Wealth distribution and the Lorenz curve: A finitary approach

**Enrico Scalas · Tijana Radivojević ·
Ubaldo Garibaldi**

Abstract We use three stochastic games for the wealth of economic agents which may be at work in a real economy and we derive their statistical equilibrium distributions. Based on a heuristic argument, we assume that the expected observed wealth distribution is a mixture of these three distributions. We compare the Lorenz curves obtained from this conjecture with the empirical curves for a set of countries.
JEL: C80, D31

Enrico Scalas
Department of Mathematics
University of Sussex
Brighton BN1 9QH, UK
and
BCAM - Basque Center for Applied Mathematics
Alameda de Mazarredo 14, 48009 Bilbao, Basque Country, Spain
E-mail: enrico.scalas@unipmn.it

Tijana Radivojević
BCAM - Basque Center for Applied Mathematics
Alameda de Mazarredo 14, 48009 Bilbao, Basque Country, Spain
E-mail: tradivojevic@bcamath.org

Ubaldo Garibaldi
IMEM-CNR, Physics Department
Genoa University
via Dodecaneso 33, 16146 Genoa, Italy
E-mail: garibaldi@fisica.unige.it

1 Introductory considerations

The recent book by T. Picketty [Picketty 2014] started a worldwide debate on wealth inequality and, among other positive effects, it may shift the focus of economic theory towards the study of distributional properties. In order to understand inequality, the Lorenz curve is one of the most used tools [Lorenz 1905]. This curve represents the cumulative fraction of total wealth versus the cumulative fraction of wealth-receiving units. To be more specific, assume we are given a positive (or non-negative) random variable U representing the wealth of an economic agent and m sampled values $\{u_l\}_{l=1}^m$ that we take ordered so that $u_l \leq u_{l+1}$. Then the Lorenz curve is the piecewise continuous curve linearly interpolating the points $\{(\hat{F}_l, \hat{L}_l)\}_{l=0}^m$, where $(\hat{F}_0, \hat{L}_0) = (0,0)$, and for $l = 1, \ldots, m$, $\hat{F}_l$ is the cumulative fraction of wealth-receiving units,

$$\hat{S}_l = \sum_{k=1}^{l} u_k \tag{1}$$

is the cumulative wealth and

$$\hat{L}_l = \hat{S}_l / \hat{S}_m \tag{2}$$

is the corresponding cumulative fraction of total wealth. These equations can be used to determine the empirical Lorenz curve from a dataset of wealths. Let us further assume that the wealth is a multiple of some unit of wealth and that we guess the distribution function $p(j, \phi) = \mathbb{P}(U = j|\phi)$, where $j \in \mathbb{N}$ is the integer numerical value of wealth and ϕ is a vector of unknown parameters. The empirical Lorenz curve, can be compared with the Lorenz curve of the distribution $p(j, \phi)$ that is given by the curve connecting the points (F_k, L_k), $k = 0, \ldots, n+1$, where n is the integer for which $F_{n+1} = 1$ (which is also the total wealth) and

$$F_{k+1} = \mathbb{P}(U \leq k) = \sum_{j=0}^{k} p(j, \phi), \tag{3}$$

$$S_{k+1} = \sum_{j=0}^{k} j p(j, \phi), \tag{4}$$

with $S_{n+1} = \sum_{j=1}^{n} j p(j, \phi)$, and

$$L_{k+1} = S_{k+1}/S_{n+1}, \tag{5}$$

and with $(F_0, L_0) = (0,0)$. Then, one can use his/her favorite statistical method to determine the vector of parameters ϕ which better fits the empirical distribution by minimizing the *distance* between the empirical points $\{(\hat{F}_l, \hat{L}_l)\}_{l=0}^m$ and the model points $\{(F_k, L_k)\}_{k=0}^{n+1}$.

A large part of the literature on fitting the Lorenz curve focuses on looking for suitable *phenomenological* equations that can be used to represent the empirical data in the better available way. Here we quote a selection of the

papers devoted to this problem, the interested reader can use these references as an entry point to the literature. [Kakwani and Podder 1976] suggest a new parameterization of the curve and use four different methods to estimate the Lorenz curve from grouped data. [Sarabia et al. 1999] introduce a general method to create parametric families of Lorenz curves. Hybrid additive and multiplicative models are proposed by [Ogwang and Gouranga Rao 2000]. [Rhode 2009] presents a one-parameter family of curves and [Helene 2010] uses another family of curves, linear in the parameters.

We depart from this tradition by proposing a fit of the Lorenz curve based on a wealth probability distribution justified according to the finitary probabilistic approach introduced in a recent book ([Garibaldi and Scalas 2010]). We consider the wealth composed of n tokens or coins to be distributed to g agents. We introduce three basic stochastic games which could be at work in an economy and, for two of them, we derive their invariant distributions which are also equilibrium distributions. A heuristic argument leads us to conjecture that the expected observed wealth distribution is a mixture of these two distributions and of the distribution of a third, degenerate process, and we compare the Lorenz curves obtained from this conjecture with those from an available database of empirical data (for the database see [Davies et al. 2009]).

2 Finitary models for the wealth distribution

In a recent note [Garibaldi et al. 2013], we proposed a model of the distribution of wealth based on the interplay of three stochastic games. Our approach is inspired by Masanao Aoki's methods (see [Aoki 1996], [Aoki 2004] and [Aoki and Yoshikawa 2011]) and was discussed at length in a book published in 2010 [Garibaldi and Scalas 2010]. These games schematically represent realistic mechanisms acting in the economy, they are:

1. The Bennati-Drăgulescu-Yakovenko (BDY) game [Scalas et al. 2006];
2. the taxation-and-redistribution (TAR) game [Garibaldi et al. 2007];
3. the Zipf-Simon-Yule (ZSY) game [Garibaldi et al. 2006].

The BDY game gives the background noise of the economy. It is a decentralized mechanism in which two agents interact and one of them "steals" part of the wealth of the other. In the TAR game, wealth is removed according to a probability distribution proportional to the wealth of agents and redistributed according to a Pólya transition probability. Finally, the ZSY game includes failures and redistribution of wealth in favor of the Rich.

Let us now define each game. Before that, we recall some basic definitions. In our finitary approach, wealth is represented by "coins" representing wealth tokens that can be exchanged between and among agents. We consider g agents with a total wealth of n coins. The occupation vector $\mathbf{n} = (n_1, \ldots, n_g)$ represents a state of the system where the i-th agent has n_i coins. Given such a state, we can count the number of agents with 0 coins, 1 coin, 2 coins, etc., and we define the *wealth distribution* using the vector $\mathbf{z} = (z_0, \ldots, z_n)$, where

z_j is the number of agents with j coins. The following constraints must be satisfied

$$\sum_{i=1}^{g} n_i = n, \qquad (6)$$

$$\sum_{j=0}^{n} z_j = g \qquad (7)$$

and

$$\sum_{j=0}^{n} j z_j = n. \qquad (8)$$

The BDY game runs as follows. At the beginning of the game, each agent is endowed e.g. with the same number of coins n/g; any other initial state will lead to the same equilibrium distribution. Then a loser is drawn uniformly among all the agents with at least one coin and a winner is drawn uniformly among all agents (the winner and the loser may coincide). The loser gives a coin to the winner and the game is repeated again and again. This game simulates an unfair transaction between two agents, where the winner *steals* part of the wealth of the loser. Those who believe in (neo)classical marginalism will say that such a transaction is not possible or it is out of economic equilibrium. However, given the presence of asymmetric information in economic transactions, the BDY mechanism is not necessarily uncommon. In the BDY game, indebtedness is impossible: an agent without coins cannot be selected as the loser. If $\mathbf{n} = (n_1, \ldots, n_g)$ denotes the agent description before the game step, the occupation vector after the step can be written as $\mathbf{n}_i^j = (n_1, \ldots, n_i - 1, \ldots, n_j + 1, \ldots, n_g)$ (notice that this is a particular case, where $i < j$, but one can also have $i \geq j$). The corresponding transition probability is

$$\mathbb{P}(\mathbf{Y}(t+1) = \mathbf{n}_i^j | \mathbf{Y}(t) = \mathbf{n}) = \frac{1 - \delta_{n_i,0}}{g - z_0(\mathbf{n})} \frac{1}{g}, \qquad (9)$$

where $z_0(\mathbf{n})$ is the number of agents with zero coins before the game step and $\delta_{n_i,0}$ is used, somewhat redundantly, to denote the fact that the loser cannot be selected among the agents with zero coins. The homogeneous Markov chain defined by (9) and by the initial state described above is reversible and the invariant measure can be found by means of detailed balance. It turns out that, the probability of any of the possible states $\pi(\mathbf{n}) = \mathbb{P}(\mathbf{Y} = \mathbf{n})$ is proportional to the number of agents with at least one coin

$$\pi(\mathbf{n}) = C(g - z_0(\mathbf{n})). \qquad (10)$$

The Markov chain defined by (9) is irreducible and aperiodic, therefore, equation (10) also gives the equilibrium probability distribution. In the limit $n \gg g \gg 1$, the marginal expected wealth distribution corresponding to (10) is

approximately described by a geometric distribution of parameter g/n, as discussed in [Scalas et al. 2006] and in [Garibaldi and Scalas 2010]

$$\frac{\mathbb{E}(Z_j)}{g} \simeq \left(1 - \frac{g}{n}\right)^j \frac{g}{n}, \tag{11}$$

given that one has

$$\mathbb{E}(Z_i) = g\mathbb{P}(Y_1 = i) \tag{12}$$

for equidistributed agents.

In the TAR game, a coin is removed from an agent with probability proportional to its wealth and it is given to another agent with probability depending on its wealth and on a redistribution weight α_i. To be more specific, at each step of the game, again one has a transition from $\mathbf{n}$ to $\mathbf{n}_i^j$, but this time the transition probability is

$$\mathbb{P}(\mathbf{Y}(t+1) = \mathbf{n}_i^j | \mathbf{Y}(t) = \mathbf{n}) = \frac{n_i}{n} \frac{\alpha_j + n_j - \delta_{i,j}}{\theta + n - 1}, \tag{13}$$

where $(\alpha_1, \ldots, \alpha_n)$ are the weights for redistribution, and $\theta = \sum_{i=1}^g \alpha_i$. Again, the Markov chain with this transition probability is reversible and we can derive the invariant probability distribution using detailed balance and leading to the Pólya distribution

$$\pi(\mathbf{n}) = \frac{n!}{\theta^{[n]}} \prod_{j=1}^g \frac{\alpha_j^{[n_j]}}{n_j!}, \tag{14}$$

where $x^{[n]}$ denotes the Pochhammer symbol for the rising factorial $x^{[n]} = x(x+1)\cdots(x+n-1)$. If we put all the agents on a par, defining $\alpha = \theta/g$ as the common weight, we can have two situations. If $\alpha > 0$, then richer agents are favored in the redistribution, whereas for $\alpha \to \infty$ and $\theta \to \infty$ in a way such that $g = \theta/\alpha$ remains constant, then all the agents have the same probability of receiving the coin. Finally, one can also consider the case $\alpha < 0$ in which poorer agents receive the coin with higher probability. In this case, the rising factorial has to be replaced by the falling factorial $x_{[n]} = x(x-1)\cdots(x-n+1)$. If we now apply equation (12) to (14), further assuming that all the weights are the same ($\alpha_i = \alpha = \theta/g$), we find the following expected wealth distribution:

$$\frac{\mathbb{E}(Z_j)}{g} = \frac{n!}{\theta^{[n]}} \frac{\alpha^{[j]}(\theta - \alpha)^{[n-j]}}{j!(n-j)!}. \tag{15}$$

In the limit $n \gg g \gg 1$, and with $\theta \gg 1$, equation (15) can be approximated by the following negative binomial distribution, where $\chi = n/\theta = n/(g\alpha)$

$$\frac{\mathbb{E}(Z_j)}{g} \simeq \frac{\alpha^{[j]}}{j!} \left(\frac{1}{1+\chi}\right)^\alpha \left(\frac{\chi}{1+\chi}\right)^j. \tag{16}$$

The ZSY game is introduced as a mechanism to generate heavy tails. As we consider a conservative model of the economy, with g and n constant in

time, we modify the original algorithm proposed in [Garibaldi et al. 2006] as follows. An agent with at least one coin is selected at random and all his coins are removed. This step can be considered as a failure. Then the coins of the failed agent are redistributed to the other agents with probability proportional to their wealth. With this mechanism, agents without coins cannot receive further coins. This is the limit of the Simon-Yule process where the probability of innovation u vanishes. For $u = 0$, one has a degenerate process leading to a state where a single agent owns all the coins after $g - 1$ steps. However, this degenerate process can be used without problems in a conservative system, if only one step is run preceded and followed by many BDY steps, as in the simulation which will be described in the next session. The ZSY redistribution procedure favors the richer agents who get a higher fraction of the coins, but all the agents with at least one coin are on a par with respect to the failure mechanism. In Monte Carlo analyses of this process (see [Garibaldi et al. 2006, Garibaldi and Scalas 2010, Garibaldi et al. 2013]), we found that the expected wealth generated by this process can be approximated by a Yule distribution, in other words we have

$$\frac{\mathbb{E}(Z_j)}{g} \simeq \rho B(j, \rho + 1), \qquad (17)$$

where $B(x, y)$ is the Beta function and $\rho > 0$. Contrary to equations (11) and (16), this result is not exact and the approximation is not controlled. What we can say is that the *Ansatz* given by (17) seems to work.

3 Data analysis

3.1 A conjecture

In [Garibaldi et al. 2013], we simulated the three games in order to reproduce what happens in a real economy, using so-called *supermoves*. On each basic period (a *day*), one BDY step is run. At the end of each *month*, corresponding to l days, we have a ZSY failure and, finally, at the end of each *year*, corresponding to k months a TAR move is run with the mechanism of block taxation (namely with the removal of $m > 1$ coins before redistribution, see [Garibaldi et al. 2007] and [Garibaldi and Scalas 2010]). This simulation is the realization of non-homogeneous Markov chain (however notice that, if the process is only sampled after each super move, the resulting chain is homogeneous). Depending on when we decide to observe the outcome of the simulation, we can observe different distributions. The three moves can be weighted so that they move the same average number of coins. This model will be the subject of a detailed study in a future forthcoming paper. However, we can see that the simulation has three exact limits. If the block taxation moves the total number of coins n, statistical equilibrium is immediately reached after the TAR move and the distribution is given by equation (16); this result is proven in [Garibaldi et al. 2007] as well as in section 8.2 of [Garibaldi and Scalas 2010]. If the number of days in a month is very large (the number of BDY steps is

very large), the equilibrium distribution after each month is given by equation (11). If the number of months is very large with a month having n/g BDY steps as well as a final ZSY failure, the limiting distribution at the end of the year and before taxation is given by the limiting distribution of the combination of a ZSY step followed by n/g BDY which is approximately given by equation (17) with $\rho \simeq 1 + g/n \gtrsim 1$.

Hence, we simply conjecture that the empirically observed expected wealth distribution can be described by a mixture of the three distributions introduced in the previous subsection and given by equations (11), (16), and (17). Therefore, the expected wealth distribution should be given by the following equation

$$p(j, \phi) = \frac{\mathbb{E}(Z_j)}{g} \simeq p_1 \left(1 - \frac{g}{n}\right)^j \frac{g}{n} + p_2 \frac{\alpha^{[j]}}{j!} \left(\frac{1}{1+\chi}\right)^\alpha \left(\frac{\chi}{1+\chi}\right)^j + p_3 \, \rho B(j, \rho+1),$$
(18)

where ϕ is the vector of unknown parameters and $p_1 + p_2 + p_3 = 1$.

3.2 The data set

The data set that we use to fit the Lorenz curve according to equation (18) is taken from [Davies et al. 2009]. In that paper, wealth is defined as the sum of real assets and financial assets minus debt. The data set covers the wealth of family units in Canada for 1999 (8 points), in Switzerland for 1997 (6 points), and in the United States for 2001 (12 points), personal wealth in China for 2002 (10 points), and in France for 1994 (4 points), household wealth in Germany for 1998 (10 points), in India for 2002–2003 (12 points), in Italy for 2000 (6 points), in Japan for 1999 (10 points), and in Spain for 2002 (8 points). The respective Lorenz curves are represented in Figure 1 where the empirical data are plotted as colored stars or circles as described by the legend where Gini indices are also given.

3.3 The fitting procedure

In order to fit the points (F_k, L_k) of the conjectured Lorenz curve based on equation (18) to the real data $(\hat{F}_l, \hat{L}_l)$ we use an ordinary least square procedure with the following objective function

$$d^2(\mathbf{L}(\phi); \hat{\mathbf{L}}) = \sum_{l=0}^{m} (L(\phi)_{k_l} - \hat{L}_l)^2,$$
(19)

to be minimized with respect to ϕ, the vector of parameters of the distribution, namely $\phi = (p_1, p_2, p_3, \lambda, \alpha, \rho)$ with $\lambda = g/n$ and $\alpha = \theta/g$ (remember also that $\chi = 1/(\lambda\alpha)$). The constraints on the weights are $p_1 + p_2 + p_3 = 1$ and $0 \le p_i \le 1$
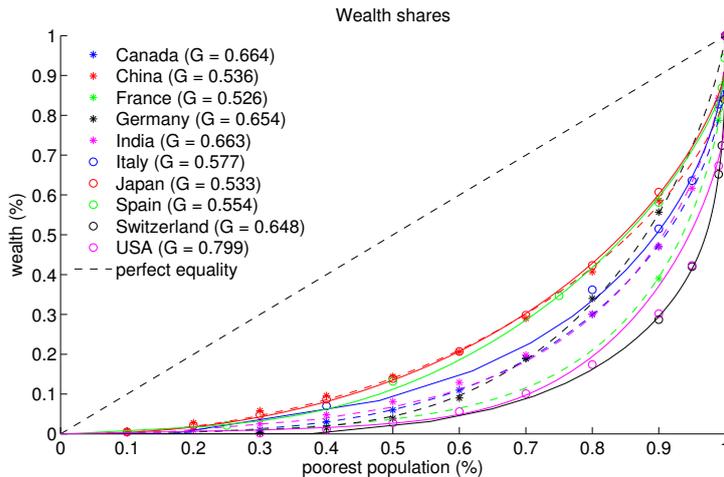
Fig. 1: (Color on-line) Wealth shares by country from the empirical data (colored stars and circles) and Lorenz curves obtained from the model (colored dotted lines and solid lines). Empirical Gini indices are calculated as twice the area between the empirical Lorenz curve (not displayed) and the line of perfect equality.

for $i = 1, 2, 3$. The constraints on the parameters of the distributions are $\lambda > 0$, $\alpha > 0$ and $\rho > 0$. In equation (19), k_l is given by

$$k_l = \arg \min_k |F_k - \hat{F}_l|, \tag{20}$$

and $L(\phi)_{k_l}$ represent the ordinate of the theoretical Lorenz curve. The results of the minimization procedure are given in Table 1. The lines in Figure 1 represent the model Lorenz curves with the parameters given by the fit. An eye inspection of this figure gives a first hint on the goodness of this fit. This is further investigated in Figure 2. In Figures 2 (a) and 2 (b), we plot the cases in which the square distance given by equation (19) is minimal and maximal, respectively. However, one can observe that the model estimate of the Gini index is worse for Switzerland than for Spain. This is due to the fact that the data of the empirical Lorenz curve for Switzerland do not include the points below 90% of the population. This fact is further illustrated in Figures 2 (c) and 2 (d), where the cases with the best and worst model estimate of the Gini index are plotted. Again, in the case of France, the lower part of the empirical Lorenz curve is not there. From Table 1, analyzing d^2, we can say that the fit between the empirical curve and the model curve is satisfactory in all the investigated cases with d^2 in the range $10^{-5} - 10^{-2}$. On the contrary, the model estimate of the Gini index is significantly different from the empirical estimate for France and Switzerland for the reason discussed above.

Given that we do not have full information on the distribution of wealth, with the previous procedure, we are looking for 5 parameters in order to fit a

Country	p_1	p_2	λ	α	ρ	d^2	G_{data}	G_{model}
Canada	0.54	0.28	0.070	0.256	0.841	$1.650 \cdot 10^{-4}$	0.664	0.690
China	0.45	0.48	0.018	1.658	0.327	$5.603 \cdot 10^{-4}$	0.536	0.563
France	0.008	0.57	0.049	0.465	0.938	$1.120 \cdot 10^{-4}$	0.526	0.755
Germany	0.35	0.30	0.022	0.875	1.934	$2.594 \cdot 10^{-4}$	0.654	0.663
India	0.40	0.22	0.106	0.292	1.236	$8.901 \cdot 10^{-4}$	0.663	0.678
Italy	0.44	0.20	0.225	0.446	1.334	$3.539 \cdot 10^{-4}$	0.577	0.639
Japan	0.66	0.30	0.016	1.211	0.322	$1.264 \cdot 10^{-4}$	0.533	0.554
Spain	0.06	0.55	0.071	2.748	1.050	$20.30 \cdot 10^{-4}$	0.554	0.575
Switzerland	0.40	0.38	0.154	0.070	0.943	$3.132 \cdot 10^{-5}$	0.648	0.819
USA	0.16	0.51	0.003	0.768	0.669	$6.200 \cdot 10^{-3}$	0.799	0.813

Table 1: Values of the model parameters (p_1, p_2, λ, α and ρ) obtained in the OLS fitting procedure, together with the square distance d^2 between the empirical data and the model from equation (19), and Gini indices G calculated both from the data and from the model as twice the area between the respective Lorenz curves and the line of perfect equality. Note that $p_3 = 1 - p_1 - p_2$.

Lorenz curve with a small number of points. In the case of France, the number of points is just 4 and it is even smaller than the number of parameters. There is a risk of over fitting. In order to investigate this issue, we can repeat the minimisation procedure setting the parameter ρ of the Yule distribution at its theoretical value $\rho = 1 + \lambda$. The results of this new fit are presented in Table 2. Now, this second model is a restriction of the previous one. Then one can use the F statistic defined as $F = (m - 5)(d_4^2 - d_5^2)/d_5^2$ in order to assess if the fit with 5 parameters is significantly better than the fit with 4 parameters, where d_5^2 is the square distance with 5 parameters, d_4^2 is the square distance with 4 parameters and m is the number of data points [Box 1953]. This statistic has the Fisher-Snedecor F-distribution of parameters $(1, m - 5)$. Table 3 contains the value of the statistic obtained from the data and the critical values at three significance levels. The null hypothesis that the model with 5 parameters does not provide a significantly better fit than model with 4 parameters is rejected if the statistic is larger than the critical value. For most of the countries, it turns out that the 5-parameter model is not significantly better than the 4-parameter model. Before concluding this section, we would like to stress that, in principle, this model has only three free parameters. In fact if the size of the total wealth n as well as the number of households or individuals g are known, λ must be fixed at the value g/n. Therefore, the only free parameters are p_1, p_2 and α. Unfortunately, in this study, we do not have these pieces of information.

4 Discussion

In the previous section, we have shown that the *Ansatz* given by equation (18) is able to satisfactorily reproduce the empirical Lorenz curve for the available dataset. Unfortunately, the empirical wealth distribution is not available in the data set. Fitting this distribution is a more demanding test for a model.

Country	p_1	p_2	λ	α	d^2	G_{data}	G_{model}
Canada	0.28	0.14	0.002	0.598	$5.0 \cdot 10^{-3}$	0.664	0.708
China	0.23	0.18	0.003	1.360	$5.7 \cdot 10^{-3}$	0.536	0.538
France	0.22	0.23	0.006	0.632	$2.0 \cdot 10^{-3}$	0.526	0.841
Germany	0.52	0.32	0.004	0.728	$2.5 \cdot 10^{-3}$	0.654	0.695
India	0.32	0.21	0.0005	0.519	$2.2 \cdot 10^{-2}$	0.663	0.704
Italy	0.10	0.45	0.006	0.860	$6.1 \cdot 10^{-3}$	0.577	0.737
Japan	0.41	0.29	0.002	0.983	$2.6 \cdot 10^{-3}$	0.533	0.556
Spain	0.13	0.10	0.004	1.105	$4.8 \cdot 10^{-3}$	0.554	0.611
Switzerland	0.31	0.28	0.010	0.568	$3.2 \cdot 10^{-2}$	0.648	0.905
USA	0.09	0.008	0.008	1.177	$2.0 \cdot 10^{-2}$	0.799	0.774

Table 2: Values of the model parameters (p_1, p_2, λ and α) obtained in the OLS fitting procedure, together with the square distance d^2 between the empirical data and the model from equation (19), and Gini indices G calculated both from the data and from the model as twice the area between the respective Lorenz curves and the line of perfect equality. Note that $p_3 = 1 - p_1 - p_2$.

Country	F_{data}	F_c (5%)	F_c (1%)	F_c (10%)
Canada	9.8	10.1	34.1	*5.5*
China	1.8	6.6	16.3	4.1
Germany	1.7	6.6	16.3	4.1
India	3.4	5.6	12.2	3.6
Italy	16.2	161.4	4052.2	39.9
Japan	3.9	6.6	16.3	4.1
Spain	0.5	10.1	34.2	5.5
Switzerland	1041	*161.4*	4052.1	*39.9*
USA	0.3	5.6	12.2	3.6

Table 3: $\hat{F}$ statistics, calculated as $\frac{(d_4^2 - d_4^2)}{d_5^2/(m-5)}$, are compared with critical value of F-distribution with $(1, m-5)$ degrees of freedom for significance levels of 5%, 1% and 10%. For France, it is not possible to calculate the critical values. The null hypothesis (that the model with 5 parameters does not provide a significantly better fit than model with 4 parameters) is rejected in the cases reported in italics.

In the insets of Figure 2, we present double logarithmic plots of the wealth distributions obtained from the fits according to our *Ansatz*. They have the desirable property that their tails follow a power law.

An important feature of the wealth distribution is the presence of indebtedness and negative wealth. Indeed, given that in [Davies et al. 2009], the wealth of an agent is defined as real assets plus financial assets minus debt, one can have individuals, families or households with negative wealth. In order to take properly into account such cases, one must use more elaborate agent-based models, where every agent is characterized by a balance sheet and the time-evolution is stock-flow consistent as in [Raberto et al. 2012].

Our future work on this subject will focus on a detailed study of the simulation described in Section 3 and initially studied in [Garibaldi et al. 2013].
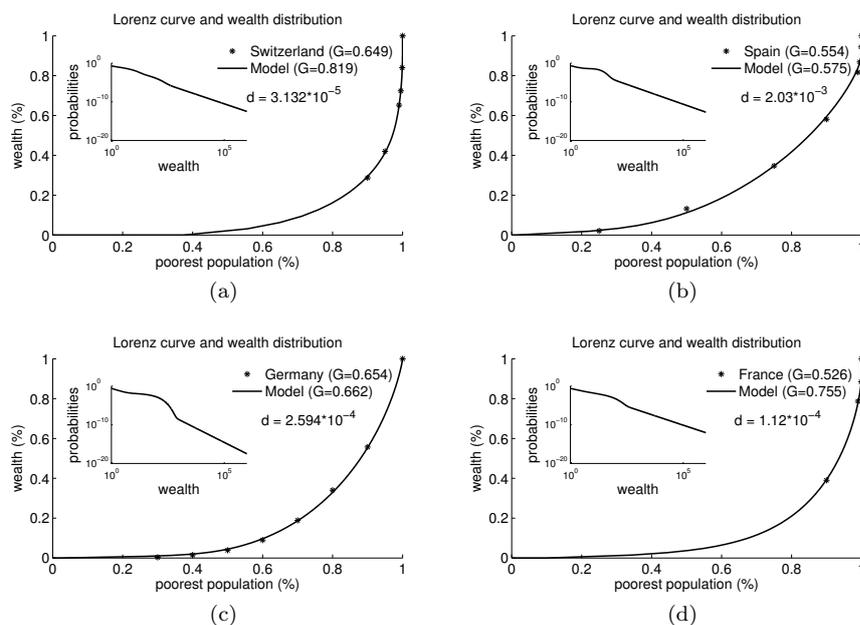
Fig. 2: Lorenz curve fitting: the empirical data are represented by stars, the solid lines are the model fits and the insets contain the corresponding wealth probability distributions from the model. The best fit in terms of distance d is obtained for Switzerland (a) and the worst for Spain (b). In the terms of Gini index (G) the best fit is obtained for Germany (c) and the worst for France (d).

References

[Aoki 1996] Aoki M (1996) New Approaches to Macroeconomic Modeling: Evolutionary Stochastic Dynamics, Multiple Equilibria, and Externalities as Field Effects, Cambridge University Press, Cambridge UK.

[Aoki 2004] Aoki M (2004) Modeling Aggregate Behavior and Fluctuations in Economics: Stochastic Views of Interacting Agents, Cambridge University Press, Cambridge UK.

[Aoki and Yoshikawa 2011] Aoki M, Yoshikawa H (2011) Reconstructing Macroeconomics: A Perspective from Statistical Physics and Combinatorial Stochastic Processes, Cambridge University Press, Cambridge UK.

[Box 1953] Box G E P (1953) Non-Normality and Tests on Variances, Biometrika, **40**(3/4):318–335.

[Davies et al. 2009] Davies J B, Sandström S, Shorrocks A, Wolff E N (2009) The world distribution of household wealth in Davies J B (Ed.) (2009) Personal Wealth from a Global Perspective (Wider Studies in Development Economics), Oxford University Press, Oxford UK.

[Kakwani and Podder 1976] Kakwani N C, Podder N (1976) Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations, Econometrica, **44**(1): 137–148.

[Garibaldi et al. 2006] Garibaldi U, Costantini D, Donadio S, Viarengo P (2006) Herding and clustering in economics: The Yule-Zipf-Simon model, Computational Economics **27**(1): 115–134.

[Garibaldi et al. 2007] Garibaldi U, Scalas E, Viarengo P (2007) Statistical equilibrium in simple exchange games II - The redistribution game, The European Physical Journal B, **60**: 241–246.

[Garibaldi and Scalas 2010] Garibaldi U, Scalas E (2010) Finitary Probabilistic Methods in Econophysics, Cambridge University Press, Cambridge UK.

[Garibaldi et al. 2013] Garibaldi U, Radivojević T, Scalas E (2013) Interplay of simple stochastic games as models for the economy, Working Paper available at `http://www.bcamath.org/documentos_public/archivos/publicaciones/GRS.pdf`.

[Helene 2010] Helene O (2010) Fitting Lorenz curves, Economic Letters **108**: 153–155.

[Lorenz 1905] Lorenz M O (1905) Methods of measuring the concentration of wealth, Publications of the American Statistical Association **9** (70): 209–219.

[Ogwang and Gouranga Rao 2000] Ogwang T, Gouranga Rao U L (2000) Hybrid models of the Lorenz curve, Economic Letters **69**: 39–44.

[Picketty 2014] Picketty T (2014) Capital in the XXI Century, Harvard University Press, Cambridge MA.

[Raberto et al. 2012] Raberto M, Teglio A, Cincotti S (2012) Debt, deleveraging and business cycles: An agent-based perspective, Economics - The Open-Access, Open-Assessment E-Journal **6**(27): 1–49.

[Rhode 2009] Rhode N (2009) An alternative functional form for estimating the Lorenz curve, Economics Letters **105**: 61–63.

[Sarabia et al. 1999] Sarabia J-M, Castillo E, Slottje D J (1999) An ordered family of Lorenz curves, Journal of Econometrics **91**: 43–60.

[Scalas et al. 2006] Scalas E, Garibaldi U, Donadio S (2006) Statistical equilibrium in simple exchange games I - methods of solution and applications to the Bennati-Dragulescu-Yakovenko (BDY) game, The European Physical Journal B, **53**: 267–272.