Parameter Evaluation for a Statistical Mechanical Model for Binary Choice with Social Interaction

Alex Akwasi Opoku, Godwin Osabutey, and Charles Kwofie

Mathematics and Statistics Department, University of Energy and Natural Resources, P. O. Box 214, Sunyani, Ghana

Correspondence should be addressed to Alex Akwasi Opoku; alex.opoku@uenr.edu.gh

Received 1 October 2018; Accepted 17 February 2019; Published 4 March 2019

Copyright © 2019 Alex Akwasi Opoku et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper we use a statistical mechanical model as a paradigm for educational choices when the reference population is partitioned according to the socioeconomic attributes of gender and residence. We study how educational attainment is influenced by socioeconomic attributes of gender and residence for five selected developing countries. The model has a social and a private incentive part with coefficients measuring the influence individuals have on each other and the external influence on individuals, respectively. The methods of partial least squares and the ordinary least squares are, respectively, used to estimate the parameters of the interacting and the noninteracting models. This work differs from the previous work that motivated this work in the following sense: (a) the reference population is divided into subgroups with unequal subgroup sizes, (b) the proportion of individuals in each of the subgroups may depend on the population size \( N \), and (c) the method of partial least squares is used for estimating the parameters of the model with social interaction as opposed to the least squares method used in the earlier work.

1. Introduction

Education provides people with the knowledge and skills that can lead to better employment opportunities and a better quality of life. The educational level attained by an individual explicitly determines the occupational choice of that individual. All these attainments and choices of individuals are made under certain socioeconomic conditions such as peers, neighbours, family members, wealth quintile of the individual, gender, residence, etc. Those who reside in the rural areas with the poorest wealth quintile are more likely to have low education [1]. Here our focus is on how the collective behaviour of a reference group of individuals is determined by the intricate interactions among the individuals [2–5]. Collective behaviour such as self-organization has been observed in biological, ecological, and socioeconomical systems [6–8].

The collective behaviour of a large group of individuals may undergo sudden changes due to slight variations in the socioeconomic structure of the group. For instance a change in the pronunciation of a language due to a small immigrant population and a substantial decrease in crime rate is a result of actions taking by the authorities [9, 10]. This abrupt change in macroscopic behaviour caused by the changes in interactions among constituents is referred to as phase transition from the statistical mechanics literature. Phase transition has been shown to exist for some classes of spin models designed to explain the phenomenon of ferromagnetism [11, 12]. The simplest spin model within this class is the mean-field Ising model proposed in [11]. This model is tractable and has seen several applications in social sciences [13], finance [14], chemistry [15], and ecology [16]. An interesting family, which has naturally emerged in applications, is a multispecies version of the mean-field Ising model for studying magnetism in anisotropic materials [17]. This model has seen social science applications in recent works of Contucci, Gallo, and Barra [18–20].

The above and the works in [2–5] highlight the need for importing statistical mechanical models into the social science to offer insights into how social interactions determine social outcomes. Multipopulation Curie-Weiss model serves as a paradigm for certain binary discrete choice where individuals are to choose between two options, say to stay in school or drop out of school, use medicated mosquito net
The authors of [18–20] considered a multigroup Curie-Weiss model where the fraction of individuals in each subgroup of the population is a constant that is independent of the total population size. In particular, the work in [19] provides an estimation procedure to estimate the parameters in a multigroup Curie-Weiss model for suicidal tendencies and mode of marriage in Italy. The authors used least squares estimation procedure to estimate the parameters of their model. The study considered residence as the only socioeconomic attribute. In this work we adapt partial least squares estimation procedure to estimate the parameters for a multigroup Curie-Weiss model for educational attainment for five developing countries. Here we use the socioeconomic attributes of residence and gender. That is, we study how the collective choice of educational attainment of a group of individuals is influenced by their gender and place of residence. Our focus is on educational attainment of individuals in developing countries due to our interest in comparing the estimates from the different countries. Because of this we wanted countries that share similar socioeconomic attribute from different parts of the globe. We chose Dominican Republic from North America, Kenya from East Africa, Egypt from North Africa, Ghana from West Africa, and Indonesia from Asia. The choice of these countries is based on the availability of data from the demographic and health survey program.

The rest of the paper is organized as follows: Section 2 addresses generalities on Curie-Weiss model, its multipopulation version, and the parameter estimation procedure for the model. A case study on how educational choices is influenced by gender and residence is presented in Section 3. Section 4 discusses the main findings of the case study and we conclude the work in Section 5.

2. The Curie-Weiss Model

The Curie-Weiss model is made up of an energy function (Hamiltonian) that assigns interaction energies to spin configurations. This energy function takes the form

$$H_N(\sigma) = \frac{1}{2N} \sum_{i,j=1}^{N} J_{ij} \sigma_i \sigma_j + \sum_{i=1}^{N} h_i \sigma_i,$$  

(1)

where $\sigma_i = \pm 1$. The energy function consists of two parts, namely, the interaction part modulated by the interaction strengths $J_{ij}$ and the external field part also controlled by $h_i$. The interaction between neighbouring spins tends to induce alignment of the neighbours if $J_{ij} > 0$; i.e., both neighbours will prefer to be either +1 or both −1 if $J_{ij} > 0$. Otherwise the neighbours will prefer to assume different spin values. The parameter $h_i$ is the external magnetic field applied to site $i$. If the magnetic field $h_i$ is positive it favours +1 spin values. On the other hand if the field is negative it favours −1 spin values. Hence, for each site $i$, the external field contributes to the energy function by a term of $h_i \sigma_i$ [25]. This model is of a mean-field type as each spin interacts with the rest of the spins. The effect of all the other spins on any given spin is approximated by the average effect of the rest, and this makes computations easy. In this paper we will use the Curie-Weiss model as a benchmark model for discrete choice with social interaction. More precisely, we are interested in how educational attainment choices of individuals from some selected countries are influenced by the interaction among the individuals and the socioeconomic attributes of gender and residence. For related applications of statistical mechanical models to social science we refer the reader to [2–5, 26].

2.1. Multipopulation Curie-Weiss Model. This work uses partial least squares estimation procedure developed in [27] for the multipopulation Curie-Weiss model to estimate parameters in an interaction based logit model for educational attainment in some selected developing countries. Here our key assumption is that individuals with the same socioeconomic attributes tend to behave the same way but people with different socioeconomic backgrounds behave differently. This assumption helps us to reparametrise the parameters in the Hamiltonian (1). Thus we will focus on finding a suitable parametrisation for the interaction coefficient $J_{ij}$ and a systematic procedure that allows us to estimate the parameters characterizing the model from data. From our discrete choice model each individual $i$ is assigned $k$ socioeconomic attributes,

$$a_i = (a_i^{(1)}, a_i^{(2)}, \ldots, a_i^{(k)}), \quad \text{with } a_i^{(j)} \in \{0, 1\}.  \tag{2}$$

Therefore a population of size $N$ can be partitioned into $2^k$ groups that do not overlap. Each of the groups is identified by one of the elements of $\{0, 1\}^k$. Let $I_{Ng}$ be the set of individuals in partition $g$ for $g = 1, \ldots, 2^k$ and $|I_{Ng}| = N_g$. Therefore $N = N_1 + \cdots + N_{2^k}$. Let $I_N = \bigcup_{g=1}^{2^k} I_{Ng}$ with $I_{Ng} \cap I_{Ng'} = \emptyset$, for $g \neq g'$. We suppose further that, for each $g = 1, \ldots, 2^k$,

$$y_N^g = \frac{N_g}{N} \tag{3}$$

$$\text{and } y^g = \lim_{N \to \infty} y_N^g..$$

Our assumption above implies that individuals in the same group or partition are characterised by the same socioeconomic attribute; i.e., individuals with the same socioeconomic attributes are characterised as a group or partition. Therefore, it follows from our assumption above that all individuals in a partition or a group $g$ have the same private incentive $\bar{h}_g$ and for any pair of groups $g$ and $g'$, $J_{gg'} = J_{gg'}$. 

While sleeping at night or not etc., subject to the constraints of their socioeconomic environments. Here our key assumption is that individuals with the same socioeconomic attributes tend to behave the same way but people with different socioeconomic attributes may behave differently [19].
for every \( i \in I_{N_g} \) and \( l \in I_{N_{g'}} \). It follows from this assumption and (1) that

\[
H_N (\sigma) = \frac{1}{2N} \sum_{g=1}^{2^k} \sum_{g'=1}^{2^k} \left( \sum_{i \in I_{N_g}} \sum_{l \in I_{N_{g'}}} J_{gl} \sigma_i \sigma_l \right) + \sum_{g=1}^{2^k} h_g \sigma_g
\]

Here

\[
m_q^g = \frac{1}{N_g} \sum_{i \in I_{N_g}} \sigma_i
\]

is the average decision for the individuals in group or partition \( g \). Our model has now been changed from individual choices to group choices. Note that \( H_N \) returns the level of satisfaction for the entire population. \( J_{gg'} \) measures the influence group \( g \) has on group \( g' \) and also it is the social incentive of groups \( g \) and \( g' \) to interact. When \( J_{gg'} \) is positive, then it implies that the groups are satisfied if their empirical means have the same signs; otherwise the empirical means of the groups prefer to have different signs; i.e., it is not encouraging or rewarding for the groups to interact with one another. See Figure 1 for an artistic impression of the \( J_{gg'} \)'s. \( h_g \) is the private incentive of the group \( g \), describing how the group is satisfied with itself.

Suppose the spins \( \sigma_1, \sigma_2, \ldots \) are independent and identically distributed sequence of random variables with

\[
P (\sigma_1 = 1) = P (\sigma_1 = -1) = \frac{1}{2}
\]

We denote by \( P_N \) the corresponding product measure on \( \Omega_N = \{-1, 1\}^N \). The equilibrium state \( \omega_N \) associated with the Hamiltonian \( H_N \) in (4) is given by

\[
\omega_N (\sigma) = \frac{e^{H_N (\sigma) / (2N)}}{Z_N}, \quad \text{for } \sigma \in \Omega_N
\]

where

\[
Z_N = \sum_{\sigma \in \Omega_N} e^{H_N (\sigma) / (2N)} P_N (\sigma) = \int_{[-1, 1]^d} e^{NF_N (mN)} dQ_N (m)
\]

is the partition function of the model and \( Q_N \) is the law of empirical means \( m_N = (m_N^1, \ldots, m_N^{2^k}) \in [-1, 1]^d \) under \( P_N \) and we have set \( d = 2^k \). In (8) we have used that

\[
F_N (m_N) = \sum_{g=1}^{d} \sum_{g'=1}^{d} \frac{1}{2} J_{gg'} \gamma_{g, g'} m_N^g m_N^{g'} + \sum_{g=1}^{d} h_g \gamma_{g} m_{g}^g
\]

Note that, for any \( m \in [-1, 1]^d \),

\[
F (m) = \sum_{g=1}^{d} \sum_{g'=1}^{d} \frac{1}{2} J_{gg'} \gamma_{g, g'} m^g m^{g'} + \sum_{g=1}^{d} h_g \gamma_{g} m_{g}^g
\]

In what follows \( \mu, h, \) and \( \bar{J} \) shall be as follows

\[
\gamma = (\gamma^1, \ldots, \gamma^d),
\]

\[
\mu = (\overline{h_1}, \ldots, \overline{h_d}) \in \mathbb{R}^d
\]

and

\[
\bar{J} = (J_{g, g'})_{1 \leq g, g' \leq d}
\]

The pressure function of the model is then given by

\[
p_N = \frac{1}{N} \log Z_N
\]

The large \( N \) behaviour of the model is governed by the pressure function. It is known from [28] that the thermodynamic limit

\[
p (\gamma, h, \bar{J}) = \lim_{N \to \infty} \frac{1}{N} \log Z_N
\]

exists. The proof for the case \( \gamma_{g}^N = \gamma^g \), for all \( N \), was earlier given in [18, 20].

**Theorem 1.** For any choice of the parameters \( \gamma, h, \) and \( \bar{J} \), the limiting pressure admits the following variational representation

\[
p (\gamma, h, \bar{J}) = \sup_{m \in [-1, 1]^d} \left[ F (m) - \sum_{g=1}^{d} \gamma^g I (m^g) \right]
\]

where

\[
I (m^g) = \frac{1 - m^g}{2} \log (1 - m^g) + \frac{1 + m^g}{2} \log (1 + m^g)
\]

and \( F \) is given in (10).

The \( m^g \)’s that maximize (14) have coordinates \( m^g \) that satisfy the self-consistency equations

\[
m^g = \tanh \left( U_g \right), \quad g = 1, 2, \ldots, 2^k,
\]

where \( U_g \) is the limiting \( U_{N,g} \) given in (4); i.e.,

\[
U_g = \sum_{g'=1}^{2^k} \frac{1}{2} J_{gg'} \gamma^{g'} \gamma^g m^{g'} + h_g \gamma^g
\]
In this work we will consider the attributes of gender and residence, i.e., \( k = 2 \). This implies that we have two socioeconomic attributes and each takes two possible values; therefore we have 4 partitions. In what follows we put \( J_{gg'} = (\mathcal{J}_{gg'}/2)\gamma^g \gamma^{g'} \) and \( h_g = \gamma^g \mathcal{h}_g \)

The \( U_g \)'s for our case become

\[
U_1 = J_{11} m_1^1 + J_{12} m_2^1 + J_{13} m_3^1 + J_{14} m_4^1 + h_1 \\
U_2 = J_{21} m_1^2 + J_{22} m_2^2 + J_{23} m_3^2 + J_{24} m_4^2 + h_2 \\
U_3 = J_{31} m_1^3 + J_{32} m_2^3 + J_{33} m_3^3 + J_{34} m_4^3 + h_3 \\
U_4 = J_{41} m_1^4 + J_{42} m_2^4 + J_{43} m_3^4 + J_{44} m_4^4 + h_4
\]

The proof of this theorem is given in [29]. In [18] the case \( k = 1 \) of this model was studied under the assumption that \( \gamma^g_N = \gamma^g \), for all \( N \). The model's thermodynamic limit was proved and a rigorous derivation and some analytical properties of the model's solution were given. In particular, it was shown there that the model factorizes completely, so that all the information about the equilibrium states of the model is captured by the self-consistency equations (16).

This allows us, in particular, to write the probability that the \( r \)th individual in group \( g \) will choose \( \sigma_r \in \{-1,+1\} \) to be given by

\[
p(\sigma_r) = \frac{e^{U_g^r}}{e^{U_g^r} + e^{-U_g}}.
\]

From here we observe that the expected value of \( r \)th individual's choice is given by

\[
\mathbb{E}(\sigma_r) = \frac{e^{U_g^r}}{e^{U_g^r} + e^{-U_g}} - \frac{e^{-U_g^r}}{e^{U_g^r} + e^{-U_g}} = \tanh(U_g).
\]

Therefore, the self-consistency equations (16) imply that, for every \( g = 1, 2, \ldots, d \),

\[
m_g = \mathbb{E}(\sigma_r) = \tanh(U_g).
\]

Note from (21) that \( m_g \) is the average decision level for the individuals in group \( g \). Further, the expression for \( U_g \)'s in (18) is a linear regression model with respect to parameters \( J_{gg'} \) and \( h_g \)’s. The private incentive part \( h_g \) is also modelled as a linear regression of the attributes (2) as follows

\[
h_g = \sum_{j=1}^{k} \alpha_j g^j + \alpha_0.
\]

where \( k \) is the number of attributes, \( \alpha_j \)'s are the relative weights that individuals associate to their socioeconomic attributes, and \( \alpha_0 \) is the homogeneous private incentive for all the individuals. In our case \( k = 2 \). It follows from here that \( J_{gg'} \), \( \alpha_j \) and \( \alpha_0 \) are the parameters to be estimated.

In the case \( J_{gg'} = 0 \) for every \( g, g' \) we get the noninteracting version of the model. In that case we have that

\[
U_1 = h_1, \\
U_2 = h_2, \\
U_3 = h_3, \\
U_4 = h_4.
\]

2.2. Estimation. Due to the factorization property of the model in the thermodynamic limit, it follows from (21) that the \( m_g^N \), defined in (5), is an estimator for \( m_g \) given in (21). Note that, for large enough \( N \), \( \gamma^g_N \) is approximately equal to \( \gamma^g \). Therefore, in what follows we will suppress the \( N \) dependence of \( \gamma^g_N \). Thus \( J_{gg'} \) and \( h_g \) will be independent of \( N \).

Therefore, it follows from (21) that \( \tanh(U_g) \) is the model's prediction and \( m_g^N \) is the observed quantity which is estimated by the empirical average choice \( m_g^N \). The model parameters are then estimated with the help of least squares method. Thus we need to find the parameter values which minimize

\[
\sum_g \left[ m_g - \tanh(U_g) \right]^2,
\]

where \( m_g \) is the estimated average choice of group \( g \). The function \( \tanh(U_g) \) is nonlinear and will make computations very cumbersome. Berkson back in the nineteen-fifties, when developing a statistical methodology for bioassay, encountered a problem of this nature [30]. This stimulus-response kind of experiment bears a close relationship to the natural kind of applications for a model of social behaviour, such
as linking stimuli given by incentive through policy and media, to behaviour responses on the part of individuals in a population. Furthermore, the same approach is used by statistical mechanics, for example, within the context of finding the proper order parameter for a given Hamiltonian [31]. Since $U_g$ is a linear function of the models parameters, and tanh is an invertible function, then the above error estimation is reduced to the following

$$\sum_g \left[ \text{arctanh} \left( \overline{m}_g \right) - U_g \right]^2.$$ \hspace{1cm} (25)

The method breaks down or becomes inefficient whenever the average choice $m^g$ of group $g$ is equivalent to $+1$ or $-1$, since $\text{arctanh}(\pm1) = \pm\infty$. If the sizes of the groups are large enough this challenge or situation will be escaped. In the next section we apply the above estimation procedure to estimate the parameters for educational choices against the socioeconomic attributes of gender and place of residence.

In the interacting case, the independent variables are correlated. Due to this the least squares method is not appropriate. The partial least squares estimation is used in that case.

3. Case Study

The data used in this case study is taken from five different developing countries, namely, Ghana, Kenya, Egypt, Dominican Republic, and Indonesia. Our choice of these countries is based on our desire to compare how interaction influences choices made by individuals from countries with similar characteristics and on availability of data. Our data was taken over a six-year period from reports gathered by the Demographic and Health Surveys Program. These reports are national representative surveys of individuals in the various countries sponsored by USAID, UNICEF, UNFPA, UNDP, The Global Fund, ILO, Danida, and other national bodies. Our interest is in how place of residence and gender influence educational choices of individuals.

Under this section we will be looking at data coming from Ghana Demographic and Health Survey, Egypt Demographic and Health Survey, Encuesta Demográfica De Salud República Dominicana, and Indonesia Demographic and Health Survey [1, 21–24]. Individuals in this report are categorized according to their socioeconomic attributes such as level of education, wealth quintiles, place of residence (rural or urban), gender (male or female), age, etc. With the help of the attributes of place of residence and gender, the sampled population is categorized into four groups. Here, individuals with the same attributes are put together in a group. The key assumption here is that individuals with the same attributes will tend to behave the same way. With these choices of attributes we define

$$a_i^{(1)} = \begin{cases} 1, & \text{if individual } i \text{ is a female} \\ 0, & \text{if individual } i \text{ is a male} \end{cases}$$ \hspace{1cm} (26)

and

$$a_i^{(2)} = \begin{cases} 1, & \text{if individual } i \text{ is in urban} \\ 0, & \text{if individual } i \text{ is in rural} \end{cases}$$ \hspace{1cm} (27)

Note that if $a_i^{(1)} = 0$, then it means that that attribute has no effect on the private incentive of the individual. In the sequel we will analyse for each of the five countries the effect of social interaction and lack of it on educational choices.

We will employ the above estimation procedure and the reparameterisation of the parameters of the model according to the attributes of place of residence and gender to study educational choices made by the people in these countries. Table 1 contains the data used for the various countries.

3.1. Educational Attainment. Here we are looking at groups of individuals that have been partitioned according to these two binary attributes of gender $a_g^{(1)}$ and residence $a_g^{(2)}$. We would want to investigate how residence and gender influence choices of educational attainment for each of the countries selected. Under this case study we will be looking at individuals choosing between having some level of education against no education. With this we code the choice of the $i^{th}$ individual as follows:

$$\sigma_i = \begin{cases} +1, & \text{if individual } i \text{ has some education} \\ -1, & \text{if individual } i \text{ has no education} \end{cases}$$ \hspace{1cm} (28)

Since the data is over different period of time we will sum our quantities of interest over the different years. For instance the $U_y^g$ associated with Ghana for year $y = 1998, 1993, 1998, 2003, 2008, 2014$ is given by

$$U_y^g = \sum_{g' = 1}^{J} \sum_{y' = 1}^{J} \{ \sum_{i \in I_y^g} \sigma_i \} \overline{m}_y^g + H^g$$ \hspace{1cm} (29)

and $U_y^g = \sum_{g = 1}^{J} U_y^g$.

Note that

$$\overline{m}_y^g = \frac{1}{N_y^g} \sum_{i \in I_y^g} \sigma_i = \frac{1}{N_y^g} \left( \sum_{i \in I_y^g} N_{y, Edu}^g - N_{y, N.Edu}^g \right),$$ \hspace{1cm} (30)

where $N_{y, Edu}^g$ is the number of people in group $g$ with some education for year $y$ and $N_{y, N.Edu}^g$ is the number of individuals in group $g$ with no education for year $y$ and $N_y^g$ is the number of individuals in group $g$ for year $y$. With this the averages $\overline{m}_y^g$ for the various years and countries can be calculated from the data set. The estimation for the parameters of the model is obtained by the use of the R statistical software.

Tables 2 and 3, respectively, contain the results for the model with interaction and the model with no interaction.

4. Discussion

The estimates for the noninteracting model will be followed by discussion on the interacting model. In the above case
Table 1: Data for the selected countries grouped according to residence and gender. Taken from [1, 21–24].

| Country       | Year | 1988 | 1993 | 1998 | 2003 | 2008 | 2014 |
|---------------|------|------|------|------|------|------|------|
| Ghana         |      |      |      |      |      |      |      |
| Residence/Gender | Female | Male | Female | Male | Female | Male | Female | Male | Female | Male |
| Urban         | 1523 | 226  | 3096 | 2621 | 1739 | 547  | 4841 | 3865 | 8830 | 7385 | 9063 | 7763 |
| Rural         | 2965 | 717  | 5784 | 5493 | 3104 | 999  | 5944 | 5511 | 10453 | 9743 | 8715 | 8005 |
| Egypt         |      |      |      |      |      |      |      |      |
| Residence/Gender | Female | Male | Female | Male | Female | Male | Female | Male | Female | Male |
| Urban         | 12996 | 11886 | 16079 | 15703 | 16659 | 16573 | 19850 | 19289 | 16648 | 16290 | 18918 | 18676 |
| Rural         | 13638 | 13110 | 19032 | 18023 | 20927 | 20994 | 26480 | 25259 | 21244 | 20267 | 29818 | 28632 |
| Kenya         |      |      |      |      |      |      |      |      |
| Residence/Gender | Female | Male | Female | Male | Female | Male | Female | Male | Female | Male |
| Urban         | 1236 | 157  | 2253 | 2108 | 2726 | 2931 | 3099 | 3051 | 3257 | 2997 | 89931 | 19729 |
| Rural         | 5914 | 1013 | 12420 | 14207 | 12667 | 11567 | 12316 | 11774 | 12805 | 11884 | 38677 | 35652 |
| Dominican Republic |      |      |      |      |      |      |      |      |
| Residence/Gender | Female | Male | Female | Male | Female | Male | Female | Male | Female | Male |
| Urban         | 9207 | 7697 | 10435 | 9338 | 1682 | 1486 | 30726 | 29026 | 37934 | 35686 | 13185 | 12418 |
| Rural         | 4969 | 5731 | 5701 | 6374 | 853 | 942 | 15249 | 16971 | 15576 | 16560 | 4399 | 4854 |
| Indonesia     |      |      |      |      |      |      |      |      |
| Residence/Gender | Female | Male | Female | Male | Female | Male | Female | Male | Female | Male |
| Urban         | 17560 | 16989 | 21216 | 20919 | 19951 | 19390 | 29786 | 29377 | 30740 | 31502 | 38557 | 37583 |
| Rural         | 38338 | 37672 | 46540 | 45964 | 46529 | 45873 | 33159 | 32889 | 40461 | 41714 | 39024 | 36901 |
Table 2: Educational Attainment: Estimates for the interacting model.

| Parameter | Ghana | Kenya | Egypt | Dominican R. | Indonesia |
|-----------|-------|-------|-------|--------------|-----------|
| $\hat{J}_{11}$ | -0.03 | -0.03 | -0.00 | 0.05 | -0.11 |
| $\hat{J}_{12}$ | 0.12 | 0.10 | 0.12 | -0.06 | -0.01 |
| $\hat{J}_{13}$ | -0.01 | 0.00 | 0.01 | -0.01 | -0.004 |
| $\hat{J}_{14}$ | 0.20 | 0.13 | -0.15 | 0.49 | -0.67 |
| $\hat{J}_{21}$ | 0.08 | 0.11 | 0.02 | 0.23 | -0.19 |
| $\hat{J}_{22}$ | 0.06 | 0.03 | 0.14 | 0.23 | 0.07 |
| $\hat{J}_{23}$ | -0.01 | 0.00 | 0.01 | -0.01 | -0.004 |
| $\hat{J}_{24}$ | 0.20 | 0.13 | -0.15 | 0.49 | -0.67 |
| $\hat{J}_{31}$ | -0.14 | -0.10 | 0.18 | -0.70 | 0.34 |
| $\hat{J}_{32}$ | -0.17 | 0.03 | -0.01 | -0.41 | 1.19 |
| $\hat{J}_{33}$ | -0.03 | 0.00 | -0.03 | 0.23 | -0.03 |
| $\hat{J}_{34}$ | 0.37 | 0.11 | -0.29 | 1.11 | 0.02 |
| $\hat{J}_{41}$ | -0.47 | -0.72 | 0.27 | 0.25 | 0.82 |
| $\hat{J}_{42}$ | 0.21 | -0.01 | -0.59 | -0.72 | 0.71 |
| $\hat{J}_{43}$ | 0.66 | -0.88 | -0.56 | 0.65 | -1.58 |
| $\hat{J}_{44}$ | 0.95 | -0.04 | -0.72 | -3.52 | 0.49 |
| $\hat{\alpha}_1$ | -0.13 | -0.26 | -0.41 | 0.04 | -0.36 |
| $\hat{\alpha}_2$ | 0.52 | 0.43 | 0.44 | 0.39 | 0.42 |
| $\hat{\alpha}_0$ | 0.43 | 0.84 | 0.62 | 0.75 | 1.05 |

Table 3: Educational Attainment: Estimates for the non-interacting model.

| Parameters | Ghana | Kenya | Egypt | Dominican R. | Indonesia |
|-----------|-------|-------|-------|--------------|-----------|
| $\hat{\alpha}_1$ | -0.29 | -0.26 | -0.41 | 0.04 | -0.36 |
| $\hat{\alpha}_2$ | 0.52 | 0.43 | 0.44 | 0.39 | 0.42 |
| $\hat{\alpha}_0$ | 0.43 | 0.84 | 0.62 | 0.75 | 1.05 |

Table 4: Groups and their interpretations.

| Group Number | Interpretation                        |
|--------------|---------------------------------------|
| 1            | Females in urban area                 |
| 2            | Males in urban area                   |
| 3            | Females in rural area                 |
| 4            | Males in rural area                   |

study there are four groups and these groups are explained in Table 4.

4.1. Noninteracting Case. When the sum of the $\hat{\alpha}_i$’s of the private incentive of a group is positive then that group will make a choice favouring having some education and when the sum is negative then the group will make a choice favouring no education. The estimates for this case have been collected in Table 3.

The sum of the $\hat{\alpha}_i$’s are all positive except for Egypt having a negative sum. The $\hat{\alpha}_i$’s of Dominica Republic are all positive signifying that the attribute of gender and residence has a significant contribution to the private incentive of the people in Dominican Republic when it comes to educational attainment. For the same country, the estimate of the private incentive for gender $\hat{\alpha}_1$ is close to zero. Implying gender does not really contribute to the private incentive of individuals towards educational attainment when interaction is absent in Dominican Republic. The estimates for $\hat{\alpha}_1$ for the rest of the countries are all negative indicating that being a female favours the no education choice.

All the selected countries have positive estimates for the private incentive of residence $\hat{\alpha}_2$. This implies that residing in the urban area has a positive effect on the choice of attaining some level of education in those countries.

The base private incentive $\hat{\alpha}_0$ has positive values for all the five countries, signifying that individuals will prefer the choice of having some level of education. In fact, the base private incentive part is greater than the sum of the estimates for the private incentives associated with gender and residence. Thus choices under the noninteracting model are dominated by the base private incentive.

4.2. Interacting Case. The interacting model has a utility function that consists of both social and private incentives. Whenever the coefficient of the social incentive $\hat{J}_{gg’}$ for groups $g$ and $g’$ is positive, individuals in groups $g$ and $g’$ prefer to imitate themselves, while when $\hat{J}_{gg’}$ is negative, the average choices of individuals in groups $g$ and $g’$ will prefer to have different signs. Here conformity is not rewarded. When the $\hat{\alpha}_i$’s of the private incentive of a group sum up to a positive value, then at the private level individuals in that group will make a choice favouring the attainment of some level of education and when it sums up to a negative value
### Table 5: Variance of $m^g$ and $U^g$ explained by the latent vectors and RMSEP for Ghana.

| Latent Vectors | Percentage of Explained Variances for $m^g$ | Cumulative Percentage of Explained Variances for $m^g$ | Percentage of Explained Variances for $U^g$ | Cumulative Percentage of Explained Variances for $U^g$ | RMSEP |
|----------------|---------------------------------------------|------------------------------------------------------|---------------------------------------------|------------------------------------------------------|-------|
| 1              | 33.4                                       | 33.34                                                | 82.7                                       | 82.7                                                 | 0.1772|
| 2              | 6.13                                       | 39.53                                                | 12.18                                      | 94.88                                                | 0.1549|
| 3              | 25.89                                      | 65.42                                                | 1.57                                       | 96.45                                                | 0.1081|
| 4              | 28.21                                      | 93.63                                                | 0.15                                       | 96.60                                                | 0.0860|

### Table 6: Variance of $m^g$ and $U^g$ explained by the latent vectors and RMSEP for Kenya.

| Latent Vectors | Percentage of Explained Variances for $m^g$ | Cumulative Percentage of Explained Variances for $m^g$ | Percentage of Explained Variances for $U^g$ | Cumulative Percentage of Explained Variances for $U^g$ | RMSEP |
|----------------|---------------------------------------------|------------------------------------------------------|---------------------------------------------|------------------------------------------------------|-------|
| 1              | 49.37                                       | 49.37                                                | 87.07                                      | 87.07                                                | 0.1347|
| 2              | 33.1                                        | 82.47                                                | 0.35                                       | 87.42                                                | 0.1362|
| 3              | 16.73                                       | 99.20                                                | 0.73                                       | 88.15                                                | 0.3740|
| 4              | 0.62                                        | 99.82                                                | 10.64                                      | 99.79                                                | 0.4133|

### Table 7: Variance of $m^g$ and $U^g$ explained by the latent vectors and RMSEP for Egypt.

| Latent Vectors | Percentage of Explained Variances for $m^g$ | Cumulative Percentage of Explained Variances for $m^g$ | Percentage of Explained Variances for $U^g$ | Cumulative Percentage of Explained Variances for $U^g$ | RMSEP |
|----------------|---------------------------------------------|------------------------------------------------------|---------------------------------------------|------------------------------------------------------|-------|
| 1              | 33.60                                       | 33.60                                                | 82.06                                      | 82.06                                                | 0.1794|
| 2              | 10.64                                       | 44.24                                                | 12.99                                      | 95.05                                                | 0.1337|
| 3              | 18.29                                       | 62.53                                                | 4.24                                       | 99.29                                                | 0.08637|
| 4              | 20.86                                       | 83.39                                                | 0.13                                       | 99.42                                                | 0.06150|
| 5              | 11.54                                       | 94.93                                                | 0.14                                       | 99.56                                                | 0.04821|

### Table 8: Variance of $m^g$ and $U^g$ explained by the latent vectors and RMSEP for Dominican Republic.

| Latent Vectors | Percentage of Explained Variances for $m^g$ | Cumulative Percentage of Explained Variances for $m^g$ | Percentage of Explained Variances for $U^g$ | Cumulative Percentage of Explained Variances for $U^g$ | RMSEP |
|----------------|---------------------------------------------|------------------------------------------------------|---------------------------------------------|------------------------------------------------------|-------|
| 1              | 38.36                                       | 38.36                                                | 51.38                                      | 51.38                                                | 0.2401|
| 2              | 23.28                                       | 61.64                                                | 6.80                                       | 58.18                                                | 0.2291|
| 3              | 24.13                                       | 85.77                                                | 14.71                                      | 72.89                                                | 0.1619|
| 4              | 13.09                                       | 98.86                                                | 22.26                                      | 95.15                                                | 0.1041|
| 5              | 0.45                                        | 99.31                                                | 2.93                                       | 98.08                                                | 0.07877|

### Table 9: Variance of $m^g$ and $U^g$ explained by the latent vectors and RMSEP for Indonesia.

| Latent Vectors | Percentage of Explained Variances for $m^g$ | Cumulative Percentage of Explained Variances for $m^g$ | Percentage of Explained Variances for $U^g$ | Cumulative Percentage of Explained Variances for $U^g$ | RMSEP |
|----------------|---------------------------------------------|------------------------------------------------------|---------------------------------------------|------------------------------------------------------|-------|
| 1              | 28.06                                       | 28.06                                                | 82.7                                       | 59.24                                                | 0.2275|
| 2              | 43.05                                       | 71.11                                                | 12.18                                      | 61.36                                                | 0.2128|
| 3              | 20.56                                       | 91.67                                                | 1.57                                       | 72.27                                                | 0.1835|
| 4              | 8.03                                        | 99.70                                                | 0.15                                       | 98.17                                                | 0.06593|
then individuals in the group will make a choice favouring no education. The estimates for this case are found in Table 2.

Note that $J_{11}$ is the interaction strength of females in an urban area interacting with themselves. The estimate for $J_{11}$ is negative for all the countries except for Dominican Republic, where the estimate is positive. This signifies that imitation is not rewarding in all the countries when it comes to females in urban areas interacting with themselves, while such an interaction is rewarding in the Dominican Republic. The estimate $J_{22}$ is positive for all the five countries. Thus males in urban area interacting with themselves could reinforce either attainment of some level of education or no education. $J_{34}$ is also positive for all the countries except Egypt. It can be observed that Dominican Republic has a higher estimate than the other four countries. This implies that females in the rural areas in the Dominican Republic have greater influence on males in the rural areas when it comes to educational attainment. The estimate for Egypt is negative. $J_{44}$ is positive for Ghana and Indonesia but it is negative for Kenya, Egypt, and Dominican Republic.

For $J_{41}$, Indonesia has the highest estimate. This indicates that in Indonesia males in rural area have greater influence on females in urban area. Ghana and Kenya have negative estimates for this parameter.

The estimate $\bar{\alpha}_1$ is positive for Egypt, which indicates that gender has a significant effect or influence on the educational choices made by individuals in that country. On the contrary, the estimates are negative for the remaining countries.

4.3. Model Diagnostics and Validation. Under this section our interest is to look at how well our interacting model fits the data. The partial least squares (PLS) method was used to estimate the parameters of the interacting model found in (18). From that equation the $U_g$ is used as our dependent variable and $m^p$ as the independent variable. The PLS method analyses (predicts) a set of dependent variables from a set of independent variables (predictors). In this method, orthogonal factors, called latent vectors, are generated from the independent variables and those factors arranged in decreasing order of their eigenvalues. The choice of the number of latent vectors to use is dependent on the vectors that can best explain the covariance between the independent and the dependent variables [32].

Our results and analyses are performed in the R statistical software, version 3.5.2. Tables 5 to 9 indicate the amount of variances explained in the independent variables ($m^p$) and the dependent variable ($U_g$) for each of the selected countries in terms of the latent vectors used in the prediction together with their root mean square error of prediction (RMSEP).

It is observed from Table 5 that the first three latent vectors explain 96.45% of variance in the $U_g$ and 65.42% of variance in the $m^p$, for the Ghanian case. The variances explained in the dependent variable are good enough to be kept for modelling or prediction but considering the variances explained in the independent variable suggests that the first three latent vectors are not good enough. When the four latent vectors are considered together, it is observed that 93.63% and 96.60% of variances in the independent and dependent variables are explained, respectively. The RMSEP is observed to be decreasing as the number of latent vectors is increased in all the countries with the exception of the Kenya case. Similar trends are for the remaining four countries in Tables 6, 7, 8, and 9.

5. Conclusion

The above study gives credence to the potential of statistical mechanical models to socioeconomic applications, as has been suggested by the authors of [2–5]. The statistical mechanical formulation allows us to use interactions among individuals to explain the macroscopic behaviour of such socioeconomic systems. Usually, what the socioeconomicist takes as their building blocks turn out to be the limiting objects that result from statistical mechanical models. In fact, the statistical mechanical model studied here is a special case of a class of probabilistic choice models in the socioeconomic literature called the Luce or the multinomial logit models. There are other probabilistic choice models such as multinomial-probit models and the generalized extreme-valued models [33]. The question here is what are the statistical mechanical analogues of these models? The understanding gleaned from these statistical mechanical models will go a long way to aid the proper understanding of the socioeconomic models. Further, starting from the statistical mechanical models we can properly reparametrise the parameters of the socioeconomic models to aid their estimation. In fact, all these should result in building a software to make these models accessible to practitioners.

Data Availability

The data used in this study are survey data carried out by the Demographic and Health Survey program in developing countries. The data used are found in [1, 21–24].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors thank the staff at the Mathematics and Statistics Department of University of Energy and Natural Resources for their support and kindness during the period this paper was written.

References

[1] Ghana Statistical Service (GSS), Ghana Health Service (GHS), and ICF International. Ghana Demographic and Health Survey 2014. Rockville, Maryland, USA: GSS, GHS, and ICF International, (1988, 1993, 1998, 2003, 2008, 2015).
[2] W. A. Brock and S. N. Durlauf, “Discrete choice with social interactions,” Review of Economic Studies, vol. 68, no. 2, pp. 235–260, 2001.
[3] S. N. Durlauf, “How can statistical mechanics contribute to social science?” Proceedings of the National Academy of Sciences
of the United States of America, vol. 96, no. 19, pp. 10582–10584, 1999.

[4] W. A. Brock and S. N. Durlauf, Discrete choice with social interaction I: Theory, NBER Working Paper Series, 1995.

[5] D. McFadden, “Economic choices,” American Economic Review, vol. 91, no. 3, pp. 351–378, 2001.

[6] G. B. Ermentrout and D. H. Terman, Mathematical foundations of neuro-science, vol. 35, Springer Science and Business Media, 2010.

[7] P. Turchin and A. D. Taylor, “Complex dynamics in ecological time series,” Ecology, Ecological Society of America, vol. 73, no. 1, pp. 289–305, 1992.

[8] W. C. Chen, “Nonlinear dynamics and chaos in a fractional-order financial system,” Chaos, Solitons & Fractals, vol. 36, no. 5, pp. 1305–1314, 2008.

[9] T. Kuran, Now Out of Never, World politics, 1991.

[10] P. Ball, Critical Mass, Arrow books, UK, 2004.

[11] P. Dai Pra, G. Giacomin, and D. Regoli, “Noise-induced periodicity: some stochastic models for complex biological systems,” in Mathematical Models and Methods for Planet Earth, vol. 6, pp. 25–35, Springer, 2014.

[12] F. Collet, P. Dai Pra, and M. Formentin, “Collective periodicity in mean-field models of cooperative behavior,” Nonlinear Differential Equations and Applications NoDEA, vol. 22, no. 5, pp. 1461–1482, 2015.

[13] F. Collet, P. Dai Pra, and E. Sartori, “A simple mean field model for social interactions: dynamics, fluctuations, criticality,” Journal of Statistical Physics, vol. 139, no. 5, pp. 820–858, 2010.

[14] R. Frey and J. Backhaus, “Pricing and hedging of portfolio credit derivatives with interacting default intensities,” International Journal of Theoretical and Applied Finance, vol. 11, no. 6, pp. 611–634, 2008.

[15] A. Di Biasio, E. Agliari, A. Barra, and R. Burioni, “Mean-field cooperativity in chemical kinetics,” Theoretical Chemistry Accounts, vol. 131, no. 3, pp. 1–14, 2012.

[16] I. Volkov, J. R. Banavar, S. P. Hubbell, and A. Maritan, “Neutral theory and relative species abundance in ecology,” Nature, vol. 424, no. 6952, pp. 1035–1037, 2003.

[17] P. Schiffer and I. Daruka, “Two-population model for anomalous low-temperature magnetism in geometrically frustrated magnets,” Physical Review B, vol. 56, no. 21, pp. 13712–13715, 1997.

[18] I. Gallo and P. Contucci, “Bipartite mean field spin systems. Existence and solution,” Mathematical Physics Electronic Journal, vol. 14, 2008.

[19] I. Gallo, A. Barra, and P. Contucci, “Parameter evaluation of a simple mean-field model of social interaction,” Mathematical Models and Methods in Applied Sciences, vol. 19, pp. 1427–1439, 2009.

[20] I. Gallo, An equilibrium approach to modelling social interaction [Ph.D. thesis], University of Bologna, Bologna, Italy, 2009.

[21] Ministry of Health and Population [Egypt], El-Zanaty and Associates [Egypt], and ICF International 2015. Egypt Demographic and Health Survey 2014. Cairo, Egypt and Rockville, Maryland, USA: Ministry of Health and Population and ICF International (1992, 1995, 2000, 2005, 2008, 2014).

[22] Kenya National Bureau of Statistics (KNBS) and ICF Macro, 2010. Kenya Demographic and Health Survey 2008-09. Calverton, Maryland: KNBS and ICF Macro (1989, 1993, 1998, 2003, 2008, 2014).

[23] “Centro de Estudios Sociales y Demográficos (CESDEM) y ICF International, 2014,” Encuesta Demografica y de Salud 2013. Santo Domingo, Republica Dominicana: CESDEM y ICF International (1991, 1996, 1999, 2002, 2007, 2013).

[24] Statistics Indonesia (Badan Pusat StatistikBPS), National Population and Family Planning Board (BKKKB), and Kementerian Kesehatan (KemenkesMOH), and ICF International 2013. Indonesia Demographic and Health Survey 2012. Jakarta, Indonesia: BPS, BKKBN, Kemenkes, and ICF International (1991, 1994, 1997, 2002-2003, 2007, 2012).

[25] J. V. Selinger, Introduction to the Theory of Soft Matter, Soft and Biological Matter, Springer International Publishing, Switzerland, 2016.

[26] P. Contucci, I. Gallo, and S. Ghirlanda, “Equilibria of culture contact derived from ingroup and outgroup attitudes,” https://arxiv.org/pdf/0712.1119.pdf.

[27] K. S. Ng, “A simple explanation of partial least squares,” Draft, 2013.

[28] A. A. Opoku and G. Osabutey, “A multi-population Curie-Weiss model for binary choice with social interaction,” in preparation.

[29] A. A. Opoku and G. Osabutey, “Multipopulation spin models: a view from large deviations theoretic window,” Journal of Mathematics, vol. 2018, Article ID 9417547, 13 pages, 2018.

[30] J. Berkson, “A statistically precise and relatively simple method of estimating the bio-assay with quantal response, based on the logistic function,” Journal of the American Statistical Association, vol. 48, no. 236, pp. 565–599, 1953.

[31] A. Barra, “The mean field Ising model through interpolating techniques,” Journal of Statistical Physics, vol. 145, pp. 234–261, 2008.

[32] H. Abdi, Partial Least Squares Regression and Projection on Latent Structure Regression (PLS Regression), John Wiley & Sons, Inc. WIREs Comp Stat, 2010.

[33] D. McFadden, “Econometric models for probabilistic choice among products,” The Journal of Business, vol. 53, no. 3, pp. S13–S29, 1980, Part 2: Interface Between Marketing and Economics.
