Qual Quant (2012) 46:451–470
DOI 10.1007/s11135-010-9380-2

Education systems and the formation of societal consensus on justice

Rüya Gökhan Koçer · Herman van de Werfhorst

Received: 28 April 2010 / Accepted: 1 November 2010 / Published online: 26 November 2010

Abstract We scrutinize whether the consensus on economic justice prevailing in a society is shaped by institutions, especially by education systems. We argue that social cohesion is ensured by the magnitude of consensus on justice rather than its content. Thus, we take the magnitude of consensus as our dependent variable abstracting it from its content. We examine the impact of various institutions on this variable by using set-theory based arguments, bootstrapping and multivariate models. The findings suggest that the sense of justice in society is significantly shaped by the institutional characteristics of the education system.

Keywords Sense of justice · Education systems · Legitimation of inequalities · Multivariate analysis · Bootstrapping · Set theory

1 Introduction

The sense of justice is an essential component of any human society that ensures social cohesion. In practice the sense of justice implies a ranking and legitimation of inequalities: considering some of them fair and the others unfair (Alves and Rossi 1978, p. 542).

There are two groups of empirical work related to evaluation of inequalities: first, there is a literature on individual attitudes towards redistribution policies (e.g., Alesina and Angeletos 2005; Osberg and Smeeding 2006). Second, there are studies that scrutinize the way in which individuals judge the fairness of earnings (for example: Alves and Rossi 1978; Jasso 1989; Kunovich and Slomczynski 2007). While studies on attitudes towards redistribution point out individuals' characteristics (self-interest, beliefs, perception) as the source of their evaluations of inequalities, the research on fairness of income argues that there are societal consensuses that guide individuals as they judge the fairness of given income levels. In this picture, the former school considers any consensus on fairness in a society as an aggregation

R. G. Koçer (✉) · H. van de Werfhorst
Department of Sociology and Anthropology, University of Amsterdam, Uva-Spinhuis,
OZ ACHTERBURGWAL 185, 1012 DK, Amsterdam, The Netherlands
e-mail: r.g.kocer@uva.nl

of individual attitudes, that is, the outcome of some exogenous variables, that emerge at micro level. The latter, on the other hand, shows that the basis of judgments on inequality is not the individual level but fails to account for the factors that generate the consensus, which determines individuals' judgments on ranking of inequalities.

In this paper, we subscribe to the idea that the sense of justice results from societal consensuses and we construct a set-theory based multivariate model in order to reveal factors that shape these consensuses. For this purpose we develop a method which allows us to make an important distinction between two concepts which hitherto remained integrated: the *average norm* about fairness that exist in a given society or sub population, and the extent to which there is *consensus* on this norm. We argue that those societies where there is wide support for a particular norm of fairness would enjoy more social cohesion regardless of the norm itself. Because the real function of a consensus on justice is that it keeps the society together. This is accomplished by the magnitude of consensuses rather than their content. Hence we take the magnitude of consensus as our dependent variable abstracting it from its content.

Our central claim is that societal or subgroup level consensuses which inform justice evaluations are shaped by institutional characteristics, especially by education systems of societies. Following up on earlier research on educational institutions and inequality, we examine whether external differentiation and vocational orientation of education systems affect the magnitude of consensus that guides individuals as they make judgments about fairness of inequalities. Differentiation in separate tracks early in the school career is known to affect equality of opportunity negatively, whereas a strong vocational orientation of the system improves labour market integration of school leavers (Brunello and Checchi 2007; Van de Werfhorst and Mijs 2010; Müller and Gangl 2003). We investigate whether these institutional features have repercussions on the legitimation of inequalities, and thus, on the sense of justice.

We use the International Social Survey Programme (ISSP). In the ISSP survey collected in 1999 respondents were interviewed about their opinions on inequality. Building on an approach developed by Osberg and Smeeding (2006) on attitudes towards inequality, we develop a new measurement of consensus among (sub-)populations. This measurement is then used as a dependent variable in order to examine the impact of education systems on the formation of societal consensus on justice. We control for relevant institutional characteristics, including the level of income inequality and labour market coordination.

We first examine the existing literature to reveal the shortcomings of considering the judgment on fairness as the outcome of individuals' characteristics and show the strengths and weaknesses of existing studies that consider societal consensuses as the basis of these judgments. Secondly, we outline our hypothesis that it is the institutional structure and especially the education system that shapes the consensus on justice. Thirdly, we develop a measure that allows us to express consensus quantitatively. Finally, we present some multivariate models which validate our hypothesis. The paper ends with the discussion and implications of findings.

2 Literature review

One can argue that the income distribution in a society is the reflection of the stratification which is at least partly legitimized (Hermkens and Boerman 1989, pp. 201–202) whereas redistribution policies are the remedy for those aspects of this stratification that are considered unfair. Thus, individuals' judgment about an income distribution and their attitude towards redistribution policies are both crucial indicators of the sense of justice prevailing in a soci-

ety. One may examine both of these indicators in order to comprehend the sources of justice evaluations.

2.1 Attitudes towards redistribution policies

The scrutiny of the attitudes towards redistribution policies mainly focuses on characteristics of individuals in order to explain various attitudes towards redistribution policies (such as support or opposition). This line of thought is premised on two assumptions: firstly, individuals' attitudes are assumed to be shaped by self-interest, for example, it is argued that individuals who support redistribution policies are those who are most likely to be the beneficiaries of these policies (Linos and West 2003, p. 399). However, there is also "strong evidence that financial self-interest is an insufficient explanation for redistributive attitudes" (Fong 2001, p. 227). Secondly, the belief regarding the nature of the relationship between effort and outcome is considered to be the main determinant of the attitudes towards redistribution: those who believe the profound role of luck for advancement in society are most likely to support redistribution policies (Alesina and Angeletos 2005, p. 963).

The impact of macro level variables on individuals' attitudes towards redistribution policies has also been studied. Certain institutional characteristics of societies such as the type of welfare regime and features of the production system are considered to influence individuals' judgments (Linos and West 2003; Esping-Andersen 1997). However, the envisaged source of judgments is still self-interest: "workers with specific skills" are, for example, expected to be "supportive of additional social protection" (Linos and West 2003, p. 395) presumably due to the risk that market fluctuations might reduce the demand for their specific skills and oblige them to seek financial aid. Thus, the scrutiny of the impact of macro level variables on attitudes towards redistribution is actually based on individual level self-interest, which is either considered as an intrinsic property of human or remains as an exogenous variable. Consequently, the source of justice evaluations remains unexplained.

Institutional structures are also thought to be influencing individuals' judgment because of the close link between human perception and cognition: any persistent state of existence would be gradually normalized by individuals so that the judgment about 'what ought to be' would be increasingly shaped by 'what is' (Marshall et al. 1999, p. 351, Homans 1974). However, this idea inevitably leads to a contradictory theory of elites: some individuals (i.e. elite minority) may shape certain institutional structures, such as employment procedures, in order to, for example, favor merit based selection so as to generate more efficient market outcomes. The other individuals (i.e., non-elite majority), on the other hand, gradually internalize the logic and consequences of these institutional arrangements and adjust their sense of justice accordingly; considering 'what is' as just. But what makes the elites elite? If there are intrinsic individual qualities that determine whether one would be an elite then, macro level structures, which explain the sense of justice, would still be generated by micro level intrinsic qualities that are unaccounted for (i.e., some individuals have an independent sense of judgment).[1]

Thus, if we seek the factors that shape the sense of justice in a society by appealing to individual level, the source of individual level judgments would remain unexplained, leading

[1] Another question is this: what determines which consideration (of elites) shapes the institutional structure? Why, for example, would the elite desire to create economically efficient institutions rather than the establishment of ethically sound structures? if it is the environment that determines the guiding principle then we are back to the denial of agency: all individuals at different levels are shaped by macro structures, if it is not the environment then we have to extend the elite theory: one needs to identify some elites of elites which determine the guiding principle of elites.

to the ontological position that the prevailing sense of justice is simply the aggregation of individuals' attitudes which are, in turn, generated by factors that are intrinsically present (i.e., self-interest) or shaped by contingencies (i.e., beliefs) or determined by elites whose source of evaluation remains obscure.

However, there is a more serious failure: in fact one may argue that there is no individual level in justice evaluations. Whenever a particular individual characteristic is pointed out as a reason for a certain attitude (supporting welfare regime because of having a very specialized occupation that might be easily lost in crisis) we actually look at a set of individuals who are members of a special subgroup in the society. More formally, we look at the elements of a set which is generated by intersections of many other sets: {women with specialized occupation}, {men with specialized occupation}, {urban dwellers with specialized occupation}, {young people with specialized occupation}. The crucial point here is that all other possible intersections of these and other sets (which contain same individuals) may also influence the justice evaluations with different logics than the one which is pointed out. Consequently, the motives of subgroups which remain in several intersections (i.e. almost all subgroups) would be ambiguous: one can invent many 'individualistic characteristics' in order to account for a sense of justice prevailing in a particular subgroup. Thus, focusing on individual level is essentially looking at a single consensus emerging in a single intersection set by disregarding all other possible intersection sets that contain the same individuals, and thereby portraying a blurred picture as if it is crystal clear.

2.2 Fairness of income research

The other strand in the justice studies, that is, individuals' judgment on the fairness of earnings, seems to have tackled with the problem of macro–micro connection more directly by asking the question of "whether there is consensus on the norms that govern the distribution of rewards in a society"(Hermkens and Boerman 1989, p. 202). The answer seems to be affirmative: "there is a clear pattern" in society "which reflects the existence of a normative consensus on fairness" (Alves and Rossi 1978, p. 562). Indeed this normative consensus appears to be rather robust: although there are distinctions between subgroups in the way in which they judge fairness of incomes (Hermkens and Boerman 1989, p. 212) vignette based studies do show that differences in the characteristics of individuals who make the judgment on fairness of a given income do explain only a small part of the variation (Alves and Rossi 1978, p. 562). In short, individuals in a given society, regardless of their individual or group characteristics, are guided by a normative consensus (or by a set of normative consensuses) as they make judgments about the fairness of a given income.

These consensuses are expressed qualitatively, for example, there is an agreement in the USA that occupational attainments are to be valued and rewarded more than the educational attainments of individuals (Hermkens and Boerman 1989, p. 212) and in evaluation of households' income male partners' characteristics should be taken into account more than those of the female partner (Alves and Rossi 1978, pp. 562–563). Against this background, one may argue that not only the individuals' judgment on fairness of a given income distribution but also their attitude towards redistribution policies may be influenced by some consensuses which are articulated as common principles. Indeed there are several such principles like desert (Marshall et al. 1999, p. 350), contribution (Fong 2001, p. 226) and reciprocity (Fong 2001, p. 242) that might be guiding individuals as they evaluate redistribution policies regardless of their individual circumstances.

In this paper, we also subscribe to the idea that the sense of justice results from societal consensuses. However, in the literature the source of these consensuses remains unaccounted

for. We don't know which characteristics of societies affect or form these consensuses. Moreover, although it is argued that there are some variations in consensuses across subgroups, the extent to which these variations suffice to generate several societal level consensuses remains obscured. Thus we also don't know whether there are contradictory consensuses in the same society which lead to conflicting judgments on fairness and legitimation of inequalities. Moreover, the empirical work which revealed the existence of consensuses on justice does not go beyond a single country. Consequently, we are informed about the guiding principles of individuals, for example in the USA, but we don't have any systematic inquiry across countries which might allow us to see whether these principles are universal.

We argue that the failure to answer these questions, at least partly results from lack of appropriate conceptualization of the notion of consensus. In order to seek the sources of consensus one should differentiate the *average norm* about fairness and the extent to which there is *consensus* on this norm among members of a given society or sub-population. We claim that it is this second property, that is, the magnitude of consensus, that needs to be explained by appealing to institutional characteristics of societies.

3 Theoretical framework: education and consensus on justice

Education is an important factor in the dynamics of inequality. Cross-national differences in inequalities can be explained, to a large extent, by institutional characteristics of education systems. Education systems do not only grant diplomas and thereby create legitimate rankings between individuals but they also shape the collective knowledge and perception by promoting certain patterns of reasoning and discrediting others (Meyer 1977). In this way, education influences the way in which people evaluate their environment and themselves. Therefore we argue that, in any society the institutional structure that regulates the distribution of knowledge and qualifications, that is, the education system, should be the main factor that shapes the consensuses which guide individuals as they make evaluations about justice. We identify two institutional features of education systems that are crucial for social stratification.

First, education systems differ in the extent to which they separate students into different classes and school types in compulsory education. This institutional characteristic of *differentiation* has been shown to affect inequality; education systems that differentiate students during the early phase of secondary education into separate schools and school types are characterized by larger dispersions of academic achievement, and stronger impacts of social class and ethnicity on achievement (Brunello and Checchi 2007; Hanushek and Wössmann 2005; Van de Werfhorst and Mijs 2010).

Second, the *vocational orientation* of education systems is relevant for inequality because it affects the transition from school to work. In education systems that are strongly vocationally oriented, such as Germany and the Netherlands, the transition from school to work runs more smoothly than in systems that are less vocationally oriented. This has been demonstrated at the individual level in terms of hazard rates of finding employment after leaving school, and at the aggregate level with regard to youth unemployment rates (Breen 2005; Müller and Gangl 2003; Shavit and Müller 1998). Moreover, as Brunello and Checchi (2007) showed, equality of educational opportunity is not negatively affected by a strong vocational orientation. These findings suggest that a vocational orientation has a strong 'inclusive' effect. Yet, other findings with regard to the wage returns to education indicate that vocationally oriented systems have larger educational returns, which has been explained by the more limited supply of tertiary degrees in those systems (Wolbers 2007).

Against this background, one may expect that increasing differentiation in education systems would probably lead to lower levels of consensus, and thus, would be detrimental for social cohesion. On the other hand, there might be a positive impact of increasing vocational orientation both on the consensus and social cohesion due to its inclusion effect. However, given the findings that vocationally oriented systems create larger wage returns, it may be expected that this positive effect of the vocational orientation is mainly found if the level of income inequality is not too high.

We constructed two country-level indices in order to capture these dimensions of education systems[2]: The *differentiation* variable is created by a factor analysis using information on three variables, ranked to a proportional score before the factor analysis was carried out: the age of first selection (reverse coded), the number of tracks available to a typical fourteen-year-old student, and the length of the tracked curriculum as a proportion of total length of secondary education. The resulting scale accounted for 82% of the variance in the dataset with $N = 34$ countries, with an Eigenvalue of 2.45. The scale was z-standardized again for the 34 countries. *Vocational orientation* variable is based on a single indicator: the percentage of students within upper secondary education enrolled in a vocational track. Upper secondary vocational enrollment is a common indicator of the vocational orientation of a country, and is available for a large number of countries (e.g. Shavit and Müller 1998). This variable was also z-standardized for 40 countries of the extended country-level dataset.

Although we argue that education systems play a crucial role in the formation of consensuses on ranking and legitimation of inequalities, we also acknowledge that judgments on perceived inequalities should be, at least to some extent, function of the factual inequalities and the institutions that regulate the economy. We use *Gini coefficient* as a concise measure of factual inequality in a country. However, the quantification of the regulation of the economy is not so easy: regulation is a complex phenomenon which may be accomplished through many different ways across countries. Fortunately, thanks to the varieties of capitalism literature, we have tools that can be used for this task. In this literature a distinction is made between two ideal type regulatory regimes: first, coordinated market economies, where the market is tamed through a set of regulations and coordination between agents is ensured by strategic institutional arrangements, and second, liberal market economies where the market is, to a large extent, allowed to function without hindrance and is expected to generate efficient coordination by itself (Hall and Soskice 2001). Obviously, no country can be classified as a pure case of either ideal type. Instead, a wide range of institutional arrangements should be analyzed for each country in order to identify their position in between these two types. Kunovich and Slomczynski (2007) developed a succinct measure, *coordination*, which allow us to express the position of countries quantitatively in this continuum[3]: low values indicate prevalence of market based coordination and high values indicate strategic coordination. But this measure leaves out the strength of organized labor, which may be crucial in generation and/or rectification of factual inequalities. Thus, we enhance the *coordination* measure further by including *trade union density* into our models. In short, we use gini coefficient,

[2] A common problem in comparative research of individual and country-level data is that countries are ranked based on their more or less 'coincidental' appearance in the microlevel (survey) dataset. We avoid this problem by first gathering information on a maximum number of countries relying on OECD statistics. See www.oecd.org.

[3] This measure is outcome of a factor analysis of 12 variables: highest marginal personal income tax rate, highest marginal corporate tax, government final consumption expenditure, difficulty of hiring workers index, difficulty of firing workers index, cost of firing workers, rigidity of working hours index, number of start-up procedures to start a business, gini, time to resolve insolvency, number of procedures to register property, stock market relative to the baking sector in the financial system.

coordination index and trade union density as three indicators that allow us to capture the magnitude of factual inequality and degree of interference with the market in a country. Our main purpose however, is not to speculate about the impact of these variables, but to capture the real impact of education on justice evaluations by taking other crucial factors also into account.

Thus, we establish a model, which explains the consensus in society on ranking and legitimation of inequalities by education system (differentiation and vocational orientation), regulation of economy (coordination and union density) and factual inequality (gini coefficient) in 15 countries. However, in order to construct such a model we need to develop a sensible measure of consensus and device a modeling strategy.

4 Creating an index of consensus

We want to distinguish two features of the sense of justice: first, the norm about the fairness of inequalities, and second, the magnitude of consensus that exists in a society or subpopulation on this norm. For this purpose, we start from a measurement of individual justice evaluations developed by Osberg and Smeeding (2006). Based on a set of questions in the ISSP survey, these authors created a measurement for individual sense of fairness of the perceived income distribution (see also Jasso 1999 for a similar approach). In the ISSP survey a set of ten professions are listed and respondents are asked to mention their view on the actual and ideal wage for each given profession[4]. In other words, respondents are asked about their perception of "what is" for ten different professions' income and then requested to replace this with "what ought to be". Osberg and Smeeding, by using these answers, proposed to establish the following equation in order to quantify individuals' sense of overall fairness of the perceived income distribution:

$$Y^{ought} = \beta_0 + \beta_1 Y^{is} \tag{1}$$

The crucial element here is the slope coefficient β_1 which summarizes respondents' sense of fairness in income distribution. Obviously, $\beta_1 = 1$ means that according to the respondent every single profession is paid what it deserves, that is, 'what is' equals to 'what ought to be', thus existing inequality in income distribution is fair. On the other hand, $\beta_1 > 1$ hints that, according to the respondent, some professions (which are already in the relatively high-paid part of the distribution) are paid less than they deserve thus there should be increase in the higher end of the income distribution while the others (which are already in the relatively low-paid part of the distribution) should receive even less.[5] In other words, the existing inequality in the income distribution is not sufficiently unequal to be fair thus more inequality is required. Finally, $\beta_1 < 1$ indicates that the respondent is of the opinion that those professions, which are currently paid relatively high wages should receive less, and the others, which receive relatively low wages should receive more. This means that existing inequality in the income distribution is unfair and more equality is required to reach fairness.

[4] A skilled worker in a factory, a doctor in general practice, chairman of a large corporation, a lawyer, a shop assistant, the owner-manager of a large factory, a judge in the highest court, an unskilled worker in a factory, a cabinet minister in the national government, someone in respondents' own occupation.

[5] Those professions whose perceived actual income is lower than $x = \frac{\beta_0}{1-\beta_1}$ are relatively low-paid and those professions whose perceived actual income is higher than this value are relatively high paid professions. This cut-off point is obtained by evaluating the value of x at the intersection between "what is" equals to "what ought to be" line, that is, $Y^{ought} = Y^{is}$, and the line created by the respondents' answers, that is, $Y^{ought} = \beta_0 + \beta_1 Y^{is}$.

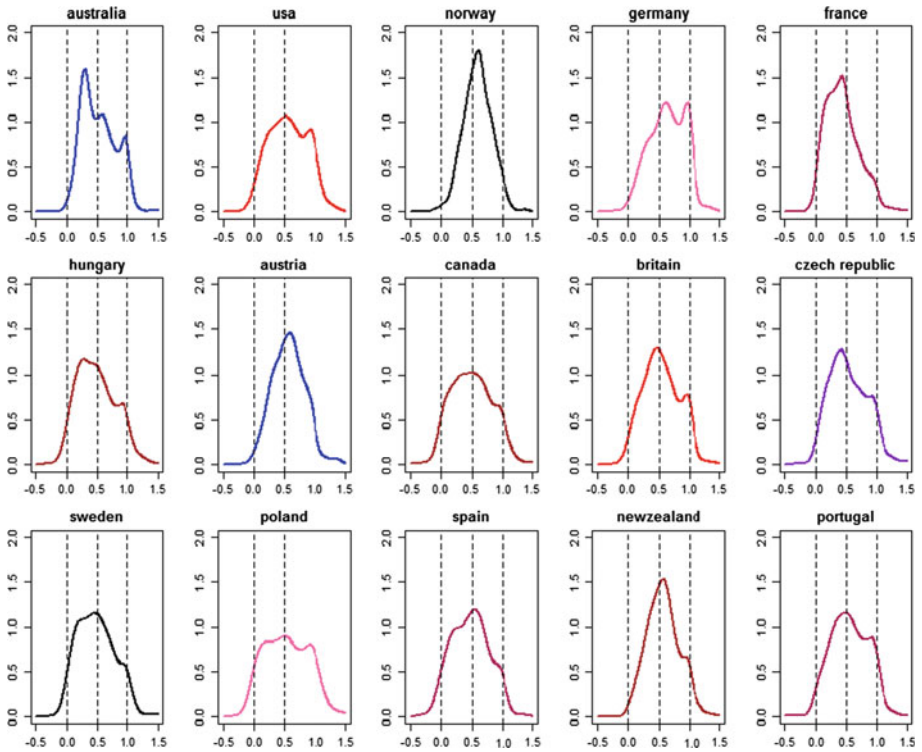

Fig. 1 Densities of leveling coefficients for 15 countries

Obviously, β_1 reveals the extent to which a respondent desires a leveling in the income distribution. Hence, it may be called *leveling coefficient*. The kernel density of leveling coefficients, when estimated on country basis, provide a concise description of the general view of the way in which a society evaluates the fairness of the perceived income distribution: the modalities in the density distribution reflect various groups in the society which agree on the required magnitude of leveling in order to reach a fair income distribution. Obviously, if the population's preferred earnings distribution is compatible with the perceived distribution, one would expect just a single modality to appear on the point 1 in the density distribution, implying that majority of people in the given society think that income distribution "as it is" is "as it ought to be". Therefore, modalities in a country-based kernel density of leveling coefficients when they appear on points different from 1 reveal a collectively agreed sense of unfairness that is to be corrected by a degree of leveling.

Figure 1 depicts the densities of leveling coefficients that we have estimated for 15 countries for which we have data on institutional variables. The figure shows that most countries have a modality below 1. This indicates that the public opinion that in order to attain fairness in income distribution those who receive low wages should receive more and, those who receive high wages should receive less. One can also see that some countries have more than one modality, indicating that large subgroups exist in these societies each of which has its own common opinion about the nature and the extent of leveling that is required in order to attain fairness in income distribution.

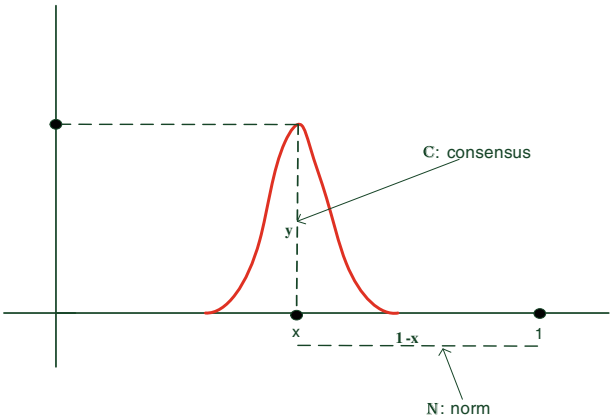

Fig. 2 Actual norm and the consensus depicted in a density of leveling coefficients

We advance the usage of modalities in the leveling coefficient density by distinguishing their two different properties: the "actual norm" and "consensus". We argue that modalities are created by combination of two values:

> **[i]** The distance between the center of modality on the x-axis and 1, that is, $1 - \mathrm{X}$, can be perceived as the deviation of the modality from the ideal norm ($\mathrm{x} = 1$), which represents the judgment that "what is" equals to "what ought to be" thus $1 - \mathrm{X}$ can be used as proxy for "actual norm" of those individuals who create the modality. **[ii]** The maximum height of the modality, that is, the y-value of its top point, provides a sense of the relative size of the consensus among individuals who have the actual norm captured by $1 - \mathrm{x}$ *relative* to the entire society.

Thus, we argue that y : "magnitude of consensus" and $1 - x$: " actual norm" are two crucial components of any normative consensus (see Fig. 2). Magnitude of consensus is our dependent variable. Here two possible concerns should be addressed.

First, from a theoretical perspective separating the consensus from the norm may appear counter-intuitive. However, we argue that the real meaning of justice is not its content but its prevalence: those societies where there is wide support for a particular norm would have high level of social cohesion regardless of the norm. Because the real use of justice is that it keeps the society together and this is the function of the magnitude of consensuses rather than their content. Hence, we take the magnitude of consensus as our dependent variable abstracting it from its norm.

Second, from a technical perspective, it is important to note that y is a measure whose meaning is closely related to the range of the given density, which may vary strongly across populations (i.e., subgroups, countries) and susceptible to the influence of extreme values. Therefore, two y values obtained from two different densities are not directly comparable. In order to make a meaningful comparison between various y values, one should use $\frac{1}{y}$ in multivariate statistical models. Because the expression $\frac{1}{y}$ captures an information that remains constant across populations: the *relative* size of any given y value in relation to one unit of x-axis. Thus, this measure renders various y values comparable by linking the measurements in y-axis to *an anchor* in x-axis. However, this has an implication: when $\frac{1}{y}$ is used as a dependent variable in a model the impact of any given independent variable on y would not be constant. For the model to be estimated is an implicit function of y, that is,

$$\frac{1}{y} = \alpha_0 + \alpha_1 x_1 + \cdots + \alpha_i x_i + \cdots + \alpha_k x_k + \varepsilon \tag{2}$$

which leads to the estimate

$$\frac{1}{\hat{y}} = \hat{\alpha}_0 + \hat{\alpha}_1 x_1 + \cdots + \hat{\alpha}_i x_i + \cdots + \hat{\alpha}_k x_k. \tag{3}$$

Obviously, this indirect model allows us to use an additive error structure which simplifies the estimation procedure. But, we are interested in $\hat{y}$ not $\frac{1}{\hat{y}}$ and

$$\hat{y} = \frac{1}{\hat{\alpha}_0 + \hat{\alpha}_1 x_1 + \cdots + \hat{\alpha}_i x_i + \cdots + \hat{\alpha}_k x_k} \tag{4}$$

Here, however, the impact of any given independent variable on the dependent variable would include the values of all independent variables adjusted by their coefficients[6]:

$$\frac{\partial \hat{y}}{\partial x_i} = \frac{-\hat{\alpha}_i}{(\hat{\alpha}_0 + \hat{\alpha}_1 x_1 + \cdots + \hat{\alpha}_i x_i + \cdots + \hat{\alpha}_k x_k)^2} \tag{5}$$

This is a desired property. Because institutions and norms (which are our independent variables) do not influence the outcome (magnitude of consensus) in isolation. They permanently and implicitly interact with each other as they generate the magnitude of consensus. Indeed Equation 5 captures this interaction in an intuitive way: the effect of each institution is tamed by the effects of other institutions. Thus we estimate 'constant' effects, which incorporate this feature, by equating all x_i's to their overall mean:

$$\frac{\partial \hat{y}}{\partial x_i} = \frac{-\hat{\alpha}_i}{(\hat{\alpha}_0 + \hat{\alpha}_1 \bar{x}_1 + \cdots + \hat{\alpha}_i \bar{x}_i + \cdots + \hat{\alpha}_k \bar{x}_k)^2}. \tag{6}$$

But when we examine the way in which a variable x_i is influenced by an explicitly interacting variable x_k we evaluate

$$\frac{\partial \hat{y}}{\partial x_i} = \frac{-(\hat{\alpha}_i + \hat{\alpha}_{k+1} x_k)}{(\hat{\alpha}_0 + \hat{\alpha}_1 \bar{x}_1 + \hat{\alpha}_2 \bar{x}_2 .. + \hat{\alpha}_i x_i + \hat{\alpha}_k x_k + \hat{\alpha}_{k+1} x_i x_k)^2} \tag{7}$$

which reveals the impact of x_i on consensus as x_i and x_k change simultaneously.

5 Multivariate analysis I: model specification

As depicted in Fig. 1, in some countries one can identify several consensuses each of which is centered on a different actual norm. This observation hints the way in which macro level consensus(es) emerge: institutional conditions influence all subgroups in a slightly different way and together with the specific characteristics of these groups (such as group's actual

[6] It might be useful to reveal this more explicitly in a simple model let $\frac{1}{y} = \alpha_0 + \alpha_1 x_1$. This can be rewritten as $y = \frac{1}{\alpha_0 + \alpha_1 x_1}$ then we proceed with the usual definition of derivative as limit: $\frac{\partial y}{\partial x_1} = \lim_{\Delta x \to 0} \left(\frac{\Delta y}{\Delta x} \right) = \lim_{\Delta x \to 0} \left(\frac{y(x+\Delta x) - y(\Delta x)}{\Delta x} \right) = \lim_{\Delta x \to 0} \left[\frac{1}{\Delta x} \left(\frac{1}{\alpha_0 + \alpha_1 (x_1 + \Delta x)} - \frac{1}{\alpha_0 + \alpha_1 x_1} \right) \right] = \lim_{\Delta x \to 0} \left[\frac{1}{\Delta x} \left(\frac{\alpha_0 + \alpha_1 x_1 - \alpha_0 - \alpha_1 x_1 - \alpha_1 \Delta x)}{(\alpha_0 + \alpha_1 x_1 + \alpha_1 \Delta x)(\alpha_0 + \alpha_1 x_1)} \right) \right] = \lim_{\Delta x \to 0} \left[\frac{1}{\Delta x} \left(\frac{-\alpha_1 \Delta x}{(\alpha_0^2 + \alpha_0 \alpha_1 x_1 + \alpha_1 x_1 \alpha_0 + (\alpha_1 x_1)^2 + \alpha_1 \Delta x \alpha_0 + \alpha_1 \Delta x x_1)} \right) \right] = \lim_{\Delta x \to 0} \left[\frac{1}{\Delta x} \left(\frac{-\alpha_1 \Delta x}{(\alpha_0^2 + 2\alpha_0 \alpha_1 x_1 + (\alpha_1 x_1)^2 + \Delta x (\alpha_1 \alpha_0 + \alpha_1 x_1)} \right) \right] = \lim_{\Delta x \to 0} \left[\frac{1}{\Delta x} \left(\frac{-\alpha_1 \Delta x}{(\alpha_0 + \alpha_1 x_1)^2 + \Delta x (\alpha_1 \alpha_0 + \alpha_1 x_1)} \right) \right] = \lim_{\Delta x \to 0} \left[\frac{-\alpha_1}{(\alpha_0 + \alpha_1 x_1)^2 + \Delta x (\alpha_1 \alpha_0 + \alpha_1 x_1)} \right] = \frac{-\alpha_1}{(\alpha_0 + \alpha_1 x_1)^2}$ $\frac{\partial y}{\partial x_1} = \frac{-\alpha_1}{(\alpha_0 + \alpha_1 x_1)^2}$ hence is the outcome in the main text.

norm), they generate a particular consensus for each subgroup which are, due to overlaps between subgroups, aggregate into few (or single) consensus values at macro level.

Thus, in order to create a multivariate model which would explain the influence of institutional factors on consensus, one should take into account the fact that these factors generate different magnitudes of consensus for different subgroups which, in turn, generate the aggregate consensuses. We argue that, by estimating leveling-coefficient densities for a set of partly overlapping and sufficiently distinct subgroups, one may scan the range of consensus values prevailing in a society and approximate to the variance which generates the aggregate level consensus(es). Therefore, as a first step in our analysis, we choose the following 16 partly overlapping sets (*constituting sets*), estimate their leveling-coefficient densities and find two values for each set, that is, magnitude of consensus in the set: SET_c and actual norm of the set: SET_a

5.1 Constituting sets:

A={individuals with university education} $\xrightarrow{generates}$ $\boldsymbol{A}_c$ and $\boldsymbol{A}_a$| **B**={male employees with full time job} $\xrightarrow{generates}$ $\boldsymbol{B}_c$ and $\boldsymbol{B}_a$| **C**={individuals with secondary education who are employed in part-time jobs} $\xrightarrow{generates}$ $\boldsymbol{C}_c$ and $\boldsymbol{C}_a$| **D**={all male persons} $\xrightarrow{generates}$ $\boldsymbol{D}_c$ and $\boldsymbol{D}_a$| **E**={individuals with university education who are employed in full time job} $\xrightarrow{generates}$ $\boldsymbol{E}_c$ and $\boldsymbol{E}_a$| **F**={ individuals who are not trade union members} $\xrightarrow{generates}$ $\boldsymbol{F}_c$ and $\boldsymbol{F}_a$| **G**={individuals in the age between 18 and 25} $\xrightarrow{generates}$ $\boldsymbol{G}_c$ and $\boldsymbol{G}_a$| **H**={ individuals in the age between 25 and 40} $\xrightarrow{generates}$ $\boldsymbol{H}_c$ and $\boldsymbol{H}_a$| **I**={individuals above 40} $\xrightarrow{generates}$ $\boldsymbol{I}_c$ and $\boldsymbol{I}_a$| **J**={individuals residing in urban areas} $\xrightarrow{generates}$ $\boldsymbol{J}_c$ and $\boldsymbol{J}_{a,}$| **K**={individuals residing in sub-urban areas} $\xrightarrow{generates}$ $\boldsymbol{K}_c$ and $\boldsymbol{K}_a$| **L**={individuals residing in rural areas} $\xrightarrow{generates}$ $\boldsymbol{L}_c$ and $\boldsymbol{L}_a$| **M**={ individuals with lowest income (4th quarter)} $\xrightarrow{generates}$ $\boldsymbol{M}_c$ and $\boldsymbol{M}_a$| **N**={individuals with low income (3rd quarter)} $\xrightarrow{generates}$ $\boldsymbol{N}_c$ and $\boldsymbol{N}_a$| **P**={individuals with high income (2nd quarter)} $\xrightarrow{generates}$ $\boldsymbol{P}_c$ and $\boldsymbol{P}_a$| **R**={ individuals with highest income (1st quarter)} $\xrightarrow{generates}$ $\boldsymbol{R}_c$ and $\boldsymbol{R}_a$.

Obviously, as mentioned above, these constituting sets contain idealized subgroups and the actual subgroups in society always emerge at the intersections of (some of) these sets.

For example, let ϕ be the set containing individuals who are above 40, with a university education, residing in urban areas and having a high (2nd quarter) income. It is clear that $\phi = I \cap A \cap J \cap P$ (see Fig. 3). Thus, the magnitude of consensus in ϕ as well as the actual norm is quite possibly contained within the range of consensus and actual norm values of the intersecting constituting sets, namely $\phi_c \in [\min(I_c, A_c, J_c, P_c), \max(I_c, A_c, J_c, P_c)]$ and similarly $\phi_a \in [\min(I_a, A_a, J_a, P_a), \max(I_a, A_a, J_a, P_a)]$ (see Fig. 4).

More importantly, if one also includes the largest magnitude of consensus and the corresponding actual norm at the aggregate level (i.e. in Fig. 1) into the sequence of norm and consensus values generated by the constituting sets in each country (as approximations to the consensus and norm values of the entire society or universal set, that is, $\mathrm{S}_c$and S_a: see two small crosses in Fig. 4 and "S:{entire society}" in Fig. 3) then, it is possible to argue that consensus and norm values of some subgroups which are not directly deducible from the constituting sets may still be contained by intervals that could be generated by the norm and consensus values of the constituting sets and S_c and S_a values.

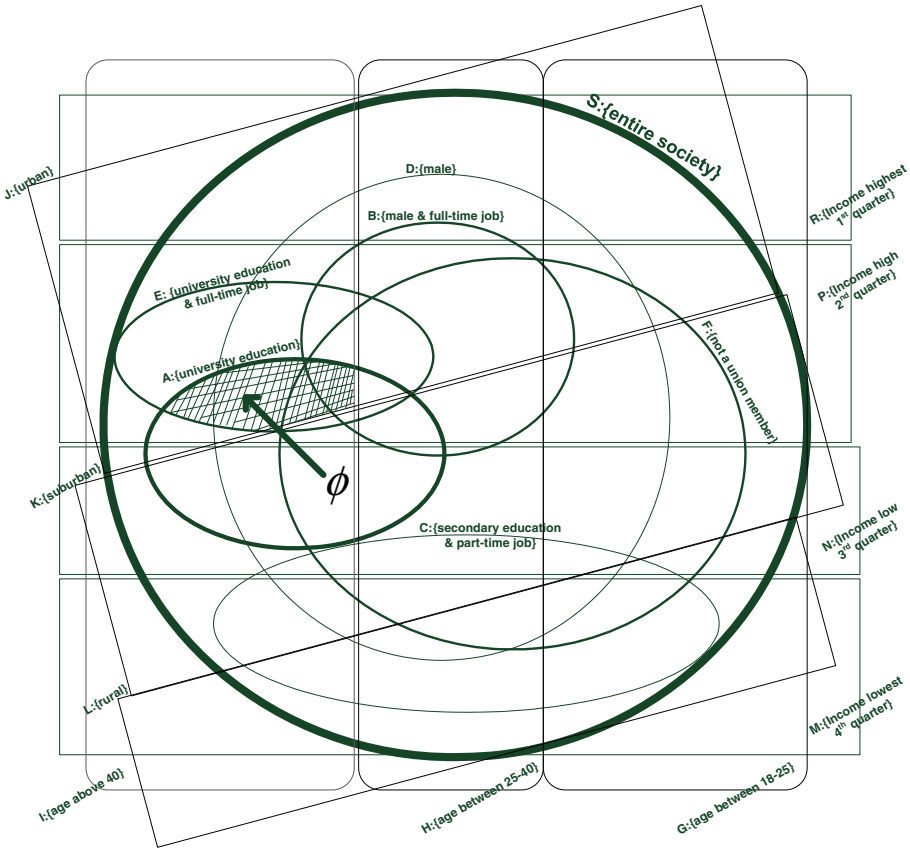

Fig. 3 Constituting Sets and their intersections

For example, the subgroup within ϕ which only contains females (ϕ^W) equals to $I \cap A \cap J \cap P \cap D^c$. Obviously, D^c(complement of D: all females) is not among the constituting sets and thus we don't have D_c^c and D_a^c values. However, given that$D^c \subset S$, one may still claim that, at least in some countries, $\phi_c^w \in [\min(I_c, A_c, J_c, P_c, S_c), \max(I_c, A_c, J_c, P_c, S_c)]$and, $\phi_a^w \in [\min(I_a, A_a, J_a, P_a, S_a), \max(I_a, A_a, J_a, P_a, S_a)]$.

On the basis of this logic we argue that the norm and consensus values generated by our 16 constituting sets together with $S_c$and S_a may scan the entire range of norm and consensus values of subgroups (that emerge at least in some intersections of constituting sets) that generate the aggregate level values in at least some countries.

Thus, for each of the 15 countries which are included in this analysis (see Fig. 1) we estimate separate leveling-coefficient densities for 16 subgroups captured by our constituting sets together with the largest consensus magnitude at aggregate level S_c and corresponding aggregate norm value S_a in each country. Then, we plot these norm values (1-x values: $\boldsymbol{S}_a, \boldsymbol{A}_a, \boldsymbol{B}_a, \boldsymbol{C}_a, \boldsymbol{D}_a, \boldsymbol{E}_a, \boldsymbol{F}_a, \boldsymbol{G}_a, \boldsymbol{H}_a, \boldsymbol{I}_a, \boldsymbol{J}_a, \boldsymbol{K}_a, \boldsymbol{L}_a, \boldsymbol{M}_a, \boldsymbol{N}_a, \boldsymbol{P}_a, \boldsymbol{R}_a$) and 1/consensus values ($\boldsymbol{1/S}_c, \boldsymbol{1/A}_c, \boldsymbol{1/B}_c, \boldsymbol{1/C}_c, \boldsymbol{1/D}_c, \boldsymbol{1/E}_c, \boldsymbol{1/F}_c, \boldsymbol{1/G}_c, \boldsymbol{1/H}_c, \boldsymbol{1/I}_c, \boldsymbol{1/J}_c, \boldsymbol{1/K}_c, \boldsymbol{1/L}_c, \boldsymbol{1/N}_c, \boldsymbol{1/P}_c, \boldsymbol{1/R}_c$) obtained from all of these densities into a *norm & consensus*$^{-1}$ space. This generates 255 data points which are depicted in Fig. 5.

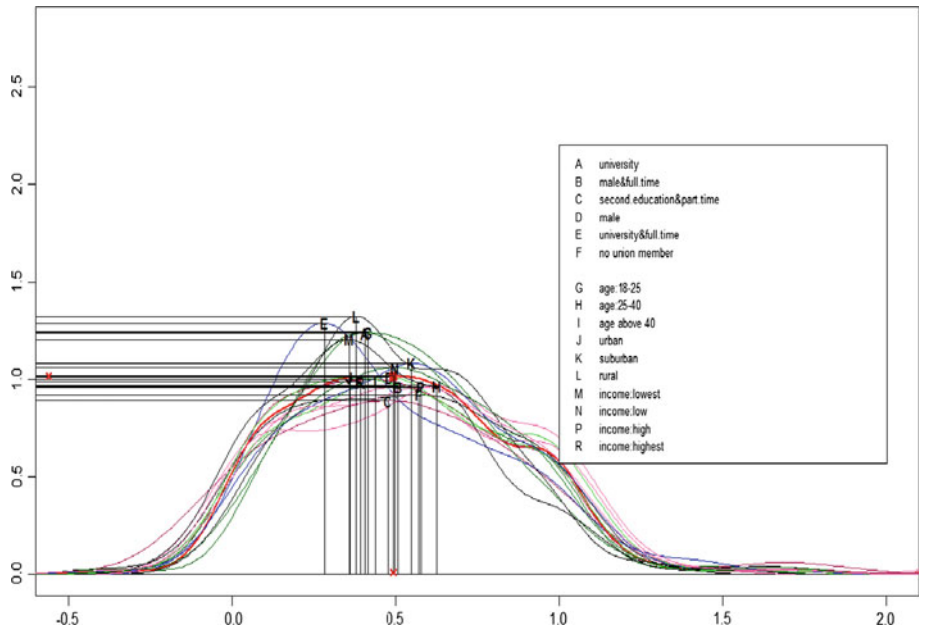

Fig. 4 Leveling Coefficient densities and norm & consensus values of constituting sets in Canada

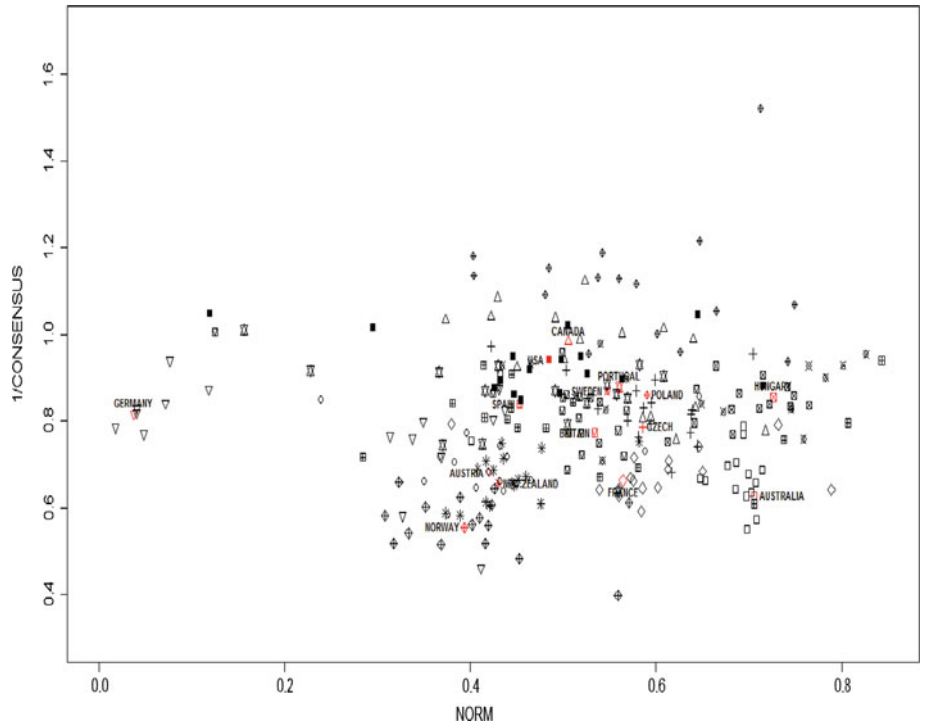

Fig. 5 Norm & Consensus^{-1} space, (note: each symbol specifies the country that is associated with it once)

The values in y-axis in this space will be the values of our dependent variable. On the other hand the values in x-axis, that is, norm values, are our loaded identifiers: they identify and differentiate each subgroup/constituting set uniquely. All group related factors that influence the magnitude of subgroup consensuses are assumed to be loaded into these identifiers. In this way a set of institutional values together with these distinctive norms generate different values for the dependent variable for each subgroup.[7] This is intuitively sound because different subgroups in a society (such as men and women) are influenced in different ways from an identical set of institutions (such as wage determination).

However, one may still argue that we cannot ensure that there would be a reasonable number of intersections between our substituting sets meaningfully filled in all countries. For neither distinctiveness nor overlap of the constituting sets do remain identical across countries, therefore in one country they may reflect the subgroups (whose norm and consensus values generate the aggregate level) quite well while in another country they may not (the problem of *under-representativeness*). Similarly, some constituting sets may reflect splinter groups in some countries thus they may have too much influence (the problem of *over-representativeness*) in any multivariate analysis (imagine, for example, the effect of points in Fig. 5 which remain apart from the main cloud). *Therefore, our choice of constituting sets (actually any particular choice[8]) is, to some extent, arbitrary* and may cause under-representativeness and/or over-representativeness. In order to overcome these problems and to strengthen our analysis we devised an iterative modeling approach, namely instead of using all 255 available values for a single model, we construct 1000 models and use the central values of the coefficient and probability distributions generated by the most reasonable subset of these 1000 models in order to build decisive models. We argue that the procedure that we use solves under-representativeness and over-representativeness problems and allows us to approximate to the reality as it is captured by most reasonable intersections of constituting sets. This procedure is as follows: **[I]** Select 181 observations from the norm/consensus^{-1} space at random. **[II]** Estimate $\frac{1}{y} = \alpha_0 + \alpha_1 x_1 + \cdots + \alpha_i x_i + \cdots + \alpha_k x_k + \varepsilon$ by OLS model (and also by GLM)[9] by regressing consensus^{-1} values in this set against the corresponding institutional variables and group norms. **[III]** Estimate Cook's distances and eliminate the entry with the largest value (to solve over-representativeness). **[IV]** Re-estimate the model with remaining 180 observations and collect the coefficients, their standard deviations and corresponding null probabilities. **[V]** Repeat first four steps 1000 times and generate 1000 models. **[VI]** Choose those models whose residual distribution approach normality quite well

[7] There are 15 distinct values for each institutional variable and 255 distinct norm values. These 15 institutional values, as they interact with norms, generate 255 distinct realizations of the dependent variable. Thus, one may construct a "255 dimensional hyperspace" for residuals.

[8] It should be noted that we undertook the same analysis with a smaller number of constituting sets (6 instead of 16) and obtained results which lead to same interpretations and significance structure with different numerical values. We argue that by increasing the number of constituting sets we fine tune the impact of institutional structures by covering their impact on more distinctive sub-groups.

[9] It is important to note that besides OLS we also estimated each model by using GLM /gamma regression with an inverse link function. For both y and $\frac{1}{y}$ are random variables (derived from the heights of modalities) which generate outcomes which are always positive and this renders the gamma distribution more suitable than the normal distribution for them. However, for two reasons we present here the OLS based outcomes: firstly, GLM/gamma-inverse link based outcomes concur with OLS based outcomes to the extent that none of our interpretations mentioned in the text requires revising. Secondly, the overall empirical distribution of $\frac{1}{y}$ locally resembles to normal (kurtosis = 2.62, skewness = −0.02) when we disregard outcomes from the long tail. This is reasonable because the impact of the entries in the long tail is reduced due to the simulation procedure above which ensures that influential entries are excluded and rare entries have limited impact (note that the entries in the long tail are both influential and rare). GLM based outcomes can be provided on request.

(to solve under-representativeness).[10] **[VII]** Construct the distributions of coefficients and corresponding null probabilities of the selected models. **[VIII]** Find out the central values of these distributions: for coefficients use mean values and for probabilities use mode values. **[IX]** Determine the significant variables in accordance with these central values of coefficient and probability distributions. **[X]** Report the outcomes.

6 Multivariate analysis II: outcomes

Table 1 summarizes three different multivariate models each emerged out of separate implementation of the procedure we outlined above.

In Model I we only look at the main effects of our independent variables on consensus. Here, union density, gini coefficient and differentiation have negative and significant effects but vocational orientation and coordination have positive and significant impacts. As hypothesized above, vocational orientation has a positive and significant effect on consensus despite the fact that we included coordination and union density in our model which might have some shadow effects that could be mistakenly attributed to vocational orientation if left out of the model. On the other hand, differentiation has significant negative impact on consensus, a finding which hints that differentiation of education system may be detrimental for social cohesion. In Model II we also included the interactions between gini coefficient and education related independent variables in order to ensure that findings of Model I remain robust when the degree of inequality in society is allowed to directly interfere with the education system. As depicted in Table 1 neither signs nor significance of independent variables of Model I do change in Model II. Only there is an increase in the null probability of differentiation variable. The important thing here is that inequality in society seems to have a significant and negative dynamic with vocational orientation as illustrated by the interaction between gini coefficient and vocational orientation.

Finally, in Model III we also introduced direct interactions between inequality in society and two political economy variables, namely, coordination and union density, in order to be sure that it is not the implicit existence of these dynamics which shape the impact of education related variables. The outcome is revealing: now differentiation is not significant anymore suggesting that its effect in previous models may be attributed to the interaction between political economy and inequality. Once these implicit dynamics are included, differentiation in education system becomes negatively significant only when it interacts with inequality as illustrated by the significant interaction between gini coefficient and differentiation. However, from the perspective of our inquiry, the interesting thing is that vocational orientation remains positively significant in Model III despite the inclusion of other interactions and it is still negatively interacting with the inequality in society. This finding is of crucial importance for our analysis, therefore, we also reveal its meaning visually by allowing vocational orientation and gini coefficient to simultaneously vary while keeping all other variables at their mean values (see equation 7). 3D and 2D graphs in Fig. 6 show the outcome: the relationship between vocational orientation and inequality is drawn from three different angles in 3D and this reality is projected into 2D to assure clarity: one can see that vocational ori-

[10] For selecting best approximating models we used a two-criterion procedure: first, we draw a perfect normal distribution onto the plot of residual distribution. At each x-point we estimate the absolute value of the difference between y values of perfect normal density and residual density. We, then, estimate the mean value of this difference. We specify the range of this mean for 1000 models and select those models which fall into initial 1/5 of the range. Second, from this initial selection we choose those models with $|\text{skewness}| < 0.1$ and $2.9 < \text{kurtosis} < 3.1$.

Table 1 Multivariate Models

Constant effect $\frac{\partial \hat{y}}{\partial x_i} = \frac{-\bar{\hat{\alpha}}_i}{(\bar{\hat{\alpha}}_0+\bar{\hat{\alpha}}_1\bar{x}_1+\cdots+\bar{\hat{\alpha}}_i\bar{x}_i+\cdots+\bar{\hat{\alpha}}_k\bar{x}_k)^2}$	MODEL I 37 models' average R^2: 0.73			MODEL II 23 models' average R^2: 0.76			MODEL III 13 models' average R^2: 0.83		
Variables	Degrees of freedom of constituting models : 173			Degrees of freedom of constituting models : 171			Degrees of freedom of constituting models :169		
	$\frac{\partial \hat{y}}{\partial x_i}$	Probability	Significance	$\frac{\partial \hat{y}}{\partial x_i}$	Probability	Significance	$\frac{\partial \hat{y}}{\partial x_i}$	Probability	Significance
Intercept	0.005	0.512		0.008	0.167		0.026	0.000	**
Union density	−0.006	0.000	**	−0.007	0.000	**	−0.015	0.004	**
Differentiation	−0.144	0.000	**	−0.300	0.077	*	0.092	0.258	
Vocational orientation	0.105	0.000	**	0.591	0.000	**	0.478	0.000	**
Norm	−0.074	0.467		−0.105	0.128		0.006	0.865	
Gini Coefficient	−0.035	0.000	**	−0.034	0.000	**	−0.027	0.000	**
Coordination	0.005	0.060	.	0.002	0.275		−0.089	0.000	**
Gini Coefficient × Vocational orientation				−0.015	0.002	**	−0.013	0.000	**
Gini Coefficient × Differentiation				0.004	0.425		−0.006	0.026	*
Gini Coefficient × Coordination							0.003	0.000	**
Gini Coefficient × Union density							0.000	0.036	*

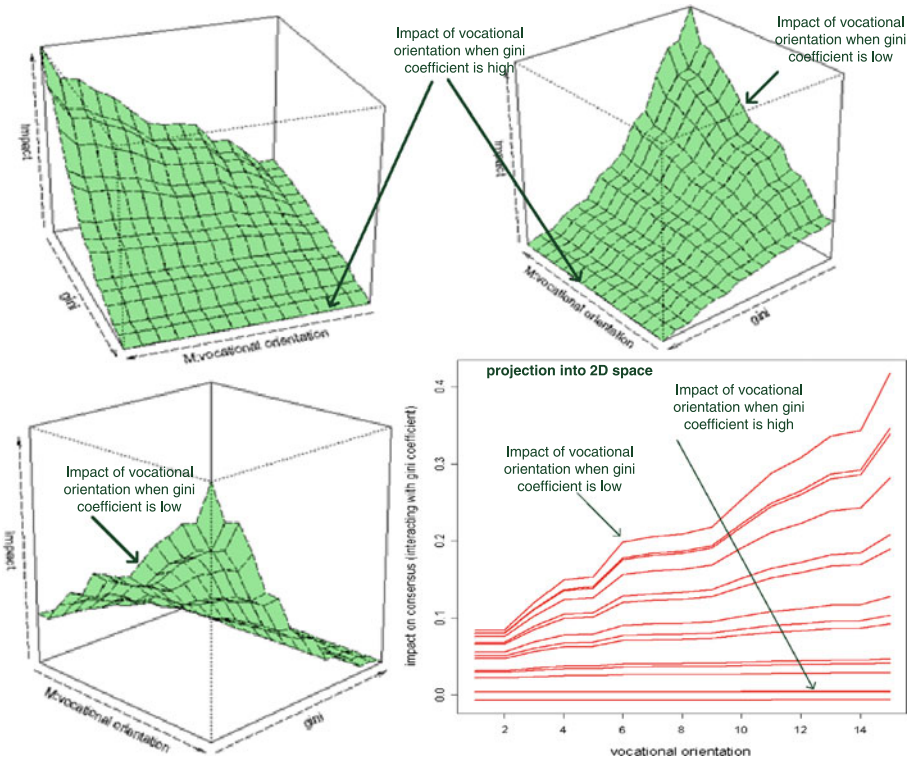

Fig. 6 Visualized interaction between vocational orientation and gini coefficient

entation (designated with M to indicate that it is the main effect) has increasingly positive impact on consensus so long as the inequality is low. However, as the inequality increases the effect of vocational orientation decreases. As depicted in Fig. 6 at 2D representation, increasing inequality first causes a decline in the rate of increase in the impact of vocational orientation on consensus and at some point this impact becomes a constant and finally it is reduced almost to zero when inequality is at its highest.

7 Discussion and conclusion

The multivariate analysis supports our expectation that education system is an important factor that influences the magnitude of consensus that guide individuals as they rank and legitimize inequalities. This validates our central claim that the sense of justice in a society is significantly shaped by the education system.

Two institutional features of education systems, that is, differentiation and vocational orientation, appear to have opposite effects on consensus; while the former has negative impact which vanishes when political economy variables interact with factual inequality, the latter has positive and consistently significant impact which remains robust.

We also see that the impact of education system is strongly conditioned by factual inequality and regulatory institutions: for example, vocational orientation is an important factor in the information of consensus but its effect is gradually tamed as the inequality increases. We

think that this indicates a possibility of institutional failure; strong vocational orientation, by facilitating smooth transition from school to work and providing equal opportunities, may enhance consensus and there by ensure social cohesion. However, under the conditions of high inequality, performing these two functions would be difficult, and consequently, education system, despite the strength of its vocational orientation, would fail to generate or enhance social cohesion. Similarly, the differentiation component also appears to be closely linked with the dynamics of inequality but also with those of regulatory institutions: strong differentiation implies a narrow set of opportunities and has a negative impact for consensus. Although this effect may disappear in highly coordinated economies, our analysis shows that high inequality may regenerate the negative impact of differentiation on social cohesion.

In short, our analysis shows that, a harmonious society cannot be built solely through economic institutions; the way in which education is institutionalized is also very crucial.

However, it is necessary to note that, high levels of inequality appears to prevent the formation of a strong consensus on justice, and thus, seriously undermines social cohesion despite all institutional arrangements.

Appendix A Basic data

Distinct elements Intersecting subsets	Australia 1064 ↓	Austria 667 ↓	Canada 732 ↓	Czech 1179 ↓	France 1191 ↓	Germany 467 ↓	Hungary 560 ↓	N.Zealand 759 ↓	Norway 252 ↓	Poland 775 ↓	Portugal 863 ↓	Spain 212 ↓	Sweden 889 ↓	Britain 540 ↓	USA 606 ↓	Total distinct →10756 Same Set totals
University	216	138	155	236	241	89	112	158	85	162	181	71	184	108	124	2260
Male.fulltime	217	130	143	239	243	91	104	147	46	148	173	42	179	101	112	2115
Secondary.parttime	71	31	37	77	77	19	25	40	17	41	53	12	54	24	27	605
Male	495	293	328	546	551	198	236	339	109	344	396	97	408	228	261	4829
University.fulltime	89	46	51	101	102	20	32	53	19	53	67	18	70	30	37	788
No.union	314	208	230	336	338	163	183	234	102	238	265	82	271	180	188	3332
Age: 18–25	181	173	175	186	186	157	163	175	139	175	177	102	179	161	165	2494
Age: 25–40	220	124	141	236	241	84	98	149	27	154	171	25	173	96	110	2049
Age: >40	644	356	401	736	742	216	287	420	79	431	500	79	521	271	319	6002
Urban	567	363	393	627	637	255	297	406	156	416	465	136	475	286	328	5807
Suburban	320	193	215	355	357	143	167	227	71	231	257	56	266	159	177	3194
Rural	104	58	67	116	116	29	45	69	7	70	80	6	85	45	50	947
Income: first	298	262	132	309	234	115	144	158	63	383	167	72	193	241	149	2920
Income: second	279	146	258	283	591	125	137	222	63	6	410	49	271	51	178	3069
Income: third	230	103	175	300	133	109	138	223	63	191	143	27	184	135	137	2291
Income: fourth	257	156	167	287	233	117	140	137	63	195	143	64	241	113	142	2455
Different Set totals	4502	2780	3068	4970	5022	1930	2308	3157	1109	3238	3648	938	3754	2229	2504	55913

For each individual in "distinct elements" row the following equation is estimated
$Y^{ought} = \beta_0 + \beta_1 Y^{is}$
The resulting 10756 coefficients re-appear several times in various cells (thus generating intersections) and create 55913 entries. From this basic data we calculate the magnitudes of 255 consensuses presented in Fig. 5

References

Alesina, A., Angeletos, G.M.: Fairness and redistribution. Am. Econ. Rev. **95**(4), 960–980 (2005)
Alves, W.M., Rossi, P.H.: Who should get what? Fairness judgments of the distribution of earnings. Am. J. Sociol. **84**(3), 541–564 (1978)
Breen, R.: Explaining cross-national variation in youth unemployment: market and institutional factors. Eur. Sociol. Rev. **21**(2), 125–134 (2005)
Brunello, G., Checchi, D.: Does school tracking affect equality of opportunity? New international evidence. Econ. Policy **22**(52), 781–861 (2007)
Esping-Andersen, G.: Hybrid or unique? The Japanese welfare state between Europe and America. J. Eur. Soc. Policy **7**(3), 179–189 (1997)
Fong, C.: Social preferences, self-interest, and the demand for redistribution. J. Public Econ. **82**(2), 225–246 (2001)
Hall, P.H., Soskice, D.: Varieties of Capitalism: The Institutional Foundations of Comparative Advantage. Oxford University Press, Oxford (2001)
Hanushek, E.A., Wössmann, L.: Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. National Bureau of Economic Research (NBER) Working Paper, No. W11124 (2005)
Hermkens, P.L.J., Boerman, F.A.: Consensus with respect to the fairness of incomes: differences between social groups. Soc. Justice Res. **3**(3), 201–215 (1989)
Homans, G.C.: Social Behavior: Its Elementary Forms. Revised edition. Harcourt Brace Jovanovich, New York (1974)
Jasso, G.: Self-interest, distributive justice, and the income distribution: a theoretical fragment based on St. Anselm's postulate. Soc. Justice Res. **3**(3), 251–276 (1989)
Jasso, G.: How much injustice is there in the World? Two new justice indexes. Am. Sociol. Rev. **64**(1), 133–168 (1999)
Kunovich, S., Slomczynski, K.M.: Systems of distribution and a sense of equity: a multilevel analysis of meritocratic attitudes in post-industrial societies. Eur. Sociol. Rev. **23**(5), 649–663 (2007)
Linos, K., West, M.: Self-interest, social beliefs, and attitudes to redistribution. Re-addressing the issue of cross-national variation. Eur. Sociol. Rev. **19**, 393–409 (2003)
Marshall, G., Swift, A., Routh, D., Burgoyne, C.: What is and what ought to be: Popular beliefs about distributive justice in thirteen countries. Eur. Sociol. Rev. **23**(5), 349–367 (1999)
Meyer, J.W.: The effects of education as an institution. Am. J. Sociol. **83**(1), 55–76 (1977)
Müller, W., Gangl, G.: Transitions from education to work in Europe: the integration of youth into EU labour markets. Oxford University Press, Oxford (2003)
Osberg, L., Smeeding, T.: Fair inequality? Attitudes toward pay differentials: the United States in comparative perspective. Am. Sociol. Rev. **71**(3), 450–473 (2006)
Shavit, Y., Müller, W.: From school to work: a comparative study of educational qualifications and occupational destinations. Oxford University Press, Oxford (1998)
Van de Werfhorst, H.G., Mijs, J.J.B.: Achievement inequality and the institutional structure of educational systems: a comparative perspective. Ann. Rev. Sociol. **36**, forthcoming (2010)
Wolbers, M.H.J: Patterns of labour market entry: a comparative perspective on school-to-work transitions in 11 European countries. Acta Sociologica **50**(3), 189–210 (2007)