Efficient Statistical Equilibria in Markets

Kurt Jörnsten and Jan Ubøe

Norwegian School of Economics and Business Administration
Helleveien 30, N-5045 Bergen, Norway.

ABSTRACT. In this paper we will study statistical equilibria in commodity markets where agents have a specified utility attached to every transaction in their offer sets. A probability measure on the product of all offer sets is called benefit efficient if market transactions with higher total benefit are more probable. We will characterize all such probability measures and show how this defines a new family of statistical equilibria in commodity markets. If agents are indifferent with respect to utility, these equilibria reduce to the classical entropy maximizing states. Moreover, we show how to construct what we call the most likely explanation for a set of observed commodity prices.

Keywords: Commodity markets, statistical equilibria, efficient probability measures
Jel codes: D40, D50, G10

1. Introduction

Foley (1994) initiated a new theory of statistical equilibria in commodity markets. His basic idea was to show how the classical theory of statistical mechanics developed by Boltzmann, Maxwell and Gibbs could be incorporated into a setting where different types of agents carry out transactions in a commodity market. Assuming that there are many agents of each type and that agents of the same type are indistinguishable, any feasible transaction can be carried out in a large number of different ways. Some states will be more probable that others, however, and the point of view in statistical mechanics is that we can expect that the combined probability distribution will get a strong peak at the most probable state. Hence it is quite unlikely that we will observe states that are much different from this state.

While Foley´s approach is novel with respect to commodity markets, the idea of exploiting statistical mechanics in economics is by no means new. Horowitz and Horowitz (1968) used entropy and markov processes to study competition in the brewing industry. Theil (1969) discussed entropy in the formation of political parties. In the same time period Georgescu-Roegen presented economic models based on entropy and the second law of thermodynamics, see Beard and Lozada (1999).

In a more recent application Krebs (1997) shows how entropy theory can be used to construct statistical equilibria in one-step forward looking models, and generally there now seems to be a quite substancial number of authors with an interest to entropy constructions. In the recent literature one can also find numerous references to entropy used as a measure of statistical fitting. Such applications are, however, somewhat on the side of the major issues discussed in our paper and will not be referenced here.

Entropy maximizing has been studied intensively by spatial economists for a large period of time. Wilson (1967) was the first to show that transportation models could be derived from entropy maximizing principles, and his theory has ever since been an important topic for further studies. The number of applications are too numerous to be mentioned here, we refer instead to the classical textbook by Sen and Smith (1995) and the references therein.

Transportation models based on entropy maximizing principles are commonly referred to as gravity models. They have found widespread use, and today practically every road planning office are running program packages based on such applications. In spite of this success story, the critics of these models have not been silent. As Brøcker (1989) puts it: "Neither do people behave like Newtonian masses nor like molecules bumping into each other under certain macroeconomic conservation constraints". The basic point of view is that models of this kind must be derived from behavioral principles. As it turns out, however, gravity models are surprisingly robust, and it is well known that they can be derived in many different ways. Anas (1983) was the first to show that gravity models can be derived from random utility theory, and we refer to Erlander and Stewart (1990) for a number of different derivations of these models. Of particular interest is the derivation from cost efficiency principles, see Erlander and Smith (1990). In their paper Erlander and Smith show that gravity models can be derived from cost efficiency principles. Their basic hypothesis is that transportation states with smaller total costs should be more probable, and it turns out that there are very few probability measures with this property. More precisely: if transportation states with smaller total costs are more probable, the model must be of gravity type.

Foley (1994) assumes that all feasible transactions are equally probable. In the absence of information this is of course a reasonable hypothesis. In many cases, however, it seems likely to assume that some information is present and that agents value the different transactions in their offer sets somewhat differently. In this paper we will show how Foley´s framework can be extended to incorporate effects of this sort. Our framework is based on Foley´s original framework, and we wish to extend this framework to a setting where all agents of the same type have a specified utility attached to every transaction in their offer sets. In doing so we will rest heavily on the construction of cost efficient probability measures in Erlander and Smith (1990). In our case we will argue that transactions with a higher total benefit should be more probable, and then we can use the core argument in Erlander and Smith (1990) to construct a representation of all benefit efficient probability measures. In doing so we will exploit a simplified construction of these measures as it is presented in Jörnsten et al. (2004).

Our paper is organized as follows: In Section 2 we present the general framework and give a brief sketch of Foley´s original construction. As we wish to include additional effects, we have chosen to modify the notation to better suit our purposes. We hope that this is not too confusing for the reader. In Section 3 we introduce utilities and give the precise definition of benefit efficiency. The highlight is Theorem 3.2 which provides an explicit characterization of all benefit efficient probability measures. In Section 4 we consider an explicit example to discuss the various consequences of Theorem 3.2. When agents put no emphasis on the different utilities, we will see that the only benefit efficient states is the one provided by maximum entropy, i.e., the state given by Foley´s construction. As more emphasis is put on utility, agents (according to our model) are typically concentrated at the more beneficial states. In Section 5 we will return to the general discussion. We will show that efficient probability measures always exist, and that they in all but degenerate cases can be interpreted in terms of unique entropy prices. In Section 5 we also discuss inverse problems: Assume that we have observed prices in the market. Is it then possible to find a set of utilities that offers an explanation for these prices. We argue that such explanations can be found. Moreover, if we search for an explanation with the largest possible entropy, such an explanation can be expected to be unique. In Section 5 we also discuss an example where we observe negative entropy prices, but where the sign of these prices can be reversed if we can persuade a new type of agent to enter into the market. Such agents can be interpreted as arbitrageurs; they can only be persuaded to enter if they are paid sufficiently much for their entry. Finally in Section 5 we offer some concluding remarks.

As some basic proofs are quite theoretical, we have placed parts of this material in the appendix. It is our hope that this will facilitate reading of the paper.

2. The framework

In this paper we will assume that there are K types of commodities, T types of agents and that all agents of type t have the same offer set $O_t \subset \mathbb{R}^K$. The explicit meaning of offer sets is explained through the following example.

EXAMPLE 2.1

Assume that there are $K = 4$ commodities and $T = 2$ types of agents. All agents of type 1 have the offer set

$$O_1 = \{(4, -2, 3, -4), (8, -4, 6, -8), (0, 0, 0, 0)\} \tag{2.1}$$

and all agents of type 2 have the offer set

$$O_2 = \{(-1, 1, 0, 3), (-3, 1, -3, 1), (-8, 4, -6, 8), (0, 0, 0, 0)\} \tag{2.2}$$

The interpretation is as follows. If an agent of type 1 performs the transaction $(4, -2, 3, -4)$ it literally means that he uses 4\$ to buy commodity 1, sells commodity 2 to receive 2\$, and so on. We enumerate the various transactions according to their position above, and let F_{ij} denote the event that an agent of type i does transaction j. An example of a *feasible* transaction is then

- That one agent of type 1 does transaction F_{11}
- That one agent of type 2 does transaction F_{21} and that another agent of type 2 does F_{22}
- All other agents take no action, i.e., choose F_{13} or F_{24} according to type

This particular combination clears the market. This market can of course be cleared in a multitude of different ways. The particular combination F_{11}, F_{21}, F_{22} involves coordination between 3 agents, and it seems more easy to carry out transactions of the form F_{12}, F_{23}. If F_{12} improves the position of agents of type 1, it seems not unlikely that F_{22} might improve the position even more. Hence it is relatively easy to point out scenarios where the different feasible actions have different probabilities.

General framework

In general we will assume that there are N_t agents of type t, and that there are $N = \sum_{t=1}^{T} N_t$ agents altogether. We order the agents according to type, and define

$$I_1 = \{1, 2, \ldots, N_1\}, \qquad I_2 = \{N_1 + 1, N_1 + 2, \ldots, N_1 + N_2\}, \qquad \text{etc} \tag{2.3}$$

Hence agent i is of type t if and only if $i \in I_t$. Letting $x^i \in \mathbb{R}^K$ denote the transaction carried out by agent i, $1 \le i \le N$, we define a market transaction $\mathbf{x} \in \mathbb{R}^{KN}$ by

$$\mathbf{x} = (x^1, x^2, \ldots, x^N) \tag{2.4}$$

A market transaction $\mathbf{x}$ is feasible when it clears the market for all commodities, i.e.

$$\sum_{i=1}^{N} x_k^i = 0 \qquad \forall k = 1, \ldots, K \tag{2.5}$$

Agents of the same type are assumed to be indistinguishable, and a statistical equilibrium is obtained in a situation when there is a strong tendency that many agents of the same type use the same transaction. Hence given any feasible transaction $\mathbf{x}$ it is important to keep track of the number of agents of type t that carries out a particular transaction $x \in \mathbb{R}^K$. To this end we define for $x \in \mathbb{R}^K$

$$f_t^{\mathbf{x}}[x] = \sum_{i \in I_t} \mathcal{X}[x^i = x] \qquad \text{where} \qquad \mathcal{X}[x^i = x] = \begin{cases} 1 & \text{if } x^i = x \\ 0 & \text{otherwise} \end{cases} \tag{2.6}$$

We make the following observation which will be of some use in the sequel:

PROPOSITION 2.2

A market transaction $\mathbf{x}$ *is feasible if and only if*

$$\sum_{t=1}^{T} \sum_{x \in O_t} f_t^{\mathbf{x}}[x] x = 0 \tag{2.7}$$

PROOF

(2.8)

$$\sum_{t=1}^{T} \sum_{x \in O_t} f_t^{\mathbf{x}}[x] x = \sum_{t=1}^{T} \sum_{x \in O_t} \sum_{i \in I_t} \mathcal{X}[x^i = x] x = \sum_{t=1}^{T} \sum_{i \in I_t} \sum_{x \in O_t} \mathcal{X}[x^i = x] x = \sum_{t=1}^{T} \sum_{i \in I_t} x^i = \sum_{i=1}^{N} x^i$$

$\square$

DEFINITION 2.3

Since agents of the same type are indistinguishable, we say that two market transactions $\mathbf{x}$ and $\mathbf{y}$ are equivalent if

$$f_t^{\mathbf{x}}[x] = f_t^{\mathbf{y}}[x] \qquad \forall t \in T, x \in O_t \tag{2.9}$$

Given t and $\mathbf{x}$ it is important to see in how many ways the N_t agents of type t can be rearranged keeping

$$f_t^{\mathbf{x}}[x] = f_t^{\mathbf{y}}[x] \qquad \forall x \in O_t \tag{2.10}$$

By definition there are $f_t^{\mathbf{x}}[x]$ agents of type t that does transaction x. Throughout the paper we will assume that all offer sets are finite. Hence there is altogether

$$\frac{N_t!}{\prod_{x \in O_t} f_t^{\mathbf{x}}[x]!} \tag{2.11}$$

rearrangements of these agents giving an equivalent outcome. We can state the following proposition:

PROPOSITION 2.4

Given any market transaction $\mathbf{x}$*, there are altogether*

$$\prod_{t=1}^{T} \frac{N_t!}{\prod_{x \in O_t} f_t^{\mathbf{x}}[x]!} \tag{2.12}$$

market transactions within the equivalence class defined by $\mathbf{x}$*.*

We now assume that all $x \in O_t$ are equally probable. To find the most probable equivalence class, we must then search for a set of integers $\{f_t[x] | t = 1, \ldots, T, x \in O_t\}$ with the following properties

(2.13)

$$f_t[x] \geq 0 \qquad \sum_{x \in O_t} f_t[x] = N_t \qquad \sum_{t=1}^{T} \sum_{x \in O_t} f_t[x]x = 0 \qquad \forall t = 1, \ldots, T, x \in O_t$$

and such that

$$\prod_{t=1}^{T} \frac{N_t!}{\prod_{x \in O_t} f_t[x]!} \tag{2.14}$$

is as large as possible. If such a set can be found, it is then easy to verify that we can find a feasible market transaction $\mathbf{x}^*$ such that

$$f_t^{\mathbf{x}^*}[x] = f_t[x] \qquad \forall t = 1, \ldots, T, x \in O_t \tag{2.15}$$

The equivalence class of $\mathbf{x}^*$ is hence the most probable set of transactions. We observe that N_t is exogenous, and taking logarithms we can conveniently rephrase the problem as follows:

Optimization problem

Given $N_t \in \mathbb{N}, O_t \subset \mathbb{R}^K$ $t = 1, \dots, T$, find a set of integers $\{f_t[x] | t = 1, \dots, T, x \in O_t\}$ such that
(2.16)

$$f_t[x] \geq 0 \qquad \sum_{x \in O_t} f_t[x] = N_t \qquad \sum_{t=1}^{T} \sum_{x \in O_t} f_t[x]x = 0 \qquad \forall t = 1, \dots, T, x \in O_t$$

and such that

$$-\sum_{t=1}^{T} \sum_{x \in O_t} \ln[f_t[x]!] \tag{2.17}$$

is as large as possible.

The basic idea in Foley (1994) is now to assume that the number of agents performing every legitimate transaction is so large that we can use Stirling's approximation to the factorial function, and replace this integer optimization problem with a classical entropy maximizing problem. If so, he can appeal to the Kuhn-Tucker theorem and see that the maximum entropy problem has a unique solution which can be described as follows:

There exist a set of equilibrium prices $\pi_t^* \in \mathbb{R}^K, t = 1, \dots, T$ such that

$$f_t^*[x] = N_t \cdot \frac{\exp[-\pi_t^* \cdot x]}{\sum_{y \in O_t} \exp[-\pi_t^* \cdot y]} \qquad \forall x \in O_t \tag{2.18}$$

Hence Foley concludes that if all transactions are equally probable, then there exists a set of entropy prices defining a statistical equilibrium for the market. In the next section, however, we will consider a more general version of the problem.

3. Efficiency

We now assume that all agents of type t has a given utility $U_t[x]$ associated with each transaction $x \in O_t$, and we define a total benefit function $B : \mathbb{R}^{KN} \to \mathbb{R}$ by

$$B[\mathbf{x}] = \sum_{t=1}^{T} \sum_{x \in O_t} f_t^{\mathbf{x}}[x] \cdot U_t[x] \tag{3.1}$$

Let P_t denote an arbitrary probability measure on O_t. The efficiency principle can be stated as follows:

A probability measure $\mathbf{P} = \prod_{t=1}^{T} P_t$ is called *benefit efficient* if and only if

$$B[\mathbf{x}] \leq B[\mathbf{y}] \Rightarrow \prod_{t=1}^{T} \prod_{x \in O_t} P_t[x]^{f_t^{\mathbf{x}}[x]} \leq \prod_{t=1}^{T} \prod_{x \in O_t} P_t[x]^{f_t^{\mathbf{y}}[x]} \tag{3.2}$$

which says that transactions with higher total benefit should be more probable. We emphasize that in these formulas the superscripts $f_t^{\mathbf{x}}[x]$ and $f_t^{\mathbf{y}}[x]$ are exponents, i.e., that

$$P_t[x]^{f_t^{\mathbf{x}}[x]} = \underbrace{P_t[x] \cdot P_t[x] \cdots P_t[x]}_{f_t^{\mathbf{x}}[x] \text{ times}} \tag{3.3}$$

$$P_t[x]^{f_t^{\mathbf{y}}[x]} = \underbrace{P_t[x] \cdot P_t[x] \cdots P_t[x]}_{f_t^{\mathbf{y}}[x] \text{ times}} \tag{3.4}$$

Taking logarithms (3.2) can conveniently be rewritten as follows:

$$B[\mathbf{x}] \leq B[\mathbf{y}] \Rightarrow \sum_{t=1}^{T} \sum_{x \in O_t} \ln[P_t[x]] \cdot f_t^{\mathbf{x}}[x] \leq \sum_{t=1}^{T} \sum_{x \in O_t} \ln[P_t[x]] \cdot f_t^{\mathbf{y}}[x] \tag{3.5}$$

Clearly it is possible to order all $P_t[x]$ in a single sequence of length M which we identify with $\mathbf{P}$, and correspondingly we can order all $f_t^{\mathbf{x}}[x]$ and $f_t^{\mathbf{x}}[x]$ in two sequences $\mathbf{f}^{\mathbf{x}}$ and $\mathbf{f}^{\mathbf{y}}$ of the same length M. With this notation (3.5) can be written in the compressed form

$$B[\mathbf{x}] \leq B[\mathbf{y}] \Rightarrow \ln[\mathbf{P}] \cdot \mathbf{f}^{\mathbf{x}} \leq \ln[\mathbf{P}] \cdot \mathbf{f}^{\mathbf{y}} \tag{3.6}$$

Here we have used the convention that we apply a function to a vector applying the function to each component of the vector. Hence

$$\ln[\mathbf{P}] = (\ln[P_1], \ldots, \ln[P_M]) \tag{3.7}$$

This convention will be used in the sequel without further remarks.

The next issue is to incorporate the market restrictions into the framework. We note that

$$\sum_{x \in O_t} f_t[x] = N_t \qquad \sum_{t=1}^{T} \sum_{x \in O_t} f_t[x] x = 0 \qquad \forall t = 1, \ldots, T, x \in O_t \tag{3.8}$$

are all linear restrictions on $\mathbf{f}$. Hence it is possible to find a matrix A and a vector $\mathbf{b}$ such that (3.8) is equivalent to the statement

$$A\mathbf{f} = \mathbf{b} \tag{3.9}$$

Next we note that in the definition of the benefit function B, it is only the frequencies that are important. Hence if $\mathbf{x}$ and $\mathbf{y}$ are two market transactions with the same frequencies $\mathbf{f}$, then

$$B[\mathbf{x}] = B[\mathbf{y}] \tag{3.10}$$

and by a slight abuse of notation we define

$$B[\mathbf{f}] = B[\mathbf{x}] \qquad \text{where } \mathbf{x} \text{ is any element in the equivalence class defined by } \mathbf{f} \tag{3.11}$$

Activity equivalence

We say that two frequency vectors $\mathbf{f}$ and $\mathbf{g}$ are activity equivalent under A if $A\mathbf{f} = A\mathbf{g}$. Hence if there exist at least one feasible market transaction $\mathbf{x}$, then (3.8) is equivalent to the statement $A\mathbf{f} = A\mathbf{f}^{\mathbf{x}}$. This leads us to our central definition:

DEFINITION 3.1

A probability vector $\mathbf{P}$ is called benefit efficient under A if for all feasible frequency vectors $\mathbf{f}$ and $\mathbf{g}$ that are activity equivalent under A

$$B[\mathbf{f}] \leq B[\mathbf{g}] \Rightarrow \ln[\mathbf{P}] \cdot \mathbf{f} \leq \ln[\mathbf{P}] \cdot \mathbf{g} \tag{3.12}$$

We will now characterize all benefit efficient probability measures. The following theorem applies.

THEOREM 3.2

Assume that there exist at least one strictly positive feasible market transaction, and that $\mathbf{P}$ *is a probability measure that is benefit efficient under the activity matrix* A *in (3.9) with reference to the benefit function defined in (3.1). Then there exist constants* $u_1, \ldots u_{T+K} \in \mathbb{R}, u_{T+K+1} \in \mathbb{R}^+$ *such that*

$$\ln[\mathbf{P}] = (u_1, u_2, \ldots, u_{T+K})A + u_{T+K+1}\mathbf{U} \tag{3.13}$$

Conversely any probability measure defined by (3.13) is benefit efficient under A.

PROOF

See the appendix.

□

The result in Theorem 3.2 is written in a compact notation, and it is not straightforward to see what are the real implications of this result. To clarify these issues, we will in the next section discuss a few explicit examples to see what is the real contents of this result.

4. Interpretations and examples

We will now return to the setting in Example 2.1, i.e., we consider $K = 4$ commodities and $T = 2$ types of agents. The strong linear dependence between the elements in the offer sets, however, puts us in a position where unique entropy prices cannot be found. The problem is degenerate, and we will hence consider a slightly extended version to remove the degeneracy:

Throughout this section we will assume that there are $N_1 = 15\,000$ agents of type 1 and $N_2 = 8\,000$ agents of type 2. We let f_1, f_2, f_3, f_4, f_5 denote the number of agents of type 1 that carries out actions

$$F_{11} = (4,-2,3,-4), F_{12} = (8,-4,6,-8), F_{13} = (0,0,1,-1)$$
$$F_{14} = (-1,1,-1,1), F_{15} = (0,0,0,0)$$

respectively, and correspondingly we let $f_6, f_7, f_8, f_9, f_{10}, f_{11}$ be the number of agents of type 2 that carries out actions

$$F_{21} = (-1,1,0,3), F_{22} = (-3,1,-3,1), F_{23} = (-8,4,-6,8)$$
$$F_{24} = (0,0,-6,6), F_{25} = (1,-1,1,-1), F_{26} = (0,0,0,0)$$

The 6 feasibility restrictions in (2.13)/(3.8) can then be expressed as follows.

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 4 & 8 & 0 & -1 & 0 & -1 & -3 & -8 & 0 & 1 & 0 \\ -2 & -4 & 0 & 1 & 0 & 1 & 1 & 4 & 0 & -1 & 0 \\ 3 & 6 & 1 & -1 & 0 & 0 & -3 & -6 & -6 & 1 & 0 \\ -4 & -8 & -1 & 1 & 0 & 3 & 1 & 8 & 6 & -1 & 0 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \\ f_9 \\ f_{10} \\ f_{11} \end{bmatrix} = \begin{bmatrix} 15\,000 \\ 8\,000 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{4.1}$$

The optimal solution of (2.17) can easily be computed, giving maximum entropy at

$$\mathbf{f}^* = (2\,628, 1\,568, 4\,484, 1\,917, 4\,403, 1\,080, 1\,080, 2\,342, 747, 1\,917, 834) \tag{4.2}$$

The corresponding entropy prices are as follows

$$\pi^* = (0.92, 1.73, -0.38, -0.36) \tag{4.3}$$

We observe that the entropy price of commodity 3 and 4 are both negative. The reason is of course that we are dealing with a model with non-free disposal. One way to handle this is to consider an alternative formulation where we admit partial clearing of the market. If we replace (2.16) with

(4.4)
$$f_t[x] \geq 0 \qquad \sum_{x \in O_t} f_t[x] = N_t \qquad \sum_{t=1}^{T} \sum_{x \in O_t} f_t[x]x \leq 0 \qquad \forall t = 1, \ldots, T, x \in O_t$$

we will get non-negative entropy prices. This, however, will not provide a completely satisfactory solution to our problem as all commodities that does not strictly clear the market must have an entropy price equal to zero due to complementary slackness.

What happens if we instead consider benefit efficient probabilities? To illustrate this we will assume that the agents has an exogenously given utility vector

$$\mathbf{U} = (1, 2, 1, 1, 0, 1, 1, 2, 1, 1, 0) \tag{4.5}$$

We let A be coefficient matrix on the left hand side of (4.1). Theorem (3.2) states that all benefit efficient probability vectors can be expressed on the form

$$\ln[\mathbf{P}] = (u_1, u_2, u_3, u_4, u_5, u_6)A + u_7\mathbf{U} \tag{4.6}$$

Seemingly this seems to leave the system with 7 degrees of freedom, and hence an abundant set of such vectors. If we assume that the numbers of agents is very large, however, the set of feasible frequency vectors that can arise from such probabilities is highly restricted. In fact 6 degrees of freedom must be used to fulfill the feasibility restrictions, which leaves the system with only one additional degree of freedom.

If we assume that $u_7 \geq 0$ is given, we hence get a parameterized set of frequencies defined in terms of the value of u_7. This parameter has a very striking interpretation; it defines how much impact the utility vector U has on the equilibrium distribution.

To carry out the calculations, we must try to find a solution of the non-linear system

$$A \exp[(u_1, u_2, u_3, u_4, u_5, u_6)A + u_7\mathbf{U}] = (\frac{15\,000}{23\,000}, \frac{8\,000}{23\,000}, 0, 0, 0, 0) \tag{4.7}$$

This makes up a special system of transcendent equations, which under normal circumstances would be very difficult to solve. A closer inspection of this system, however, reveals that is has some very attractive mathematical properties. To explain this we take a closer look on the third equation which can be rearranged as follows:

$$\begin{aligned} &4e^{u_1+4u_3-2u_4+3u_5-4u_6+u_7} + 8e^{u_1+8u_3-4u_4+6u_5-8u_6+2u_7} \\ &= e^{u_2-u_3+u_4+3u_6+u_7} + 3e^{u_2-3u_3+u_4-3u_5+u_6+u_7} + 8e^{u_2-8u_3+4u_4-6u_5+8u_6+2u_7} \end{aligned} \tag{4.8}$$

The important observation is the following: The left hand side is an increasing function of u_3 and the right hand side is a decreasing function of u_3. Hence given any values for $u_1, u_2, u_4, u_5, u_6, u_7$, this equation has a unique solution. It turns out that this principle is true in general: Given all parameters except u_i, then equation i in (4.5) has a unique solution, see the Appendix for a complete proof of the general case. The system can then be solved by the following algorithm:

Given u_7, we put

i) $(u_1^0, u_2^0, u_3^0, u_4^0, u_5^0, u_6^0) = (0, 0, 0, 0, 0, 0)$
ii) Given $u_2^0, u_3^0, u_4^0, u_5^0, u_6^0$, find u_1^1 such that equation 1 is satisfied.
iii) Given $u_1^1, u_3^0, u_4^0, u_5^0, u_6^0$, find u_2^1 such that equation 2 is satisfied.

$\vdots$

vii) Given $u_1^1, u_2^1, u_3^1, u_4^1, u_5^1$, find u_6^1 such that equation 6 is satisfied.
viii) Repeat steps ii)-vii) until the system comes to rest at an equilibrium.

This provides an exceptionally fast numerical algorithm for this problem, and solutions can be found in only a few seconds. The basic idea is very similar to idea behind the well known Bregman balancing algorithm that is commonly used in spatial economics, see Bregman (1967).

Case 1: $u_7 = 0$. In this particular case the agents do not care about the utility vector $\mathbf{U}$ which may in fact be arbitrary. Using the algorithm above, we find

$$\mathbf{f} = (2\,628, 1\,568, 4\,484, 1\,917, 4\,403, 1\,080, 1\,080, 2\,342, 747, 1\,917, 834) \tag{4.9}$$

If we compare this solution with the max entropy solution given in (4.2), we see that these solutions coincide. Hence if the agents are insensitive with respect to the benefits from trading, there is a unique frequency vector that is benefit efficient, and this vector coincides with the max entropy solution. See Theorem 5.1 for a discussion of the general case.

Case 2: $u_7 = 1$. In this case the utility vector $\mathbf{U} = (1, 2, 1, 1, 0, 1, 1, 2, 1, 1, 0)$ is crucial to get a well defined result. Using the algorithm above, we find

$$\mathbf{f} = (2\,750, 3\,415, 5\,174, 1\,447, 2\,214, 515, 515, 4\,532, 862, 1\,447, 128) \tag{4.10}$$

We observe that the no-transaction states 5 and 11 are now considerably less attractive. Moreover agents of type 1 are now much more likely to prefer transaction 2 which is the one that gives the most profit, and correspondingly agents of type 2 are more likely to prefer transaction 8.

Case 3: $u_7 = 6$. Due to the exponential nature of the problem, a value $u_7 = 6$ is very large, implying, e.g., a ratio $e^{12} \approx 150\,000$ between transactions 2 and 5. Using the algorithm above, we find

$$\mathbf{f} = (984, 6\,222, 7\,622, 17, 155, 0, 0, 6\,714, 1\,270, 15, 0) \tag{4.11}$$

We see that this case is quite extreme. No agents of type 2 carries out transactions F_{21}, F_{22} and F_{26} which are less profitable for these agents. It may seem surprising that transaction $F_{24} = (0, 0, -6, 6)$ is not void. The reason for this, however, is obvious: one transaction of this type allows for no less than 6 transactions of type F_{13}. Hence we would see a considerable loss in entropy if F_{24} is void.

5. General issues and entropy prices

We have seen that when agents put more and more emphasis on utility we get a gradual change in distribution. The next issue, however, is how this affects entropy prices. It is not clear that the setting in Theorem 3.2 can be interpreted in the sense of entropy, but a convenient reformulation of the problem clarifies the picture. Consider an entropy problem of the form:

Optimization problem

Given $\beta \geq 0$ find a feasible transaction $\mathbf{f}$ such that $B[\mathbf{f}] \geq \beta$ and such that the entropy

$$-\sum_{i=1}^{M} f_i \ln[f_i] \tag{5.1}$$

is as large as possible.

It is then easy to see that the characterization in Theorem 3.2 is nothing but the Kuhn-Tucker conditions for this non-linear problem. Since $\mathbf{f} \mapsto \mathbf{f}\ln[\mathbf{f}]$ is convex and all restrictions are linear, we have a unique optimal value, and in all but degenerate cases there exists unique shadow prices $(u_1, \ldots, u_{T+K+1})$. Hence u_{T+K+1} can be interpreted as the entropy cost associated with the gain of one unit of utility, and the vector

$$\pi = (-u_{T+1}, -u_{T+2}, \ldots, -u_{T+K}) \tag{5.2}$$

defines entropy prices for the K commodities. If $\beta = 0$, this problem reduces to the classical entropy problem studied by Foley. This explains why the solutions in (4.2) and (4.9) coincide. These observations can be summarized as follows:

THEOREM 5.1

Let $\mathbf{U} \in \mathbb{R}^M$ be any exogenously given utility vector and let

$$\mathbf{P} = \exp[(u_1, \ldots, u_{T+K})A + u_{T+K+1}\mathbf{U}] \tag{5.3}$$

For any $u_{T+K+1} \in \mathbb{R}_+$ there exist $u_1, \ldots, u_{T+K}$ such that the vector

$$\mathbf{f} = N \cdot \mathbf{P} \tag{5.4}$$

is feasible. Moreover

• $\pi = (-u_{T+1}, -u_{T+2}, \dots, -u_{T+K})$ *defines the entropy prices for the* K *commodities.*
• *If the gradients of the binding restrictions to (5.1) are linearly independent, then the* $u_1, \dots, u_{T+K}$ *are unique.*
• *If* $u_{T+K+1} = 0$*, then* $\mathbf{f}$ *given by (5.4) coincides with the classical entropic maximum.*

EXAMPLE 5.2

We now return to the example studied in Section 4, to see what entropy prices we get in that case. As in case 2, we assume that $u_7 = 1$ and that

(5.5) $$\mathbf{U} = (1, 2, 1, 1, 0, 1, 1, 2, 1, 1, 0)$$

It turns out that the entropy prices defined in Theorem 5.1 are as follows

(5.6) $$\pi = (1.41, 2.99, -0.5, -0.66)$$

These entropy prices are, however, still negative. Hence we need to rephrase our question: Is it possible to find a utility vector $\mathbf{U}$ yielding strictly positive entropy prices in this case? As it turns out we can do considerably better than that. We consider an inverse problem: Assume that we have observed, e.g., a price vector

(5.7) $$\pi^{\text{observed}} = (1.0,\ 2.0,\ 0.5,\ 0.5)$$

in the market. Is it possible to find a utility vector $\mathbf{U}$ replicating these prices? The answer is yes, and one possible such $\mathbf{U}$ (using $u_7 = 1$) is

(5.8) $$\mathbf{U} = (0.36, 0.35, 3.55, 0.35, 2.08, 7.00, 2.50, 0.00, 3.96, 0.55, 0.01)$$

This offers one possible explanation why a market of this sort may give rise to positive entropy prices. We can see, e.g., that when agents have a strong preference for transaction nr 6, i.e., $F_{21} = (1, -1, 0, 3)$, they highly value the trade in commodity 4, reversing the sign of the entropy price in (4.3). The solution is of course not unique; we have essentially 11 degrees of freedom in our choice of $\mathbf{U}$ and only 7 of these are used in the replication. Hence there will always be a large number of different explanations for the observed prices in (5.7). Explanations with high entropies are, however, more likely that the others. Hence we should search for an explanation with maximum entropy, and such an explanation we can expect to be unique (when units are chosen). In general we formulate the framework as follows:

DEFINITION 5.3

Assume that we have observed commodity prices

(5.9) $$\pi^{\text{observed}} = (\pi_1, \dots, \pi_K)$$

in the market. The *most likely explanation* for these prices is a utility vector $\mathbf{U}^*$ such that

(5.10) $$\mathbf{f} = N \cdot \exp[(u_1, \dots, u_{T+K})A + u_{T+K+1}\mathbf{U}^*]$$

• is feasible
• $\pi^{\text{observed}} = (-u_{T+1}, -u_{T+2}, \dots, -u_{T+K})$
• the entropy of $\mathbf{f}$ is as large as possible

Remark

If $u_{T+K+1} \neq 1$, we can define a new utility vector $\tilde{\mathbf{U}} = u_{T+K+1}\mathbf{U}^*$. Clearly $\tilde{\mathbf{U}}$ will be optimal if $u_{T+K+1} = 1$. Hence we can assume that $u_{T+K+1} = 1$ without loss of generality. This only corresponds to a choice of utility units.

The solution in (5.8) was found using the Nelder-Mead simplex algorithm with the origin and the unit vectors as initial points. If we perform a more refined search taking maximum entropy into account, we get

$$\mathbf{U}^* = (0.63, 1.33, 0.71, 0.37, 2.93, 5.52, 1.02, 5.04, 0.66, 0.11, 0.75) \tag{5.11}$$

increasing the entropy from $-192\,694$ in (5.8) to $-189\,248$ in (5.11). If we compare this vector with the utility vector from (5.8), i.e.

$$\mathbf{U} = (0.36, 0.35, 3.55, 0.35, 2.08, 7.00, 2.50, 0.00, 3.96, 0.55, 0.01) \tag{5.12}$$

we can notice some characteristic features. Both vectors puts strong emphasis on component nr 6, with the implications we discussed above. The optimal vector in (5.11), however, is much more evenly distributed. This is what we expect. If all the components of $\mathbf{U}$ are equal, we get the solution in (4.2), i.e., the global entropic maximum.

Arbitrageurs

Let us return to the case where $\mathbf{U} = (1, 2, 1, 1, 0, 1, 1, 2, 1, 1, 0)$ and $u_7 = 1$. As we have seen, this leads to some negative entropy prices. Assuming that $\mathbf{U}$ is fixed, it turns out that we still can obtain positive entropy prices if new agents are introduced in the system. One example of this sort can be described as follows: Let us introduce $N_3 = 3\,000$ agents of type 3 in the system. Agents of type 3 can carry out the transactions

$$F_{31} = (0, 0, 0, 1), \; F_{32} = (0, 0, -1, 0), \; F_{33} = (0, 0, 0, 0) \tag{5.13}$$

For simplicity we assume that F_{32} and F_{33} has utility 0. The basic question is now; What is the minimum utility on $F_{31} = (0, 0, 0, 1)$ giving non-negative entropy prices? It turns out that if the utility of doing transaction F_{31} is $U_{31} = 3.12$, then we obtain entropy prices

$$\pi = (0.89, 1.46, 0, 0.02) \tag{5.14}$$

and strictly positive entropy prices are obtained if we increase this utility beyond that point.

We can think of agents of type 3 as arbitrageurs in the system. They are initially unwilling to participate in the market. If they are paid sufficiently, however, they can be persuaded to enter. As we have just seen, such an entry may give rise to strictly positive entropy prices in a situation where no such prices can be found initially. Hence it might be beneficial for all parties to pay agents of type 3 a certain fee for their entry.

6. Concluding remarks

In this paper we have showed how the classical entropy framework can be extended to cases where agents can attach different utilities to the transactions in their offer sets. If all such utilities are equal, our solutions coincide with those suggested by the classical theory. While the final results coincide in this case, our framework offers an alternative derivation which we believe is appealing from an economic point of view. We have showed that the classical equilibrium is the only state that is consistent with an efficient probability measure, i.e., a measure where states with a larger total utility are more probable.

Foley´s original approach rested on the assumption that all transactions are equally likely. In our extended approach no such conditions are needed. Quite the contrary, we have studied situations where we assumed that market prices were different from the ones obtained from the classical entropy models. In that setting we showed how to construct a set of utilities that offers an explanation to the deviation. Moreover, we suggested that one should search for a set of utilities with the largest possible entropy, and this set we would like to interpret as the most likely explanation for the observed prices.

The usefulness of our approach must eventually be judged from its empirical explanatory power. This raises a number of topics for future research. In particular we believe it would be interesting to apply the suggested framework to currency markets. Here bundles of currencies are traded in a manner that may be quite suitable for entropy modeling. In the last few years there has also been increasing interest for entropy models in marketing, see, e.g., Phillips (1994), and we suggest that the ideas in our paper can be useful in that setting as well.

A slight limitation of our framework is that it can only be applied to situations where there is a reasonably large number of agents and where the number of different transactions is not too large. While the number of agents should be large, there is no reason to assume astronomical numbers like the ones commonly used in particle physics. Within the field of regional science similar models have been applied to situations with moderate numbers of agents with great success, and we expect that our approach can be used in such cases as well. Moreover, the advances of modern computer science now admits applications with fairly large offer sets. Balancing algorithms of Bregman type have long since been used in situations more than 25 000 different equations/inequalities, see, e.g., Herman et al (1978). Hence we suggest that offer sets with a similar or even higher order of magnitude may be within reach.

7. Appendix

Proof of Theorem 3.2

Assume that $\mathbf{P}$ is benefit efficient under A. Choose any feasible frequency vector $\mathbf{g} > 0$, and consider the LP-problem

$$
\begin{aligned}
&\max_{\mathbf{f}} \ln[\mathbf{P}] \cdot \mathbf{f} \\
&A\mathbf{f} = A\mathbf{g}, \qquad B[\mathbf{f}] \le B[\mathbf{g}], \qquad \mathbf{f} \ge 0
\end{aligned}
\tag{7.1}
$$

This particular LP-problem must have the solution $\mathbf{f}^* = \mathbf{g}$, for if not the pair $\mathbf{f}^*, \mathbf{g}$ would violate (3.12). For simplicity of notation we let $m = T + K$ and $n = M$. If we define an extended matrix

$$W = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} & U_1 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} & U_n \end{bmatrix} \tag{7.2}$$

the dual problem of (7.1) can be stated as follows:

$$\begin{aligned} &\min_{\mathbf{u}=(u_1,\ldots,u_{m+1})} \mathbf{g}W\mathbf{u} \\ &W\mathbf{u} \leq \ln[\mathbf{P}] \qquad u_1, \ldots u_m \in \mathbb{R}, u_{m+1} \in \mathbb{R}^+ \end{aligned} \tag{7.3}$$

Since $\mathbf{f}^* > 0$ in (7.1), all slack variables in the dual problem must be zero, i.e.,

$$\ln[\mathbf{P}] = (u_1, u_2, \ldots, u_m)A + u_{m+1}\mathbf{U} \tag{7.4}$$

Conversely if $\mathbf{P}$ is defined by (3.13), let $\mathbf{f}$ and $\mathbf{g}$ be any two feasible frequency vectors that are activity equivalent under A. Then

$$\begin{aligned} \ln[\mathbf{P}] \cdot \mathbf{g} - \ln[\mathbf{P}] \cdot \mathbf{f} &= ((u_1, u_2, \ldots, u_m)A + u_{m+1}\mathbf{U})(\mathbf{g} - \mathbf{f}) \\ &= (u_1, u_2, \ldots, u_m)(A\mathbf{g} - A\mathbf{f}) + u_{m+1}(B[\mathbf{g}] - B[\mathbf{f}]) \end{aligned} \tag{7.5}$$

Hence if $B[\mathbf{f}] \leq B[\mathbf{g}]$, then $\ln[\mathbf{P}] \cdot \mathbf{g} - \ln[\mathbf{P}] \cdot \mathbf{f} \geq 0$, i.e., $\ln[\mathbf{P}] \cdot \mathbf{f} \leq \ln[\mathbf{P}] \cdot \mathbf{g}$.

□

Assume in general that there are K different commodities and T types of agents, with $N_1, \ldots, N_T > 0$ agents of each type. Let M be the total number of available transactions, and let $\mathbf{U} = (U_1, U_2, \ldots, U_M)$ be given utilities from each different transaction. Consider the i-th component of the left hand side of the equation

$$A \exp[(u_1, \ldots, u_T, u_{T+1}, \ldots, u_{T+K})A + u_{T+K+1}\mathbf{U}] = (N_1, \ldots, N_T, \underbrace{0, \ldots, 0}_{K \text{ times}}) \tag{7.6}$$

i.e.

$$\begin{aligned} &(A \exp[(u_1, \ldots, u_T, u_{T+1}, u_{T+K})A + u_{T+K+1}\mathbf{U}])_i \\ &= \sum_{j=1}^{M} A_{ij} \exp\left[u_i A_{ij} + \sum_{\substack{k=1 \\ k \neq i}}^{T+K} u_k A_{kj} + u_{T+K+1} \cdot U_j \right] \end{aligned} \tag{7.7}$$

Note that equation (7.6) does not define a probability measure $\mathbf{P}$. It is, however, straightforward to find a probability measure that is consistent with the frequencies in (7.1). The following scaling principle explains this.

PROPOSITION 7.1

Given $u_1, \ldots, u_T, u_{T+1}, \ldots, u_{T+K}$ define

$$v_i = \begin{cases} u_i + \ln[C] & \text{if } 1 \le i \le T \\ u_i & \text{if } T < i \le T+K \end{cases} \tag{7.8}$$

Then

$$\exp[(v_1, \ldots, v_{T+K})A + v_{T+K+1}\mathbf{U}] = C \cdot \exp[(u_1, \ldots, u_{T+K})A + u_{T+K+1}\mathbf{U}] \tag{7.9}$$

PROOF

Observe that if $1 \le i, j \le T$, there is exactly one $A_{ij} = 1$ and all other such A_{ij} are zero. The effect of adding $\ln[C]$ to all those u_i is hence that one and only one term in (7.7) is affected, i.e.

(7.10)
$$\exp\left[v_i A_{ij} + \sum_{\substack{k=1 \\ k \ne i}}^{T+K} v_k A_{kj} + v_{T+K+1} \cdot U_j\right] = \exp\left[\ln[C] + u_i A_{ij} + \sum_{\substack{k=1 \\ k \ne i}}^{T+K} u_k A_{kj} + u_{T+K+1} \cdot U_j\right]$$

□

COROLLARY 7.2

If we can find $u_1, \ldots, u_{T+K}$ such that (7.6) is satisfied, we can find a consistent probability measure $\mathbf{P}$ using $C = 1/\sum_{i=1}^{T} N_i$ in Proposition 7.1.

PROPOSITION 7.3

For all $i = 1, \ldots, K+T$, then given any values of $u_1, \ldots, u_{i-1}, u_{i+1}, u_{T+K+1}$ the i-th component of equation (7.6) has a unique solution.

PROOF

First assume that $1 \le i \le T$. Then all A_{ij} are either 0 or 1. In this case the function in (7.7) is strictly increasing from 0 to ∞, and hence given $u_1, \ldots, u_{i-1}, u_{i+1}, u_{T+K+1}$ we can find a unique u_i such that

$$(A \exp[(u_1, \ldots, u_{T+K})A + u_{T+K+1}\mathbf{U}])_i = N_i \tag{7.11}$$

Next assume that $T < i \le T+K$. Then we can split the sum in (7.7) into two different parts according to the sign of A_{ij}, i.e.

$$\begin{aligned} &(A \exp[(u_1, \ldots, u_{T+K})A + u_{T+K+1}\mathbf{U}])_i \\ &= \sum_{\substack{j=1 \\ A_{ij} \ge 0}}^{M} A_{ij} \exp\left[u_i A_{ij} + \sum_{\substack{k=1 \\ k \ne i}}^{T+K} u_k A_{kj} + u_{T+K+1} \cdot U_j\right] \\ &- \sum_{\substack{j=1 \\ A_{ij} < 0}}^{M} |A_{ij}| \exp\left[-u_i |A_{ij}| + \sum_{\substack{k=1 \\ k \ne i}}^{T+K} u_k A_{kj} + u_{T+K+1} \cdot U_j\right] \end{aligned} \tag{7.12}$$

We have assumed in general that there exist at least one feasible transaction. Hence there must be at least one $A_{ij} > 0$ and at least one $A_{ij} < 0$ (if not commodity i does not admit any transactions and can be ignored from the outset). Letting u_i pass from $-\infty$ to $+\infty$ the first sum in (7.12) will hence be strictly increasing from 0 to $+\infty$, while the second sum in (7.12) will be strictly increasing from $-\infty$ to 0. The whole expression will be strictly increasing from $-\infty$ to $+\infty$, and given $u_1, \ldots, u_{i-1}, u_{i+1}, u_{T+K+1}$ we can find a unique u_i such that

$$(A \exp[(u_1, \ldots, u_{T+K})A + u_{T+K+1}\mathbf{U}])_i = 0 \tag{7.13}$$

□

The next corollary is an immediate consequence of Proposition 7.3.

COROLLARY 7.4

If the algorithm: Given u_{T+K+1}, put

i) Put $u_i^0 = 0$ for all $i = 1, \ldots, T + K$

ii) Given $u_2^0, \ldots, u_{T+K}^0$, find u_1^1 such that first component of (7.6) is satisfied.

iii) Given u_1^1, u_3^0, u_{T+K}^0, find u_2^1 such that the second component of (7.6) is satisfied.

$\vdots$

vii) Given $u_1^1 \ldots, u_{T+K-1}^1$, find u_{T+K}^1 such that the last component of (7.6) is satisfied.

viii) Repeat steps ii)-vii) until the system comes to rest at an equilibrium.

converges, it gives a solution to (7.6)

One relatively important question is left without a precise answer. We are presently unable to prove that the algorithm above always converges. From Theorem 5.1 and Proposition 7.1 we know that there exists a solution to (7.6), and that this solution is usually unique. Still we cannot completely exclude the possibility that the algorithm above diverges. This problem is common to several widely used algorithms, so it should not exclude this algorithm from practical use.

REFERENCES

Anas, A. 1983. "Discrete choice theory, information theory and the multinomial logit and gravity models", *Transportation Research B*, 17, 13–23.

Beard T. R., and G. A. Lozada (1999). *Economics, Entropy and the Environment: The Extraordinary Economics of Georgescu-Roegen*, Edward Elgar, Cheltenham.

Bregman, L. M. 1967. "The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming", *USSR Computational Mathematics and Mathematical Physics*, vol. 7, 200–217.

Brøcker, J. 1989. "Partial equilibrium theory of international trade and the gravity model", *Papers of the Regional Science Association*, vol. 66, 7–18.

Erlander, S., and T. E. Smith. 1990. "General representation theorems for efficient population behavior", *Applied Mathematics and Computation*, 36, 173–217.

Erlander, S., and N. F. Stewart. 1990. *The gravity model in transportation analysis - theory and extensions*, VSP, Utrecht.

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

Herman, G. T., A. Lenth, and P.H. Lutz (1978). "Relaxation methods for image reconstruction", *Communications of the ACM*, vol. 21, no. 2, 152–158.

Horowitz A., and I. Horowitz (1968). "Entropy, Markov Processes and Competition in the Brewing Industry", *The Journal of Industrial Economics*, vol. 16, no. 3, 196–211.

Jörnsten, K., I. Thorsen and J. Ubøe. 2004. "Replication/prediction problems in the journey to work", *Environment and Planning A*, vol. 36, 347–364.

Krebs T. (1997). "Statistical equilibrium in one-step forward looking economic models", *Journal of Economic Theory*, vol. 73, no. 2, 365–394.

Phillips F. Y. 1994. "Brand Shifting in Hierarchical Markets: A New Model and Information Equivalency", *The Journal of the Operational Research society*, vol. 45, no. 8, 901–915.

Sen, A. and T. E. Smith. 1995. *Gravity Models of Spatial Interaction Behavior*, Springer-Verlag, Berlin.

Theil H. 1969. "The Desired Political Entropy", *The American Political Science Review*, vol. 63, no.2, 521-525.

Wilson, A. G. 1967. "A statistical theory of spatial distribution models", *Transportation Research*, vol. 1, 253–269.