Accepted Manuscript

Correlation between Gini index and mobility in a stochastic kinetic model of economic exchange

Maria Letizia Bertotti, Amit K. Chattopadhyay, Giovanni Modanese

PII: S2211-3797(17)30323-6
DOI: http://dx.doi.org/10.1016/j.rinp.2017.05.031
Reference: RINP 748

To appear in: *Results in Physics*

Please cite this article as: Bertotti, M.L., Chattopadhyay, A.K., Modanese, G., Correlation between Gini index and mobility in a stochastic kinetic model of economic exchange, *Results in Physics* (2017), doi: http://dx.doi.org/10.1016/j.rinp.2017.05.031

Results in Physics (2017)

Contents lists available at ScienceDirect

Results in Physics

journal homepage: www.elsevier.com/locate/results-in-physics

Correlation between Gini index and mobility in a stochastic kinetic model of economic exchange

Maria Letizia Bertotti[a,1], Amit K. Chattopadhyay[b,2], Giovanni Modanese[a,*]

[a] *Free University of Bolzano-Bozen, Bolzano 39100, Italy*
[b] *Aston University, Birmingham B4 7ET, UK*

ARTICLE INFO

Article history:
Received ...
Received in final form ...
Accepted ...
Available online ...

Communicated by ...

Keywords: Langevin, Gini, Ornstein-Uhlenbeck, Fokker-Planck, Hurst

ABSTRACT

Starting from a class of stochastically driven kinetic models of economic exchange, here we present results highlighting the correlation of the Gini inequality index with the social mobility rate, close to dynamical equilibrium. Except for the canonical-additive case, our numerical results consistently indicate negative values of the correlation coefficient, in agreement with empirical evidence. This confirms that growing inequality is not conducive to social mobility which then requires an "external source" to sustain its dynamics. On the other hand, the sign of the correlation between inequality and total income in the canonical ensemble depends on the way wealth enters or leaves the system. At a technical level, the approach involves a generalization of a stochastic dynamical system formulation, that further paves the way for a probabilistic formulation of perturbed economic exchange models.

1. Introduction

This article reports results obtained from an application of a statistical mechanics model to a socio-economic problem, in an area that is currently attracting much interest, namely the relation between the Gini index of a country and its social mobility.

The Gini index G is a widespread measure of income inequality in a society expressed as a non-dimensional ratio of the relative mean absolute difference of income between two income classes to double their mean [1]. The social mobility M can be identified with multiple definitions [2], but in essence it is defined as the probability for an individual to pass to the upper income class in a given unit time, averaged over all classes. Empirical evidence shows a clear correlation between these two quantities, namely it is found that mobility reduces when inequality rises, thus implying a negative correlation between G and M [3, 4]. This correlation, nicknamed the "Great Gatsby Curve" [5], is important since it means that the increase of inequality (as presently observed in several countries) tends to be a self-reinforcing phenomenon, unless it is complemented by suitable social policies. It should also be stressed that this correlation holds for societies at near equilibrium, while it may be different in phases of strong economic growth [6].

The analytical structure of the model employed is a derivative of our well established kinetic model [8] that we briefly recall below. The key feature of this macroeconomic model was its ability to allow the computation of the income distribution of an idealized society as a macroscopic feature emerging from elementary microscopic interactions between individuals of different income classes. Using extensive numerical simulations and related analysis, we arrive at a clear confirmation of the negative correlation between G and M, when in equilibrium. In line with our recent work incorporating the impact

*Corresponding author: Tel.: +39-0471-017134; fax: +39-0471-017009; *e-mail:* giovanni.modanese@unibz.it (Giovanni Modanese)
[1]E-mail: marialetizia.bertotti@unibz.it
[2]E-mail: a.k.chattopadhyay@aston.ac.uk

of randomness in monetary transactions [14], we now introduce the vital stochastic element in the dynamics which then leads us to a set of Langevin equations driven by additive or multiplicative noise. Due to initial and boundary condition related constraints controlling the stochastic dynamics, this leads to nontrivial analysis of constrained socioeconomic dynamics that we study under the following separate ambits: Case A: total population conservation, and Case B: total population conservation and total income conservation.

We find that this potentially leads to a variety of possible cases. Using thermodynamic analogy, one may consider both a canonical ensemble (system in contact with the outside world) or otherwise a microcanonical (isolated system, with conserved total income), driven separately by additive and/or multiplicative noise, which could be drawn from an ensemble of distributions (e.g. white Gaussian noise or Ornstein-Uhlenbeck noise with memory effects).

For the case (Case A) of a fluctuating total income (canonical ensemble), it is also possible to compute the correlation between the Gini index G and the total income μ, defined as the sum of the class populations each multiplied by the average income of the class (see Sect. 2). This is of interest to economists because it is often debated whether an increase in a countrys GDP has generally a positive or negative influence on inequality [7]. Towards this end, our recent results show that the G-μ correlation depends on the nature of noise, being positive for additive noise and negative for multiplicative noise (in a certain range of G). In this sense, it is less general than the G-M correlation. We discuss this result in Sect. 3.

The manuscript is organized as follows. In Sect. 2, the original model is briefly recalled, followed by Sect. 3 where results obtained in Sect. 2 are analyzed statistically. Sect. 4 contains our conclusions and outlook.

2. Mathematical model

Our model [8] is based upon a system of n kinetic nonlinear differential equations. The variables are the populations of n income classes which make up a society of individuals interacting among themselves and possibly with the external world. Typically n varies between 10 and 100. The equations are strongly coupled and can only be solved numerically, due to the complexity which arises from nonlinearity. They contain adjustable constant coefficients which describe the transition probabilities between income classes due to economic interactions. Unlike in linear master equations, however, dynamics is also crucially determined by the probabilities of encounters, which are proportional to the products of the class populations.

The subdivision of the total population into classes is an essential feature of the discretized kinetic theory, and allows to recast the Boltzmann equation as a system of ordinary differential equations, instead of a partial differential equation. While for, say, a gas of physical particles, this subdivision may seem artificial, in economics the use of income classes is quite frequent and especially natural when considering taxation. Also the empirical studies on social mobility employ a subdivision of society in income classes. The choice of n in our model is a matter of convenience, but it is possible to use more or less classes without any substantial changes in the results, provided the simulations are run with the same total income.

The solution of the coupled equations represents the time evolution of the class populations $x_i(t), i = 1, ..., n$. It turns out that, starting from arbitrary initial conditions, the system independently evolves towards an equilibrium configuration, namely the equilibrium income distribution, which depends only on the initial total income and on the model parameters. From the equilibrium configuration, one can evaluate several quantities of economic interest, like the Gini inequality index, the social mobility index, the Pareto exponent (for a possible extended non-Gaussian tail in the probability density function), etc.

In [9] the income distributions obtained from our model have been fitted with the Kaniadakis κ-distribution, a very general three-parameter distribution which gives an excellent description of empirical data both with and without Pareto tails. The Pareto exponents of our distributions turn out to be typically in the range $2.5 < a < 3.5$, where a is the power-law exponent ($f \sim r^{-a}$, where f is the income distribution and r is the income variable). The value of a depends on the parameters defining the microscopic interactions.

This kind of equations can also clearly be applied to model various physical systems, where interactions are present which depend on the product densities or concentrations, like for instance in chemical kinetics, crystal growth, etc. [10].

In more sophisticated versions of the model, redistribution terms (appearing as third-ordered correction terms in the dynamics) have been introduced. From the economic perspective, these terms represent the effects of taxation and welfare benefits [11]. A network structure can also be implemented to parlay inhomogeneous interactions that are directed along certain preferred links [12]. Here we refer to the basic version of the model, constituted out of quadratic terms, and represented as follows

$$\frac{dx_i(t)}{dt} = \sum_{h,k=1}^{n} C_{hk}^{i} x_h(t) x_k(t) - \sum_{h,k=1}^{n} C_{ik}^{h} x_i(t) x_k(t), \quad (1)$$

for $i = 1, 2, ...n$, where the constant coefficients C_{hk}^i, satisfying for any fixed h and k the condition $\sum_{i=1}^{n} C_{hk}^i = 1$, express the probability that an individual of the h class will belong to the i class after a direct interaction with an individual of the k class; they define all the features of the model, as described in detail in the cited works [8, 9], and allow a large degree of flexibility. The class populations x_i are normalized to 1.

All analyzes are based on the previously alluded definition of Gini index [1]. Our definition of equilibrium mobility is essentially the weighted average, over all classes, of the probability for an individual to be promoted to the upper class in the unit time [13, 14]. Note that the model allows the definition and computation of mobility because the equilibrium is dynamical, i.e., it is obtained as a balance of up and down transitions along the income ladder.

Let us now consider a system which is at equilibrium, but is also subjected to random fluctuations of the populations, representing monetary transactions with stochastic perturbations. Such fluctuations can be of two different kinds: canonical fluctuations, with conservation of the total population but not of the total income (case A), and micro-canonical, with conservation both of the total population and of the total income (case B). In economic literature, these are more popularly referred to as "exogenous" and "endogenous" fluctuations, respectively. Case A refers to a system that can interact with the external world, as is represented in the case of import/export of goods and capitals, or in the case of incoming/outgoing tourism. Case B, on the other hand, refers to fluctuations that occur when the system is isolated and its total wealth remains unaffected in the dynamic equilibrium state. In this case, the income distribution changes not only due to transitions with fixed probabilities based on the C^i_{hk} coefficients, but also due to random transitions caused by temporary, uncontrolled factors.

The Langevin equation with additive noise is written in general

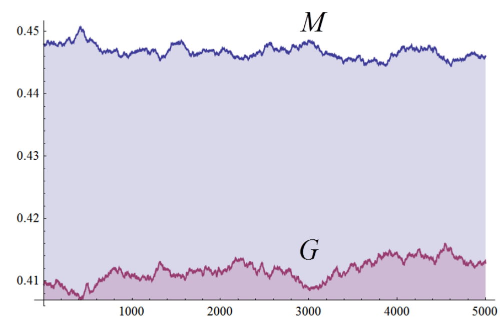
$$dx_i = D^{(1)}_i(x)dt + \sum_j D^{(2)}_{ij}(x)\xi_j \Gamma \sqrt{dt}, \tag{2}$$

where the deterministic term $D^{(1)}_i(x)$ is as in eq. (1), ξ_i denotes n independent Gaussian stochastic variables and the matrix $D^{(2)}$ in the canonical case is such as to transform the ξ_i only by imposing $\sum_{i,j} D^{(2)}_{ij}\xi_j = 0$, while in the micro-canonical case it must also satisfy the conditions $\sum_{i,j} r_i D^{(2)}_{ij}(x)\xi_j = 0$ (see [14]). Here r_i is the income of the class i, taken equal to $10i$ in the present computations. The total income is given by $\mu = \sum_{i=1}^n x_i r_i$.

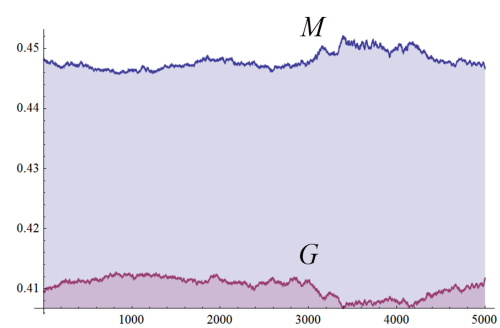
The case of multiplicative noise is more involved, because it intuitively relates to the expectation that random variations of the populations are proportional to the population themselves. (Note that multiplicative noise has been previously used in the Bouchaud-Mezard model [15, 16], where it represents the effect of investment in the stock market. The evolution equations of that model, however, are linear, and do not represent binary interactions like the kinetic equations; there are no income classes and the time evolution of the income of each individual depends on the income of the other $(N-1)$ individuals, but not on its own income.) In the canonical case with multiplicative noise, the matrix $D^{(2)}$ is as given in [14], while for the micro-canonical case with multiplicative noise the transformation needed is more complex and will be presented in detail in a forthcoming paper [17].

3. Results

The Langevin equations are solved numerically with a discrete-time algorithm based on a Taylor-Euler scheme of discretization. At each time step, a noise vector ξ_i is generated and suitably normalized; the noise is added to the deterministic model, either as an additive or a multiplicative term (proportional to the population), with tunable amplitude Γ. The deterministic term represents the time dynamical evolution with fixed probability, while the stochastic term represents the effects of internal and external perturbations. A control loop ensures that the variables x_i remain always positive after the stochastic



(a) Micro-canonical, additive noise; M and G versus time t

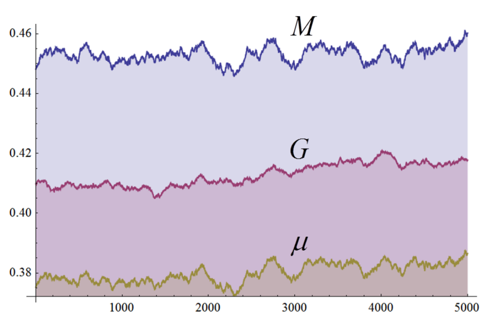


(b) Micro-canonical, multiplicative noise; M and G versus time t

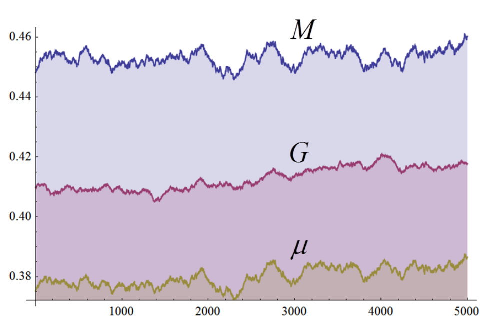
Fig. 1. Samples of time series of mobility M and Gini index G in the micro-canonical case. A negative correlation is clearly visible (compare Table 1, first column). For graphical reasons, the real value of M has been multiplied by 800.

Table 1. Correlations R_{GM} (Gini-mobility) and $R_{G\mu}$ (Gini-total income) computed with initial total income $\mu = 30$ and noise amplitude $\Gamma = 0.001$. Averages of 50 realizations, each of 5000 integration steps.

	Conserved total income μ (micro-canonical)	Non-conserved total income μ (canonical)
Additive noise	$R_{GM} = -0.79 \pm 0.02$	$R_{GM} = 0.53 \pm 0.05$ $R_{G\mu} = 0.68 \pm 0.04$
Multiplicative noise	$R_{GM} = -0.90 \pm 0.06$	$R_{GM} = -0.58 \pm 0.05$ $R_{G\mu} = -0.43 \pm 0.05$



(a) Canonical, additive noise; M and G versus time t



(b) Canonical, multiplicative noise; M and G versus time t

Fig. 2. Samples of time series of mobility, Gini index and total income in the canonical case (compare Table 1, second column). For graphical reasons, the real value of M has been multiplied by 800 and that of μ divided by 80. The multiplicative case shows a negative correlation between M and G, in agreement with the micro-canonical case and the empirical data; the G-μ correlation is also negative. The additive case is anomalous, showing positive correlations between G, M and μ.

variations. When $\Gamma = 0$, the system evolves towards an equilibrium income distribution which depends on μ but not on the initial conditions. In order to arrive at a solution for the stochastic model with $\Gamma \neq 0$, the deterministic equilibrium is taken as the initial condition, which also ensures faster convergence. (Examples of these time-series are shown in Figs. 1 and 2, where the correlations between M, G and μ are also qualitatively visible.) For each population variable x_i we obtain a stochastic time series, typically comprising $\sim 10^4$ values. We study the standard features of these time series, like statistical mean and variance, time auto-correlation and Hurst exponent. However, our main interest is in the time series of integral quantities, like the total income μ and the Gini inequality index G. For white Gaussian noise, these resemble closely a Wiener process with time auto-correlation decreasing over a scale of $\sim 10^3$ steps and $H \sim 0.50$.

The correlation R_{GM} between G and M (see Table 1) is negative both for additive and multiplicative noise (except in the additive-canonical case). On the other hand, $R_{G\mu}$ is positive for additive noise and negative for multiplicative noise. Note that all temporal autocorrelation functions are computed at the same time points. The values given in Table 1 are estimated for a total income $\mu = 30$, in non-dimensional units that are related to the Gini index G through a deterministic formulation already discussed in [8], e.g. $\mu = 30$ corresponds to $G = 0.41$. The correlation $R_{G\mu}$ with multiplicative noise has constant sign in a range of μ and G of a few percent around these values. For its behavior in a wider range see [17].

The introduction of an Ornstein-Uhlenbeck noise (with autocorrelation time τ) in the dynamics does not substantially affect the μ-G and M-G correlations, the most visible effect being that on the Hurst exponent of the time series, which then approaches the value 1, suggesting a Wiener process with memory. Work is ongoing on the computation of the values of the μ-G correlation in the phase space τ-Γ, that may lead to hidden phase transitions.

For a correct interpretation of these results, it is important to notice that canonical additive random variations in the class populations have a more prominent effect on the rich classes, whose populations are much smaller. Moreover, additive population variations in the rich classes cause much bigger variations in the total income. This explains why the $R_{G\mu}$ correlation is positive in the canonical additive case. In fact, in this case, the sign of the total correlation is dominated by the behavior of the rich classes. Any increase in the total income μ due to an external inflow is strongly correlated with an increase of the population of the rich classes, and therefore with an increase in the value of G. Also, the mobility increase which leads to a positive R_{GM} correlation is exclusive of the relatively richer classes. On the other hand, multiplicative noise invokes a better balance across income classes, hence is more realistic.

The reason why the micro-canonical R_{GM} correlation is negative even with additive noise can be explained as follows: here any population variation in the rich classes will precipitate a greater variation in the other classes, due to an avalanche effect if the total income is still to be conserved. In practice, this will redistribute the population variations occurring in the rich classes, thereby redistributing wealth.

4. Conclusions, outlook

In this work, we have proposed a Langevin-type stochastic formulation of economic exchange based on a set of discrete stochastic kinetic equations describing a system with n interacting classes and their populations $x_1, ..., x_n$. Given the transition probabilities between these classes, the deterministic equations lead us to the dynamical equilibrium configuration of the system. Other quantities of interest, like for instance the relaxation times, the initial conditions which lead to equilibrium, etc. could also be estimated starting from this structure. What the noise-driven formulation does is to allow us to assess possibilities of deviation from general "deterministic" trends which are intrinsically driven by transient fluctuations in market assets. In turn, this strongly depends on the nature of noise perturbing such market trends, with possibilities of a global economic avalanche perpetrated through local modulations, a *butterfly-effect* analogy.

The augmented models defined in Eqs. (1) and (2) can cater to realms outside the purview of the immediate economic contexts presented here. In the kinetic equations, the transition probabilities are proportional to products $x_i x_j$ of populations, which express the probability of encounters. In the original Boltzmannian formulation of the microscopic kinetic theory, such encounters were thought of as collisions between molecules; in chemistry-based applications, the encounters are reversible reactions between n reagents with concentrations $x_1, ..., x_n$, etc. The model then describes which kind of encounters could occur, and which quantities are conserved in these encounters, and so on.

For instance, in [10], the variables $x_1, ..., x_n$ denote the masses of n grains of a ceramic or metallic powder which grow to produce a crystal aggregate under the effect of an applied temperature and pressure. The total mass of the grains is constant, but in certain conditions the equilibrium solution of the kinetic equations describes grains of homogeneous size, while in other conditions it shows that some grains grow much more than others. The equations may also contain noise terms, which can randomly drive the system from one regime to the other. (Phenomenologically, the noise originates from local temperature and density fluctuations of the material.)

In any case, the noise should preserve, completely or partially, those dynamical quantities which are conserved in the deterministic equations. This may be easy to implement for small n, but is non-trivial for large n. The relevant literature do offer recipes for the introduction of noise into constrained systems, also for high dimensional values of n. For instance, in the Brownian motion of large molecules, or polymers, certain degrees of freedom are regarded as frozen, while others are free to fluctuate [18]. In the latter case the constraint is imposed explicitly in the deterministic equations through Lagrange multipliers, and is also taken into account in the Langevin or Fokker-Planck equations. The same applies to the case of a brownian motion of a particle in n dimensions constrained to move on a line or a surface [19]. On the other hand, the deterministic kinetic equations are such that certain quantities are conserved dynamically, not through explicit constraints. Therefore, the introduction of noise requires intrinsically different methods, compared to deterministic modeling. In this paper we have proposed a possible technique, along with some concrete applications and preliminary results.

References

[1] P.M. Dixon, J. Weiner, T. Mitchell-Olds and R. Woodley, Bootstrapping the Gini Coefficient of Inequality, Ecology 68, 1548-1551 (1987).

[2] M. Corak, Income inequality, equality of opportunity, and intergenerational mobility, J. Econ. Perspect. 27, 79-102 (2013).

[3] D. Andrews and A. Leigh, More inequality, less social mobility, Appl. Econ. Lett. 16, 1489-492 (2009).

[4] D. Spannagel and S. Broschinski, Reichtum in Deutschland waechst weiter, WSI Report 17/2014 (in German).

[5] J. Jerrim and L. Macmillan, Income inequality, intergenerational mobility, and the Great Gatsby Curve: is education the key? Social Forces 94, 505-533 (2015).

[6] P. Aghion, E. Caroli and C. Garcia-Penalosa, Inequality and economic growth: the perspective of the new growth theories, J. Econ. Lit. 37, 16151660 (1999).

[7] D. Dollar and A. Kraay, Growth is Good for the Poor, Journal of economic growth 7, 195-225 (2002).

[8] M.L. Bertotti and G. Modanese, From microscopic taxation and redistribution models to macroscopic income distributions, Physica A 390, 3782-3793 (2011).

[9] M.L. Bertotti and G. Modanese, Exploiting the flexibility of a family of models for taxation and redistribution, Europ. Phys. J. B, 85, 261 (2012).

[10] D.T. Hristopulos and A. Muradova, Kinetic model of mass exchange with dynamic Arrhenius transition rates, Physica A 444, 95-109 (2016).

[11] M.L. Bertotti and G. Modanese, Microscopic models for welfare measures addressing a reduction of economic inequality, Complexity 21, 89-98 (2016).

[12] M.L. Bertotti and G. Modanese, Discretized kinetic theory on scale-free networks, Europ. Phys. J. ST, 225, 1879-1891 (2016).

[13] M.L. Bertotti and G. Modanese, Economic inequality and mobility in kinetic models for social sciences, Europ. Phys. J. ST, 225, 1945-1958 (2016).

[14] M.L. Bertotti, A.K. Chattopadhyay and G. Modanese, Stochastic Effects in a Discretized Kinetic Model of Economic Exchange, Physica A 471, 724-732 (2017).

[15] J.P. Bouchaud and M. Mezard, Wealth condensation in simple model of economy, Physica A 282, 536-545 (2000).

[16] A.K. Chattopadhyay, G.J. Ackland and S. K. Mallick, Income and poverty in a developing economy, Europhys. Letts. 91, 58003 (2010).

[17] M.L. Bertotti, A.K. Chattopadhyay and G. Modanese, Economic inequality and mobility for stochastic models with multiplicative noise, http://lanl.arxiv.org/abs/1702.08391.

[18] E.J. Hinch, Brownian motion with stiff bonds and rigid constraints, J. Fluid Mech. 271, 219-234 (1994).

[19] M. Schwartz and Y. Navot, Stochastic dynamics of constrained systems, Physica A 245, 517-522 (1997).