Horizontal Equity

Roberto Galbiati, Pietro Vertova

Horizontal Equity

By Roberto Galbiati† and Pietro Vertova‡

†*European University Institute, Italy* ‡*University of Bergamo and Bocconi University, Milan*

Final version received 24 July 2006.

The principle of horizontal equity (HE) is considered as a basic principle to follow in order to design and evaluate a redistributive policy. However, the theoretical debate has not yet clarified what exactly horizontal equity is. In this paper we aim to clarify the meaning of the principle of HE and its normative content. In particular, we establish the true status of the two fundaments of the principle of HE, defining what is meant by 'equals' and then treating these 'equals' 'equally'. Our analysis brings forth a new and more appropriate definition of the principle of HE.

Introduction

The principle of non-discrimination is widely acknowledged as a fundamental normative principle in the design and implementation of public policies. In the US Constitution this principle is formalized by the 'equal protection clause', maintaining that the laws of a State must treat in the same way all individuals in equal conditions and circumstances. In the context of fiscal policies, the non-discrimination principle is identified by the principle of horizontal equity (HE), which claims that a redistributive policy should provide a fiscally equal treatment of 'equals'.

Scholars can glean much about the meaning and operationalization of HE from Richard Musgrave (1959, 1976, 1990) and from more recent works by Louis Kaplow (1989, 2000), Peter Lambert (1998, 2004) and others, which address philosophical, methodological and measurement issues. However, there are many unresolved issues still surrounding the HE notion, and more in general surrounding the non-discrimination principle. As some scholars have recently pointed out (e.g. Auerbach and Hasset 1999; Kaplow 1989, 2000), the theoretical debate has not yet clarified the properties and the normative content of this principle.

A fundamental unsolved problem concerns the exact meaning of horizontal equity. What is the principle of horizontal equity? What is its normative content? What does HE ask of scholars and policy-makers? As Auerbach and Hasset argue, 'From Musgrave (1959) on, there is a general agreement that horizontal equity is important, but little agreement on quite what it is' (1999, p. 1). A second and related question about horizontal equity concerns its normative support. Why should we care about this principle? Does it have a firmly grounded normative basis, as pointed out by Musgrave (1990), or does it have no normative support whatsoever, as suggested by Kaplow (2000)? This question calls for a conclusive answer. Indeed, as Kaplow observes, any application or measurement of HE would be meaningless or even harmful for policy-makers if HE were indeed a bad normative principle: 'Horizontal equity should not be measured and new measures of social welfare should not be deployed until we know what we are trying to measure and why' (Kaplow 2000, p. 22).

In this paper we propose a systematic analysis of the concept of horizontal equity, giving a comprehensive picture of the fundamentals and suggesting a clear-cut way to understand the relevant issues. Our contribution is twofold. First, we make clear the exact, and perhaps counter-intuitive, meaning of horizontal equity. Second, we analyse the normative issues related to HE. In the process of clarifying the concept of horizontal equity, we are indebted to some philosophical intuitions highlighted by Peter Westen (1990) in his treatise about the concept of equality.

Our basic reasoning is as follows. Any policy must necessarily select some individual characteristics to be considered as 'relevant' for the redistributive process (the so-called 'identification problem'). This selection is a sufficient condition to justify subdividing the population into groups of individuals defined as 'equals'. As we show, a policy cannot treat in a discriminatory way the individuals it defines as 'equals'. This means that 'equal treatment' of individuals defined as 'equals' is not a normative problem, but a condition that is always satisfied by any policy. The definition of 'equals' is a normative choice exactly *because* it selects those individuals who will be treated equally and those who will not. Since a certain policy may violate horizontal equity only if 'equals' are defined in a 'normatively non-appropriate' way, whereas there is no normative problem in the 'equal treatment' of 'equals' once they are properly defined, the principle of horizontal equity should be redefined as follows: 'Individuals treated equally by a policy should be those who are deemed normatively as equals.'

From this analysis it follows that the normative debate should focus on the 'appropriate' definition of 'equals' from a normative point of view, rather than on whether or not horizontal equity is a good or bad normative principle (which is the main issue tackled by Musgrave 1990 and Kaplow 2000).

Some important implications follow from this analysis with respect to the meaning of the measurement of horizontal inequity (HI). In recent decades several indices have been proposed in order to measure if and to what extent HE is violated by a certain fiscal policy or reform. (For a broad survey, see Lambert 1998.[1]) But most of these works do not pose a crucial question about the meaning of the measurement of HI. Our paper shows that that the proper role of an HI index is to measure the extent to which a policy deviates from a criterion of HE (that is to say, from a definition of 'equals') that is different from the one adopted by the policy under examination. Therefore tax analysts should clearly specify why they adopt a certain definition of 'equals' and should justify it from a normative point of view.

The paper is organized as follows. In the first section we analyse the identification problem (the definition of 'equals') and the 'treatment of equals'. In Section II we make clear the meaning of horizontal equity. Finally, in Section III we draw some concluding remarks.

I. Selection and Treatment of 'Equals'

The principle of HE claims that a policy should treat 'equals' 'equally'. A question immediately arises from this formulation: who are the 'equals'? As a matter of fact, two individuals are never 'equals' in all possible senses, but only with respect to some selected variables. This means that the definition of 'equals' requires choosing a set of variables considered as 'relevant' for the definition itself. Hence a straightforward definition of the principle of HE is the one proposed by Atkinson and Stiglitz (1980): 'The principle of horizontal equity states that those who are in all relevant senses identical should be treated identically' (p. 353). The selection of the relevant variables to define 'equals' is clearly a normative choice, as outlined also by Atkinson and Stiglitz: 'The implementation of this principle raises several issues. The first is the definition of "relevant" ... there is undoubtedly scope for differences of opinion, and historically standards have altered' (Atkinson and Stiglitz 1980, p. 353).

An example may be useful to clarify the normative problem existing in the selection of the relevant variables. Suppose that a redistributive policy must be designed for a population of three individuals (A, B and C). Individuals A and B have an income of 100, whereas individual C has an income of 50. Furthermore, suppose that individuals B and C suffer from bad health, whereas individual A has good health. Assuming that income is a relevant variable, should health be considered a second 'relevant variable'? If the answer is 'no', then health will not be considered as a variable deserving attention with regard to the redistributive aim. It follows that individuals A and B, who are characterized by the same income but different health conditions, will be considered 'equals'. Furthermore, the inequality between individuals A and B will be considered only in terms of their income, and not also in terms of their health. If instead 'health' is considered a good reason for discriminating among individuals, it will be selected as a relevant variable: A and B will not be considered as 'equals', and the inequality between A and C will be evaluated in terms of their health as well as their incomes. This example shows how the selection of the relevant variables has clear normative implications with respect to the relative fiscal treatment of the individuals in a population.

Another typical example concerns the choice of the household equivalence scale when designing an optimal income tax. As pointed out by Lambert, 'Household equivalence scales are typically invoked to identify the equals (and unequals), a value judgment which becomes centrally important for the analysis' (2004, pp. 1–2).

All in all, the choice of the relevant variables providing a definition of 'equals' is a normative choice, and any policy selecting the variables considered as 'relevant' refers explicitly or implicitly to a normative criterion. Now, since in the design implementation or evaluation of any policy the normative problem to select 'equals' is inevitably faced, one should follow the following general normative principle: the selection of 'equals' should be made considering all and only those variables that are 'normatively appropriate' for the redistributive process. Obviously, different value judgments may suggest different ideas of what 'normatively appropriate' means. Below we will better understand the normative meaning of the selection of 'equals'.

We can now provide a formal definition of 'redistributive policy' and analyse how a redistributive policy treats individuals defined as 'equals'. We distinguish between a 'theoretical' policy and an 'actual' one. By 'theoretical policy' we mean the one emerging from a given design of the government (i.e. the formal rules as they are written in codes and laws) and needing to be implemented; by 'actual policy' we mean the one emerging from the actual implementation of a theoretical design. This distinction is relevant as long as the actual outcomes differ from the desired theoretical ones because of wrong or illegal behaviour of the individuals that a given policy addresses (e.g. tax evasion) and/or of the bureaucratic system that implements the designed policy.

Suppose that at time 0 a theoretical redistributive policy is designed by the government in a community of $N \geqslant 2$ individuals. At time 0 each individual i, where $i = 1, \ldots, N$, can be identified on the basis of the values assumed by her Q 'individual characteristics', where $Q \in \mathcal{N}$ and $\mathcal{N}$ is the set of natural numbers. Then we can define Q sets $X_q \subseteq \Re$, where $q = 1, \ldots, Q$ and $\Re$ is the set of real numbers, each composed of all the values assumed by the qth characteristic in the population at time 0. Moreover, we define the set of these sets: $X = \{X_1, \ldots, X_Q\}$. A *theoretical redistributive policy* is a double selection made by the government:

1. the selection of a set of *relevant variables*, i.e. Z observable individual characteristics, $Z \leqslant Q$, that the government considers as relevant for the redistributive process; according to this selection, a subset of X is defined $X^Z = \{X_1, \dots, X_Z\}$;[2]
2. the selection of a *redistributive function*[3] f: $D \to C$, where $D = X_1 \times \cdots \times X_Z$, $C = \hat{X}_1 \times \cdots \times \hat{X}_Z$ and $\hat{X}_z \subseteq \Re$ $(z = 1, \dots, Z)$ is the set of all the values (or probability distributions of the values) 'theoretically'[4] assumed by the zth characteristic in the population at time 1, that is to say after the design of the redistributive policy.

Hence a theoretical redistributive policy can be seen as a selection of two distinct elements: some 'relevant variables' and a 'redistributive function'.[5] Both selections result explicitly or implicitly from the formal laws and codes as designed by the government. Now suppose that at time 0 a theoretical redistributive policy (as defined above) is implemented by the government. We define as 'actual redistributive policy' a double selection emerging from the implementation of a theoretical policy:

1. the selection of a set of *actually relevant variables*, i.e. $S \in \mathcal{N}$ individual characteristics, where $S \leqslant Q$, as 'actually relevant' when a given theoretical policy is implemented. The set of the S 'actually relevant' characteristics may differ from the set of the Z characteristics resulting as relevant in the theoretical design. According to this selection, a subset X^s of X is defined;
2. the selection of an *actual redistributive function* g: $D' \to C'$, where $D' = X_1 \times \cdots \times X_S$, $C' = \hat{X}_1 \times \cdots \hat{X}_S$ and $\hat{X}_s \subset \Re$, with $s = 1, \dots, S$, is the set of all the values actually assumed by the sth characteristic in the population at time 1, that is to say after the implementation of a given theoretical redistributive policy.

Given the selection of the relevant variables, two individuals characterized by the same values of these selected variables are defined by the policy as 'equals in all relevant senses', whereas two individuals differing in one or more characteristics are defined as 'unequals'. This means that, whatever relevant variables are selected, two individuals defined as 'equals in all relevant senses' are represented by the same element in the domain of the redistributive function. Hence, by the same definition of function, two individuals selected as 'equals in all relevant senses' cannot be discriminated between. This means that any policy must necessarily treat 'equally' the individuals it defines as 'equals in all relevant senses'.

This result could appear counterintuitive. Indeed, it seems plausible that a redistributive policy can treat individuals defined as 'equals' in a discriminatory way. However, the following simple reasoning will persuade the sceptics. Suppose that a given redistributive policy defines two individuals as 'equals in all relevant senses' according to a certain selection of the relevant variables, and suppose *ad absurdum* that these individuals reach two different outcomes after the design of the policy. How can this redistributive policy discriminate between them? Which individual will obtain one outcome and which individual the other? The only possible answer is that this policy is discriminating on the basis of one or more initial characteristics that are not considered in the selection of the variables according to which they are considered as 'equals'. However, by selecting a different set of relevant variables, the redistributive policy is defining 'equals' in a different way with respect to the initial definition (*reductio ad absurdum*). It follows that a redistributive policy can discriminate only between two individuals defined as 'unequals'.[6] The same result is pointed out by Westen (1990) in his philosophical analysis of the concept of equality: 'Every rule treats "equals" "equally" because every rule necessarily prescribes identical treatment for the persons it defines as identical' (Westen 1990, p. 125).

One could object to this conclusion by referring to the case of a random policy, e.g. random assignments of income transfers to individuals. Indeed, one could argue that in this case two individuals defined as 'equals' could reach two different outcomes *ex post* because of different random draws from the lottery implemented by the redistributive policy. But also in this case, the individuals discriminated by the random policy cannot be logically defined as 'equals'. Indeed, a random policy includes among the relevant individual variables the random draws obtained by the individuals from the lottery. Therefore, two individuals identical in all the other relevant senses but differing in their lottery outcomes are defined as 'unequals', and for this reason can be discriminated between.

In conclusion, any redistributive policy (whether theoretical or actual) treats 'equally' the individuals it selects as 'equals in all relevant senses'. Obviously, when a theoretical redistributive policy is implemented, it is possible that the actual outcomes differ from the desired theoretical ones. This means that an actual policy can treat differently 'equals in all relevant senses' as defined by the theoretical policy. It is straightforward that this is possible only if the set of the actually relevant characteristics is different from the set of the theoretically relevant characteristics.

II. The Meaning of Horizontal Equity

When is horizontal equity violated by a policy? This question is quite ambiguous. Indeed, the classical definition of HE states that 'equals' should be treated 'equally', but it does not specify if such 'equals' are those defined by the policy under examination or the 'normatively appropriate equals'. However, our previous analysis suggests that the only meaningful way to think of horizontal inequity is the latter. Indeed, if we assume that 'equals' are those defined by the policy under examination (hence considering as 'solved' the related normative problem), there is no way to obtain horizontal inequity. In such a case, horizontal equity should be considered as a property always satisfied by any policy and not as a normative principle. It is instead meaningful to adopt a different view of horizontal inequity, summed up in the following statement:

> Horizontal equity is violated by a policy when 'equals', as defined by selection of all and only the normatively relevant individual characteristics, are discriminated between by that policy.

We can now understand the conditions under which a policy generates horizontal inequity. Suppose that there exists a certain definition of 'equals' deriving from the 'normatively appropriate' selection of the relevant variables. Suppose that all the other definitions of 'equals' are 'non-appropriate' from a normative point of view. In the previous section we argued that any policy treats 'equally' the individuals it defines as 'equals'. This means that a certain policy (theoretical or actual) can treat in a discriminatory way the 'normatively appropriate equals' only if that policy provides a different definition of 'equals'. Therefore, since horizontal inequity can emerge only if 'equals' are defined by selecting a 'normatively non-appropriate' set of individual characteristics, the only normative problem existing in the 'equal treatment' of 'equals' consists in the definition of 'equals'.

Notice that the selection of the 'normatively appropriate variables' is not a strictly necessary condition for a policy to be horizontally equitable. Suppose that a policy considers as relevant all and only the normatively appropriate variables y. In this case it is possible that some 'normatively appropriate equals' will be treated unequally by the policy. But this may not be so; for example, it is possible that not all individuals differ with respect to the variable y. However, it is worth pointing out that, if we want to guarantee that 'equals' as defined in a normatively appropriate way are treated 'equally' by a policy, regardless of the distribution of the characteristics in the population or the redistributive function, then the selection of the 'normatively appropriate relevant variables' becomes a necessary condition as well. Obviously, it is possible that a theoretical policy does not violate horizontal equity but that the corresponding actual policy does so. This can happen only if the definition of 'equals' provided by the actual policy is different from the theoretical definition and the latter is normatively appropriate whereas the former is not. Nevertheless, the contrary is also possible. Take for example an apartheid regime, where black people are fiscally discriminated against but they refuse to pay taxes that exceed the amount levied on white people. In such a situation one would probably agree that selection of the 'actually relevant' variables is more justified from a normative point of view than selection of the theoretically relevant ones.

The previous analysis allows us to argue that the principle of HE as traditionally stated (that 'equals should be treated equally') is misleading. Indeed, it presumes that there are two normative problems: the first (which is generally underestimated) is the definition of 'equals'; the second is that of horizontal equity, would consist in the 'equal treatment' of 'equals'. However, as we have shown above, once the first problem is solved, the second disappears. Indeed, once the 'normatively appropriate equals' are selected by a certain policy, the individuals so defined cannot be discriminated between: this policy is necessarily horizontally equitable. Therefore, the only normative problem concerns the definition of 'equals', and the only meaningful normative claim is that 'equals' should be defined by selecting all and only the 'normatively appropriate' characteristics.

Let us now point out the normative meaning of the selection of 'equals'. The logical result obtained in the previous section, i.e. that individuals defined as 'equals' by a policy are necessarily treated 'equally' by that policy, can be re-formulated as follows: the individuals that a policy treats 'equally' are those identified as 'equals'. Hence the normative content of the selection of 'equals' (i.e. of horizontal equity) is clear: when a policy selects the 'relevant variables', it is choosing which individuals have to be treated 'equally' and which individuals can be discriminated against.

In conclusion, the normative problem of horizontal equity consists in the definition of individuals that are to be treated 'equally' and not in the 'equal treatment' of 'equals'. Accordingly, the principle of HE should be redefined as follows:

> The individuals treated equally by a policy should be those who are deemed normatively equals.

III. Concluding Remarks

This paper aims to clarify the meaning of horizontal equity. Our first claim is that selection of the variables providing a definition of 'equals' is a normative choice. Furthermore, since individuals defined by a policy as 'equals' are always treated 'equally' by that policy, the 'equal treatment' of 'equals', once they are defined in the normatively appropriate way, is not a normative problem. However, a policy may create horizontal inequity when, defining its 'equals' in a normatively non-appropriate way, it treats differently 'equals' defined in the appropriate way. This means that the principle of horizontal equity should be redefined as follows:

> The individuals treated equally by a policy should be those who are deemed normatively equals.

Our analysis allows us to understand the meaning of the use of an HI index. Technically, such an index could be useful in determining whether, and to what extent, a given redistributive policy discriminates between two individuals considered as 'equals' according to a selection of the relevant variables *different from* the selection made by that redistributive policy. Alternatively, an HI index may be used to measure whether and to what extent the actual redistributive policy discriminates between 'equals' as defined by the theoretical redistributive policy, and whether and to what extent a different implementation would decrease or increase this measure.

Concerning the normative meaning of an HI index, the analysis developed above suggests that an HI index reveals a true violation of horizontal equity only if the definition of 'equals' given by the policy under examination is not 'appropriate' from a normative point of view, whereas the definition of 'equals' provided by the researcher is appropriate. Therefore, a researcher aiming to test whether and to what extent horizontal equity is violated by a policy should justify from a normative point of view (i) why the definition of 'equals' provided by the theoretical (or actual) policy is not appropriate, and (ii) why the definition of 'equals' proposed by the researcher is normatively appropriate. It follows that the measurement of horizontal equity is not only a statistical issue but also a normative problem. Moreover, our reasoning suggests that another sensible empirical exercise may consist of inferring the definition of 'equals' implicit in a certain redistributive process in order to evaluate its normative appropriateness.[7]

Our analysis shows that the normative debate should focus on the problem of selecting the individuals a policy should treat 'equally' and not about whether or not a policy should treat the 'equals' equally (which is a meaningless view of horizontal equity). As Peter Westen points out,

> It follows that the controlling question will always be 'which rule, of the many available rules, is the one by which people ought to be measured and compared?' That, however, is a question for which answers must be sought not in conceptions of equality (which coincide with, rather then precede, the formulation of rules), but in external theories of justice and justification. (Westen 1990, p. 125)

ACKNOWLEDGMENTS

We are largely indebted to Peter Lambert for important suggestions and discussions on various drafts of the paper. For valuable comments we are grateful to Richard A. Musgrave, Marianna Belloc, Simone D'Alessandro, Massimo D'Antoni, Carlo Devillanova, Dora Kadar, Michele Di Maio, Francesco Drago, Marc Fleurbaey, Matteo Rizzolli, Alessandra Rossi, Ernesto Savaglio, Stefano Vannucci, Alberto Zanardi, seminar participants at the Universities of Lisbon (Nova), Pavia, Siena and Rome (LaSapienza) and two anonymous referees.

NOTES

1. New indices have been recently proposed by Lambert and Ramos (1997), Dardanoni and Lambert (2001) and Auerbach and Hasset (2002).
2. For notational convenience, and without loss of generality, we assume that the variables selected as relevant are ordered as the first Z among the Q individual characteristics.

3. This function may be deterministic or random. It is random when there exists at least one redistributed characteristic that is a random variable in the co-domain, so that the elements of the co-domain are random vectors. The redistributive function is deterministic when it is not random.
4. These values are only 'theoretically' and not 'actually' assumed. Indeed, any theoretical policy needs to be implemented, and the outcome resulting from the implementation may differ from the theoretical one.
5. In the context of taxation, this means the choice of a tax base and the selection of a tax schedule, respectively.
6. This result is not surprising: any social discrimination is based on the belief that some individuals are morally different because of one or more characteristics (e.g. race, sex, ideology).
7. We thank an anonymous referee for suggesting this point.

REFERENCES

ATKINSON, A. B. and STIGLITZ, J. E. (1980). *Lectures on Public Economics*. New York: McGraw-Hill.

AUERBACH, A. J. and HASSET, K. A. (1999). A new measure of horizontal equity. NBER Working Paper no. 7035.

——— and ——— (2002). A new measure of horizontal equity. *American Economic Review*, **92**, 1116–25.

DARDANONI, V. and LAMBERT, P. J. (2001). Horizontal equity comparisons. *Social Choice and Welfare*, **18**, 799–816.

KAPLOW, L. (1989). Horizontal equity: measures in search of a principle. *National Tax Journal*, **42**, 139–54.

——— (2000). Horizontal equity: new measures, unclear principle. John M. Olin Center for Law, Economics and Business, Discussion Paper no. 279, Harvard Law School.

LAMBERT, P. J. (1998). Horizontal inequity: some new perspectives. CES Working Paper no.162, University of Munich.

——— (2004). Income taxation and equity. *Baltic Journal of Economics*, **4**, 39–54.

——— and RAMOS, X. (1997). Vertical redistribution and horizontal inequity. *International Tax and Public Finance*, **4**, 25–37.

MUSGRAVE, R. A. (1959). *The Theory of Public Finance*. New York: McGraw-Hill.

——— (1976). Optimal taxation, equitable taxation or second best taxation. *Journal of Public Economics*, **6**, 3–16.

——— (1990). Horizontal equity, once more. *National Tax Journal*, **43**, 113–22.

WESTEN, P. (1990). *Speaking of Equality: An Analysis of the Rhetorical Force of 'Equality' in Moral and Legal Discourse*. Princeton, NJ: Princeton University Press.