Integration of the Walrasian Paradigm into the Statistical Paradigm

Alexis Akira Toda*

This Version: September 24, 2009

Abstract

Walrasian and statistical equilibrium have distinct methodologies and the relation between the two equilibrium concepts has been unknown. We reformulate the concept of statistical equilibrium and prove its existence under much less restrictive assumptions than in the previous literature. As a corollary, we show that Walrasian equilibrium can be interpreted as a special case of statistical equilibrium, so the latter is "general general equilibrium." Considering the pervasive use of Walrasian equilibria in economic analysis, statistical equilibrium theory can explain even richer economic phenomena, in particular horizontal inequality such as the income distribution.

1 Introduction

For centuries in human history, considerable inequality in income and wealth distributions has been observed. Figure 1 shows the histogram of the 2008 US family income together with the estimated gamma density, which apparently fits quite well. Where do such regularities come from? Pareto (1896) first quantitatively analyzed income distributions and discovered his celebrated Pareto distribution. Champernowne (1953) proposed a theoretical model

*Department of Economics, Yale University. alexisakira.toda@yale.edu

deriving the Pareto distribution, but Salem and Mount (1974) empirically showed that the gamma distribution fits better. Despite the relevance of the mechanism underlying horizontal inequality, however, economists have not yet come up with a general, universally accepted theory that explains it. We hope to provide one in this paper.

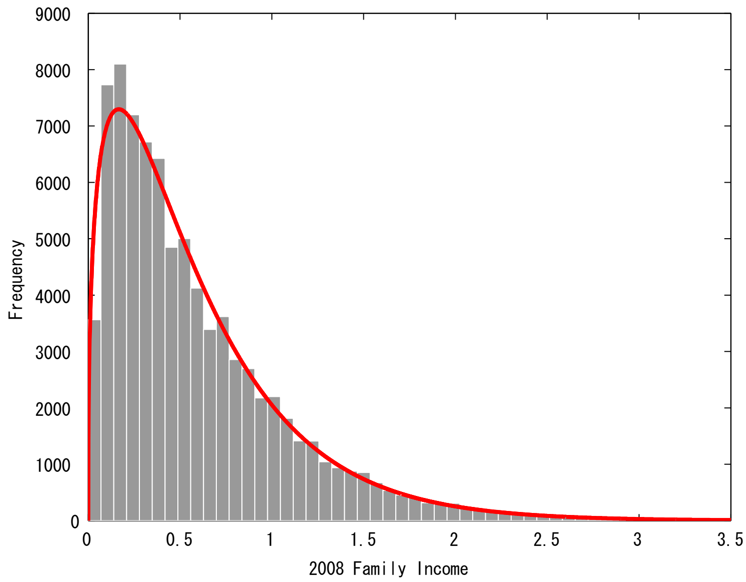

Figure 1: Histogram and estimated gamma density for 2008 income, horizontal axis in $100,000.

Statistical equilibrium theory, first developed by Foley (1994) and then expanded by Toda (2009), is such a candidate. The term "statistical equilibrium" has sporadically appeared in the economics literature in different contexts (Simon, 1959; Green and Majumdar, 1975; Grossman and Stiglitz, 1980; Krebs, 1997; Silver et al., 2002). Although the definition of statistical equilibrium varies across these authors, a common aspect is that the equilibrium concept is a distribution, not a point. The purpose of statistical equilibrium theory in the sense of Foley (1994); Toda (2009) is to explain the distribution of economic variables—such as consumption, income, or labor supply—in a large economy in a general equilibrium context. The trading process is a "black box" (Foley, 2003) except for that agents have private perceptions about their trade opportunities, which are called *offer sets.* If the system is large and random, the distributions of outcomes typically become the entropy-maximizing distributions, as in the case of the energy distribution of gas molecules in a given volume.

Thus statistical equilibrium theory has a different set of axioms from Walrasian equilibrium theory. The axioms of the latter are, simply put, as follows:

1. agent optimization,

2. market clearing,

3. rational expectations (informational consistency),

4. absence of market power.

In statistical equilibrium theory, we replace Axiom 1 by

1'. entropy maximization,

and we may drop Axiom 4 in some cases.

Foley (1994) considered entropy maximization, market clearing, and the absence of market power to be the axioms of statistical equilibrium theory. Since, in general, agents' perception of their trade opportunities depends on the information they have, Toda (2009) introduced the concept of endogenous offer sets, i.e., offer sets are not fixed but change as agents gather information. Thus, Toda (2009) modified the equilibrium concept such that agents make rational expectations, namely the information they initially have and the information they collect from the market coincide in equilibrium: informational consistency is part of the new definition.

One major issue within statistical equilibrium theory has been the relationship between Walrasian and statistical equilibria. Foley (2003) made a heuristic argument that Walrasian equilibria might be viewed as an asymptotic approximation to statistical equilibria. The present paper is the first contribution in the literature that rigorously resolves this issue. We reformulate the concept of statistical equilibrium and prove its existence in order to clarify the connection between Walrasian and statistical equilibria. It turns out that Walrasian equilibria are special cases of statistical equilibria. Since Walrasian equilibrium theory is completely contained in statistical equilibrium theory (statistical equilibrium is "general general equilibrium"), the latter can explain any phenomena that the former can, but not vice versa. For instance, Walrasian equilibrium theory fails to explain horizontal inequality; statistical equilibrium theory can.

The structure of the remaining of the paper is as follows. Section 2 defines the basic concepts in statistical equilibrium theory. These concepts originate from Foley (1994); Toda (2009) but we slightly modify them in order to derive a general theory in later sections.

Section 3 is the core of this paper and contains our major contributions. We prove the existence of statistical equilibria in Theorems 3.5 and 3.7 under much less restrictive assumptions than in Toda (2009). We immediately obtain Corollary 3.9, which states that Walrasian equilibria can be interpreted as special cases of statistical equilibria. Thus, we provide a rigorous mathematical foundation to Foley's observation that Walrasian equilibria might be viewed as an asymptotic approximation to statistical equilibria. In fact, the Walrasian paradigm is integrated into (but not *onto*) the statistical paradigm.

Section 4 discusses some applications and Section 5 concludes. In Appendix A, we briefly review the basic concepts in information theory. Although the material is standard, we felt the urge to include it since most economists are not familiar with information theory. Appendix B deals with mathematical details.

2 Definitions

In this section, we define the basic concepts in statistical equilibrium theory. The definition of a *statistical economy* is formally the same as in Toda (2009) but we give a new interpretation. We slightly modify the equilibrium concept proposed in Toda (2009) to derive a general theory in later sections.

2.1 Economy

Given agents, Foley (1994); Toda (2009) start by defining *offer sets* of agents, which are interpreted as the sets consisting of all feasible and acceptable transactions. In Foley (1994), offer sets are finite sets and he puts a weight on each point; Toda (2009) considers more general sets that can be turned into regular Borel measure spaces,[1] referred to as *offer spaces*.[2] This direction of definition has a shortcoming, for we do not know how to choose the measures from preferences.

To overcome this ambiguity, we reverse the direction: we start from a regular Borel measure (offer space) for each agent type and define its support[3] to be the offer set of that type. Formally, we define as follows.

Definition 2.1 (statistical economy)**.** The object

$$\mathcal{E} = \left\{ \mathcal{I}, \{w_i\}_{i \in \mathcal{I}}, \{\mu_{i,p}\}_{i \in \mathcal{I}, p \in \Delta^{C-1}} \right\}$$

is called a *statistical economy* if

- $\mathcal{I} = \{1, 2, \ldots, I\}$ is the set of agent types,
- w_i is the weight (ratio) of type i agents, so $w_i > 0$ and $\sum_{i=1}^{I} w_i = 1$,
- $\Delta^{C-1} = \left\{ p \in \mathbb{R}_+^C : \|p\|_1 = 1 \right\}$[4] is the set of the scarcity parameter p,[5] where C denotes the number of commodities,
- Given $p \in \Delta^{C-1}$, $\mu_{i,p}$ is a regular Borel measure of type i agents over their trade opportunities $x \in \mathbb{R}^C$. *Offer set* $X_{i,p}$ is the support of $\mu_{i,p}$,

[1] See Folland (1999) for the concepts in measure theory that appear in this paper.

[2] Note that the formulation of Foley (1994) is a special case of Toda (2009) since weights on a finite set can be regarded as a counting measure, which is regular Borel.

[3] The support of a regular Borel measure μ on a second-countable topological space X is defined as follows. Let $\mathcal{U}$ be a countable base of the topology (for the case of $X = \mathbb{R}^C$, it suffices to take the family of all open balls with rational radii and centers with rational coordinates) and $S = X \setminus \bigcup_{U \in \mathcal{U} : \mu(U) = 0} U$. Since $\mathcal{U}$ is the countable base of the topology, it follows that S is closed, $\mu(S) = \mu(X)$, and $\mu(U \cap S) > 0$ whenever U is open and $U \cap S \neq \emptyset$. S is called the *support* of μ and is denoted by $S = \operatorname{supp} \mu$.

[4] Hereafter, $\|\cdot\|_p$ will refer to the L^p norm.

[5] This set is often referred to as the "price simplex" in Walrasian equilibrium theory, but p may or may not be the price vector. For details, see Toda (2009).

and the measure space $(X_{i,p}, \mu_{i,p})$ is called the *offer space*.

Notice the flexibility of our definition: since we have put no structure whatsoever on $\mu_{i,p}$, Definition 2.1 can be used to describe many different situations. For instance, the economy might be a pure exchange economy or a production economy; agents might be consumers, producers, arbitrageurs, or the government, etc; agents might be price takers or might have market power; the market can be well-functioning or can have a lot of friction; agents might or might not be rational, and so on.

We interpret that the measures $\{\mu_{i,p}\}$ are the complete description of agents' prior information over the transactions, for instance the degree of perceived uncertainty or liquidity, technological feasibility, preference, likelihood of accepting a transaction that has been offered, and so on. Following Toda (2009), we do not answer to the question of how offer spaces are determined: anything goes as far as they are regular Borel measures. Simply put, the primitives of a statistical economy are agents and their beliefs over their trade opportunities.

The natural interpretation of the offer set $X_{i,p}$ is the set of net transactions that agents expect to be engaged in with positive probability. At this point we deviate from the interpretation of the previous literature, Foley (1994); Toda (2009), where offer sets were viewed as the sets of net transactions that agents are willing to execute. Dropping acceptability from the definition is not as unnatural as it seems at the first glance. For instance, quite often investors incur losses in the stock market (the trades were unacceptable *a posteriori*). However, investors were *a priori* well aware of the possibility of losing: they simply invested in the stock market in the hope of making profits.

Although in application $\mu_{i,p}$ are typically absolutely continuous with respect to the Lebesgue measure, we may sometimes wish to treat the counting measure or the product measure of such measures. Definition 2.2 fixes the terminology for such measures.

Definition 2.2 (atomic/non-atomic parts)**.** Let

$$\mathcal{E} = \left\{ \mathcal{I}, \{w_i\}_{i\in\mathcal{I}}, \{\mu_{i,p}\}_{i\in\mathcal{I}, p\in\Delta^{C-1}} \right\}$$

be a statistical economy with offer sets $X_{i,p} = \operatorname{supp}\mu_{i,p}$. If $\mu_{i,p}$ has the form $\mu_{i,p} = \mu_{i,p}^n \times \delta_{x_{i,p}^a}$, where δ_x denotes the counting measure on the point x, and $X_{i,p}$ has the form $X_{i,p} = X_{i,p}^n \times \left\{ x_{i,p}^a \right\}$ accordingly, then type i agents are said to be *atomic* for the commodities included in $x_{i,p}^a$ and *non-atomic* for others. We partition $\{1, 2, \dots, C\} = A_i \cup N_i$ such that A_i, N_i denote the set of indexes of commodities for which type i agents are atomic and non-atomic, respectively.

The natural interpretation of atomic agents for a commodity is that the agents have so much information about that commodity market that they believe they can trade optimally (such as the utility maximizing trade subject to the budget constraint), thus reducing their offer sets to single points, namely the optima.[6]

2.2 Equilibrium

As the interpretation of the economy is modified, so is the definition of the equilibrium. Before mentioning the new definition, we remark that the *entropy* of a collection of probability density functions $h = \{h_i\}_{i\in\mathcal{I}}$ (with respect to the reference measure $\mu_{i,p}$) and the associated *average transaction* are defined by

$$H_p[h] := -\sum_{i=1}^{I} w_i \int h_i \log h_i \mathrm{d}\mu_{i,p}, \tag{2.1}$$

$$\bar{x}_p[h] := \sum_{i=1}^{I} w_i \int x h_i \mathrm{d}\mu_{i,p}, \tag{2.2}$$

[6] We introduced the concept of atomic agents simply for the sake of mathematical completeness. Since we shall use only non-atomic agents in the applications in Section 4, the reader may, without losing the essence, focus to the case where the atomic part is empty.

respectively.[7] We follow Jaynes (1968); Toda (2009) to understand (2.1). Suppose that there are n_i agents belonging to type i, and let the total number of agents be $n = \sum_i n_i$ and the weight be $w_i = n_i/n$. In general, if two random variables X, Y are independent, we have $H(X, Y) = H(X) + H(Y)$ (see Cover and Thomas (2006)). Therefore, if agents are acting independently, the economy-wide entropy is

$$H = -\sum_{i=1}^{I} n_i \int h_i \log h_i \mathrm{d}\mu_{i,p}.$$

Dividing this expression by n, we obtain the per capita entropy (2.1). The rationale of (2.2) is the law of large numbers: if each type has a large number of agents acting independently (or a type is atomic), then the economy-wide per capita average transaction converges to (2.2) almost surely. Hence, the implicit assumption is that non-atomic agents act independently conditional on p and the action of atomic agents.

Next, we give the new definition of statistical equilibrium. Here, we classify statistical equilibria into two categories: *genuine* and *degenerate* equilibria. Genuine equilibria are precisely what Toda (2009) defined to be statistical equilibria. Degenerate equilibria are intuitively the asymptotic limit of genuine ones and correspond to Walrasian equilibria when the theory is applied to standard Walrasian economies. By introducing the notion of degenerate statistical equilibria, we can integrate the Walrasian equilibrium concept into the statistical one as we shall see later in Corollary 3.9.

Definition 2.3 (statistical equilibrium)**.** Let

$$\mathcal{E} = \left\{ \mathcal{I}, \{w_i\}_{i\in\mathcal{I}}, \{\mu_{i,p}\}_{i\in\mathcal{I}, p\in\Delta^{C-1}} \right\}$$

be a statistical economy with offer sets $X_{i,p} = \operatorname{supp} \mu_{i,p}$.

- A collection of probability density functions $h = \{h_i\}_{i\in\mathcal{I}}$ over offer sets and a pair of vectors $(p, \pi) \in \Delta^{C-1} \times \mathbb{R}^C_+$ are called a *genuine statistical equilibrium* if

[7] Unless otherwise specified, the integral sign always means the integration over the whole space.

1. $\{h_i\}_{i\in\mathcal{I}}$ solves the maximum entropy program (MEP) associated with the scarcity parameter p, i.e., $h = \{h_i\}_{i\in\mathcal{I}}$ solves

$$\max H_p[h] \text{ subject to } \bar{x}_p[h] \leq 0, \tag{2.3}$$

2. π is the *entropy price* for the MEP, i.e., π is the Lagrange multiplier to (2.3),[8]

3. π and p are collinear.[9]

The factor $T \in [0, \infty]$ such that $\pi = \frac{1}{T}p$ is referred to—using an analogy from physics—as the *economic temperature*.[10][11]

- A scarcity parameter p and a collection of points $\{x_i\}_{i\in\mathcal{I}}$, where $x_i \in \operatorname{cl}\operatorname{co} X_{i,p}$[12] for all i, are called a *degenerate statistical equilibrium* if

  4. $\sum_{i=1}^{I} w_i x_i \leq 0$, and $p_c = 0$ if $\sum_{i=1}^{I} w_i x_{ic} < 0$,
  5. for all $i \in \mathcal{I}$ and $x \in X_{i,p}$, we have $p'x \geq 0$.

- A genuine statistical equilibrium or a degenerate statistical equilibrium is simply called a *statistical equilibrium.*

Genuine equilibria are precisely what Toda (2009) defined to be statistical equilibria, which have already been justified in Toda (2009) by rational expectations or as stationary distributions in a dynamic economy. Thus we only need to justify the definition of degenerate equilibria. The rationale we provide is the analogy from Walrasian equilibrium theory. If p denotes the

[8]Since (2.3) is an optimization problem in some functional space, it is not obvious that the standard Karush-Kuhn-Tucker theorem applies. Appendix A in Toda (2009) provides the relevant theorems.

[9]Here we allow the possibility of $\pi = 0$ (thus $T = \infty$). This can happen when all goods are "bads." For instance, if the offer sets of all agents consist of single points in $-\mathbb{R}^C_{++}$, the solution of (2.3) is trivially autarkic and the entropy price (Lagrange multiplier) π is 0.

[10]π does not appear in the definition of degenerate equilibria. In that case we *define* $T = 0$.

[11]Although we have used the term "temperature," making an analogy between physics and economics (which has a long history, for reference refer to Samuelson (1966); Saslow (1999); Smith and Foley (2008) among others) is not our purpose.

[12]$\operatorname{cl} A$ and $\operatorname{co} A$ denote the closure and the convex hull of A respectively.

price system and $X_{i,p}$ is the set of transactions that are at least as desirable as the Walrasian excess demand, then conditions 4 and 5 precisely define the Walrasian equilibrium. To see this, condition 4 implies that the market clears and the price of a commodity in excess supply is zero; condition 5 implies that agents are optimizing.

A natural question arising from Definition 2.3 is whether the two equilibrium concepts—genuine and degenerate equilibria—are mutually exclusive. In general the answer is negative. However, if at least one $\mu_{i,p}$ is absolutely continuous with respect to the Lebesgue measure (which is almost always the case in application), then the two concepts are mutually exclusive.[13]

3 Main Results

Our main results are divided into four parts. First, we show that solving the minimum log-partition program (MLPP, which we shall define shortly in Definition 3.2) is sufficient for solving the maximum entropy program (MEP). This proposition has already been proved (Toda, 2009, Proposition 3.3), but our proof is simpler since we do not use Gâteaux derivatives. We also improve the sufficient condition for the existence of a solution to MLPP from what was known before (Toda, 2009, Proposition 3.4). These results are crucial to prove the equilibrium existence theorems.

Second, we state and prove the equilibrium existence theorems. In these theorems, we weaken most of the assumptions of our previous result (Toda, 2009, Theorems 4.1 and 4.2) and strengthen others (such as the uniform boundedness of offer sets). Note, however, that the present formulation has a broader range of applicability. Indeed, we can apply these theorems to prove the existence of Walrasian equilibria in a standard competitive environment.

Third, we briefly discuss the computational aspect of statistical equilibria and provide a result that is general enough to analyze many economic situa-

[13] To see this, if a statistical equilibrium is genuine as well as degenerate, then by conditions 1 and 5, we must have $p'x = 0$ almost surely with respect to the probability measure induced by the equilibrium distributions. However, this is impossible since the hyperplane $p'x = 0$ has Lebesgue measure 0.

tions, yet special enough to obtain the equilibrium in an almost closed-form solution.

Fourth, we deduce a striking corollary: *Walrasian equilibria are special cases of statistical equilibria.* This corollary gives a rigorous mathematical foundation to the observation in Foley (2003):

> [...] there may be a sense in which Walrasian equilibrium can be viewed as an asymptotic approximation to statistical equilibrium.

3.1 MEP and MLPP

Proposition 3.1. *Let $\{\mu_i\}_{i\in\mathcal{I}}$ be a collection of regular Borel measures on $\mathbb{R}^C$. Suppose that there exist a collection of probability density functions $\{h_i^*\}_{i\in\mathcal{I}}$ and $\pi\in\mathbb{R}^C_+$ such that $h_i^*(x)=C_i\mathrm{e}^{-\pi' x}$ for some $C_i>0$,*

$$\bar{x}[h^*]:=\sum_{i=1}^I w_i\int xh_i^*(x)\mathrm{d}\mu_i\le 0,$$

and $\pi'\bar{x}[h^*]=0$. Then, $\{h_i^*\}_{i\in\mathcal{I}}$ solves the MEP (2.3).

Proof. Let $h=\{h_i\}_{i\in\mathcal{I}}$ be any collection of probability density functions such that $\bar{x}[h]\le 0$. Applying Lemma B.1 for $a=h_i$ and $x=h_i^*$, we obtain

$$\begin{aligned}H[h]&=-\sum_{i=1}^I w_i\int h_i\log h_i\mathrm{d}\mu_i\le\sum_{i=1}^I w_i\int\left(h_i^*-h_i-h_i\log h_i^*\right)\mathrm{d}\mu_i\\&=-\sum_{i=1}^I w_i\int h_i\log h_i^*\mathrm{d}\mu_i=\sum_{i=1}^I w_i\int h_i(\pi' x-\log C_i)\mathrm{d}\mu_i\\&=\pi'\bar{x}[h]-\sum_{i=1}^I w_i\log C_i\le-\sum_{i=1}^I w_i\log C_i.\end{aligned}$$

Since equality occurs if and only if $h_i(x)=h_i^*(x)$ μ_i-a.e. for all i, we are done. □

Toda (2009) defines the minimum log-partition program as follows.

Definition 3.2. Let $\{\mu_i\}_{i\in\mathcal{I}}$ be a collection of regular Borel measures on $\mathbb{R}^C$. The functions

$$Z_i(\xi) := \int \mathrm{e}^{-\xi' x}\mu_i(\mathrm{d}x),$$
$$Q(\xi) := \sum_{i=1}^{I} w_i \log\left(\int \mathrm{e}^{-\xi' x}\mu_i(\mathrm{d}x)\right)$$

are called the *partition function* and the *log-partition function* respectively. The optimization problem

$$\min_{\xi\geq 0} Q(\xi) \tag{3.1}$$

is called the *minimum log-partition program* (MLPP).

Proposition 3.3. *Let $\{\mu_i\}_{i\in\mathcal{I}}$ be a collection of regular Borel measures on $\mathbb{R}^C$ such that μ_i is supported on $X_i \subset \mathbb{R}^C$. Suppose that $Q(\xi) < \infty$ for some $\xi \in \mathbb{R}^C_+$. Then, the MLPP (3.1) has a solution if*

$$\left(\sum_{i=1}^{I} w_i \operatorname{co} X_i\right) \cap (-\mathbb{R}^C_{++}) \neq \emptyset. \tag{3.2}$$

Conversely, if

$$\left(\sum_{i=1}^{I} w_i \operatorname{co} X_i\right) \cap (a-\mathbb{R}^C_{++}) = \emptyset \tag{3.3}$$

for some $a \gg 0$, then the MLPP (3.1) has no solutions.

Proof. Since Q is lower semi-continuous by Proposition B.2, we only need to know the behavior of Q as $\xi \to \infty$.

If (3.2) holds, take $\epsilon > 0$ and $x_i \in \operatorname{co} X_i$ such that $\sum_{i=1}^{I} w_i x_i \leq -\epsilon\mathbf{1} \ll 0$. By Carathéodory's theorem (Rockafellar, 1970, p. 155), we can express x_i as a convex combination of at most $C+1$ points: $x_i = \sum_{k=1}^{C+1} \alpha_i^k x_i^k$, where $x_i^k \in X_i$, $\alpha_i^k \geq 0$, and $\sum_{k=1}^{C+1} \alpha_i^k = 1$. For each i and k, take an open neighborhood U_i^k of x_i^k small enough such that $\sum_i w_i \sum_k \alpha_i^k U_i^k \subset -\mathbb{R}^C_{++}$. By assumption, we have $\mu_i(U_i^k \cap X_i) > 0$ for all i, k. Thus, we obtain

$$v_i(\xi) := \inf\{\xi' x : x \in X_i\} \leq \min_k \xi' x_i^k \leq \sum_{k=1}^{C+1} \alpha_i^k \xi' x_i^k = \xi' x_i. \tag{3.4}$$

By (3.4) and Proposition B.3, it follows that if $\xi \geq 0$ and $\|\xi\|_1 = 1$, then

$$\lim_{t\to\infty} \frac{1}{t} Q(t\xi) = -\sum_{i=1}^{I} w_i v_i(\xi) \geq -\xi' \sum_{i=1}^{I} w_i x_i \geq \xi' \epsilon \mathbf{1} = \epsilon \|\xi\|_1 = \epsilon. \tag{3.5}$$

By (3.5), we get $Q(t\xi) \to \infty$ as $t \to \infty$. Therefore, we may restrict our search of the infimum of $Q(\xi)$ to a compact set, but since $Q(\xi)$ is lower semi-continuous, the minimum is attained.

Conversely, suppose that (3.3) holds. Since $\sum_{i=1}^{I} w_i \operatorname{co} X_i$ and $a - \mathbb{R}^C_{++}$ are both convex and the latter contains an interior point, by the separating hyperplane theorem (Rockafellar, 1970, p. 97), there exists $\xi \in \mathbb{R}^C_+ \setminus \{0\}$ such that $\xi' x \geq \xi' a$ for all $x \in \sum_{i=1}^{I} w_i \operatorname{co} X_i$. Since $\xi > 0$ and $a \gg 0$, we have $\xi' a > 0$. By a similar argument as above, we obtain

$$\lim_{t\to\infty} \frac{1}{t} Q(t\xi) = -\sum_{i=1}^{I} w_i v_i(\xi) \leq -\xi' a < 0.$$

Consequently, $Q(t\xi) \to -\infty$ as $t \to \infty$. Since by definition $Q(\xi) > -\infty$, we get $\inf_{\xi \geq 0} Q(\xi) = -\infty$; however, the infimum is never attained. □

Corollary 3.4. *Let everything be as in Proposition 3.3 and suppose that for all $i \in \mathcal{I}$, $\int \mathrm{e}^{-\xi' x} \mathrm{d}\mu_i$ is differentiable with respect to ξ under the integral sign.*[14] *If (3.2) holds, then the MEP (2.3) has a solution given by $h_i(x) = \mathrm{e}^{-\pi' x} / Z_i(\pi)$, where $\pi \in \mathbb{R}^C_+$ is the solution of the MLPP (3.1). Furthermore, if $h = \{h_i\}_{i \in \mathcal{I}}$, then we have $H[h] = Q(\pi)$.*

Proof. By Proposition 3.3, the MLPP (3.1) has a solution $\xi = \pi$. Let

$$\mathcal{L}(\xi, \lambda) = \sum_{i=1}^{I} w_i \log \left(\int \mathrm{e}^{-\xi' x} \mathrm{d}\mu_i \right) - \lambda' \xi$$

be the Lagrangian of (3.1), where $\lambda \in \mathbb{R}^C_+$ is the Lagrange multiplier. Since $Q(\xi)$ is convex by Proposition B.2, the Karush-Kuhn-Tucker condition implies

$$\sum_{i=1}^{I} w_i \frac{\int x \mathrm{e}^{-\pi' x} \mathrm{d}\mu_i}{\int \mathrm{e}^{-\pi' x} \mathrm{d}\mu_i} = -\lambda \tag{3.6}$$

[14] See Proposition B.5 in Toda (2009) for a sufficient condition.

and $\lambda'\pi = 0$. If we define $C_i^{-1} = Z_i(\pi) = \int \mathrm{e}^{-\pi' x} \mathrm{d}\mu_i > 0$, then by (3.6), all assumptions of Proposition 3.1 are satisfied. Thus, the MEP has a solution.

Substituting $h_i = \mathrm{e}^{-\pi' x}/Z_i(\pi)$ into (2.1) and using (3.6) and $\lambda'\pi = 0$, we immediately obtain $H[h] = Q(\pi)$. □

3.2 Existence of Statistical Equilibrium

The main theorems of this paper are Theorems 3.5 and 3.7 below. We first prove Theorem 3.5, which shows the existence of statistical equilibria when all measures $\{\mu_{i,p}\}$ are finite: assuming finiteness simplifies the argument. In practice, we typically wish to deal with infinite measures such as the Lebesgue measure. By assuming that the measures grow at most exponentially, Theorem 3.7 guarantees the existence of statistical equilibria in such cases.

Theorem 3.5. *Let $\mathcal{E} = \left\{\mathcal{I}, \{w_i\}_{i\in\mathcal{I}}, \{\mu_{i,p}\}_{i\in\mathcal{I}, p\in\Delta^{C-1}}\right\}$ be a statistical economy with offer sets $X_{i,p} = X_{i,p}^n \times \left\{x_{i,p}^a\right\} = \operatorname{supp}\mu_{i,p}$, where $x_{i,p}^a$ and $X_{i,p}^n$ denote the atomic and non-atomic part, respectively. Let $\mu_{i,p} = \mu_{i,p}^n \times \delta_{x_{i,p}^a}$ accordingly. Consider the following assumptions:*

1. *for all $i \in \mathcal{I}$ and $p \in \Delta^{C-1}$, we have $\mu_{i,p}(\mathbb{R}^C) < \infty$,*
2. *for all $i \in \mathcal{I}$, the partition $\{1, 2, \ldots, C\} = A_i \cup N_i$ of commodities for which type i agents are atomic and non-atomic is independent of $p \in \Delta^{C-1}$,*
3. *$X_{i,p}$ is uniformly bounded below, i.e., there exists $a \in \mathbb{R}^C$ such that $x \geq a$ for all $i \in \mathcal{I}$, $p \in \Delta^{C-1}$, and $x \in X_{i,p}$,*
4. *for all $p \in \Delta^{C-1}$, we have*
   - *(a) for all $i \in \mathcal{I}$, $\inf\{p'x : x \in X_{i,p}\} \leq 0$,*
   - *(b) $\sum_{i=1}^I w_i \inf\{p'x : x \in X_{i,p}\} < 0$,*
5. *for all $i \in \mathcal{I}$, the mapping $p \mapsto x_{i,p}^a$ is continuous,*
6. *for all $i \in \mathcal{I}$, the following conditions are satisfied:*

(a) *there exists a fixed finite regular Borel measure* μ_i^n *on* $\mathbb{R}^{N_i}$ *such that the Radon-Nikodym derivative* $\frac{\mathrm{d}\mu_{i,p}^n}{\mathrm{d}\mu_i^n}$ *(as a function of* p*) is continuous in* p *for* μ_i^n*-a.e.* x*,*

(b) *for all* $p \in \Delta^{C-1}$*, there exists a measurable function* $g_{i,p}$ *and a neighborhood* U_p *of* p *such that* $\left|\frac{\mathrm{d}\mu_{i,q}^n}{\mathrm{d}\mu_i^n}\right| \leq g_{i,p}$ *for all* $q \in U_p$ *and* μ_i^n*-a.e.* x*, and we have* $\int \left(1 + \sum_{c \in N_i} |x_c|\right) g_{i,p} \mathrm{d}\mu_i^n < \infty$*,*

7. *for all* $i \in \mathcal{I}$*, the graph of the correspondence* $p \mapsto \operatorname{cl}\operatorname{co} X_{i,p}$ *is closed.*

Then,

- *Under assumptions 1–3, 4a, and 5–7,* $\mathcal{E}$ *has a statistical equilibrium.*
- *Under assumptions 1–3, 4b, 5, and 6,* $\mathcal{E}$ *has a genuine statistical equilibrium but no degenerate equilibria.*

Only assumptions 3, 4a, and 4b are economically relevant; the others are either about boundedness (of transactions and integrals) or continuity, which are necessary for the rigorous mathematical argument but are difficult to interpret in terms of economics.

Assumption 3 means that agents' net transactions are bounded below. This is a constrained form of free disposal: agents are able to throw away undesired commodities, but only up to a certain finite amount. It may also be interpreted as limited arbitrage (Chichilnisky, 1995). This assumption is not restrictive in practice at all, for in the world there is only a finite amount of everything and hence there is no reason to think that agents wish to dispose of commodities beyond that limit.

Assumption 4a implies that agents are "realistic" in the sense that they essentially put a positive subjective probability on their budget set $\left\{x \in \mathbb{R}^C : p'x \leq 0\right\}$ in Walrasian environment, and assumption 4b is its variant. Assumption 4a is crucial, for the market can never clear if all agents perceive trade opportunities only for tremendously large amounts.

Assumption 6b is equivalent to

6b'. for all $i \in \mathcal{I}$ and $p \in \Delta^{C-1}$, there exists a neighborhood U_p of p such that

$$\int \left(1 + \sum_{c \in N_i} |x_c|\right) \sup_{q \in U_p} \left|\frac{\mathrm{d}\mu_{i,q}^n}{\mathrm{d}\mu_i^n}\right| \mathrm{d}\mu_i^n < \infty.$$

To see why, see Proposition B.4.

The outline of the proof of Theorem 3.5 is as follows. Since we know from Propositions 3.1, 3.3 and Corollary 3.4 that solving MLPP is enough for maximizing entropy, we wish to construct a correspondence from p to the Lagrange multiplier for MLPP and show the existence of a fixed point (which was precisely the idea in Toda (2009)). However, Proposition 3.3 shows that MLPP may not always have a solution. To overcome this difficulty, we bound MLPP by a constant and define a quasi equilibrium concept.

Definition 3.6 (b-quasi equilibrium)**.** Let $b > 0$. The pair of vectors $(p, \pi) \in \Delta^{C-1} \times \mathbb{R}_+^C$ is said to be a *b-quasi equilibrium* if

1. $\xi = \pi$ solves

$$\min Q_p(\xi) := \sum_{i=1}^{I} w_i \log \left(\int \mathrm{e}^{-\xi' x} \mathrm{d}\mu_{i,p} \right) \text{ subject to } \xi \geq 0, \|\xi\|_1 \leq b,$$

2. π and p are collinear.

Thus, the definition of a b-quasi equilibrium is the same as that of a degenerate statistical equilibrium except for that the former solves the MLPP bounded by b. We show by a standard fixed point argument that a b-quasi equilibrium always exists. Thus, we can take a sequence of b-quasi equilibria such that $b \to \infty$. We then show that either some b-quasi equilibrium becomes a genuine statistical equilibrium, or a subsequence of b-quasi equilibria converges to a degenerate statistical equilibrium.

Proof of Theorem 3.5. We first note that by assumption 3, we have $\mathrm{e}^{-\xi' x} \leq \mathrm{e}^{-\xi' a}$ for all $\xi \geq 0$, so $\int \mathrm{e}^{-\xi' x} \mathrm{d}\mu_{i,p} < \infty$ by assumption 1. Thus, together with assumptions 5 and 6, we can show that the integrals $\int \mathrm{e}^{-\xi' x} \mathrm{d}\mu_{i,p}$ are continuously differentiable with respect to ξ by applying Lebesgue's convergence theorem to the non-atomic parts and by direct computation for the atomic parts.[15]

Step 1. *For all $b > 0$, a b-quasi equilibrium exists.*

Since $\mu_{i,p}$ is a finite Borel measure, by assumption 3 we have $Q_p(\xi) < \infty$ whenever $\xi \geq 0$. By Proposition B.2, $Q_p(\xi)$ is convex and lower semi-continuous in ξ. Since the set $X_b := \left\{\xi \in \mathbb{R}^C_+ : \|\xi\|_1 \leq b\right\}$ is compact and convex, the set

$$\Pi(p) := \arg\min_{\xi \in X_b} Q_p(\xi)$$

is nonempty, compact and convex. If $0 \in \Pi(p)$ for some $p \in \Delta^{C-1}$, since p and 0 are collinear, obviously $(p, 0)$ is a b-quasi equilibrium (actually a genuine statistical equilibrium).

Assume $0 \notin \Pi(p)$ for all p. By the remark at the beginning of the proof, $Q_p(\xi)$ is continuous in (p, ξ) on $\Delta^{C-1} \times \mathbb{R}^C_+$. Thus, by Berge's maximum theorem (Berge, 1959, p. 116), $\Pi : \Delta^{C-1} \rightrightarrows X_b$ is upper semi-continuous. Define $\Phi(p) := \{\xi / \|\xi\|_1 : \xi \in \Pi(p)\}$. Since $0 \notin \Pi(p)$, $\Phi(p)$ is well-defined.

Let us show that $\Phi : \Delta^{C-1} \rightrightarrows \Delta^{C-1}$ is nonempty, compact, convex and upper semi-continuous. $\Phi(p) \neq \emptyset$ is trivial. Since $\Phi(p)$ is the intersection of Δ^{C-1} (a convex set) and the convex cone generated by $\Phi(p)$, it is convex. If $p_n \to p$, $q_n \in \Phi(p_n)$ and $q_n \to q$, take a sequence $\{\xi_n\} \subset \Pi(p_n)$ such that $q_n = \xi_n / \|\xi_n\|_1$. Since $\{\xi_n\} \subset X_b$ and X_b is compact, $\{\xi_n\}$ has a convergent subsequence $\xi_{n_k} \to \xi$. Since $p \mapsto \Pi(p)$ is upper semi-continuous, we have $\xi \in \Pi(p)$, so $q = \xi / \|\xi\|_1 \in \Phi(p)$. Thus, $p \mapsto \Phi(p)$ is upper semi-continuous. In particular, by letting $p_n = p$ for all n, it follows that $\Phi(p)$ is closed, but since $\Phi(p) \subset \Delta^{C-1}$, it is compact.

By Kakutani's fixed point theorem, there exists $p^* \in \Delta^{C-1}$ such that $p^* \in \Phi(p^*)$. Thus, there exists $t > 0$ such that $tp^* \in \Pi(p)$, so (p^*, tp^*) is a b-quasi equilibrium.

Step 2.

[15] Lebesgue's convergence theorem does not apply to the counting measure in our context, which is why we needed to introduce atomic parts and treat them separately.

Let $\{b_n\}_{n=1}^{\infty} \subset (0, \infty)$ be a monotone increasing sequence tending to ∞. By passing to a subsequence if necessary, we may assume that for each n, there exists a b_n-quasi equilibrium (p_n, π_n) such that p_n converges to some $p \in \Delta^{C-1}$. If $\pi_n \in \arg\min_{\xi \geq 0} Q_{p_n}(\xi)$ for some n, then by Corollary 3.4 (p_n, π_n) is a genuine statistical equilibrium. Therefore, without loss of generality we may assume that for all n, $\min_{\xi \geq 0} Q_{p_n}(\xi)$ has no solutions. Let

$$\mathcal{L}(\xi, \lambda_n, \theta_n) = \sum_{i=1}^{I} w_i \log \left(\int \mathrm{e}^{-\xi' x} \mathrm{d}\mu_{i,p_n} \right) - \lambda_n' \xi + \theta_n (\|\xi\|_1 - b_n)$$

be the Lagrangian of $\min_{\xi \in X_{b_n}} Q_{p_n}(\xi)$, where $\lambda_n \in \mathbb{R}_+^C$ and $\theta_n \geq 0$ are Lagrange multipliers. By the Karush-Kuhn-Tucker theorem, we obtain

$$\sum_{i=1}^{I} w_i \frac{\int -x \mathrm{e}^{-\pi_n' x} \mathrm{d}\mu_{i,p_n}}{\int \mathrm{e}^{-\pi_n' x} \mathrm{d}\mu_{i,p_n}} - \lambda_n + \theta_n \mathbf{1} = 0 \tag{3.7}$$

and $\lambda_n' \pi_n = 0$. Since (p_n, π_n) is not a statistical equilibrium but a b_n-quasi equilibrium, $\|\pi_n\|_1 \leq b_n$ is binding for all n.

Let us show that $\lim_{n \to \infty} \theta_n = 0$. Let $\pi_n = t_n p_n$. Since the constraint $\|\pi_n\|_1 \leq b_n$ is binding, we have $t_n = \langle t_n p_n, \mathbf{1} \rangle = \langle \pi_n, \mathbf{1} \rangle = b_n$. Multiplying $\pi_n = t_n p_n$ as an inner product to (3.7) and dividing both sides by $t_n > 0$, it follows from $\lambda_n' \pi_n = 0$ that

$$\sum_{i=1}^{I} w_i \frac{\int -p_n' x \mathrm{e}^{-t_n p_n' x} \mathrm{d}\mu_{i,p_n}}{\int \mathrm{e}^{-t_n p_n' x} \mathrm{d}\mu_{i,p_n}} + \theta_n = 0. \tag{3.8}$$

Regard $Q_p(t\xi) = \sum_{i=1}^{I} w_i \log \left(\int \mathrm{e}^{-t\xi' x} \mathrm{d}\mu_{i,p} \right)$ as a function of t. Since Q_p is convex, $Q_p'(t\xi)$ is increasing in t. Take any $t > 0$ and choose n sufficiently large such that $t_n > t$. Then, by (3.8) we obtain

$$\theta_n = -Q_{p_n}'(t_n p_n) \leq -Q_{p_n}'(t p_n) = \sum_{i=1}^{I} w_i \frac{\int p_n' x \mathrm{e}^{-t p_n' x} \mathrm{d}\mu_{i,p_n}}{\int \mathrm{e}^{-t p_n' x} \mathrm{d}\mu_{i,p_n}}. \tag{3.9}$$

Letting $n \to \infty$ in (3.9), by the remark at the beginning of the proof, it follows that

$$\limsup_{n\to\infty} \theta_n \le \sum_{i=1}^{I} w_i \frac{\int p'x \mathrm{e}^{-tp'x} \mathrm{d}\mu_{i,p}}{\int \mathrm{e}^{-tp'x} \mathrm{d}\mu_{i,p}} = -Q_p'(tp). \tag{3.10}$$

Since t is arbitrary in (3.10), for any $s > 0$, take $t_s > 0$ such that

$$\frac{Q_p(sp) - Q_p(0)}{s} = Q_p'(t_s p),$$

which is of course possible by the mean value theorem. Then, (3.10) becomes

$$\limsup_{n\to\infty} \theta_n \le -\frac{Q_p(sp) - Q_p(0)}{s}. \tag{3.11}$$

Letting $s \to \infty$ in (3.11), by Proposition B.3 and assumption 4a or 4b we obtain

$$\limsup_{n\to\infty} \theta_n \le \sum_{i=1}^{I} w_i \inf \{p'x : x \in X_{i,p}\} \le 0. \tag{3.12}$$

Since $\theta_n \ge 0$, we have $\theta_n \to 0$.

Step 3. *Under assumptions 1–3, 4a and 5–7, if $\mathcal{E}$ has no genuine statistical equilibria, then $\mathcal{E}$ has a degenerate statistical equilibrium.*

Define the sequence $\{x_i^n\}_{n=1}^{\infty} \subset \mathbb{R}^C$ by

$$x_i^n = \frac{\int x \mathrm{e}^{-t_n p_n' x} \mathrm{d}\mu_{i,p_n}}{\int \mathrm{e}^{-t_n p_n' x} \mathrm{d}\mu_{i,p_n}}.$$

Since $\mathrm{e}^{-t_n p_n' x} / \int \mathrm{e}^{-t_n p_n' x} \mathrm{d}\mu_{i,p_n}$ is a probability density function, $x_i^n \in \operatorname{co} X_{i,p_n}$. Since by assumption 3 we have $x_i^n \ge a$, it follows from $\lambda_n \ge 0$ and (3.7) that

$$a \le (1 - w_i)a + w_i x_i^n \le \sum_{i=1}^{I} w_i x_i^n \le \theta_n \mathbf{1} \implies a \le x_i^n \le \frac{\theta_n \mathbf{1} - (1 - w_i)a}{w_i}.$$

Since $\theta_n \to 0$, by taking a subsequence we may assume $x_i^n \to x_i$ for all i and for some $x_i \in \mathbb{R}^C$; hence $\sum_{i=1}^{I} w_i x_i^n \to \sum_{i=1}^{I} w_i x_i \le 0$. By assumption 7, $p \mapsto \operatorname{cl}\operatorname{co} X_{i,p}$ is closed, so we obtain $x_i \in \operatorname{cl}\operatorname{co} X_{i,p}$ for all i. By (3.8), we have $\sum_{i=1}^{I} w_i p_n' x_i^n = \theta_n$, so letting $n \to \infty$ we get $p' \sum_{i=1}^{I} w_i x_i = 0$. Hence, if

$\sum_{i=1}^{I} w_i x_{ic} < 0$, it must be $p_c = 0$ and condition 4 of Definition 2.3 holds.

By (3.12), we obtain $\sum_{i=1}^{I} w_i \inf \{p'x : x \in X_{i,p}\} = 0$, but by assumption 4a it must be $\inf \{p'x : x \in X_{i,p}\} = 0$ or $p'x \geq 0$ for all $x \in X_{i,p}$. Thus, condition 5 of Definition 2.3 holds. Therefore $(\{x_i\}_{i\in\mathcal{I}}, p)$ is a degenerate statistical equilibrium.

Step 4. *Under assumptions 1–3, 4b, 5 and 6, $\mathcal{E}$ has a genuine statistical equilibrium but no degenerate equilibria.*

If $\mathcal{E}$ has no genuine equilibria, by (3.12) and $\lim_{n\to\infty} \theta_n = 0$, we obtain $\sum_{i=1}^{I} w_i \inf \{p'x : x \in X_{i,p}\} = 0$. However, this contradicts to assumption 4b; hence a genuine equilibrium exists. There are no degenerate equilibria because assumption 4b and condition 5 of Definition 2.3 are mutually exclusive. □

There are two crucial ideas in the proof of Theorem 3.5. First, we deal with b-quasi equilibria instead of statistical equilibria and take limits. Second, we take the limit of $t_n p_n$ in (3.9) stepwise, namely first with respect to p_n and then t_n. Usually this kind of "partial limit argument" is unjustifiable, but in this case it is legitimate because we are concerned only with an upper bound as in (3.10) and Q is convex.

Theorem 3.7 below guarantees the existence of statistical equilibria when at least one of the measures $\{\mu_{i,p}\}$ is infinite.

Theorem 3.7. *Let everything be as in Theorem 3.5. For all $\epsilon \in \mathbb{R}^C_+$, define the measure $\mu^\epsilon_{i,p}$ by $\mu^\epsilon_{i,p}(\mathrm{d}x) = \mathrm{e}^{-\epsilon' x}\mu_{i,p}(\mathrm{d}x)$. Suppose that*

1' for all $p \in \Delta^{C-1}$, the following conditions are satisfied:

(a) for all $\epsilon \gg 0$ and $i \in \mathcal{I}$, $\mu^\epsilon_{i,p}(\mathbb{R}^C) < \infty$,

(b) for all $\epsilon \geq 0$ such that $\epsilon_c = 0$ for some $c \in \bigcup_{i\in\mathcal{I}} N_i$, there exists $i \in \mathcal{I}$ such that $\mu^\epsilon_{i,p}(\mathbb{R}^C) = \infty$,

6b' for all $p \in \Delta^{C-1}$, there exists a measurable function $g_{i,p}$ and a neighborhood U_p of p such that $\left|\frac{\mathrm{d}\mu^n_{i,q}}{\mathrm{d}\mu^n_i}\right| \leq g_{i,p}$ for all $q \in U_p$ and μ^n_i-a.e. x, and for all $\epsilon \gg 0$ we have $\int \mathrm{e}^{-\sum_{c\in N_i} \epsilon_c x_c} g_{i,p} \mathrm{d}\mu^n_i < \infty$.

Then, in addition to these assumptions,

- *Under assumptions 2, 3, 4a, 5, 6a, and 7 of Theorem 3.5, $\mathcal{E}$ has a statistical equilibrium.*
- *Under assumptions 2, 3, 4b, 5, and 6a of Theorem 3.5, $\mathcal{E}$ has a genuine statistical equilibrium but no degenerate equilibria.*

Remark. By the definition of $\mu_{i,p}^{\epsilon}$, assumption 1'a is trivially implied by assumption 6b'. However, in order to draw a parallel between Theorems 3.5 and 3.7, we mentioned assumption 1'a.

Proof. For all $\epsilon \gg 0$, define the statistical economy

$$\mathcal{E}^{\epsilon} := \left\{\mathcal{I}, \{w_i\}_{i\in\mathcal{I}}, \left\{\mu_{i,p}^{\epsilon}\right\}_{i\in\mathcal{I},p\in\Delta^{C-1}}\right\}.$$

Since $\mu_{i,p}^{\epsilon}$ is a finite measure by assumption 1'a, and since assumption 6b for $\mathcal{E}^{\epsilon}$ is implied by assumption 6b' for $\mathcal{E}$, by Theorem 3.5, $\mathcal{E}^{\epsilon}$ has a statistical equilibrium. If the equilibrium is degenerate, since $\mu_{i,p}$ and $\mu_{i,p}^{\epsilon}$ have a common support and the support (offer set) is all that matters in the definition of degenerate equilibria (see Definition 2.3), $\mathcal{E}$ also has a degenerate equilibrium. Hence, for all $\epsilon \gg 0$, we may assume that $\mathcal{E}^{\epsilon}$ has a genuine statistical equilibrium.

Let $\{\epsilon_n\} \subset \mathbb{R}_{++}^{C}$ be a sequence such that $\epsilon_n \to 0$ as $n \to \infty$. Let p_n and $\pi_n = t_n p_n$ be the equilibrium scarcity parameter and the entropy price of $\mathcal{E}^{\epsilon_n}$, respectively. By passing to a subsequence if necessary, we may assume $p_n \to p \in \Delta^{C-1}$ and $t_n \to t \in [0, \infty]$ as $n \to \infty$.

If $t_n \to \infty$, by the same argument as in Step 3 of the proof of Theorem 3.5, we can show that a degenerate statistical equilibrium exists.

If $t_n \to 0$ or $t_n \to t < \infty$ and $p_c = 0$ for some $c \in \bigcup_{i\in\mathcal{I}} N_i$, take any $\xi \in \mathbb{R}_{+}^{C}$ such that $\xi_c > 0$ for all $c \in \bigcup_{i\in\mathcal{I}} N_i$. Since $\pi_n = t_n p_n$ solves the MLPP, we obtain

$$\sum_{i=1}^{I} w_i \log\left(\int \mathrm{e}^{-t_n p_n' x - \epsilon_n' x} \mathrm{d}\mu_{i,p_n}\right) \leq \sum_{i=1}^{I} w_i \log\left(\int \mathrm{e}^{-\xi' x - \epsilon_n' x} \mathrm{d}\mu_{i,p_n}\right). \quad (3.13)$$

Taking $\liminf$ of (3.13), by Fatou's lemma and assumption 1'b, the left-hand side of (3.13) is bounded below by ∞. By assumption 6b, we may apply Lebesgue's convergence theorem to the non-atomic part and directly compute the atomic part of the right-hand side of (3.13) to obtain

$$\infty \leq \sum_{i=1}^{I} w_i \log \left(\int \mathrm{e}^{-\xi' x} \mathrm{d}\mu_{i,p} \right).$$

However, this contradicts to assumption 1'b. Therefore, it must be $t_n \to t \in (0, \infty)$ and $p_c > 0$ for all $c \in \bigcup_{i \in \mathcal{I}} N_i$. Again, letting $n \to \infty$ in (3.13), by either Lebesgue's convergence theorem or direct computation, we obtain

$$\sum_{i=1}^{I} w_i \log \left(\int \mathrm{e}^{-tp' x} \mathrm{d}\mu_{i,p} \right) \leq \sum_{i=1}^{I} w_i \log \left(\int \mathrm{e}^{-\xi' x} \mathrm{d}\mu_{i,p} \right), \tag{3.14}$$

or $Q_p(tp) \leq Q_p(\xi)$. (3.14) holds for all $\xi \in \mathbb{R}_+^C$ such that $\xi_c > 0$ for all $c \in \bigcup_{i \in \mathcal{I}} N_i$, but by assumption 1'b, it trivially holds even if $\xi_c = 0$ for some $c \in \bigcup_{i \in \mathcal{I}}$. Thus, (3.14) holds for all $\xi \in \mathbb{R}_+^C$, and as such tp solves the MLPP associated with p. Therefore, (p, tp) is a genuine statistical equilibrium.

If assumption 4b of Theorem 3.5 holds, by the same argument as in the proof of Theorem 3.5, $\mathcal{E}$ has no degenerate statistical equilibria. □

Since we have proved two existence theorems in this paper—Theorems 3.5 and 3.7—and Toda (2009) proved two others, it is worth mentioning the relationship between the four existence theorems. The fact is that none of the four theorems imply another, which can be seen as follows. First, Theorem 3.5 is independent from others because Theorem 3.5 assumes that all the measures are finite, whereas others assume that at least one measure is infinite. Second, Theorem 3.7 is not implied by the two in Toda (2009) because the latter two assume $\left(\sum_{i=1}^{I} w_i \operatorname{co} X_{i,p} \right) \cap (-\mathbb{R}_{++}^C) \neq \emptyset$, which the former does not. Third, Theorem 3.7 does not imply the two in Toda (2009) because the offer sets are uniformly bounded below in Theorem 3.7 but not necessarily so in Toda (2009). Finally, the two theorems in Toda (2009) are clearly independent.

Therefore, we have four different existence theorems of statistical equilibrium and thus the reader is free to choose which one to apply to specific models. However, in our view Theorem 3.7 is the most useful because the assumptions are weak as well as economically intuitive, and we generally wish to allow infinite measures.

3.3 Computation of Statistical Equilibrium

Since the exact evaluation of the integral $\int \mathrm{e}^{-\xi' x} \mathrm{d}\mu_i$ is impossible unless the offer set $X_i = \operatorname{supp} \mu_i$ has a simple structure such as a translation of the positive orthant $\mathbb{R}_+^C$ (or more generally, a polyhedron), Toda (2009) discusses a numerical algorithm to compute statistical equilibria.

Here we provide a result that is general enough to analyze many economic situations, yet special enough to obtain the equilibrium in an almost closed-form solution. In this case the offer spaces are simplest, namely the product measure of the counting measure and the Lebesgue measure restricted to translations of the positive orthant.

Corollary 3.8. *Let everything be as in Theorem 3.5. Suppose that for all $i \in \mathcal{I}$ and $p \in \Delta^{C-1}$, there exists a point $x_{i,p} = (x_{i,p}^n, x_{i,p}^a) \in \mathbb{R}^{N_i} \times \mathbb{R}^{A_i} = \mathbb{R}^C$ with the following properties:*

1. *the offer space $(X_{i,p}, \mu_{i,p})$ has the form $X_{i,p} = (x_{i,p}^n + \mathbb{R}_+^{N_i}) \times \left\{ x_{i,p}^a \right\}$ and $\mu_{i,p} = \mu_{i,p}^n \times \delta_{x_{i,p}^a}$, where $\mu_{i,p}^n$ is the restriction of the Lebesgue measure on $X_{i,p}^n = x_{i,p}^n + \mathbb{R}_+^{N_i}$ and $\delta_{x_{i,p}^a}$ is the counting measure,*
2. *for all $p \in \Delta^{C-1}$, we have*
   - *(a) $p' x_{i,p} \leq 0$ for all $i \in \mathcal{I}$,*
   - *(b) $p' \sum_{i=1}^{I} w_i x_{i,p} < 0$.*
3. *for all $i \in \mathcal{I}$, the mapping $p \mapsto x_{i,p}$ is continuous.*

Then,

- *Under assumptions 1, 2a, and 3, $\mathcal{E}$ has a statistical equilibrium.*
- *Under assumptions 1, 2b, and 3, $\mathcal{E}$ has a genuine statistical equilibrium but no degenerate equilibria.*

Furthermore, if $N_i \neq \emptyset$ for some i, then the economic temperature T is finite and the equilibrium can be obtained by solving the $C+1$ equations $\sum_{c=1}^{C} p_c = 1$ and

$$\forall c \in \bigcup_{i\in\mathcal{I}} N_i, \quad T \sum_{i:c\in N_i} w_i = -p_c \sum_{i=1}^{I} w_i x_{ic,p}, \tag{3.15a}$$

$$\forall c \in \bigcap_{i\in\mathcal{I}} A_i, \quad \sum_{i=1}^{I} w_i x_{ic,p} \leq 0,\ p_c \geq 0,\ p_c \sum_{i=1}^{I} w_i x_{ic,p} = 0, \tag{3.15b}$$

with $C+1$ unknowns $p \in \Delta^{C-1}$ and $T \geq 0$. In particular, if $A_i = \emptyset$ for all $i \in \mathcal{I}$ (thus all agent types are non-atomic for all commodities), then (3.15) becomes

$$\forall c, \quad T = -p_c \sum_{i=1}^{I} w_i x_{ic,p}. \tag{3.16}$$

Proof. If $N_i = \emptyset$ for all $i \in \mathcal{I}$, we can easily verify that $\mathcal{E}$ satisfies all assumptions of Theorem 3.5. (The least obvious assumption is the uniform boundedness assumption 3, but this follows from assumption 3 and the compactness of Δ^{C-1}.) If $N_i \neq \emptyset$ for some $i \in \mathcal{I}$, $\mathcal{E}$ satisfies all assumptions of Theorem 3.7. Thus, in either case there exists a statistical equilibrium. In the case of a degenerate equilibrium, (3.15) holds by setting $T = 0$. In the case of a genuine equilibrium, let p and π be the equilibrium scarcity parameter and entropy price, respectively. The log-partition function is

$$\begin{aligned} Q_p(\xi) &= \sum_{i=1}^{I} w_i \log \left(\prod_{c\in N_i} \int_{x_{ic,p}}^{\infty} \mathrm{e}^{-\xi_c x_c} \mathrm{d}x_c \times \prod_{c\in A_i} \mathrm{e}^{-\xi_c x_{ic,p}} \right) \\ &= \sum_{i=1}^{I} w_i \log \left(\prod_{c\in N_i} \frac{1}{\xi_c} \mathrm{e}^{-\xi_c x_{ic,p}} \times \prod_{c\in A_i} \mathrm{e}^{-\xi_c x_{ic,p}} \right) \\ &= -\sum_{i=1}^{I} w_i \left(\xi' x_{i,p} + \sum_{c\in N_i} \log \xi_c \right). \end{aligned} \tag{3.17}$$

If $N_i \neq \emptyset$ for some i, then by (3.17) we have $Q_p(0) = \infty$. Thus, the economic temperature defined by $\pi = \frac{1}{T} p$ must be finite. Then, (3.15) can be obtained by applying the Karush-Kuhn-Tucker theorem to the minimization of $Q_p(\xi)$ subject to $\xi \geq 0$ and using the relation $\pi = \frac{1}{T} p$ or $1/\pi_c = T/p_c$ imposed by

Definition 2.3. □

Since $\sum_{i=1}^{I} w_i x_{ic,p}$ is the infimum average transaction of commodity c, (3.16) implies that the value of the infimum average transaction evaluated at the equilibrium scarcity parameter p is common across all commodities and that their absolute values are equal to the economic temperature. If $T = 0$, then (3.15) is the condition of degenerate statistical equilibrium. Hence degenerate equilibria are indeed "degenerate" in the sense that the economic temperature is lowest, namely absolute zero.

3.4 Relationship between Walrasian and Statistical Equilibria

Now we are ready to answer the deep question: *how are Walrasian and statistical equilibria related?* Although Corollary 3.9 below is almost trivial by Theorem 3.5, it has a strong philosophical implication: *Walrasian equilibrium theory is contained in statistical equilibrium theory.*

Corollary 3.9. *A Walrasian equilibrium is a statistical equilibrium. More precisely, let $\mathcal{E} = \{\mathcal{I}, \{u_i\}, \{e_i\}\}$ be an endowment economy, where $u_i : \mathbb{R}_+^C \to \mathbb{R}$ is a continuous, locally non-satiated utility function of type i agents (with ratio $w_i > 0$), and the endowments satisfy $e_i \gg 0$ for all i. Then,*

1. *there exists a statistical economy $\mathcal{E}'$ such that all Walrasian equilibria of $\mathcal{E}$ are statistical equilibria of $\mathcal{E}'$,*
2. *the existence of Walrasian equilibria can be shown by using statistical equilibrium theory.*

Proof. Take $b > 0$ such that $\sum_{i=1}^{I} e_i \leq b\mathbf{1}$. Let $X_b = [0, b]^C$ and

$$X_{i,p} = \left\{ y \in X_b : u_i(y) \geq \max_{p'x \leq p'e_i} u_i(x) \right\} - e_i$$

be the set of transactions that make type i agents at least as well off as the best consumption bundle under the budget constraint. Obviously, $X_{i,p}$ is a compact set contained in $[-b, b]^C$ and the correspondence $p \mapsto X_{i,p}$ is closed.

Let μ be the Lebesgue measure on $\mathbb{R}^C$ and define $\mu_{i,p}$ by $\mathrm{d}\mu_{i,p} = \chi_{X_{i,p}} \mathrm{d}\mu$,[16] so $\mu_{i,p}$ is the restriction of the Lebesgue measure on $X_{i,p}$. Then, $\mathcal{E}' = \{\mathcal{I}, \{w_i\}, \{\mu_{i,p}\}\}$ is a statistical economy.

Now, any Walrasian equilibrium of $\mathcal{E}$ is a degenerate statistical equilibrium of $\mathcal{E}'$ by Definition 2.3, for in Walrasian equilibrium the price of the commodity in excess supply must be zero, and by local non-satiation agents must spend all their income. Thus, the first part of Corollary 3.9 is shown.

For the second part, since all agent types are non-atomic, assumptions 2 and 5 of Theorem 3.5 are trivial. Assumption 3 follows from $X_{i,p} \subset [-b, b]^C$, and assumption 4a is satisfied because the value of the Walrasian demand is at most the budget. Assumption 7 is satisfied since the correspondence $p \mapsto X_{i,p}$ is closed and $X_{i,p}$ is compact. Since u_i is locally non-satiated, the indifference curve has measure zero, so $\chi_{X_{i,p}}$ is continuous in p for μ-a.e. x, thus assumption 6 is satisfied.

Since all assumptions of Theorem 3.5 are satisfied, there exists a statistical equilibrium. Independent of whether the equilibrium is degenerate or not, the aggregate average transaction is nonpositive and the average transaction of type i agents belongs to $X_{i,p}$ because $X_{i,p}$ is compact. These points clearly define a Walrasian equilibrium. □

An alternative proof of Corollary 3.9 is to introduce atomic agents whose offer sets are the Walrasian excess demand functions (assuming they are unique) and apply Corollary 3.8. This proof is essentially identical to the standard ones as in Debreu (1987).

As is obvious by Corollary 3.9, we made no convexity assumptions in establishing the existence of statistical equilibria. Therefore, in proving the existence of Walrasian equilibria, we may completely dispose of convexity provided that the number of agents within each type is sufficiently large: in fact, by Carathéodory's theorem, it suffices to have at least $C + 1$ agents in each type.

This result neither implies nor is implied by the result of Aumann (1966). Our result rests on the assumption of a finite number of types, which Aumann

[16] χ_A denotes the characteristic function of A.

(1966) does not assume; on the other hand, we do not need the assumption of preference saturation that Aumann (1966) requires.

4 Applications

In this section, we provide a few applications of statistical equilibria.

4.1 Information Centralization as Government's Role

The commodity space is $\mathbb{R}^C$. There is only one agent type, consumers with initial endowment $e \in \mathbb{R}^C_{++}$ and Cobb-Douglas utility function $u(x_1, \dots, x_C) = \sum_{c=1}^C \alpha_c \log x_c$, where $\alpha_c > 0$ and $\sum_{c=1}^C \alpha_c = 1$. The Walrasian demand given price p is obviously $x_c(p) = \frac{\alpha_c w}{p_c}$, where $w = \sum_{c=1}^C p_c e_c$ denotes the wealth. Let $0 \le r \le 1$ be the "safety margin" and define the offer set by

$$X_p = \left\{ x \in \mathbb{R}^C : \forall c,\ x_c \ge (1-r)x_c(p) - e_c \right\}.$$

That is, the offer set is a translation of the positive orthant $\mathbb{R}^C_+$ and agents perceive the possibility of trading quantities smaller than the Walrasian demand by the factor $1-r$ due to uncertainty, limited liquidity, or whatever reasons. Assume that the measure is the Lebesgue measure restricted on X_p.

This example satisfies all assumptions of Corollary 3.8.[17] By (3.16), we obtain

$$\forall c, \quad T = -(1-r)\alpha_c w + p_c e_c. \tag{4.1}$$

Summing up (4.1) with respect to c and using the definition of w, we obtain $T = \frac{rw}{C}$. Substituting this back into (4.1) and using $\sum_{c=1}^C p_c = 1$, we finally

[17] Strictly speaking, the offer set is not well-defined when $p_c = 0$ for some c but it is easy to justify by a "box argument" as in the proof of Corollary 3.9.

get

$$w = \left[\sum_{c=1}^{C} \frac{1}{e_c}\left(\frac{r}{C} + (1-r)\alpha_c\right)\right]^{-1}, \tag{4.2a}$$

$$T = \frac{rw}{C}, \tag{4.2b}$$

$$p_c = \frac{1}{e_c}\left(\frac{r}{C} + (1-r)\alpha_c\right) w. \tag{4.2c}$$

As r tends to 0, by (4.2), both w and p are of the order $O(1)$ and $T = O(r)$. The entropy price $\pi = \frac{1}{T}p$ is $O(r^{-1})$. Since the equilibrium distribution is exponential with the exponent $\mathrm{e}^{-\pi' x}$, its variance is of the order $O(r^2)$. Hence the smaller the "safety margin" r is, i.e., the less uncertain and the more liquid the economy, the smaller horizontal inequality. In the limit of $r \to 0$, the equilibrium distribution converges in probability to the Walrasian transaction and horizontal *equality* is attained.

This simple example has an interesting policy implication. It is a common practice in many countries to impose progressive income tax to mitigate the unequal distribution of wealth, but according to statistical equilibrium theory, horizontal inequality is a necessity in an economy with some uncertainty or limited liquidity, in which the offer sets diverge from the "ideal" Walrasian ones. If we were to believe statistical equilibrium theory, redistributive tax policies fail to achieve their objectives.

If a government seeks to reduce horizontal inequality, there are two solutions. The first one—which we do not recommend—is to become a totalitarian state and enforce people specific transactions. The second one is to gather and publish information so that people can bring their offer sets as close as possible to the "ideal" situation, that is the Walrasian offer sets. Thus, from a normative point of view, an important role of the government is *information centralization*.

4.2 Gamma Distribution as a Model of Income Distribution

Let $\mathbb{R}^C$ be the commodity space and assume that one of the agent types is consumers. Suppose that the consumers' offer set is the positive orthant $\mathbb{R}^C_+$ and the measure is

$$\mu(\mathrm{d}x) = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_C^{\alpha_C} \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_C,$$

the Lebesgue measure multiplied by the power function. Then the partition function is

$$Z(\xi) = \prod_{c=1}^{C} \int_0^\infty x_c^{\alpha_c} \mathrm{e}^{-\xi_c x_c} \mathrm{d}x_c = \prod_{c=1}^{C} \xi_c^{-1-\alpha_c} \Gamma(1+\alpha_c).$$

Hence, the density of the equilibrium distribution of transactions for consumers (with respect to the Lebesgue measure) is

$$h(x)\frac{\mathrm{d}\mu}{\mathrm{d}x} = \prod_{c=1}^{C} \frac{\pi_c^{1+\alpha_c}}{\Gamma(1+\alpha_c)} x^{\alpha_c} \mathrm{e}^{-\pi_c x_c}, \tag{4.3}$$

the gamma density, where π is the entropy price. Since the expression (4.3) is multiplicatively separable, the distribution of transactions of any commodity is again a gamma distribution and is independent from other commodities. Thus, in an empirical study we may sometimes separate one commodity market from others and perform a partial equilibrium analysis.

The gamma distribution has been documented to fit the income distribution well (Salem and Mount, 1974), which can also be seen in Figure 1.[18] In Silver et al. (2002), the equilibrium distribution of transactions between agents with Cobb-Douglas preference becomes a gamma distribution, but this is an agent-based model. In contrast, our theory has a firm micro-foundation and yet deduces the same result.

[18] The gamma density in Figure 1 has been estimated by maximum likelihood.

5 Conclusion

We reformulated the concept of statistical equilibrium advanced by Foley (1994); Toda (2009), proved its existence and showed that it is "general general equilibrium." Our theory potentially has a broad range of applicability because: our formulation of statistical equilibrium theory is abstract; the assumptions to ensure the existence of equilibria are weak as well as economically intuitive; there exists a simple algorithm to numerically obtain the equilibria as proposed in Toda (2009); our theory contains Walrasian equilibrium theory as a special case.

Although statistical equilibrium theory (statistical economics) is more general than Walrasian equilibrium theory (neoclassical economics), this fact of course does not invalidate the use of Walrasian equilibrium theory. For instance, neoclassical economics provides us with strong positive results such as the Pareto efficiency of equilibria. Therefore, the two fields simply have different philosophies and should be applied wisely depending on the problem at hand.

A natural future research topic is to test statistical equilibrium theory empirically. In Section 4, we suggested that the equilibrium distribution for consumers be gamma. Since all equilibrium distributions (not necessarily the gamma distribution) arising in statistical equilibrium theory belong to the exponential family, the parameters can be easily estimated by maximum likelihood. Since Walrasian equilibrium theory has no predictive power regarding the distribution of outcome (e.g., income), statistical equilibrium theory has stronger implications in comparative statics, which can be exploited in empirical analyses.

A challenge of theoretical interest is to provide the micro-micro-foundation of statistical equilibrium theory. In the present formulation we took entropy maximization as axiom, but under what conditions do the ergodic distributions (of, say, a Markov chain) become the entropy-maximizing ones? A different but related theoretical issue is to consider a tâtonnement process converging to the equilibrium. Since the log-partition function $Q(\xi)$ is always convex, we can expect such a process to converge globally. If this could be done, we will not only obtain another computational algorithm but also a constructive proof of the existence of equilibria as hinted in Toda (2009).

We can introduce time and make the economy dynamic. Consider an economy with a single agent type that lives to the next period with probability $\beta < 1$. Let Ω_t be the information set at t. Then, the dynamic entropy maximization program can be written in the form

$$H(\Omega_t) = \max_h \left\{ -\int h \log h \mathrm{d}\mu_t + \beta \,\mathrm{E}[H(\Omega_{t+1})|\Omega_t] \right\}.$$

As such, we can possibly define a concept that should be naturally called *recursive statistical equilibrium*. Since the introduction of time necessarily makes the number of commodities infinite, yet another important issue is to prove the existence of equilibria in economies with infinitely many commodities. Methods as in Bewley (1972) might be helpful.

A Review of Information Theory

A good starting point to expose oneself to the materials of information theory that is related to statistical equilibrium theory is Chapters 2, 3, 12 of Cover and Thomas (2006). In the information theory literature, the entropy of a discrete distribution $\mathbf{p} = (p_1, \ldots, p_n)$ is defined by

$$H(\mathbf{p}) = -\sum_{i=1}^{n} p_i \log p_i \tag{A.1}$$

and the entropy of a continuous distribution with probability density function $f(x)$ is defined by

$$H(f) = -\int_{-\infty}^{\infty} f(x) \log f(x) \mathrm{d}x, \tag{A.2}$$

if the integral converges. (It is customary to define $0 \log 0 = 0$.) Although in practice discrete or continuous distributions are most often used, there are other probability distributions. Hence, if μ is a measure and P is a probability measure that is absolutely continuous with respect to μ, we say that the Radon-Nikodym derivative $f = \frac{\mathrm{d}P}{\mathrm{d}\mu}$ is the probability density function of

P with respect to μ, and we define the entropy of P (with respect to the reference measure μ) to be

$$H(f) = -\int f \log f \mathrm{d}\mu, \tag{A.3}$$

if the integral converges. Obviously, (A.1) and (A.2) are special cases of (A.3) when μ is either the counting measure or the Lebesgue measure on $\mathbb{R}$.

For the time being assume that $X_1, \dots, X_n$ are discrete i.i.d. random variables with probability mass function $p(x)$. By the law of large numbers, we can relate the logarithm of the joint probability $p(X_1, \dots, X_n) = p(X_1) \cdots p(X_n)$ to the entropy as

$$\begin{aligned} -\frac{1}{n} \log p(X_1, \dots, X_n) &= -\frac{1}{n} \sum_{i=1}^{n} p(X_i) \xrightarrow{\text{a.s.}} -\mathrm{E}[\log p(X)] \\ &= -\sum_{x} p(x) \log p(x) = H(\mathbf{p}). \end{aligned}$$

Hence we define the set

$$A_\epsilon^{(n)} := \left\{ (x_1, \dots, x_n) : \left| H(\mathbf{p}) + \frac{1}{n} \log p(x_1, \dots, x_n) \right| < \epsilon \right\}$$

to be the *typical set*. Letting $H = H(\mathbf{p})$, we can show (see Cover and Thomas (2006) for details) that the number of elements in the typical set is approximately e^{nH}. Hence the maximum entropy principle can be interpreted as making the typical set as large as possible or obtaining the macrostate that has the most microstates.

B Mathematical Appendix

Lemma B.1. *Let us define $0 \cdot (\pm\infty) = 0$. Then, for all $a, x \geq 0$, we have $x - a \log x \geq a - a \log a$ with equality if and only if $x = a$.*

Proof. If $a = 0$, the claim is obvious by definition. If $a > 0$, let $f(x) =$

$x - a\log x$. Obviously, $f(0) = \infty$ and f is C^∞ on $(0,\infty)$. Since

$$f''(x) = \left(1 - \frac{a}{x}\right)' = \frac{a}{x^2} > 0,$$

f is strictly convex. Since $f'(a) = 1 - \frac{a}{a} = 0$, f attains its unique minimum at $x = a$. Therefore,

$$x - a\log x = f(x) \geq f(a) = a - a\log a. \qquad \square$$

Proposition B.2. *Let μ be a regular Borel measure on $\mathbb{R}^C$ such that $\mu(\mathbb{R}^C) > 0$. Then,*

$$f(\xi) := \log\left(\int \mathrm{e}^{-\xi' x}\mu(\mathrm{d}x)\right)$$

is convex and lower semi-continuous.

Proof. See Proposition B.4 in Toda (2009). $\square$

Proposition B.3. *Let μ be a regular Borel measure on $\mathbb{R}^C$ such that $\mu(\mathbb{R}^C) > 0$ and supported on $X = \operatorname{supp}\mu$. If $\int \mathrm{e}^{-t\xi' x}\mathrm{d}\mu < \infty$ for some $t > 0$, then*

$$\lim_{t\to\infty}\frac{1}{t}\log\left(\int \mathrm{e}^{-t\xi' x}\mathrm{d}\mu\right) = -\inf\left\{\xi' x : x \in X\right\}.$$

Proof. The proof consists of two steps.

Step 1. *For all $\xi \in \mathbb{R}^C$, let $v(\xi) = \inf\left\{c : \mu(\left\{x \in \mathbb{R}^C : \xi' x \leq c\right\}) > 0\right\}$. Then,*

$$\lim_{t\to\infty}\frac{1}{t}\log\left(\int \mathrm{e}^{-t\xi' x}\mathrm{d}\mu\right) = -v(\xi). \tag{B.1}$$

Let $E_n = \left\{x \in \mathbb{R}^C : \xi' x \leq n\right\}$. If $\mu(E_n) = 0$ for all n, then

$$\mu(\mathbb{R}^C) = \mu\left(\bigcup_{n=1}^\infty E_n\right) = \lim_{n\to\infty}\mu(E_n) = 0,$$

which is a contradiction. Therefore $\mu(E_n) > 0$ for some n, hence $v(\xi) < \infty$.

Let us prove (B.1) when $v(\xi) > -\infty$. Let $v = v(\xi) \in \mathbb{R}$ and define

$$X^+ = \left\{x \in \mathbb{R}^C : \xi' x \geq v\right\},\ X^- = \left\{x \in \mathbb{R}^C : \xi' x < v\right\},\ X_n = \left\{x \in \mathbb{R}^C : \xi' x \leq v - \frac{1}{n}\right\}.$$

Since $X^- = \bigcup_{n=1}^{\infty} X_n$ and $\mu(X_n) = 0$ by the definition of $v(\xi)$, we have $\mu(X^-) = 0$. Obviously, $X^{\pm}$ are disjoint and $X^+ \cup X^- = \mathbb{R}^C$, so $\mu(X^+) = \mu(\mathbb{R}^C) > 0$. Fix $t_0 > 0$ such that $\int \mathrm{e}^{-t_0 \xi' x} \mathrm{d}\mu < \infty$. Then, for all $t > 0$ we obtain

$$\int \mathrm{e}^{-t\xi' x} \mathrm{d}\mu = \mathrm{e}^{-tv} \int_{X^+} \mathrm{e}^{-t(\xi' x - v)} \mathrm{d}\mu. \tag{B.2}$$

Denote the integral over X^+ in (B.2) by $I(t)$. Since $\xi' x \geq v$ for $x \in X^+$, $\mathrm{e}^{-t(\xi' x - v)}$ is decreasing in t, so $I(t)$ is decreasing in t. (In particular, $0 < I(t) < \infty$ for $t \geq t_0$.) Hence for $t \geq t_0$ we obtain

$$\frac{1}{t} \log \left(\int \mathrm{e}^{-t\xi' x} \mathrm{d}\mu \right) \leq -v + \frac{1}{t} \log I(t_0). \tag{B.3}$$

Letting $t \to \infty$ in (B.3), we obtain

$$\limsup_{t \to \infty} \frac{1}{t} \log \left(\int \mathrm{e}^{-t\xi' x} \mathrm{d}\mu \right) \leq -v. \tag{B.4}$$

Take any $\epsilon > 0$ and let $A = \{x \in X^+ : \xi' x \leq v + \epsilon\}$. By assumption and the definition of $X^{\pm}$, we have $\mu(A) = \mu(\left\{x \in \mathbb{R}^C : \xi' x \leq v + \epsilon\right\}) > 0$. By taking a compact subset of A if necessary, by the regularity of μ we may assume $0 < \mu(A) < \infty$. Therefore we obtain

$$\begin{aligned} \frac{1}{t} \log \left(\int \mathrm{e}^{-t\xi' x} \mathrm{d}\mu \right) &\geq \frac{1}{t} \log \left(\mathrm{e}^{-t(v+\epsilon)} \int_A \mathrm{e}^{-t(\xi' x - v - \epsilon)} \mathrm{d}\mu \right) \\ &\geq -(v + \epsilon) + \frac{1}{t} \log \mu(A). \end{aligned} \tag{B.5}$$

Letting $t \to \infty$ in (B.5) and noting that $\epsilon > 0$ is arbitrary, we obtain

$$\liminf_{t \to \infty} \frac{1}{t} \log \left(\int \mathrm{e}^{-t\xi' x} \mathrm{d}\mu \right) \geq -v. \tag{B.6}$$

By (B.4), (B.6), we obtain (B.1).

If $v(\xi) = -\infty$, for arbitrary $n > 0$ define $X_n = \{x \in X : \xi' x < -n\}$. By the definition of $v(\xi)$, we have $\mu(X_n) > 0$. Then, we obtain the same result as (B.5) with A replaced by X_n and $v + \epsilon$ replaced by $-n$. Since n is arbitrary, (B.1) is obvious.

Step 2. *For all $\xi \in \mathbb{R}^C$, we have $v(\xi) = \inf\{\xi' x : x \in X\}$.*

Let $u(\xi) = \inf\{\xi' x : x \in X\}$. For all $\epsilon > 0$, there exists an $x_0 \in X$ such that $u(\xi) + \epsilon > \xi' x_0$. Since $\xi' x$ is continuous in x, there exists an open neighborhood U of x_0 such that $x \in U$ implies $\xi' x < u(\xi) + \epsilon$. Since $\mu(U \cap X) > 0$ by assumption, it follows that $v(\xi) \le u(\xi) + \epsilon$. Since $\epsilon > 0$ is arbitrary, we obtain $v(\xi) \le u(\xi)$.

On the other hand, by the definition of v, we have $\mu(\{x \in X : \xi' x \le v(\xi) + \epsilon\}) > 0$. In particular, there exists an $x_0 \in X$ such that $\xi' x_0 \le v(\xi) + \epsilon$. Therefore,

$$u(\xi) = \inf\{\xi' x : x \in X\} \le \xi' x_0 \le v(\xi) + \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we obtain $u(\xi) \le v(\xi)$. Therefore $u(\xi) = v(\xi)$. □

Proposition B.4. *Let (X, μ) be a measure space and Y be a separable topological space. Let $f : X \times Y \to \mathbb{R} \cup \{\pm\infty\}$ be such that for all $y \in Y$, $f(\cdot, y) : X \to \mathbb{R} \cup \{\pm\infty\}$ is measurable, and for μ-a.e. $x \in X$, $f(x, \cdot) : Y \to \mathbb{R} \cup \{\pm\infty\}$ is lower semi-continuous. Then, $g(x) = \sup_{y \in Y} f(x, y)$ is measurable.*

Proof. Let D be a countable dense subset of Y and define $h(x) = \sup_{y \in D} f(x, y)$. Since D is countable, h is measurable. Since $D \subset Y$, we have $h(x) \le g(x)$. Therefore, in order to show that g is measurable, we only need to show that $g(x) \le h(x)$ μ-a.e. Fix any $x \in X$ such that $f(x, \cdot)$ is lower semi-continuous and take any $\epsilon > 0$. The set $U = \{y \in Y : f(x, y) > g(x) - \epsilon\}$ is open because $f(x, \cdot)$ is lower semi-continuous and is nonempty by the definition of g. Since $D \subset Y$ is dense, there exists $y_1 \in U \cap D$. Therefore,

$$g(x) - \epsilon < f(x, y_1) \le h(x),$$

so letting $\epsilon \to 0$ we obtain $g(x) \le h(x)$. □

References

AUMANN, R. J. (1966): "Existence of Competitive Equilibria in Markets with a Continuum of Traders," *Econometrica*, 34, 1–17.

Berge, C. (1959): *Espaces Topologiques*, Paris: Dunod, English edition: *Topological Spaces*. MacMillan, 1963.

Bewley, T. F. (1972): "Existence of Equilibria in Economies with Infinitely Many Commodities," *Journal of Economic Theory*, 4, 514–540.

Champernowne, D. G. (1953): "A Model of Income Distribution," *Economic Journal*, 63, 318–351.

Chichilnisky, G. (1995): "Limited arbitrage is necessary and sufficient for the existence of a competitive equilibrium with or without short sales," *Economic Theory*, 5, 79–107.

Cover, T. M. and J. A. Thomas (2006): *Elements of Information Theory*, Hoboken, NJ: John Wiley & Sons, second ed.

Debreu, G. (1987): *Theory of Value*, Cowles Foundation Monograph 17, Yale University Press.

Foley, D. K. (1994): "A Statistical Equilibrium Theory of Markets," *Journal of Economic Theory*, 62, 321–345.

——— (2003): "Statistical Equilibrium in Economics: Method, Interpretation, and an Example," in *General Equilibrium: Problems and Prospects*, ed. by F. Petri and F. Hahn, London and New York: Routledge, chap. 4.

Folland, G. B. (1999): *Real analysis: modern techniques and their applications*, Hoboken, NJ: John Wiley & Sons, second ed.

Green, J. R. and M. Majumdar (1975): "The Nature of Stochastic Equilibria," *Econometrica*, 43, 647–660.

Grossman, S. J. and J. E. Stiglitz (1980): "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 70, 393–408.

Jaynes, E. T. (1968): "Prior Probabilities," *IEEE Transactions on Systems Science and Cybernetics*, SSC-4, 227–241.

Krebs, T. (1997): "Statistical Equilibrium in One-Step Forward Looking Economic Models," *Journal of Economic Theory*, 73, 365–394.

Pareto, V. (1896): *Cours d'Economie Politique*, Lausanne: F. Rouge.

Rockafellar, R. T. (1970): *Convex Analysis*, Princeton, NJ: Princeton University Press.

Salem, A. B. Z. and T. D. Mount (1974): "A Convenient Descriptive Model of Income Distribution: The Gamma Density," *Econometrica*, 42, 1115–1127.

Samuelson, P. A. (1966): "Structure of a Minimum Equilibrium System," in *The Collected Scientific Papers of Paul A. Samuelson*, ed. by J. E. Stiglitz, Cambridge, MA: MIT Press, 651–686.

Saslow, W. M. (1999): "An Economic Analogy of Thermodynamics," *American Journal of Physics*, 67, 1239–1247.

Silver, J., E. Slud, and K. Takamoto (2002): "Statistical Equilibrium Wealth Distributions in an Exchange Economy with Stochastic Preferences," *Journal of Economic Theory*, 106, 417–435.

Simon, H. A. (1959): "Theories of Decision-Making in Economics and Behavioral Science," *American Economic Review*, 49, 253–283.

Smith, E. and D. K. Foley (2008): "Classical Thermodynamics and Economic Equilibrium Theory," *Journal of Economic Dynamics and Control*, 32, 7–65.

Toda, A. A. (2009): "Existence of a Statistical Equilibrium for an Economy with Endogenous Offer Sets," *Economic Theory*, DOI: 10.1007/s00199-009-0493-6.