Statistical Equilibrium Methods in Analytical Political Economy

Ellis Scharfenaker

Statistical Equilibrium Methods in Analytical Political Economy

Ellis Scharfenaker
Department of Economics, University of Utah
ellis.scharfenaker@economics.utah.edu

Abstract

Economic systems produce robust statistical patterns in key sate variables including prices and incomes. Statistical equilibrium methods explain the distributional proper- ties of state variables as arising from specific institutional and behavioral postulates. Two traditions have developed in political economy with the complementary aim of conceptualizing economic processes as irreducibly statistical phenomena but differ in their methodologies and interpretations of statistical explanation. These conceptual differences broadly mirror the methodological divisions in statistical mechanics, but also emerge in distinct ways when considered in the context of social sciences. This paper surveys the use of statistical equilibrium methods in analytical political economy and identifies the leading methodological and philosophical questions in this growing field of research.

1   Introduction

Market based economic systems are complex decentralized systems that give rise to spontaneous order in the formation of statistical moments such as average prices and incomes. Economic outcomes are statistical in nature and typically consist of central tendencies as well as fluctuations around these tendencies. The fact that economic systems produce stable regularities that persist for extended periods of time permits theory formation around the explanation and prediction of these regularities. Economic theory attempts to explain the existence, location, and evolution of these modalities in terms of individual and collective actions within a framework of social and economic institutions. Statistical variation in economic outcomes poses several conceptual issues for economic theory. The first deals with the nature of statistical fluctuations and whether or not they are endogenous to the system and thus in need of explanation, or if they are exogenous stochastic variation irrelevant to theory formation. The second difficulty concerns the appropriate methodological approach to deal with statistical variation in economic outcomes at various levels of abstraction and aggregation.

Orthodox economic theory born out of the marginalist doctrines of Walras and Marshall, as well their modern descendant general equilibrium theory, tends to abstract away from statistical thinking instead emphasizing the determination of stable fixed-point equilibrium properties of an abstract completely specified economic system. Neo-Ricardian and Marxian political economy has also tended to reason in terms of abstract notions of uniform equilibrium prices and incomes despite adopting an essentially statistical vision. Real experience with markets and analysis of market data, however, reveals that economic variables such as incomes and commodity transaction prices exhibit significant variation over time and space giving rise to statistical distributions of these variables. The reduction of the statistical distribution of prices to the modal or mean price simplifies economic theory greatly, but perforce neglects the existence of disequilibrium states and the determination of higher order moments that leaves many fundamental questions about the determination of equilibrium unresolved.

In contrast to orthodox theory, however, the classical political economy of Smith, Ricardo, and Marx recognized the essentially statistical nature of economic processes and used it as foundation for conceptualizing equilibrium. Smith was the first to articulate a dynamic theory to account for a tendential equalization of prices and incomes and most classical political economists in the nineteenth century adopted some notion of "gravitational" equilibrium. The equilibrium rate of profit and equilibrium wage rate were the centers of gravity that reflected a uniform rate of profit for capital and uniform wage rate for workers. These

abstractions were then used as basis for explaining other economic phenomena, including prices, growth, distribution, social conflict, and economic policy. From the classical perspective, equilibrium was always treated as a dynamic gravitational process determined by strong stabilizing feedbacks arising from various institutional structures. Explaining the endogenous variation of prices and incomes was as essential to economic theory as was explaining the central tendencies. Marx was often keen to emphasize that a capitalist economy in absence of these endogenous fluctuations would cease to function altogether.

While essentially adopting a complexity perspective of economic systems (Foley, 2003; Scharfernaker & Yang, 2020), classical political economists were unequipped with the modern scientific tools used to study complex many-bodied systems resulting in fragmented analytical methods discordant with their theoretical outlook (Farjoun & Machover, 1983; Langston, 1984). The issues surrounding the transformation of values into prices is perhaps the best illustration of the kind of difficulties that arose due to the dissonance between theory and methods working at different levels of abstraction and aggregation in classical political economy.

Recent advances in the application of statistical mechanics and information theory to economics has revealed the new directions for the range of applicability and validity of classical political economic theory. These methods identify the equilibrium of a system's state variables as irreducibly statistical and thus driven to equilibrium probability distributions. The statistical equilibrium approach identifies the stable configuration of these distributions as revealing important information about the relevant institutional and behavioral constraints that produce both the system's central tendencies as well the generative dynamics that shape the higher moments of these distributions.

The statistical equilibrium approach fills a lacuna in contemporary political economy by undertaking a complete statistical description of social and economic phenomena. This approach is squarely aligned with the theoretical and methodological principles of classical political economy because it explicitly models the fluctuations in state variables such as prices, profit rates, and wages, that arise from the decentralized nature of market activity and the purposive behavior of individuals, such as workers and capitalists, interacting in complex non-additive ways within the institutions of capitalism. While statistical equilibrium thinking has had a productive effect in political economy in general, there is significant disagreement on the methodological basis and interpretation of statistical explanation. Two traditions have emerged in statistical political economy that broadly parallel those in statistical mechanics: the ergodic camp which reasons in terms of infinite time averages and the macroscopic camp that reasons in terms of system-level information. These methodologies share many of the same philosophical conundrums and dichotomies that exist in statistical

3

mechanics, but also pose many distinctive and unresolved questions in their application to social sciences. This paper surveys the role of statistical equilibrium thinking in analytical political economy as well as the many unresolved interpretive and philosophical issues attendant with this growing field of research.

2 Statistical Equilibrium

Statistical equilibrium methods in physics were developed in the late 19th century by Boltzmann (1871), Maxwell (1860) and Gibbs (1902) and were originally applied to relatively simple thermodynamic problems such as deriving the phenomenal laws of gas behavior from the hypothesis that gases were constituted from particles such as molecules and atoms. Statistical mechanics uses known microscopic properties of systems in order to make probabilistic and statistical assertions about the macroscopic behavior of systems. The key theoretical breakthrough in statistical mechanics is the idea of ensemble reasoning. The ensemble represents a collection of all conceivable microscopic configurations of a system's N members across all the possible individual states each one of them may take. Which microscopic configurations are possible depends on the laws governing the functioning of the system, such as Liouville's equation that describes its Hamiltonian dynamics and thermodynamic laws such as the conservation of energy. Each microscopic configuration generates a unique macroscopic state or frequency distribution describing the relative occupancy of each individual state across all members of the system. The entropy of a macroscopic state is simply a combinatorial measure quantifying how many different microscopic configurations support any given macroscopic state. The distribution or macroscopic state with maximum entropy across all possible distributions is simply the one with the greatest support across the microscopic configurations allowed by the functioning of the system. In systems where $N >> 1$, the combinatorial dominance of that distribution is overwhelming, effectively ensuring that no matter where a system's evolution started, it will statistically tend toward microscopic configurations supporting that distribution—which we understand as a statistical equilibrium macroscopic state.

Physicists in the twentieth century (Jaynes, 1957) discovered that these same methods can be applied fruitfully to a wide range of systems including biological and social systems (Kapur, 1989; Kapur & Kesavan, 1992) particularly when viewed through the theoretical scope of information theory (Caticha, 2007; Golan, 2018). The goal-oriented purposive behavior of individuals in social situations where people interact through specific institutions, like private property, are from a statistical modeling point of view exactly like the constraints of physical laws that characterize the statistical properties of physical systems. Statistical

equilibrium modeling in economics, however, does not imply that social systems are represented by a "social physics" subject to immutable "social laws," that individual social beings are equivalent to purposeless particles, or that social science is ultimately reducible to the physical sciences. Statistical equilibrium methods in economics attempts to identify the relevant systemic constraints acting on a system at any point in time that shape observed distributional outcomes of social and economic interactions.

While the economy as a whole is not in any kind of stable equilibrium most agents do traverse similar dynamic paths over and over again creating a type of "punctuated equilibrium," that may persist for extended periods of time. Social and economic institutions shape individual actions and social interactions in remarkably robust ways. Capitalist institutions of private property and market mediated exchange produce imperatives for labor to work under relatively uniform conditions which inevitably shapes domestic habits as well. While there may be significant heterogeneity in the type of work done, individuals tend not to deviate far from their average patterns. The same can be said of firms producing, distributing, and marketing products, which also tends to be highly routine.[1]

Capitalism also creates systemic regularities *independently* of the details of micro-level behavior, and often in spite of the intentions governing that behavior. These systemic regularities can be broadly understood as reflecting the fact that markets generate prices that reduce all details to simple quantifications, and that all individuals interact interdependently though social and economic institutions. There is an important distinction and interplay between the behavioral and systemic constraints that determine the macroscopic characteristics of capitalist systems. Statistical equilibrium methods allow one to gain insights about the functioning of the system from understanding the routine and repetitive aspects of reality rather than by focusing on the much less frequent and much less structured periods of transition. This is not to downplay the importance of transitional periods, which can be highly disruptive to and fundamentally transformative of previously stable economic patterns, but the routine patterns of economic life that are invariably shaped by social and economic institutions are likely to reassert themselves, even if in novel ways.

The probabilistic and statistical nature of statistical mechanics implies that it is inextricably tied to the conceptual issues associated with the interpretation of probabilities. Understanding exactly what is being asserted in a statistical equilibrium model will depend on the general conception of the meaning of statistics and probability. Due in part to these unresolved issues, two separate traditions developed in statistical mechanics with complementary aims but distinct methodologies which Hobson (1971) refers to as the *ergodic* and

---

[1]Punctuated equilibrium is treated in non-equilibrium thermodynamics as local or momentary equilibrium arising in systems undergoing small changes to macroscopic parameters.

macroscopic approaches.

The ergodic view of statistical mechanics follows the perspectives of Maxwell and Boltzmann and derives the macroscopic behavior of a system from the evolution of micro-kinetic behavior of weakly-interacting particles. Methodologically, the ergodic approach studies the infinite time averages of state-space variables. The macroscopic approach to statistical mechanics uses known microscopic properties to study the macroscopic behavior of many-body systems based on probabilistic reasoning alone. This view follows the insights of Gibbs and later Jaynes (Jaynes, 1967; Jaynes, 1957) who identified Shannon (1948) information theory as a logical basis for making any probabilistic assertions about incompletely specified systems.

The lack of a universally accepted account of the theory of statistical mechanics[2] suggests that while these methods may be useful in understanding complex economic systems, there is little expectation that economists will agree on a single correct interpretation and methodology in their applications of statistical mechanics to economics or even find agreement as to what an acceptable statistical explanation amounts to.

3   Statistical Thinking in Political Economy

Statistical thinking was foundational to classical political economic theory. Adam Smith constructed his economic theory of natural prices upon a "gravitational" equilibrium logic that attempted to explain both the central tendencies of prices as well as the ceaseless and endogenous fluctuations that generate these tendencies as statistical moments. In his discussion of value and prices, Smith makes an important methodological distinction between natural price and market price of a commodity. The market prices are the amount of money for which any particular commodity exchanges at any particular moment. These prices fluctuate in time and space due to persistent imbalances of supply and demand. Such fluctuations in prices and incomes were understood as a natural outcome of the perennial difficulties of decentralized and specialized production meeting the material and social needs of reproduction. Central to his line of reasoning was that underlying the complex process of production and exchange was a common organizing logic of competition that generated negative stabilizing feedbacks that regulated market prices around natural prices. Marx adopted Smith's notion of gravitational equilibrium and stressed the statistical and combinatorial logic of natural prices that made them an essentially macroscopic phenomenon.

---

[2]As Jaynes notes, "statistical mechanics has become a queer hybrid, in which the practical calculations are always based on the methods of Gibbs; while in the pedagogy virtually all one's attention is given to repeating the arguments of Boltzmann."(Jaynes, 1967, pp. 87). See Sklar (1993) for a good overview of the historical and philosophical issues in foundations of statistical mechanics.

While classical political economy tended to reason in terms of a statistical methodology, no formal analytical methods consistent with this reasoning were advanced. Neo-Ricardian (Kurz & Salvadori, 1995) and Marxian traditions (Shaikh, 2016) have also avoided explicit statistical consideration in theory formation instead favoring analysis based on the abstract long-run equilibrium behavior of capitalist economies where all higher moments are assumed away.

In the 1980s Farjoun and Machover (1983) (FM) heavily criticized the paradigm in these traditions arguing that the impossibility of uniform prices, wages, and profit rates severely limited the interpretive and explanatory aims of Neo-Ricardian and Marxian theory. Instead, they argued that prices and incomes should be treated as random variables that are driven not to a deterministic uniformity, but to time-invariant stationary probability distributions with general forms that are "theoretically ascertainable" and "empirically verifiable." FM's work was a groundbreaking attempt to construct a non-deterministic theoretical framework for the foundations of political economy that avoided an equilibrium logic unable to account for the observed distribution of economic variables. FM called for an explicit statistical methodology in political economy consistent with the central themes of Smith, Ricardo, and Marx and pointed towards statistical mechanics as satisfying this methodological criterion. While FM offer a compelling argument for the relevance of statistical equilibrium methods in political economy, their particular application of these methods to political economic theory may have raised more questions than they answered (Scharfenaker & Semieniuk, 2017).

A diverse range of recent contributions have since renewed interest in statistical equilibrium approaches in political economy and significantly extend the aims of FM. These contributions are complimentary in their aims of introducing an explicit statistical methodology to political economy, but differ in their methodological prescriptions (dos Santos, 2020; Golan, 2018; Reddy, 2020; Scharfernaker & Yang, 2020). One tradition studies the parametric statistical distributions of economic outcomes explained as ergodic solutions to stochastic processes. This tradition has a long history in economics dating back to Gibrat (1931), Kalecki (1945) Champernowne (1953) and Simon (1955) who looked to formalize the empirical findings of Pareto (1987a, 1987b) on the distribution of personal income and wealth and has experienced a resurgence in political economy (Alfarano, Milaković, Irle, & Kauschke, 2012; Cottrell, Cockshott, Michaelson, Wright, & Yakovenko, 2009; Shaikh, 2020; Shaikh, Papanikolaou, & Wiener, 2014) with the development of econophysics (Lux, 2016; Mantegna & Stanley, 1999; McCauley, 2009; Rosser Jr., 2008b; Yakovenko, 2007). The econophysics approach to political economy argues in terms of well-specified micro-kinetic dynamics punctuated by stochastic variation in order to derive the statistical equilibrium of the system as the limiting ergodic distribution. Specifically, "the time evolution of an economic system

7

is represented by an aperiodic, irreducible Markov chain and the distribution of relevant quantities is given by the invariant distribution of the Markov chain."(Garibaldi & Scalas, 2010, pp.224) This ergodic approach is most often adopted in economics and finance through a class of stochastic models called drift-diffusion models which identifies the sources of statistical effects in random disturbances. While no explicit position is ever really taken on the origins of these disturbances, that is, no specifically tychist or external interventionist position is taken, the ontological interpretation of randomness is almost always philosophically implicit and methodologically explicit.

The other tradition adopts the complexity perspective of social systems but identifies the complex nature of such systems as arising from incomplete information (dos Santos, 2020; dos Santos & Scharfenaker, 2019; dos Santos & Wiener, 2019; Foley, 1994; Scharfenaker & Foley, 2017; Scharfenaker & Semieniuk, 2017; Scharfernaker & dos Santos, 2015; Scharfernaker & Yang, 2020; Yang, 2018a, 2018b). This approach originates in the ideas of Edwin T. Jaynes (Jaynes, 1957) who identified the generality of Claude Shannon's (Shannon, 1948) work in information theory as a justification for the study of statistical mechanics. Jaynes emphasized the inferential nature of all real scientific problems as arising from incomplete information and identified entropy as a measure of information that is to be understood as a description of a state of knowledge about a system. Jaynes' Maximum Entropy Principle (MEP) derives the statistical equilibrium distribution by maximizing Shannon entropy $\mathcal{H} = -\int f(x) \log f(x) dx$ subject to the normalization of probabilities $\int_x f(x) dx = 1$ and any other information one has about the relevant constraints acting on the system. The statistical equilibrium distribution is thus the combinatorially most probable and least biased representation of the system's state compatible with the imposed constraints. No assumptions about ergodicity or deterministic micro-kinetic dynamics are necessary. Rather, this approach emphasizes the combinatorial and Laplacian logic of making inferences when faced with incomplete information about the detailed dynamics of a system. From this perspective, randomness is always understood as epistemological representation of our lack of knowledge about a particular phenomenon. The maximum-entropy principle of inference is thus a general principle for which statistical mechanics is one particular application.

4   Ergodic Approaches

4.1   Background

The ergodic approach to studying the distributional outcomes of economic variables reasons in terms of well-defined micro-kinetic processes that describe the general behavior of

individual agents coupled with ongoing random transitions between states. In early thermodynamic theory the ergodic approach begins with a single particle in a multi-dimensional space in which a single point represents the momentum and position of the particle at a particular time. The evolution of this point through the phase space is taken as a representation of the entire system and the equilibrium properties of the system's state variables are computed by calculating the infinite time averages. In modern thermodynamics probabilities are not attached to the state of a single particle but rather to the state of the entire system. Thus, the system's variables can be consider as stochastic variables and the interactions between particles can be taken into account.

The ergodic viewpoint locates the meaning of the probability of the system to be in a particular phase space configuration as the fraction of time the system spends in that phase to the total (infinite) time it spends in all phases.[3] Characteristically this is done by coarse-graining the phase space into a set of discrete states and then postulating a constant probabilistic law describing the transitions between states. For this reason the ergodic perspective makes statistical and probabilistic assertions about measurable frequencies and that what is measured is a time average "long enough" for the equilibrium properties of the system to be established. How long is long enough is an endless source of controversy. Fluctuations arise by superimposing perpetual "statistical randomness" to the micro-dynamical laws that are posited as governing the system. Despite the considerable differences in the interpretations and philosophical perspectives on the foundations of non-equilibrium statistical mechanics, there does seem to be unification around a common set of mathematical methods that Uffink (2007) refers to as the framework of *stochastic dynamics*. Stochastic dynamics replace a deterministic dynamical characterization of the evolution of the state of system for an explicit probabilistic law of evolution and have been widely adopted in economics and finance and more recently in political economy. In economic applications the economic process is frequently modeled with stochastic differential equations, typically Itô processes driven by Brownian motion of the form:

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t \tag{1}$$

Here X_t is a state vector of process variables, W_t is Wiener process, or Brownian motion vector, μ is vector-valued drift-rate function, and σ is a matrix-valued diffusion-rate function. The Wiener process ensures that as a continuous time stochastic process X_t only undergoes small normally distributed changes and is independent of its past behavior ensuring

---

[3]Boltzmann's motivation for the ergodic approach was in part "to eliminate the seeming arbitrariness of the probabilistic hypothesis used earlier to derive equilibrium features."(Sklar (1993) pp.45)

the Markov property. This assumption of fixed transition probabilities leads, along with a "master equation" such as the Fokker-Plank equation, to the kinetic equations describing the monotonic movement to the equilibrium stationary distribution of the state vector X_t. The Fokker-Plank equation for the density $f(x, t)$ of the random variable X_t is:

$$\frac{\partial}{\partial t} f(x, t) = -\frac{\partial}{\partial x} \mu(x, t) f(x, t) + \frac{\partial^2}{\partial x^2} D(x, t) f(x, t) \tag{2}$$

where the diffusion coefficient $D(x, t) = \sigma(x, t)^2/2$. The stationary distribution is defined as one for which $f(x, t)$ is constant over time and is reached as $t \to \infty$ for all initial conditions. The stationary distribution $f^*(x)$ is the solution to the Fokker-Plank equation when $\frac{\partial}{\partial t} f(x, t) = 0$, implying the following condition:

$$\frac{\partial}{\partial x} \mu(x) f^*(x) = \frac{\partial^2}{\partial x^2} D(x) f^*(x) \tag{3}$$

Such stochastic drift-diffusion models represent the thermalization (equilibration) process for a single particle. Because the equilibrium equation is a single equation defined by two unknowns, solving for $\mu(x)$ and $D(x)$ is an underdetermined problem. This underdetermination is customarily dealt with by fixing a constant diffusion $D(x) = D$ which eliminates one degree of freedom. Methodological and interpretive issues with the ergodic approach to the foundations of statistical mechanics are discussed in Callender (1999), Sklar (1993) and Redhead (1995). These issues, as they relate to the study of economic systems, are discussed below in section 6 as well as in dos Santos (2020).

4.2   Applications to Political Economy

There have been several important contributions since Farjoun and Machover extending ergodic statistical equilibrium reasoning to political economy.[4] One of the most novel papers that emerged in the last ten years that applied these statistical equilibrium methods to political economy was Alfarano et al. (2012). This paper recast a long tradition in the industrial organization literature on firm growth rates (Bottazzi & Secchi, 2006; Bottazzi, Li, & Secchi, 2019; Dosi, 2007; Ijiri & Simon, 1977; Simon & Bonini, 1958; Steindl, 1965) in terms of classical political economy. Alfarano et al. propose from a classical perspective that the persistent entry and exit of firms in search of the highest rate of profit generates a negative feedback that stabilizes the profit rate distribution into a statistical equilibrium

---

[4]See Föllmer (1974) for an early ergodic approach to modeling equilibrium in economic systems.

distribution. They initially derive the model by applying Jaynes' maximum entropy principle and attempt to map entry and exit behavior of firms to a single moment constraint by interpreting competition among capitalist firms as generating a constraint on the dispersion of profit rates around the average. The MEP with this constraint is expressed as:

$$\begin{aligned}
&\underset{f(x) \geq 0}{\text{Max}} - \int f(x) \log f(x) dx \\
&\text{subject to} \int f(x)\, dx = 1 \\
&\text{and} \int f(x) \left| \frac{x-\mu}{\sigma} \right|^{\alpha} dx = 1
\end{aligned} \quad (4)$$

The solution is given by the exponential power or Subbotin distribution (Subbotin, 1923), which is a three three parameter unimodal symmetric distribution capable of generating a wide range of distributions including the Normal (for $\alpha = 2$) and Laplace ($\alpha = 1$):

$$f(x) = \frac{1}{2\sigma\Gamma(1+1/\alpha)} e^{-\frac{1}{\alpha} \left| \frac{x-\mu}{\sigma} \right|^{\alpha}} \quad (5)$$

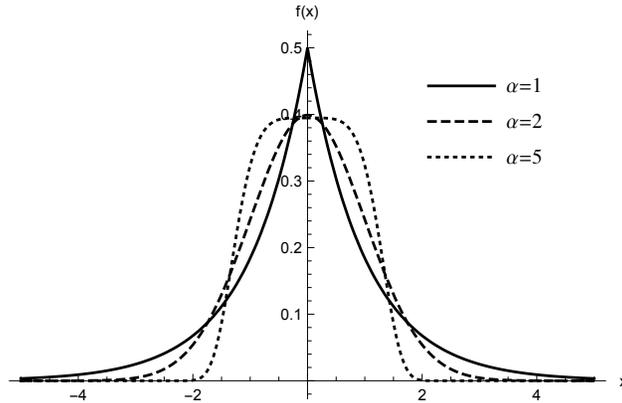and appears in Figure 1 for a range of parameters.

Figure 1: Subbotin distribution with $\alpha = 1, 2, 5$, $\mu = 0$, and $\sigma = 1$.

The authors argue, however, that the principle of maximum entropy does not permit them to interpret the model parameters in any economically meaningful way and thus turn towards stochastic methods. They introduce a stochastic drift-diffusion model capable of generating a stationary Subbotin distribution on which they base subsequent research (Mundt, Alfarano,

11

& Milaković, 2016, 2020) and is of the form:

$$dX_t = \frac{D}{2\sigma}\text{sign}(X_t - \mu)\left|\frac{X_t - \mu}{\sigma}\right|^{\alpha-1} dt + \sqrt{D}dW_t \tag{6}$$

Assuming zero correlation between idiosyncratic shocks and the individual firm's profit rate fixes the diffusion function as a constant $D(x) = D$ and permits an expression of the drift function, which is used as the basis for interpretation of the model parameters. The empirical fits to firm-level data for long-lived firms are compelling and offer solid evidence against the normality of the profit rate distribution. While Alfarano et al. (2012) provides a convincing argument for looking at capitalist competition as a basis for understanding the generative mechanism behind the statistical equilibrium profit rates and growth rates and a parsimonious model to predict profit rate distributions, several of their claims and assumptions are worth considering in critical detail.

First, their claim that the principle of maximum entropy does not permit them to interpret the model parameters in any economically meaningful way is a consequence of their model choice and is not true for of maximum entropy models in general (see Golan (2018) for a thorough treatment on maximum entropy models). Typically, the parameters of a maximum entropy distribution correspond the the Lagrange multipliers associated with each constraint and thus have a direct interpretation as the marginal amount of information a constraint contributes to the reduction of the entropy of the statistical equilibrium distribution. From an economic perspective, the maximum entropy problem is exactly like the type of constrained optimization problems one routinely encounters in microeconomic theory and thus the model parameters can be understood readily as the "shadow-price" representing the constraint. If the constraints are clearly articulated from economic theory the parameters of the model should be directly interpretable.

Second, it is not clear that the stochastic methods align with the authors theoretical position. They emphasize that the non-normality of the profit rate distribution violates the assumption of statistical independence leading them to consider the interactions of capitalist firms instead of the detailed characteristics of the individual firms. But it is unclear whether the interactions of competing capitalists are well captured by an analogy to the collision of weakly-interacting particles in a thermodynamic system. There is good reason to believe competition among firms generates rather complex interactions and long-ranging interdependencies of firm profit rates which violate Markov and multinomial assumptions (see 6 below).

Lastly, the authors ultimately put themselves in an awkward and unnecessary position

by conflating the unobservable entry and exit of capital in specific markets with the observable birth and death of firms in a particular dataset. It is not clear that the birth and death of firms is theoretically relevant to the argument of profit rate equalization. A single firm can, and for large established firms almost always do, contemporaneously compete in many different markets. Because firm balance sheets do not capture the detailed reallocation of capital from one of line of production to another the entry and exit of capital is unobservable. It might also be argued that the theoretically relevant notion of a "market" makes the detailed dynamics of the allocation of capital unobservable in principle. This point is important because it underscores the sources of incomplete information that make the economic problem underdetermined and ill-posed. If the observable birth/death of firms determines the equalization of profit rates then the problem of inferring the joint statistical equilibrium distribution (from which the marginal distribution of profits rates can be derived) is not underdetermined (see Scharfenaker and Foley (2017)). Instead of identifying the underdetermination of the problem in entry/exit the authors shift the focus to the unobservable interactions of competing firms which understandably leads them to consider stochastic explanations.

Similar to Mundt et al. (2020), Shaikh and Jacobo (2019) and Shaikh (2020) take an explicit ergodic perspective to profit-rate and wage-rate equalization through the use of stochastic models. Like Alfarano et al. Shaikh posits a negative feedback mechanism that stabilizes incomes and profit rates and attributes the stochastic element to positive and negative shocks. Importantly noting that these shocks ensure that intentions of individual workers and capital need not be realized. To represent the process of turbulent arbitrage Shaikh attributes the drift term in a stochastic equation to the entry and exit movements induced by differences in variables (i.e. wage and profit-rate differentials) that leads to the equalization of the variables to their means and the diffusion term to the effects of ongoing shocks that leads to a persistent distribution of the variables.

Regarding profit rates, he argues that a good representation of the inter-industry competitive process is captured by the regulating capital which is well represented by the linear Ornstein–Uhlenbeck process:

$$dr_t = -\theta(r_t - 1)dt + \sqrt{\sigma}dW_t \tag{7}$$

This stochastic equation generates a stationary Gaussian distribution. In contrast to Alfarano et al. Shiakh looks at industry profit rates and argues for the normality of the statistical equilibrium profit rate distribution based on Eq. 7. Following Yakovenko and

Silva (2005) and Shaikh et al. (2014) Shaikh identifies the distribution of income as mixture of distributions, one representing the bulk of labor income and the other property income, representing the tail of the distribution. He argues that the same classical logic of profit rate equalization through turbulent arbitrage can explain wage rate equalization, but that the strict positivity of labor incomes requires the use stochastic models that are logarithmic in wages w_t. He offers two such models, one that is log-linear and the other that is log-log form:

$$
\begin{aligned}
d\log w_t &= -\theta(w_t - 1)dt + \sigma dW_t \\
d\log w_t &= -\theta \log w_t dt + \sigma dW_t
\end{aligned}
\tag{8}
$$

These equations give rise to Log-Normal and Gamma stationary distributions. For property income $\pi = a\rho$ which is equal to the product of the stock of financial assets a and the rate of return ρ he derives the log-linear and log-log models:

$$
\begin{aligned}
d\log \pi_t &= (\gamma - \theta)(\rho_t - 1)dt + \sigma dW_t \\
d\log \pi_t &= -(\gamma - \theta)\log \rho_t dt + \sigma dW_t
\end{aligned}
\tag{9}
$$

These equations give rise to the Log-Normal and Power Law stationary distributions. While Shaikh does not commit to a specific parameterization, he notes that both forms appear to fit the profit rate and income data reasonably well. As a methodological point, Shaikh argues in favor of the ergodic approach because it permits prediction of economic phenomena as the result of explicit deterministic economic processes. While the stationary distributions turn out to be entropy maximizing, there is no good reason to think entropy maximization should be an aim in itself. Reddy (2020) favors this position as well.

Again, the question remains whether stochastic differential equations represent the relevant economic processes they are posited to and whether the conditions that sustain the statistical equilibrium should be taken into explicit consideration in the relevancy of the theory. From this perspective complex economic processes such as profit rate equalization is presented in a highly reduced form that attributes all "interactions" to a stochastic Wiener process. The drift-diffusion model postulates mean reversion and persistent dispersion driven by small independent Normally distributed memoryless shocks acting upon an individual particle. Statistical equilibria defined by drift-diffusion models thus requires all individuals follow trajectories that over time trace the same distributional form as the population cross section. The logical implication is the reduction of the functioning of the system to an

ensemble of equivalent individuals collectively doing the same thing. This approach immediately raises the question about individualist reductionism and the actual social content of the theory (dos Santos, 2020).

5   Macroscopic Approaches

5.1   Background

The macroscopic approach dispenses away with notions of ergodicity in favor of probabilistic foundations for inferring the statistical equilibrium distribution. Formally the macroscopic approach begins with Boltzmann's original derivation of entropy as a combinatorial measurement of the number of micro-states consistent with macroscopic constraints, but was fully developed by Gibbs as a radically different view from the ergodic perspective.

Consider a system with N individual components each defined by ν states, which may be observable or unobservable. A complete state-space description would be $\mathbf{X}_{\{N,\nu\}} = \{\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^\nu\}$ where $\mathbf{x}^j = \{x_1^j, x_2^j, \cdots, x_N^j\}$ is a complete description of state variable j across all N individual components. In statistical mechanics N may be the number of particles in an ideal gas and $\nu = 6$ represents the position and momentum of each particle in three-dimensional space. In an economic setting N may represent the total number of individuals or households in an economy defined by states of income, education, preferences, race, etc. or it could represent the total number of firms characterized by their states of size, age, performance, etc..[5] The complete description of the system would requite we specify all $N\nu$ degrees of freedom. Microfoundations in orthodox economic theory hopes to derive aggregate properties of the system by having a completely specified (possibly dynamic) theory of the evolution of these $\nu(t)$ states across all N agents. When N or ν is large a full description of the system is computationally unobtainable.

The physicist Ludwig Boltzmann realized that despite the impossibility of a full microscopic description of the system, it was possible to derive aggregate macroscopic properties by using a statistical approach that summarized the system with a coarse-grained ensemble. By partitioning the state space of $\mathbf{X}_{\{N,\nu\}}$ one obtains a macroscopic description of the system characterizing the coarse-grained frequency of individuals across discrete states $\mathbf{F}(\nu) = \{f(\mathbf{x}^1) \cdots f(\mathbf{x}^\nu)\}$ where $f(\mathbf{x}^j) = \{n^j(1), n^j(2), ..., n^j(K)\}$, K is the number of coarse-grained bins, and $n^j(k)$ is the number of individuals in bin k across state space j. $\sum_k n_k^j = N$

---

[5]Theil (1967) awkwardly took money as the unit of analysis rather than individuals in his derivation of the Theil index of inequality. As discussed in dos Santos (2020) units of money are indistinguishable and thus Theil's error was analogous to the error Bose (1924) and Einstein (1925) recognized in the use of conventional multinomial statistics to the analysis of ensembles of Bosons.

is the total number of individual components distributed across state j. The precise state j and identity of each individual within each bin is a description of the microstate of the system and the histogram $f(\mathbf{x}^j)$ describes the macrostate which is the distribution of state j over the N individual components. Many different configurations of individual components in state j will lead to the same distribution of individuals over K bins so that any macrostate will correspond to many microstates. The number of ways a particular macrostate can be realized is the multiplicity of the system which is measured by the multinomial coefficient:

$$W = \binom{N}{n_K^j} = \frac{N!}{n_1^j! n_2^j! \cdots n_K^j!}.$$

Using Stirling's approximation for large N, $\log(N!) \approx N \log(N) - N$ for $N >> 1$, the logarithm of the multiplicity is expressed as the entropy of the system:

$$H = \frac{\log(W)}{N} = -\sum_{k=1}^{K} p_k^j \log(p_k^j),$$

where $p_k^j = \frac{n_k^j}{N}$. If all N components had identical state j the macrostate of the system can only be realized in one way since all components share the same bin. In this case the multiplicity is $W = 1$ and the entropy is minimized at $H = 0$. When $n_1^j = n_2^j = ... = n_K^j$ individual components are partitioned equally across all levels of state j and $f(x^j) = \frac{1}{K}$ for all k and entropy is maximized at $H = \log(K)$. As the system's components become more spread out or evenly distributed across the K bins the ways in which a particular histogram (macrostate) can be realized or configured increases dramatically. In this sense entropy is a measure of the dispersion of components and is bounded by the degenerate distribution (minimum entropy) and the uniform distribution (maximum entropy).

In 1948 Claude Shannon developed his theory of communication and axiomatically derived the entropy equation as a measure of the information content of a random variable. The physicist Edwin T. Jaynes (Jaynes, 1967; Jaynes, 1957, 1979) recognized Shannon's theorem as providing a logical basis for understanding the entropy concept in statistical mechanics. In a series of articles over several decades Jaynes argued (Jaynes, 1983; Jaynes, 1979) that Shannon's information theory answered many of the open questions in statistical physics by clarifying the epistemological nature of the probabilistic assertions in statistical mechanics, that is, that probabilities are an expression of what we know about a system, not an ontological feature of a system itself.

Incorporating Shannon's theorem into statistical mechanics led Jaynes to propose the maximum entropy principle of inference. Jayne's MEP is an operational rule for applying Laplace's Principle of Insufficient Reason (or, as Keynes put it, the Principle of Indifference). It states that when making inferences about any system in which our knowledge does not permit deductive reasoning, it is reasonable to expect to observe behavior in line with distributions or macroscopic states achieving maximum entropy across all macroscopic states or distributions compatible with our knowledge. The distribution that maximizes entropy is both the most probable distribution in the sense that it can be realized in the greatest number of ways and is the most unbiased distribution as it only takes into account the information put into the calculation. Jaynes maintained that because Gibbsian equilibrium distributions maximize the multiplicity of a macrostate under certain constraints it can be rationalized in terms of the MEP. From the MEP perspective, the foundations of statistical mechanics can be justified without any appeal to the detailed dynamical features of a system or disputable properties of ergodicity "if we are willing to pay the price. The price is simply that we must loosen the connections between probability and frequency."(Jaynes, 1979, pp.26) Of course, the Jayesian perspective is not without controversy. Much of the disagreement is ultimately tied to the broader disputes over the Bayesian interpretation of probability theory and the questions surrounding transformation groups, though others are concerned with the general limitations of the MEP (Seidenfeld, 1986; Uffink, 1996).

5.2 Applications to Political Economy

A diverse range of recent contributions to political economy have adopted the macroscopic approach to statistical modeling. Drăgulescu and Yakovenko (2000) provided initial theoretical considerations to macroscopic constraints that generate exponential distributions in economic outcomes by arguing for the relevance of certain conservation principles exchange-based economic systems. The authors posit a conservation of money in exchange which generates the exponential distribution as the statistical equilibrium distribution of income. Such conservation principles have become standard in the econophysics literature and have been criticized and discussed at length (for example in dos Santos (2017), Gallegati, Keen, Lux, and Ormerod (2006), Jr. (2016), Rosser Jr. (2008a, 2008c), Scharfernaker and Yang (2020)). While there are many economic situations in which such conservation principles can be rationalized as a relevant macroscopic constraint, for example concerning the allocation of a scarce fixed resource, the argument does not hold much water when applied more complex social processes such as those that concern the distribution of incomes and profit rates.

More recently, Scharfenaker and Semieniuk (2017) attempted to distance explanations

of statistical equilibrium from crude physical analogies and instead situate the statistical equilibrium model in well articulated economic theory through the use of Jaynes' MEP. The authors take up the question of the determinants of the equilibrium profit rate distribution and propose a maximum entropy model that could broadly capture the statistical features of the profit rate distribution with a structural break. Two models are considered, one constraining the mean of the absolute value of profit rates:

$$
\begin{aligned}
&\underset{f(x)\geq 0}{\text{Max}} -\int_{-\infty}^{\infty} f(x)\log f(x)dx \\
&\text{subject to} \int_{-\infty}^{\infty} f(x)\,dx = 1 \\
&\text{and} \int_{-\infty}^{\infty} f(x)\,|x-\mu|\,dx = c
\end{aligned}
\tag{10}
$$

which gives rise to the Laplace distribution and is fit to the pre-neoliberal period and two, doing so piecewise around the central moment:

$$
\begin{aligned}
&\underset{f(x)\geq 0}{\text{Max}} -\int f(x)\log f(x)dx \\
&\text{subject to} \int_{-\infty}^{\infty} f(x)\,dx = 1 \\
&\text{and} \int_{-\infty}^{\mu} f(x)(\mu-x)dx = c_1 \\
&\text{and} \int_{\mu}^{\infty} f(x)(x-\mu)dx = c_2
\end{aligned}
\tag{11}
$$

which generates the Asymmetric Laplace distribution and is fit to the neoliberal period. Like Alfarano et al. (2012) the authors maintain an implicit classical theory of competition and only consider the reduced-form single moment constraint in their model allowing for an additional moment in the later years to account for persistent asymmetries in the profit rate distribution.

The economic interpretation of the constraints is developed in dos Santos (2017), Scharfenaker and dos Santos (2015) and dos Santos and Scharfenaker (2019), which applies the MEP to modeling the statistical equilibrium distribution of Tobin's q. The proposed maximum entropy program considers the two relevant macroscopic constraints as the first moment $\mathbb{E}(x) = c_1 \in \mathbb{R}$, and the absolute first moment $\mathbb{E}(|x|) = c_2 > 0$, which equivalently gives rise to the Asymmetric Laplace distribution as the statistical equilibrium. These constraints are rationalized in the following way. First, the state variable (profit rates or q) is a socially

scaled variable meaning that individual values are normalized by average or social measures of themselves and generate constraints on the mean. For example, the average profit rate is determined by actions of all capitalist firms interacting through competition and is imposed upon all capital as macroscopic constraint. Second, the constraint on the absolute first moment is shown to reflect the obstacles to realizing arbitrage opportunities, which can and do arise in a multitude of ways. The parameters are shown to be directly interpretable as aggregate measures of speculative behavior, the informational content of prices, and of the allocative performance of capital markets.

In a separate line of work, dos Santos and Yang (2020) argue for critical considerations about the nature of interactions of capitalist firms. They argue that capitalist competition likely produces complex long-ranging interactions that rule out ergodicity and the use of simple stochastic differential equations as well as Shannon entropy as the correct entropy measure. As formally detailed by Thurner, Hanel, and Klimek (2018) complex systems characterized by strong interactions are not generally representable as multinomial systems because their phase-space volume is not conserved. The non-conservation of phase-space volume implies that the competitive process effectively rules out some configurations of the system (i.e. micro-states) as it evolves in time. It turns out that for such non-ergodic or self-organizing critical systems the separation axiom of information theory is not valid and as Hanel and Thurner (2013) show the correct measure of entropy takes the generalized non-additive $c - d$ form:

$$S_{c,d} \propto \sum_i^W \Gamma(d + 1, 1 - c \log p_i) \tag{12}$$

where Γ is the incomplete gamma function and c and d are the so-called "scaling exponents." The Standard Boltzmann-Gibbs-Shannon entropy is recovered for $(c, d) = (1, 1)$ and the entropy related to stretched exponential systems arise under $(c, d) = (1, d)$.

dos Santos and Yang (2020) follow the same classical interpretation of profit rate equalization whereby the persistent movement of capital from "low- to high-profitability undertakings push all rates of return toward an emergent, general rate of return r_g that acts as the benchmark measure of the opportunity cost of capital against which all rates of return are measured." They interpret this "social scaling" of the profit rate as generating a first moment constraint along of the lines of dos Santos (2017). While both dos Santos and Yang (2020) and dos Santos and Scharfenaker (2019) share the same formal constraint, they differ in the entropy functional used in the MEP resulting in different statistical equilibrium distributions. Maximizing the generalized $1 - d$ entropy subject to a piecewise first-moment constraint on

profit-rate differentials $x = r - r_g$ results in the four-parameter double stretched-exponential distribution:

$$
f(x) \propto
\begin{cases}
e^{(1-(\lambda_l x + \mu_l))^{1/d_l}} & \text{if } x < 0 \\
e^{(1-(\lambda_r x + \mu_r))^{1/d_r}} & \text{if } x > 0
\end{cases}
\tag{13}
$$

The authors fit this model to firm-level profit rates gathered from the Amadeus dataset which contains over 20 million individual European enterprises across fifteen countries. The model is reported as capturing 99.1% of the informational content of the data according to Soofi's (Soofi & Retzer, 2002) informational indistinguishability index indicating very good measures of fit.

The generalized entropy approach raises several immediate questions. First, while generalized entropies are extensive mappings for systems where strong interactions ensure Shannon entropy is non-extensive the information ensuring the non-extensivity of a system may not be easily testable or testable at all. What exactly is meant by "long-ranging" and "strong" interactions in the context of specific economic phenomena is not clear. Second, the authors assume the $1 - d$ entropy functional based on observed profit rate data which appears to have heavy tails and argue that these empirical regularities justify the use of generalized non-extensive entropies. While there is certainly a compelling argument to be made for the "observational approach" to economic theory that begins with the identification of statistical regularities in the data, it does not follow that that these regularities uncover the relevant institutional structures that define specific moment constraints or even the appropriate entropy functional.

Upon closer inspection it appears that many of the statistical equilibrium models discussed presuppose the existence of a known closed-form distribution with specific mathematically necessary constraints that can be rationalized in terms of economic theory. Economic theory appears to follow as an afterthought to the identification of the moment constraints necessary to derive a presumed well-fitting closed-form model. But, from the MEP there is no *a priori* reason to consider the subset of known closed-from distributions as candidate statistical equilibrium models. This does not suggest that no useful economic interpretations or even theory development around presumed constraints is possible. But it does imply that such a strategy is necessary limited to a possibly arbitrary subset of models.

Scharfenaker and Foley (2017) attempt to overcome this limitation and propose a model in the Jaynesian tradition that attempts to reason from explicit economic theory to the statistical equilibrium distribution without presupposing the distributional outcome. The next section reviews this model in detail due its distinctiveness and growing use in statistical

equilibrium applications.

5.3 Quantal Response Statistical Equilibrium

The quantal response statistical equilibrium (QRSE) model introduced by Scharfenaker and Foley ([2017](#)) is a maximum entropy model that explains the statistical distribution of social outcomes by considering the fundamental sources of variation that arise in competitive decentralized economic interactions. This model was first used to explain the equalization of profit rates (Scharfenaker & Foley, [2017](#)) from a classical Smithian perspective, and has been used to model induced technical change (Yang, [2018a](#)), fluctuations in housing markets (Ömer, [2018](#), [2020](#)), asset price fluctuations (Blackwell, [2018](#); Scharfernaker, [2019](#)), and international competition in labor markets (Wiener, [2020](#)). The main idea behind the QRSE model is to consider a system in which an outcome, $x \in \mathcal{X}$, is brought into statistical equilibrium by the purposive actions, $a \in \mathcal{A}$, of participants in an institutional structure the generates negative stabilizing feedbacks. An equilibrium of the system is represented by a joint probability distribution, $f(a, x)$. This equilibrium joint distribution determines the marginal distributions $f(a) = \int f(a, x) \, dx$, $f(x) = \sum_{\mathcal{A}} f(a, x)$, and conditional distributions, $f(a|x) = \frac{f(a,x)}{f(x)}$, $f(x|a) = \frac{f(a,x)}{f(a)}$, which are interpretable as the causal forces constituting the statistical equilibrium.

The first component of the model represents the behavior of the typical agent in terms of the probability that she will choose a particular action a conditional on the aggregate variable x and is expressed in the conditional distribution $f(a|x)$. The second component of the model reflects the impact of the action on the aggregate variable which is expressed in the conditional distribution $f(x|a)$.

Letting the action set be the binary case $\mathcal{A} = \{\bar{a}, \underline{a}\}$, where $\bar{a}$ represents the positive action (e.g. a firm entering a market) and $\underline{a}$ the negative action (e.g. a firm exiting a market). The joint entropy can be written in terms of the marginal outcome conditional action frequencies:

$$
\begin{aligned}
\mathcal{H}(f(a, x)) &= -\left(\int f(\bar{a}, x) \log(f(\bar{a}, x)) \, dx + \int f(\underline{a}, x) \log(f(\underline{a}, x)) dx \right) \\
&= -\int f(x) \log(f(x)) dx - \int f(x) \left(f(\bar{a}, x) \log(f(\bar{a}, x)) + f(\underline{a}, x) \log(f(\underline{a}, x)) \right) dx \\
&= \mathcal{H}(f(x)) + \int f(x) \mathcal{H}(f(a|x)) dx
\end{aligned}
$$

$$(14)$$

If there is a non-zero impact of the action on the social outcome then $f(x|a) \neq f(x)$. In many economic interactions the impact of actions on social outcomes limits the difference in the conditional expected outcome. This feedback is a central theme in classical political economy and represents the unintended consequences of individual actions on social outcomes. The negative feedback of actions on the social outcome constrains the difference between the expected outcomes conditional on the action to be $\delta > 0$. In general this difference arises from where the actions impact the outcome and can be expressed in the following inequality

$$\mathbb{E}\left(\underline{a}|x\right) \leq \mathbb{E}(x) \leq \mathbb{E}(\bar{a}|x)$$

$$\int f\left(\underline{a}, x\right) x \, dx \leq \xi \leq \int f(\bar{a}, x) x \, dx \tag{15}$$

$$\int f\left(\bar{a}, x\right)(x - \xi) \, dx - \int f(\underline{a}, x)(x - \xi) \, dx \geq 0$$

Maximizing entropy will make this difference infinite unless we constrain it to be some positive value δ. Letting $\Delta f(a|x) = f(\bar{a}|x) - f\left(\underline{a}|x\right)$ we can write inequality 15 as:

$$\int \Delta f(a|x)(x - \xi) dx \leq \delta \tag{16}$$

Maximizing the joint distribution subject to this negative feedback constraint and normalization we get:

$$\underset{f(x) \geq 0}{\text{Max}} - \int f(x) \log(f(x)) dx + \int f(x) \mathcal{H}(f(a|x)) dx$$

$$\text{subject to} \int f(x) \, dx = 1 \tag{17}$$

$$\text{and} \int \Delta f(a|x) f(x)(x - \xi) dx \leq \delta$$

The Lagrangian associated with this programming problem is:

$$\mathcal{L}(f(x); \lambda, \gamma) = -\int f(x) \log(f(x)) dx + \int f(x) \mathcal{H}(f(a|x)) dx - \lambda \left(\int f(x) \, dx - 1\right)$$

$$- \gamma \left(\int \Delta f(a|x) f(x)(x - \xi) dx - \delta\right) \tag{18}$$

22

The first-order conditions are sufficient to characterize the solution for $\gamma > 0$:

$$f(x) = \frac{1}{\mathcal{Z}(x, \gamma)} e^{\mathcal{H}(f(a|x))} e^{-\gamma(\Delta f(a|x))(x-\xi)} \tag{19}$$

Where $\mathcal{Z}(x, \gamma) = \int e^{\mathcal{H}(f(a|x))} e^{-\gamma(\Delta f(a|x))(x-\xi)} \, dx$ is the normalizing constant, or partition function, which has no known closed-form solution.[6]

Scharfenaker and Foley (2017) define this function as the Quantal Response Statistical Equilibrium (QRSE) distribution. The final model closure is specifying the conditional distribution $f(a|x)$, which describes the behavioral response of an agent to the outcome x. From a CPE perspective the decentralized and uncoordinated nature of capitalist markets ensures that while individual actions may be optimizing, such as capital seeking the highest rate of profit, there is no guarantee that optimizing actions lead to optimizing outcomes. Neither capitalists nor workers are assured the maximum wage rate or profit rate just because they believed they entered the most remunerative line of production.

5.3.1 Entropy constrained behavior

Shannon's theorem is a useful and intuitive point of departure for modeling the uncertainty in a decision environment. It quantifies the amount of choice variability for a given decision environment and is an increasing function of the number of choices available. Because Shannon entropy is a measure of uncertainty it can provide a meaningful probabilistic description of a decision environment by being incorporated into the quantal behavior of agents as a behavioral constraint. Introducing a minimum uncertainty (entropy) constraint to any decision problem has a simple intuitive interpretation as representing the perceptual and control limitations that affect the agent's ability to determine environmental signals and its own behavioral determinations in a given decision environment. Assuming the typical agent conditions their quantal choices $a \in \mathcal{A}$ on perceived signals from their social environment $x \in \mathcal{X}$ e.g. profit rate and wage rate differentials, we can summarize individual behavior as responding to a payoff function $v(a, x) : \mathcal{A} \times \mathcal{X} \to \mathbb{R}$. The typical agent's mixed strategy in this scenario is defined by the probability distribution $f(a|x) : \mathcal{A} \to \mathbb{R}_{\geq 0}$ over the actions that maximizes the expected payoff

$$\mathbb{E}(v(a, x)) = \sum_{\mathcal{A}} f(a|x)v(a, x) \tag{20}$$

---

[6]This partition function under certain transformations reduces to the so called "sophomore's dream" (Borwein, Bailey, & Girgensohn, 2004) which has no closed form solution.

Maximizing expected payoff without any further constraints results in the typical von Neumann-Morgenstern result where the individual always put all weight on the payoff-maximizing action and zero weight on any other action. As a probability measure on $\mathbb{R}$, this is characterized by the Heaviside unit step function:

$$\theta(a|x - \hat{x}) = \begin{cases} 1 & \text{if } x \geq \hat{x} \\ 0 & \text{if } x < \hat{x} \end{cases} \tag{21}$$

Differentiating results in the degenerate delta distribution that puts all weight on the action with the maximum expected payoff $\hat{x}$:

$$f(a|x) = \delta(a|x - \hat{x}) \tag{22}$$

Imposing a constraint on the minimum entropy of the mixed strategy distribution introduces variation in choice akin to Sims' "rational inattention" approach (Sims, 2003).

$$
\begin{aligned}
\underset{\{f(a|x) \geq 0\}}{\text{Max}} \; & \mathbb{E}(v(a, x)) \\
\text{subject to } & \sum_{\mathcal{A}} f(a|x) = 1 \\
& -\sum_{\mathcal{A}} f(a|x) \log(f(a|x)) \geq H_{min}
\end{aligned}
\tag{23}
$$

The Lagrangian associated with this programming problem is:

$$
\begin{aligned}
\mathcal{L}(f; \lambda, T) = & -\sum_{\mathcal{A}} f(a|x) v(a, x) - \lambda \left(\sum_{\mathcal{A}} f(a|x) - 1 \right) \\
& + T \left(\sum_{\mathcal{A}} f(a|x) \log(f(a|x)) - H_{min} \right)
\end{aligned}
\tag{24}
$$

The first-order conditions imply that agent's behavior conditional on the outcome is described by the logit quantal response (LQR) function:

$$f(a|x) = \frac{e^{\frac{v(a,x)}{T}}}{\sum_{\mathcal{A}} e^{\frac{v(a,x)}{T}}} \tag{25}$$

The dual representation of this problem maximizes the entropy of the agent's mixed strategy subject to the payoff of choosing an action being at some minimum.

$$
\begin{aligned}
&\underset{\{f(a|x) \geq 0\}}{\text{Max}} - \sum_{\mathcal{A}} f(a|x) \log(f(a|x)) \\
&\text{subject to } \sum_{\mathcal{A}} f(a|x) = 1 \\
&\text{and } \mathbb{E}(v(a,x)) \geq V_{min}
\end{aligned}
\tag{26}
$$

We can interpret this macroscopic dual problem as maximizing our uncertainty about the typical agent's mixed strategy $f(a|x)$ subject to the constraint that agents require a minimum payoff to act, which might be interpreted as a type of participation constraint. The dual solution is

$$
f(a|x) = \frac{e^{\beta v(a,x)}}{\sum_{\mathcal{A}} e^{\beta v(a,x)}}
\tag{27}
$$

In equation (25) the "behavioral temperature" T is the shadow price of obtaining information and represents the unit informational entropy cost of increasing expected payoff. In general, as T declines, we come closer and closer to the zero-entropy solution as the probability of any non-maximizing action goes to zero because for $f(a|x) = 1$ the entropy $H = -\sum_{\mathcal{A}} f(a|x) \log(f(a|x)) = 1 \cdot log(1) = 0$. This implies that as long as $T > 0$ there is a positive probability for each action an agent will make. As $T \to 0$, the distribution reduces to singleton maximal strategy. In general T might be functionally dependent on the payoff x, though that case is not explored here.

The simplest example of a choice setting is one in which only two alternative actions are available as identified above. In this case, the conditional action frequencies can be expressed as logistic equations. Letting $\Delta v(a,x) = v(\bar{a}, x) - v(\underline{a}, x)$,

$$
f(\bar{a}|x) = \frac{e^{\frac{v(\bar{a},x)}{T}}}{e^{\frac{v(\bar{a},x)}{T}} + e^{\frac{v(\underline{a},x)}{T}}} = \frac{1}{1 + e^{\frac{-\Delta v(a,x)}{T}}}
\tag{28}
$$

$$
f(\underline{a}|x) = 1 - f(\bar{a}|x) = \frac{1}{1 + e^{\frac{\Delta v(a,x)}{T}}}
\tag{29}
$$

As Foley (2020a) points out, the entropy constrained approach to individual choice theory does violate the assumptions of consistency and completeness of preferences. But, rather than

try to rationalize the quantal response distribution in order to preserve these assumptions, the information theoretic approach highlights the uncertainty of choice and has important implications for economic theory that are not emphasized in conventional choice theory. Information theory gives meaning and interpretation to the residual indeterminacy of behavior associated with the LQR distribution. Approaching the problem of individual choice from an information theoretic perspective leads to the very important conclusion that what is usually interpreted as random noise is in fact relevant information about the choice setting.

The introduction of the informational constraint also recovers the logit-quantal response distribution found in the random utility literature without imposing any ad hoc distributional constraints on the errors that affect decision making, for example that they are Extreme Value Type 1 distributed. Rather than treating the distribution over the individual's mixed strategy as arising from random error of noisy players or oscillating utility, the informational entropy constrained model gives meaning to the observed dispersion of behavior as the relative payoff of different actions.

5.4 The QRSE Distribution with LQR Closure

The Quantal Response Statistical Equilibrium Distribution with the Logit Quantal Response closure allows the expression of equation 19 in terms of a payoff function $v(a, x)$ and behavioral parameter T.

$$f(x) \propto e^{\mathcal{H}_T(\Delta v(a,x))} e^{-\gamma \tanh\left(\frac{\Delta v(a,x)}{2T}\right)(x-\xi)} \tag{30}$$

Where

$$\mathcal{H}_T(\Delta v(a, x)) = -\left(\frac{1}{1+e^{-\frac{\Delta v(a,x)}{T}}} \log\left(\frac{1}{1+e^{-\frac{\Delta v(a,x)}{T}}} \right) + \frac{1}{1+e^{\frac{\Delta v(a,x)}{T}}} \log\left(\frac{1}{1+e^{\frac{\Delta v(a,x)}{T}}} \right) \right) \tag{31}$$

is the entropy of the conditional action distribution. In the simplest scenario the payoff is symmetric for each action taken such that $v(\bar{a}, x) = v(\underline{a}, x)$ and equal to the outcome x expressed as a differential from the agent's expected average payoff, or fundamental valuation of x, $\mathbb{E}(x) = \mu$. For example, the payoff x may represent the real wage earned by a worker for entering a particular line of employment or the profit rate realized by a capitalist for competing in a particular market and μ is equal to the expected rate of return for entering the market. If agent's expectations are correct then $\mu = \xi$. In this case Eq. 32 reduces to

the symmetric QRSE distribution

$$f(x) \propto e^{\mathcal{H}_{T,\xi}(x)} e^{-\gamma \tanh\left(\frac{x-\xi}{T}\right)(x-\xi)} \tag{32}$$

The kernel of the QRSE distribution with symmetric linear payoffs appears in Figure 3.
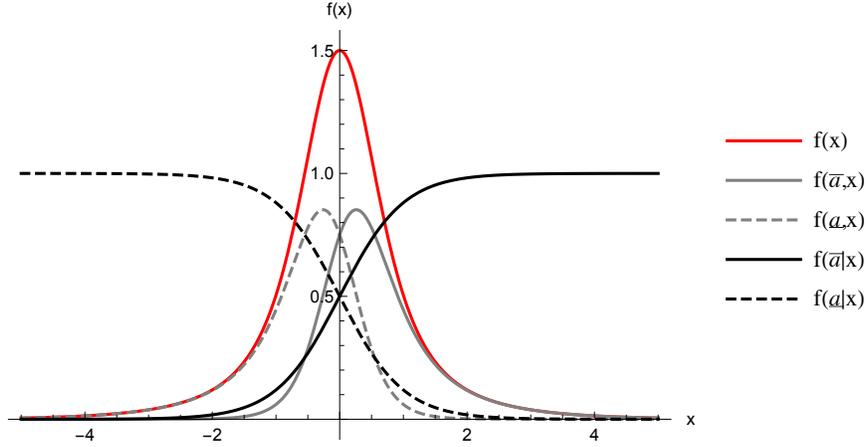

Figure 2: QRSE marginal, joint, and conditional frequencies for symmetric linear payoffs with $T = 1$, $\gamma = 1$ and $\xi = 0$.

In the limiting case where $T \to 0$ agents will always chose the action $\bar{a}$ whenever $x > \xi$ and $\underline{a}$ whenever $x < \xi$ and Eq. 25 reduces to the step function 21 simplifying Eq. 32 to:

$$\hat{f}(x) \propto e^{-\gamma|x-\xi|} \tag{33}$$

which is the kernel of the Laplace distribution. The QRSE framework thus explains Laplace distributed fluctuations of the outcome variable as arising not from agents' informational constraints, but from the institutional mechanisms that lead to a feedback of their actions on the outcome. Because the conditional response probabilities reduce to a degenerate distribution while the marginal frequencies of the outcome stabilize to a Laplace distribution it is clear to see that the fluctuations of x are determined jointly by the institutional feedback constraint and the conditional response probabilities. In contrast to orthodox theory, which tends to neglect the institutional constraints that engender social interactions while also assuming zero-entropy behavior, the QRSE model incorporates both institutional constraints that operate at the macroscopic level and bounded rationality constraints on the individual that operate at the microscopic level.

While the Laplace distribution implies a zero-entropy constraint consistent with orthodox approaches, it may be difficult in many situations to distinguish between $T = 0$ and $T \approx 0$.

When T is very low $f(x)$ may appear to be Laplace distributed, but the same conclusions cannot be drawn from the knife-edge case when $T = 0$ (Blackwell, 2018; Foley, 2020a; Scharfernaker, 2019). Thus, the widely used Laplace distribution (Bottazzi & Secchi, 2006; Kotz, Kozubowski, & Podgórski, 2001; Scharfenaker & Semieniuk, 2017; Stanley et al., 1996; Williams, Baek, Li, Park, & Zhao, 2017), for statistical equilibrium modeling may be better understood for those cases with behaviorally relevant theoretical foundations as a quantal response statistical equilibrium with very low informational constraints.

The other limiting case is when $T \to \infty$ and actions are conditionally independent of the outcome. In this case response probabilities become uniform. As $\gamma \to \infty$ as well, indicating $\delta \to 0$, the difference between the expected outcomes conditional on the action goes to zero and Eq. 32 reduces to:

$$f(x) \propto e^{-\frac{\gamma(x-\xi)^2}{T}} \tag{34}$$

which is the kernel of the Normal distribution. Thus the Laplace and Normal distributions are the behaviorally limiting cases of the QRSE distribution (see Blackwell (2018) for further statistical details of the QRSE distribution). The QRSE distribution with symmetric linear payoffs shares many of the same characteristics of the Subbotin distribution discussed above, but has the benefit of model interpretability.

5.5 Sources of Asymmetries in Quantal Response Statistical Equilibrium

There are several ways that the QRSE distribution generalizes to produce asymmetric frequencies in the distribution of economic outcome $f(x)$. In many cases the unconditional expectation of the outcome $\mathbb{E}(x)$ is determined at the system level. For example, the average rate of profit in the classical theory of competition is not determined by any individual capitalist, but rather emerges as systemic constraint that all capitalists face. Under these circumstances the outcome x can be understood as a "socially scaled" variable (dos Santos, 2017), which is expressed as a constraint on the average outcome.

Maximizing the joint distribution subject to the negative feedback constraint and normalization we get:

$$\underset{f(x)\geq 0}{\text{Max}} - \int f(x)\log(f(x))dx + \int f(x)\mathcal{H}(f(a|x))dx$$
$$\text{subject to} \int f(x)\,dx = 1$$
$$\text{and} \int f(x)x\,dx = c \tag{35}$$
$$\text{and} \int \Delta f(a|x)f(x)(x-\xi)dx \leq \delta$$

The solution is

$$f(x) \propto e^{\mathcal{H}_T(\Delta v(a,x))}e^{-\gamma\tanh\left(\frac{\Delta v(a,x)}{2T}\right)(x-\xi)}e^{-\phi x} \tag{36}$$

Where $\gamma > 0$ is the Lagrange multiplier associated with the feedback constraint and ϕ is the Lagrange multiplier associated with the mean constraint.

As with the Asymmetric Laplace used in Scharfenaker and Semieniuk (2017) and dos Santos and Scharfenaker (2019) an equivalent specification to 35 is to constrain the feedback of the action on the outcome in piecewise fashion, which would indicate the possibility of asymmetric impacts of the action on the outcome. For example, the effect of entering a market on the profit rate may end up reducing the profit rate more than the increase in the profit rate in the market exited. The asymmetric impact constraint is

$$\underset{f(x)\geq 0}{\text{Max}} - \int f(x)\log(f(x))dx + \int f(x)\mathcal{H}(f(a|x))dx$$
$$\text{subject to} \int f(x)\,dx = 1$$
$$\text{and} \int f(x)f(\bar{a}|x)(x-\xi)dx = \bar{\delta} \tag{37}$$
$$\text{and} \int f(x)f(\underline{a}|x)(\xi-x)dx = \underline{\delta}$$

The solution, with Lagrange multipliers $\bar{\gamma}$ and $\underline{\gamma}$, is

$$f(x) \propto e^{\mathcal{H}_T(\Delta v(a,x))}e^{-\bar{\gamma}\frac{v(\bar{a}|x)}{2T}(x-\xi)}e^{-\underline{\gamma}\frac{v(\underline{a}|x)}{2T}(x-\xi)} \tag{38}$$
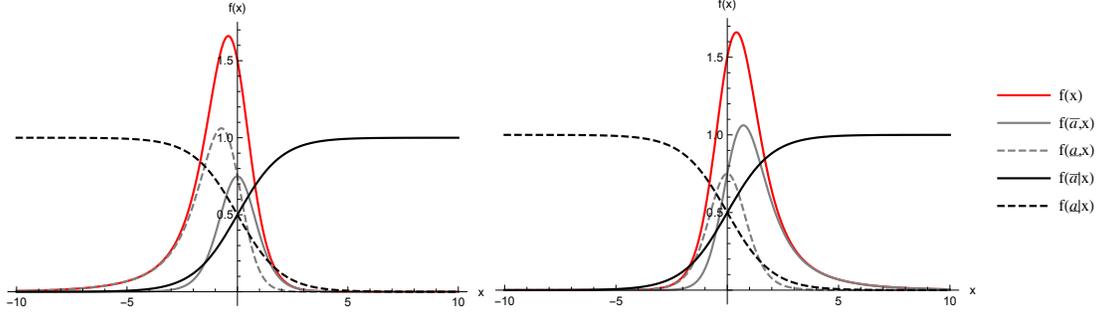
which appears in Figure 3.

Figure 3: Asymmetric QRSE marginal, joint, and conditional frequencies for symmetric linear payoffs with $T = 1$, $\gamma = 1$, $\xi = 0$, $\phi = 0.5$ (left) and $\phi = -0.5$ (right) .

As is common with asymmetric unimodal distributions that can be expressed piecewise, some algebraic manipulation shows the equivalence between the two approaches. In this case the piecewise constraint can be expressed as:

$$f(x) \propto e^{\mathcal{H}_T(\Delta v(a,x))} e^{-\left(\frac{\bar{\gamma}+\underline{\gamma}}{2}\right) \tanh\left(\frac{\Delta v(a,x)}{2T}\right)(x-\xi)} e^{-\left(\frac{\underline{\gamma}-\bar{\gamma}}{2}\right)x} \tag{39}$$

which shows the equivalence with Eq. 36 with $\gamma = \frac{\bar{\gamma}+\underline{\gamma}}{2}$ and $\phi = \frac{\underline{\gamma}-\bar{\gamma}}{2}$.

As detailed in Foley (2020b) and Scharfernaker (2019) the asymmetric effects may also emerge from differences in agents' expectations μ and the actual average value of the market rate ξ. Setting $\phi = 0$ asymmetries emerge only through the interaction of μ and ξ.

$$f(x) \propto e^{\mathcal{H}_T(x)} e^{-\gamma \tanh\left(\frac{x-\mu}{T}\right)(x-\xi)} \tag{40}$$

When $\mu = \xi$ agent's expectations are correct or self fulfilling. Agents' actions are determined by the perception of the social outcome through their estimated "fundamental" value μ. However, there is no reason to assume that agents' expectations of the fundamental μ will be correct. When $\mu - \xi \neq 0$ expectations are unfulfilled incentivizing agents to revalue their estimate of the fundamental through some type of market-based punishment for acting incongruous with the market.

This formulation is equivalent to specifying asymmetric costs for taking an action (Foley & Blackwell, 2019) which implies $v(\bar{a}, x) + v(\underline{a}, x) = \mathcal{C}(a|x)$. For example, there may be higher costs associated with entering a market than with exiting a market. With a simple linear additive cost $\mathcal{C}(a|x) = c$ the Asymmetric QRSE can be expressed as:

$$f(x) \propto e^{\mathcal{H}_{T,\mathcal{C}(a|x)}(x)} e^{-\gamma \tanh\left(\frac{(x-\xi)+\frac{c}{2}}{T}\right)(x-\xi)} \tag{41}$$

which is equivalent to Eq. 40 with $\mu = \xi + \frac{c}{2}$. Thus an additive cost offers an alternative interpretation of unfulfilled expectations, or what Sargent (2017) identifies as the equivalence between "mistaken beliefs" and the price of risk. Combining these sources of asymmetry the Generalized QRSE distribution is

$$f(x) \propto e^{\mathcal{H}_{T,\mathcal{C}(a|x)}(x)} e^{-\gamma \tanh\left(\frac{\Delta v(a,x)+\mathcal{C}(a|x)}{2T}\right)(x-\xi)-\phi x} \tag{42}$$

The Generalized QRSE distribution is a four-parameter distribution that shares many of the the relevant qualitative features with the Asymmetric Subbotin Distribution Bottazzi and Secchi (2011), Mundt and Oh (2019), but has the benefit of being behaviorally founded and directly interpretable without recourse to stochastic differential equations.

5.6 Applications to the Profit Rate and Income Distribution

Scharfenaker and Foley (2017) derive the QRSE distribution in the context of profit rate equalization and show that the model is able to account for the fat tails and asymmetries observed in the profit rate distribution and permits inference about the most probable behavior of firms conditional on profit rate differentials. The difference between the QRSE approach and other statistical equilibrium models of profit-rate equalization is that it derives the constraints from classical political economic theory directly rather than from mathematical necessity of a predetermined closed-form distribution. This approach thus permits direct economic interpretation of the model parameters.

The QRSE model is also able to account for the statistical equilibrium income distribution. In considering the relevance of the QRSE model to wage rate equalization it is important to consider how the two variables differ. Smith argued the wage rate was bounded from below by the level of subsistence. The center of gravity for the average wage become a highly debated topic throughout the 19th century ranging from the "iron law of wages" that continuously pushed labor back to physiological subsistence to Marx's theory of the value of labor power as the costs of reproducing the commodity labor power under average conditions. Capital accumulation and economic growth tends to pull the wage above this level, but there is some conceivable minimum beyond which workers' standard of living would deteriorate appreciably. The mobility of labor and equalization of wage rates in Smith's thought experiment and the equalization of rates of surplus value in Marx, tends to be more difficult to conceptualize than the mobility of capital primarily due to the life-cycle effects of individual workers. While a capitalist might conceivably encounter short-run obstacles to investing in any line of production, such as construction of buildings, start-up costs, mo-

bilization of labor, etc. the mobility of capital, particularly financial capital, makes profit rate equalization appear feasible at the individual and social level of production. Individual workers, however, have a more difficult time moving between occupations due to a multitude of barriers such as training costs, geography, access to education, language, and issues of race and gender. Smith's notion of the tendential equalization of wages, however, explicitly abstracts away from the individual life cycle and focuses on the ways social labor can adapt to needs of social reproduction. The key concept for Smith and Marx, is that labor effort is comparable in terms of its effects in production across workers and fungible between different tasks. From this perspective, the observed heterogeneity of career, education, and training paths of individuals implies equality of expected outcomes across individuals.

A classical statistical equilibrium theory of the distribution of wages specifically requires we abstract away from the individual and instead consider the mass of social labor as homogeneous and fungible. From this perspective, the distribution of wages is an *ex post* outcome of social labor in the abstract adapting to the needs of social reproduction through persistent balancing of the advantages and disadvantages of different employments. The main point in a statistical equilibrium model of wage rate equalization is that the abstract labor of individual workers can be aggregated in social production despite the subjective quality of labour effort, and the variety of specific tasks, skills, training, and so forth that constitute concrete labour.

A statistical model of wage rate equalization then considers the decisions of a typical or "average" worker to enter a particular line of employment and the aggregate effect of these decisions on the distribution of wages including the formation of a central tendency and the persistent fluctuations around this tendency due to the decentralized nature of the division of labor in a capitalist system. In the view of Smith and Marx this average worker is an abstraction that is determined at any point in time by the general circumstances of social reproduction. The other important difference between the profit rate distribution and the income distribution is that the profit rate is, in the language of statistical mechanics, an *intensive* variable that does not scale with the size of the system. Thus the relevant dynamics are those of a negative feedback. The wage, on the other hand, is an *extensive* variable that does scale with the size of the system. The growth of wages, however, is an intensive variable that stabilizes through negative feedbacks. To put the QRSE model on the right footing then, consider a variable $y = e^x$ for which one can make the transformation of variables to derive the Log-Quantal Response Statistical Equilibrium (Log-QRSE) distribution:

$$f(x) \propto \frac{1}{y} e^{\mathcal{H}_T(\log y)} e^{-\gamma \tanh\left(\frac{\log y - \xi}{T}\right)(\log y - \xi)} e^{-\phi \log y} \tag{43}$$

with

$$\mathcal{H}_T(\log y) = -\left(\frac{1}{1+e^{-\frac{\log y - \xi}{T}}} \log\left(\frac{1}{1+e^{-\frac{\log y - \xi}{T}}} \right) + \frac{1}{1+e^{\frac{\log y - \xi}{T}}} \log\left(\frac{1}{1+e^{\frac{\log y - \xi}{T}}} \right) \right) \quad (44)$$

This four-parameter distribution is plotted in Figure 4.
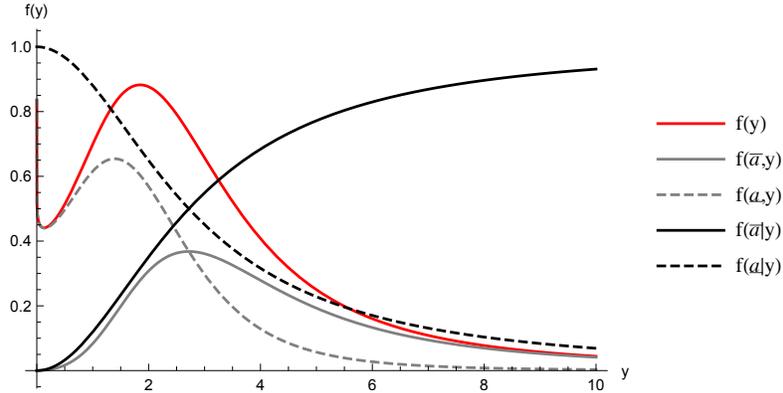

Figure 4: Log-QRSE marginal, joint, and conditional frequencies for symmetric linear payoffs with $T = 0.5$, $\gamma = 0.95$, $\mu = 0$, $\phi = 0$.

It can easily be shown that the Log-QRSE contains the Log-Normal distribution as a limiting case, which has a long tradition of use in the study of income distribution (Champernowne & Cowell, 1998; Gibrat, 1931; Kalecki, 1945; Kanbur & Venkatasubramanian, 2020; Venkatasubramanian, 2017). It is also capable of capturing bi-modalities, which have been shown to be characteristic of the the income distribution (Scharfernaker & Schneider, 2019; Schneider, 2013). Like the Generalized QRSE model, the Log-QRSE also exhibits characteristics of widely used, but poorly theorized statistical equilibrium models of income distribution. Most notably this includes the Generalized Beta of the Second Kind (GB2), the go-to parametric distribution for describing the distribution of income based entirely on statistical fit criteria (Burkhauser, Feng, Jenkins, & Larrymore, 2008; Feng, Burkhauser, & Butler, 2006; Jenkins, 2009). The Log-QRSE the offers a compelling alternative to these purely fit-based models that is theoretically grounded in the classical political economic theory of wage rate equalization.

5.7 Considerations on QRSE

There are at least two broad, but related criticisms that have been raised about the QRSE framework. The first is that the model attributes all distributional outcomes "as reflecting the limited capacity for information processing of the typical or average individual."(dos Santos & Yang, 2020) The second, concerns the limits of the model as a representative agent

model. As detailed above, it should be clear that there is an important interplay between the various constraints that define the QRSE model. In the limiting zero-entropy scenario when agents' behavior is deterministically described by the step function the statistical equilibrium distribution is either the symmetric (or asymmetric) Laplace distribution and not the degenerate Delta distribution. The reason is that the institutional feedback constraint also contributes to explaining distributional regularities. Only when the institutional feedback constraint is "turned-off" can all variation be attributed to the probabilistic assumption about individual choice. The model importantly predicts that without the feedback generated by the presumed institutions the outcome is Gaussian, which is exactly what would be predicted by the central limit theorem assuming that the decisions of agents are independent and additive. From the QRSE perspective, agents interact through institutions which generate non-Gaussian outcomes.

To the second point, it seems reasonable to think that observed distributional outcomes arise for a multitude of reasons including unpredictable shocks, bounded rationality, institutional structures, and the heterogeneity of individuals. Different models emphasize different sources of variation. For example, the concept of a homogeneous labor market naturally leads to explanations of the variation in income as arising from individual characteristics and not from institutional structures of capitalism. This point was elegantly demonstrated in Foley (1994) who derived the statistical equilibrium distribution of market exchange that arises with transactions at non-Walrasian prices. The implications are far reaching, but not all that surprising. As is articulated in the fundamental welfare theorems, the Walrasian market is incapable of generating inequality in outcomes that is not already present in endowments, technologies, and preferences. Under the Walrasian approach, households of the same type (same preferences and endowment) will always receive the same bundle of consumption goods in equilibrium leading to the "equal treatment" property of such systems. When exchange at non-equilibrium prices become possible identical agents end up with different final consumption bundles in the statistical equilibrium. The market thus introduces horizontal inequality among traders of the same type and the welfare theorems no longer hold. By emphasizing the institutional constraints that are indifferent to heterogeneous individual characteristics the model reveals the inequities structured in market mediated exchange. Only by assuming heterogeneity of agents (in terms of endowments, technologies, and preferences) does general equilibrium theory generate Pareto efficient outcomes.

Observed variation in economic outcomes is undeniably the product of heterogeneous individual characteristics as well as institutional structures. The QRSE incorporates this information by summarizing it in behavioral averages. The average can be determined in a large (possibly infinite) number of ways, but the maximum entropy QRSE distribution ensures

34

that the predictive joint statistical equilibrium distribution is overwhelmingly probable. The QRSE model does not predict any deterministic paths traversed by a representative agent. Instead, it provides a complete dynamic description of all individual behavior through the ensemble averages. In other words, the QRSE details a *probabilistic* description of individual dynamics based on behavioral averages. Because QRSE summarizes all behavioral information in terms of averages, the behavioral parameters μ and T only reflect heterogeneity of agents indirectly. However, there may many circumstances in which the behavioral averages differ in important ways across different types of agents. For example, a statistical equilibrium distribution of Tobin's q could be considered as determined by the investment decisions of corporate managers and outside investors each with investment decisions conditioned by conceptually distinct institutional and social environments. What would make this model interesting and distinct from the homogeneous QRSE model is there are two heterogeneous groups interacting: speculators pricing the stock taking the firm's investment/financial behavior as given, and firm management deciding on investment/financial strategy taking the stock market as given. These two participants are likely to operate in qualitatively different decision environments distinguished, for example, by their time horizon. Other applications could include the study of labor market outcomes between different groups as in Wiener (2020).

One advantage to the QRSE framework that makes it distinct among the various approaches discussed in this survey is that it offers a "meso-level" account of the determinations of statistical equilibrium. The determination of the distribution of outcomes in the QRSE model is neither reducible to simple strategic interactions among individual participants in a game as with conventional quantal response equilibrium or to the reduced-form macroscopic framework that opaquely subsumes interactions and feedbacks into macro-level constraints. Instead, it uniquely integrates the two through the meso-level interaction of the average agent's behavior with the macroscopic constraints that act upon all agents.

6   Open Questions

There are many unresolved issues surrounding the use of statistical equilibrium methods in political economy. Given the open questions that remain in the mature field of statistical mechanics from which they come, there may be little hope these issues will be resolved in economics. However, our understanding of how these methods operate in social sciences is evolving and given the history of cross-fertilization of social and physical disciplines, it does not seem unreasonable to expect completely original developments in economics. At this point at least three important issues stand out as deserving critical attention.

The first concerns the relevance of ergodic approaches to economic systems. While the issues surrounding this question are essentially the same that afflict statistical mechanics, they emerge in a distinct, possibly more compelling way when considered in the scope of social sciences. Ultimately, these questions concern when it is legitimate to make the assumptions necessary to derive the ergodic distribution. What considerations are given to substantiating the Markov property? Can one be assured that the stochastic process is in fact a homogeneous Markov process, or, equivalently, obeys a so-called master equation? In the ergodic approach the Wiener process is the prime driver of the stochasticity of the partial differential equation that describes the motion of a representative particle subject to a large number of small molecular shocks. The statistical effects of the equilibrium distribution thus broadly follow the Gaussian model of random error which originated as a way of explaining the variation in observed planetary motions as errors arising from telescopic and microscopic observations. If the disturbances producing the statistical effects are independent and have finite variance one can appeal to the central limit theorem to make distributional conclusions. The question then is how compelling is the central limit theorem for explaining the distributional regularities that emerge in complex social systems? Is it possible to distinguish between a movement due to Brownian motion and movement due to other effects, such as increasing returns, purposive behavior, or institutional constraints that are independent of individual interactions and intentions? As dos Santos argues, this approach provides "poor conceptual foundation for grappling with the economic or social significance of statistical regularities" because there is "no scope for explicit consideration of interactions between individuals that may define irreducibly systemic or macro-level regularities involving all moments of observed distributions."(dos Santos, 2020, pp.8-9)

Superimposing probabilistic principles on micro-dynamical laws permits the desired derivation of the posited kinetic determinations, but does this approach offer a statistical explanation of observed macroscopic regularities or merely determine a set of conditions that guarantee the presumed deterministic laws that govern the system? This problem emerges in thermodynamics when the hypothesis of molecular chaos (that the collision of two particles is uncorrelated with their individual velocities and independent of their positions) is introduced in order to guarantee Boltzmann's equation applies to the kinetic changes of non-equilibrium thermodynamic systems. But it remains an open question as to whether the assumption of a continuous rerandomization of the dynamic evolution of an ensemble where the Hamiltonian dynamics are repeatedly "turning on" and "turning off", which ensures a Markovian representation of the system, is even consistent with the micro-dynamical laws of the system (Callender, 1999). As Roman Frigg argues, "The main problem with this approach is that its probabilistic laws are put in 'by hand' and are not derived from

the underlying dynamics of the system; that is, it is usually not possible to derive the probabilistic laws from the underlying deterministic evolution and hence the probabilistic laws are introduced as independent postulates. However, unless one can show how the transition probabilities postulated in this approach can be derived from the Hamiltonian equations of motion governing the system, this approach does not shed light on how thermodynamical behaviour emerges from the fundamental laws governing a system's constituents."(Frigg, 2008)

In an economic context, the Markov assumption generally seems doubtful. Capitalism is about accumulation and the ways in which accumulation manifests tend to be highly path dependent (Arthur, 1994; Bowles, 2006). Imposing the same dynamic law on every economic entity but allowing for idiosyncratic variation does not seem a convincing postulate when considering important path dependencies. Appeals to the central limit theorem may be appropriate for certain phenomena, but given the fact that many economic outcomes are the result of social beings interacting in complex non-additive ways within institutional structures, central limit theorems seems less convincing as an explanation.

Lastly, it should be worth considering Jaynes' point that "even if one had succeeded in proving [the] ergodic theorems rigorously and universally, the result would have been established only for time averages *over infinite times*; whereas the experiments which verify Gibbs' rules measure time averages only over finite times. Thus, a further mathematical demonstration would in any event be necessary, to show that these finite time averages have sufficiently approximated their limits for infinite times... such an additional demonstration *cannot* be given"(Jaynes (1967) pp.94). For the same reason that the frequency interpretation cannot establish a meaningful interpretation of probabilities from finite samples, the ergodic approach is forced to substitute the mathematical concept of the limit for the non-existent infinite sample of observations. Thus, while the ergodic or stochastic representation of statistical equilibria is a parsimonious and mathematically convenient representation of economic processes, it seems ultimately to introduce even more unnecessary complications in applications to social systems than exist in statistical mechanics. The fact that we do not and can not even possess strong enough behavioral postulates to establish an analog to the equations of motion or principles of conservation for an economic system has already ruled out statistical equilibrium thinking in many economists' minds (Davidson, 1996; Mirowski, 1991). It may ultimately be worth seriously evaluating Jaynes' goading claim that "It appears to be a quite general principle that, whenever there is a randomized way of doing something, then there is a nonrandomized way that delivers better performance but requires more thought."(Jaynes, 2003, p. 497)

The second issue concerns the complex nature of economic systems and how it changes

our understanding about the validity of different approaches. For example, what entropy functional is the correct one to use? Are generalized entropies necessary or do they introduce information unwarranted and thus violate Jaynes' MEP? As Golan (2018) argues, "Much work has been done to investigate the large-sample properties of the generalized entropy function... But the issue we most often have to deal with is limited and possibly noisy information, which did not come from an experiment and may be ill-behaved." Thus, when it comes to including implicit information, like that which would warrant generalized entropy functional, "it would be better to introduce this information via priors... that way we know exactly what information we introduced and how these priors were constructed."(Golan, 2018, pp.391)

A broader related question concerns the meaning of interactions in social and economic settings and how to represent these in a statistical equilibrium model. The stochastic approach attributes interaction to the kinetic theory of gasses in which molecules interact only through collision or possibly further ranging potential interactions. The macroscopic approach abstracts away from detailed dynamics and subsumes interactions into reduced-form models capturing the most probable configuration of the system given macroscopic constraints. Is it important to make interactions explicit, for example by distinguishing between a movement due to Brownian motion and movement due to other effects in a stochastic model, or by introducing explicit interaction constraints in a macroscopic model? Making such distinctions, however, may be difficult because equilibrium systems are not easily decomposed into causal links as everything at some level is both a cause and effect of everything else. Nevertheless it may be important to distinguish between simple additive memoryless interactions and long-ranging path dependent interactions and the statistical patters each are capable of generating.

Finally, it seems some consideration should be given to the level of generality in different approaches. For example, the combinatorial justification of entropy requires that N must be large enough to justify Stirling's approximation while Shannon's axiomatic entropy requires no such requirement and can be applied to any problem of inference where information is incomplete. There is debate surrounding the generalities of different conceptual foundations of entropy (Caticha, 2012; Golan, 2018; Gorban, Gorban, & Judge, 2010; Niven, 2007), and it may be worth considering these issues in terms of the types of questions economists hope to answer. Ultimately it may be that the decision of what method to use should just depend on the information we have about the problem that faces us.

The third issue concerns the meaning of statistical explanation in economics. Sklar (1993) offers a thoughtful discussion on this issue in statistical mechanics and many of the points are valid in a social setting as well because both areas share similar features and analytical tools.

However, the logic of explanation of statistical equilibrium in social systems is bound to be different from the logic governing physical systems due to nature of intentions and social interactions. Thus, we might ask whether in explaining the statistical equilibrium features of profit rate equalization it is acceptable to reduce the process to a stochastic differential equation where all statistical effects are driven by Brownian motion. Similarly, is it sufficient to reduce the causal determining process of entry and exit to a single moment constraint such as the dispersion of profit rates or mean profit rate in a maximum entropy problem? A recent paper by Reddy (2020) argues "that methods from statistical physics can help in providing explanations of economic outcomes only if they include adequate attention to processes and moreover characterize these in an appropriate way, recognizing their specifically social character." And "This requires attention to "non-mechanical" processes of interaction, inflected by power, culture, institutions etc." It thus seems necessary to keep both the individual and the interaction end of social phenomena firmly in view when constructing theories and models rather than to subsume them under opaque mechanical processes. The QRSE model seems to address this issue because it articulates the process of profit rate equalization explicitly in the conditional distributions. By arguing that a "process must be specified in order to provide the causal insight that is needed for there to be an explanation (for instance, through an appeal to wage and profit dynamics giving rise to a Fokker-Planck equation)" Reddy (2020) seems to miss the point that processes can be specified statistically as well as deterministically.

Ultimately, many of these questions are related to the deeper questions about the meaning of statistical and probabilistic assertions in general and thus the meaning of probabilities. While both ergodic and macroscopic approaches are capable of producing models that can explain the distributional regularities of social outcomes, it is worth considering the scope of phenomena each are capable of handling. The ergodic approach shares the same qualitative and quantitative features with the frequency definition of probability, while the macroscopic approach tends to favor the Laplacian (or Bayesian) interpretation. The frequency approach is widely seen as a special case of the more general Laplacian approach due to the necessary existence of prior probabilities.[7] Thus, while both approaches may be capable of explaining some phenomena, the ergodic approach may reach its limits faster as the problems it is applied to become more sophisticated. Both approaches may be valid and useful in different contexts and one is not necessarily "wrong" to use stochastic approaches and central limit theorems if the problem at hand permits such methods. But, in dealing with complex social phenomena one always faces underdetermined ill-posed problems with incomplete informa-

---

[7]Though the introduction of such statistical techniques as "regularization" have aided in the reach of classical statistics.

tion that does not easily permit frequency interpretations of probabilities or the existence of infinite time averages. There seem to be many good reasons to favor more general approaches that avoid unnecessary technical and interpretive difficulties.

7   Conclusion

Statistical methods began in the human and social sciences and are a natural basis for situating social and economic theory on firm mathematical and quantitative foundations. The nature of economic data requires a coherent statistical methodology if there are hopes of satisfying the minimum criteria of what counts as an explanation of real economic phenomena. A theory should, at the least, be able to reproduce the important features of observable economic data. Once those features have been identified they define the complementary set of indeterminate residual features. Different theories amount to different choices about what salient reproducible features of the data are important and consequently what remains as the residual indeterminacy. Theory that concerns only the central moments of the statistical distributions of economic outcomes is necessarily limited to explaining only a small subset of features of far more complex social phenomena. The statistical equilibrium approach recognizes the statistical nature of social and economic outcomes and attempts to explain a broad set of statistical features as arising from specific economic processes. Some theories emphasize such processes deterministically and superimpose variation as arising from ontological randomness. Other theories emphasize such processes statistically deriving the higher moments of the model by identifying the appropriate set of systemic constraints. While there are still many unresolved questions about how statistical mechanics can meet specific predictive and explanatory aims of economic theory, it offers a general methodology for conceptualizing complex systems that can be configured in an astronomical number of ways and makes considerable progress on our ability to model and understand the distributional outcomes of social and economic processes.

The move from conventional equilibrium thinking in economics to statistical equilibrium thinking in economics brings with it a richer set of concepts and possibilities for prediction and interpretation of social phenomena. Equilibrium prices and incomes traditionally understood as abstract centers of gravity are modified to statistical regularities with clearly articulated distributional consequences. The statistical equilibrium approach provides a probabilistic rather than deterministic account of social phenomena and thus differs in its explanatory aims and fundamental postulates. Because the explanatory goals of the statistical approach are fundamentally different than traditional equilibrium methods in economics, the principle analytical resource of probability theory must be evaluated in the context of

these goals. Conventional econometric accounts of statistical variation in social phenomena are an inadequate framework for conceptualizing statistical equilibrium. The ergodic and macroscopic approaches both offer some degree of explanation, but differ in important methodological and interpretive ways.

References

Alfarano, S., Milaković, M., Irle, A., & Kauschke, J. (2012). A statistical equilibrium model of competitive firms. *Journal of Economic Dynamics and Control*, *36*(1), 136–149.

Arthur, W. B. (1994). *Increasing returns and path dependence in the economy*. University of Michigan.

Blackwell, K. (2018). *Entropy constrained behavior in financial markets: A quantal response statistical equilibrium approach to financial modeling* (Doctoral dissertation, The New School for Social Research).

Boltzmann, L. (1871). Über das wärmegleichgewicht zwischen mehratomigen gasmolekülen. *Wiener Berichte*, *63*, 397–418.

Borwein, J. M., Bailey, D. H., & Girgensohn, R. (2004). *Experimentation in mathematics: Computational paths to discovery*. A K Peters/CRC Press.

Bose, S. N. (1924). Planck's law and the hypothesis of light quanta. *Zeitchrift fur Phyik*, *26*, 178–181.

Bottazzi, G., & Secchi, A. (2006). Explaining the distribution of firm growth rates. *The RAND Journal of Economics*, *37*(2), 235–256.

Bottazzi, G., Li, L., & Secchi, A. (2019). Aggregate fluctuations and the distribution of firm growth rates. *Industrial and Corporate Change*, *28*(3), 635–656.

Bottazzi, G., & Secchi, A. (2011). A new class of asymmetric exponential power densities with applications to economics and finance. *Industrial and Corporate Change*, *20*(4), 991–1030.

Bowles, S. (2006). *Microeconomics: Behavior, institutions and evolution*. Princeton University Press.

Burkhauser, R. V., Feng, S., Jenkins, S., & Larrymore, J. (2008). Estimating trends in us income inequality using the current population survey: The importance of controlling for censoring. *Journal of Economic Inequality*, *9*(3), 393–415.

Callender, C. (1999). Reducing thermodynamics to statistical mechanics: The case of entropy. *The Journal of Philosophy*, *96*(7).

Caticha, A. (2007). *Information and entropy*. arXiv:0710.1068.

Caticha, A. (2012). *Entropic inference and the foundations of physics* (E. Monograph commissioned by the 11th Brazilian Meeting on Bayesian Statistics, Ed.). Sao Paulo: University of São Paulo Press.

Champernowne, D. G., & Cowell, F. A. (1998). *Economic inequality and income distribution.* Cambridge: Cambridge University Press.

Champernowne, D. (1953). A model of income distribution. *Economic Journal, 63*(250), 318–351.

Cottrell, A. F., Cockshott, P., Michaelson, G. J., Wright, I. P., & Yakovenko, V. M. (2009). *Classical econophysics.* Routledge.

Davidson, P. (1996). Reality and economic theory. *Journal of Post Keynesian Economics, 18*(4), 479–508.

dos Santos, P. L. (2017). The principle of social scaling. *Complexity, 2017*(Article ID 8358909).

dos Santos, P. L. (2020). Statistical equilibria in economic systems: Socio-combinatorial or individualist-reductionist characterizations? *European Physical Journal Special Topics.*

dos Santos, P. L., & Scharfenaker, E. (2019). Competition, self-organization, and social scaling—accounting for the observed distributions of tobin's q. *Industrial and Corporate Change, 28*(6), 1587–1610.

dos Santos, P. L., & Wiener, N. (2019). Indices of informational association and analysis of complex socio-economic systems. *Entropy, 21(4)*(367).

dos Santos, P. L., & Yang, J. (2020). Arbitrage, information, and the competitive organization of distributions of profitability. *Advances in Complex Systems, Forthcoming.*

Dosi, G. (2007). Statistical regularities in the evolution of industries. a guide through some evidence and challenges for the theory. In F. Malerba & S. Brusoni (Eds.), (pp. 153–186). Cambridge University Press.

Drăgulescu, A., & Yakovenko, V. M. (2000). Exponential and power-law probability distributions of wealth and income in the united kingdom and the united states. *Physica A, 299*, 213–221.

Einstein, A. (1925). Quantum theory of the monoatomic ideal gas. *Sitzungsberichte der Preussischen Akademie der Wissenchaften.*

Farjoun, F., & Machover, M. (1983). *Laws of chaos: A probabilistic approach to political economy.* Verso.

Feng, S., Burkhauser, R. V., & Butler, J. (2006). Levels and long-term trneds in earnings inequality: Overcoming current population survey censoring problems using the gb2 distribution. *Journal of Business and Economic Statistics, 24*(1), 57–62.

Foley, D. K. (1994). A statistical equilibrium theory of markets. *Journal of Economic Theory, 62*(2), 321–345.

Foley, D. K. (2003). *Unholy trinity: Labor, capital, and land in the new economy.* New York, NY: Routledge.

Foley, D. K. (2020a). Information theory and behavior. *European Physical Journal Special Topics.*

Foley, D. K. (2020b). Unfulfilled expectations: One economist's history. In A. Arnon & W. Young (Eds.), *Expectations: Theory and applications from historical perspectives.* Springer International.

Foley, D. K., & Blackwell, K. (2019). *A general quantal response statistical equilibrium model.* Unpublished Notes.

Föllmer, H. (1974). Random economies with many interacting agents. *Journal of Mathematical Economics 1*, 51–62.

Frigg, R. (2008). A field guide to recent work on the foundations of statistical mechanics. In *The ashgate companion to contemporary philosophy of physics.* Routledge.

Gallegati, M., Keen, S., Lux, T., & Ormerod, P. (2006). Worrying trends in econophysics. *Physica A*, *370*, 1–6.

Garibaldi, U., & Scalas, E. (2010). *Finitary probabilistic methods in econophysics.* Cambridge University Press.

Gibbs, J. W. (1902). *Elementary principles in statistical mechanics.* New York: C. Scribner.

Gibrat, R. (1931). *Les inégalités économiques.* Paris: Librairie du Rucueil Sirey.

Golan, A. (2018). *Foundations of info-metrics: Modeling, inference and imperfect information.* New York, NY: Oxford University Press.

Gorban, A. N., Gorban, P. A., & Judge, G. (2010). Entropy: The markov ordering approach. *Entropy*, *12*, 1145–1193.

Hanel, R., & Thurner, S. (2013). Generalized (c,d)-entropy and aging random walks. *Entropy*, *15*, 5324–5337.

Hobson, A. (1971). *Concepts in statistical mechanics.* New York, NY: Gordon and Breach.

Ijiri, Y., & Simon, H. (1977). *Skew distributions and the sizes of business firms.* North-Holland: Amsterdam.

Jaynes, E. T. (2003). *Probability theory: The logic of science.* Cambridge University Press.

Jaynes, E. T. (1967). Foundations of probability theory and statistical mechanics. In M. Bunge (Ed.), *Delaware seminar in the foundations of physics.* Springer-Verlag.

Jaynes, E. T. (1983). *Papers on probability, statistics, and statistical physics* (R. Rosenkrantz, Ed.). Reidel.

Jaynes, E. T. (1957). Information theory and statistical mechanics. *The Physical Review*, *106*(4), 620–630.

Jaynes, E. T. (1979). Where do we stand on maximum entropy? In R. D. Levine & M. Tribus (Eds.), *The maximum entropy formalism*. MIT Press.

Jenkins, S. P. (2009). Distributionally-sensitive inequality indices and the gb2 income distribution. *Review of Income and Wealth*, *55*(2), 392–298.

Jr., J. B. R. (2016). Entropy and econophysics. *European Physical Journal Special Topics*, *225*, 3091–3104.

Kalecki, M. (1945). On the gibrat distribution. *Econometrica*, *13*(2), 161–170.

Kanbur, R., & Venkatasubramanian, V. (2020). Occupational arbitrage equilibrium as an entropy maximizing solution. *European Physical Journal Special Topics*.

Kapur, J. N. (1989). *Maximum entropy models in science and engineering*. New York, NY: John Wiley & Sons.

Kapur, J. N., & Kesavan, H. K. (1992). *Entropy optimization principles with applications*. Academic Press Inc.

Kotz, S., Kozubowski, T., & Podgórski, K. (2001). *The laplace distribution and generalizations: A revisit with new applications to communications, economics, engineering, and finance*. Birkhauser Verlag GmbH.

Kurz, H. D., & Salvadori, N. (1995). *Theory of production: A long-period analysis*. Cambridge University Press.

Langston, R. (1984). A new approach to the relation between prices and values. In *Ricardo, marx, sraffa: The langston memorial*. London: Verso.

Lux, T. (2016). Applications of statistical physics methods in economics: Current state and perspectives. *The European Physical Journal Special Topics*, *225*, 3255–3259.

Mantegna, R. N., & Stanley, H. E. (1999). *An introduction to econophysics: Correlations and complexity in finance*. Cambridge, UK: Cambridge University Press.

Maxwell, J. C. (1860). Illustrations of the dynamical theory of gases. *Philosophical Magazine*, *4*(19).

McCauley, J. L. (2009). *Dynamics of markets: The new financial economics*. Cambridge University Press.

Mirowski, P. (1991). *More heat than light: Economics as social physics, physics as nature's economics*. Cambridge University Press.

Mundt, P., Alfarano, S., & Milaković, M. (2016). Gibrat's law redux: Think profitability insterad of growth. *Industrial and Corporate Change*, *25*(4), 549–571.

Mundt, P., Alfarano, S., & Milaković, M. (2020). Exploiting ergodicity in forecasts of corporate profitability. *Journal of Economic Dynamics and Control*, *111*.

Mundt, P., & Oh, I. (2019). Asymmetric competition, rish, and return distribution. *Economics Letters*, *179*, 29–23.

Niven, R. K. (2007). Origins of the combinatorial basis of entropy. In *Aip conference proceedings* (Vol. 954, *133*).

Ömer, Ö. (2018). Dynamics of the us housing market: A quantal response statistical equilibrium approach housing market: A quantal response statistical equilibrium approach. *Entropy*, *20*(11).

Ömer, Ö. (2020). Maximum entropy approach to market fluctuations as a promising alternative. *European Physical Journal Special Topics*.

Pareto, V. (1987a). Aggiunta allo studio della curva delle entrate. *Giornale degli Economisti*, *2*(14), 15–26.

Pareto, V. (1987b). *Cours d'économie politique* (Rouge, Ed.). Lausanne.

Reddy, S. (2020). Statistical mechanics approaches in economics: What is an explanation? *The European Physical Journal Special Topics*.

Redhead, M. (1995). *From physics to metaphysics*. Cambridge University Press.

Rosser Jr., J. B. (2008a). Debating the role of econophysics. *Nonlinear Dynamics, Psychology, and Life Sciences*, *12*(3), 311–323.

Rosser Jr., J. B. (2008b). Econophysics. In *The new palgrave dictionary of economics* (2nd ed., pp. 1625–1643). Palgrave MacMillan.

Rosser Jr., J. B. (2008c). Econophysics and economic complexity. *Advances in Complex Systems*, *11*(5), 745–760.

Sargent, T. J. (2017). Risk aversion or mistaken beliefs? In *David k. backus memorial lecture*. NYU Stern.

Scharfenaker, E., & Foley, D. (2017). Quantal response statistical equilibrium in economic interactions: Theory and estimation. *Entropy*, *19*(9), 444.

Scharfenaker, E., & Semieniuk, G. (2017). A statistical equilibrium approach to the distribution of profit rates. *Metroeconomica*, *68*(3), 465–499.

Scharfernaker, E. (2019). *Implications of quantal response statistical equilibrium* (Working Paper No. 2019-07). Univeristy of Utah.

Scharfernaker, E., & dos Santos, P. L. (2015). The distribution and regulation of tobin's q. *Economic Letters*, *137*, 191–194.

Scharfernaker, E., & Schneider, M. P. A. (2019). *Labor market segmentation and the distribution of income: New evidence from internal census bureau data* (Working Paper No. 2019-08). Univeristy of Utah.

Scharfernaker, E., & Yang, J. (2020). Maximum entropy economics. *The European Physical Journal Special Topics*.

Schneider, M. P. A. (2013). Evidence for multiple labor market segments: An entropic analysis of us earned income, 1996-2007. *Journal of Income Distribution*, *22*(2), 60–98.

Seidenfeld, T. (1986). Entropy and uncertainty. *Philosophy of Science, 53*, 467–491.

Shaikh, A. (2016). *Capitalism: Competition, con ict, crises*. Cambridge, UK: Cambridge University Press.

Shaikh, A. (2020). The econ in econophysics. *European Physical Journal Special Topics*.

Shaikh, A., & Jacobo, E. (2019). *Economic arbitrage and the econophysics of income inequality* (tech. rep. No. 1902). New School Economic Papers.

Shaikh, A., Papanikolaou, N., & Wiener, N. (2014). Race, gender and the econophysics of income distribution in the usa. *Physica A, 415*, 54–60.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal, 27*, 379–423, 623–656.

Simon, H. A. (1955). On a class of skew distributiuon functions. *Biometrika, 42*(3-4), 425–440.

Simon, H. A., & Bonini, C. P. (1958). The size distribution of business firms. *American Economic Review, 48*(4), 607–617.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics, 50*(3), 665–690.

Sklar, L. (1993). *Physics and chance: Pilosophical issues in the foundationations of statistical mechanics*. Cambridge University Press.

Soofi, E. S., & Retzer, J. J. (2002). Information indicies: Unification and applications. *Journal of Econometrics, 107*, 17–40.

Stanley, M., Amaral, L., Buldyrev, S., Havlin, S., Leschhorn, H., Maass, P., ... Stanley, H. (1996). Scaling behavior in the growth of companies. *Nature, 379*(6568), 804–806.

Steindl, J. (1965). *Random processes and the growth of firms*. London: Griffin.

Subbotin, M. (1923). On the law of frequency of error. *Mathematicheskii Sbornik, 31*, 269–301.

Theil, H. (1967). *Economics and information theory*. North-Holland Publishing Company.

Thurner, S., Hanel, R., & Klimek, P. (2018). *Introduction to the theory of complex systems*. Oxford University Press.

Uffink, J. (1996). The constraint rule of the maximum entropy principle. *Studies in History and Philosophy of Modern Physics, 27*, 47–79.

Uffink, J. (2007). Compendium of the foundations of classical statistical physics. In J. Butterfield & J. Earman (Eds.), *Philosophy of physics*. North-Holland Publishing Company.

Venkatasubramanian, V. (2017). *How much inequality is fair? mathematical principles of a moral, optimal, and stable capitalist society*. Columbia Univeristy.

Wiener, N. M. (2020). Labor market segmentation and immigrant competition: A quantal response statistical equilibrium analysis. *Entropy, 22*(7), 742.

Williams, M. A., Baek, G., Li, Y., Park, L. Y., & Zhao, W. (2017). Global evidence on the distribution of gdp growth rates. *Physica A*, *468*, 750–758.

Yakovenko, V. M., & Silva, A. C. (2005). Two-class structure of income distribution in the usa: Exponential buld and power-law tail. In A. Chatterjee, S. Yarlangadda, & B. K. Chakrabarti (Eds.), *Econophysics of income and wealth distributions* (pp. 15–23). Milan: Springer.

Yakovenko, V. M. (2007). Econophysics, statistical mechanics approach to. In R. A. Meyers (Ed.), *Encyclopedia of complexity and system science*. Springer.

Yang, J. (2018a). A quantal response statistical equilibrium model of induced technical change in an interactive factor market: Firm-level evidence in the eu economies. *Entropy*, *20*(3).

Yang, J. (2018b). Information theoretic approaches in economics. *Journal of Economic Surveys*, *32*(3).