Central European Economic Journal

sciendo

ISSN: 2543-6821 (online)

Journal homepage: http://ceej.wne.uw.edu.pl

Yanling Guo, Friedrich L. Sell

Equilibrium and Convergence in Income Distribution: The Case of 28 European Countries in the Recent, Turbulent Past (1995–2019)

To cite this article

Guo, Y., Sell, F.L. (2021). Equilibrium and Convergence in Income Distribution: The Case of 28 European Countries in the Recent, Turbulent Past (1995–2019). Central European Economic Journal, 8(55), 95-125.

DOI: 10.2478/ceej-2021-0007

To link to this article: https://doi.org/10.2478/ceej-2021-0007

Central European
Economic Journal

Yanling Guo, Friedrich L. Sell

Universität der Bundeswehr München/Bundeswehr University Munich, Institut für Ökonomie und Recht der globalen Wirtschaft, Werner-Heisenberg-Weg 39, 85577 Neubiberg, Germany, corresponding author: friedrich.sell@unibw.de

Equilibrium and Convergence in Income Distribution: The Case of 28 European Countries in the Recent, Turbulent Past (1995–2019)

Abstract

The authors developed a political economy equilibrium framework for personal income distribution. In the beginning, they set up a theoretical model which was rooted in status theory. With this concept, one may explain a certain or optimal degree of inequality in society and define a steady state to which inequality can converge. By taking the aggregated Gini coefficient due to a collective decision process, deviations from the steady state due to shocks are allowed. A return to equilibrium is feasible with speed which is compatible with the collective decisionmaking process. The authors then conducted an empirical analysis of personal income distribution in 28 European nations for the period before, during and after the great recession of 2009/2010 and the Euro crisis of 2010/2015 (1995–2019). Not surprisingly, they found inequality convergence in the data. However, the speed of convergence is not the same for all countries.

Keywords

inequity convergence | equilibrium | personal income distribution | redistributive policies | collective decision making

JEL Codes

D31, D63, D72

1 Introduction

The central hypothesis which underlies this article is the concept of equilibrium in personal income distribution. The famous economist Vilfredo Pareto (Pareto, 1895) is already known to have believed that the social distribution of personal incomes moves towards a stable equilibrium over time. He based his statement on the observation that personal income dispersion fluctuates neither internationally nor inter-temporally. Much later, Hans Jürgen Ramser (Ramser, 1987) identified stationarity in the secondary distribution of personal incomes (i.e. the income net after government intervention with taxes and transfers) and not in the primary distribution of incomes (out of the market process). Recent studies support this empirical finding (Genc, Miller & Rupasingha, 2011).

The existing skewness of (personal) income distribution may be interpreted as a display of social preferences, thereby implying that preserving a specific degree of income inequality is intentional (Blümle, 1992, p. 224) and not arbitrary. Although distributional justice continues to be a fundamental goal of economic policies, it does not focus strictly on achieving a perfect equitable income distribution. Both short- and long-term scenarios accept the unwarranted existence of a certain degree of inequitable income distribution in society (Blümle, 1992, p. 225). In reality, such equilibrium would seldom be achieved to a full extent, although the policymakers have good reasons to push towards the 'steady state' and thereby help reach convergence.

When it comes to the question of breaking down this idea to a possible empirical analysis, the world seems to be more complex: Will countries be moving always towards a similar equilibrium level of inequality? This question is all but trivial. Works by Esping-Andersen (1990, 1994, 1998), by Esping-Andersen and Myles (2009) and by Hall and Soskice

(2001) suggest something different and highlight the fact that among developed countries there exist long-run differences in the institutional structure and in so-called 'institutional complementarities in the macroeconomy' (Hall & Gingerich, 2004), which can also be conceived as different equilibria in income distribution. These different equilibria may be attributed not only to the differences in the institutional set-up, but also to the observed variety of historical experiences in the respective countries. Taking into account the role of business in national economies and the fact that there is more than one path to economic success ('liberal market economies' vs. 'coordinated market economies') explain the further differences. However, different historical experiences and/or unalike institutional structures become much less important, once countries share a longer period of a common economic, social and political history, as is the case for the member countries in the EU. This insight holds even if countries realise to a different and necessarily often changing extent what Esping-Andersen calls the 'three worlds of welfare capitalism' (liberal, conservative, social-democratic systems; see Tiemann, 2006 for a critical empirical evaluation of this non-undisputed concept). Furthermore, it is understood that the institutional complementarity and implicit coordination between a 'supra-national' European Central Bank (ECB) on the one hand, and national institutions, such as national unions, pursuing a 'national' wage policy on the other hand, is a common challenge for members of the Eurozone.

The article is organised as follows: after a brief review of relevant literature, we present our equilibrium model in personal income distribution. An exhaustive empirical part—considering a period of extreme economic turbulence—follows, whose findings support the concept of steady state and convergence in personal income distribution. The article concludes with a discussion of the theoretical and empirical results of the present study, the limitations of the analysis and the possibilities for future research endeavours.

2 A Brief Review of the Relevant Literature

In principle, there exist two strands of the literature which are relevant to the subject under consideration in the present study. While one of the two perspectives focuses on the relationship between an *economic crisis* and personal income distribution, the other questions the existence of *convergence* in personal income distribution. However, to date, no study is available in the literature which connects these two perspectives, as the present study does. In both of these research perspectives, it is essential to (1) address the ex-post Gini coefficient or (2) at least compare the Gini coefficient on market income over time with the former.

In a pre-crisis study (Van Kerm & Alperin, 2013), the authors reported that the arrangement of the countries of the world in descending order of annual income inequality for the period of 2003–2007 puts "Portugal and Baltic states (such as Estonia, as reported by the authors) at the top, and most Scandinavian countries (such as Finland and Denmark, as reported by the authors) at the bottom" (p. 937). This result was overly unspecific. The research papers published by Dolls, Fuest and Peichl (2011), De Beer (2012) and Kaitila (2013) concerned the issue of economic crisis and income distribution in Europe. Dolls et al. (2011) conducted two controlled experiments (simulations) of macro shocks to income and employment and observed that "both shocks lead to higher differences between the Gini coefficients based on equivalent disposable and market income" (ibid., p. 240). This effect applied to all the 19 European countries considered by the authors in their study. De Beer (2012), in our view, utilised too short a period (2008–2009) to conclude that "the economic crisis has not so far led to a general widening of income disparities and a rise in poverty" (ibid, p. 23).

The second group of contributions has a considerable tradition and follows the seminal paper of Ravallion (2003): as a long-time member of World Bank research groups, his variety of natural interest concentrated on developing countries' fate. However, his 2003 paper also delivered important methodological aspects for measuring (conditional and unconditional) convergence in personal income distribution. The critical hypothesis that Ravallion tested is whether the trend in inequality depends on its initial level (ibid., p. 352). He finds inequality convergence, "with a tendency for within-country inequality to fall (rise) in countries with initially high (low) inequality" (ibid., p. 355).

A specific example of follow-up investigations is the contribution of Alfani and Ryckbosch (2015). In their long-run historical perspective (1500–1800), they compared changes in inequality between central and northern Italy on the one hand and the southern and northern Low Countries on the other hand. Similarly, Martinez-Carrion and Maria-Dolores (2017) explored inequality and regional convergence in Spain and Italy for the long period of 1850–2000. Lessmann and Seidel (2015) investigated regional inequality (based on Gini indexes and other measuring instruments) and convergence based on much more recent satellite night-time light data for a vast sample of developed and developing countries. Sell (2015, p. 15–20) also analysed both developed and developing countries and obtained as a result that "globalisation and possibly other forces linked to the revolution in communication and information technologies have contributed to an almost worldwide convergence in the distribution of personal incomes. More precisely, one can say that developing (developed) countries' income distribution has become more equal (unequal)" (ibid, p. 16).

A smaller section of this body of literature deals directly with the convergence issue applied to overall and/or parts of Europe (for example, the EU and/or the Eurozone) and the recent past. For example, Paas and Schlitte (2007) analysed a cross-section of 861 EU-25 regions from 1995 to 2003 to detect between- and within-country disparities in income distribution—measured by the Theil index. Melchior (2008) studied regional inequality and convergence in Europe (1995–2005), that is, for a period which preceded exactly the financial and world economic crisis of 2008/2009. She found that for the EU-27 as a whole, there was a modest increase in within-country regional inequality, but convergence across countries (ibid, p. 31).

In 2013, Kaitila, in turn, reported a result that was considered close to the findings of the present study, although a slightly different approach (involving the calculation of the sigma convergence) was followed in that study. According to Kaitila's findings, "For the EU–15 (a little less for the EU–27, the authors), we found that the national Gini coefficients have converged considerably during these (1995–2011, the authors) years" (ibid., p. 14). A most recent contribution to the EU inequality subject stems from Savoia (2019): his sample covers the years 1989–2013 (i.e. he necessarily misses part of the Eurozone crisis episode) and countries from the so-called 'NUTS 2' regions. He finds a clear tendency that supports the results of Kaitila: "inequality is converging, but to a higher level" (ibid., p. 29).

A minor methodological remark: It is surprising to notice that regressions between GDP per capita and Gini coefficients of disposable income are seldom run for the enormous projects concerning European inequalities, and the focus of these projects remains restricted to the role of redistributive policies (Medgyesi & Toth, 2009, p. 135; Paulus, Figari & Sutherland, 2009, p. 154). An obvious step towards convergence remained undetected between 2000 and 2005, a period when the EU-27 was without Cyprus, Malta, Slovakia and Luxembourg. As stated by a previous study, "the level of inequality at the beginning of the period does not seem to influence the direction and the magnitude of the change in inequality" (Medgyesi & Toth, 2009, p. 140). This finding, however, appears flawed. The discrepancy is attributed to the addition of countries to the EU, which turned EU-15 into EU-25 in 2004 and into EU-27 later.

Therefore, the question of how economic crisis, equilibrium and convergence in personal income distribution could be addressed scientifically in a comprehensive approach remains demonstrated so far. We describe this task in the following sections.

3 Introduction of a Theoretical Framework for Equilibrium in the Income Distribution

It is surprising to see that irrespective of the definition of income, the economy in question or the time period under consideration, the distribution of incomes is skewed positively (i.e. skewed towards the right-hand or steep on the left-hand).

This fact has significant consequences for the parameters of the density function, for which the maximum value, which is referred to as the modus (y_{mo}) and is the most frequent event, would usually be located to the left of the median (y_{me}), and the latter, in turn, is located to the left of the arithmetic mean (y_{ar}). The characteristics of this kind of density function we depict in Figure 1.

The consequences of this are extensive. According to Blümle (2005), most economic agents would receive an above-modus income. Based on this observation, the agents would have the impression of being well-paid, and therefore their attitude towards a redistribution (the existing distribution) of incomes should be quite critical (benevolent). The density function depicted in Figure 1 may be approximated, rather accurately, using a log-normal distribution of incomes, which is represented as follows:

$Y = \exp(X)$ with $X = N(\mu, \sigma^2)$

The expected or similar average wage rate is then obtained using the following expression (Beichelt & Montgomery, 2003, pp. 46–8):

$$E(y) = y_a = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

Taking the full differential of the above-mentioned expression from left to right yields the following expression:

$$dE(y) = dy_a = (d\mu + \sigma d\sigma)\exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

Proposition 1: *An increase in σ would shift the arithmetic mean to the right.*

Furthermore, we consider the following expression:

$$y_{mo} = \exp(\mu - \sigma^2)$$

Taking the full differential of the above-mentioned expression yields the following expression:

$$dy_{mo} = (d\mu - 2\sigma d\sigma)\exp(\mu - \sigma^2)$$

Proposition 2: *An increase in σ would shift the modus to the left.*

Finally, we obtain the following expression:

$$y_{me} = \exp(\mu)$$

$$dy_{me} = \exp(\mu)$$

Proposition 3: *An increase in σ would not affect the median. This holds true for any $\sigma^2 > 0 : y_{mo} < y_{me} < y_{ar}$.*

In Figure 2, an increase in the standard deviation, σ, has the effect derived in Propositions 1–3: the modus is shifted to the left and the arithmetic mean is shifted to the right while the position of the median remains unchanged.

It is now assumed that an increase in inequality or a higher concentration of incomes is perceived by an individual i as a loss of utility. The utility function of the individual i then reads as follows:

$$U_i = U_i(y_i - y_{mo}; \sigma)$$

where $\frac{\partial U_i}{\partial y_{mo}} < 0; \frac{\partial U_i}{\partial \sigma} < 0.$

Assuming the law of diminishing increases of damage, one obtains:

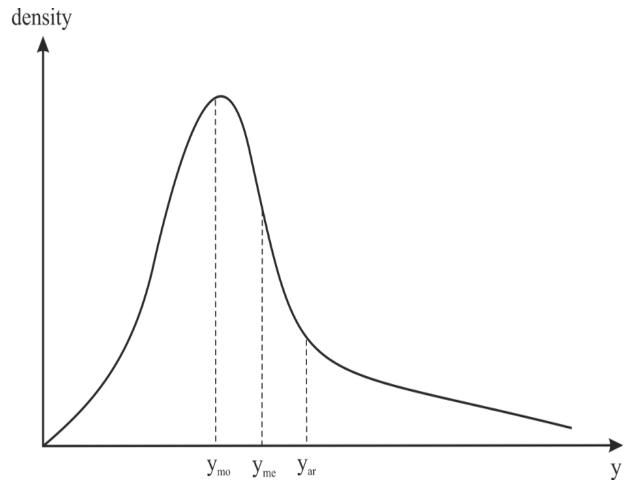

Fig. 1. The time-invariant distribution pattern of personal incomes. *Source*: Blümle (2005)

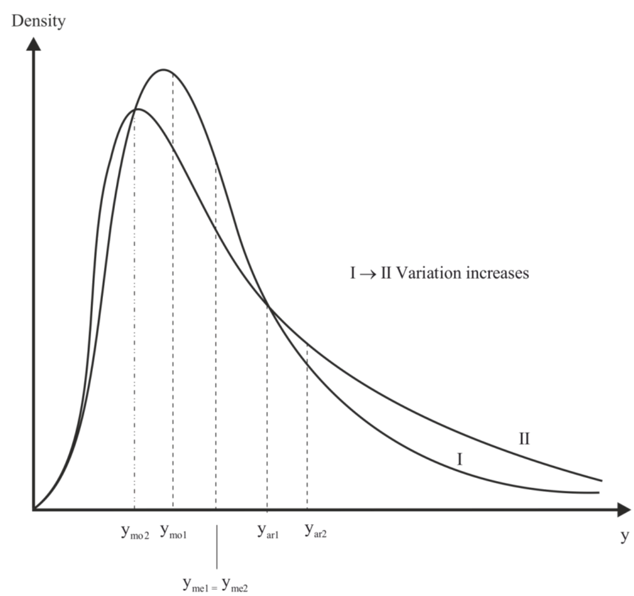

Fig. 2. Increasing the standard deviation in the distribution of personal incomes. *Source*: Sell (2015)

$$\frac{\partial^2 U_i}{\partial y_{mo}^2} > 0; \frac{\partial^2 U_i}{\partial \sigma^2} > 0.$$

Therefore, the corresponding iso-damage curves are concave. It is essential to apply a kind of budget constraint to determine an optimal solution. Such a budget constraint may be found in the properties of the log-normal distribution. Its properties reveal that an increasing dispersion of incomes does not alter the median of the distribution (Proposition 3), reducing and shifting the modus to the left (Proposition 2). These findings also indicate a likely increase in households' share with an income above the (new) modus whenever the concentration of incomes, as measured by the standard deviation σ, increases.

The following expression was used to determine the mathematical solution for the Cobb–Douglas utility function which is presented in Figure 3:

$$U_i = \sqrt{y_i - y_{mo}}\sqrt{1/\sigma}$$

In the diagram depicted in Figure 3, one may identify and locate the equilibrium in the personal income distribution. The modus (y_{mo}) and the dispersion of incomes (σ) are allocated along the axes. The non-linear budget constraint, representing the log-normal distribution of incomes, has been labelled VV. This schedule is confronted with a troop of iso-damage curves (I_i), which are concave towards the origin of the coordinate system. The farther these curves are located from the origin, the higher is the individuals' loss of utility.

Point P signals towards a situation where a preferably low iso-damage curve is tangential to VV. In a sense, P represents an equilibrium in the income distribution. Note that point P stands for what Chiang (1984, p. 231–232) labels a 'goal equilibrium'. The "equilibrium state is defined as the optimum position for a given economic unit (a household, a business firm, or even an entire economy) and in which the said economic unit will be deliberately striving for the attainment of that equilibrium" (ibid, p. 232). A 'nongoal equilibrium', on the contrary, "dictates an equilibrium state ... in which ... opposing forces (demand and supply, for example, the authors) are just balanced against each other, thus obviating any further tendency to change". (ibid, p. 231).

In comparison, the points Q and R represent the suboptimal solutions. Although Q and R fulfil the

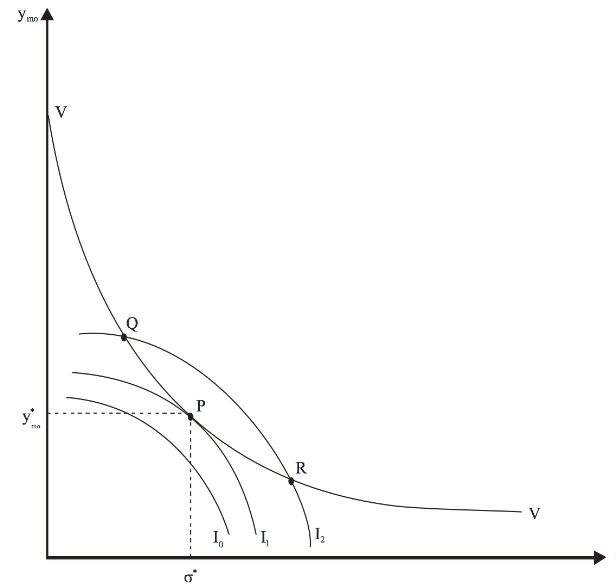

Fig. 3. Equilibrium in the distribution pattern of personal incomes. *Source*: Sell (2015)

'budget constraint' of the log-normal distribution, they are located on the less favourable iso-damage curve I_2. The equilibrium level of inequality for the individual and the society, economy and political system as a whole is now also being derived formally. As after learning the optimal solution's intuition through the graphical analysis, the study has to proceed to a straightforward mathematical solution of what has been described ahead.

The individual *i* receives his own income, y_i, and the distributional parameters μ and y_{mo} are provided. Although the distributional parameter σ is also provided to him, with the knowledge that this parameter of income dispersion could be altered through a re-distributional policy by the government, he may deliberate on which σ would be the best for him. He would accordingly suggest his personal desired level of income dispersion, σ^*_i, to the decision makers, thereby attempting to influence the ultimately determined level of income dispersion, σ*, in his favour. Of course, hardly anyone can calculate the exact value of his σ^*_i, and in the usual policy-making process, the individuals cannot choose the final decision in a very exact manner. However, suppose that everyone knows by and large what is best for him, and the final decision can roughly reflect the collective decisions, then our theoretical result is still a good approximation, around which the real decisions would centre.

First, the personal desired level of income dispersion, $\sigma^*_{i,}$ is considered. The maximisation problem may be expressed as follows:

$$\max U_i s.t. y_{mo} = \exp\left(\mu - \sigma^2\right)$$

Inserting the expression for U_i in the above-mentioned equation, one obtains the following expression:

$$\max \sqrt{y_i - y_{mo}} \sqrt{1/\sigma} s.t. y_{mo} = \exp\left(\mu - \sigma^2\right)$$

The Lagrangian for this maximisation problem is:

$$L_i = U_i - \lambda\left(y_{mo} - \exp\left(\mu - \sigma^2\right)\right) =$$
$$= \sqrt{y_i - y_{mo}} \sqrt{1/\sigma} - \lambda\left(y_{mo} - \exp\left(\mu - \sigma^2\right)\right)$$

Taking the first-order condition (FOC) yields the following expression:

$$\frac{\partial L_i}{\partial \sigma} = -\frac{\sqrt{y_i - y_{mo}}}{2} \sigma^{-\frac{3}{2}} - 2\lambda\sigma \exp(\mu - \sigma^2) = 0 \quad (1)$$

$$\frac{\partial L_i}{\partial y_{mo}} = -\frac{\sqrt{1/\sigma}}{2}(y_i - y_{mo})^{-\frac{1}{2}} - \lambda = 0 \Rightarrow$$
$$\Rightarrow \lambda = -\frac{\sqrt{1/\sigma}}{2}(y_i - y_{mo})^{-\frac{1}{2}} \quad (2)$$

Inserting Eq. (2) in Eq. (1) yields the following expression:

$$(y_i - y_{mo})^{-\frac{1}{2}} \exp(\mu - \sigma^2)\sqrt{\sigma} - \frac{\sqrt{y_i - y_{mo}}}{2} \sigma^{-\frac{3}{2}} = 0$$

Substituting y_{mo} with $\exp\left(\mu - \sigma^2\right)$ yields the following expression:

$$(y_i - \exp(\mu - \sigma^2))^{-\frac{1}{2}} \exp(\mu - \sigma^2)\sqrt{\sigma} -$$
$$- \frac{\sqrt{y_i - \exp(\mu - \sigma^2)}}{2} \sigma^{-\frac{3}{2}} = 0$$

Multiplying both the sides with $(y_i - \exp(\mu - \sigma^2))^{\frac{1}{2}} \sigma^{\frac{3}{2}}$ yields the following expression:

$$\exp(\mu - \sigma^2)\sigma^2 - \frac{y_i - \exp(\mu - \sigma^2)}{2} = 0 \Rightarrow \sigma^*_i = f(y_i, \mu)$$

Therefore, the desired level of income dispersion by the individual i, represented by σ^*_i, is determined uniquely by his income, y_i, and the distributional parameter, μ. The ultimately determined overall level of income dispersion, σ^*, depends on all the individually desired levels, σ^*_i, and on the exact policy decision-making process. In a process in which all the individual desires are assigned same weightage, the ultimately determined overall level of income dispersion, σ^*, is the parameter which maximises social welfare in the form of aggregated utilities. Therefore, one obtains the following optimisation problem:

$$\max \sum_{i=1}^{n} U_i \text{ s.t. } y_{mo} = \exp(\mu - \sigma^2) \text{ and } \bar{y}_a = \exp(\mu + \frac{1}{2}\sigma^2)$$

In the aggregated utility function, n denotes the number of individuals in the society. In addition to the known constraint $y_{mo} = \exp(\mu - \sigma^2)$, one obtains another constraint $\bar{y}_a = \exp(\mu + \frac{1}{2}\sigma^2)$, in which $\bar{y}_a$ denotes the value of average income, which is constant because it was assumed that the redistribution does not change the total income, and therefore, the average income, which is expressed as follows:

$$\sum_{i=1}^{n} U_i = \sum_{i=1}^{n} \sqrt{y_i - y_{mo}} \sqrt{1/\sigma} = \sqrt{1/\sigma} \sum_{i=1}^{n} \sqrt{y_i - y_{mo}}$$

When the utility function is represented by a first-order Taylor series approximation around y_a, one obtains the following expression:

$$\sum_{i=1}^{n} U_i = \sqrt{1/\sigma} \sum_{i=1}^{n} (\sqrt{\bar{y}_a - y_{mo}} + \frac{y_i - \bar{y}_a}{2\sqrt{\bar{y}_a - y_{mo}}}) =$$
$$= n\sqrt{1/\sigma} \sqrt{\bar{y}_a - y_{mo}}$$

The Lagrangian for the maximisation problem now becomes the following:

$$L = n\sqrt{1/\sigma} \sqrt{\bar{y}_a - y_{mo}} - \lambda_1(y_{mo} - \exp(\mu - \sigma^2)) -$$
$$- \lambda_2(\bar{y}_a - \exp(\mu + \frac{1}{2}\sigma^2))$$

Taking the FOC yields the following:

$$\frac{\partial L}{\partial \sigma} = -\frac{n}{2}\sigma^{-\frac{3}{2}}\sqrt{\bar{y}_a - y_{mo}} - 2\sigma\lambda_1 \exp(\mu - \sigma^2) +$$
$$+\sigma\lambda_2 \exp(\mu + \frac{1}{2}\sigma^2) = 0 \quad (3)$$

$$\frac{\partial L}{\partial y_{mo}} = -\frac{n}{2}\frac{\sqrt{1/\sigma}}{\sqrt{\bar{y}_a - y_{mo}}} - \lambda_1 = 0 \quad (4)$$

$$\frac{\partial L}{\partial \mu} = \lambda_1 \exp(\mu - \sigma^2) + \lambda_2 \exp(\mu + \frac{1}{2}\sigma^2) = 0 \quad (5)$$

Substituting $\sigma\lambda_2 \exp(\mu + \frac{1}{2}\sigma^2) = -\sigma\lambda_1 \exp(\mu - \sigma^2)$ from Eq. (5) in Eq. (3) yields the following:

$$\frac{\partial L}{\partial \sigma} = -\frac{n}{2}\sigma^{-\frac{3}{2}}\sqrt{\bar{y}_a - y_{mo}} - 3\sigma\lambda_1 \exp(\mu - \sigma^2) = 0 \quad (6)$$

Substituting $\lambda_1 = -\frac{n}{2}\frac{\sqrt{1/\sigma}}{\sqrt{\bar{y}_a - y_{mo}}}$ from Eq. (4) in Eq. (6) yields the following:

$$-\frac{n}{2}\sigma^{-\frac{3}{2}}\sqrt{\bar{y}_a - y_{mo}} + 3\sigma\frac{n}{2}\frac{\sqrt{1/\sigma}}{\sqrt{\bar{y}_a - y_{mo}}}\exp(\mu - \sigma^2) =$$
$$= -\frac{n}{2}\sigma^{-\frac{3}{2}}\sqrt{\bar{y}_a - y_{mo}} + 3\frac{n}{2}\frac{\sqrt{\sigma}}{\sqrt{\bar{y}_a - y_{mo}}}\exp(\mu - \sigma^2) = 0$$

Dividing both sides by $n/2$ and then multiplying both the sides with $\sigma^{\frac{3}{2}}\sqrt{\bar{y}_a - y_{mo}}$ yields $-(\bar{y}_a - y_{mo}) + 3\sigma^2 \exp(\mu - \sigma^2) = 0$, which is equivalent to $-(\bar{y}_a - y_{mo}) + 3\sigma^2 y_{mo} = 0$, because $y_{mo} = \exp(\mu - \sigma^2)$.

Dividing both the sides by y_{mo}, one obtains the following expression:

$$-(\bar{y}_a / y_{mo} - 1) + 3\sigma^2 = 0$$

One should notice that both the constraints imply $y_{mo} = \bar{y}_a \exp(-\frac{3}{2}\sigma^2)$ or $\bar{y}_a / y_{mo} = \exp(\frac{3}{2}\sigma^2)$, and therefore, the above-mentioned equation becomes $-(\exp(\frac{3}{2}\sigma^2) - 1) + 3\sigma^2 = 0$, which determines σ^* to be equal to approximately 0.9. (The exact value can be produced by Mathematica using the command Solve [-(Exp[3/2*s^2]-1)+3*s^2==0,s,Reals].)

Of course, external shocks will destroy any earlier equilibrium and lead to deviations from the former equilibrium level of inequality. The question as to which income groups will favour a subsequent redistribution policy depends on the external shock's nature. If the shock tends to an increase (decrease) in the concentration of personal incomes, modern theory of inequity (equity) aversion (Bolton & Ockenfels, 2000; Sell, 2015) would suggest that the group of inequity (equity) -averse agents will push government to progressive (regressive) redistribution policies. It is important to note that our model, now as it stands, captures only the preference of inequity aversion.

4 Empirical Findings

4.1 Some methodological issues ex-ante

The theoretical section of this article describes the development of a model which demonstrated that the (collective) choice of the preferred variance of income distribution results from a trade-off between the preference for one's above-modus income, which is expressed as the first root term in the utility function, and the preference for a low concentration of the income distribution in the society, which is expressed as the second root term in the utility function. A testable hypothesis derived from the model was that there exists a long-term equilibrium value of the income distribution variance, towards which the society converges.

However, to conduct the empirical test, two issues had to be resolved first. The first issue was that the empirical data often contain Gini coefficients instead of the variance of the income distribution. The second issue was that it sounds counter-intuitive that the societies at different time points in history and those at different developmental stages should all converge to 'the' optimal variance of the income distribution, which would be clear later, is also not supported by the data.

The first issue could be conveniently resolved because of the assumed log-normal distribution of the incomes, which encompasses several real-world features (including a left-steep/right-skewed income distribution with a modus lower than the median and a median lower than the mean), thereby serving as a good approximation of the real-world data. Since the log-normal distribution serves as a fair approximation of the empirical distribution of incomes, it is a well-known fact that the Gini coefficient is a monotonically increasing function of the variance or its root, the standard deviation σ. More precisely, $Gini = 2\Phi\left(\sigma / \sqrt{2}\right) - 1$ with $\Phi(x)$ being the standard normal distribution. Therefore, the convergence to a particular value of variance is equivalent to the convergence towards a particular Gini coefficient of the income distribution.

The second issue could also be resolved conveniently by introducing a preference parameter into the developed basis model. The utility function would then become:

$$U_i = \sqrt{y_i - y_{mo}}\left(\sqrt{1/\sigma}\right)^{\gamma}$$

The introduced parameter γ stands for the weightage which the individuals of a society assign to the preference for an equal-income distribution (or against a higher σ), relative to the preference for a larger above-modus own income. The larger the value of γ, the lower would be the equilibrium standard deviation σ^* (see Appendix A1). Although it is possible that in addition to varying across regions, σ^* also changes over time due to shifts in preferences, it should remain relatively stable, such that an equilibrium state of σ^* would nonetheless be observed in the absence of the underlying force driving the change.

To test our model, we need more restrictive assumptions about γ. Given that it represents the relative preference of more equal income over a higher own income, or solidarity over competition, we postulate that γ does not only change slowly over time, but that it also converges to similar values for countries which are closely linked each other. Thus, the Gini steady-state value in an integration process would not diverge but rather converge to similar values. This fact makes our model prediction similar to the neoclassical growth model (NGM), which predicts a convergence of Gini coefficients of various countries linked via free trade. However, the difference between our model and the NGM is that we are explicitly dealing with after-tax income distribution. Thus, the parameter of preference γ plays an essential role through the collective decision process. Furthermore, our model is about steady-state values of Gini, not Gini per se. Thus, we do not strictly assume that Gini's change would be (negatively) proportional to its initial value. The following empirical test shows that our model better fits the data. When modelling Gini's change, we also assume that a negative relationship exists between the initial value and the change of *Gini*. However, we also allow different countries to have a different speed of change; this makes our model similar to club convergence theory (CCT), which focuses on country-specific factors. The difference between our model and CCT is that we have more strict assumptions about *Gini*; namely, *Gini* instead converges among countries linked to each other. Altogether, this makes our model empirically testable.

4.2 Data presentation and descriptive analysis of *Gini* data

European data on income distribution stem from "EU Statistics on Income and Living Conditions (EU-SILC) (2015).

The *Gini* data in particular are from Eurostat (Gini coefficients of equalised disposable income—EU-SILC survey). We have analysed the Gini coefficients from the 27 EU member countries plus UK, with annual data from 1995 to 2019. Due to missing data, the panel is unbalanced. All figures and tables of this section can be found in the Appendix.

Figure 4 plots the Gini coefficients against the calendar years; each of the 28 countries is on a separate line. Missing data are plotted as a dotted line in between using linear interpolation. The year 2008 is marked with a grey vertical line. Figure 5 rearranges the data by plotting the Gini coefficients against membership years in the EU (negative numbers correspond to years before joining the EU) instead of plotting against the calendar years. The year of joining the EU is plotted as a grey vertical line through the zero point. The year 2008 now appears as a short grey vertical line through the corresponding data points. The Gini coefficients of the countries, as one can see in Figure 5, converge (are getting closer) to each other, the longer the respective EU membership.

This convergence looks further clearer when we plot the Gini coefficients minus the group mean against membership years in the EU in Figure 6. The group mean is computed as the mean value of all EU members in each year, namely, Belgium, Denmark, Germany, Finland, France, Greece, Ireland, Italy, Luxembourg, the Netherlands, Austria, Portugal, Spain, Sweden and the United Kingdom from the beginning on, then plus Estonia, Latvia, Lithuania, Malta, Poland, Slovakia, Slovenia, Czechia, Hungary and Cyprus from 2004 onwards. Bulgaria and Romania joined in 2007 and Croatia in 2013. Because the dataset ended in 2019, the UK has always been counted as an EU member.

Whether this convergence is due to the EU membership, as a first glance on the graphs would suggest, or rather due to the worldwide observed convergence in Gini coefficients as documented in Ravallion 2001, Ravallion 2003 and Savoia 2019, is unclear. To distinguish the two types of convergence, we also plotted Gini coefficients minus the overall mean, namely, the mean value of all 28 countries each year, in Figure 7. Here, the same degree of convergence can be detected. Thus, the observed convergence may also be part of a worldwide inequality convergence process, derived from the NGM, as pointed out in Ravallion 2003. Figure 7 is almost identical to Figure 6. The convergence process is not very well visible in Figure 4 because the old member countries in our data are from the beginning (in 1995) near the overall mean; thus, they remain hidden among the data points and do not stand out. When plotted against years of membership in the EU instead of calendar years in Figures 5–7, the old member countries are put on the right side while the new member countries are on the left side; then the convergence becomes more visible. That the old member countries are from the beginning onwards closer to the overall mean may be because they have been more extended members in the EU, but it could also be because they have been longer integrated into the world market economy.

To investigate why Figure 7 is almost identical to Figure 6, which share the same *x*-axis, but differ in the *y*-axis (one represents difference to the overall mean; meanwhile the other represents the difference to the EU-wide mean), we also plotted the overall mean and the EU-mean Gini coefficients for each year in Figure 8. Figure 8 shows that the two means are very close to each other and that the difference is at most about 0.3. That the two means are almost identical in each year has both technical and non-technical reasons. Technically, from 2013 onwards, all 28 countries were members in the EU; thus, the two means have to be identical after 2013. In the first two years, data from all new member countries are missing; therefore, the two means are also identical in our sample. However, from 1997 to 2012, some non-technical reasons explain the closeness of the two means. The reason, we speculate, is that the EU has been from the beginning onwards integrated into the world economy. Thus, the development of Gini coefficients within the EU has not been substantially different from the development of Gini coefficients outside the EU, nor does joining the EU significantly change the trajectory of each member country's *Gini*. Whatever the reason may be, our sample data cannot distinguish between convergence within EU and convergence as part of the worldwide story. Given that the convergence phenomenon—in general motivated by the NGM—has been more extensively studied in the literature, we opt to pursue this to increase the comparability of our contribution with the relevant papers in the literature.

In Figures 5–7, we can also see that the Gini coefficients of the old member countries are closer to the overall mean, without having the same value, and that they show less convergence towards the overall mean. This may be because they converge more slowly, and they do this because they are closer to the steady state as predicted by the NGM (which predicts the same long-run steady state for all countries). It may also well be that the countries implied have slightly different steady states and are pending around them as pointed out by CCT. CCT explains that some long-lasting differences among countries are due to some hard-to-change underlying country-specific institutional factors. Hence, countries that are similar concerning important institutional factors converge better. Since the steady states are quite close, and the change is small, one cannot determine which theory is approximately more correct. What we can observe, however, is that if there were some hard to change differences among the EU countries, then they are relatively small, and being in the EU did not force the long-term members to all converge to the same *Gini* value. Whether the latter is due to slightly different steady states or to steadily occurring small asymmetric shocks—causing deviation from 'the' steady state—remains unknown.

4.3 In-depth analysis of Gini coefficients

The NGM predicts a negative relationship between the change in and Gini coefficients' initial value as documented in Ravallion (2001), Ravallion (2003) and Savoia (2019). Our data mostly confirm the prediction of the convergence theory, as shown in Figure 9. Since our data set is unbalanced, i.e. not every country has the same t_0 and t_T, we obtain the average annual change $\left(Gini_T - Gini_0\right)/\left(t_T - t_0\right)$ instead of the absolute change $Gini_T - Gini_0$, considering Gini coefficients of each country from year t_0 to t_T. These are plotted against the initial Gini coefficients, $Gini_0$, in Figure 9 (a very similar plotting can be obtained when we use the time coefficient β in the equation $Gini = \alpha + \beta t$ instead of $\left(Gini_T - Gini_0\right)/\left(t_T - t_0\right)$ for each country). Eurostat has changed the source/methodology several times in the data pool. When considering these changes by adding a dummy for each change, Figure 9 remains mostly unchanged. Savoia (2019) has studied this negative relationship for varying samples out of the data set. The missing data problem seemingly naturally divides our data set into several subsamples for each country; therefore, we also plotted the data for these subsamples in Figure 10. The negative relationship remains, though R^2 necessarily decreases due to the existence of more noise in the data.

It would be interesting to see if this negative relationship remains if we break the data into as many subsamples as possible, namely, when each subsample only encompasses two years. Then, we are able to plot the annual change of the Gini coefficient, $\Delta Gini_t := Gini_t - Gini_{t-1}$, against the Gini coefficient in the previous year, $Gini_{t-1}$. This is what we did in Figure 11. Here, we also took the methodological changes mentioned above into account and we only computed $\Delta Gini_t$ when both $Gini_t$ and $Gini_{t-1}$ were collected using the same source/methodology and when no data points were missing. Although much more noise is now included, the negative trend remains intact, which again confirms NGM. Of course, this result is also perfectly compatible with our model, which predicts non-divergence for interlinked countries.

The convergence hypothesis, widely represented in the literature, is usually expressed as a negative relationship between the change in the respective Gini coefficient and its initial value. The higher the initial value, the lower the expected change (the Gini coefficient increases slower or it decreases faster). Alternatively, one could formulate this negative relationship as follows: the higher the initial value above the overall mean, the lower the change. Graphically spoken, one would plot $\Delta Gini_t^i$ against $Gini_{t-1}^i - \sum_{j=1}^{N} Gini_{t-1}^j$, instead of against $Gini_{t-1}^i$, with i and j being the respective country indexes, and N being the number of EU country members. The results, which can be requested from the authors, remain relatively similar, with almost the same slope and R^2. Only the intercept, and of course also the notation of the x-axis, do change. Given that the conventional formulation is more straightforward and more widely recognised, we opt to no longer employ our alternative formulation.

Convergence in the Gini coefficients can also be explained by the fixed effects (FE) approach or by country dummies, an idea similar to CCT. When $\Delta Gini_t$ is the dependent variable, the country dummies capture the country-specific effects, explaining possible differences in average annual change of Gini coefficients between countries. Figure 12 shows the coefficients for the country dummies arranged on a line in the bottom part. In the unconditional regression A(2) $Gini = \alpha C(country)$, with $C(country)$ denoting the 28 used country dummies, most coefficients not being significantly different from zero and R^2 being slightly higher at a level of 0.048. However, the adjusted R^2 is negative due to the usage of 28 independent variables. This result is in line with our model explaining convergence towards one's own steady state, where we based our findings on an aggregated optimisation behaviour. In the steady-state neighborhood, the change of the Gini coefficient should be insignificantly different from zero. The very low R^2 is alarming, which points to the fact that our data are most of the time and also in the majority of cases not in a steady state. Hence, our steady-state model can only explain a relatively small part of the variation in the Gini coefficient. Nevertheless, the same is true for the NGM captured in Figures 9–11; with an ever shorter period used of the subsamples, R^2 decreases from 0.260 to 0.030 due to the existence of more noise, though the coefficient is always significantly negative as predicted and the adjusted R^2 is only slightly less than R^2.

To see how similar two parameters are to each other, one could directly look at the parameters and ignore their standard errors, as we did in the bottom part of Figure 12 for the country dummies. Alternatively, one can look at their pairwise t-statistics, $\left(\alpha^i - \alpha^j\right) / SE\left(\alpha^i - \alpha^j\right)$, namely the difference between two dummy coefficients divided by its standard error. It is challenging to plot the pairwise t-statistics as 28 points, though, because the standard errors are not all the same and therefore $\left(\alpha^i - \alpha^j\right) / SE\left(\alpha^i - \alpha^j\right)$ cannot be computed as sum of $\left(\alpha^i - \alpha^k\right) / SE\left(\alpha^i - \alpha^k\right)$ and $\left(\alpha^k - \alpha^j\right) / SE\left(\alpha^k - \alpha^j\right)$ with k being the index for the third country. Consequently, one needs a space of up to 27 dimensions to represent 28 points fully. We made use of a machine learning technique called multidimensional scaling (MDS) algorithm to reduce the dimensions while preserving the distance information, here in form of the absolute value of pairwise t-statistics, as good as possible, and plot the 28 points on a line in the upper part of Figure 12. The closer the two points are, the smaller is the absolute value of their pairwise t-statistic and the more unlikely it is that their coefficients are different from each other. The presentation in the upper part, according to pairwise t-statistics, is not as precise as the presentation of the dummy coefficients in the bottom part of Figure 12. This fact is due to information lost in the dimension reduction, but it corresponds better to the concept of similarity, here defined as how likely it is that two coefficients cannot be distinguished from each other. This finding adds more information to the graphic. Ultimately, the two presentations look quite similar to each other.

It is interesting to see what happens when regressing $\Delta Gini_t$ both on the previous Gini coefficient and on the country dummies. Doing so increases R^2 to 0.173 and the adjusted R^2 to 0.122. The large increase in R^2 and the significance of all estimates suggest that our empirical model which combines the convergence prediction of NGM with FE/CCT is a viable strategy. The coefficient reflecting the impact of the previous *Gini* is, furthermore, negative and significant. The country dummies have now all become significant. Of course, the interpretation is somewhat different. The country dummies now represent country-specific effects, except those that can be somewhat explained by their previous *Gini* values. If we plot the countries according to their dummies, then the figure looks slightly different from before, as shown in Figure 13. That the pairwise *t*-statistic for model A(3) can have an absolute value of up to 6.72 is impressive. The NGM, as far as we know, does not predict significant between-country differences in the change of *Gini* value except those that can be explained by the country-specific initial state of *Gini* value. This assumption would be very strict and is not supported by the estimates from model A(3), where some country dummies are significantly different from each other. The CCT, which allows for between-country differences in the change of *Gini* value, would be preferably supported by the estimation result of A(3). However, the country dummies only capture factors unexplained by the initial state of *Gini* value of each subsample encompassing only two consecutive periods, namely $Gini_{t-1}$, and do not precisely correspond to the hard-to-change factors such as institutions. By modelling the aggregated *Gini* value as resulting from a collective decision, we allow *Gini* to deviate from the steady-state value. When shocks occur, a return to equilibrium is feasible at speed that is compatible with the collective decision-making process, which may not be the same for all countries. Hence, the model is compatible with the data.

Our parameter γ does not affect $\Delta Gini_t$. Hence, $\Delta Gini_t$ should be zero at the steady state while Gini is some positive number as determined by γ . $\Delta Gini_t$ is in model A(3) and treated as also depending on $Gini_{t-1}$ to describe the off-equilibrium data points better. The steady-state value of $Gini_{t-1}$ is the same as the steady-state value for $\Delta Gini_t$, which is around 30, as can be seen in Figure 5; therefore, the predicted $\Delta Gini_t$ around the steady-state value for each country is the respective country dummy minus roughly 0.249 × 30 and thus lies between about –1.5 and 1.5, divided by now higher standard errors. Thus, many computed predicted changes in the Gini coefficients around the steady states would remain statistically insignificant from zero. These results are in line with our model. Similar points can be made for model A(4), which implicitly assumes that the financial crisis preceding the great recession of 2009/2010 and the Euro crisis of 2010/2015 permanently raises the annual change in the Gini coefficient until the end of the sample period, namely 2019, or even further, by adding a dummy for being from 2008 on. The financial crisis dummy coefficient is positive and significant; the model fit increases slightly, and other coefficients do not change much. We summarise the regression results in Table 1.

The respective standard errors are given in parentheses.

Finally, we present the ADF test results for the Gini coefficients in Table 2. Since the ADF test we use cannot deal with missing values, we only use the most extended recent time series without any missing value for each country. According to Lin and Huang (2012), the NGM also implies stationarity in time series of inequality measures such as the Gini coefficient. Findings based on our data set weakly support this view. While 13 out of 28 countries are stationary according to at least one of the three alternative measures, namely, AIC, BIC and t-stat, for 9 countries the unit root hypothesis cannot be ruled out at the 10% confidence level, with any measure. Similar results (with US data) can be found in Lin and Huang (2012); they showed that the unit root hypothesis cannot be ruled out for several states if omitting structural breaks and then they also showed that by taking into account up to two structural breaks, almost all time series are stationary. Although we have a structural break candidate with the year 2008, not do we did the structural break test because the start year of our time series is in the range of 2003–2010, which makes the test less meaningful, and because our own model does not require empirical stationarity; although our model does not suggest unit root, it does not rule out that *Gini* may follow an AR(1) process with a coefficient close to 1, either. Thus, unit root cannot be ruled out empirically.

4.4 Inclusion of GDP data into the analysis

According to Lessmann and Seidel (2015), regional inequality depends on the respective per capita GDP, y, and follows a N-shape, namely, the coefficient for

Tab. 1. Dependent variable $\Delta Gini_t$

Model	A(1)	A(2)	A(3)	A(4)
Previous Gini	−0.053*** (0.013)		−0.249*** (0.030)	−0.265*** (0.030)
Financial crisis				0.321*** (0.110)
Austria		−0.210 (0.255)	6.417*** (0.837)	6.676*** (0.835)
Belgium		−0.171 (0.255)	6.588*** (0.852)	6.855*** (0.850)
Bulgaria		0.707** (0.312)	9.187*** (1.067)	9.481*** (1.063)
Croatia		−0.267 (0.389)	7.329*** (0.989)	7.497*** (0.982)
Cyprus		0.246 (0.324)	7.888*** (0.973)	8.109*** (0.968)
Czechia		−0.143 (0.312)	6.075*** (0.807)	6.201*** (0.802)
Denmark		0.294 (0.283)	6.528*** (0.800)	6.741*** (0.796)
Estonia		−0.663** (0.292)	7.619*** (1.039)	7.932*** (1.036)
Finland		0.085 (0.261)	6.342*** (0.796)	6.553*** (0.793)
France		−0.205 (0.268)	7.022*** (0.910)	7.318*** (0.908)
Germany		−0.020 (0.261)	7.017*** (0.886)	7.278*** (0.883)
Greece		−0.259 (0.249)	8.180*** (1.048)	8.548*** (1.047)
Hungary		−0.100 (0.292)	6.679*** (0.865)	6.875*** (0.860)
Ireland		−0.271 (0.255)	7.475*** (0.968)	7.806*** (0.966)
Italy		−0.175 (0.261)	7.775*** (0.993)	8.110*** (0.991)
Latvia		−0.071 (0.312)	8.871*** (1.121)	9.172*** (1.117)
Lithuania		−0.060 (0.302)	8.694*** (1.097)	9.002*** (1.093)
Luxembourg		0.076 (0.255)	7.023*** (0.874)	7.302*** (0.872)
Malta		0.071 (0.312)	6.972*** (0.885)	7.142*** (0.879)
Netherlands		−0.115 (0.261)	6.558*** (0.844)	6.811*** (0.841)
Poland		−0.473 (0.302)	7.269*** (0.979)	7.511*** (0.974)
Portugal		−0.281 (0.255)	8.541*** (1.094)	8.926*** (1.093)
Romania		−0.179 (0.312)	8.331*** (1.071)	8.605*** (1.066)
Slovakia		−0.408 (0.324)	5.829*** (0.813)	5.960*** (0.808)
Slovenia		0.006 (0.292)	5.880*** (0.761)	6.017*** (0.757)
Spain		0.025 (0.261)	8.351*** (1.037)	8.710*** (1.036)
Sweden		0.369 (0.292)	6.662*** (0.809)	6.847*** (0.805)
United Kingdom		0.118 (0.283)	8.179*** (1.011)	8.528*** (1.010)
Constant	1.482*** (0.389)			
Observations	**479**	**479**	**479**	**479**
***R*-squared**	**0.033**	**0.048**	**0.173**	**0.189**

$^{***}P < 0.01$, $^{**}P < 0.05$, $^{*}P < 0.1$.

Tab. 2. ADF test result

Country	Stationary (AIC)	Stationary (BIC)	Stationary (t-stat)	Trend sta. (AIC)	Trend sta. (BIC)	Trend sta. (*t*-stat)	Nobs
Austria	**	**	**	**	**	**	17
Belgium		**					17
Bulgaria							14
Croatia						***	10
Cyprus						*	15
Czechia							15
Denmark	***	***	***				17
Estonia	***	***	*				16
Finland							16
France	**	**	**				15
Germany	**	**		***	***	***	15
Greece							17
Hungary		**	*	**	**	**	15
Ireland				***	***	***	16
Italy				***	***	***	15
Latvia	**	**	**	***	***	***	15
Lithuania							15
Luxembourg				*	*		17
Malta	*	*	*	*	*	*	15
Netherlands	***	***	***	***	***	***	15
Poland	**	**	**				15
Portugal				**	**	**	16
Romania		**	**				13
Slovakia							14
Slovenia							15
Spain							16
Sweden	***	***	***	**	**	**	16
United Kingdom							14

$^{***}P < 0.01$, $^{**}P < 0.05$, $^{*}P < 0.1$.

Tab. 3. Dependent variable $Gini_t$

Model	B(1)	B(2)	B(3)	B(4)
logGDPpc	210.401** (101.18)			463.717*** (146.76)
(logGDPpc)²	−22.158** (10.25)			−46.512*** (15.17)
(logGDPpc)³	0.765** (0.35)			1.541*** (0.52)
previous *Gini*		0.950*** (0.01)		0.351*** (0.05)
Austria			−253.595** (104.02)	−1753.78*** (482.91)
Belgium			272.008*** (104.02)	−1437.44*** (479.71)
Bulgaria			−1451.324***(108.02)	−1902.24*** (385.60)
Croatia			500.411* (283.47)	−1291.01** (557.56)
Cyprus			−377.443*** (145.40)	−1706.16*** (489.32)
Czechia			171.240 (145.40)	−1537.05*** (512.88)
Denmark			−535.120*** (118.87)	−1892.70*** (480.87)
Estonia			323.865*** (100.53)	−1448.45*** (508.94)
Finland			123.248 (100.53)	−1460.13*** (480.53)
France			−265.087** (109.12)	−1737.01*** (486.63)
Germany			−387.554*** (117.02)	−1825.68*** (491.62)
Greece			87.160 (104.02)	−1331.06*** (460.23)
Hungary			−332.505*** (104.51)	−1832.71*** (500.85)
Ireland			155.019 (113.32)	−1520.09*** (458.52)
Italy			−291.261** (120.55)	−1686.74*** (480.19)
Latvia			459.139*** (145.40)	−1280.34** (507.37)
Lithuania			−487.852*** (117.02)	−1868.91*** (510.14)
Luxembourg			−488.867*** (104.02)	−1838.25*** (454.26)
Malta			−181.419 (145.40)	−1882.67*** (541.84)
Netherlands			58.677 (104.51)	−1571.00*** (476.53)
Poland			421.547*** (117.02)	−1298.42*** (501.97)
Portugal			686.918*** (110.50)	−1085.10** (483.61)
Romania			−15.311 (183.48)	−1518.38*** (488.56)
Slovakia			596.187*** (162.52)	−1212.39** (526.13)
Slovenia			−199.579* (104.51)	−1764.69*** (504.61)
Spain			−240.605** (100.53)	−1695.74*** (486.51)
Sweden			−578.102*** (113.37)	−1972.66*** (481.94)
United Kingdom			324.630*** (113.35)	−1358.44*** (484.78)

Continued **Tab. 3**. Dependent variable $Gini_t$

Model	B(1)	B(2)	B(3)	B(4)
Austria*TIME			0.140*** (0.05)	0.122** (0.05)
Belgium*TIME			−0.122** (0.05)	−0.035 (0.05)
Bulgaria*TIME			0.739*** (0.05)	0.198* (0.11)
Croatia*TIME			−0.233* (0.14)	−0.109 (0.13)
Cyprus*TIME			0.203*** (0.07)	0.099 (0.07)
Czechia*TIME			−0.073 (0.07)	0.012 (0.07)
Denmark*TIME			0.279*** (0.06)	0.192 (0.06)
Estonia*TIME			−0.145*** (0.05)	−0.030 (0.06)
Finland*TIME			−0.048 (0.05)	−0.024 (0.05)
France*TIME			0.146*** (0.05)	0.115** (0.05)
Germany*TIME			0.207*** (0.06)	0.159*** (0.06)
Greece*TIME			−0.027 (0.05)	−0.087* (0.05)
Hungary*TIME			0.179*** (0.05)	0.160*** (0.05)
Ireland*TIME			−0.062 (0.06)	0.008 (0.07)
Italy*TIME			0.161*** (0.06)	0.090* (0.06)
Latvia*TIME			−0.210*** (0.07)	−0.112 (0.07)
Lithuania*TIME			0.260*** (0.06)	0.180** (0.08)
Luxembourg*TIME			0.257*** (0.05)	0.165*** (0.05)
Malta*TIME			0.104 (0.07)	0.185** (0.08)
Netherlands*TIME			−0.016 (0.05)	0.032 (0.05)
Poland*TIME			−0.194*** (0.06)	−0.105** (0.07)
Portugal*TIME			−0.324*** (0.06)	−0.209*** (0.05)
Romania*TIME			0.025 (0.09)	0.006 (0.10)
Slovakia*TIME			−0.284*** (0.08)	−0.149* (0.08)
Slovenia*TIME			0.111** (0.05)	0.125** (0.05)
Spain*TIME			0.136*** (0.05)	0.095** (0.05)
Sweden*TIME			0.300*** (0.06)	0.231*** (0.06)
UK*TIME			−0.145*** (0.06)	−0.073 (0.05)
constant	−624.115* (331.96)	1.450*** (0.43)		
Observations	**427**	**427**	**427**	**427**
R-squared	**0.185**	**0.913**	**0.933**	**0.946**

***$P < 0.01$, **$P < 0.05$, *$P < 0.1$.

y is positive, it is negative for y^2 and again positive for y^3. For our data of equalised disposable income *Gini*, we got similar results. Both in model B(1), where only functions of GDP per capita and a constant are independent variables, and in model B(4), where the country dummies and their interaction with time are added, we can confirm the N-shape found in Lessmann and Seidel (2015). This fact is somewhat surprising given that the real GDP per capita may further increase in the future without an upper bound while *Gini* is upper-bounded. Our real per capita GDP data stem from Eurostat. They cover 20 years, from 2000 to 2019, and contain only very few elements of missing data for the 28 countries. The data in Lessmann and Seidel (2015) are from 2001 to 2010 and worldwide. Our confirmative result may have something to do with the overlapping 10 years (2001–2010) and with the fact that EU countries economically make up a substantial part of the world economy. On the other hand, as a distributional measure of incomes, the Gini coefficient may well depend on income, and the variation of income should be a good candidate for explaining especially off-equilibrium Gini coefficients.

In the previous subsection, we found that β in $\Delta Gini_t^i = c + \beta Gini_{t-1}^i$ is negative, as predicted by NGM. This implies that in $Gini_t^i = c + \beta Gini_{t-1}^i$ should be between 0 and 1, which is confirmed by estimation model B(2). Moreover, the 95% confidence interval for β in B(2) is [0.922, 0.978] and the R^2 is 0.913, which is quite useful for over 400 observations, as summarised in Table 3.

The respective standard errors are given in parentheses.

In estimation model B(3), we regress $Gini_t^i$ on country dummies and their interaction term with time, which is similar to running a regression $Gini_t = c + \beta t$ for each country separately.

The country dummies' estimates are not very relevant (currently, they stand for the predicted *Gini* in the calendar year 0, if the linear trend lasts that long). The estimates for their interaction term with the time give a vital hint: for if the *Gini* is around its steady state, a significant value points to a time-changing *Gini*. We get more significant values for the estimated parameter dummy*TIME in B(3) than for the country dummies in A(2), possibly because the sample here is shorter of time periods. The GDP data are from 2000 onwards; then, in order to preserve some data for the lag of *Gini*, the starting year of the sample is 2001. Combining the estimates of the dummies and their interaction terms, we could calculate a predicted value of *Gini* either for the first year or for the last year, according to CCT. The prediction for UK 2019, which is missing in the sample, is, for instance, 324.630 + 2019x (−0.145) = 31.875, which is quite close to, for example, the 29.700 of Germany in this year, despite their difference of 10 at the start of the sample in 2001 (25.000 for Germany vs. 35.000 for UK).

The R^2 for model B(3) is 0.933, also quite good for over 400 observations. In B(4), we included all the independent variables considered so far in this subsection: functions of real per capita GDP, the lag of Gini, the country dummies and their interaction terms with time. The coefficients for the first to the third power of log of GDP p.c. all have the correct sign and are significant. The coefficients for the lag of Gini coefficient are also significant and lie in the interval (0, 1), as predicted. The coefficients for the interaction terms are less significant than their pendant in B(3), which suggests that the off-equilibrium *Gini* values are mostly due to off-equilibrium real per capita GDPs and off-equilibrium past *Gini* values. Hence, the annual change of *Gini* would be in the majority of cases insignificantly different from zero when these variables are controlled for. These findings are compatible with our own model and also with NGM and CCT.

Taking all together, the data support our model quite well. Our regression results are also supportive for the convergence theorem derived from the NGM, though our less restrictive assumptions in line with the CCT can increase the model fit significantly. Our findings also suggest that the CCT can fit our data only if the predicted differences between countries are not too high (the right-hand part in Figure 5 suggests a roughly ± 5 band somewhere around somewhat less than 30 for the EU countries' *Gini* values.). The achieved results based on the used data sample finally confirm the N-shape of income (real per capita GDP in log term) determined by *Gini*.

The empirical model of an N-shaped income-determined Gini coefficient is interesting for our model because income was included in our model but then cancelled out after the optimisation and aggregation. This cancelling-out effect is intended because, as mentioned earlier, the per capita GDP, or its log, is not bounded, while *Gini* is bounded (between 0 and 100 in our sample). By ruling out income as a determinant for *Gini*'s steady state through model construction, we eliminate a potential plausibility problem. However, the empirical result does suggest that income, or some function of it, may play an essential role in deviation from and return to the steady-state process, which we have not explicitly modelled. For the future, we may develop an even more comprehensive model, which focuses on the determination of the steady state and concerns itself with the deviation from the steady state due to shocks and then possibly describes the back to the steady-state converging process.

5 Conclusions and the Scope for Future Research

Convergence in inequality is a big issue in empirically oriented research on personal income distribution. The majority of contributions in our empirical part test the NGM/convergence theory and CCT, as well as our approach. While the first suggests that market forces (international and intranational trade, factor movements, etc.) push the tendency towards convergence, the latter (and theoretical strands close to it; see the introduction) explains why primarily institutional factors—similar within clubs, but not similar outside—tend to put forward some sort of 'conditional convergence' in income inequality. Our approach is close to CCT but stresses much more the political decision process as a determinant for the steady state in income distribution.

Our empirical research tends to support our hypothesis, but the findings can be differentiated as follows:

Before controlling for the impact of the financial crisis (and its aftermath) and possible assimilation—a sort of 'institutional assimilation process' is presumably significant among members of the EU—of the Gini coefficients, an insignificantly-different-from-zero change in the Gini coefficient was obtained for the majority of countries. *After* controlling for the assimilation and optionally for the financial crisis effect, all countries show a significant deviation from our simple modelling, ignoring country-specific effects. This deviation, however, is quite similar among the countries under study. Only for a small number of countries can one observe statistically significant differences. Therefore, we have not only found convergence in inequality—despite the significant economic turbulence during the period of observation (1995–2019)—but we have also found that there are non-negligible country-specific effects that can contribute to improving the fit of the model.

That convergence in the Gini coefficients can also be explained by the FE approach or country dummies, which are similar to CCT. Our empirical findings reveal that combining the convergence prediction of NGM with FE/CCT is a viable strategy.

We ran additional regressions that detected the impact of time trend, previous Gini coefficients, country dummies and their interaction terms and of real per capita GDP on the actual inequality of personal incomes. Furthermore, we could confirm the Lessmann and Seidel result (2015), according to which regional inequality depends on the respective per capita GDP, y, and follows an N-shape. Finally, as the NGM also implies stationarity in time series of inequality measures such as the Gini coefficient, we implemented ADF tests to inquire the existence of stationarity. Our results weakly support this view.

The present study's limitations are related to the availability of consistent and comparable data on the personal income distribution's statistical moments, such as the modus of incomes. Consequently, it was not possible to test, directly, all the implications of the theoretical model. Future research on personal income distribution in Europe could intend to deepen the knowledge regarding the convergence-stimulating effects of the economic crises or the change in the institutional setting and could also interconnect the analysis of personal income distribution with the development of macroeconomic shares of total income (profits and wages).

Declarations

Funding

None

Conflicts of interest/Competing interests

The authors declare that they have no conflict of interest.

Availability of data and material

All data were extracted from a publicly available data source.

Code availability

All

Authors' contributions

All authors contributed to the study conception and design. The second author performed a theoretical foundation and model setup, model derivation and empirical analysis were provided by the first author. The second author wrote the first draft of the manuscript, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

References

Alfani, G., & Ryckbosch, W. (2015). VOX[EU] / CEPR. Income inequality in pre-industrial Europe | VOX, CEPR Policy Portal (voxeu.org).

Beichelt, F. E., & Montgomery, D. C. (2003). *"Teubner Pocket Book of Statistics" (Teubner-Taschenbuch der Stochastik: Wahrscheinlichkeitstheorie, Stochastische Prozesse, Mathematische Statistik)* (1st ed.). Wiesbaden, Germany: Teubner.

Blümle, G. (1992). "Personal income distribution as a mode of expression for equilibrium in income distribution. A comment on the relativization of a goal for personal income distribution" (Personelle Einkommensverteilung als Ausdruck eines Verteilungsgleichgewichts. Anmerkungen zur Relativierung des Ziels der personellen Einkommensverteilung). In H. Mäding, F. L. Sell, & W. Zohlnhöfer (Eds.), *"Economics by Order of Political Consulting–Basic Questions and Scopes" (Die Wirtschaftswissenschaft im Dienste der Politikberatung–Grundsatzfragen und Anwendungsbereiche)* (pp. 209–225). Berlin, Germany: Duncker & Humblot.

Blümle, G. (2005). *Theory and Policy of Income Distribution* (Verteilungstheorie und Verteilungspolitik). Freiburg i. Br, Germany: Mimeo.

Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *The American Economic Review*, 90(1), 166–193.

Chiang, A. C. (1984). *Fundamental Methods of Mathematical Economics* (3rd ed.). Auckland, New Zealand: McGraw-Hill International Editions.

De Beer, P. (2012). *The impact of the crisis on earnings and income distribution in the EU,* ETUI Working Paper 2012.01, Brussels.

Dolls, M., Fuest, C., & Peichl, A. (2011). Automatic stabilizers, economic crisis and income distribution in Europe. In Immervoll, H., Peichl, A. and Tatsiramos, K. (Eds.): *Who loses in the downturn? Economic Crisis, Employment and Income Distribution Research in Labor Economics, Vol. 32,* Emerald Group Publishing Limited, Bingley,pp. 227–255.

Esping-Andersen, G. (1990). *The Three Worlds of Welfare Capitalism.* Princeton, NJ: Princeton University Press.

Esping-Andersen, G. (1994). *After the golden age: the future of the welfare state in the new global order.* UNRISD Occasional Paper: World Summit for Social Development, No. 7, Geneva.

Esping-Andersen, G. (1998). The three political economies of the welfare state. In J. O'Connor (Ed.): *Power Resources and the Welfare State* (pp. 123–153). Toronto, Canada: University of Toronto Press.

Esping-Andersen, G., & Myles, J. (2009). Economic inequality and the welfare state. In W. Salverda (Ed.), *The Oxford Handbook of Economic Inequality* (pp. 639–664). Oxford, UK: Oxford University Press.

EU Statistics on Income and Living Conditions (EU-SILC) (2015).

Genc, I., Miller, J. R., & Rupasingha, A. (2011). Stochastic convergence tests for US regional per capita personal income: some further evidence; a research note. *The Annals of Regional Science,* 46(2), 369–377.

Hall, P. A., & Gingerich, D. W. (2004). *Varieties of capitalism and institutional complementarities in the macroeconomy.* MPIfG Discussion Paper No. 04/5, Cologne.

Hall, P. A. & Soskice, D. (2001). *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage.* Oxford, UK: Oxford University Press.

Household Finance and Consumption Network. (2013a). The Eurosystem household finance and consumption survey: Methodological report for the first wave. ECB Statistical Paper Series, No. 1.

Household Finance and Consumption Network. (2013b). The Eurosystem household finance and consumption survey: Results from the first wave. ECB Statistical Paper Series, No. 2.

Kaitila, V. (2013). *Convergence, income distribution, and the economic crisis in Europe*. ETLA Working Paper No. 14, June, Helsinki.

Lessmann, C., & Seidel, A. (2015). *Regional inequality, convergence, and its determinants – A view from outer space*. CESifo Working Paper No. 5322, Munich.

Lin, P. & Huang, H. (2012). Convergence in income inequality?: evidence from panel unit root tests with structural breaks. *Empirical Economics, Vol. 43 (1), pp. 153-174.*

Martinez-Carrion, J. M., & Maria-Dolores, R. (2017). Regional inequality and convergence in Southern Europe. Evidence from height in Italy and Spain, 1850-2000. *Revista de Economia Aplicada, 25*(74), pp. 75–103.

Medgyesi, M., & Toth, I. (2009). Economic growth and income inequalities. In T. Ward et al. (Eds.), *European Inequalities. Social Inclusion and Incomes Distribution in the European Union* (pp. 131–152). Budapest, Hungary: Tarki Social Research Institute.

Melchior, A. (2008). *Regional inequality and convergence in Europe, 1995–2005*. Case Network Studies & Analyses No. 374, Case, Warsaw.

OECD Income Distribution Database (2020). Retrieved from http://stats.oecd.org/Index.aspx?DataSetCode=IDD.

Paas, T., & Schlitte, F. (2007). *Regional income inequality and convergence processes in the EU-25*. HWWI Research Paper 1-11, Hamburg.

Pareto, V. (1895). La Courbe de la Répartition de la Richesse. *Giornale degli Economiste*, January.

Paulus, A., Figari, F., & Sutherland, H. (2009). The effects of taxes and benefits on income distribution. In T. Ward, O. Lelkes, H. Sutherland, & I. Gyorgy Toth (Eds.), *European Inequalities. Social Inclusion and Incomes Distribution in the European Union* (pp. 153–175). Budapest, Hungary: Tarki Social Research Institute.

Ramser, H. J. (1987). *Theory of Income Distribution (Verteilungstheorie).* Berlin, Heidelberg and New York: Springer Verlag.

Ravallion, M. (2003). *Inequality convergence. Economics Letters,* 80(3), pp. 351–356.

Savoia, F. (2019). *Income inequality convergence across EU regions.* LIS Cross-National Data Center in Luxembourg, Working Paper No. 760, July.

Sell, F. L. (2015). *The New Economics of Income Distribution.* Cheltenham, UK and Northampton, USA: Edward Elgar Publishing.

Tiemann, A. (2006). *Stability in G. Esping-Andersen's "The three worlds of capitalism".* LIS Working Paper No. 449, Luxembourg.

Van Kerm, P., & Alperin, M. N. P. (2013). Inequality, growth and mobility: The intertemporal distribution of income in European countries 2003–2007. *Economic Modelling,* 35(3), 931–939.

Appendix

A1. Derivation of aggregated utility

The aggregated utility function becomes the following:

$$\sum_{i=1}^{n} U_i = n\sqrt{1/\sigma^{\gamma}}\sqrt{\overline{y}_a - y_{mo}}$$

Inserting $y_{mo} = \exp(\mu - \sigma^2)$ and $\overline{y}_a = \exp(\mu + \frac{1}{2}\sigma^2)$ yields $\sum_{i=1}^{n} U_i = n\sqrt{1/\sigma^{\gamma}}\sqrt{\overline{y}_a(1-\exp(-\frac{3}{2}\sigma^2))}$.

Note that n and $\overline{y}_a$ are both positive constants such that

$$\max\sum_{i=1}^{n} U_i \equiv \max\sqrt{1/\sigma^{\gamma}}\sqrt{1-\exp(-\frac{3}{2}\sigma^2)} \equiv \max\frac{1-\exp(-\frac{3}{2}\sigma^2)}{\sigma^{\gamma}}$$

Taking the FOC. yields
$-\gamma\sigma^{-\gamma-1}(1-\exp(-\frac{3}{2}\sigma^2)) + \sigma^{-\gamma}\exp(-\frac{3}{2}\sigma^2)3\sigma = 0$,
and multiplying both sides with
$\sigma^{\gamma+1}\exp(\frac{3}{2}\sigma^2)$ yields $-\gamma(\exp(\frac{3}{2}\sigma^2)-1) + 3\sigma^2 = 0$.

Denoting the value of FOC at equilibrium as F, we then obtain

$$\frac{d\sigma^*}{d\gamma} = -\frac{\partial F/\partial\gamma}{\partial F/\partial\sigma} = \frac{\exp(\frac{3}{2}\sigma^2)-1}{\partial F/\partial\sigma}$$

Since $\sigma^2 > 0$, the numerator is positive, because it is evaluated at the maximum and the denominator is negative; thus, σ^* decreases when γ increases:
$\frac{d\sigma^*}{d\gamma} < 0$.

A2. Figures of Section 4

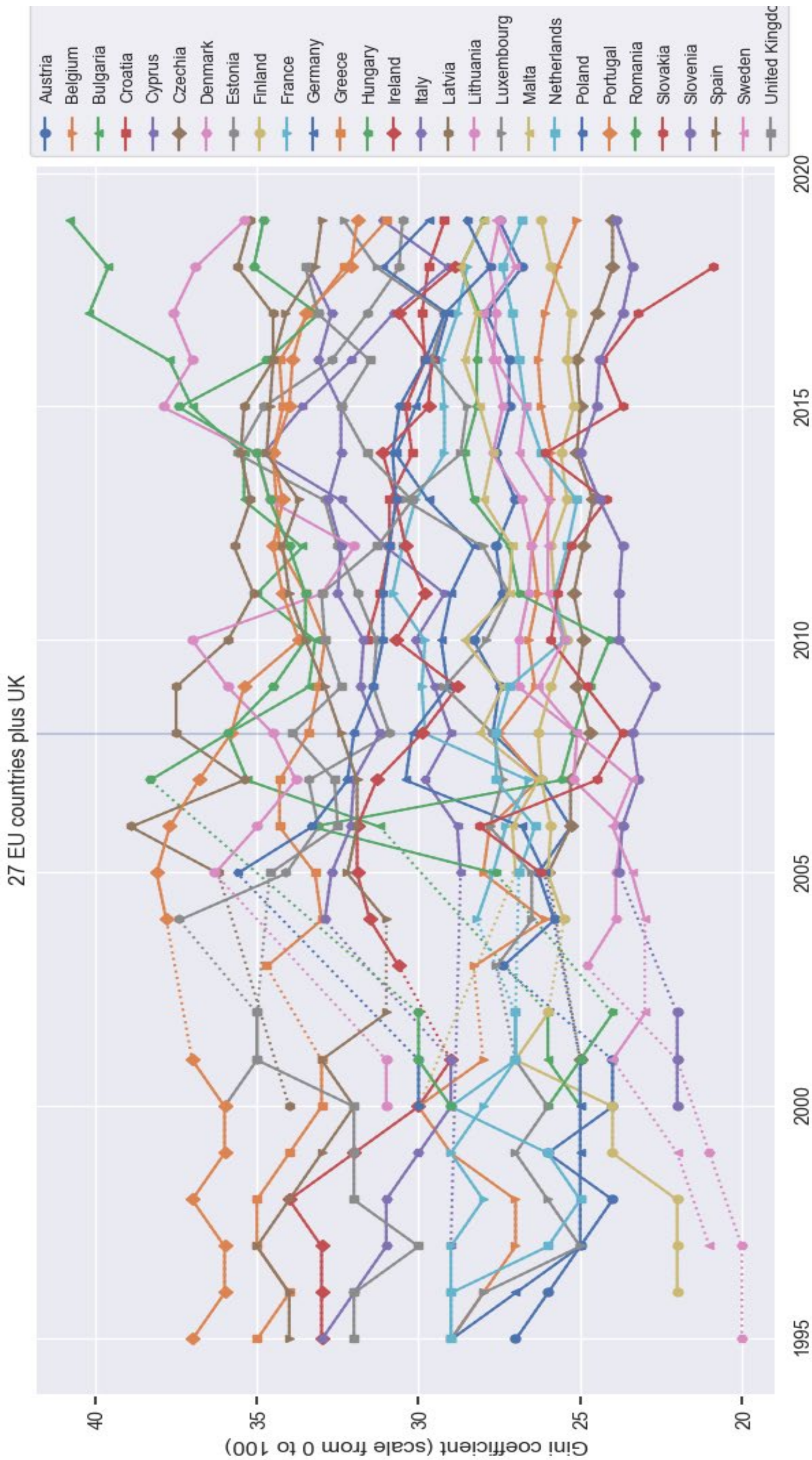

Fig. 4. (SOURCE: EUROSTAT, OWN DEPICTION)

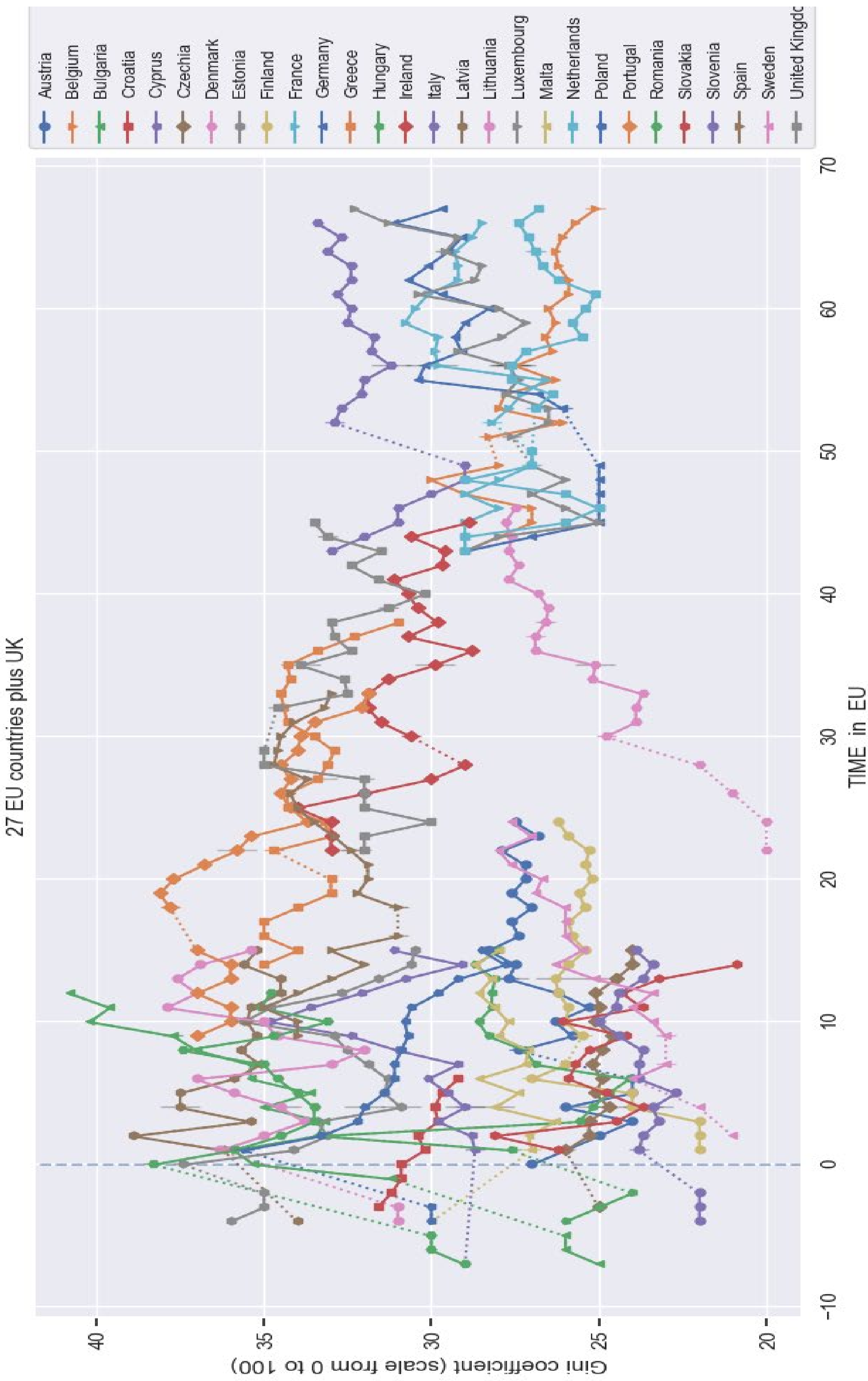

Fig. 5. (SOURCE: EUROSTAT, OWN DEPICTION)

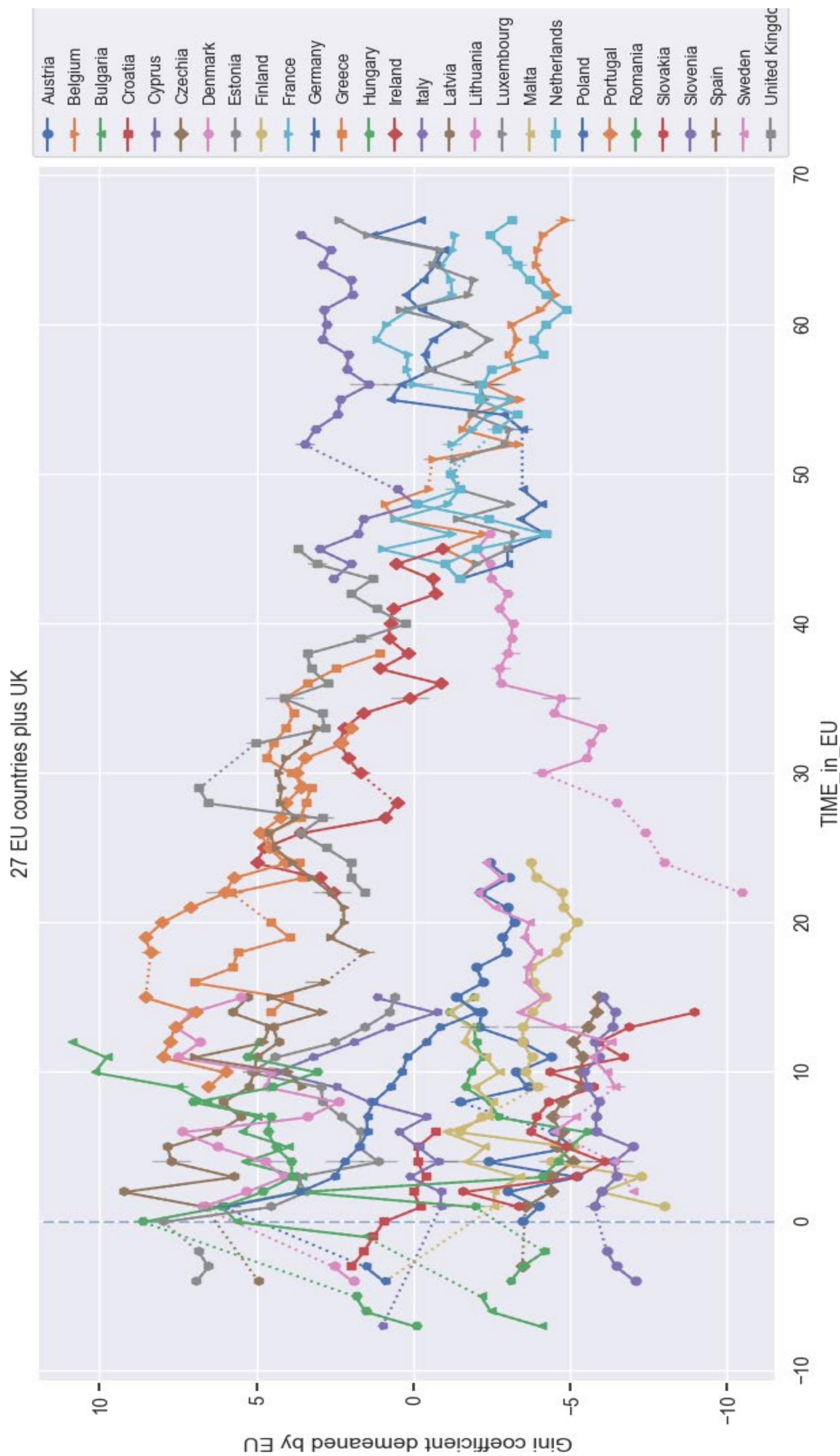
27 EU countries plus UK
Gini coefficient demeaned by EU
TIME_in_EU
10
5
0
-5
-10
-10
0
10
20
30
40
50
60
70
Austria
Belgium
Bulgaria
Croatia
Cyprus
Czechia
Denmark
Estonia
Finland
France
Germany
Greece
Hungary
Ireland
Italy
Latvia
Lithuania
Luxembourg
Malta
Netherlands
Poland
Portugal
Romania
Slovakia
Slovenia
Spain
Sweden
United Kingdo

Fig. 5.

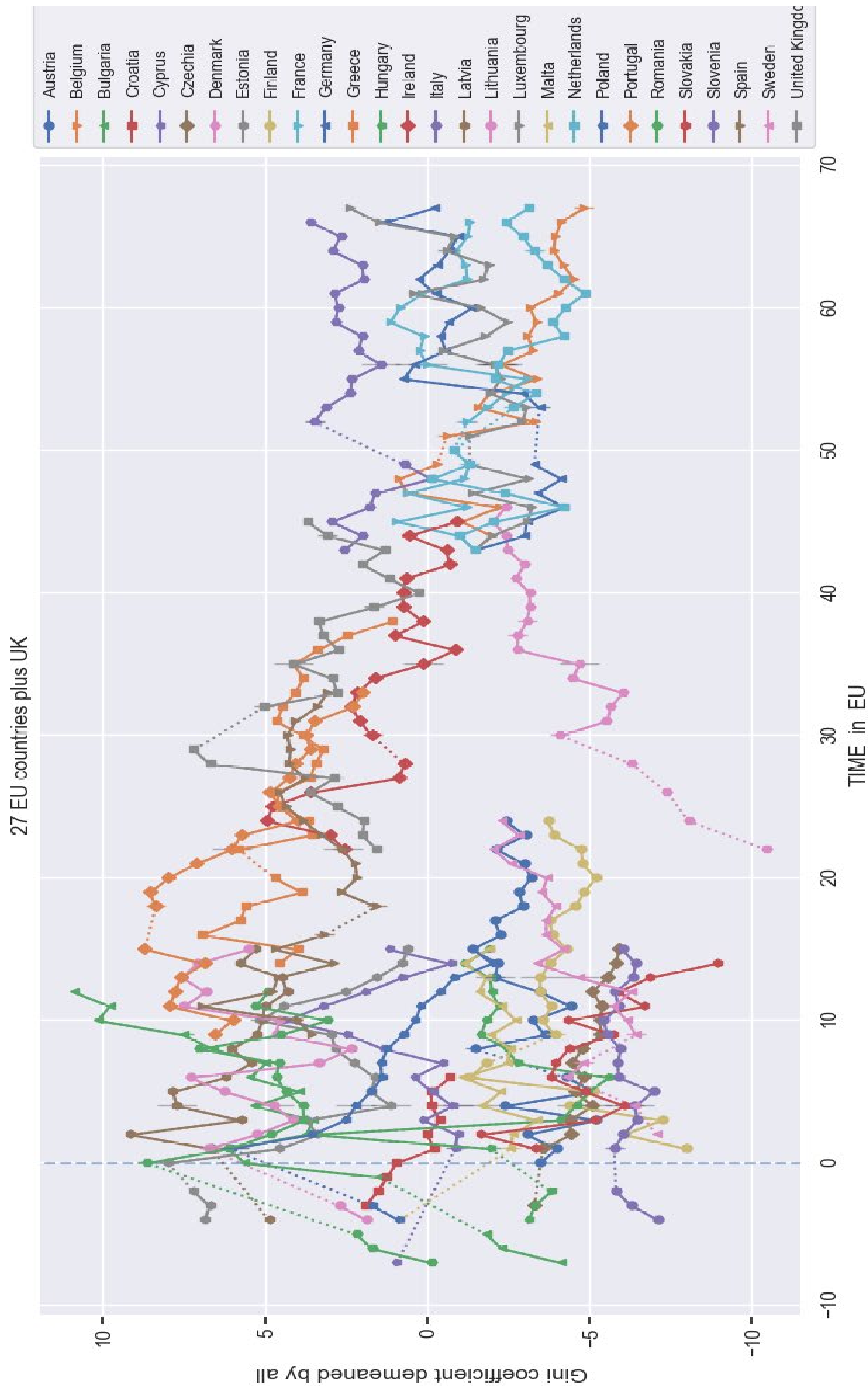

Fig. 7. (SOURCE: EUROSTAT, OWN DEPICTION)

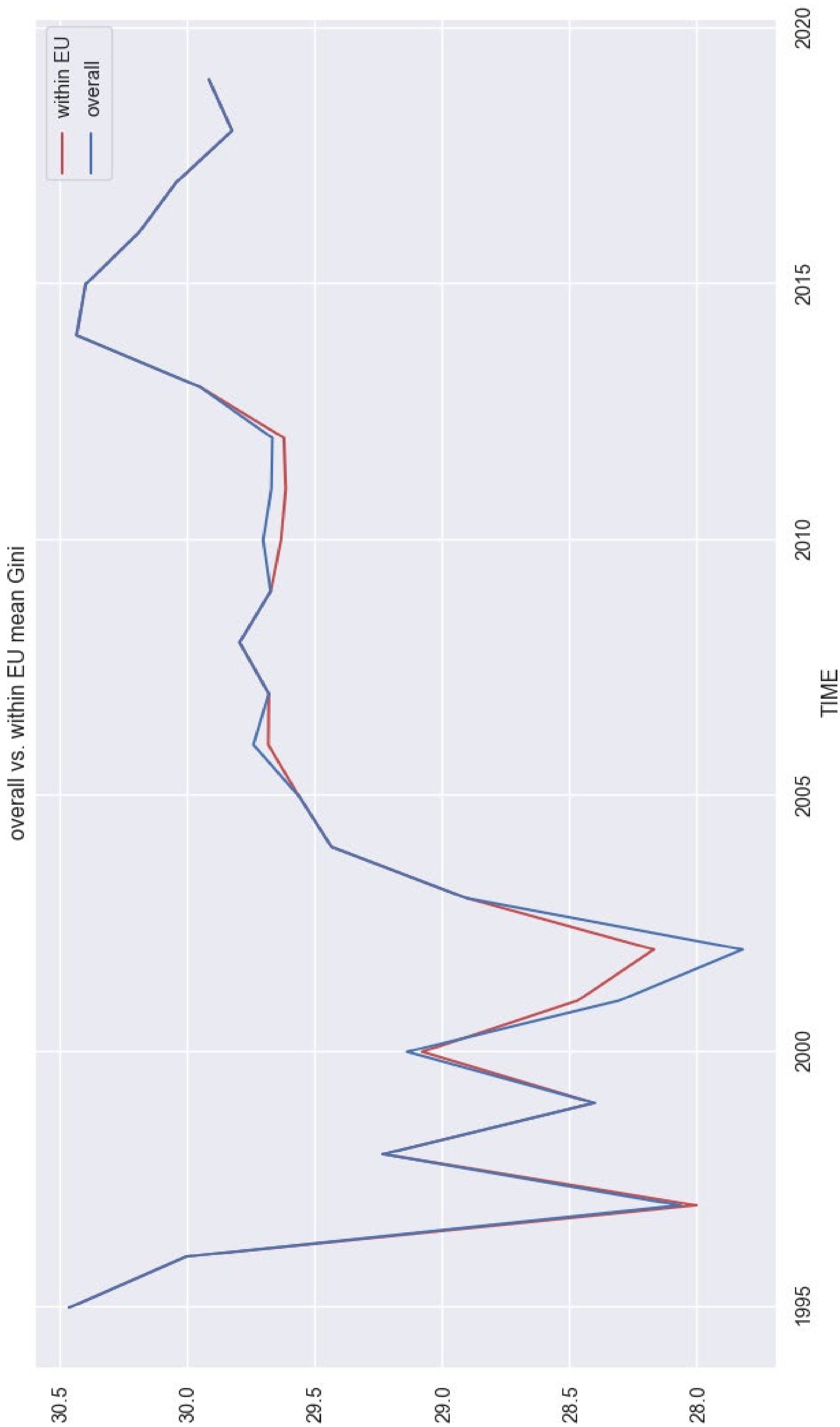

Fig. 8. (SOURCE: EUROSTAT, OWN DEPICTION)

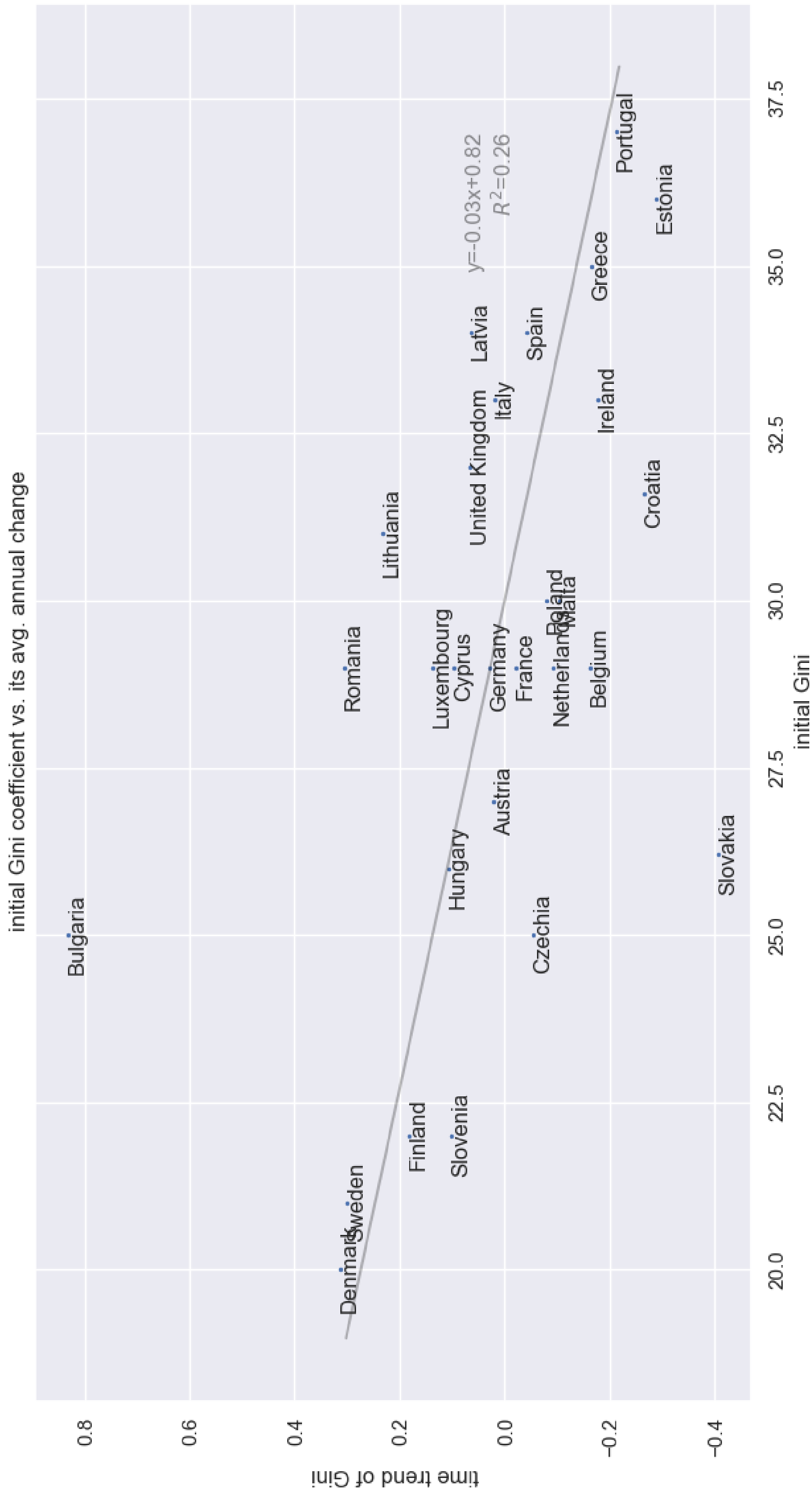

Fig. 9. (SOURCE: EUROSTAT, OWN DEPICTION)

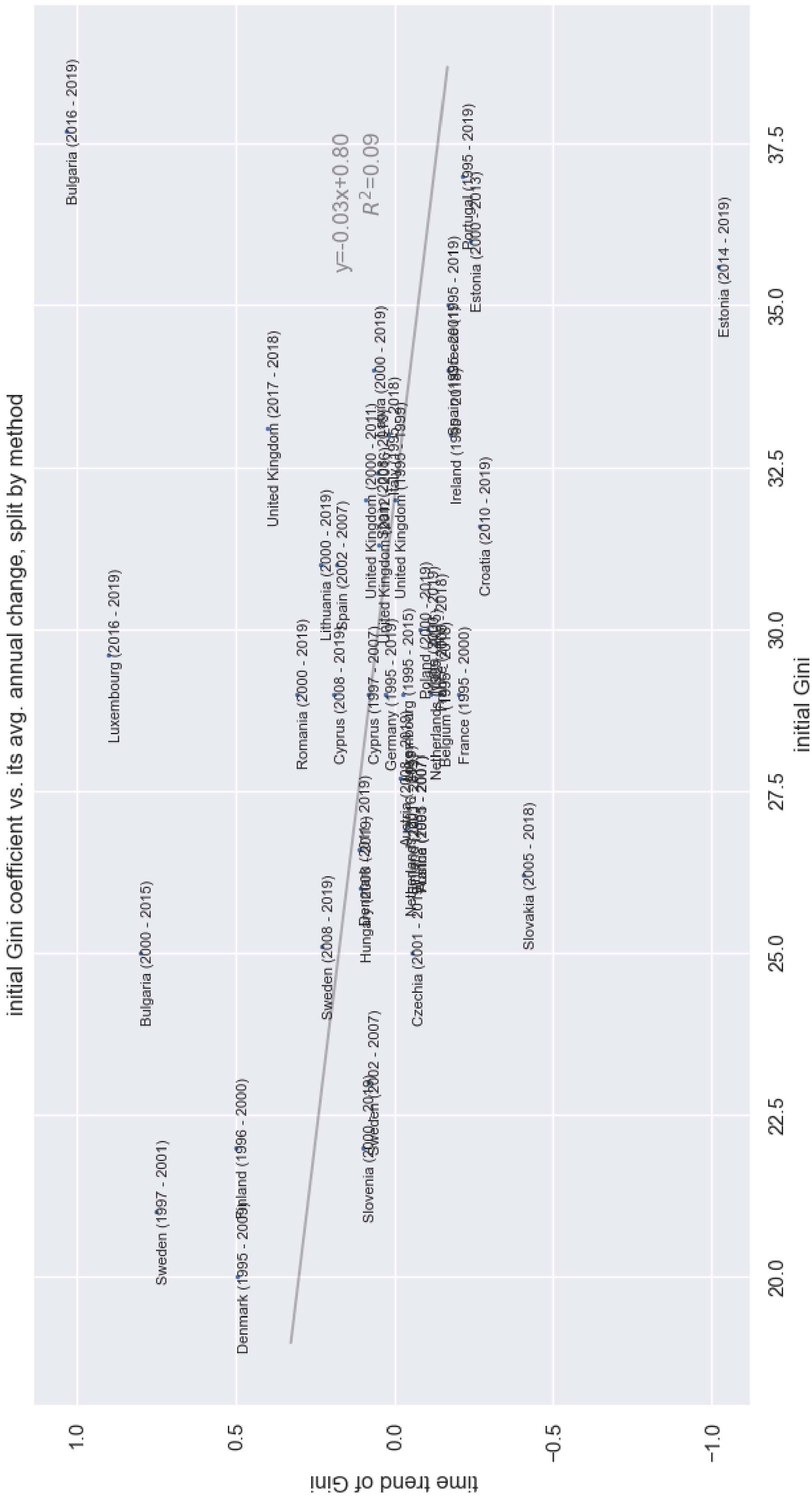

Fig. 10. (SOURCE: EUROSTAT, OWN DEPICTION)

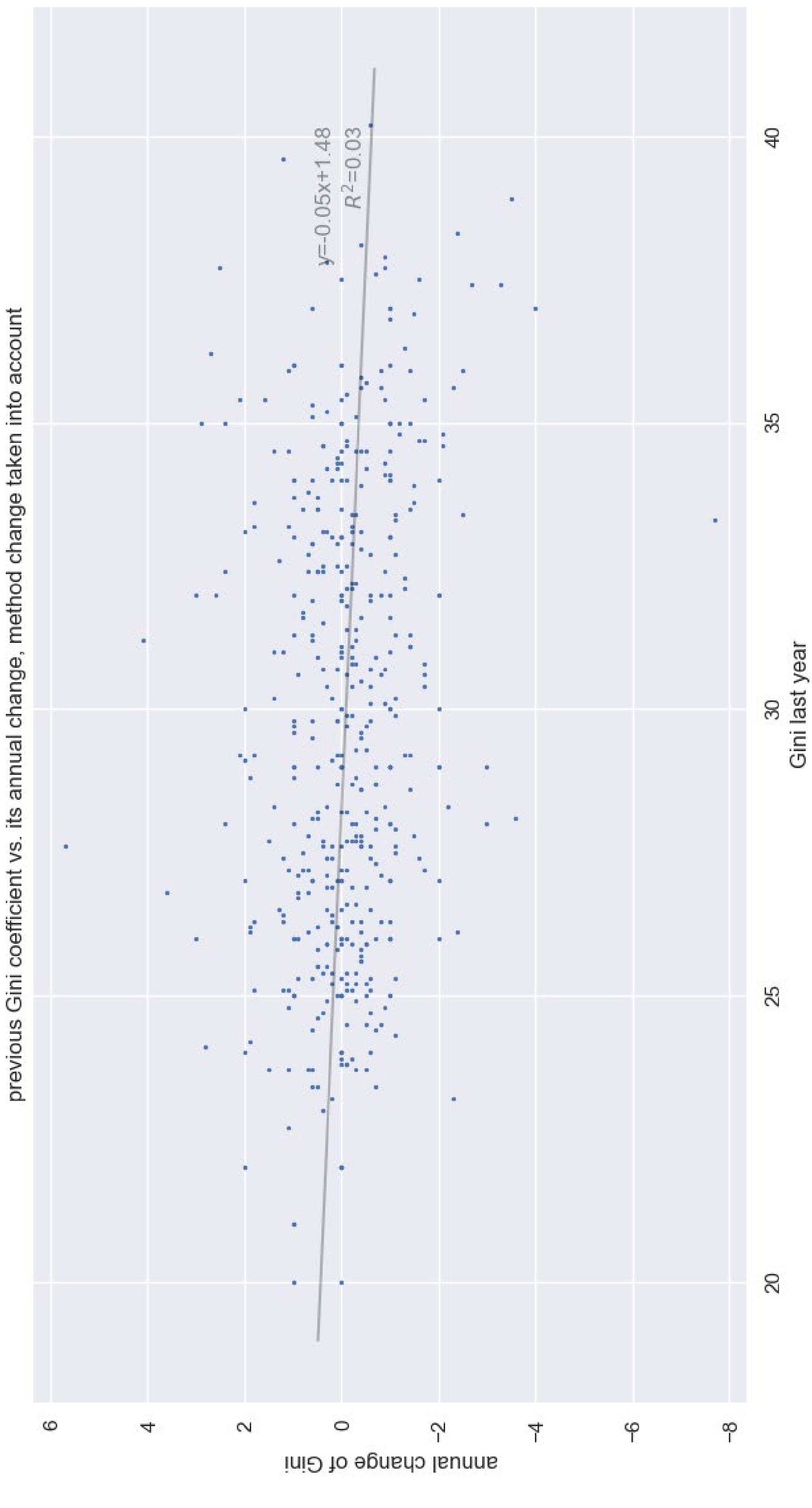

Fig. 11. (SOURCE: EUROSTAT, OWN DEPICTION)

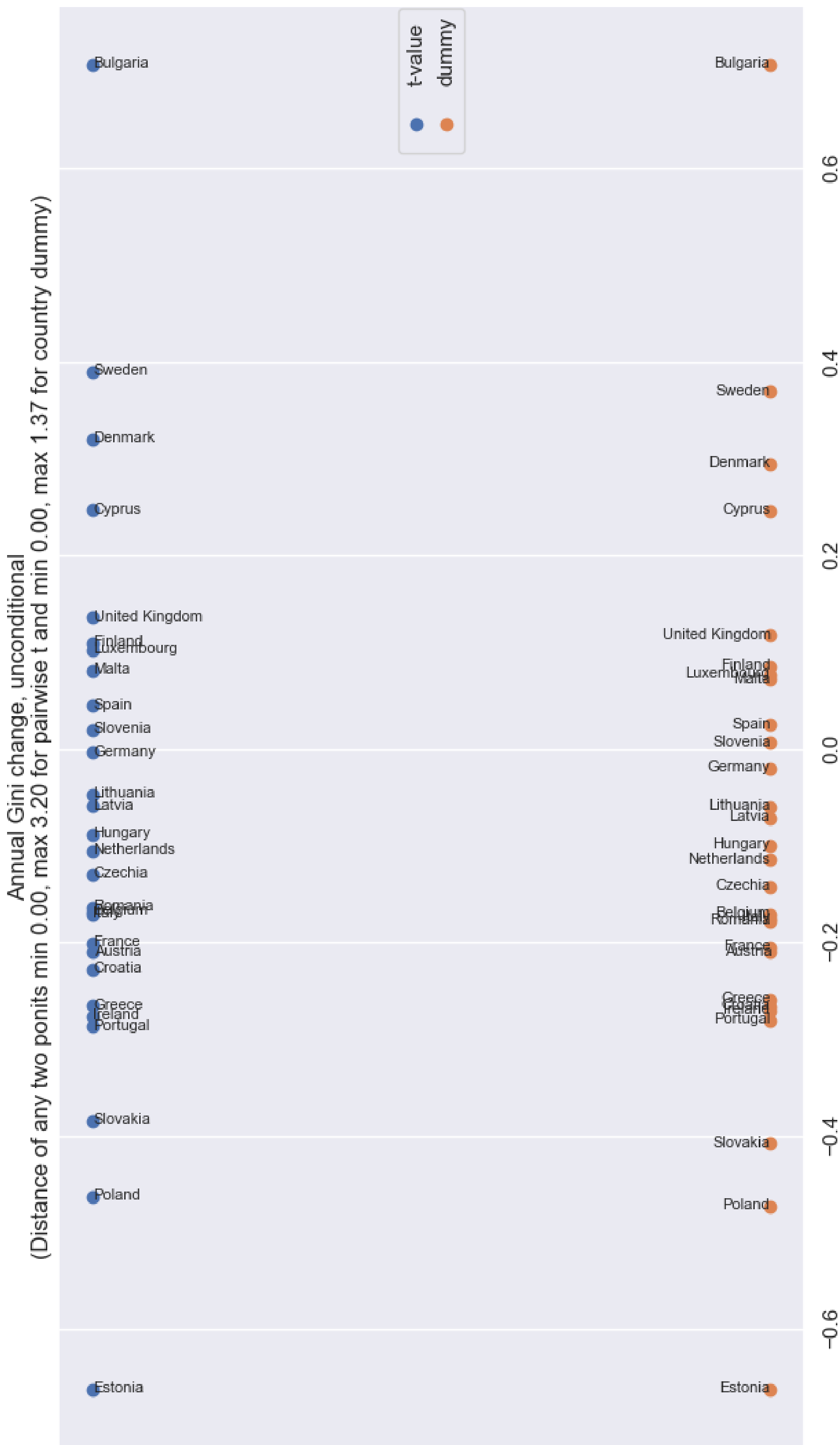

Fig. 12. (SOURCE: EUROSTAT, OWN DEPICTION)

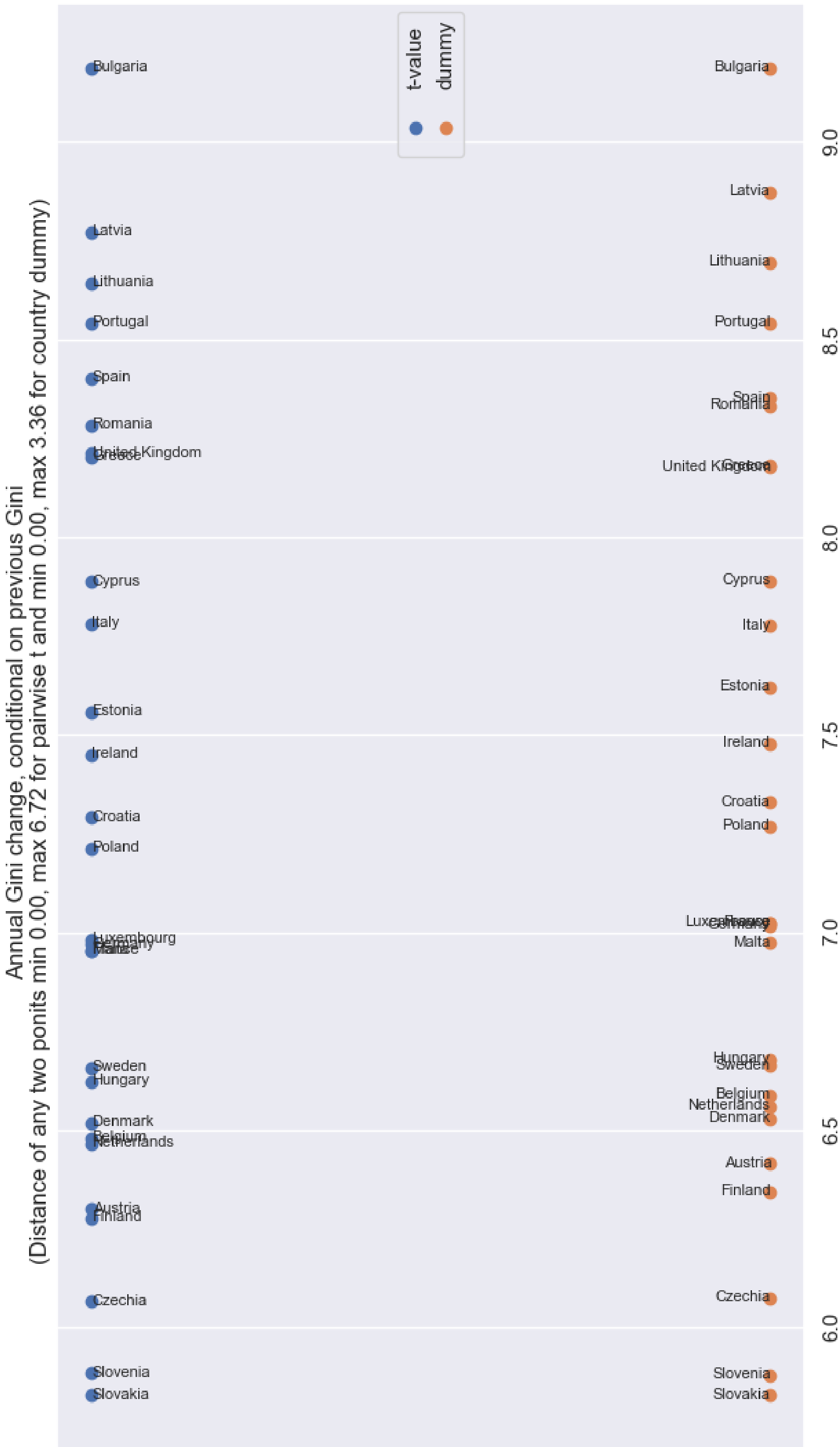

Fig. 13. (SOURCE: EUROSTAT, OWN DEPICTION)