] Yu, Y., Berger-Wolf, T. Y., Saia, J., et al. (2010). Finding spread blockers in dynamic networks. In *Advances in Social Network Mining and Analysis*, pages 55–76. Springer.

[Zhao et al., 2013] Zhao, L., Cui, H., Qiu, X., Wang, X., and Wang, J. (2013). Sir rumor spreading model in the new media age. *Physica A*, 392(4):995–1003.

Chapter 8

Diplomatic Relations in a Virtual World

8.0 Preface

The following paper is accepted for publication, with minor revisions, in the A1 journal *Political Analysis*. It fits in this dissertation as an in-depth investigation of the research possibilities of virtual worlds for testing economic and social theories. We illustrate this by testing variations and extensions of Structural Balance Theory, using the highly detailed data of alliance relationships in the virtual world Eve Online.

My contributions to this paper were providing the data on Eve Online alliance relationships, expanding the analysis software for tracking propagation through temporal webs (see chapter 6) with extra functionality supporting weighted Social Balance Theory and generating benchmark data for comparing networks, writing the community detection software revealing emergent and dynamically changing coalitions in the Eve Online alliance data, writing the section in the paper on the triadic network transformation, and the empirical analysis of the Eve Online alliance relationships data.

8.1 Abstract

We apply variations and extensions of structural balance theory to analyze the dynamics of geopolitical relations using data from the virtual world *Eve Online*. The highly detailed data enables us to study the interplay of alliance size, power, and geographic proximity on the prevalence and conditional behavior of triads built from empirical political alliances. Through our analysis we reveal the degree to which the behaviors of players conform to the predictions of structural balance

theory and whether our augmentations of the theory improve these predictions. In addition to studying the time series of the proportions of triad types, we investigate the conditional changes in triad types and the formation of polarized political coalitions. We find that player behavior largely conforms to the predictions of a multipolar version of structural balance theory that separates strong and weak configurations of balanced and frustrated triads. The high degree of explanatory power of structural balance theory in this context provides strong support for both the theory and the use of virtual worlds in social science research.

8.2 Introduction

In order to better understand the dynamics of real world political interactions we analyse the characteristics and dynamics that naturally emerge within a virtual world political environment. We first describe an expanded version of structural balance theory that we use as a guiding principle governing the dynamics of political relations. We then evaluate the degree to which the virtual world dynamics cohere to the predictions of this theory in multiple ways. We further describe how differences between the virtual and real world affect the translation of social theories between virtual and real world environments.

Data regarding the stability and dynamics of real-world political relations are scarce because the number of real-world countries is small and because political relations tend to be relatively stable over the span of history for which reliable data are available. This paucity of data on the dynamics of real world political relations makes it neigh impossible to test general theoretical claims about the drivers of change in international relations. Datasets such as the Correlates of War [CorrelatesOfWar, 2019] and the conflict in the middle east [Economist, 2015], for example, are too sparse to provide strong support for general sociopolitical theories. Furthermore, changes in external factors, such as technology and economic activity, alter the context in which political changes occur, thus further complicating or even invalidating the analysis of real world political relations across across time and space. We address these problems by considering an alternative body of empirical data on political relations that does not suffer from these limitations. In particular we employ data on political relations from a virtual world called *Eve Online* to study the dynamics of political relations [CCP, 2019].

The core of our analytical approach derives from Structural Balance Theory [Heider, 1946, Cartwright and Harary, 1956, Harary, 1959]. The principle of balance theory is that interpersonal relations have valences – some people you like and some you dislike – and we can capture those as positive and negative links in a social network. Structural balance theory (sometimes "social balance theory" or simply "balance theory") provides a characterization of when such a signed social network is balanced – and if not, then how frustrated it is. The implication

is that frustrated relations are more likely to change than balanced ones, and so the theory implicitly also provides a theory about the dynamics of signed social networks. There are already many variations of this methodology, and we further extend it to explore patterns of frustration dynamics in political networks as they evolve.

The community structure of the *Eve Online* alliance standings network reveals multiple mutually competing coalitions. By treating this virtual world as a multipolar system engaged in structural balance dynamics, we can explain many patterns observed in the data. Deviations from the predictions of structural balance theory (like persistent strong frustration) reveal the limitations of applying the theory to a conflict-driven game context [Belaza et al., 2017, Belaza et al., 2019]. Still, the high degree of explanatory power of structural balance theory in this context provides strong support for both the theory and the use of virtual worlds in social science research.

8.3 Data: *Eve Online* Virtual World

Appreciating the limited data available on real-world political and/or trade networks and large-scale conflicts, we explore the appropriateness of data collected from a massively multiplayer online game called *Eve Online* (EVE). The advantages of using data from an immersive, sandbox computer game [Squire, 2008] are: (1) there is more data, (2) the data are reliable, (3) events in a computer game occur more frequently than in the real world, and (4) most of the activity is already quantified in a consistent way. One concern about data from a computer game is that it may lack sophistication because the game world is necessarily a simplified version of the real world. Along these lines, one may worry that the behaviors of players reflect constraints and interventions from the developers, rather than natural diplomatic actions. We desire data that is simpler than real-world data in some respects, yet still complicated in the right ways to be useful.

Few games provide the sociopolitical structure, variety of activities, and realistic motivations necessary to be candidates for serving this role. *Eve Online* does seem to provide all the right ingredients to generate realistic (and yet constrained) dynamics in large hierarchical political structures. Players of EVE engage in activities such as mining, harvesting, research, industry, trading, couriering, protection, piracy, and politics in an open-ended sandbox environment. Players are free to join corporations, which can form into alliances, which are in turn part of implicit coalitions in a multi-tiered hierarchy of political arrangements. Alliance membership determines friends and foes, which regulates conflicts on a scale ranging from single-player economic sabotage to prolonged territorial war involving thousands of players.

Because we wish to use this data to understand why and how certain political,

economic, and social activities occur, it is important to ensure that the underlying forces driving them – and not just aggregate patterns – match up with corresponding real-world motivations. The political actions we examine are generated by human players reacting to each other, and to an environment with fixed and known rules. The game offers players a wide range of economic activities [Mildenberger, 2013]. Because these activities are connected in a complex interplay of dependencies, players inside the game are motivated in similar ways as those real-world actors whose behavior we wish to study. Players invest time and real-world money to acquire virtual resources in the game world. These resources are then used to enhance the players' capabilities, influence and/or status. When players lose these resources in conflict or through mistake, the time, effort and money they used to acquire them are permanently lost. This feature creates genuine scarcity and risk-aversion that fuels realistic economic behaviors [BBCNews, 2014, Goh, 2018, Hoefman et al., 2019].

The constraints imposed by a virtual world are also a blessing. Although the scale of the game is large, with hundreds of thousands of players (in fact larger than some countries), it is still manageable. Alliances set sales taxes for their stations and income taxes for their members, but there is just a single rate for each, rather than an encyclopedic tax code. Political standings between pairs of alliances are captured by a single value between -10.0 and $+10.0$, rather than a complicated mixture of treaties, international laws, and historical conventions. All the necessary ingredients for complex socio-politico-economic dynamics exist, but they are all simplified. A virtual world is like a computer simulation, except with humans instead of algorithms as the driving force. What we get is a "frictionless pulley" type of environment in which to test social and economic theories without the noise, confounding factors and size limitations that plague real world datasets.

8.3.1 Further Details of *Eve Online*

There are between 400,000 and 500,000 accounts created for playing *Eve Online*. On average 33,000 accounts are signed on simultaneously at any given time during our analysis period [Eve-Offline, 2018]. The number of participating players fluctuates, and many advanced players have multiple characters and/or accounts: the total number of registered characters exceeds 10 million. EVE's core player base is extremely dedicated: there are multiple annual game-related conferences, some drawing thousands of attendees from all over the world.

Corporations. All characters belong to a corporation, some of which are default non-player character (NPC) corporations run by the game itself, and others which are owned and run by players. Players can create and close corporations, so the number of corporations and their memberships change over time; however, there

are consistently around 380,000 distinct corporations. Existing corporations have between 1 and roughly 12,000 members – many of the successful, long-lived corporations have memberships in the thousands [EveWho, 2017]. The leader of a corporation has complete power to determine corporate policies and represent the corporation, although the duties can also be shared among a body of directors and the leader can be replaced through a shareholder vote [UniWiki, 2017].

Alliances. Corporations can unite to form alliances.[1] The largest alliance has more than 28,000 player members combined from more than 500 corporations. Most successful and long-lived alliances have thousands of player members and dozens of corporations [EveWho, 2017]. The player who is elected to run the alliance has the authority to decide which other alliances, corporations and/or individual players are enemies, and which are allies. Standings towards other alliances, corporations and individuals are set on a scale from -10.0 (sworn enemy) to +10.0 (close ally); or they can remain unset. A standings value of 0.0, indicating a neutral disposition, is distinct from having standings unset. Alliance standings are directed, and need not be reciprocal (see section 8.4.1). Within the game, ships belonging to an enemy alliance show up as red, indicating to players that they are potential targets, and that they are free to engage in aggression themselves. Friendly alliance ships appear blue, and ships in alliances with unset standings appear white (which matches the various other neutral objects like stations and asteroids). Alliance members who attack friendly players (colloquially called "blues") are typically punished by their own alliance, by paying a fine or being kicked out. In some cases, these actions can trigger a large-scale conflict.

Geography. The game universe consists of 7930 solar systems, located in three-dimensional space and connected by a network of (mostly) short-range transportation channels, connected to fixed-location gates within each system. Although some expensive ships are capable of ignoring the connections network by directly jumping to systems within a certain range, most travel is done by jumping to neighboring systems through the gate network. Travel within a solar system may be just from gate to gate, but potential destinations include various stations, planets, moons, asteroid belts, landmarks, and mission areas. The setup is similar to a network of islands, well-connected to nearby islands via bridges. As a consequence, travel distance is best captured by network distance, but Euclidean distance is still appropriate for the physical separation of regions on the map.

Security and Sovereignty. Of the 7930 solar systems, 5201 are normally navigable. These are further divided into three categories by their security level: null

[1] Corporations can remain independent, and some do, but the sizes of non-alliance corporations are roughly one-fifth the size across a ranked comparison.

security (3294), low security (817), and high security (1090). In addition, there are another 230 unreachable systems, used as a company testbed. The remaining 2499 systems are reachable using special equipment/skills (so-called "Wormhole space"). Of the 3294 null security systems, 2712 are conquerable by players. In these conquerable systems, alliances can hold sovereignty (ownership) over a system's stations and resources.

Sovereignty allows an alliance to extract resources from the system, for example through mining operations or taxation, as well as to exert control over access to its stations. Many alliances do not hold sovereignty over any systems; for example, because they operate solely within high-security space or in abnormal wormhole space. Although these alliances also set standings to one another, we may assume these standings convey less information about the mutual political relations, since the parties involved can neither engage in territorial conflict with one another, nor block others from accessing their stations. We therefore perform separate analyses for (1) all alliances with more than 200 members and (2) alliances holding sovereignty over at least one system at some time during the sample period (all these alliances turn out to have more than 200 members).

Coalitions. Alliances sometimes further coalesce into so-called coalitions: sets of alliances that cooperate for purposes of mutual protection and coordinated attack. Coalitions are emergent social superstructures. Their existence was neither planned, nor foreseen by the game designers, and there is no official designation or in-game support for these structures. Coalition members typically maintain strong positive relations to one another, and tend to share the same set of enemies. Despite there being no official record of coalition membership in the game, nor any direct effect of coalition membership to the game mechanics, the existence of these social superstructures can be confirmed through player chat logs and online forums. In section 8.6 we show that the structure of the coalitions mirror the polarized groups predicted by structural balance theory – with some noteworthy differences. The mere existence of emergent self-organized coalitions of political alliances in the game lends credence to the applicability of balance theory to this virtual world and on the usefulness of virtual worlds to further study balance theory.

8.4 Methods: Structural Balance Theory

As stated earlier, the core of our analytical approach is based on structural balance theory [Heider, 1946, Cartwright and Harary, 1956, Harary, 1959]. The technical aspects of balance theory have evolved over time, and we further develop the method in this paper. Originally, researchers focused on all cycles in the social network, and classified networks as balanced if all cycles had an even number of negative links – otherwise they were unbalanced. [Hart, 1974] includes two mea-

sures from [Harary et al., 1965]: β, the proportion of semicycles that are positive; and λ, the minimal number of edge removals or flips necessary to balance the network. Starting with [Abell, 1968] it has become standard practice to examine only triads (aka: triangles, connected triplets, 3-cycles) in the network (although also see [Facchetti et al., 2011]). Typically, the triads are considered balanced (or stable) when there are zero or two negative edges, and unbalanced (or unstable) if there are either one or three negative edges. However, it can be more insightful to look at all four types of triads separately (see Figure 8.1). Rather than classifying a network as balanced or unbalanced, we are interested in the proportions of each type of triad.

From a static perspective, we can determine whether the proportion of frustrated triads is smaller or larger than expected compared to random signed networks. However, since balance theory is really about network dynamics, it has become common to measure aggregate frustration over time. This approach has been applied to both simulations [Hummon and Doreian, 2003, Antal et al., 2006] and empirical networks [Leskovec et al., 2010, Szell et al., 2010, DuBois et al., 2011]. Balance theory implies that not only should the frustrated triads constitute a small proportion of the total network, but also their prevalence in the system should exhibit a decreasing trend (unless some event injects new information into the system). Although the theory operates at the level of individuals triads, because triads share edges, the change of a single edge can set in motion a series of subsequent edge changes, as the frustration cascades throughout the network. [Bramson et al., 2017b].

The game offers a great deal of anecdotal evidence suggesting that players face situations and choices matching the premises of social balance. When an alliance sets a positive or negative standing to another alliance, this has a direct effect on the behaviors of everybody in both alliances (often many thousands of players), as well as indirect effects on other alliances. One such example is described in an anecdote from the history of the BRUCE alliance:

> When BRUCE first arrived in Y4Y7, a constellation in southern Syndicate at the invitation of COE, they were forced to also become friends with Anarchy Empire – a rather unpleasant group of pirates that COE had befriended. BRUCE held a closer philosophy to that of the OSS at the time and became blueboxed [i.e. had positive standing] to both OSS and the COE/AE pair. This was a strange step as the OSS faction and COE/AE faction were hostile to each other, and led almost immediately to problems in the area.
>
> This very early diplomatic challenge for BRUCE was, in their own words, handled poorly. BRUCE ended up dumping the OSS bluebox less than a week after gaining it because they, understandably, did not want to fire on COE whom invited them to Syndicate. De-

> spite BRUCE's loathing of Anarchy Empire's philosophies they valued their word to COE more. [Soon after]...war erupted between BRUCE/COE and the OSS. [EVE-history.net, 2011].

Such a story strongly indicates that theories for the dynamics of international relations in general, and structural balance theory specifically, have gainful application to alliance relations in EVE.

8.4.1 Making the Edges Symmetric

Although early versions of balance theory were based on directed networks, and [Hart, 1974] proposed a measure of frustration as the proportion of balanced 2- and 3-semicycles, nearly all work on structural balance theory uses undirected networks. In EVE, alliance leaders set alliance standings toward other alliances, so standings are directed edges. Although not completely symmetric, they are nearly so with an average matched sign reciprocity of 84% (see Supplementary Materials for a plot of reciprocity over time).[1] Near symmetry has been observed in other signed networks as well; for example, [Facchetti et al., 2011] reports that the directed edges in Epinions, Slashdot, and WikiElections signed networks are nearly symmetric. The reason for symmetry is clear in the game's context: if X is an ally of Y, while Y is an enemy of X, then players in Y may get penalized if they return fire when players in X attack them. Such reasoning/motivation would be present in any aggression-modulating political standings.

During the timeframe of our analysis, alliances set standings using values between -10 and 10 in increments of 5. In the game, the standing value between a player's alliance and that of others' ships/assets are displayed as colors: +10/very friendly = dark blue, +5/friendly = light blue, 0/neutral = white, -5/hostile = orange, and -10/very hostile = red. Previously, any value between -10 and 10 was allowed, but what matters to most players is whether the standing is positive, negative, zero, or unset. While other, specific values could be used as a signaling device among alliance leadership, they were rarely used in practice and so the system was simplified.

Although we incorporate several weighting schemes for triads in Section 8.5.1 we do not incorporate the standing weights (i.e., 5 vs 10) in the current analysis. To determine the valence for our undirected network, we set the links to be negative if either direction is negative/neutral, and positive if both directions are positive, or one is positive and the other is unset.[2]

[1] One coalition called Provi-Bloc has a nonstandard policy of setting standings known as "Not Red, Don't Shoot" (NRDS). Rather than leaving most standings unset, Provi-Bloc sets enemy standing to anybody who isn't explicitly an ally. This difference in interpretation leads to a situation where a number of mutual relations appear asymmetric in the dataset but are actually symmetric in practice. These can be seen as red bands of negative standings in Figure 8.9.

[2] Neutral standings (value 0) are considered negative because in EVE, any player who isn't flagged

8.4.2 Strong and Weak Frustration

As mentioned above, we examine the proportions of all four types of triads, rather than classifying them as balanced or unbalanced. We consider the triple negative triads as being weakly frustrated (essentially neither frustrated nor balanced) compared to strongly frustrated single negative triads [Davis, 1967, Szell et al., 2010]. We can also consider triple positive triads as being more strongly balanced than the weaker "mutual enemy" single positive triads. The reasoning here is that if a strongly balanced triad changes via a single edge flip, then it would necessarily turn into a strongly frustrated one. However, a weakly balanced triad may become strongly or weakly frustrated, the latter of which exerts less pressure on the stability of the system of standings. Figure 8.1 shows a breakdown of all four types of triads according to this looser version of balance theory.

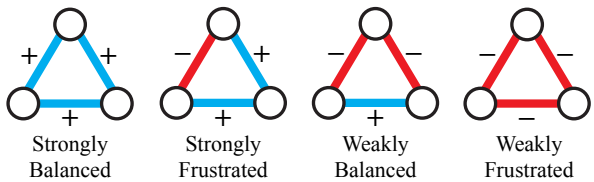


Figure 8.1: The four distinct triad types of balance theory and their looser categorization into weaker and stronger levels of frustration.

Compared to the tenets of classic balance theory, one specific deviation we expect to see is that triple negative triads will be over-represented. Or, more precisely, that the original structural balance theory overstates the amount of frustration because triple negative triads do not actually represent an unstable situation. Essentially, classic balance theory implicitly assumes that there are only two real teams, and the bulk of early work focused on this "bipolar world" [Waltz, 1964, Waltz, 1979]. In this bipolar case, one pair of nodes in the triple negative triad would find it preferable to team up against a mutual enemy, thus converting into a weakly balanced triad; the so-called common enemy effect. We see temporary cooperative behavior like this in the World Wars [Antal et al., 2006], and this dynamic certainly does occur in EVE. But EVE also exhibits long-standing, three-way mutual antagonism among independent political entities [Davis, 1967, Axelrod and Bennett, 1993]. Separating strongly and weakly frustrated triads allows one to capture the structure and dynamics of multi-polar political systems in a more refined way.

as friendly is likely to be attacked. This policy, employed by most players in the game, is called NBSI, for "Not Blue, Shoot It". Under NBSI, players with unset standings are equally likely to be met with hostility, and setting the standing to 0 makes it clear that such hostilities are permitted, without explicitly declaring the target an enemy.

8.4.3 Triadic Network Transformation

Our analysis below first looks at the levels of frustration through daily time-slices of the network structure of positive and negative links; however, we delve deeper to better understand triad *dynamics* and contingent behaviors. Although solutions for studying network dynamics such as RSIENA [Snijders, 2017] already exist, they rely on the dyadic structure of networks. Because we are interested in the dynamics of the triadic superstructure that emerges from the structural balance interpretation of the underlying network data, we transform the network of alliance relationships into a triadic network. This transformation is illustrated in Figure 8.2, and explained in more detail in [Bramson et al., 2017b] where a temporal network version of this novel triadic network representation is used to identify frustration cascades.

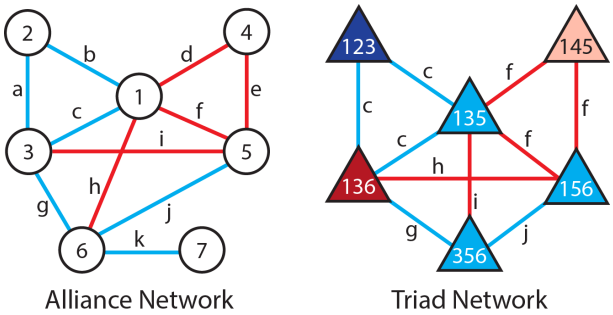


Figure 8.2: Left: an alliance network consisting of 7 nodes connected by positive and negative standing edges. Right: (part of) the derived triadic network reflecting which edges are shared between pairs of triadic nodes. Dark blue indicates strongly balanced, light blue is weakly balanced, pink is weakly frustrated, and the red node is strongly frustrated.

To build a triadic network, we generate a *triadic node* for each triplet of alliances in the network. These triadic nodes store the relevant properties for the state of the triad that we are interested in (e.g., how many positive and negative edges or whether it is frustrated), allowing us to track triad changes over time, as well as apply weights to the triads based on their properties (see below). Connections between triadic nodes are the edges of the original alliance network that are shared by pairs of triadic edges, so each edge of the original network can appear many times in the triadic network (see edge labels in Figure 8.2).

Given a dyadic network of N nodes, the derived triadic node network contains a number of triadic nodes equal to:

$$\frac{N!}{3! * (N-3)!} = \frac{N * (N-1) * (N-2)}{6}.$$

For our dataset of 607 alliances, the resulting triadic network contains approximately 37 million triadic nodes and approximately 1,375 trillion possibly edges.

This is far greater than most systems and software can handle. To make the above achievable, we wrote our own analysis toolkit in C++ that exploits the structure embedded inside the triadic network. Because we know how the triads are constructed, we also know how they are connected. In a fully created triadic network, each edge in the alliance network is replicated (N-2) times to connect (N-2) triads. Each triadic node is connected to (N-3) other triads per edge, for a total of 3 * (N-3) other triads. We construct the triadic network in computer memory using object-oriented software techniques: the triadic nodes and the edges of the original network are the objects in the software model, and the triadic nodes are interconnected through the edge objects via memory addresses.

8.5 Results: Structural Balance in EVE

Figure 8.3 shows the stacked proportions of triad types for sovereign alliances in EVE from 2/4/2015 to 4/17/2016. The proportions for large alliances (those with more than 200 members) are very similar (see Supplementary Materials). An alliance is included in every time period if it satisfies the inclusion condition (being large and/or sovereign) at any point during the analysis period. However, if the alliance was formed/dissolved during the analysis period it will (obviously) not be part of any triads before/after that time.

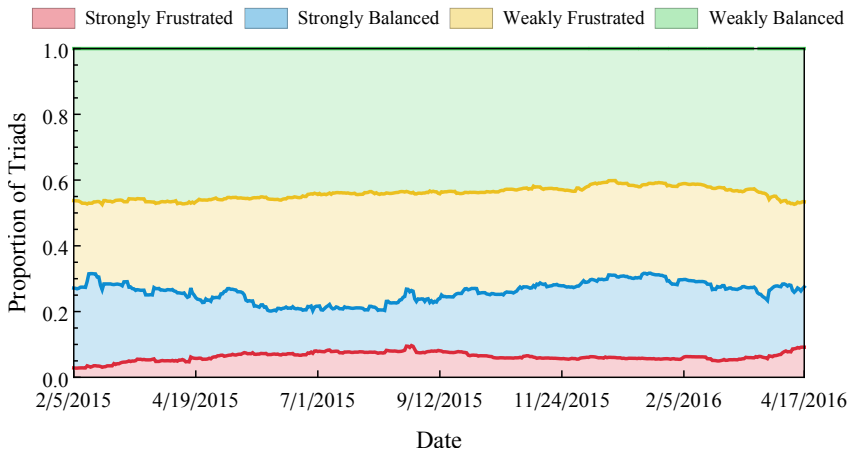


Figure 8.3: The stacked daily unweighted proportions of each triad type among alliances that hold sovereignty during our time frame.

Although the proportions clearly fluctuate over time, they are remarkably stable in light of the volatile game environment [Belaza et al., 2017, Belaza et al., 2019]. Considering the mean over time, the strongly balanced triads make up 19.69% of the total, while the strongly frustrated triads make up only 6.3%. Ac-

cording to the bipolar interpretation of classic balance theory, that first number is too small and the second too large. We also see a persistent level of 29.94% weakly frustrated and 44.07% weakly balanced triads. Nearly two-thirds of the triads are balanced, but that also means one-third of the triads are consistently unbalanced.

Although coalitions are not officially recognized within the game, the unofficial data [Chuggi and Sky, 2017] report that strongly balanced triads are mostly (but not exclusively) internal links within coalitions – mutually friendly relations are what makes it a coalition. Coalition members nearly all share the same enemies as well as friends, so every alliance in coalition A is aggressive to both alliances X and Y in coalition B, thus creating a large number of weakly balanced triads. In addition, nearly every alliance in coalition A is an enemy with all alliances in both coalitions B and C, and this generates the large number of triple negative (weakly frustrated) triads. The existence of coalitions thus helps make sense of why these two kinds of weak triads typically make up around three-quarters of the triads in the system.

The frequency distributions of alliances by type of triad (see Figure 8.3) are similar for sovereign and large alliances, indicating that the same general forces are operating in both contexts. Furthermore, the results in both cases support a modified structural balance analysis: grouping the weakly and strongly frustrated triads together drastically overestimates the number of triad changes we should expect to see. That is, strongly frustrated triads always make up less than 10% of the total, thus indicating that these arrangements are indeed avoided. The large percent of weakly frustrated (triple negative) triads support our conjecture that they do not actually inject frustration into a multi-polar political system like this one.

In order to establish a baseline for comparison we randomly assign the valence properties of the network edges while maintaining the same network structure and numbers of positive and negative links. This is repeated 100 times for each day's network, and we capture the time series of proportions of each triad type. A plot of the stacked mean proportions is shown in Figure 8.4 (a line plot showing the mean and three standard deviations is shown in the Supplementary Materials, but the variance is small). Due to the large numbers of negative links in the system, randomly assigning their location dramatically increases the number of strongly frustrated triads (from 6.3% (sov) to 17.22%) while dramatically decreases the number of strongly balanced triads (from 19.96% (sov) to 2.32%). Although less dramatic, there is a substantial increase in the triple negative weakly frustrated triads as well (29.94% (sov) to 37.05%). Clearly the political dynamics of EVE add meaningful structure to the distribution of edge valences in a way that reflects the social tensions of strong frustration and presence of friendly coalitions.

The number of large alliances climbed from 311 at the beginning, peaked at 348 in the middle of the dataset, and dropped back to 312 at the end (see time series plot in Supplementary Materials). However, there are actually 607 unique alliances

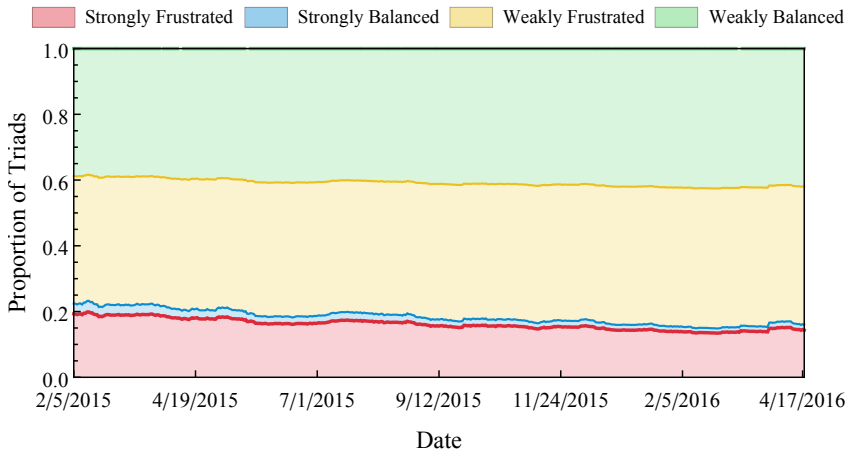


Figure 8.4: The stacked mean proportions of triad types across 100 samples of randomly assigning positive or negative weights while keeping both the structure of the network and the numbers of positive and negative links equal to the empirical data for each day.

that meet the criterion of having 200+ members within our analysis period, and all 607 of these alliances are included throughout our analysis. The number of triads among the large alliances grew steadily (though non-monotonically) from 50,265 to 170,296 during our analysis period. Although some increase can be attributed to alliances that come into existence and quickly become large during our analysis, such a dramatic growth in triads can only be explained by an increase in the edge density of the standings network.

Compared to large alliances, the numbers of sovereign alliances instead saw consistent growth from 83 at the beginning to 154 at the end with 282 unique alliances holding sovereignty over at least 1 system within our analysis period. Again, we include all 282 of these alliances throughout the analysis; i.e., the triads they form with other sovereign alliances are included at time t even if that alliance doesn't hold sovereignty at time t as long as it exists at time t and holds sovereignty at some point during our analysis period. The number of triads among these alliances grew from 10,490 to 45,373 during our analysis period. Comparing the ratio of active alliances to total unique alliances we can see that there is more churn among the sovereign alliances, thus reflecting that sovereignty is a more competitive status.

Despite these changes in alliance numbers and participants, the overall fluctuation in proportions is small. Collectively these results support our hypothesis that the alliance politics in EVE are consistent with multi-polar balance theory. On the other hand, we expected to find (1) even less strong frustration, (2) more dynamics in frustration over time rather than a steady level, and (3) a larger difference

between sovereign and large alliances' strong frustration (which are roughly equal here).

In order to explain these discrepancies we extend the analysis in two directions. First, in Section 8.5.1 we start from the intuition that not all alliances matter equally for systemic frustration and weigh the triads based on a variety of importance indicators. Then in Section 7.4.3 we dive into the contingent behavior of the alliance triads to evaluate whether resolving frustration is actually driving edge changes in the triads.

8.5.1 Extension 1: Weighted Structural Balance

In the previous analysis we discovered more strong frustration than expected, and nearly similar levels for large and sovereign alliances. Here we consider that not all alliances are equally important in assessing systemic frustration levels. We use properties of the components of each triad to weight that triad's contribution to the counts of each type of triad. For example, alliances that are far away from each other may form a strongly frustrated triad without it actually interfering with their activities, and so it should be weighted less in computing total system frustration. A small alliance may create a strongly frustrated triad with two large ones, who may not even notice. Concretely, we consider the following properties as weights:

1. [W1] the number of member players (M_i),
2. [W2] the number of systems over which an alliance has sovereignty (S_i),
3. [W3] Euclidean distance between two alliances (D_{ij}),
4. [W4] directed edge standings weights (W_{ij}), and
5. [W5] a combination of all four of the above factors.

Our reasoning for (1) and (2) is that the larger the alliance, the greater the impact its involvement in a frustrated situation will have, and hence the more pressure it will experience to alleviate the frustration [McDonald and Rosecrance, 1985]. Distance clearly plays a role in the importance of political ties [Neumayer and Plümper, 2010, Sommerer and Tallberg, 2019], and we propose that being far apart spatially (3) implies little *de facto* inconvenience, and hence less pressure to alleviate the frustration. If the roughly 6% of triads that are strongly frustrated involve small alliances, or alliances that are far away from each other, then the actual frustration in the system would be lower than the unweighted analysis above suggests. Although our construction of the triads uses symmetrified edges to determine the triad type, we use the sum of the absolute value of the weights of the 6 included directed edges to determine the standings weight (4). The directed edges have

weights between -10 and 10, so this weight amplifies triads among close friends (coalition members), sworn enemies, and mixtures of the two.

For a given triad, properties (1) and (2) are features of each node while property, (3) is a feature of each edge, and (4) is a property of the triad itself. For the combined weighting (5), we use the arithmetic mean of the four values, but for weights (1-3) we use the geometric mean of the three values for each triad. We focus on the geometric mean because it will assign a small weight to a triad if any of the three values is small, making it a natural match for this application. The larger alliances A and B are, the less likely they are to care about their relations with small alliance C. The arithmetic mean, on the other hand, will put a medium weight on a triad combining large and small alliances, which may overestimate the importance of the triad in the overall system. Our analysis of the effects on balance still uses the proportions of triads (or in this case the proportion of weights on triads), so they are effectively normalized.

8.5.1.1 Calculating the Weights.

Although the calculations of our weights are straightforward and unambiguous, here we provide the details to ensure total transparency. Consider a triad made up of three alliances A, B, and C. First we standardize each property ρ for each node or edge i to be between 0 and 1 using the standard method:

$$\frac{\rho_{i,t} - \min\rho}{\max\rho - \min\rho} .$$

That is, each value at each period is regularized by the minimum and maximum values over the entire time series so that changes in the values across time are not influenced by changes in the maximum or minimum values at a particular time. This is important because alliances gain and lose members, gain and lose systems, and move locations over time. The triad weighting for the number of members (W1) is calculated by

$$\mathcal{M} = \sqrt[3]{M_A \cdot M_B \cdot M_C}$$

and weighting by the number of systems an alliance holds sovereignty over (W2) similarly becomes

$$\mathcal{S} = \sqrt[3]{S_A \cdot S_B \cdot S_C} .$$

For distance weightings we use the Euclidean distance between the coordinates of the centroids of the systems each alliance holds sovereignty over:

$$D_{ij} = \sqrt{(\bar{x}_i - \bar{x}_j)^2 + (\bar{y}_i - \bar{y}_j)^2 + (\bar{z}_i - \bar{z}_j)^2},$$

where $\bar{x}$, $\bar{y}$, and $\bar{z}$ are the mean of the x, y, and z coordinates of the alliances' systems at that time. The distances are then standardized by the maximum and

minimum distances described above. The closer the alliances are to one another, the stronger the weight should be, so the distance weighting (W3) is calculated using

$$\mathcal{D} = 1 - \sqrt[3]{D_{AB} \cdot D_{BC} \cdot D_{CA}}.$$

(Below we also analyze a distance weight using the two closest systems.) The standings weights (W4) are calculated from the directed edges weights (W_{ij}) composing it as

$$\mathcal{W} = \frac{1}{60} \sum_{i=1,j=1,i\neq j}^{3,3} |W_{ij}|.$$

Neutral edges add a standings weight of 0, while both strong friends and strong enemies add 10 for each of up to 6 edges, hence the $1/60$ normalization factor. For the combined weighting (W5) we use the arithmetic mean of these four weights:

$$\mathcal{C} = \frac{\mathcal{M} + \mathcal{S} + \mathcal{D} + \mathcal{W}}{4}.$$

Because each weight is already rescaled to the $[0, 1]$ range, the combined measure is always within that range, without being again rescaled to fill that range.

8.5.1.2 Weight-Adjusted Frustration.

For each of the five weightings above, we recalculate the number of each of the four types of triads by summing up the weights of the triads on each day. Although the sums of the weights are always smaller than the sums of the counts, by looking at the proportions of each type of triad we can directly assess their relative incidence in a parsimonious way. Figures 8.5 through 8.8 show the triad proportions using each weighting scheme (solid line) compared to the unweighted sovereign proportions (dotted line) for sovereign alliances (plots for large alliances appear in the Supplementary Materials). Tables 8.1 and 8.2 summarize the differences in average proportions for both large and sovereign alliances for comparison along with a statistical measure of the significance of that difference.

For sovereign alliance triads weighting by membership (Figure 8.5) yields 0.86% fewer strongly frustrated and 2.37% fewer strongly balanced triads. This indicates slightly less participation in coalitions, and slightly stronger pressure to alleviate situations that would be awkward for its members. The total effect of membership-weighting is more than twice the size for sovereign alliances than large alliances; however, a stricter adherence to balance theory is not clearly the best explanation.

Weighting by sovereignty (Figure 8.6) has nearly zero (−0.01%) effect on strong frustration for the sovereign alliances, and the effect on large alliances is only marginally positive (+0.57%). Weighting the sovereign alliance triads by the number of systems they hold sovereignty over has an effect similar in total size

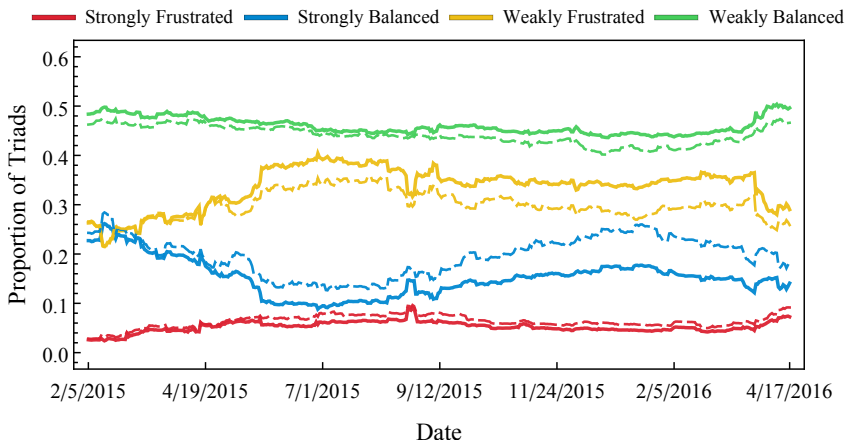


*Figure 8.5: The daily **membership-weighted** proportions of each triad type among sovereign alliances (bold line) compared to the unweighted proportions (dashed line).*

to membership-weighting, but in the reverse direction. For example, instead of a −2.37% change for strongly balanced triads, with sovereignty weighting we see a +3.53% increase. This means that alliances with large populations are less likely, but alliances with large territories are more likely, to be in the same coalition.

The total effects of sovereignty for large alliances (+20.3%) are nearly three times the effect on the sovereign alliances (+7.07%). So among large alliances (and recall that the sovereign alliances are also large), holding sovereignty has the largest impact on triad types (especially strongly balanced at +9.59%), although conditional on being sovereign the effects are smaller (+3.53%). The effect of sovereignty-weighting on the strongly frustrated triads is too small to tout it as a validation of balance theory, however the sensitivity of large alliances to sovereignty lends strong support to the idea that impacting alliance members with alliance-induced frustration does affect political decision-making within EVE.

For the distance-weighted triads (Figure 8.7), the main effect is more strongly balanced (+2.57%) and fewer weakly frustrated (−2.14%) triads. Although the total effect of distance is small for sovereign alliances (+5.43%), it is essentially negligible for large alliances (+0.91%). We initially expected distance to have a larger impact, but two factors appear to mitigate the effect. First, the territory of some alliances is very spread out, and as a result, our use of the mean location places that alliance in a location that is unrepresentative for the actual situation, sometime even placing them in a spot where they don't own any territory. Second, although there do exist clusters of nearby systems owned by a coalition of friendly alliances, when such a cluster is under a focused attack, the attacking alliance(s) often set up a base adjacent to the enemy's territory, or even relocate there al-

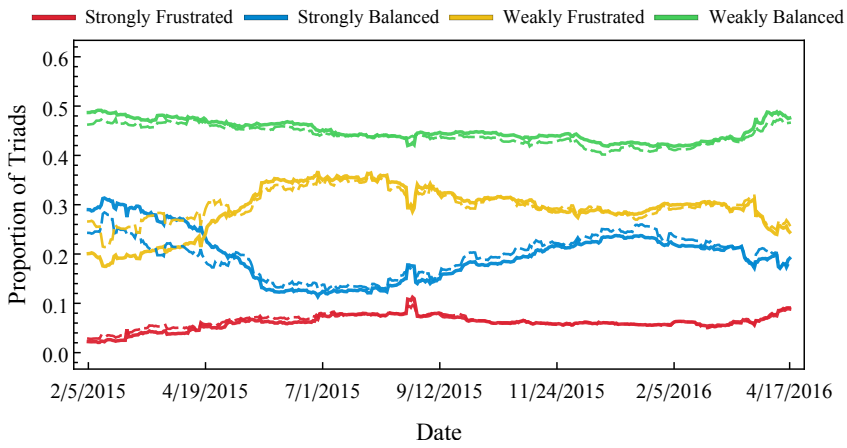


*Figure 8.6: The daily **sovereignty-weighted** proportions of each triad type among sovereign alliances (bold line) compared to the unweighted proportions (dashed line).*

together. Combined, we can see that the mean location of an alliance's systems may be a poor indicator of who its neighbors are, and alliances often have large numbers of friendly **and** enemy neighbors.

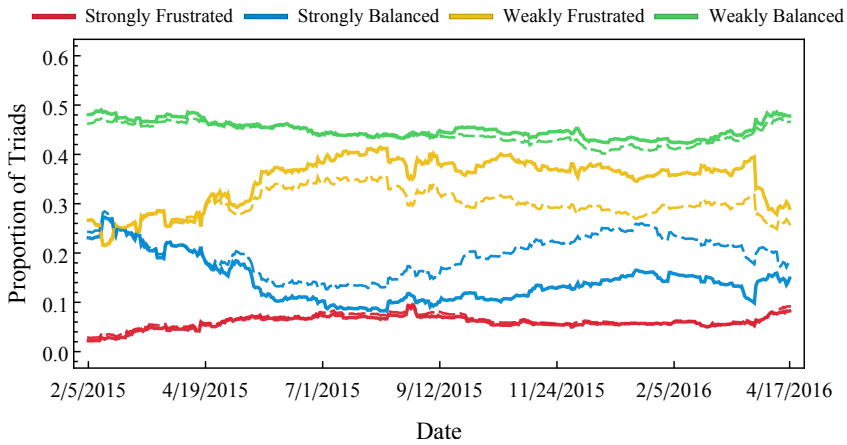


*Figure 8.7: The daily **distance-weighted** proportions of each triad type among sovereign alliances (bold line) compared to the unweighted proportions (dashed line).*

As an alternative to the centroid-based distance measure, we also investigate a distance weight using the distance between the two closest systems owned by each alliance. There is evidence that closer physical relationships are positively correlated with increased violence [Maoz et al., 2007, Gibler and Braithwaite, 2013] due to territorial disputes or long-standing cultural differences. What we find is that the

closest-system distance generates triad proportions very similar to centroid-based distance but slightly closer on average to the unweighted triads. The combined weight incorporating closest-system distance also reveals a slightly smaller change from unweighted triads (details in Supplementary Materials). For either version, distance has the weakest overall effect. The high mobility of alliance locations in EVE implies that, unlike geographically-bounded states, close proximity (by centroid or by nearest system) does not generate substantially increased conflict or partnership.

We do see a slight increase in strongly balanced triads to reflect close-knit and closely located coalition clusters. Distance-weighting decreases the weakly frustrated triads for the same reason; triads among three mutually antagonistic alliances are more likely when all three alliances inhabit different sectors of the map. There is a limit to this effect, because alliances that never meet are also unlikely to specify each other as enemies. Note that the difference in strongly frustrated triads for either distance weight (for large and sovereign alliances) are the only cases where the difference is not statistically significant at the 99% level. So the idea that actual system frustration is systemically reduced (or increased) when reducing the weight of triad members who are far from each other does not pan out.

The standings weights (Figure 8.8) provide the strongest support for balance theory in two ways: (1) the total effect is the largest for sovereign alliances and second largest for large alliances, and (2) the direction of the effects all correspond to those implied by the theory. That is, it is only for the standings weightings (and the combined weighting in virtue of them) that we see more of both types of balanced triads and less of both types of frustrated triads. The largest effects occur on the triple positive and triple negative triads which lends support to the existence of a multipolar political structure among coalitions.

The plots of the time series reveal that the effects of weights depend on dynamic features of the world. However, to ease further comparison, Table 8.1 shows the mean differences in proportions of all triad types between the unweighted and each weighted analyses for the large alliances, while Table 8.2 does so for the sovereign alliances. In addition to the magnitudes of the changes, we also report the p-value of the Kolmogorov-Smirnov two-sample test comparing each weighted series to the unweighted series [Wilcox, 1997]. This nonparametric test was chosen due to the non-normal (multi-modal) distributions of the differences (see Supplementary Materials).

We do not show the time series for the combined weighting here (see Supplementary Materials), but the results in the tables are enough to understand the most interesting points. The changes for the combined weights are *smaller* than most of the single weightings. For the large alliances this is particularly striking: despite the sovereign weighting having a 20% impact, the combined weighting (of which

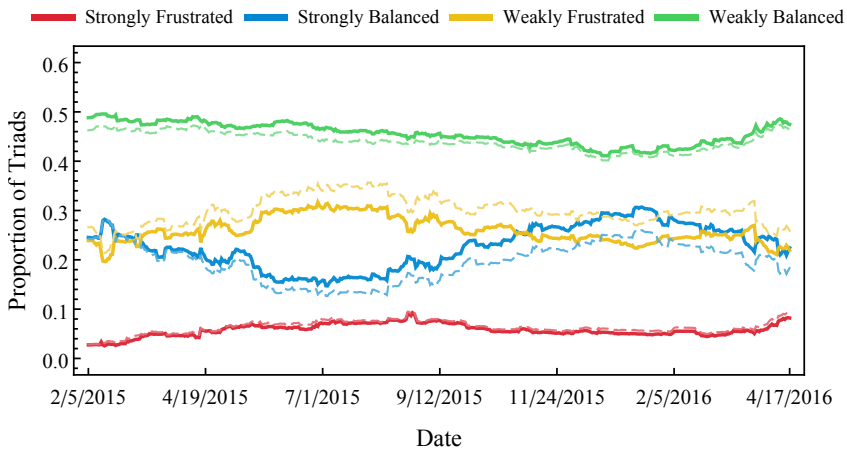


*Figure 8.8: The daily **standings-weighted** proportions (bold line) of each triad type among sovereign alliances compared to the unweighted proportions (dashed line).*

sovereignty makes up 1/4) only has a 3.45% impact. What this means is that, in those triads for which sovereignty had a big impact, the other weights were negligible/countervailing, and similarly for other weightings. Rather than finding that closely located, large alliances with many members and/or territories are strongly bonded (or some other combination effect), we find for both large and sovereign alliances that the four weightings are strongly complimentary. That said, the inclusion of the standings weights does pull the combined weights in the right direction to conform to the predictions of balance theory.

The largest surprise for the weighted triads proportion analysis is the lack of a clear effect on the strongly frustrated triads. Not only are the changes among the smallest, only the strongly frustrated triads reveal insignificant changes from weighting (for distance). If any of the weightings had brought strong frustration to near zero, this would have offered decisive support for our modified structural balance in this context. Although finding persistent frustration across all the weightings is consistent with empirical data on political relations [McDonald and Rosecrance, 1985], to more deeply explore the source of this structural frustration we move on to the next extension.

8.5.2 Extension 2: Conditional Behavior of Alliances

The previous analysis revealed a persistent level of strong frustration in the network, belying the tenet of balance theory that frustrated triads indicate relations that need to be straightened out, and that we should expect to see decreasing levels over time. One explanation in our domain is that there exists a constant stream

weighting	strongly balanced	strongly frustrated	weakly balanced	weakly frustrated	sum of abs differences
membership	+0.97	−0.62	−0.95	+0.60	+3.15
	($1.231e^{-8}$)	(0.000)	(0.000)	(0.000)	
sovereignty	+9.59	+0.57	−4.59	−5.57	+20.33
	(0.000)	($8.890e^{-12}$)	(0.000)	(0.000)	
distance	+0.17	−0.06	+0.28	−0.40	+0.91
	($4.304e^{-7}$)	(0.166)	($1.860e^{-8}$)	(0.000)	
standing	+2.18	−0.62	+1.81	−3.37	+7.97
	(0.000)	(0.000)	(0.000)	(0.000)	
combined	+0.99	−0.26	+0.74	−1.46	+3.45
	(0.000)	($9.281e^{-8}$)	(0.000)	(0.000)	

Table 8.1: The mean percent change (proportions ×100) between the unweighted triads and each type of weighted triad for the ***large alliances****. p-values of the Kolmogorov-Smirnov two-sample test appear below each value to indicate the significance of the effect.*

of exogenous events that push triads into frustrated political relations. That is, the level of frustration may be rather consistently around 6%, but the particular alliances involved in frustrated triads may be changing due to exogenous influences. This is especially plausible considering the churn we saw in the set of alliances holding sovereignty at a given time. Here we test the ideas that (1) alliances in frustrated triads tend to resolve those triads into balanced configurations, and (2) that the persistence of strong frustration occurs primarily through an influx of new relations.

As the anecdote about the BRUCE alliance in Section 8.4 illustrated, frustrated triads can lead to problems in collective action. Many factors typically go into the decision-making process of which allies to keep and which to drop, but structural balance theory implies that the change which minimizes local frustration

weighting	strongly balanced	strongly frustrated	weakly balanced	weakly frustrated	sum of abs differences
membership	−2.37 (0.000)	−0.86 (0.000)	+1.05 ($6.050e^{-11}$)	+2.18 (0.000)	+6.45
sovereignty	+3.53 (0.000)	−0.01 (0.007)	−0.19 ($2.953e^{-7}$)	−3.33 (0.000)	+7.07
distance	+2.57 (0.000)	+0.14 (0.036)	−0.58 (0.000)	−2.14 (0.000)	+5.43
standing	+3.00 ($1.110e^{-16}$)	−0.55 ($1.110e^{-16}$)	+1.47 (0.000)	−3.93 (0.000)	+8.95
combined	+2.55 (0.000)	−0.17 (0.001)	+0.27 (0.001)	−2.66 (0.000)	+5.65

*Table 8.2: The mean percent change (proportions ×100) between the unweighted triads and each type of weighted triad for the **sovereign alliances**. p-values of the Kolmogorov-Smirnov two-sample test appear below each value to indicate the significance of the effect.*

is the most likely, even if this is not explicitly part of the decision process. Keep in mind that alliances can only change their own standings to other alliances, a change which may or may not be (but usually is) reciprocated. The triad network is extremely dense, and a change in one edge to balance a frustrated triad can (and usually does) propagate across the triad network, inducing more triads to become frustrated [Bramson et al., 2017b].

In the previous two analyses we looked at changes in the proportions of triad types. Here we look at individual triads, and the propensity of each kind of triad to change to each other kind. Balance theory implies that balanced triads will tend to stay balanced, and strongly frustrated triads will be either avoided in the first place, or be resolved. Because the proportion of strongly frustrated triads remains fairly steady throughout the analysis period, the hypothesis we test here is that the

existing, frustrated triads are indeed balancing out, but there are other, exogenous sources of frustration that are being injected into the system, causing the whole to maintain a systemic, aggregate level of frustration.

8.5.2.1 Triad Dynamics.

Here we focus on sovereign alliances, to determine whether the behavioral predictions of balance theory materialize at the micro level, and to what degree the pressure to resolve frustration is affected by alliance sizes and distances. Our triadic network framework allows us to examine specific changes in the state of each triad in the system. Unlike previous analyses that examine changes in the system through dyadic changes in the network, using our triadic network construction we directly examine the triadic dynamics of the system. First we examine differences in the creation, persistence, and dissolution rates for each type of triad in Table 7.1. Dissolution of a triad means that at least one of the three symmetrified edges is removed (requiring both directed edges no longer exist), while triad creation implies a link formation that suffices to create a new triad.

	strongly balanced	strongly frustrated	weakly balanced	weakly frustrated
triad creation	21.71	8.98	44.52	24.79
triad persistence	19.74	6.43	43.83	30.22
triad dissolution	20.12	9.68	43.13	27.07

Table 8.3: Summary of results showing the percentages of unweighted sovereign alliance triad types when they are created, that persist in the system over time, and when they are removed. This only includes adding and removing triads through link creation and destruction; i.e., excluding nodes entering/leaving the system.

The strongly frustrated triad creation rate (8.98%) is slightly higher than its persistence rate (6.43%), and slightly lower than its dissolution rate (9.68%). This confirms that there is a proportional influx of frustration when links/triads are created, and a proportional outflux of strong frustration through removed links/triads. The proportions of both types of balanced triads are highly similar across entrance, persistence, and removal, but the weakly frustrated triads have a 30% persistence level, despite a 25% creation rate. This implies that triads are becoming weakly frustrated from the three other types. Next we examine conditional changes in triad types in Table 8.4; i.e., what kind of triad turns into what other kind of triad.

from	to strongly balanced	to strongly frustrated	to weakly balanced	to weakly frustrated	to nonexistent
strongly balanced	99.21	0.32	0.07	0.00	0.4
strongly frustrated	0.56	98.11	0.68	0.02	0.65
weakly balanced	0.01	0.07	99.42	0.22	0.27
weakly frustrated	0.00	0.00	0.22	99.62	0.15

Table 8.4: Percentage summary of unweighted triad type changes, including changes through the deletion of edges, for sovereign alliances.

More than 98% of the triads stay in the same state from day to day. The stability in the proportions of triad types we saw in the time series above can be attributed to the proportions of triads when created combined with the high persistence of triad types. That said, note that strongly frustrated triads are the least persistent and most likely to become nonexistent. Here we again see the relatively high dissolution of strongly balanced triads (more below) and the greatest stability occurring for weakly frustrated triads.

We now examine changes in triad types conditional on there being a change. We show these conditional triad change proportions, including triads that dissolve, in Table 7.3. (The analogous tables for large alliances and weighted triads can be found in the Supplementary Materials.)

The first observation is the size of the proportions of triads that are dissolved through link removal (right-most column of Table 7.3). It is worth noting that strongly balanced triads are more likely to be dissolved altogether, rather than to change type. The second observation is that when a triad does change its state, it is proportionally much more likely to change by a single valence flip. Because we use daily data, more than one edge change can occur in a single time step. And although this does happen, we find that 94.6% of changes (excluding deletion) are a single edge flip away.

This second observation is especially interesting in the context of the first one. Because a single edge change will always bring a strongly balanced triad to the strongly frustrated state, the fact that strongly balanced triads are disproportionately dissolved rather than changed can be explained by the avoidance of the frustration it would otherwise create. Comparing this dynamic to the weakly balanced triads further supports this conclusion: weakly balanced triads are more than three

from	to strongly balanced	to strongly frustrated	to weakly balanced	to weakly frustrated	to nonexistent
strongly balanced	—	40.43	8.84	0.03	50.7
strongly frustrated	29.33	—	35.65	0.9	34.12
weakly balanced	1.99	12.69	—	37.87	47.45
weakly frustrated	0.03	0.45	59.3	—	40.22

Table 8.5: Percentage summary of unweighted sovereign alliance triad type changes, including changes through the deletion of edges. Because we use daily data, it is possible for triads to change multiple edge valences in one iteration.

times more likely to transition to a weakly frustrated state (37.87%) than to a strongly frustrated state (12.69%). Dissolving the triad is more than four times more likely (47.45%) than changing to a strongly frustrated state. Although the persistence of strongly frustrated triads seems to conflict with the predictions of balance theory, we find that when there are changes, strongly frustrated triads tend to be preferentially avoided. This result provides reasonable support for balance theory as a partial explanation of triad dynamics.

8.5.2.2 Weighted Triad Dynamics.

Here we augment the unweighted analysis with a brief treatment of the effects of weights on triad changes. The tables for triad change rates for all weightings, for both large and sovereign alliances, appear in the Supplementary Materials. Here we present tables for the *differences* in change rates between the unweighted analysis and the combined weights, for creation, persistence and dissolution (Table 8.6), and for conditional changes (Table 8.7), both using the data from sovereign alliances. Although we present these only for the combined weights, the pattern in changes are similar across the individual weights (available in the Supplementary Materials) with only specific numerical difference by weight.

Table 8.6 shows that the total effect of the weightings is small in most places, but there are some interesting observations. First, creation and persistence of strongly frustrated triads have very small negative effects (less than 0.3%), but the combined weighting produces a 4.75% increase in their dissolution. Second, the persistence effect on both strongly and weakly balanced triads is positive, while the effect on weakly and strongly frustrated triads is negative. Thus the marginal

effect of the weights is an increased adherence to balance theory with respect to the kinds of triads that persist. Third, the effect is strongest for triad dissolution, where strongly balanced triads are 8.59% more likely, and weakly frustrated triads are 11.73% less likely, to be dissolved when the combined weights are applied. The strong changes in triad dissolution, however, are not clearly indicative of a stronger or weaker adherence to balance theory as a result of the weightings.

	strongly balanced	strongly frustrated	weakly balanced	weakly frustrated
triad creation	+2.22	−0.28	−0.12	−1.83
triad persistence	+2.84	−0.21	+0.20	−2.82
triad dissolution	+8.59	+4.75	−1.61	−11.73

Table 8.6: Summary of results showing percentage changes resulting from applying the combined weighting to sovereign alliance creation, persistence, and removal rates (i.e., combined weighting rates minus unweighted rates). This only includes adding and removing triads through link creation and destruction; i.e., excluding nodes entering/leaving.

We next examine our hypothesis that the inclusion of weights increases the predictive power of structural balance theory on triad changes. Table 8.7 reports the changes from Table 7.3 by including the combined weighting. The top row reveals that strongly balanced triads are 2.81% *more* likely to become strongly frustrated, and 3.28% *less* likely to become nonexistent. Strongly frustrated triads are 4.93% *more* likely to become strongly balanced, and 3.93% *less* likely to become non-existent.

The effect on transitions from strongly balanced and strongly frustrated triads is too small to make any strong claims, and the effect on other transitions is even smaller. However, we do note that these effects are not in the directions we would expect from balance theory. So, although the unweighted conditional triad changes did lend support to balance theory, our addition of weights to amplify these changes had little effect, and the overall effect does not strengthen our support for the predictions of structural balance.

from	to strongly balanced	to strongly frustrated	to weakly balanced	to weakly frustrated	to nonexistent
strongly balanced	—	+2.81	+0.48	+0.00	−3.28
strongly frustrated	+4.93	—	−0.93	−0.08	−3.93
weakly balanced	+0.63	+1.96	—	−0.03	−2.54
weakly frustrated	+0.01	+0.08	+3.95	—	−4.04

Table 8.7: Summary of results of the differences in the proportions of triad type changes between combined-weighted and unweighted of sovereign alliances conditional on there being a change, including though the deletion of edges.

8.6 Coalitions and Polarization

In addition to a tendency towards alleviating systemic frustration, balance theory also predicts a tendency for political entities to cluster into mutually friendly coalitions, separated by negative links. This kind of signed network clustering is often referred to as "polarization", and received a lot of attention in early studies of balance theory. The term "polarization" takes on many senses [Bramson et al., 2017a], and there are various measures for each of those senses [Bramson et al., 2016]. Although polarization is most often measured on distributions of values on a scale (e.g. of political opinions or beliefs) [Esteban and Ray, 1994], one can also measure polarization on spatial or networked values [Maoz, 2006, Esteban and Ray, 2008]. Importantly, one must distinguish between polarization of attributes across a network and polarization in the network structure itself.

Most social and political analyses are concerned with the effect of network structure on the polarization of social, political, and economic attributes. However, structural balance has a long tradition of analyzing structural polarization directly [Cartwright and Harary, 1956, Hart, 1974]. The earliest literature focused on determining whether a network could be split into two "polarized" groups. Later work also examined how well networks were split into (possibly more than two) polarized groups [Kulakowski, 2007, Doreian and Mrvar, 2009]. We are interested in the latter; specifically, we are interested in using the EVE standings data to determine how well balance theory's predictions about group formation describe the observed system of political relations in this virtual world.

As discussed in Section 8.3.1, the alliances in EVE form unofficial coalitions: political super-entities that are not part of the game's mechanics, yet are widely

recognized by players. Coalition dynamics include tacit internal no-conflict pacts, and joint strategizing among member alliances. We use player-reported, daily coalition data from [Chuggi and Sky, 2017] in the form of unofficial, yet widely used, coalition maps to assess how well the network structure of actual alliance standings corresponds to these player-reported coalitions.

Note, however, that this data is collected and reported by players, which has several limitations. For one, there are instances where alliances we know to hold sovereignty on a given day are unrepresented on the maps for that day. Our analysis here is limited to the alliances for which we actually have coalition membership data. These exclusions could be because the alliances are too small to bother; are widely known to the players to be closely associated with other alliances; or are simply not members of one of the existing coalitions. In any case, no reasons for their exclusion are provided. Also, the provided maps only cover sovereign space (the part of the map where ownership is possible), whereas coalitions between non-sovereign alliances also do exist. Non-sovereign alliances occasionally also participate in the coalitions reported in the data. This isn't a major problem because sovereign space is our main focus (for reasons explained above), but it does limit the option of expanding the analysis to include other parts of the EVE universe.

Because coalitions do not officially exist inside the game, only player-reported, hand-drawn coalition maps are available, we are unable to directly perform a systematic analysis of coalition dynamics across time. Despite these limitations, these player-reported maps are the best record of alliances' membership in coalitions, and we can use them to perform a preliminary test of the hypothesis that structural balance guides the formation of polarized coalitions.

8.6.1 Detecting Coalitions

We use the same standings data here as was used in the triad analysis, but now we are using it to assess polarization patterns in the network. Figure 8.9 displays the adjacency matrices of standings using the same intuitive color-scheme the players see in the game. On the left it is among the 38 sovereign alliances known to be involved in coalitions on 2/4/2015 (left), and on the right the 80 sovereign alliances in coalitions on 4/17/2016. By sorting the rows by (Canberra) similarity, we can easily discern a block structures of friendly links that is typical of network community structure and reminiscent of polarization in classical balance theory networks. To formalize the detection of coalitions we apply two categories of methods: (1) network community structure methods and (2) vector distance-based measures. For all methods we apply them to the directed standings, the reverse directed standings, and the unweighted symmetric matrix.

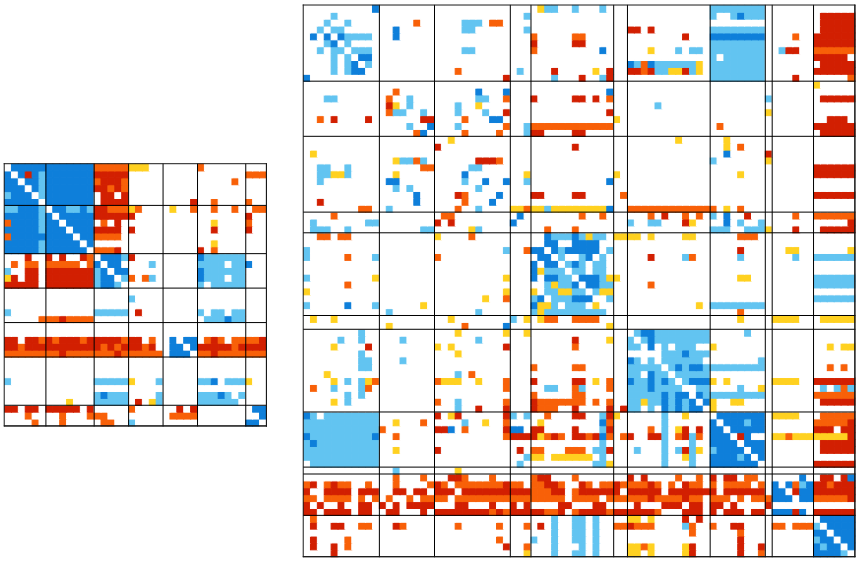


Figure 8.9: Plots of the adjacency matrices of alliance standings on 2/4/2015 (left) and 4/17/2016 (right) between the sovereign alliances included in [Chuggi and Sky, 2017]'s alliance and coalition maps. Blocks of dark and light blue (standings of 10 and 5 respectively) represent clusters of alliances that are densely interconnected with positive links. Negative links are represented by red (-10) and orange (-5), and neutral relationships by yellow (0) cells. White space indicates unset standings. The black mesh lines indicate coalitions discovered through clustering by Canberra distance on in-edges, one of the measures we compare to the player-reported coalition data.

8.6.1.1 Network Community Detection

Community detection algorithms are sophisticated techniques from network theory, specifically designed to find clusters of nodes that are densely connected internally, and sparsely connected externally. In the past decade, several research groups have turned to the challenge of detecting communities in signed social networks [Doreian and Mrvar, 2009, Traag and Bruggeman, 2009, Anchuri and Magdon-Ismail, 2012, Amelio and Pizzuti, 2013, Chen et al., 2014, Esmailian and Jalili, 2015], but these methods are not yet well-tested or easily accessible. One simple and obvious method to detect network communities is based solely on the positive links [Yang et al., 2007]. Although this technique ignores the repulsive force that negative links should have between communities, it suffices for our purposes because, in practice, alliances with no set positive standings act hostile towards each other by default. We include six different algorithms for identifying community structure in the networks (Hierarchical, Centrality, Vertex Moving, Modularity, Spectral, and Clique Percolation [Wolfram Research, 2019]).

8.6.1.2 Measures of Distance

Vector-based methods take each row (or column) as a vector of values, and apply a standard distance metric to perform pairwise tests of similarity. We applied the following list of distance measures here: Canberra, Euclidean, Normalized Squared Euclidean, Squared Euclidean, Cosine, Manhattan, Bray Curtis, Damerau-Levenshtein, Hamming, Correlation, and Chessboard [Wolfram Research, 2019].

8.6.1.3 Measure of Accuracy

After using a vector distance or network community method to partition the alliances into proposed coalitions, we need to determine how accurately they matched our best reference for the "real" coalitions. Naturally, a perfect match (getting "full points") occurs when each member of a discovered coalition is a member of that coalition in the player-reported data as well. In our approach, matches start with a full score and are then penalized for mismatches.

For each discovered coalition, we find the distribution of real coalitions for those members. We then identify the real coalition with plurality within the discovered coalition, and use it as the matching real coalition.[1] For each member of a discovered coalition that is not in the same real coalition as the plurality, we reduce the accuracy by one point. The points are then normalized by the number of alliances to create a percent accuracy score.

Note that even though this method does not directly penalize the splitting of a real coalition into two discovered coalitions, because we restrict the number of discovered coalitions to the number of real coalitions (in the case of vector distance methods), this error is still penalized insofar as it forces the displaced alliances to be associated with some other (incorrect) coalition. Thus, this measure penalizes each mismatched alliance exactly once.

Because setting the desired number of partitions is unavailable in our software (Mathematica 10.4 [Wolfram Research, 2019]), the number of detected partitions will vary, and this generally (though not necessarily) has a negative impact on accuracy. As a result, what we actually measure is *how many alliances in a detected coalition are not associated with a different real coalition*. This over-reports the accuracy with regards to a perfect sorting into the correct coalitions (in terms of edit distance). However, the technique suffices for our current exploratory purposes of comparing partition methods, and checking adherence to structural balance theory predictions.

[1] In the case of ties we choose the first coalition by index, because the decrease in matching score is the same regardless of which tied coalition is chosen.

8.6.1.4 Results of Detecting Coalitions

The mean accuracy among the best five methods for the out-edges, in-edges, and symmetric edges are 0.727, 0.756, and (by coincidence) again 0.756, respectively. These very similar scores belie the fact that the five best-performing methods for each category are actually quite different. When analyzing out-edges (the values actually set by alliance leaders), the network community detection methods are the most accurate (see Table 8.8, full tables available in the Supplementary Materials). The best method, hierarchical community detection, only mismatched 4 of the 38 alliances.

	First Day	Last Day	
Distance Measure	Out-Edges	Out-Edges	Mean
Hierarchical Community*	0.895	0.675	0.785
Centrality Community*	0.763	0.675	0.719
Vertex Moving Community*	0.763	0.675	0.719
Modularity Community*	0.763	0.663	0.713
Canberra Distance	0.737	0.663	0.700

Table 8.8: The top five ranked accuracy methods used to identify the coalitions reported by [Chuggi and Sky, 2017] using the rows of the directed weighted standings matrix (i.e., the out-edges). Network-based community detection algorithms are marked with an asterisk.

When clustering on the in-edges, we find that the vector-based measures outperform the network measures. Furthermore, the best measure (Canberra Distance) achieves scores of 0.868 and 0.763 for the first and last day respectively; this is better on average than Hierarchical Community Detection on the out-edges. We find that the vector-based measures also perform better on the symmetric data, yet the ranking of the measures is very different. For example, Canberra Distance drops from #1 to #6, and Euclidean Distance jumps from #13 to #2. Although the performance of any given measure may vary widely among the three data representations, the overall performance of the suite of measures is consistent (making it difficult to choose a single best method).

One form of this consistency is that nearly every measure on all three representations performs better on the first day than on the last day. Recall Figure 8.9,

showing the adjacency matrices; not only are there many fewer alliances on the first day, but they are also more clearly organised in a block structure. Some alliances form strong coalitions, and these are almost universally discovered. Other coalitions are weakly bound and more fluid. Upon analysis, some of the actual coalition members seem out of place and perhaps about to leave; thus some inaccuracy can be attributed to performing a static analysis on a changing, dynamic situation.

Using these techniques, we find a match between predicted and player-reported coalitions of roughly $70-80\%$. This result indicates that the unofficial, organically player-created coalitions in the game correspond to the emergent meta-structures predicted by balance theory reasonably well. It is rare to see such a clear demonstration of a theory of political organization.

8.6.2 Polarization among Coalitions

We now move from the accuracy of the coalitions found to the analysis of the link valences among coalitions. Figures 8.10 and 8.11 show network diagrams, both based on the standings on the first day of our data. The nodes represent the alliances, and the edge colors again reflect the directed standings between the alliances. The clusters, shown in Figure 8.10, reflect the real coalitions, while in Figure 8.11 the clusters represent the communities discovered by hierarchical partitioning based on positive edges.

The first thing to notice are several small coalitions at the top of Figure 8.10, joined solely by positive edges. These would naturally be considered as one single cluster by any network clustering algorithm. Two of these small coalitions are also mostly positively connected to the cluster on the bottom right. But with very few exceptions, the remaining inter-coalition edges are negative, and the remaining intra-coalition edges are positive. So, although some of the coalitions could be fused, by and large the coalitions created by players conform to the predictions of balance theory.

That said, we can also consider balance theory to claim that social ties will be partitionable (but not necessarily de facto partitioned) into groups of internally positive and externally negative relations. For this, we look at the results of hierarchical clustering in Figure 8.11, where we discover four clusters with only four negative edges within any of the clusters, and two positive edges between clusters (out of a total of 543 edges). So both the actual and the discovered coalitions are highly polarized, as predicted by balance theory.

Although we can't make any bold claims based on this preliminary analysis, the results here are sufficiently strong to merit deeper study. Because the coalitions are not part of the game mechanics, there are no official rules guiding their formation or behavior. There is no official need for political organizations above

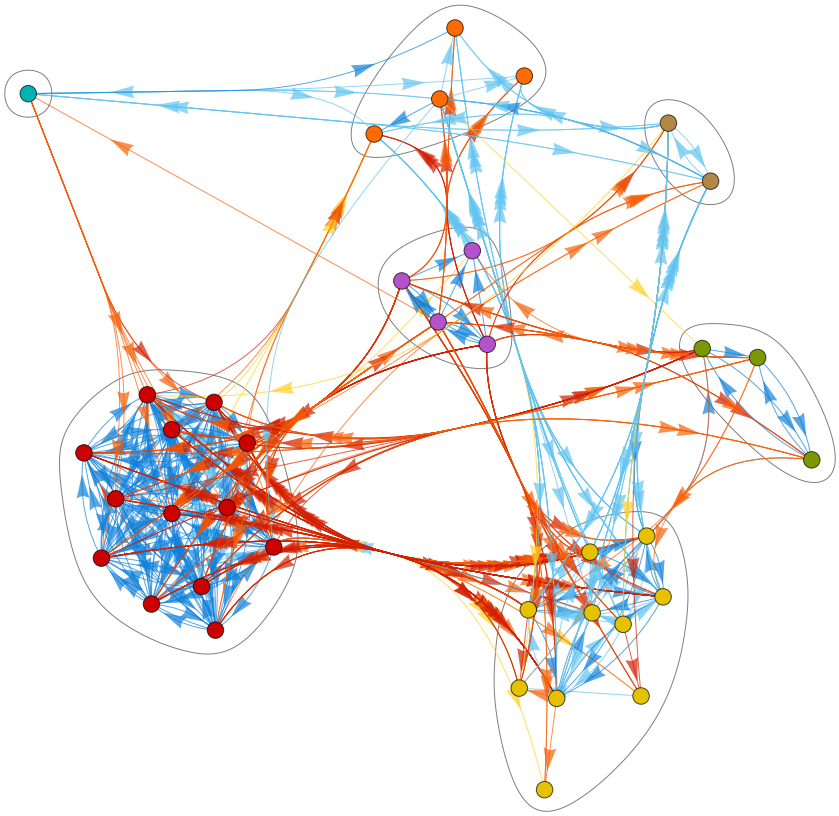


Figure 8.10: A network diagram of the alliance standings within each coalition on 2/4/2015 according to [Chuggi and Sky, 2017]. Alliance nodes are grouped by their collation, and the edges are colored by their directed standings as in Figure 8.9. Additional visualizations are available in the Supplementary Materials.

the alliance level to exist. Demonstrating that not only do coalitions emerge, but that the emergent coalitions are highly polarized, is a big win for balance theory.

8.7 Conclusions

Our analysis of political relationships between alliances in the virtual world of *Eve Online* revealed mixed support for balance theory. The persistence of strongly negative triads in Section 8.5.1 goes against balance theory, while the fact that these strongly negative triads consistently make up the smallest proportion of triads conforms with what we would expect based on the theory. The analysis of contingent behaviors in Section 7.4.3 showed that strongly frustrated triads are preferentially

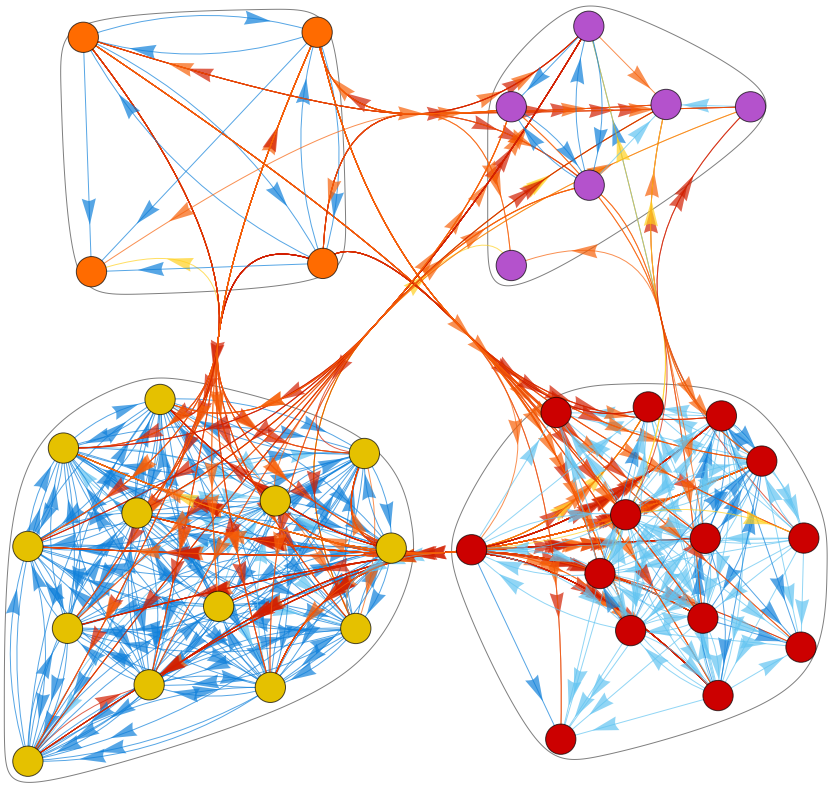


Figure 8.11: A network diagram of the coalitions discovered by hierarchical clustering on the positive edges, based on the standings data on 2/4/2015.

avoided, but that they are still tolerated more than expected. Section 8.6 makes the case that the mere existence of coalitions is already in support of balance theory. The clear polarization of the empirical and discovered coalitions provides the best support.

Our findings also highlight the importance of considering the system as multipolar rather than bipolar; i.e., splitting the frustrated and balanced triads into weak and strong versions of each, and tracking the four distinct structures. By doing so, we find that the prevalence of mutual antagonism (triple-negative = weakly frustrated triads) is much greater than predicted by traditional balance theory. In the game, as in real politics, there are more than two factions vying for power. And although unfriendly factions may temporarily team up to defeat a stronger mutual enemy, long-term three-way animosity seems a natural occurrence in large-scale political networks.

The effects of weighting the triads by membership, sovereignty and distance is smaller than expected, implying triads of all types are somewhat evenly distributed over alliance sizes and locations. Furthermore, the combined weight effects are smaller than the individual weights, implying that at least one of the three weights is very low when another is high. One exception is weighting by sovereignty, which has an appreciable effect for large alliances, and a much smaller effect for sovereign alliances. This supports our belief that sovereignty itself is important for political relations in EVE, because of how it affects access to resources. We expected the weighted analyses to reveal a small subset of major players (e.g., large and well-managed dominating alliances), adhering more strongly to balance theory than the average participant. Although the effects of the weights were in the right direction, further investigation is required to explain the deviations of alliance behavior from the predictions of balance theory.

8.7.1 Future Work

One of our broader goals is to explore whether data such as this can provide evidence for generalized law-like features of sociopolitical systems. Other work based on the EVE data aims to demonstrate that economic principles apply, and in fact apply more cleanly, in the partially idealized reality of virtual worlds [Hoefman et al., 2019]. Balance theory is one candidate for a sociopolitical theory that should hold with similar generality. Going forward, we will extend the analysis to other sociopolitical theories, and other aspects of EVE. Potentially, and most excitingly, the cleaner and simpler data from virtual worlds such as EVE may provide the inspiration and fuel to develop novel sociopolitical theories.

For example, just after the end of our current dataset there was an event in EVE called "World War Bee", in which several smaller coalitions attacked and dissipated the largest coalition at the time (whose logo was a Bee). The gradual increase in the density of the standings network towards the end of our dataset is due to the build-up to that event, which precipitated large changes in alliance wealth and sovereignty. Many old political bonds were broken, and new ones forged, as alliances tried to enhance their position during the war, and in its immediate aftermath. Having demonstrated the usefulness of data from this virtual world to study political structural dynamics, we plan to use data from the war period to also evaluate other social and political theories [Mildenberger and Pietri, 2018].

We currently use the actual standing weights (i.e., 10 vs 5) to weight triad frustration, and the symmetric positive vs negative valences of our triad edges coincide with the binary conflict conditions faced by most players (either shoot, or don't shoot). But setting a standing to 5 instead of 10, or setting it to zero (neutral) instead of leaving it unset, are intentional, and presumably meaningful signals about the relationship. Insofar as these signals are important for the political landscape,

we would like to include the directed edge weights (including neutral edges) into the formal analysis of triad composition and dynamics. This requires a technique that does not yet exist for balance theory (although [Belaza et al., 2017] and [Belaza et al., 2019] account for neutral ties and degeneracies in triad composition). There are currently very few methods for analysing signed and weighted networks, but we anticipate developing new methods for this research thrust.

Although balance theory is a popular and well-supported model, it is not the only model for analyzing tensions in social/political relations. [Axelrod and Bennett, 1993] offer a *Landscape Theory* that utilizes features of states to partition them into coalitions. Although not explicitly a network approach, the factors considered there (trade relations, distance, etc.) could be given a signed multigraph representation and analyzed for structural polarization. Both [Leskovec et al., 2010] and [DuBois et al., 2011] provide methods to predict the sign of removed links from Wikipedia and eOpinion trust data. The former method is based on triad balancing, while the latter makes no use of structural theory, yet their methods perform approximately equally as well. Furthermore, the latter performed roughly just as well on 50% as on 95% of the data, meaning there is a lot of redundancy in the structure of those graph for determining whether pairs of individuals are likely to trust or distrust one another.

One can also explore the change and spread of policies/properties across networks through econometric methods on dyadic data. For example, [Neumayer and Plümper, 2010] examine spatial effects in bilateral investment treaties, and [Sommerer and Tallberg, 2019] perform a similar analysis for the diffusion of participatory governance. Although these dyadic models may offer improvements over monadic models [Simmons and Elkins, 2004], to understand the dynamics of signed international relations one must at least analyze triadic relationships [Maoz et al., 2007]. Further examining the statistical properties of the triad changes through something like *SIENA* [Snijders et al., 2010, Snijders, 2017] may yield deeper insight into the structural and temporal dependencies driving the dynamics. Although statistical models looking at dyadic, triadic, and even larger relationships face serious theoretical and methodological issues [Cranmer and Desmarais, 2016] compared to network analytical methods, given the richness of the EVE dataset we are also interested in exploring alternatives to balance theory for explaining the political dynamics among alliances in future work.

Overall, our results bolster the relevance of balance theory for understanding a wide range of human behaviors – including video game politics. They also demonstrate the usefulness of data from virtual worlds for evaluating social theories. People are still social beings, even when interacting in virtual worlds, as long as the incentives structures are sufficiently realistic.

References

[Abell, 1968] Abell, P. (1968). Structural balance in dynamic structures. *Sociology*, 2(3):333–352.

[Amelio and Pizzuti, 2013] Amelio, A. and Pizzuti, C. (2013). Community mining in signed networks: a multiobjective approach. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 95–99. ACM.

[Anchuri and Magdon-Ismail, 2012] Anchuri, P. and Magdon-Ismail, M. (2012). Communities and balance in signed networks: A spectral approach. In *Advances in Social Networks Analysis and Mining (ASONAM), 2012 IEEE/ACM International Conference on*, pages 235–242. IEEE.

[Antal et al., 2006] Antal, T., Krapivsky, P. L., and Redner, S. (2006). Social balance on networks: The dynamics of friendship and enmity. *Physica D: Nonlinear Phenomena*, 224(1):130–136.

[Axelrod and Bennett, 1993] Axelrod, R. and Bennett, D. S. (1993). Landscape theory of aggregation. *British Journal of Political Science*, 23(02):211–233.

[BBCNews, 2014] BBCNews (2014). Eve online virtual war 'costs $300,000' in damage.

[Belaza et al., 2017] Belaza, A. M., Hoefman, K., Ryckebusch, J., Bramson, A., van den Heuvel, M., and Schoors, K. (2017). Statistical physics of balance theory. *PLoS one*, 12(8):12–15.

[Belaza et al., 2019] Belaza, A. M., Ryckebusch, J., Bramson, A., Casert, C., Hoefman, K., Schoors, K., van den Heuvel, M., and Vandermarliere, B. (2019). Social stability and extended social balance–quantifying the role of inactive links in social networks. *Physica A: Statistical Mechanics and its Applications*, 518:270–284.

[Bramson et al., 2017a] Bramson, A., Grim, P., Singer, D. J., Berger, W. J., Sack, G., Fisher, S., Flocken, C., and Holman, B. (2017a). Understanding polarization: Meanings, measures, and model evaluation. *Philosophy of science*, 84(1):115–159.

[Bramson et al., 2016] Bramson, A., Grim, P., Singer, D. J., Fisher, S., Berger, W., Sack, G., and Flocken, C. (2016). Disambiguation of social polarization concepts and measures. *The Journal of Mathematical Sociology*, 40(2):80–111.

[Bramson et al., 2017b] Bramson, A., Hoefman, K., van den Heuvel, M., Vandermarliere, B., and Schoors, K. (2017b). Measuring propagation with temporal webs. In *Temporal Network Epidemiology*, pages 57–104. Springer.

[Cartwright and Harary, 1956] Cartwright, D. and Harary, F. (1956). Structural balance: A generalization of heider's theory. *Psychological Review*, 63(5):277.

[CCP, 2019] CCP (2019). Eve online.

[Chen et al., 2014] Chen, Y., Wang, X., Yuan, B., and Tang, B. (2014). Overlapping community detection in networks with positive and negative links. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(3):P03021.

[Chuggi and Sky, 2017] Chuggi and Sky, M. (2017). Eve null-sec coalition influence maps.

[CorrelatesOfWar, 2019] CorrelatesOfWar (2019). The correlates of war project.

[Cranmer and Desmarais, 2016] Cranmer, S. J. and Desmarais, B. A. (2016). A critique of dyadic design. *International Studies Quarterly*, 60(2):355–362.

[Davis, 1967] Davis, J. A. (1967). Clustering and structural balance in graphs. *Human Relations*, 20:181–187.

[Doreian and Mrvar, 2009] Doreian, P. and Mrvar, A. (2009). Partitioning signed social networks. *Social Networks*, 31(1):1–11.

[DuBois et al., 2011] DuBois, T., Golbeck, J., and Srinivasan, A. (2011). Predicting trust and distrust in social networks. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, pages 418–424.

[Economist, 2015] Economist, T. (2015). Friends and foes: rifts in the middle east.

[Esmailian and Jalili, 2015] Esmailian, P. and Jalili, M. (2015). Community detection in signed networks: the role of negative ties in different scales. *Scientific reports*, 5:14339.

[Esteban and Ray, 2008] Esteban, J. and Ray, D. (2008). Polarization, fractionalization and conflict. *Journal of peace Research*, 45(2):163–182.

[Esteban and Ray, 1994] Esteban, J.-M. and Ray, D. (1994). On the measurement of polarization. *Econometrica: Journal of the Econometric Society*, pages 819–851.

[EVE-history.net, 2011] EVE-history.net (2011). Bruce - eve history.

[Eve-Offline, 2018] Eve-Offline (2018). Eve-online status monitor.

[EveWho, 2017] EveWho (2017). Eve who.

[Facchetti et al., 2011] Facchetti, G., Iacono, G., and Altafini, C. (2011). Computing global structural balance in large-scale signed social networks. *PNAS*, 108(52):20953–20958.

[Gibler and Braithwaite, 2013] Gibler, D. M. and Braithwaite, A. (2013). Dangerous neighbours, regional territorial conflict and the democratic peace. *British Journal of Political Science*, 43(4):877–887.

[Goh, 2018] Goh, A. (2018). How to build a robust game economy.

[Harary, 1959] Harary, F. (1959). On the measurement of structural balance. *Behavioral Science*, 4(4):306–323.

[Harary et al., 1965] Harary, F., Cartwright, D., and Norman, R. Z. (1965). *Structural Models: An Introduction to the Theory of Directed Graphs*. John Wiley & Sons.

[Hart, 1974] Hart, J. (1974). Symmetry and polarization in the european international system, 1870-1879: a methodological study. *Journal of Peace Research*, 11(3):229–244.

[Heider, 1946] Heider, F. (1946). Attitudes and cognitive organization. *Journal of Psychology*, 21:107–122.

[Hoefman et al., 2019] Hoefman, K., Bramson, A., Schoors, K., and Ryckebusch, J. (2019). The impact of functional and social value on the price of goods. *PLoS One*, 13(11):e0207075.

[Hummon and Doreian, 2003] Hummon, N. P. and Doreian, P. (2003). Some dynamics of social balance processes: Bringing heider back into balance theory. *Social Networks*, 25(1):17–49.

[Kulakowski, 2007] Kulakowski, K. (2007). Some recent attempts to simulate the heider balance problem. *Computing in Science & Engineering*, 9(4):80–85.

[Leskovec et al., 2010] Leskovec, J., Huttenlocher, D., and Kleinberg, J. (2010). Signed networks and in social and media. In *CHI 2010: Machine Learning and Web Interactions April 10–15, 2010, Atlanta, GA, USA*.

[Maoz, 2006] Maoz, Z. (2006). Network polarization, network interdependence, and international conflict, 1816–2002. *Journal of Peace Research*, 43(4):391–411.

[Maoz et al., 2007] Maoz, Z., Terris, L. G., Kuperman, R. D., and Talmud, I. (2007). What is the enemy of my enemy? causes and consequences of imbalanced international relations, 1816–2001. *The Journal of Politics*, 69(1):100–115.

[McDonald and Rosecrance, 1985] McDonald, H. B. and Rosecrance, R. (1985). Alliance and structural balance in the international system: A reinterpretation. *Journal of Conflict Resolution*, 29(1):57–82.

[Mildenberger, 2013] Mildenberger, C. (2013). *Economics and Social Conflict: Evil Actions and Evil Social Institutions in Virtual Worlds*. Springer.

[Mildenberger and Pietri, 2018] Mildenberger, C. D. and Pietri, A. (2018). How does size matter for military success? evidence from virtual worlds. *Journal of Economic Behavior & Organization*, 154:137–155.

[Neumayer and Plümper, 2010] Neumayer, E. and Plümper, T. (2010). Spatial effects in dyadic data. *International Organization*, 64(1):145–166.

[Simmons and Elkins, 2004] Simmons, B. A. and Elkins, Z. (2004). The globalization of liberalization: Policy diffusion in the international political economy. *American political science review*, 98(1):171–189.

[Snijders, 2017] Snijders, T. A. (2017). Stochastic actor-oriented models for network dynamics. *Annual Review of Statistics and Its Application*, 4(1):343–363.

[Snijders et al., 2010] Snijders, T. A., Van de Bunt, G. G., and Steglich, C. E. (2010). Introduction to stochastic actor-based models for network dynamics. *Social networks*, 32(1):44–60.

[Sommerer and Tallberg, 2019] Sommerer, T. and Tallberg, J. (2019). Diffusion across international organizations: connectivity and convergence. *International Organization*, 73(2):399–433.

[Squire, 2008] Squire, K. (2008). Open-ended video games: A model for developing learning for the interactive age. *The ecology of games: Connecting youth, games, and learning*, pages 167–198.

[Szell et al., 2010] Szell, M., Lambiotte, R., and Thurner, S. (2010). Multirelational organization of large-scale social networks in an online world. *PNAS*, 107(31):13636–13641.

[Traag and Bruggeman, 2009] Traag, V. A. and Bruggeman, J. (2009). Community detection in networks with positive and negative links. *Physical Review E*, 80(3):036115.

[UniWiki, 2017] UniWiki (2017). Corporation mechanics 101.

[Waltz, 1964] Waltz, K. N. (1964). The stability of a bipolar world. *Daedalus*, pages 881–909.

[Waltz, 1979] Waltz, K. N. (1979). *Theory of international politics*. Waveland Press.

[Wilcox, 1997] Wilcox, R. R. (1997). Some practical reasons for reconsidering the kolmogorov-smirnov test. *British Journal of Mathematical and Statistical Psychology*, 50(1):9–20.

[Wolfram Research, 2019] Wolfram Research, I. (2019). Mathematica, version 10.0. Champaign, IL, 2019.

[Yang et al., 2007] Yang, B., Cheung, W., and Liu, J. (2007). Community mining from signed social networks. *IEEE transactions on knowledge and data engineering*, 19(10).

Chapter 9

The Impact of Functional and Social Value on the Price of Goods

9.0 Preface

The following paper was published in 2018 in the A1 journal *PLOS One* [Hoefman et al., 2018]. It fits in this dissertation as an illustration of how virtual worlds data can contribute to economic theory, by testing the relationship between the quality and the price of economic goods using market data from the virtual world Eve Online.

My contributions to this first author paper were concept, data, empirical analysis and the writing of the paper. Further contributions to the concept and the writing were made by Aaron Bramson, Koen Schoors and Jan Ryckebusch.

9.1 Abstract

According to hedonic pricing theory (HPT) market forces operate on individual characteristics of a good, and the price of a product is the aggregate of the price across those characteristics. The relationship between price and characteristics remains poorly understood because characteristic qualities are hard to quantify, people have varying levels of information about characteristics, and people have heterogeneous preferences over characteristics. By analyzing data from a large, market-driven virtual world we are able to test HPT, while largely avoiding these pitfalls. We find that a linear model with functional characteristics predicts the prices poorly, but a log-linear model performs quite well. Adding social char-

acteristics to this log-linear model improves the predictions substantially. This work strongly supports HPT and demonstrates a "rational" calculus including social value.

9.2 Introduction

The canonical model for the determination of the price of an economic good is that of supply and demand, where market dynamics establish an equilibrium price at which aggregate demand is equal to aggregate supply. While useful in many situations, it is difficult to apply the supply and demand model to the price differences observed among differentiated goods having multiple interacting qualities. For this reason it is necessary to move the focus of analysis from the supply and demand of goods to the supply and demand of qualities of those goods.

9.3 Hedonic Pricing Theory

In a seminal paper, Rosen proposed a model called hedonic pricing theory (HPT) that describes the workings of a market for differentiated goods [Rosen, 1974, Lancaster, 1966, Court, 1939]. In this model, the price of a product is the aggregate of the price consumers are willing to pay for each quality characteristic of said product. As a consequence, products with more desirable characteristics command higher prices than products that are perceived to be of lower aggregate quality. HPT is used in markets such as wine [Oczkowski, 1994, Orrego et al., 2012, Dimsom et al., 2015], housing [Goodman, 1978, Malpezzi, 2002, Sander and Polasky, 2009], art [Gérard-Varet, 1995, Renneboog and Spaenjers, 2013], cars [Andersson, 2005], and agriculture [Ready and Abdalla, 2005]. But despite this widespread use and the clear formal structure of the theory, the econometric relationship between quality and price is still not well understood due to difficulties acquiring the proper data.

While accurate information on product prices is often readily available, reliable quality information is much harder to obtain. Challenges can be summarized into six main categories. First, one would have to be able to identify all the relevant quality factors, since an underspecified model would give inaccurate predictions. Second, some quality factors (such as terroir for wine) are almost impossible to quantify. Third, consumers tend to have heterogeneous preferences, implying that what constitutes a quality factor for some may be irrelevant to the purchasing decision of others. Fourth, consumers may not only have preferences over goods' functional qualities, but also over their social qualities such as prestige and exclusivity (see further). These are often idiosyncratic, hard to measure, or even unobservable. Fifth, consumers may have preferences over goods' per-

ceived social externalities such as environmentally friendly goods [Kalafatis et al., 1999, Cronin et al., 2011], non-GMO goods [Baker and Burnham, 2001], organic goods [Yiridoe et al., 2005, Lee et al., 2013], fair trade goods [Hainmueller et al., 2015], etc. Social norms conferred via descriptive or injunctive signals [Schultz et al., 2007] may lead them to believe these goods are morally superior. Recent research [Capraro and Rand, 2018, Tappin and Capraro, 2018] shows individual preferences may depend on the framing of available options as moral or immoral. Sixth, consumers with imperfect information may fail to correctly assess quality, either because they don't have the necessary expertise or because the cost of obtaining the information is too high. These factors make real-world data on quality very noisy at best, rendering it difficult to pick up a pattern even if it is present. While this is unavoidable because the world is a complex place, it also means that it is difficult to test HPT (like many other economic and social theories) using real world data.

If it were possible to obtain perfect information about the quality for a range of products of a differentiated good, one could create a model to determine whether HPT explains the relationship between those qualities and prices. Identifying such a relationship would offer solid support for HPT as a scientific theory. And in doing so, it could provide strong evidence of consumer rationality [Simon, 1986] by demonstrating that the expenditure of resources can be strictly explained by the underlying quality of the purchased good. However, applying this degree of scientific scrutiny towards HPT has been thus far impossible.

Towards this end we analyze the less noisy environment of a virtual world called Eve Online. Virtual worlds are computer-based environments offering unique opportunities for testing economic and social theories [Castronova, 2004, Thurner et al., 2012] due to the following properties: (1) a computer keeps track of the state of everything that populates it; (2) properties of items are precisely defined; (3) data is available about the entire population of participants; (4) all participants can be given access to all relevant details, (5) players can act in a situation of (almost) perfect information; (6) participants tend to have more homogeneous preferences over the quality characteristics because virtual worlds are invariably simpler than the real world; and (7) perceived moral qualities of a good are not expected to play any role given the general amoral nature of the environment. Although most games are too simple to provide support for such an endeavor, some large-scale virtual worlds embody dynamics that are representative of the behavior we would like to study in the real world.

9.4 A Virtual Laboratory

Eve Online (EVE) is an open-ended Massively Multiplayer Online Game set in a science fiction universe created by Icelandic company CCP Games in 2003. More

than 500,000 players compete for resources and territory while engaging in a variety of professions and activities including mining, manufacturing, trading, piracy, exploration, and combat, both versus the environment and against other players. As a sandbox game [Squire, 2008], EVE provides its players with a virtual world and the tools to explore, but players have the freedom to choose what, when, and how to approach the available content, including the purchase and sale of goods.

EVE contains 12,709 distinct items that can be bought and sold between players including ships, ship modules, minerals, ammunition, blueprints, and many more. The items available and their characteristics (which players can view at any time at zero cost) are decided by CCP and only rarely adjusted. Market prices, on the other hand, are endogenously determined by the market behavior of players via a double auction system that matches buy orders with sell orders. Prices fluctuate daily around a mostly stable base (see Fig 9.1). Players can buy and sell anywhere in the virtual universe, but for reasons of efficiency market activity tends to cluster in hubs. Two thirds of all market transactions are conducted in a single central trading hub.

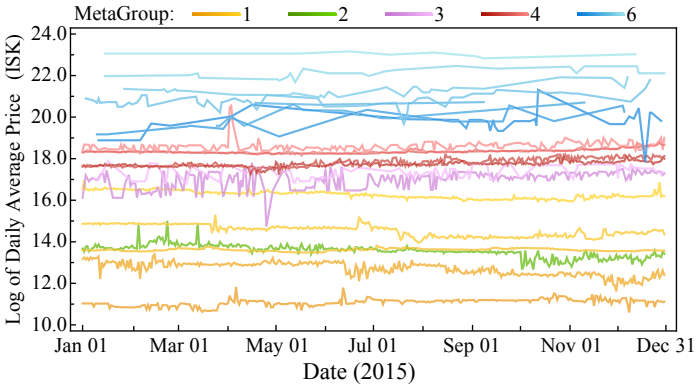


Figure 9.1: Time series of the natural log (ln) of prices in the main trading hub in 2015 for each of the 21 varieties of the Ballistic Control System module. Module prices are subject to supply and demand dynamics but maintain an overall stable price ranking. Diagrams for other items are available in the Supplemental Information.

Items and resources in the game are bought and sold with an in-game currency called the Inter-Stellar Kredit (ISK). Players earn ISK as a reward for engaging in activities such as defeating pirates, running missions, selling resources gained through mining, selling goods made through industry, offering services like courier contracts or protection to other players, or by paying real-world money. Most players spend time in the game earning ISK to purchase ships, modules, skills, etc. Advanced players even pay for their game subscription through their in-game earnings. One billion ISK is roughly equivalent to $15 USD in the period of our

analysis. Because ISK has both in-game and real-world value, and because losses in EVE are permanent, players tend to be risk averse, and this fosters realistic economic behavior.

Our analysis focuses on ship modules (henceforth 'modules') such as weapons and defensive systems that can be fitted onto ships to improve their performance. In line with the terminology of Rosen, a module *class* contains all module varieties of the same type. There are 296 different module classes in EVE. Of these we select the 20 most traded based on total market revenue for modules in 2015 (equivalent to approximately 12 million USD, 2015 is a typical year in the Eve universe), representing 35% of total module revenue. We filter out 4 of these because the game designers changed their characteristics during 2015. The remaining 16 module classes contain between 14 and 43 varieties, for a total of 396 distinct modules in our analysis. These modules bestow 1 or 2 *benefits* (Armor, Shield, Damage, Range and/or Strength) where a *higher* value implies a better item and add between 0 and 3 *constraints* (CPU, Power and/or Energy) where a *lower* value implies a better item, depending on the class. As a consequence the total number of functional quality characteristics for each module class varies from a minimum of 1 to a maximum of 5. This heterogeneity in the number of explanatory variables prevents us from analyzing the price of all modules within a single regression.

HPT predicts that modules with higher values for the benefits should be more expensive, whereas higher values for the constraints should result in a lower price. In light of the observed price consistency (recall Fig 9.1), we take the mean of the average daily price in the main trading hub over 2015 as the reference price for each module. Table 9.1 shows the characteristics and in-game reference prices for each variety of our running example module class: it has a single benefit (Damage) and a single constraint (CPU). Characteristics and prices for all module classes can be found in the Supplemental Information. The information in these tables is available to players in the game at all times.

9.5 Pricing Models with Functional Quality

HPT is agnostic about the functional form of the relationship between price and quality [Halvorsen and Pollakowski, 1981, Atkinson and Halvorsen, 1984]. For example, while people who aren't trained economists (and many who are) intuitively expect a linear relationship such that a product twice as good in overall quality would also be twice as expensive [Karelaia and Hogarth, 2008, Deane et al., 1972], several empirical hedonic pricing studies have found that nonlinear forms typically perform better than the linear relationship (see [Orrego et al., 2012] for an overview). It is well understood that the functional form is critical to the accuracy and consistency of the econometric model [Halvorsen, 1988, Brown and Ethridge, 1995, Triplett, 2004], so here we explore both a linear and non-linear relationships

		Functional Qualities		Social Qualities		
	Module Name	Damage (+%)	CPU	MetaGroup	Rarity	Price (ISK)
1	Ballistic Control System I	15.02	35	Tech I = 1	1	65 968
2	Cross-linked Bolt Array I	16.59	37	Tech I = 1	3	346 039
3	Ballistic Control System II	21.55	40	Tech II = 2	1	816 551
4	Muon Coil Bolt Array I	18.16	39	Tech I = 1	3	816 602
5	Multiphasic Bolt Array I	19.74	40	Tech I = 1	3	2 139 204
6	'Pandemonium' Ballistic Enhancement	21.33	42	Tech I = 1	3	11 818 778
7	Ballistic 'Purge' Targeting System I	18.25	30	Storyline = 3	4	26 122 686
8	'Full Duplex' Ballistic Targeting System	19.35	30	Storyline = 3	4	36 102 807
9	Domination Ballistic Control System	21.55	28	Faction = 4	3	50 382 156
10	Republic Fleet Ballistic Control System	21.55	28	Faction = 4	4	55 074 387
11	Caldari Navy Ballistic Control System	24.31	24	Faction = 4	2	92 894 448
12	Dread Guristas Ballistic Control System	24.31	24	Faction = 4	3	94 448 681
13	Khanid Navy Ballistic Control System	24.31	24	Faction = 4	4	127 593 395
14	Mizuro's Modified Ballistic Control System	22.10	31	Officer = 6	4	435 538 765
15	Hakim's Modified Ballistic Control System	22.66	34	Officer = 6	4	444 063 110
16	Gotan's Modified Ballistic Control System	23.21	36	Officer = 6	4	561 835 333
17	Tobias' Modified Ballistic Control System	23.76	39	Officer = 6	4	945 333 333
18	Kaikka's Modified Ballistic Control System	25.00	26	Officer = 6	4	1 137 267 911
19	Thon's Modified Ballistic Control System	25.69	29	Officer = 6	4	1 900 513 954
20	Vepas' Modified Ballistic Control System	26.38	31	Officer = 6	4	4 023 867 247
21	Estamel's Modified Ballistic Control System	27.07	34	Officer = 6	4	10 021 052 707

Table 9.1: Properties and mean of the daily average price in 2015 of all 21 varieties of the Ballistic Control System module.

between price and quality.

For each of the 16 module classes in our dataset we analyze the degree to which prices within the module class are explained by the relevant functional quality characteristics using ordinary least squares regression. We also test for a log-linear relationship by using the natural logarithm of the price as the dependent variable. This process yields two sets of equations relating the price P_i of each module in class $i \in \{1, 2, ..., 16\}$ to the benefits and constraints of its $1 \leq n_i \leq 5$ functional quality characteristics F_{ij}:

$$P_i = \alpha_i + \sum_{j=1}^{n_i} \beta_{ij} F_{ij} \quad \text{(Linear Functional Quality)}, \tag{9.1}$$

$$\ln P_i = \alpha_i + \sum_{j=1}^{n_i} \beta_{ij} F_{ij} \quad \text{(Log-linear Functional Quality)}. \tag{9.2}$$

Here, the coefficient α_i captures any systematic contributions to price that cannot be attributed to the quality parameters included in the model.

In the Supplementary Materials we present the goodness of fit for all 16 module classes in terms of the coefficient of determination R^2 (the proportion of the variance in the predicted module price, P_i or $\ln P_i$, that is predictable from the independent variables). We also provide two alternate measures to compare the performance of competing models: the adjusted R^2 and the Akaike Information Criterion (AIC), an entropy-based metric where a lower value indicates a better fit [Bozdogan, 1987]. In order to assess the overall performance of the models we average the performance measures across all 16 modules. These results are summarized in Table 9.2.

We find that the log-linear model captures the contribution of quality to price rather well. Averaged across all 16 modules, the log-linear model explains 79.5% of the variance in prices as measured by R^2, and the coefficients of the quality characteristics consistently have the correct sign (positive for benefits, negative for constraints). The linear model doesn't perform nearly as well, with an average R^2 of 49.6% and coefficients that often have the wrong sign.

Fig 9.2 plots the standardized residuals of the fitted values versus residuals. The values are divided by the standard deviation of prices within each module class. The linear model predicts negative prices, and the residuals demonstrate uncaptured dynamics and sometimes errors larger than the fitted values themselves. The predicted values of the log-linear model, however, are in range with the actual observed prices, and the residuals are randomly distributed (meaning the model is correctly specified).

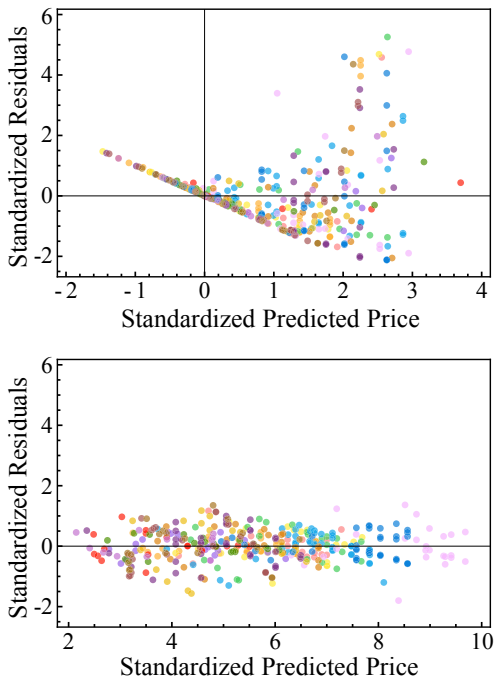


Figure 9.2: Comparison between the standardized predictions of the linear functional quality pricing model (top) and the log-linear functional quality pricing model (bottom) colored by module class (16 in total).

9.6 The Contribution of Social Value

So far our pricing models have only considered *functional quality*: those characteristics that affect the (objective) performance of a product. Our dataset includes module classes with multiple items having identical functional qualities (e.g., modules 9-10 and 11-13 in Table 9.1). Often, large price differences occur between these functionally identical modules; for example, players pay a surplus of 37% for module 13 over 11 for no functional benefit. Furthermore module 17 is functionally *inferior* to modules 11-13, yet its price is an order of magnitude higher.

These discrepancies seem to refute the core concept of HPT. If the price of a product is the aggregate of what people are willing to pay for the individual quality factors of said product, then modules with identical quality characteristics should command the same price. That is, unless we accept that other, non-functional characteristics have value too.

The frequent occurrence of modules with identical functional properties commanding different prices in our dataset allows us to isolate what has been termed "conspicuous consumption" in the literature. The willingness to pay extra for no

discernible benefit was first described by Veblen [Veblen, 1899] in 1899 as a way for the rich to demonstrate their wealth. Whether it is rational to pay for no added functional benefit is a controversial topic in economics even today. The more classical explanation is one of competition for status goods via the expenditure of wealth [Ireland, 1994, Bagwell and Bernheim, 1996] which can lead to negative consumption effects [Frank, 1985, Hopkins and Kornienko, 2004]. Interestingly, it is often those earning the least that spend the greatest fraction of their income on conspicuous consumption [Charles et al., 2009, Kaus, 2013]. The motivation for engaging in conspicuous consumption is classically explained by the comparative advantages obtained from (increased) social status, such as sexual selection [Sundie et al., 2011]. But more recent research reveals that psychological aspects like identity [Akerlof and Kranton, 2000], social belonging [Han et al., 2010] and self-esteem [Sivanathan and Pettit, 2010, Bursztyn et al., 2017] also play a role. We use the term *social value* to refer to benefits that are derived from the social confirmation of a consumer's ability to pay. Since social value exists in the interaction with other people, only characteristics that can be observed by others are candidates for this type of consumption. In EVE all modules share two characteristics that are observable by all players yet add no functional benefit to the module.

The first social variable is the Rarity of the module, which can be observed by looking at market volumes. Any positive effect of rarity on prices within a class of differentiated products cannot be attributed to the dynamics of supply and demand alone. When two products are identical in quality yet different in price, rational buyers should always prefer the cheaper alternative. This would cause demand to shift to the cheaper substitute, until an equilibrium is reached where both prices are the same even though their supply may be different. Economists refer to this as the non-arbitrage condition of any economic equilibrium [Werner, 1987]. If rarity is found significant for the price while controlling for functional quality, this implies that the item's rarity confers social value to its owner.

The second social variable, the module's so-called MetaGroup, appears as a color-coded marker on items informing players about how the item enters the game. There are six categories of MetaGroup with increasing exclusivity (MG_i): Tech I (for the most basic modules), Tech II, Storyline, Faction, Deadspace, and Officer (the most exclusive). The meta groups are correlated with functional quality by design (see Table 9.1); however, this correlation varies considerably across characteristics and module classes, and significant functional quality overlaps occur across metagroups. Specifically, we even have 39 pairs of modules from different metagroups with identical functional characteristics, thus creating a natural experiment for the role of social value.

By analyzing price differences for these functionally identical pairs we find that people pay at least 2 times, and on average 7 times, more for an identical

Pricing model	R^2	adjusted R^2	AIC
linear functional quality (Eq. 9.1)	49.6% ± 21.6%	42.5% ± 22.7%	1076.4 ± 356.8
log-linear functional quality (Eq. 9.2)	79.5% ± 10.3%	75.9% ± 12.4%	95.8 ± 27.7
log-linear functional + social quality (Eq. 9.3)	95.35% ± 2.77%	93.76% ± 4.12%	63.26 ± 20.47
log-linear functional + social quality (Eq. 9.4)	95.42% ± 2.95%	93.93% ± 3.83%	61.84 ± 23.84

Table 9.2: Performance of the four pricing models averaged across the 16 module classes.

module in a higher MetaGroup. This price difference can only be understood as a social value premium: players pay for the social status of an item in addition to (or apart from) how good the item actually is in terms of functional quality.

9.7 Pricing Models with Functional and Social Quality

Here we test the importance of the social variables along with the functional ones using econometric models of HPT.

To do this we incorporate the social quality characteristics in our pricing model by adding variables that capture their rarity and metagroup membership. To measure the contribution of rarity, we define the variable Rarity R_i using a logarithmic binning technique based on its average daily market volume (v) during 2015. Modules with the lowest rarity (i.e., $v \geq 10000$) are assigned $R = 0$, $10000 > v \geq 1000 \rightarrow R = 1$, $1000 > v \geq 100 \rightarrow R = 2$, $100 > v \geq 10 \rightarrow R = 3$, and $v < 10 \rightarrow R = 4$. We measure the contribution of MetaGroup MG_i with a discrete variable ranging from 1 (Tech I) to 6 (Officer) of increasing exclusivity as described above.

Adding parameters R_i and MG_i to our purely functional log-linear model (9.2), we obtain the log-linear social pricing model in Equation 9.3. We test the robustness of this log-linear model with an alternative model where the contribution of each MetaGroup value is measured with a separate dummy variable MG_{ik} (except for the lowest category which is captured by the regression constant α_i). This alternative model is shown in Equation 9.4.

(Log-linear Functional + Social Quality, v1)

$$\ln P_i = \alpha_i + \sum_{j=1}^{n_i} \beta_{ij} F_{ij} + \gamma_i R_i + \delta_i MG_i \,, \tag{9.3}$$

(Log-linear Functional + Social Quality, v2)

$$\ln P_i = \alpha_i + \sum_{j=1}^{n_i} \beta_{ij} F_{ij} + \gamma_i R_i + \sum_{k=1}^{6} \delta_{ik} MG_{ik} \,. \tag{9.4}$$

We find that the pricing model from Equation 9.3 that accounts for both functional value and social value explains 95.35% ± 2.77% of the observed price variance across all 16 module classes compared to 79.5% ± 10.3% for the purely functional models (see Table 9.2).

Adjusted R^2 and AIC (from 95.8 to 63.26) confirm that the addition of the social value parameters improves the performance of the model considerably. The results from the model in Equation 9.4 are nearly identical. Fig 9.3 illustrates this by comparing the fitted values of the log-linear model with the actual price observations for two example module classes. Table 9.2 summarizes the results while individual regression results for each module class are shown in the Supplementary Materials. Fig 9.4 shows the improvement in fit for the log-linear model across all modules and module classes by adding social quality to the model by plotting their normalized residuals.

9.8 Isolating Social Quality

Recall that in an HPT regression, the total price of an item is broken down into the contributions of the parameters that are included in the model. In Equation 9.3 (where MG_i is represented by a discrete variable) the effects of Rarity and MetaGroup on the total price can be isolated with 95% confidence in 8 of the 16 regressions. For these 8 module classes we employ the estimated coefficients to derive the individual contributions from each of the functional characteristics, rarity and the meta group, to the (natural log of the) predicted module price. For the model in Equation 9.4 we use the averages of the estimated coefficients for each of the MetaGroup values.

From Fig 9.5 it is clear that not only does social value contribute substantially to prices, but also that the relative contribution of social value to the price tends to rise with the more exclusive items (as should be expected). This provides clear support for the claim that social value is economically relevant and demonstrates that the economic value of these social qualities can be assessed through the lens of HPT.

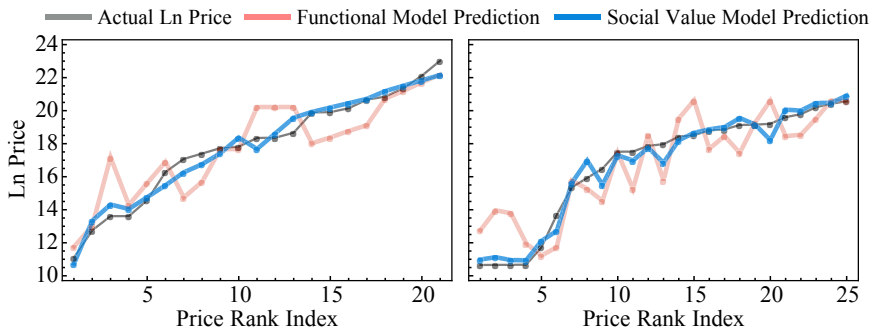


Figure 9.3: Comparison between observed prices (grey), predicted prices from the log-linear model including only functional value (red, Equation 9.2) and predicted prices from the log-linear model also including social value (blue, Equation 9.3). Left: prices from the Ballistic Control System class shown in Table 9.1 increase exponentially (here drawn in their natural logarithm), and this is captured accurately by the log-linear model including social value. Right: the 25 modules in the Large Shield Booster class exhibit flat prices till module 4, a sharp rise in prices between modules 4 and 8, and a smooth exponential rise from module 10 onwards. The log-linear model including social value captures these features relatively accurately.

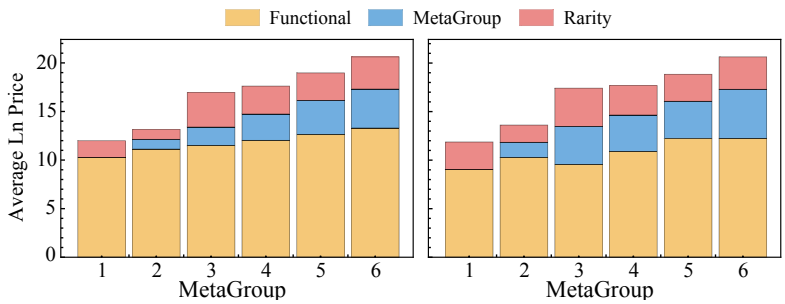


Figure 9.5: Contribution of functional value (F_{ij}) and social value (rarity R_i and meta group MG_i) to the (natural log of the) predicted price, organized per MetaGroup. Left: contribution according to the model in Equation 9.3. Right: contribution according to the model in Equation 9.4. Red and blue are the contributions of R_i and MG_i based on the estimated coefficients for these social quality parameters. Yellow is the part of the predicted module price that is unexplained by the social quality characteristics R_i and MG_i and is therefore attributed to functional characteristics F_{ij}.

9.9 Conclusions

The question of whether Hedonic Pricing Theory can predict the prices of items within a class of differentiated goods is ultimately about consumer rationality. In previous studies imperfect information, market failures, heterogeneous pref-

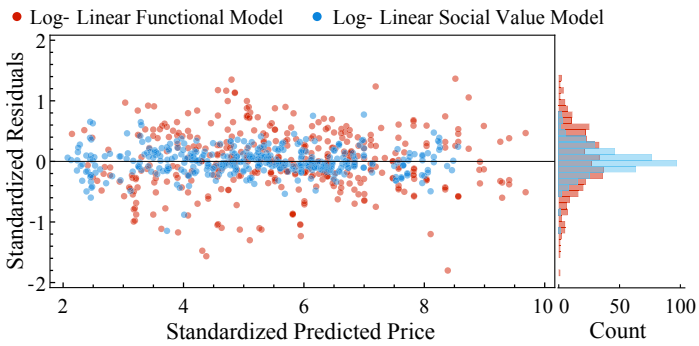


Figure 9.4: Comparison of the purely functional and social value log-linear models. Residuals of the log-linear model with (Equation 9.3) and without (Equation 9.2) social value parameters across all modules. The predicted values and residuals are standardized by dividing by the standard deviation within each module class. The addition of the social quality parameters clearly improves the accuracy of the model. On the right hand side we plot the full distribution of the standardized residuals of both models

erences, and other factors have complicated the proper empirical testing of HPT. We circumvent these shortcomings by analyzing economic data from a very rich and market-driven virtual world.

Our analysis reveals that consumers appraise the social quality of a product in the same general way as its functional quality: incremental increases in the social quality of a product lead to exponential increases in the price consumers are willing to pay for the good embodying those qualities. The extremely good fit of the log-linear social quality pricing model provides strong support for HPT as an explanation for the prices of products within a class of differentiated goods.

This has important implications for economic theory and practice. Economy theory suggests that firm investments in branding in part signal the firm's willingness to engage in a long term relationship with its clients and therefore in equilibrium reduces the information asymmetry about the product's quality on the part of the consumers, rendering them more willing to pay a price mark-up. This argument has lead observers to predict that the ongoing digital revolution, by reducing information asymmetries between firms and consumers, would erode the value of brands and shrink price mark-ups. We show that social value considerations make consumers value labels even if they have incentives to behave rationally and their information is complete. This may inform further theoretical advances on the deeper drivers of social value and suggests that top brands' price mark-ups may well survive the digital revolution, because their value hinges not only on the presence of information asymmetries, but also on social value.

Finally we may also conclude that consumers, if equipped with the proper information and incentives, tend to be very rational in their price decisions, as

long as we properly account for social quality in addition to the purely functional quality of a product. After all, humans are social animals, so it should come as no surprise that social value is real value, too.

References

[Akerlof and Kranton, 2000] Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753.

[Andersson, 2005] Andersson, H. (2005). The value of safety as revealed in the swedish car market: An application of the hedonic pricing approach. *Journal of Risk and Uncertainty*, 30(3):211–239.

[Atkinson and Halvorsen, 1984] Atkinson, S. and Halvorsen, R. (1984). A new hedonic technique for estimating attribute demand: An application to the demand for automobile fuel efficiency. *The Review of Economics and Statistics*, 66(3):417–426.

[Bagwell and Bernheim, 1996] Bagwell, L. S. and Bernheim, B. D. (1996). Veblen effects in a theory of conspicuous consumption. *The American Economic Review*, 86(3):349–373.

[Baker and Burnham, 2001] Baker, G. A. and Burnham, T. A. (2001). Consumer response to genetically modified foods: market segment analysis and implications for producers and policy makers. *Journal of Agricultural and Resource Economics*, pages 387–403.

[Bozdogan, 1987] Bozdogan, H. (1987). Model selection and akaike's information criterion (aic): The general theory and its analytical extensions. *Psychometrika*, 52(3):345–370.

[Brown and Ethridge, 1995] Brown, J. and Ethridge, D. (1995). Functional form model specification: An application to hedonic pricing. *Agricultural and Resource Economics Review*, 24(2):166–173.

[Bursztyn et al., 2017] Bursztyn, L., Ferman, B., Fiorin, S., Kanz, M., and Rao, G. (2017). *Status goods: experimental evidence from platinum credit cards*. The World Bank.

[Capraro and Rand, 2018] Capraro, V. and Rand, D. G. (2018). Do the right thing: Experimental evidence that preferences for moral behavior, rather than equity or efficiency per se, drive human prosociality. *Judgment and Decision Making*, 13(1):99–111.

[Castronova, 2004] Castronova, E. (2004). The price of bodies: A hedonic pricing model of avatar attributes in a synthetic world. *Kyklos*, 57(2):173–196.

[Charles et al., 2009] Charles, K., Hurst, E., and Roussanov, N. (2009). Conspicuous consumption and race. *The Quarterly Journal of Economics*, 124(2):425–467.

[Court, 1939] Court, A. T. (1939). Hedonic price indexes with automotive examples. *The Dynamics of Automobile Demand*, pages 98–119.

[Cronin et al., 2011] Cronin, J. J., Smith, J. S., Gleim, M. R., Ramirez, E., and Martinez, J. D. (2011). Green marketing strategies: an examination of stakeholders and the opportunities they present. *Journal of the Academy of Marketing Science*, 39(1):158–174.

[Deane et al., 1972] Deane, D., Hammond, K., and Summers, D. (1972). Acquisition and application of knowledge in complex inference tasks. *Journal of Experimental Psychology*, 92(1):20–26.

[Dimsom et al., 2015] Dimsom, E., Rousseau, P. L., and Spaenjers, C. (2015). The price of wine. *Journal of Financial Economics*, 118(2):431–449.

[Frank, 1985] Frank, R. H. (1985). The demand for unobservable and other nonpositional goods. *The American Economic Review*, 75(1):101–116.

[Gérard-Varet, 1995] Gérard-Varet, L. (1995). On pricing the priceless: Comments on the economics of the visual art market. *European Economic Review*, 39:509–518.

[Goodman, 1978] Goodman, A. C. (1978). Hedonic prices, price indices and housing markets. *Journal of Urban Economics*, 5:471–484.

[Hainmueller et al., 2015] Hainmueller, J., Hiscox, M. J., and Sequeira, S. (2015). Consumer demand for fair trade: Evidence from a multistore field experiment. *Review of Economics and Statistics*, 97(2):242–256.

[Halvorsen, 1988] Halvorsen, R. (1988). On the choice of functional form for hedonic price functions. *The Review of Economics and Statistics*, 70(4):668–675.

[Halvorsen and Pollakowski, 1981] Halvorsen, R. and Pollakowski, H. (1981). Choice of functional form for hedonic price equations. *Journal of Urban Economics*, 10:37–49.

[Han et al., 2010] Han, Y., Nunes, J., and Drèze, X. (2010). Signaling status with luxury goods: The role of brand prominence. *Journal of Marketing*, 74(4):15–30.

[Hoefman et al., 2018] Hoefman, K., Bramson, A., Schoors, K., and Ryckebusch, J. (2018). The impact of functional and social value on the price of goods. *PloS one*, 13(11).

[Hopkins and Kornienko, 2004] Hopkins, E. and Kornienko, T. (2004). Running to keep in the same place: Consumer choice as a game of status. *The American Economic Review*, 94(4):1085–1107.

[Ireland, 1994] Ireland, N. J. (1994). On limiting the market for status signals. *Journal of Public Economics*, 53:91–110.

[Kalafatis et al., 1999] Kalafatis, S. P., Pollard, M., East, R., and Tsogas, M. H. (1999). Green marketing and ajzen's theory of planned behaviour: a cross-market examination. *Journal of consumer marketing*, 16(5):441–460.

[Karelaia and Hogarth, 2008] Karelaia, N. and Hogarth, R. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin*, 134(3):404–426.

[Kaus, 2013] Kaus, W. (2013). Conspicuous consumption and "race": Evidence from south africa. *Journal of Development Economics*, 100(1):63–73.

[Lancaster, 1966] Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of political economy*, 74(2):132–157.

[Lee et al., 2013] Lee, W.-c. J., Shimizu, M., Kniffin, K. M., and Wansink, B. (2013). You taste what you see: Do organic labels bias taste perceptions? *Food Quality and Preference*, 29(1):33–39.

[Malpezzi, 2002] Malpezzi, S. (2002). Hedonic pricing models: a selective and applied review. *Housing economics and public policy*, pages 67–89.

[Oczkowski, 1994] Oczkowski, E. (1994). A hedonic price function for australian premium table wine. *Australian Journal of Agricultural and Resource Economics*, 38(1):93–110.

[Orrego et al., 2012] Orrego, M., Defrancesco, E., and Gennari, A. (2012). The wine hedonic price models in the 'old and new world': State of the art. *Revista de la Facultad de Ciencias Agrarias*, 44(1):205–220.

[Ready and Abdalla, 2005] Ready, R. C. and Abdalla, C. W. (2005). The amenity and disamenity impacts of agriculture: Estimates from a hedonic pricing model. *American Journal of Agricultural Economics*, 87(2):314–326.

[Renneboog and Spaenjers, 2013] Renneboog, L. and Spaenjers, C. (2013). Buying beauty: On prices and returns in the art market. *Management Science*, 59(1):205–220.

[Rosen, 1974] Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1):34–55.

[Sander and Polasky, 2009] Sander, H. and Polasky, S. (2009). The value of views and open space: Estimates from a hedonic pricing model for ramsey county, minnesota, usa. *Land Use Policy*, 26(3):837–845.

[Schultz et al., 2007] Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., and Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological science*, 18(5):429–434.

[Simon, 1986] Simon, H. (1986). Rationality in psychology and economics. *The Journal of Business*, 59(4):209–224.

[Sivanathan and Pettit, 2010] Sivanathan, N. and Pettit, N. (2010). Protecting the self through consumption: Status goods as affirmational commodities. *Journal of Experimental Social Psychology*, 46(3):564–570.

[Squire, 2008] Squire, K. (2008). Open-ended video games: A model for developing learning for the interactive age. *The ecology of games: Connecting youth, games, and learning*, pages 167–198.

[Sundie et al., 2011] Sundie, J. M., Kenrick, D. T., Griskevicius, V., Tybur, J. M., Vohs, K. D., and Beal, D. J. (2011). Peacocks, porsches, and thorstein veblen: Conspicuous consumption as a sexual signaling system. *Journal of personality and social psychology*, 100(4):664.

[Tappin and Capraro, 2018] Tappin, B. M. and Capraro, V. (2018). Doing good vs. avoiding bad in prosocial choice: A refined test and extension of the morality preference hypothesis. *Journal of Experimental Social Psychology*, 79:64–70.

[Thurner et al., 2012] Thurner, S., Szell, M., and Sinatra, R. (2012). Emergence of good conduct, scaling and zipf laws in human behavioral sequences in an online world. *PLOS ONE*, 7(1).

[Triplett, 2004] Triplett, J. (2004). *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes*. OECD Publishing.

[Veblen, 1899] Veblen, T. (1899). *The theory of the leisure class: An economic study of institutions*. Unwin Books.

[Werner, 1987] Werner, J. (1987). Arbitrage and the existence of competitive equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1403–1418.

[Yiridoe et al., 2005] Yiridoe, E. K., Bonti-Ankomah, S., and Martin, R. C. (2005). Comparison of consumer perceptions and preference toward organic versus conventionally produced foods: A review and update of the literature. *Renewable Agriculture and Food Systems*, 20(4):193–205.

Chapter 10

Summary, Outlook and Future Research

10.1 Summary

The work presented in this dissertation can be divided into a theoretical, a methodological, and an empirical part. Although the different chapters study a variety of research questions from both economics and social physics, they are bound together by the central thesis of Live Agent-Based Models: that virtual worlds can help us advance our understanding of people's individual behavior, as well as the complex interaction between individuals, in turn leading to better specifications for the microfoundations of macroeconomic models.

In the first theoretical chapter, we explored the state of the art of macroeconomic modeling, focusing in particular on three families of macroeconomic models: econometric, DSGE, and agent-based models. We considered whether macroeconomic models should be realistic, and whether they should be founded on microeconomic principles, so-called microfoundations. We retraced the history of macroeconomic modeling, in order to better understand the principles, strengths and weaknesses of all three types of models. We pointed out that econometric models by design do not implement microfoundations, and that as a consequence, these models cannot be used for long-term predictions, as they do not account for structural changes in the economy resulting from people's adaptations to policy changes. DSGE models and agent-based models appear to be microfounded, but closer scrutiny, and recent advances in our understanding of human cognition and the complex nature of human interaction, reveals that these microfoundations are unrealistic and/or based on an incomplete understanding of human decision-making and interaction.

In the second and third theoretical chapters, we substantiated our claim that

the current understanding of human decision-making, and the complex interaction between people, is too incomplete to provide a reliable foundation for macroeconomic models. In chapter two, using the Rational Utility Maximization Hypothesis as a starting point, we showed that even though scientists have for a long time understood that rationality is bounded, dual-system understanding of human cognition has only very recently become accepted in mainstream cognitive psychology. What new insights the "emerging science of heuristics" will lead to remains to be seen, but it is clear that without a proper understanding of how individuals actually make decisions, founding macroeconomic models on realistic human cognition will remain problematic. In the third chapter, we showed that while people tend to think in terms of linear relationships and normal distributions, many interesting phenomena, including human economic and social interaction, are characterized by non-linear behavior and "fat tail" distributions. So far, the study of complexity has not resulted in reliable macroeconomic agent-based models, perhaps because agent-based models are themselves complex non-linear systems. Due to the inherent non-linearity between the accuracy of the underlying assumptions and the emergent outcome of a complex system, we can't know how close we really are to accurate macroeconomic agent-based modeling. All we can do is keep looking.

In the fourth theoretical chapter, we proposed a novel methodology for studying human decision-making behavior, and the dynamics of the complex interaction between people. By treating virtual worlds as data-generating processes, we provide a bridge between experimental economics and real-world data. "Live Agent-Based Models" offer the best of two worlds: precise data generation, characteristic of agent-based models, driven by the behavior of "live" agents: real people. We suggested three domains of application: empirical validation of agent-based models, LAB-M bootstrapping for non-ergodic processes, and studying the virtual worlds data to enhance our understanding of human decision-making and interaction by testing existing theories, and developing new hypotheses.

In the methodological chapter that followed, we documented the software we created to study the propagation dynamics of temporal webs. We suggested two ways of optimizing memory use in the case of large networks: a State Machine implementation, applicable to temporal networks in general, and a second optimization based on the topology of our network. We also suggested multi-threaded network analysis as a way to increase analysis speed.

In the empirical chapters, we applied Live Agent-Based Models to research questions from economics and social physics. In the first empirical chapter, we studied propagation through temporal networks using Eve Online alliance data, interbank loans data from the Russian banking system, and Twitter posts regarding the H1N1 vaccine, in an approach we called temporal influence abduction. This publication contributed towards a better understanding of the methodology of temporal webs in general, and the nature of cascades through temporal webs in

particular.

In the second empirical chapter, we analyzed the dynamics of geopolitical relations, by applying variants and extensions of Structural Balance Theory to Eve Online alliance data. By transforming the alliance relationships network into a triad temporal network, we were able to study the propagation of coalition changes through time, providing strong support for Structural Balance Theory, and the use of virtual worlds data in social science research.

In the third empirical chapter, we investigated the long-standing question of the relationship between the quality of a product and the price people are willing to pay for it. By applying Hedonic Pricing Theory to Eve Online market data, we provided strong evidence indicating that people pay exponential price increases for linear increases in quality. We also revealed that people pay large surplusses for benefits that are purely social in value, raising the question of what is rational pricing behavior.

10.2 Outlook and Future Research

As a final step, we look ahead by discussing the ongoing and potential future research connected to the work presented in the preceding chapters.

The empirical chapters presented in this study are just a sample of the potential of live agent-based models as a research bed for studying economic and social theories. This potential is limited only by the researcher's ability to identify parallels for real-life behavior within large-scale virtual worlds. Herein lies somewhat of a subtle challenge: in our experience, it takes familiarity with virtual worlds and their mechanics to be able to distinguish between a useful match and a spurious similarity. We will refrain from advising researchers to play more massively multiplayer online computer games, yet there is truth in the saying that you can't recognize what you don't know.

In a follow-up paper to the work on Hedonic Pricing Theory from chapter 9, we are investigating a new theory, which we call Heuristic Pricing Theory, as an alternative explanation for the pricing behavior observed in Eve Online markets. Harking back to Herbert Simon's seminal work on heuristics, we propose that, instead of precisely quantifying every quality trait of a good - a process impossible in most real-world situations due to imperfect information - customers use a two-step heuristic process for determining price: intuitively ranking competing goods by assigning them to discrete levels on a heuristic quality scale, followed by attributing exponential price increases to each level of the scale, regardless of exact quantification. Our preliminary results are promising: statistical measures for goodness of fit of the heuristic model rival those we published for the hedonic pricing model. However, the question of what determines the slope of the exponential, which is different for each type of good, needs further conceptualisation.

Other avenues for further research include building on LAB-M's potential for empirical validation of agent-based models, as well as for bootstrapping agent-based models for non-ergodic systems.