

Preferences as a Determinant of the Optimal  
Level of Decentralisation in Health Care  
Resource Allocation  
Theoretical Insights and an Empirical Application

Carlota Maria Miranda Quintal

Submitted for the Degree of PhD

to the University of York

Department of Economics and Related Studies

July 2006

## ABSTRACT

From an economic point of view, decentralisation is expected to increase social welfare through better matching of service delivery to preferences. The latter have been a central piece of the economic rationales for decentralisation but only indirectly. Thus, at the theoretical level, the main question addressed in this dissertation is: might preferences in themselves influence the impact of decentralisation on allocative efficiency, in the context of health care resource allocation?

Regardless of which model (public choice theory or *principal-agent* theory) is used to explain the positive outcome above mentioned, in any case, the benefits generated by decentralisation depend on the assumption of variation in preferences across jurisdictions. However, there is little empirical evidence regarding this matter. Consequently, at the empirical level, the main question addressed in the current work is: does geographic variation in preferences, in the context of health care resource allocation, exist? To answer this question we developed and administered the same questionnaire (eliciting preferences) to two independent samples drawn from two Portuguese municipalities.

Within our framework, central and local decision-makers are seen as alternative *agents* acting on behalf of local populations. Given the different capabilities possessed by *agents*, decentralisation of resource allocation generates some trade-offs between objectives. Depending on the trade-offs that local populations are willing to make, they will be better-off with one or the other *agent*. Therefore, we conclude that the specific preferences held by individuals might also determine in themselves whether or not decentralisation is optimal, when compared to centralisation.

Concerning the empirical work, the principal conclusion is that the results do not corroborate the hypothesis of geographic variation in preferences, meaning that the theoretical discussion about the impact of decentralisation on allocative efficiency might have to be revisited. The empirical results further suggest that the geographical dimension of (in)equality in treatment matters to people and that a maximum opportunity cost of equality, in terms of health gain foregone, is likely to exist.

# LIST OF CONTENTS

ABSTRACT.....	2
LIST OF CONTENTS .....	3
LIST OF FIGURES .....	6
LIST OF TABLES .....	8
ACKNOWLEDGEMENTS .....	9
AUTHOR'S DECLARATION .....	10
CHAPTER 1. Introduction.....	11
CHAPTER 2. Review of literature: the core concepts of decentralisation, efficiency and preferences; theory and evidence about the impact of decentralisation on efficiency; and the evidence about preferences .....	20
2.1. Introduction.....	20
<i>Methodology</i> .....	21
2.2. The concept of decentralisation .....	22
2.3. The concept of efficiency.....	25
<i>Welfarism versus extra-welfarism</i> .....	26
<i>Health maximisation</i> .....	27
<i>Allocative efficiency and social justice</i> .....	28
2.4. The impact of decentralisation on efficiency .....	35
<i>Review of theoretical arguments</i> .....	36
<i>Review of theoretical arguments applied to the context of health care</i> ....	38
<i>Review of empirical literature applied to the context of health care</i> .....	45
2.5. The concept of preferences in the context of health care.....	49

<i>Classifying preferences according to whom respondents are asked to think about</i> .....	50
<i>Classifying preferences according to what respondents are asked to think about</i> .....	53
<i>Individual versus collective preferences</i> .....	55
2.6. What matters to people when allocating health care resources? A review of empirical evidence .....	56
<i>Preferences related to distributive justice</i> .....	56
<i>Preferences related to procedural justice</i> .....	59
<i>Preferences over attributes of health care delivery</i> .....	60
2.7. Main findings and gaps in the literature.....	61
CHAPTER 3. Might preferences in themselves influence the impact of decentralisation on allocative efficiency? A theoretical discussion.....	64
3.1. Introduction.....	64
3.2. The basic framework and assumptions .....	66
3.3. The trade-off between information regarding technical knowledge and information regarding preferences .....	73
3.4. Preferences for equality of health care provision between jurisdictions ...	80
3.5. Preferences for voice or participation in decision making.....	84
3.6. Preferences affecting the size of local health care budgets: preferences for health versus non-health-related ‘enjoyment’ .....	89
3.7. Is it optimal to allocate resources between jurisdictions based on local preferences? .....	93
3.8. Discussion .....	98
CHAPTER 4. Do preferences vary across regions? An empirical study applied to two Portuguese municipalities .....	105
4.1. Introduction.....	105
4.2. The Portuguese context.....	106
4.2.1. Some demographic and health indicators .....	106

4.2.2. The health care system.....	107
4.3. Methods.....	112
4.3.1. The questionnaire.....	112
4.3.2. The respondents.....	126
4.3.3. The analysis.....	131
4.4. Results.....	133
4.5. Discussion.....	145
CHAPTER 5. A discussion of ambiguities and confusions affecting the debate about decentralisation in the English NHS.....	150
5.1. Introduction.....	150
5.2. Brief characterisation of the NHS.....	151
5.3. A summarised account of decentralisation in the NHS: from Conservatives (1979-1997) to Labour.....	156
5.4. The ambiguous use of the concept of preferences in the context of the NHS .....	159
5.5. Variations in service provision: the confusion between variations in preferences and variations in technical information.....	162
5.6. Ambiguities regarding the role attached to participation in the NHS.....	168
5.7. Concluding comments.....	171
CHAPTER 6. Conclusion.....	173
APPENDIX A –Maps of Portugal.....	179
APPENDIX B – Questionnaire: English version.....	180
APPENDIX C – Questionnaire: Portuguese version.....	190
REFERENCES.....	200

## LIST OF FIGURES

Figure 1. <i>Notions of efficiency and equity</i> .....	28
Figure 2. <i>An adaptation of the principal-agent approach</i> .....	71
Figure 3. <i>Gain in social welfare due to information asymmetry about local preferences</i> .....	74
Figure 4. <i>Change in social welfare due to information asymmetry about preferences and technical knowledge</i> .....	75
Figure 5. <i>Changes in social welfare given information asymmetry and variation in preferences</i> .....	76
Figure 6. <i>'Resource constrained' clinical guidelines as a mechanism to pass technical knowledge into local decisions</i> .....	78
Figure 7. <i>Preferences for equal health care provision</i> .....	81
Figure 8. <i>Convergent effects of decentralisation across jurisdictions</i> .....	82
Figure 9. <i>Divergent effects of decentralisation across jurisdictions</i> .....	84
Figure 10. <i>The impact of preferences for participation or voice on the optimal level of participation / decentralisation</i> .....	86
Figure 11. <i>Different optimal levels of participation / decentralisation as the result of different preferences</i> .....	88
Figure 12. <i>Loss of social welfare created by a uniform level of decentralisation</i>	89
Figure 13. <i>Loss of health-related welfare as a negative 'externality effect' of local decisions</i> .....	92
Figure 14. <i>The impact of different preferences about health states on health gain measurement</i> .....	94
Figure 15. <i>Deriving points of indifference from trade-off questions</i> .....	122

Figure 16. <i>Bar chart of HMAX by region of residence</i> .....	133
Figure 17. <i>Bar chart of HGEO by region of residence</i> .....	133
Figure 18. <i>HMAX and HGEO by region of residence, with three categories</i> .....	134
Figure 19. <i>Bar chart of QUEST3 by region of residence</i> .....	135
Figure 20. <i>Bar chart of QUEST4 by region of residence</i> .....	135
Figure 21. <i>Bar chart of TOFF3 by region of residence</i> .....	136
Figure 22. <i>Bar chart of TOFF4 by region of residence</i> .....	138
Figure 23. <i>Bar chart of DOCGEND by region of residence</i> .....	140
Figure 24. <i>Bar chart of CHOICE by region of residence</i> .....	141
Figure 25. <i>Bar chart of THRESH61 by region of residence</i> .....	142
Figure 26. <i>Bar chart of THRESH62 by region of residence</i> .....	143
Figure 27. <i>The NHS in England</i> .....	152

## LIST OF TABLES

Table 1. <i>Variables designation and codification</i> .....	126
Table 2. <i>Respondent characteristics in sample 1 - Coimbra (N1=35)</i> .....	129
Table 3. <i>Respondent characteristics in sample 2 - Bragança (N2=35)</i> .....	130
Table 4. <i>HMAX and HGEO by region of residence: summary statistics and hypothesis test results</i> .....	134
Table 5. <i>TOFF3 and TOFF4 by region of residence: summary statistics and hypothesis test results</i> .....	137
Table 6. <i>Implied <math>r</math> and implied trade-off for TOFF3 and TOFF4</i> .....	139
Table 7. <i>Implied <math>r</math> and implied trade-off by region of residence: mean values and hypothesis test results</i> .....	140
Table 8. <i>Fisher exact test for DOCGEND and CHOICE</i> .....	141
Table 9. <i>THRESH61 and THRESH62 by region of residence: summary statistics and hypothesis test results</i> .....	143
Table 10. <i>Tests for within-respondent differences in each sample</i> .....	144



## ACKNOWLEDGEMENTS

I am grateful to my supervisor, Anthony J. Culyer, for thoughtful comments on the successive drafts that led to the current work and for drawing my attention to points of interest to the general debate about decentralisation in health care. I am also grateful to Peter C. Smith for his comments on initial versions of the theoretical discussion. This dissertation has also benefited from helpful recommendations of the examiners Paul Dolan and Richard Cookson.

For their pertinent suggestions about the questionnaire design I would like to thank Aki Tsuchiya and Pedro Ferreira. I would also like to thank the precious help received from all that filled, helped me to administer and gave their opinion about the questionnaire. I would thus like to thank Ana Abrunhosa, Carlos Lopes, Francisco Lopes, Nuno Cid, Patrícia Sá, Patrícia Silva, Paula Lopes and Rita Martins. I am grateful to Óscar Lourenço for his helpful comments about the empirical chapter.

Naturally, I retain full responsibility for opinions expressed and remaining errors.

I would like to acknowledge funding from the *Fundação para a Ciência e Tecnologia* as well as the institutional support received from the Faculty of Economics, University of Coimbra.

I am grateful to my parents, sister, brothers and friends for their support during this journey. A deep thank to my mother-in-law for looking after my children during school holidays.

Finally, I would like to thank those that share the day-to-day life with me, Carlos, Catarina and João, for their love, despite my mental and physical unavailability to be with them, in so many occasions. A special mention here to my son João for his companionship during our 3-year stay in York.

## AUTHOR'S DECLARATION

Part of the content of section 3.5 of this dissertation corresponds to the paper “Public preferences as a determinant of the optimal level of decentralisation in health care”, presented in the conference, *69th Meeting of the UK Health Economists' Study Group*, York-UK, 26-28 July, 2006.

Part of the empirical work, reported in chapter four, will be presented in the *10<sup>th</sup> National Conference of the Portuguese Health Economics Association*, Lisbon, 22-24 November, 2007, with the title: “Health maximisation *versus* equality of treatment: What is the optimal trade-off from the general public's perspective?”.

## CHAPTER 1. Introduction

Decentralisation in the health care sector has received, since the early 1990s, an increasing attention from researchers and policy makers. The World Bank's *World Development Report*, in 1993, dedicated to health, has certainly contributed to strengthen discussions about decentralisation by classifying the latter as “potentially the most important force for improving efficiency and responding to local health conditions and demands” (cf. Reich, 2002, p.1672). But the debate has not been restricted to less developed countries. Saltman and Bankauskaite (2004), for example, say that the concept of decentralisation has become a cornerstone of health policy in an increasing number of Western European countries.

The phenomenon of decentralisation has been addressed by various disciplines and it is difficult to define but one might say that in broad terms it entails the transfer of powers from a central authority to more local institutions. Decentralisation is also associated with an array of different outcomes. Within economics, one of such (positive) outcomes is the increase in social welfare or allocative efficiency through better matching of service delivery to preferences.

There have been two major lines of thought, within economics, that explain this better matching of service provision to preferences. One, linked to public choice theory, argues that locally elected governments compete with each other for constituents (who are also taxpayers), who will choose their preferred mix of services by ‘voting with their feet’ (the Tiebout effect) and moving to jurisdictions offering services more in line with their preferences. The other, related to the *principal-agent* theory, draws attention on information. In this case, the idea is that local authorities, by living closer to the people that they serve, have a better knowledge and understanding of the preferences held by these people, compared to the central authority. The concept of preferences has therefore played a relevant role in economic analyses of decentralisation but only indirectly. That is, what lead to the positive outcome above mentioned is information asymmetry (about preferences) or competition (among local decision makers) and not preferences themselves.

The nature and content of preferences have not in fact been given attention

in the literature about decentralisation. In this literature, the concept of preferences has been implicitly interpreted as consumers' preferences. However, a much broader interpretation of this concept has been proposed in the health economics literature, recognising that preferences might be society regarding and not only centred on the self. Therefore, at the theoretical level, the main question addressed in this dissertation is: are the nature and content of preferences relevant for the evaluation of the impact of decentralisation on social welfare? In other words, might preferences in themselves influence the impact of decentralisation on allocative efficiency, in the context of health care resource allocation?

Regardless of which model (public choice theory or *principal-agent* theory) is used to explain a better matching between provision and preferences, in any case, the benefits generated by decentralisation depend on the extent to which there is variation in preferences across jurisdictions. Despite the relevance of the assumption of variation in preferences, there is little empirical evidence regarding this matter, for the context of health care. Consequently, at the empirical level, the main question addressed in the current work is: does geographic variation in preferences, in the context of health care resource allocation, exist?

We start nonetheless with a systematic review of literature, whose main findings are reported in chapter two. The part of the review of literature corresponding to section 2.4 not only gives an account of the current state of the art regarding the theoretical and empirical literature about the impact of decentralisation on efficiency but it also serves as the basis to set up the framework underlying the theoretical discussion developed in chapter three. But because both concepts of decentralisation and efficiency beg for some clarification about their meaning, before reviewing the literature about the link between the two, we present in sections 2.2 and 2.3 the interpretations of these two concepts which have been mostly used in economics and health economics.

A possible taxonomy classifies decentralisation by three categories of devolved responsibilities: political, administrative and fiscal. In the first case, the focus is on the transfer of powers from the centre to democratically elected representatives of local populations. Within the administrative approach, a well known typology classifies processes of decentralisation as deconcentration, delegation, devolution, and privatisation. Fiscal decentralisation focuses on the control at lower levels of government over fund-raising and expenditures. Some

have argued that political and fiscal decentralisation have a limited applicability in the context of health care (due to the small weight that locally raised resources represent and due to the existence of administrative restrictions regarding the allocation of intergovernmental transfers). Two alternative frameworks have been proposed: the ‘*decision space*’ (Bossert, 1998), which draws on the *principal-agent* relationship and is interpreted as the range of effective choice that is allowed by the central government (the *principal*) to be utilised by local authorities (the *agents*); and the ‘arrows framework’ (Peckham *et al.*, 2005) that is a two dimensional framework that distinguishes eight tiers of decision making (ranging from ‘global’ to ‘individual’) and three types of activities (inputs; process or decisions; and outcome).

Concerning the concept of efficiency, three main interpretations have been used in health economics: X-efficiency or production efficiency; health maximisation; and allocative efficiency. Given a health production frontier, X-efficiency might be increased by moving the health distribution from any point inside the frontier to any point on the frontier. This movement, however, does not necessarily lead to the maximum total health. Thus, a popular interpretation of efficiency is health maximisation, which is met in that point on the health frontier where its slope equals -1. In this case, movements on the frontier towards this point are said to increase efficiency interpreted in this particular fashion. But health maximisation does not necessarily corresponds to the maximum welfare if the latter is affected namely by distributional considerations. Theoretical arguments supporting these varied distributional considerations are reviewed in section 2.3. In this section, we follow a classification that groups principles of justice into maximising principles, egalitarian principles, and need principles. Moreover, a further distinction is made between distributive and procedural justice. In the former case, outcomes are at the centre of discussions, while in the latter case, procedures are the core issue. Some of these considerations might conflict with each other; in particular, when there is a conflict between health maximisation and other principles, it is said that an equity-efficiency trade-off exists. It might even occur that a point inside the health frontier is preferred to a point on the frontier, on the grounds of e.g. procedural concerns. Thus, in this case, the equity-efficiency trade-off might involve not only health maximisation but also production efficiency.

Bearing in mind that efficiency and decentralisation might be interpreted in various ways, section 2.4 reviews the theoretical and empirical literature regarding the impact of decentralisation on efficiency, with a special focus on the context of health care resource allocation. There are arguments that point to a positive link between decentralisation and both production and allocative efficiency but there are also arguments that point to the opposite result. Examples of the former are greater knowledge at the local level regarding local circumstances affecting production costs, greater innovation at the local level, and a better match between service provision and preferences. Examples of the latter are the loss of economies of scale, lower technical knowledge at the local level, and the existence of externalities. On the whole, there is some ambiguity regarding the positive or negative impact of decentralisation on efficiency. This ambiguity of results is also a characteristic of empirical studies that have addressed this question.

The literature about decentralisation has used the concept of preferences but this concept has not been expressly addressed within that stream of literature, in which preferences have been implicitly interpreted as consumers' preferences. However, once we look at the literature about health care resource allocation, we conclude that this interpretation is limited in the context of health care. In fact, in this dissertation the concept of preferences is used in a broad sense. Therefore, it is important to review the literature that supports this wider interpretation, which is done in section 2.5. Three main criteria that might be used to classify preferences are identified: 'whom respondents are asked to think about'; 'what respondents are asked to think about'; and 'individual *versus* collective preferences'. A major and perhaps the most important distinction is that between personal preferences and social preferences. In the former case, what matters is what individuals prefer for themselves on whatever basis, whereas in the latter case, what matters is what individuals prefer for the society as a whole. In general terms, this distinction opposes a consumer perspective to a citizen perspective.

Expanding the concept of preferences in economic analyses might lead some to question whether this is like opening Pandora's box. Consequently, in section 2.6, the empirical literature regarding preferences in the context of health care resource allocation is reviewed, showing that a link between theories that explain preferences and preferences held by the general public does in fact exist.

Many empirical studies have been conducted in England or UK, therefore, this part of the review of literature is also used to support the discussion in chapter five. In the final section of chapter two, we identify several topics that remain under-explored in the literature, some of which are addressed in subsequent chapters.

Chapter three then addresses the central theoretical question of this dissertation: might preferences *in themselves* influence the impact of decentralisation on allocative efficiency? We begin by defining the framework on which the discussion is grounded. This framework stems from the literature reviewed in chapter two and more detailed information about it as well as about the theoretical or empirical literature supporting it is provided in section 3.2. The central assumptions are: there is variation in preferences across jurisdictions; there is imperfect information regarding preferences and local authorities are better informed about these preferences (in their respective jurisdictions) compared to the central authority; the centre has superior technical information compared to local authorities; decision-makers, both at the central and local levels, are benevolent planners that seek to maximise the social welfare of local populations. This is a possible framework but we acknowledge that, based on the review of literature, information asymmetry for example can go either way. Moreover, the literature is ambiguous regarding the impact of decentralisation on allocative efficiency as well as on production efficiency. Our interest lies on the hypothesis of both impacts diverging but we recognise that given the current state of knowledge different assumptions would be acceptable as well.

Globally, our framework might be viewed as an adaptation of the *principal-agent* approach. Traditionally, the *principal-agent* analysis applied to decentralisation is focused on the relationship between the central decision-maker (the *principal*) and local decision-makers (the *agents*). In our case, by assuming that decision-makers are benevolent planners, emphasis is given to the relationship between the central and local decision-makers, on the one hand, and local populations, on the other. Both types of decision-makers are thus seen as alternative *agents* and the idea is to compare, for each local population, the relative performance of the central and local authorities, in meeting the preferences of local people (see Figure 2). Centralisation is taken as the *status quo* (which fits the perspective adopted in most Western European countries as well as

less developed countries) against which decentralisation is evaluated. In our discussion, decentralisation is said to improve allocative efficiency if it leads to a better match, compared to centralisation, between what is achieved with resource allocation and what are the preferences regarding this same allocation.

Sections 3.3 to 3.6 deal with different trade-offs between objectives of resource allocation that might emerge with decentralisation. These objectives range from production efficiency and health maximisation to equal provision of health care; participation or voice in decision making; and reduction in inequalities in health across jurisdictions. In this context, whether or not decentralisation improves welfare depends on how people trade-off the various arguments of welfare considered in each case. We thus conclude that preferences might determine the extent to which decentralisation is optimal compared to centralisation. In sections 3.3 to 3.6, preferences affect the distribution of local budgets within jurisdictions but not the distribution of the whole health care budget among jurisdictions. Therefore, in section 3.7, we consider the possibility of (heterogeneous) local preferences about health states being used to allocate a common pool of resources among jurisdictions and discuss its implications, compared to the case in which national preferences are used. We argue that distributing resources between jurisdictions based on local preferences raises some problems but, once local budgets have been set, it is optimal from the perspective of local welfare to allocate local budgets based on local preferences.

The conclusions achieved in chapter three are obviously dependent on the assumptions presented in the beginning of the chapter; hence, in the final section (section 3.8) we point the limitations that might apply, discussing the consequences of relaxing some assumptions.

One of the previous limitations has to do precisely with the assumption of variation in preferences across jurisdictions. Thus, in chapter four we report on a small-scale empirical study that sought to address this question; the main objective of this chapter is therefore to assess whether or not there are differences in preferences, in our case, between two independent samples drawn from two distinct Portuguese municipalities.

To pursue this objective we developed and administered (method of delivery and collection) the same questionnaire to both samples. The selection of regions from which to draw the samples was partly by practical convenience and



partly judgemental, being the samples defined by quotas based on the composition of local populations (age, sex and level of education). The questionnaire comprises 6 questions (for the English version of the questionnaire see Appendix B; for the Portuguese version of the questionnaire see Appendix C) and adopting the terminology reviewed in chapter two, one might say that two types of preferences are elicited: personal preferences and social preferences. Regarding the latter, we have elicited preferences concerning the criteria of health maximisation and equality of treatment (across geographic regions) as well as the trade-off between them (using person trade-off questions). The selection of preferences to test for geographic diversity followed from the review of literature in which we concluded that the geographical dimension of (in)equality has not been addressed in empirical works; thus, this seemed to be an opportunity to contribute to current knowledge by looking at people's views about the trade-off between health maximisation and geographic equality in health care provision. This sort of trade-offs might actually acquire particular relevance in contexts of decentralisation given that the latter have been associated with geographic variations in provision. Additionally, that type of trade-off is addressed in chapter three; therefore, the empirical findings could be seen as a test regarding the pertinence, from a practical point of view, of the theoretical discussion. Descriptive statistics and graphical methods are used to explore the data, giving an account of respondents' preferences and an initial overview of differences between samples. Differences are afterwards analysed using inductive statistical methods. In all tests, the null hypothesis is that both samples have been drawn from the same population (or, from populations with the same mean, median or distribution, depending on the test), that is, to conclude that differences between samples are statistically significant one should be allowed to reject the null hypothesis.

The main conclusions of the empirical study are: there are no differences in preferences across samples (given that the statistical tests showed that basically for all variables the null hypothesis could not be rejected); people are sensitive to geographic inequality in health care provision; and people are willing to make trade-offs between the principles of health maximisation and geographic equality of treatment (most people are willing to forego between 10 and 30 per cent of total health gain to keep geographic equality of treatment in return). These results

naturally apply for the case of the specific regions and for the particular topics approached in the study, which has its own limitations. These and other limitations (e.g. method of administration and size of samples) are identified in the final section of chapter four and directions to future research are suggested.

Given the various forms that concepts such as decentralisation, efficiency and preferences might embody, there has been some lack of clarity in the use of these concepts in the literature. But this lack of clarity has also characterised the political discourse about decentralisation in health care. It is therefore important to point where ambiguities arise and further clarifications are needed. The objective of chapter five is precisely to draw attention on these ambiguities for the case of the debate about decentralisation in the English National Health Service (NHS). The focus on the English NHS is justified with the fact decentralisation in health care is one of the current policy themes in England, reflected in the political discourse of ‘shifting the balance of power’. In fact, the English NHS has been the object of numerous analyses which is comprehensible given the successive reforms that it has been subjected to, in some cases representing worldwide pioneer initiatives.

We first look at the use of the concept of preferences, in the context of the English NHS, noting that both consumer and citizen perspectives apply to this context but the distinction between the two has not always been explicit. Another ambiguity or confusion that seems to exist in the English NHS and that might affect the respective debate about decentralisation is that between variations in provision due to varying technical knowledge and variations in provision due to varying preferences. In chapter five, we use the framework proposed in section 3.2 to distinguish between the two sources of variations. This question is even more pertinent in the case of the English NHS given the aversion to variations in provision that seems to exist in England. Finally, public involvement or participation in decision making is also a current theme in the English NHS and it is also a central issue associated with decentralisation. In England, there is at a rhetorical (and legal) level an association between participation and local NHS organisations. The question is that two approaches to participation have been identified (consequentialist approach and procedural approach) and some ambiguities might arise when, for example, the public and authorities have different views about participation or, when the approach adopted by authorities is

not accompanied in practice by mechanisms to implement it. In chapter five we suggest that this sort of ambiguities exists in the English NHS as well.

To conclude the chapter, we suggest that the debate would benefit if the issues that we have addressed in regards to the NHS were made clearer, particularly, in terms of policy guidance flowing from the central government to local organisations and in terms of information directed at the general public. Although our discussion is focused on the English case, these ambiguities are likely to apply to other countries as well.

Chapter six closes this dissertation with a summary of our main results and conclusions, pointing out their major limitations and the future research agenda that might succeed the current work.

## **CHAPTER 2. Review of literature: the core concepts of decentralisation, efficiency and preferences; theory and evidence about the impact of decentralisation on efficiency; and the evidence about preferences**

### **2.1. Introduction**

The phenomenon of decentralisation has been addressed by various disciplines and is associated with an array of different outcomes; in some cases, the same outcome results from different explanations depending on the discipline at stake. Given that the current dissertation focuses on the efficiency of health care resource allocation in the context of decentralisation, this chapter reviews the theoretical as well as the empirical literature regarding the impact of decentralisation on efficiency. Moreover, our interest lies on the perspective of economics and, in particular, of health economics.

The main objectives are therefore to review the theoretical literature concerning the impact (positive or negative) of decentralisation on efficiency; the arguments used to explain that impact and how the concept of preferences relates to these arguments. We start with economic arguments in general and then focus on the particular case of health care resource allocation. In addition, the review of empirical evidence (in the context of health care) aims to reach a conclusion on whether or not the theoretical arguments are supported by the evidence. This part of the review (of theoretical and empirical literature), presented in section 2.4, not only gives an account of the current state of the art but it also serves as the basis to set up the framework (section 3.2) within which the discussion is developed in chapter three.

Because social scientists often do not share common meanings even when they use common words (Culyer, 1989b), it is important to clarify the meanings of the concepts of decentralisation and efficiency, before reviewing the literature regarding the relationship between them. Consequently, sections 2.2 and 2.3 present the interpretations of these concepts mostly used in economics and health economics.

The economic arguments, presented in section 2.4, that support a positive link between decentralisation and (allocative) efficiency make use of the concept of preferences, without explicitly addressing the concept in itself. However, several distinctions have been made in the literature about health care resource allocation, namely between self-regarding and society-regarding preferences. Hence, section 2.5 reviews the main classifications of preferences found in the health economics literature. This part of the review is important to support the (wide) concept of preferences used in the theoretical discussion, in chapter three, as well as to classify preferences in the empirical work reported in chapter four. Together with section 2.3, section 2.5 also serves to show that the examples of preferences used throughout this dissertation are not arbitrary. To show in addition that they are not empirically unfounded, section 2.6 reviews the empirical evidence about what matters to people when it comes to allocate health care resources.

In chapter five, we will also recall much of the content of sections 2.3, 2.5 and 2.6, namely to point out the lack of clarity of some concepts (like values or preferences) that have been used in the context of the English NHS and to support some parts of the discussion with available evidence for this context (e.g. preferences regarding participation in decision making).

In the final section of this chapter (section 2.7), we summarise the main findings of the review of literature and identify some gaps and areas that remain under-explored in the literature, some of which are addressed in subsequent chapters.

### *Methodology*

Data sources included electronic searches of various databases (EconLit, Medline, Wilson Social Sciences Abstracts and PsycINFO) as well as the follow-up of references in selected papers. The main search was carried out in the second semester of 2004 though more recent publications are also included, which have resulted from a few updates of the literature search (the last update took place in June 2007).

The key terms used (usually in the form of combinations that included

health or decentralisation) were: health, decentralisation, efficiency, public (or social) preferences (or values), equity, justice, fairness and public participation (or involvement).

Regarding the empirical literature about decentralisation in health care, much of the evidence has been produced for the case of developing countries. Although there might be differences in terms of key issues between developing and e.g. Western European countries, we did not exclude *a priori* the studies applied to the former context. In terms of the empirical studies about preferences in the context of health care resource allocation, we excluded empirical studies that elicited preferences from a specific perspective like the views of patients or health staff (we have nonetheless retained studies that used convenience samples such as students or academic staff provided that the stated aim was to elicit views from the perspective of the general public). Within this latter literature, there has been a discussion about the pros and cons of the various methods of preference elicitation. It is not however our objective to address this question, therefore, we also excluded studies specifically designed to approach methodological issues (there is an exception for the case of the person trade-off technique but this part of the literature is reviewed in chapter four, in the context of the empirical work reported there).

## **2.2. The concept of decentralisation**

Decentralisation has been approached from diverse perspectives and by various disciplines, leading to different interpretations. As above stated, here our objective is to present the main interpretations referred to in the health economics literature, which in turn rely on the frameworks developed by economists. It should be said that it is unlikely that any single framework apply in all circumstances and it is also common to find overlaps between categories within the same framework. The main interdisciplinary consensus regarding the concept of decentralisation is perhaps the recognition that it represents a complex phenomenon. Levaggi and Smith (2005, p.225) say that decentralisation in health care is indeed difficult to define. These authors still suggest that, in broad terms,

“it entails the transfer of powers from a central authority (typically the national government) to more local institutions”.

A possible taxonomy (used by e.g. Saltman and Bankauskaite, 2004, and Hutchinson and LaFond, 2004) classifies decentralisation by three categories of devolved responsibilities: political, administrative and fiscal. The distinction among these categories is not however clear cut.

Regarding *political* decentralisation, it has been broadly defined as the shift in the public decision making power from a more-central to a more-local authority, in particular through democratic processes. It generally assumes that local decision makers are democratically elected representatives of local populations.

*Administrative* decentralisation “is the transfer of responsibility for planning, financing, and managing certain public functions from the central government and its agencies to subordinate units or levels of government, semi-autonomous public authorities or corporations, or areawide, regional, or functional authorities” (Hutchinson and LaFond, 2004, p.5, quoting Dennis Rondinelli). Saltman and Bankauskaite (2004, p.8) say that the concept of administrative decentralisation emerged in opposition to the core tenets of Weberian bureaucracy. This form of decentralisation emphasises the greater or lower operational authority of lower level functionaries.

Within the administrative approach, Rondinelli’s (1983) four-fold typology (related to the institutions to which responsibilities are devolved) has been widely used to classify processes of decentralisation: *deconcentration* - spatial relocation of decision making, i.e. the shift of power from the central offices to peripheral offices of the same administrative structure (e.g. Ministry of Health and its district offices); *delegation* – shifts responsibility and authority to semi-autonomous agencies (e.g. a separate regulatory commission or an accreditation commission); *devolution* – shifts responsibility and authority from the central offices of the Ministry of Health to separate administrative structures still within the public administration (e.g. local governments of provinces, municipalities); and *privatisation* – transfers operational responsibilities and in some cases ownership to private providers.

A distinction between vertical decentralisation and horizontal decentralisation might also be introduced here. In the former case, authority is

transferred to lower levels of decision making (this adjusts to forms of decentralisation such as delegation or devolution). In the case of horizontal decentralisation, at each level of decision making, authority might be spread out from a central point (this notion conforms to the concept of deconcentration).

The third category of decentralisation, above introduced, is *fiscal* decentralisation. This sort of decentralisation refers to the control, at lower levels of government, over financial resources, either in terms of fund-raising and/or expenditure assignments. Saltman and Bankauskaite (2004) note that the seminal literature on fiscal decentralisation appeared under the label of ‘fiscal federalism’, given its original conceptualisation to fit the characteristics of a federal state (US).

Each of these frameworks captures only some of the diverse dimensions of decision making. For this reason, each framework naturally presents weaknesses in some aspects, while performing better in other particular features. An advantage of the administrative approach, pointed by Saltman and Bankauskaite (2004), is that it acknowledges the distinction between policy and management, recognising that policy can remain partly central even as health system management is mostly decentralised.

In the case of the four-part typology presented above, a central feature is the focus on the appropriate level for decentralising functions and responsibilities. However, although there is an implicit assumption that moving from deconcentration toward privatisation is likely to increase the range of choice allowed to local officials and managers, there is not a clear link between each form of decentralisation and what is being decentralised.

Bossert (1998) argues that the main strength of the political and fiscal decentralisation approaches is their focus on choices made by local governments, developing “clear and parsimonious theoretical propositions to explain those choices” (p.1515) but he claims that this framework has a limited applicability in the health care sector, namely in Western European countries and less developed countries. The author therefore proposes his own framework, based on the *principal-agent* theory, the latter being in fact presented by this author as an autonomous framework for the analysis of decentralisation.

Initially developed to examine choices made by managers of private corporations, the *principal-agent* approach proposes a *principal* with specified objectives and *agents* who are needed to implement activities to achieve those

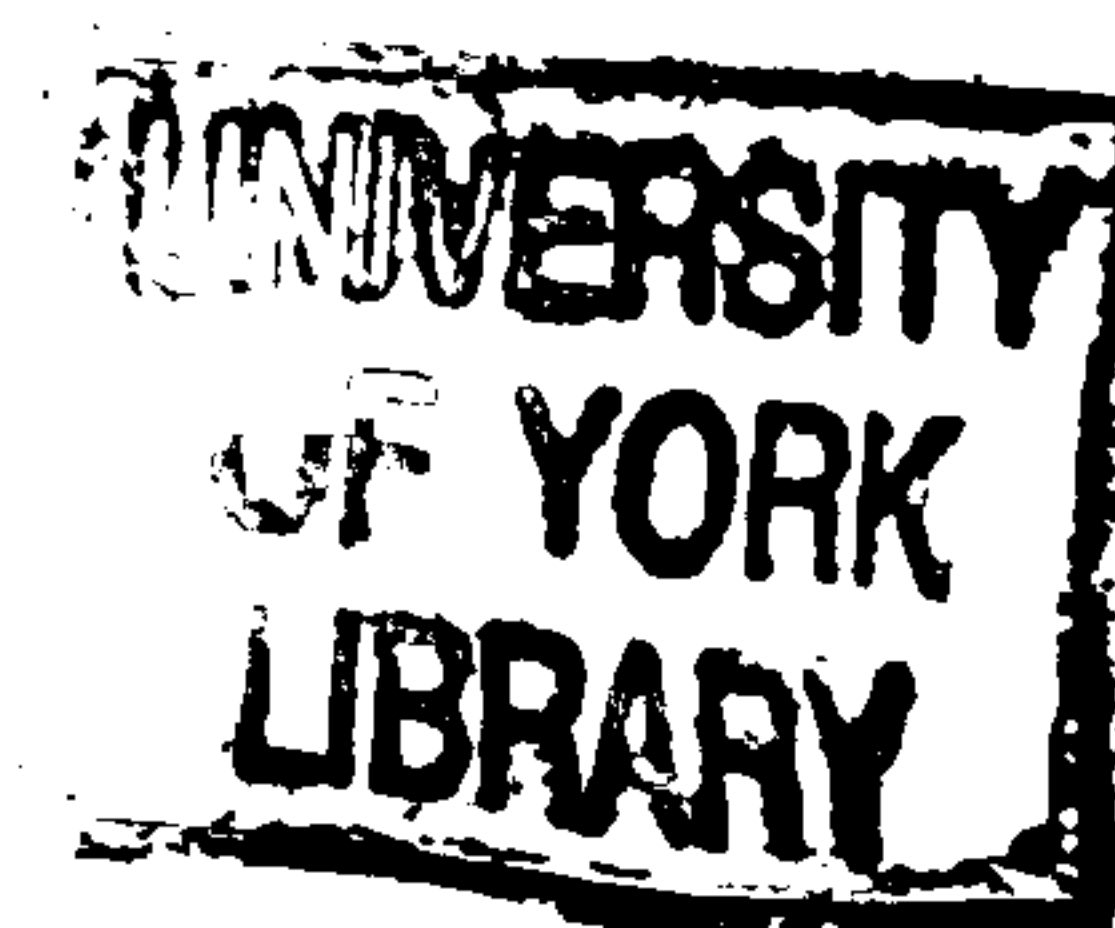


objectives. *Agents* may share some of the *principal's* objectives but they may also have additional interests and have some informational advantage over the *principal*. In order to overcome the informational asymmetry and self-interests of *agents*, the *principal* might use a scheme of incentives and sanctions. In the health sector, the Ministry of Health could be seen as a *principal*, with a range of objectives, and local authorities could be seen as *agents* who are given resources to implement general policies to achieve these objectives. Drawing on the *principal-agent* analysis, Bossert (1998) suggests the concept of '*decision space*', which is interpreted as the range of effective choice that is allowed by the central government (the *principal*) to be utilised by local authorities (the *agents*). The author considers five categories of functions (finance, service organisation, human resources, access rules and governance rules). Within each category, Bossert (1998) defines the range of choice as 'narrow', 'moderate' and 'wide'.

A framework more recently proposed by Peckham *et al.* (2005), conceptualised to fit the characteristics of the English NHS is termed by the authors as the 'Arrows Framework'. This is a two dimensional framework that distinguishes eight tiers of decision making (Global; Europe; UK; England/Scotland/Wales and Northern Ireland; Organisation, like Primary Care Trust; Subunit, like practice; and individual) and three types of activities (inputs, including funding; process or decisions; and outcome, i.e. the definition of targets). The strength of this approach is making clear 'from where to where' decentralisation occurs, including the individual as the maximum exponent of decentralisation as well as the type of decisions over which decision making power is being transferred. It also accommodates opposite movements, in terms of decentralisation/centralisation, at the same time. This framework contributes to reducing the ambiguity in the definition of decentralisation but, as the authors note, it does not establish a link between centralising or decentralising movements and performance.

### **2.3. The concept of efficiency**

In economics, one may distinguish among different kinds or levels of



efficiency. Culyer (1989a), for instance, classifies them as: technical efficiency (when physical inputs are minimised for a given level of output, or, given the amount of physical inputs, the level of output is maximized – might be defined as ‘being on an isoquant’); cost-effectiveness, sometimes termed as X-efficiency or as production efficiency (not only are physical inputs minimised for a given level of output but also the total cost is minimised – might be defined as ‘being where an isocost line is tangential to an isoquant’); and ideal output or top-level efficiency (using Culyer’s (1989a, p.35) words, this level of efficiency is achieved when cost-effective outputs are produced at a rate that is ‘socially’ optimal and allocated to individual members of society in a ‘socially’ optimal fashion, that is, when the marginal rate of (product) transformation is equal to the marginal rate of substitution).

#### *Welfarism versus extra-welfarism*

The definitions above make use of to the concept of output, which has itself been the subject of debate. That is, there is a debate about what it is that should be maximised or distributed. In this respect, there are two broad approaches that became known as welfarism *versus* extra-welfarism, two terms introduced in the literature by Sen (1977) and Culyer (1989a), respectively. Others prefer to talk about non-welfarism instead of extra-welfarism (Tsuchiya and Williams, 2001). In the context of health care, two kinds of desideratum are associated with the previous perspectives: utility and health.

Utilitarian theory is welfarist and welfarism holds that the only information that is relevant to assessing social welfare is the level of personal utility, that is, the level of utility that is perceived by the individuals as the level of their own wellbeing (Tsuchiya and Miyamoto, 2007). According to Culyer (1989a), extra-welfarism transcends traditional welfare; it does not necessarily exclude individual welfares from the judgement about the social state but it does supplement them with other aspects of individuals. Sen (1980), for instance, suggests that further information beyond utility must enter the process of comparing social states, stressing the role of ‘basic capabilities’ and ‘functionings’. A capability is the ability to achieve a certain sort of functioning,

while the space of functionings is the various things a person may value doing or being (see also Sen and Foster, 1997). Culyer (1989b) adopts the notion of ‘characteristics of people’. These characteristics may represent deprivation or desired states. Thus, commodities and characteristics of commodities are needed to remove the deprivation or to move towards the desired state.

While welfare economics literature conceptualises the social welfare function<sup>1</sup> (SWF) as a function of personal utility, the health economics literature has often defined the relevant social objective function as a function of individual health (Abasolo and Tsuchiya, 2004). Tsuchiya and Miyamoto (2006) note that “the desideratum in health economics has also been perceived to be health *per se*, not because it is valued by individuals as patients or consumers (although it may well be) but because it is valued by the public at large or the relevant decision makers”. Tsuchiya and Miyamoto (2007) conclude “social welfare in the context of public policy decision making is not a function of the utility enjoyed by constituent individuals of society as judged by themselves, but a function of social desiderata, dictated by the relevant policy context”.

### *Health maximisation*

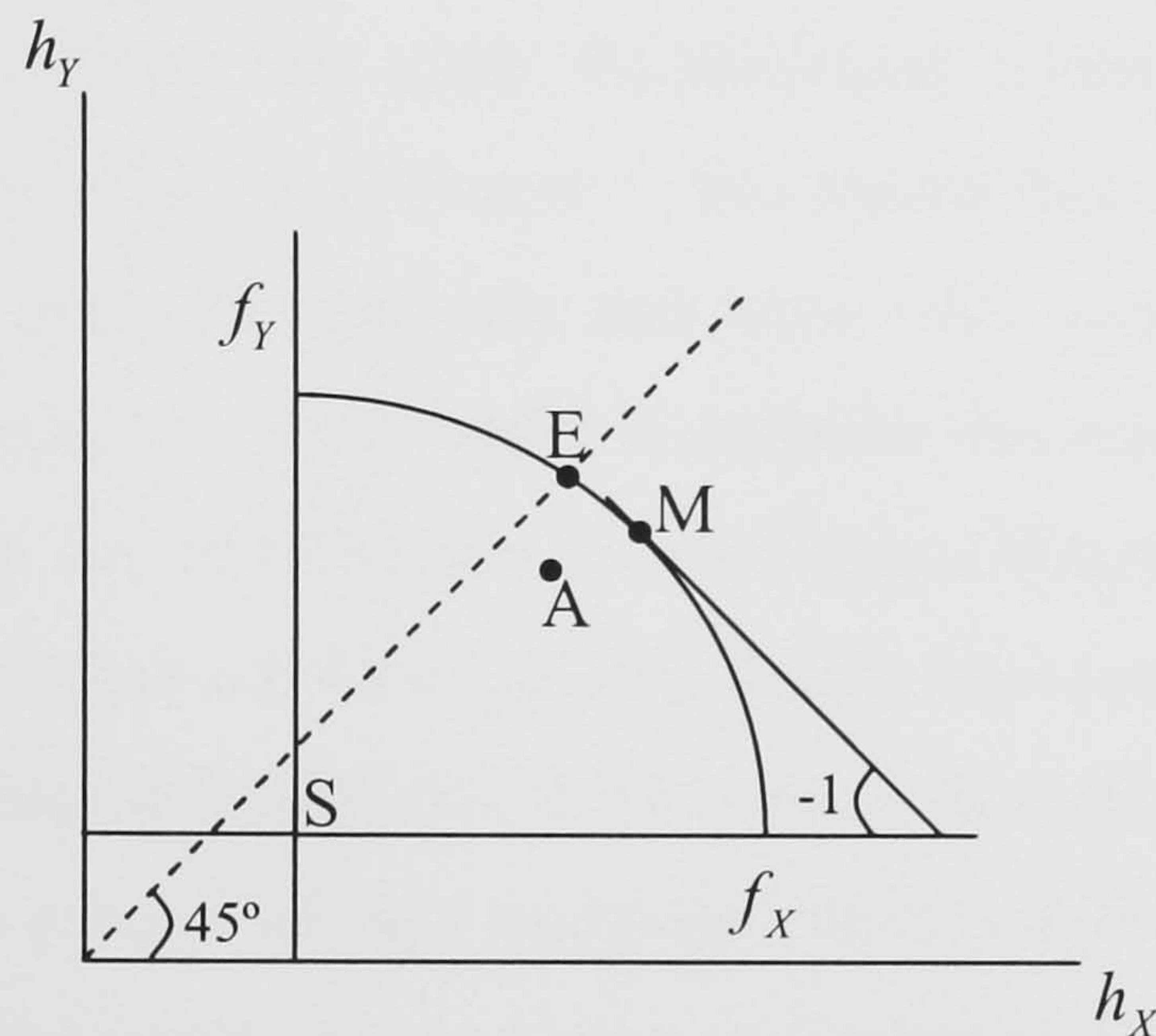
Within the predominant ‘extra-welfarist’ perspective adopted in the health economics literature, two main interpretations of efficiency can be identified: X-efficiency and health maximisation. We use Figure 1 to illustrate them.

In Figure 1,  $h_x$  and  $h_y$  cardinaly measure the health of two representative individuals or groups of individuals,  $X$  and  $Y$ . An example of a health measure is the Quality Adjusted Life Year (QALY). This measure entails two dimensions: length and quality of life. Each year of life is weighted by a quality-of-life score. This score depends on the health state associated with each year and varies typically from zero (death) to one (perfect health). In some situations, negative scores may be considered appropriate if there are health states seen as being worse than death.  $S$  is the health endowment, corresponding to the total amount of health

---

<sup>1</sup> The social welfare function is an economic device that provides a complete ordering of alternative social states. It is common to represent it in a two-dimension space by a set of indifference curves (on the conventional properties of social welfare functions, namely in the context of health, see e.g. Abasolo and Tsuchiya, 2004).

enjoyed by  $X$  and  $Y$ , in the absence of health care.  $f_Y f_X$  is the health possibility frontier and shows for a fixed level of  $X$ 's health (or  $Y$ 's health) what is the maximum amount of health that  $Y$  ( $X$ ) can obtain, given the available resources to allocate between  $X$  and  $Y$  and given the technology that transforms health care resources into health. X-efficiency can therefore be described as 'being on the health possibility frontier' and movements towards the frontier can be said to increase X-efficiency.



**Figure 1. Notions of efficiency and equity**

But efficiency in health is sometimes interpreted as 'health maximisation' (e.g. Wagstaff, 1991 and Williams, 1996, 1997b), corresponding to a particular point on the health possibility frontier,  $f_Y f_X$ . On this point, the total sum of  $X$ 's and  $Y$ 's units of health (with uniform weights attached to the health gain going for each individual) is maximised. Efficiency is met when it is not possible to further increase total health by substituting  $X$ 's health by  $Y$ 's health and vice-versa (in Figure 1, health maximisation is achieved on point M, where the slope of the frontier is -1). Further note that X-efficiency is a necessary condition to achieve health maximisation.

#### *Allocative efficiency and social justice*

Allocative efficiency or top-level efficiency is about choosing the point that represents the preferred social state. In Figure 1, the preferred social state

might coincide with point M, or it might not. It has long been acknowledged (Wagstaff, 1991) that health maximisation, as represented by point M, might not reflect all the distributional concerns which potentially determine the preferred social state. These distributional concerns have been addressed mainly under the heading of equity (sometimes the question is put as a matter of principles of justice in health care priority setting) and it is now widely accepted that equity is an important policy objective in the health care field, particularly in publicly funded health care systems.

Because accommodating other distributional concerns might imply a movement away from point M, in Figure 1, this means that a trade-off has to be made between what point M represents and what other aimed points represent. The most common approach in the health economics literature has identified this sort of trade-off as an equity-efficiency trade-off (e.g., Wagstaff, 1991; Williams 1997b; Dolan, 1998; Dolan and Olsen, 2001). Others (most notably Culyer (1995, 2006)) have nonetheless defended that there is no such trade-off as the notion of efficiency should be purged of ‘any embodied distributive equity assumptions’ and should therefore be interpreted as ‘being on the health possibility frontier’ (in which case, movements away from M but still on the frontier do not represent a departure from efficiency). But, apart from these divergences in terms of the interpretation of the equity-efficiency trade-off, all agree that top-level efficiency, i.e. the preferred social state, might depend on a range of distributional concerns beyond that of health maximisation.

There is an immense literature about social justice already applied to the context of health and health care. Although there might be diverse forms of grouping theories of justice, here we follow Cookson and Dolan (2000)<sup>2</sup>. These authors distinguish between principles of substantive justice (which are about determining who should get what health care and when) and principles of procedural justice (which are about what decision making process should be followed). They focus on substantive justice, identifying three principles more relevant to the context of health care: maximising principles (in this case, health care should be distributed in order to achieve the maximum ‘benefit’); egalitarian

---

<sup>2</sup>For a comprehensive classification of theories of equity in the distribution of health see Williams and Cookson (2000).

principles (health care should be distributed with the aim of reducing ‘inequality’); and need principles (the distribution of health care should be proportional to ‘need’). Depending on how ‘benefit’, ‘inequality’ and ‘need’ are defined, these three principles have given rise to several variants.

The most common variant of maximising principles, in the case of health and health care, interprets ‘benefit’ as the population’s health (an alternative, much less popular among health economists, would be to interpret ‘benefit’ as utility or happiness, in line with classical utilitarianism – see above discussion about welfarism *versus* extra-welfarism). We have already seen that point M, in Figure 1, conforms to maximising principles.

Egalitarian principles advocate equality but, in fact, it has been said that all theories of justice seeks equality of something and that “the real work begins with the specification of what is it that is to be equalised” (Sen, 2002, p. 660). A straightforward variant of egalitarian principles in the case of health care requires the distribution of health care resources so as to reduce inequalities in health. In Figure 1, equality of health ( $h_x = h_y$ ) is verified along the 45° line through the origin. As it can be seen, point M (health maximisation) lies off this line and, from an egalitarian perspective, it is therefore inequitable. From the set of points that form the health possibility frontier, in Figure 1, only point E meets equality of health.

Cookson and Dolan (2000) note however that most authors that defend egalitarian principles would not in fact pursue equality as a sole objective, but would rather combine equality with other goals. A combination of equality of health and, for example, health maximisation would generate a point on the frontier, in Figure 1, between M and E (see e.g. the seminal paper on the equity-efficiency trade-off, Wagstaff (1991)). A strict egalitarian position would nonetheless permit no concessions to health maximisation except as between alternatives that gave people equal levels of health. This situation is commonly represented by L-shaped indifference curves, in which there is a social improvement only if the health of the worse-off person is increased (also known as the ‘maximin’ principle). We should further note that L-shaped indifference curves have also been termed as Rawlsian indifference curves after Rawls’ theory of justice (Rawls, 1971, 1999) but, in this theory, social and economic inequalities are to be measured by an index of ‘primary goods’ (all purpose means that every

rational person needs to pursue their own ends - Williams and Cookson, 2000) in which health is not included<sup>3</sup>.

Further variants of equality of health might be identified, depending on how health is defined (see Dolan and Olsen, (2001)). One of such variants has gained considerable visibility and became known as the ‘fair innings’ argument (Williams, 1997a, 1997b), according to which everyone is entitled to a similarly long and healthy life. The implication for health care resource allocation is the prioritisation in favour of individuals in risk of failing to achieve their entitlement in terms of length and quality of life, compared to individuals that already got more than their entitlement. Generally, this principle favours the young over the old<sup>4</sup>. In the case of Figure 1, point E could represent this principle (equalisation of lifetime health) provided that  $h_x$  and  $h_y$  were measured in terms of total health (including past health).

Equality of health is not however the only kind of equality that might represent an equitable distribution. Cookson and Dolan (2000) identify a second variant of egalitarian principles which focuses on equalising people’s opportunity for lifetime health, rather than achieved levels of health to account for individual freedom of choice and autonomy in making choices that influence health. This approach introduces a difference between health inequality and health inequity; the same is to say, there might be inequalities in health that are deemed fair if, for example, they result from a personal decision not to worry about health in particular.

As Sen (2002) argues, health inequalities cannot be identified with health inequity. This author has firstly developed the capability approach (see above), precisely emphasising equality of opportunity. In the case of health, Sen (2002) distinguishes between health achievement and the capability to achieve good health (which may or may not be exercised). This author continues saying that “what is particularly serious as an injustice is the lack of opportunity that some may have to achieve health because of inadequate social arrangements” (Sen,

---

<sup>3</sup> In addition, the maximin principle following from Rawls’ theory is not driven by altruism or concern for the least advantaged members of society; instead, it is the result of risk-averse individuals choosing behind a veil of ignorance (i.e. individuals choose to maximise the position of the worst off because they do not know in advance if they will belong to this group).

<sup>4</sup> This is not the only principle that favours the young over the old: at least two other forms of ageism, ‘health maximization ageism’ and ‘productivity ageism’. have been identified in the literature (Tsuchiya *et al.*, 2003).

2002, p.660). He nonetheless acknowledges that health achievement tends to be a good guide to underlying capabilities. This might be due to the fact that ill health would rarely be entirely attributable to a person's own actions and even in the case of the so-called 'lifestyle-diseases', it is said that they are not just self-inflicted but rather caused by a mental or biological response to the environment in which the person lives (Olsen *et al.*, 2003)<sup>5</sup>. In the case of Figure 1, the egalitarian variant of equal opportunity implies that points on the frontier involving health inequality might well represent distributional concerns, which are equally defensible from an ethical point of view.

Sen (2002) further distinguishes between achievement and capability, on the one side, and the facilities socially offered for that achievement. Health care can be seen as an example (though not the only one) of these latter facilities. In fact, social justice in health and health care has also been discussed in reference to health care itself and the notion that access to health care ought to be the same for everyone is a popular distributive principle (Culyer and Wagstaff, 1993) but the concept of access too can be interpreted in various way: (i) access as utilisation of health care; (ii) access as the money and time costs incurred in receiving health care; (iii) access as the maximum attainable consumption of health care; and (iv) access as the forgone utility cost of obtaining health care.

Williams and Cookson (2000) say that the principle of equal access to health care places ethical constraints on the health possibility set, ruling out all attainable health outcomes that require unequal access to health care. Thus, in Figure 1, we might have points on the health frontier excluded from the health possibility set or we might have points inside the health possibility frontier that represent preferred social states, compared to points on the frontier, if the latter imply unequal access to health care. For example, point A in Figure 1 might be preferred over points on the frontier because in A some kind of equality of access is guaranteed. Further note that, in such cases, alternative social states might involve a trade-off where X-efficiency itself is at stake and not only health

---

<sup>5</sup> The discussion about individual responsibility has led some to claim that discrimination should occur in terms of health care financing but not in terms of health care delivery (e.g. Le Grand, 1991). Williams and Cookson (2000), for instance, accommodate Le Grand's suggestion by defining an autonomous 'theory', which they call 'equality of opportunity for health using financial means only'. Even at the finance level there might be some distinctions. Cappelen and Norheim (2006), for instance, argue that individuals should be held responsible for their choices and not for the consequences of their choices. This argument supports, they say, for example, levying taxes on tobacco instead of having cancer victims paying for their treatment.



maximisation.

There is a third set of principles of justice identified by Cookson and Dolan (2000), related to the concept of need. These authors say that distributing health care resources according to need is one of the most popular principles, namely among clinicians (perhaps because it is seen as the opposite to distributing resources according to ability to pay). But, as with the previous principles, 'need' might be interpreted in different ways, originating several variants of this principle of justice. The most common interpretation of need defines it as the degree of ill-health. Following this principle, priority in terms of resource allocation should be given to the most severely ill, in the extreme, to those facing life threatening situations. In fact, this latter case has been specifically identified as the 'rule of rescue' (the imperative people feel to rescue identifiable individuals facing avoidable death - Mckie and Richardson, 2003). These interpretations have been criticised for ignoring the expected benefit of health care (for example, a terminally ill person certainly needs health but it is hard to see how he or she needs health care if the situation is irreversible). To overcome this limitation, economists have proposed an alternative interpretation of need as capacity to benefit from health care. However, this still does not account for differences in treatment costs (for instance, if two individuals have the same capacity to benefit, according to the previous principle, both should receive the same amount of resources; however, if treatment costs vary between them, it might mean that one gets treated and the other does not). Hence, another interpretation of need that has been suggested (Culyer and Wagstaff, 1993) sees it as the expenditures required to exhaust capacity to benefit. Culyer and Wagstaff (1993) show how the allocation of resources based on the different interpretations of need can lead to different points on the health possibility frontier and how it can generate unequal distribution of health. Hence, need principles might also support the distinction between health inequality and health inequity.

Besides the consequentialist approach to social justice, which evaluates alternative social states based on outcomes, there is an alternative view according to which it matters how the different states of the world are determined. We now focus on this latter, procedural, dimension of social justice.

The works by Thibaut and Walker (1978) and Lind and Tyler (1988) are pointed as two seminal theories into procedural justice, in this case, more oriented

to the legal studies and social psychology. The former authors suggested that fair procedures influenced disputant's satisfaction with the legal system and their belief that justice has been done, regardless of the verdict, that is, regardless of whether the outcomes they received were fair or personally beneficial. In particular, these authors proposed that disputants will perceive that they have received fair treatment when they are permitted to communicate their views to those responsible for the final decision and they believe that their input facilitates their chance of obtaining desired and fair outcomes. Thibaut and Walker (1978) thus emphasised the instrumental value of having voice or control over decision making.

Lind and Tyler's (1988) group value theory in turn explains why having a 'voice' in the decision making process might also have non-instrumental value. Non-instrumentality means that even if disputants believe that communicating their views will not affect outcomes, having a 'voice' still affects their perceptions on whether they have been treated fairly. The group value model assumes that group identification is psychologically rewarding and that individuals are motivated to establish and maintain group bonds. A central claim of the theory is that being listened to is symbolically important as it reveals that group authorities value the individuals' standing in their social groups (Heuer, 2005).

In addition to voice, five other broad procedural characteristics have been identified (Wailoo and Anand, 2005; Tsuchiya *et al.*, 2005; Dolan *et al.*, 2007): neutrality<sup>6</sup> (requires that the interests of similar people within a process are treated in the same way); consistency (the roles accorded to similar people must be the same or the same decision making criteria must be applied across time and comparable decision contexts)<sup>7</sup>; accuracy (assessment of the accuracy of different types of information); reversibility (right to appeal an unfavourable decision, for example, on the grounds that there were some process violations); and transparency (requirement about transparency about what procedures are in use and whether procedural rules were followed).

Further note that in some literature, namely in the political economy and in the literature about decentralisation, in particular, the concept of participation (or

---

<sup>6</sup> In some cases, this dimension is identified as 'impartiality' (Tsuchiya *et al.*, 2005) or 'absence of vested interests' (Wailoo and Anand, 2005).

<sup>7</sup> A close relationship has been established between this dimension of procedural fairness and the notion of equality of opportunity (Leventhal, 1980, cf. Wailoo and Anand, 2005).

involvement) is usually used instead of voice. In any case, it has been acknowledged that voice or participation might take different forms. It might be interpreted in terms of informing the public or it might mean consulting the public. The highest level of participation consists of sharing the responsibility for decisions or even delegating power (see e.g. Litva *et al.*, 2002). Yet another distinction is that between a micro and a macro aspect of voice. In the former case, the individual is concerned that his or her own rights are honoured and it is about direct involvement, while in the latter case, because the expression of views by all is not feasible, the primary concern of the individual may just be that a hearing is given to some representative individuals (Anand, 2001).

There has been a gradual movement beyond consequentialism, in economics. Anand (2001) says that Amartya Sen has been the most prominent and articulate advocate of the theoretical position that outcomes are not all that matter for economic welfare (Sen, 1997, 2002). Bruno Frey and colleagues (e.g. Frey *et al.*, 1996; Stutzer and Frey, 2006) have conducted much of the empirical work into procedural preferences in economic contexts, though in this chapter we are interested in the application of procedural justice to the context of health and health care; an application that only recently has started to assume more explicit forms (e.g. Wailoo and Anand, 2005; Dolan *et al.*, 2007). In the specific context of health, Sen (2002) argues that in the face of the relevance of processes, inequalities in health care and not only in health achievement can also have relevance to social justice and to health equity. Therefore, in Figure 1, it is conceivable that points on the frontier other than point E are preferred based on procedural considerations and despite involving health inequality; it might even be the case that points inside the frontier are preferred over points on the frontier. We have said before that in these situations trade-offs between efficiency and other considerations of social justice occur not only when efficiency means health maximisation but also at the level of X-efficiency.

#### **2.4. The impact of decentralisation on efficiency**

In this section, our objective is to review the theoretical and empirical

literature about the impact of decentralisation on efficiency, in the case of health care resource allocation. Following the previous interpretations of the concept of efficiency, three forms of efficiency might be involved: X-efficiency (or production efficiency); health maximisation; and allocative efficiency. We should note nonetheless that in the literature the concept of efficiency appears in many cases with vague and ambiguous meanings. Moreover, it is acknowledged that assessing the effects of decentralisation is a challenging exercise, not only because it is difficult to distinguish between dependent and independent variables but also because the appropriate associations between them are not easily demonstrated (Bankauskaite and Saltman, 2006). Actual processes of decentralisation are most likely to be motivated by various factors and in some cases the different disciplines that have addressed the phenomenon of decentralisation offer different explanations for similar outcomes; here, our objective is to focus on arguments developed within the discipline of economics and applied to health economics.

#### *Review of theoretical arguments*

The classical theory of fiscal decentralisation is based on the core argument that public goods<sup>8</sup> that are consumed locally should also be produced locally. Decentralisation is believed to increase welfare by allowing local authorities to act in accordance with local cost structures and local preferences (Oates, 1972, 1999). Improvement in welfare might thus be interpreted in terms of both production efficiency (adjustment to local costs) as well as allocative efficiency (adjustment to local preferences). This result is based on the assumption that local authorities are better informed than central government about (heterogeneous) preferences and costs. Further, by allowing for different mixes of public services across jurisdictions, decentralisation can achieve greater allocative efficiency in the allocation of resources by forcing local governments to compete for constituents (who are also tax payers) who will choose their preferred mix of services by ‘voting with their feet’ (known as the Tiebout effect) and

---

<sup>8</sup> Goods that are *non-rival* as well as *non-excludable* (this means that consumption of the good by one individual does not reduce the amount of the good available for consumption by others; and no one can be effectively excluded from using that good). Consequently, public goods have been pointed as an example of ‘market failure’, that is, competitive markets alone cannot fully provide public goods in line with society’s wishes.

moving to jurisdictions offering services more in line with their preferences (Tiebout, 1956; Musgrave, 1959). These results are based on the assumption that local populations will be able to exercise choice and that they know their preferences and priorities better than regional or national authorities.

Bossert (1998) groups these arguments in two models or frameworks: the *principal-agent* approach and the local fiscal choice. In the former case, the main rationale for decentralisation is grounded on information asymmetry between the central and local levels of government; in the latter case, the main rationale for decentralisation is based on competition among locally elected governments. Examples of the former models can be found in Gilbert and Picard (1996)<sup>9</sup>, Robalino *et al.* (2001) and Faguet (2004); examples of the latter models are provided by e.g. Chubb (1985) and Seabright (1996). Importantly, in both approaches, diversity in provision across jurisdictions increases allocative efficiency because it is assumed that preferences also vary among jurisdictions.

In addition to the standard theory of fiscal decentralisation, another argument proposed by economists to justify a positive link between decentralisation and (production) efficiency is related to the issue of innovation, according to which autonomous local governments may be more willing to and able to experiment with new models of delivery (for a summary of economic arguments for and against decentralisation see Levaggi and Smith, 2005; for a summary of arguments generated by different disciplines see Peckham *et al.*, 2005).

So far, we have identified arguments pro decentralisation but economists have also formulated a set of arguments that might lead to the conclusion that decentralisation decreases production efficiency. One of these arguments is that economies of scale are lost with decentralisation (e.g. larger entities may be able to secure more favourable contracts with service providers - Levaggi and Smith, 2005). Additionally, inefficiencies are said to arise from the multiplication of small administrative units, that is, decentralisation increases transaction costs (Saltman and Bankauskaite, 2004; Levaggi and Smith, 2005). Also, the results

---

<sup>9</sup> Gilbert and Picard (1996) identify two effects potentially determining the precision of information on costs: on the one hand, the geographical proximity effect implies that the smaller the size of jurisdictions the more precise cost estimates will be; on the other, there is a learning effect that makes estimates on costs more precise in larger jurisdictions. Thus, the former effect pushes towards more decentralisation, while the latter pushes towards more centralisation.

suggested by the standard theory of fiscal decentralisation are valid assuming that well-qualified personnel are available at lower levels of government (Saltman and Bankauskaite, 2004). Thus, the local administrative capacity (or the lack of it) is identified as a strong determinant of the positive (or negative) effects of decentralisation (Lindaman and Thurmaier, 2002; Akin *et al.*, 2005). This argument applies to production efficiency as well as to allocative efficiency. That is, even if information about costs and preferences is greater at the local level, this might not lead to increased efficiency under decentralisation if local authorities do not possess the necessary skills and expertise to use that information in decision making (Khaleghian, 2004). Within the *principal-agent* framework this can be thought in terms of a set of capabilities of the agent, which will influence the achievement or otherwise of expected outcomes (Bossert, 1998). Competition between local governments may also be harmful rather than beneficial if, for example, jurisdictions compete on tax rates (which might lead to widespread under-provision of public services - Levaggi and Smith, 2005).

A final important economic argument against decentralisation concerns the existence of externalities or spillovers (when services provided in one jurisdiction affect citizens from another jurisdiction - Saltman and Bankauskaite, 2004; Levaggi and Smith, 2005). In this case, if these externalities are not properly accounted for by local authorities, there might be under-provision (or over-provision) of services compared to what is socially optimal, meaning that allocative efficiency is negatively affected by decentralisation.

#### *Review of theoretical arguments applied to the context of health care*

Theoretical analyses specifically developed for the case of decentralisation in health care resource allocation have mainly relied on informational issues, being therefore closer to the *principal-agent* approach than to the local public choice approach. This might be partly explained by the limitations of the application of the latter model to the context of health care. Saltman and Bankauskaite (2004), for instance, say that across most tax-funded health care systems in Western Europe not only there is a generalisation of centralised taxation but also central governments nearly always retain the authority to set

broad health strategy and often a variety of lesser health policy issues as well. Moreover, national governments nearly always maintain significant regulatory and monitoring roles. For the case of less developed countries, Bossert (1998) says that local resources are a small portion of local expenditures and intergovernmental transfers come with many administrative restrictions.

Under these circumstances there is only a limited applicability of the local public choice framework. As Lindaman and Thurmaier (2002, p.919) note, “without economic independence, either through an independent taxing authority capable of generating significant resources or no-strings block grants from the central government’s tax revenues, local autonomy will be undermined because local authorities will be accountable for budgetary decisions to the central authorities but not to their constituents”. The local public choice approach fits better scenarios of devolution and the *principal-agent* approach fits better scenarios of delegation and the latter seems to better reflect decentralisation in health care decision making.

Hurley *et al.* (1995) address precisely the issue of information, discussing how it affects the potential for decentralised structures in the health care sector to improve production and allocative efficiency compared to centralised structures. They start by distinguishing between information and knowledge. They consider that information includes e.g. facts and data, “all of which can exist, in a sense, outside human persons and which do not in and of themselves constitute usable knowledge”. On the contrary, knowledge “involves human understanding; human beings convert information into usable knowledge” (p.4). This distinction is important because information asymmetry does not necessarily mean that the access to data is asymmetric; what it might mean is that the capacity to incorporate this data into decisions is not the same across different organisations or across different levels of decision making.

The authors further identify three relevant aspects of information: its distribution -are the sources of information widely distributed among individuals or are they concentrated among a small number of individuals or locations?; its communicability - can the information be easily summarised and inexpensively transmitted within the organisation without distortion or loss of meaning?; and its degree of technicality - are specialised skills required to interpret the information? Moreover, three types of information are presented as requisites for an efficient

allocation of health care resources: (i) expert, technical information regarding the effectiveness of alternative interventions; (ii) information about ‘needs, values and preferences’; and (iii) information regarding circumstances affecting the delivery of care in each region (e.g. the geographic distribution of the population).

Hurley *et al.* (1995) also distinguish between primarily normative decisions, which often require broad consultation, and primarily technical decisions, which require that the decision maker have the requisite of knowledge and skills to evaluate alternatives under consideration. According to the authors, in health care, the latter decisions favour centralised structures given the nature of information required (technical information). Although information regarding the effectiveness and technical efficiency could be accessible to decision makers in either decentralised or centralised systems, the concentration of scarce expertise might constitute, in the short run at least, a constraint for decentralisation. Concerning normative decisions, involving the assessment of health care needs, values and preferences among the population, they tend to favour decentralisation. Partly, because decentralised settings may make certain institutional mechanisms for eliciting values and preferences of the community more feasible (including regular public meetings), and partly because this type of knowledge is likely to grow out of experience and is elusive, making it hard both to collect and to communicate without loss of meaning. Finally, they argue that the cost-effectiveness of alternative interventions is heavily dependent on site-specific factors such as input prices. Consequently, information about local circumstances tends to favour decentralisation.

Hurley *et al.* (1995) conclude that the gain in allocative efficiency associated with decentralisation is directly proportional with the degree of heterogeneity of preferences and values in the population, while the gain in production efficiency is directly proportional with the degree of variation in information about production-relevant local conditions. Hence, the impact of decentralisation on efficiency is ultimately an empirical matter.

Bossert (1998, p.1518), based on his ‘decision space’ approach (see section 2.2 above), says that decisions in each of the functional areas (finance, service organisation, human resources, access rules and governance rules) are likely to affect the system’s performance in achieving the objectives of equity, efficiency, quality and financial soundness. He acknowledges that most areas are



likely to have an impact on all the stated objectives of health care but his statements are vague, both in terms of what these objectives mean and the expected direction of the impact in each area. The author also notes that with multiple objectives, some activities may support some objectives at the expense of others. This reflects the acknowledgment of the possibility of trade-offs but it is difficult to see how this problem would be tackled within his framework, involving sixteen functions and three degrees of range of choice and given the absence of clear links between functions and objectives.

The strength of the ‘decision space’ approach is focusing on the issue of how much discretion, over what kinds of functions, is likely to lead local authorities to make choices that will achieve central government objectives but Bossert himself recognises the limitation of his approach regarding the impact of decentralisation namely on efficiency: “What decision space and what incentives work better in terms of efficiency, financial soundness, equity and quality and impact on health outcomes? Ideally the answers to these questions would form the largest section of this article. (...) Unfortunately, little comparative information is available concerning the relationship between decentralisation policies and performance” (Bossert and Beauvais, 2002, p.26).

Robalino *et al.* (2001) developed an empirical analysis that evaluates the impact of decentralisation on health outcomes. They say that decentralised decision making is expected to be more successful in allocating scarce resources to alternative interventions in order to maximise health and they suggest a theoretical framework within which this result might be achieved.

These authors start with the problem of a benevolent decision-maker aiming to maximise the population’s health (as measured by a given indicator like the infant mortality rate). They consider that within each region  $g$  in a country the outcome indicator,  $M$ , is a function of structural characteristics of the economy, represented by  $\theta$ , and the allocation of public expenditures,  $x_{gi}$ , among a set of interventions,  $I$ :  $M_g = f_g(\theta, x_{g1}, \dots, x_{gI})$ . The authors assume that health is increasing in expenditures but the marginal effect decreases with the level of expenditure. The problem to be solved by decision makers is:

$$\text{Max}_{x_{gi}} : M = n_g \sum_g f_g(\theta, x_{g1}, \dots, x_{gI}) - \lambda (\sum_g \sum_i x_{gi} - Y),$$

where  $n_g$  is the contribution of region  $g$  to the national average and  $Y$  is the total budget to be allocated.

Optimality implies that expenditures are allocated such that the marginal impact of an additional monetary unit to an intervention  $i$  in region  $g$  (adjusted by its weight  $n_g$ ) is the same across all interventions and regions. The optimal spending in intervention  $i$  in region  $g$  is given by  $x_{gi}^*$  (similarly, the optimal share of expenditure in intervention  $i$  in region  $g$  is represented by  $s_{gi}^*$ ).

Robalino *et al.* (2001) note that given local budgets,  $x_{gi}^*$  is also solution to the health maximising problem defined for region  $g$ . Nonetheless, they also stress that the allocation at the local level will generate a national optimum only if the allocation of  $Y$  among regions is optimal in the first place. Still, the authors argues that even if observed budgets are sub-optimal, their management by local governments can improve outcomes as long as local authorities have a better knowledge of  $f_g(\cdot)$  than the central authority.

A general indicator of inefficiency,  $U$ , is defined as the deviation between actual amounts allocated to each intervention/region and the respective optimal amount.  $U$  is then presented as a function of the share,  $S$ , of the total budget that is managed (i.e. allocated) by local authorities. The partial derivative of  $U$  with respect to  $S$  will depend on the relative levels of efficiency of the local and central governments in the management of public expenditures. Hence the authors postulate:  $\frac{\partial U}{\partial S} = C(c - l)$ , where  $c$  and  $l$  are indicators of the level of efficiency in

managing public resources of the central and local government, respectively.  $(c - l) < 0$  might be justified, as above noted, with a better knowledge of  $f_g(\cdot)$  by local authorities. The case where  $(c - l) > 0$  might be justified, according to the authors, with an institutional capacity at the local level low relative to the centre.

In summary, according to the previous framework there is not an unequivocal impact of decentralisation on efficiency, here interpreted as health maximisation, which in turn requires production efficiency. The argument advanced by the authors to justify a positive impact of decentralisation on efficiency is information asymmetry (better knowledge at the local level about health production functions); the argument advanced to explain an opposite outcome is basically the absence of a usual assumption of the standard theory of fiscal decentralisation: the availability of well-qualified personnel at lower levels of government.

Petretto (2000) analyses the ‘desirability of the regionalisation of a

National Health Service’; by ‘regionalisation’ the author means increasing local health services provision financed by local taxes. In his analysis, the utility of individual  $i$  is given by  $U_i = U(x_i, l_i, h(s_i, q_i))$ , where  $x_i$  is private consumption,  $l_i$  is leisure,  $h(.)$  is health,  $s_i$  is health status before care and  $q_i$  is a composite health service. The latter is a publicly-supplied (free of charge) private good; it includes a part that is guaranteed uniformly by the central government and a residual part that is determined locally. There are also some spillover effects in the sense that inhabitants of one region may obtain the service elsewhere (though the author considers that services provided by different regions are not perfect substitutes).

In this framework, an increase in regional autonomy with regard to health care expenditure and taxation is desirable for the community of region  $k$  if the marginal benefit of the public funds used for financing a supply increase in the composite health service is greater than the marginal cost of public funds. Marginal benefits and costs thus depend on consumers’ preferences (an increase in local health care has a positive impact on individual utility but it also has a negative impact since local taxes increase too and private consumption decreases). This is basically a cost-benefit analysis and is no different from the scenario of centralised provision and taxation, with a utilitarian welfare function and assuming no externalities.

Petretto (2000) further discusses another perspective to evaluate the effects of regionalisation: the perspective of overall social welfare point of view. This perspective now takes into account the spillover effects (actually, we previously saw that the existence of externalities is an argument that favours centralisation) but it also considers that spillover effects are corrected by considerations of redistributive aims (the author says that the social value of these effects is higher in regions ‘more deserving from a social point of view’). The differences between the regional and social perspective is therefore not confined to the internalisation of spillovers; the latter perspective includes non-welfarist considerations. It is not however very clear the origin of the latter considerations; the author talks about ‘federal mandates’ and ‘constitutional objectives’. In fact, in Petretto’s framework, only these restrictive mandates seem to offer a rationale for decentralisation (in the sense that a central authority is not allowed to provide varying supplementary health services across regions). Additionally, how this

framework might be used to evaluate the ‘desirability of regionalisation’ is somehow puzzling given that there are two non-comparable alternatives to measure such ‘desirability’.

Theoretical analyses in the context of decentralisation have considered that allocative efficiency is a matter of matching preferences without explicitly discussing what these preferences are; they seem to be consumers’ preferences (below in section 2.5 we discuss this issue in detail), for different health care services, for example. However, in section 2.3 we said that allocative efficiency is about choosing the preferred social state and that a variety of considerations, namely distributional and procedural concerns, could determine the preferred social state.

Therefore, we have identified two other outcomes attributed to decentralisation that might in fact impact upon allocative efficiency in health care resource allocation. One of these outcomes is the very diversity in the amount and type of services provided under decentralisation (Saltman and Bankauskaite, 2004; Levaggi and Smith, 2005; Magnussen *et al.*, 2007). If inequalities in health care provision, for instance, are considered a relevant factor to evaluate social alternative allocations, then, decentralisation, by generating variations in this regard, might negatively affect allocative efficiency.

Another outcome of decentralisation that has been identified in the literature is an increase in participation in decision making (Inman and Rubinfeld, 1997a, 1997b; Prud’Homme, 2003; Hutchinson and LaFond, 2004; Peckham *et al.*, 2005)<sup>10</sup> and we saw above in section 2.3 that procedural considerations like participation or voice might affect the ranking of social states. Consequently, decentralisation can also affect allocative efficiency via its impact on the procedural characteristic of voice. We said in the beginning of section 2.4 that our focus would be on outcomes derived within the discipline of economics but participation is an important outcome attributed to decentralisation within the democratic theory and participative democracy (Peckham *et al.*, 2005). However, we also noted in section 2.3 that economics (and health economics) has not a long tradition in dealing with procedural considerations; it is therefore almost

---

<sup>10</sup> Inman and Rubinfeld (1997b) say that participatory actions (such as ‘voting, debating, marching, picketing, contributing, and passive and armed resistance) are likely to be higher for each individual citizen when governments are smaller (because there are fewer participants).

inevitable to go beyond economics if we are to include this dimension of social justice in the evaluation of social welfare.

Actually, participation might be valued in its own right, directly contributing to people's utility (procedural approach discussed in section 2.3) but it can also be linked to the more consequentialist approach of economics. That is, it might be seen as a means of matching outcomes to the preferences of the population (consequentialist approach - indirect contribution to utility). Stutzer and Frey (2006) say that the latter perspective, focusing on the utility produced by more favourable outcomes, has been at the centre of attention in both political science and public choice analysis, while the procedural utility has been sustained for a long time now by political philosophers and the proponents of participatory democracy.

#### *Review of empirical literature applied to the context of health care*

The empirical evidence on the outcomes of decentralisation in health care is relatively small and much of it has been conducted using data from developing countries. In addition, probably due to the already mentioned difficulties in identifying and measuring independent and dependent variables, many studies are more qualitative rather than quantitative. Still, we give here a brief account of the main results that we have found in our literature search. We start with evidence regarding production-efficiency and the impact of decentralisation on health outcomes; then we focus on analyses that seem to be closer to allocative efficiency.

According to Hutchinson and LaFond (2004,p.14), “the *Gestão Plena do Sistema Municipal*, in Brazil, which entitled municipalities to receive block transfers of funds directly from the federal government for all levels of care, provided incentives for municipalities to over-invest in hospitals, laboratories, and high-tech equipment, which were then under-utilized given the municipality population sizes”. Collins *et al.* (2000, p.122) too argue that “many of the municipalities in Brazil are really too small to consider them as having the capacity to develop workable municipal health systems”. Among the factors contributing to this situation, are the lack of sufficient population size and

expertise.

Needham *et al.* (2004), argue that decentralisation has been associated with negative consequences on a specialised programme like Zambia's tuberculosis programme. They note that, in an urban setting, decentralisation may be more successful than in a rural setting, since when population density is very low (in Zambia,  $<30/\text{km}^2$ ), it is prohibitively costly to provide an acceptable level of services within reach of the sparse population.

Koivusalo (1999) says that, in Finland, local governance is well established yet there is a significant concern on the extent to which the relevant skills should be and can be maintained by municipal, provincial, and hospital administrations.

Bossert *et al.* (2003) (carefully) conclude, for the case of Zambia, that despite the decline in real per capita expenditure following decentralisation, there was little impact on the utilisation of outpatient services. They say that this may mean that the efficiency of the use of resources increased, though they stress their inability to evaluate the quality of services that may have been declining with the declining expenditures.

The recent recentralisation in Norwegian health care system (2002 reform) has been in part justified by a decrease in productivity and the loss of economies of scale (Mosca, 2006). In the same line, Magnussen *et al.* (2007, p.2135) conclude that efficiency 'certainly seems to have increased after the reform'. To reach this conclusion, the authors compare measures of technical and cost efficiency (calculated in Norway on an annual basis) for the period 1999-2000 (before) with the same measures for the period 2002-2004 (after).

Concerning the impact of decentralisation on health status, Robalino *et al.* (2001) conclude that higher fiscal decentralisation (measured by the 'ratio between expenditures managed by local governments and expenditures managed by central government') is consistently associated with lower infant mortality rates (the indicator of health outcomes used in the study) and that these benefits are particularly important for poor countries. They however note that they cannot distinguish between local governments with low and high institutional capacity and assess how decentralisation affects health outcomes in each case. Still, they argue that their results support the view that if local governments are strengthened then fiscal decentralisation is likely to improve health outcomes.

Smith, B. (1997, p.401) also points the case of Papua New Guinea, where devolution is believed to have led to “substantial improvements in the health of the population, notably a lowering of infant, childhood and maternal mortality, and increased life expectancy”.

Empirical evidence about the impact of decentralisation on allocative efficiency is almost inexistent. It is difficult to evaluate the match between delivery and preferences if the latter are not themselves assessed or even clearly defined to start with. Consequently, studies that claim to address allocative efficiency usually look at the allocation of resources between different social services and within health care the most common approach is to look at the allocation between primary and secondary care.

Homedes and Ugalde (2004) report several situations which are presented as failures in meeting allocative efficiency: in Bolivia, mayors spent health and education funds on road development; in Colombia, health funds were directed to building hospitals in municipalities, which, according to the infrastructure plans, were not needed; in Mexico, some states have diverted federal health funds to other activities; in Nicaragua, the Ministry of Health donated sacks of cement to rural households to build latrines; instead many families sold them.

Bossert and Beauvais (2002) mention the issue of distribution of resources between primary and secondary care saying that they found some evidence that local choices on expenditures in the Philippines and Uganda resulted in allocations to curative care rather than the national priority of primary care. In Zambia, the central government set percentage ceilings and floors on how much of district budgets could be spent on each level. It seems that in general the guidelines were respected (Bossert *et al.*, 2003).

Khaleghian (2004) examines the relationship between decentralisation and percentage of immunisation in 1-year old children for 140 countries and found that decentralisation is associated with higher immunisation coverage rates in low-income countries and lower coverage rates in middle-income countries. Indeed, immunisation is often used as an example of externalities in health, meaning that decentralisation could in theory decrease the level of provision.

Regarding the question of variations in service provision, there has been a tendency for regarding them not as the result of local governments' responding to variations in local preferences; instead, these variations have increasingly been

viewed as an undesirable feature of decentralised systems (Magnussen *et al.*, 2007). Variations in provision have also been associated with variations in resources in the first place. In Finland, however, municipalities have tax powers and there has been an increasing cost sharing by users of health services but despite the concerns about growing divergences between areas, it seems that it has not led to major differentials between areas in terms of service provision (Koivusalo, 1999).

Even in the absence of local taxes, if decentralisation creates the potential for local bids on global resources, then some inequalities might arise. Homedes and Ugalde (2004) argue, in respect to the Latin America in general, that: “decentralisation widened existing urban-rural and inter-regional inequities. Urban dwellers exert more political pressure than rural dwellers and, therefore, obtain a disproportional amount of health resources. (...) Geographical inequities are aggravated by the well-established fact that rural dwellers’ health needs are larger than those of urban populations”.

Varatharajan *et al.* (2004, p.49) mention, for the case of an Indian State - Kerala, the impact that active lobbying and technical expertise might have on “converting health care needs into fundable projects”. In this sort of situation, decentralisation might not only lead to a reduction in local funds but also to variations among localities depending on their ability to attract resources towards them.

Bossert *et al.* (2003) study the distribution of per capita health expenditures in Zambia, after decentralisation. Although there was a decline in expenditures, there was relatively little difference among districts when they were analysed according to degrees of urbanisation and income indicators.

In what concerns the impact of decentralisation on participation in general, Smith, B. (1997) notes that decentralisation might not always be accompanied by the necessary conditions to guarantee an effective and meaningful community participation in decisions, particularly in poor countries, where community participation is widely recognised as a problem due to factors such as political inequality and dependency, illiteracy, poverty, among others. For the case of health care resource allocation, we found no empirical evidence about this topic. There are several discussions about participation and involvement but only in one reference (Allen, 2006) there is an explicit attempt to link participation to



decentralisation. Allen (2006) looks at the link between membership of local organisations and participation, in the context of the English National Health Service. He says that historical literature concerning local and mutual organisations demonstrates that it is hard to achieve high levels of participation in such organisations. Focusing then on recent analyses of Foundation Trusts (FTs), this author reports that small numbers of people volunteered to become members of FTs and that when they were given the opportunity to vote for the board of governors, the voting turnout varied from relatively high to below of 20%; in three quarters of the elections, under 60% of the eligible patient and public members actually voted.

Peckham *et al.* (2005) carried out a review of empirical literature about the outcomes of decentralisation (including results from other disciplines besides economics). For all the outcomes included in their review, namely production and allocative efficiency, they found evidence of positive as well as negative relationships between decentralisation and these outcomes. In summing up their findings, the main message is that, apart from questions about the strength and quality of the evidence, the review demonstrates that the ‘balance of evidence is often equivocal at best or does not provide any real conclusion’ (p.90).

The impact of decentralisation on participation is not included in the ‘analysis of evidence’ in Peckham *et al.*’s (2005) report, which might be in itself a reflection of the gap in the empirical literature in this regard. In fact, participation appears in their topic related to future research, where it is said that given the ‘strong assumption made about participation and democracy being improved through decentralisation’ it is important that further research is undertaken to address this issue, namely in terms of determining what level of decentralisation is best for public involvement; how the public relates to local health agencies; and how does this relate to issues of responsiveness of local health services.

## **2.5. The concept of preferences in the context of health care**

In section 2.4, in which we explored the literature regarding the impact of decentralisation on efficiency, we saw that the relationship between

decentralisation and allocative efficiency derives from a better match between the provision of public services and preferences. And this is true regardless of which model, the public choice analysis or the *principal-agent* model, is used to evaluate that impact. Preferences are therefore at the centre of a major economic argument in favour of decentralisation<sup>11</sup>. The nature and content of these preferences have not however given rise to particular debate, within the literature about decentralisation. In fact, Frey *et al.* (2004) say that economics has been deliberately vague about what human preferences are defined over but in principle, ‘what individuals value could be anything’ (p.379).

In the literature about health care resource allocation however preferences, or values as they are often interchangeably used, have received an increasing attention. The concept of values is said to constitute now a fundamental element of policy analysis, even ‘comparable in importance to concepts such as ‘health’ or ‘evidence’ (Giacomini *et al.*, 2004, p.22). Despite this enhanced role, it is said that research into the definition (as well as the operationalisation and application) of the concept of values or preferences remains under-developed (Giacomini *et al.*, 2004; Bankauskaite and Saltman, 2006). In the current section, our objective is consequently to review the literature about health care resource allocation, summarising the main insights that have hitherto been proposed in order to clarify the meaning of preferences. We identified three main criteria that might be used to classify preferences (‘whom respondents are asked to think about’; ‘what respondents are asked to think about’; and ‘individual *versus* collective preferences’) but there are overlaps between them and there might also be some interdependence between them (for instance, the ‘whom’ might depend on the ‘what’).

#### *Classifying preferences according to whom respondents are asked to think about*

Dolan *et al.* (2003) identify three main perspectives that an individual could be asked to adopt in studies designed to elicit preferences for use in

---

<sup>11</sup> In the literature on decentralisation the term ‘local preferences’ is usually used instead of ‘public preferences’ but this stems from the above mentioned assumption of variation in public preferences across localities, thus, public preferences in a given locality are ‘local’ in the sense that they are different from other localities’ preferences.

informing resource allocation decisions in health care. These are: (i) ‘personal preferences’ – when the question is concerned with the respondent herself; (ii) ‘social preferences’ – when the question concerns people other than the respondent; and (iii) ‘socially inclusive personal preferences’ – when the question concerns both the respondent and other people.

The authors note that, under the first category, the individual is asked to consider only himself (self-regarding preferences). They further distinguish between *ex ante* and *ex post* preferences: in the former case, individuals should imagine that they face the possibility of being a patient; in the latter case, individuals are asked to imagine themselves as patients or preferences can be elicited from real patients. This latter dichotomy (hypothetical *versus* real patients) nonetheless usually arises in a specific context (health status measurement – see below).

In the second category, ‘social preferences’, respondents are detached from any self interest given that the probability of becoming a patient is null. According to Dolan *et al.* (2003), what is at stake here are individuals’ preferences about other people, which will reflect their concerns about the utility or welfare of other people as well as the distribution of utility or welfare among others. The distinction between *ex ante* and *ex post* preferences can be exemplified, the authors say, with the difference between statistical lives and identifiable lives (in the former case, there is some probability between 0 and 1 of others becoming ill; in the latter, illness is a certain outcome).

Finally, ‘socially inclusive personal preferences’ present a combination of the other two categories and respondents are asked to think about their own interests as well as the interests of others. In the *ex ante* context, individuals are asked behind a veil of ignorance (they do not know whether themselves will require health care but this possibility exists); in the *ex post* context the individual being asked might require the same treatment as others or not.

Dolan *et al.* (2003) note that under welfare economics, individuals are considered the best judges of their own welfare and this notion conforms to the perspective of ‘personal preferences’ above. Still, it might be that an individual’s utility is affected by the levels of utility enjoyed by others<sup>12</sup>. In this case, ‘socially

---

<sup>12</sup> In the presence of ‘caring externalities’, individuals may maximise their utility by giving. Culyer (1989a) terms this behaviour as ‘welfarist altruism’ (which, in some sense, can be thought as

inclusive personal preferences' might be more appropriate than 'personal preferences'. Tsuchiya and Miyamoto (2007) distinguish these two situations opposing the concept of 'personal utility' to the concept of 'individual utility': 'preferences over possible outcomes for other people or preferences over distributions where one's own absolute and relative wellbeing are not affected can be part of individual utility but not personal utility'. Dolan *et al.* (2003) say that although self-interest exists, it does not necessarily follow that it must be the basis for social welfare. They say that the perspective of 'social preferences' would seem entirely reasonable for example in the context of a tax-based health care system. The authors further note that when eliciting 'social preferences' the individual is typically placed directly in the shoes of a social decision maker. The perspective of 'social preferences' also seems in line with the shift that has occurred within the discipline of health economics from a positive study of how individuals make consumption decisions for themselves to a normative study of how society should make resource allocation decisions within the context of a publicly funded health care system (Tsuchiya and Miyamoto, 2007).

A different though related classification of preferences, depending on the different perspectives that individuals might be asked to adopt in preference elicitation, can be found in Tsuchiya and Miyamoto (2006, 2007). These authors give particular attention to the issue of health state valuation, more specifically, to health related quality of life. Recall from section 2.3 that measures of health like the QALY involve both dimensions of length and quality of life: each year of life is weighted by a quality-of-life score in which the latter can be represented by a utility function defined over health states (Tsuchiya and Miyamoto, 2007). A distinction must however be made between this utility function, used to measure health outcomes, and the utility that derives from the consumption of health care. Tsuchiya and Miyamoto (2006) define the 'informed consumer' precisely as the individual that is assumed to be rational, selfish and the best judge of how much utility health care will yield for him.

Focusing then on the utility incorporated in health status measurement, Tsuchiya and Miyamoto (2006, 2007) propose three ways of classifying individuals: the 'capable patient'; the 'informed non-patient'; and the 'informed

---

selfish: one gives away to the extent that this maximises one's own utility).

citizen'. The former two are assumed to be rational and selfish. The difference between them is that the 'capable patient' is an actual patient and therefore knows what it feels like to have a given health problem, while the 'informed non-patient' is assumed to know about ill-health states but he himself is not ill. It is in this context of health status measurement that the distinction made by Dolan *et al.* (2003) between *ex ante* and *ex post* personal preferences (or between hypothetical and actual patients) has become an issue of debate. In regards to the 'informed citizen', the individual is assumed to know what it feels like to have different problems, to be rational but, contrarily to the previous cases, is assumed to be self-less "in the sense that she will not make judgements in order to forward her own case, or to advance the case of one particular health problem over another" (Tsuchiya and Miyamoto, 2007).

Smith and Richardson (2005, p.80) in turn distinguish between 'social utility' and 'individual utility'. They say that while the former "is an 'intellectualised' preference, one thought through rationally, the latter is more directly related to personal feelings and hence much more subjective". In fact, it has been suggested that a qualification should be applied to social preferences in the sense of excluding unacceptable views. Tsuchiya and Miyamoto (2007), for example, talk about justifiable societal preferences as opposed to unacceptable views based on prejudices (like differential treatment by characteristics such as race, religion or sexual orientation); Dolan *et al.* (2003) talk about 'ethically defensible set of society-regarding preferences'; Olsen *et al.* (2003) argue that ethically unacceptable preferences, such as discrimination on the basis of race, can be distinguished from those that are defensible in terms of moral argument; Richardson and Mckie (2005) suggest that defensible principles for allocating health care should be derived in an iterative way, involving both an empirical study of population values and ethical analysis of the results.

#### *Classifying preferences according to what respondents are asked to think about*

The concept of preferences has also been used in different circumstances in terms of what respondents are asked to think about. An alternative that we have already identified concerns preferences over health states, which might be

combined with the perspectives previously reviewed. Tsuchiya and Miyamoto (2007) say that so far not many actual valuation studies (regarding health related quality of life) have used the citizen perspective. There is an ongoing debate about which of the three perspectives (the ‘capable patient’; the ‘informed non-patient’ or the ‘informed citizen’) should be used to derive quality-of-life scores (e.g. Dolan and Green, 1998; Ubel *et al.*, 2000; Smith and Richardson, 2005).

In Shiell *et al.* (1997), a distinction is made between preferences for market goods, which include health care, and ‘underlying preferences’, which are defined over fundamental aspects of life, such as health, prestige, benevolence or envy. This view draws on psychology and it is said that the basis for the distinction between preferences and ‘underlying preferences’ or ‘values’ is not the moral argument. The latter are stable and the distinction between the two types of preferences “relates to the ease with which individuals are able to formulate and express the values they hold for fundamental aspects of life” (Shiell *et al.*, 1997).

Even if one accepts that social preferences should be defensible in terms of moral argument, there are several competing and sometimes conflicting arguments. Williams (2001) says that ethical discourse is typically inconclusive because the premises on which it is based are usually contestable, in the first place. So, preferences can also be defined over a set of ethical principles, both in terms of which principles matter and given conflicts between them in terms of the trade-offs between principles. Procedural characteristics might equally be the target of preferences. Examples of these principles have been discussed in section 2.3 and several examples of empirical studies focusing on principles of justice are referred below.

Preferences might further be defined over specific health care interventions or services (e.g. Bowling, 1996). These preferences might be elicited from a personal perspective or, following the tendency in health economics to focus on the societal level, in the context of a publicly funded health care system, they might be elicited from a social perspective.

The concept of preferences can also be used to represent preferences for attributes of health care delivery, such as opening hours of facilities, doctor’s gender, time for appointment *versus* time of consultation, etc. Mooney (1998) for example argues that the processes of care themselves can be utility bearing. These are clearly personal preferences that are closer to a consumer perspective. There

are some examples of empirical analyses for the case of out-of-hours care (e.g. Gerard *et al.*, 2006). Preferences for attributes of health care delivery might be relevant even in a context of a tax-based health care system. In some cases, the success or otherwise of policies might depend on people voluntarily seeking health care services. Think for example about preventive care such as screening programmes, in which different attributes might constitute an incentive or, on the contrary, a barrier, to seek care.

### *Individual versus collective preferences*

When discussing the meaning of values, Giacomini *et al.* (2005) say that, among other things, values can stem from collective beliefs, as expressed for example through social institutions, or, they can stem from individual convictions, as expressed by opinion polls.

Mooney (2005) uses the concept of ‘communitarian’ preferences, arguing that this is different from Dolan *et al.*’s (2003) concept of ‘socially inclusive personal preferences’, given that in the latter case the individual is asked ‘qua individual’; the ‘community qua community’ is not asked. Mooney (2005) says that community would most often be a geographical entity such as a province or region or state or even country for which there is already a health service and a health service budget. The author says that to elicit the preferences of the community qua community, individuals should be asked to reflect and discuss together what values or principles they would want adopted by the health service decision makers when these decision makers determine resource allocation. The views of the NICE Citizens Council<sup>13</sup>, in England and Wales, as expressed for instance in the Institute’s report on social value judgements (NICE, 2005) can be thought as an example of collective preferences.

In the majority of situations, the concept of preferences has been used to represent ‘individual’ as opposed to ‘collective’ preferences. In empirical studies, even when discussion groups are used, after the exposition and discussion of the issues in question, respondents are asked to express their preferences individually.

---

<sup>13</sup> NICE stands for National Institute for Health and Clinical Excellence and the Citizens Council is a form of citizens’ jury (see chapter five).

## 2.6. What matters to people when allocating health care resources? A review of empirical evidence

So far we have seen that preferences might be defined over different things, including principles of distributive or procedural justice; we have reviewed the arguments used in the theoretical literature to explain why different principles or procedures might matter to people; and we have also seen that preferences might vary depending on whom individuals are asked to think about. As Frey *et al.* (2004) note individuals could value anything. This view may nonetheless pose some questions about its limits (is anything admissible?). Fehr and Fischbacher (2002) say that there is a strong convention in mainstream economics of not explaining puzzling observations by changing assumptions on preferences. “Changing preferences is said to open Pandora’s box because everything can be explained by assuming the ‘right’ preferences” (Fehr and Fischbacher, 2002, p.C30). Thus, it is important to show that the sorts of preferences used in this dissertation are not only based on the theoretical literature but they are also based on empirical evidence.

The objective of this section is therefore to review the empirical evidence regarding people’s views on different criteria and processes that might be adopted in health care resource allocation. The concept of preferences has been indistinctively used in this literature involving different contexts as identified in section 2.5. We review studies that elicit the views of the general public but, depending on the framing of questions, in some cases personal preferences might be at stake while in other cases social (citizens’) preferences might have been elicited. Studies that elicit views from the perspective of particular groups such as patients or health professionals are excluded. Below we present the main results<sup>14</sup>.

### *Preferences related to distributive justice*

Starting with health maximisation, there is evidence that the expected outcome of treatment matters to people (Mossialos and King, 1999; Roberts *et al.*,

---

<sup>14</sup> Some references generated by our literature search are, or include, themselves reviews of surveys of public preferences (Sassi *et al.*, 2001; Schwappach, 2002; Dolan and Shaw, 2003; and Dolan *et al.*, 2005).



1999; Dolan and Shaw, 2003), in particular in the context of the allocation of donor organs where health gain has been ranked first (Ratcliffe; 2000; Dolan and Shaw, 2004) or second (Neuberger *et al.*, 1998) among a set of factors used to prioritise potential recipients of organs. Roberts *et al.* (1999) concluded however that there is little support for health care programmes that provide a prognostic improvement but leave patients in relatively poor states of health and Dolan and Cookson (2000) found that although individuals are not totally unwilling to give priority to those who gain most from treatment, the end-point of health that the groups end up with after treatment seems to be relevant to people.

Concerning health inequality, there is evidence that the general public embodies some inequality aversion and is therefore willing to trade-off health gain against the reduction in health inequalities. In some cases, studies have focused on health inequality *per se* (Dolan and Robinson, 2001); in other cases, differences by social class (Lindholm and Rosen, 1998; Abasolo and Tsuchiya, 2004; Williams *et al.*, 2005, Tsuchiya and Dolan, 2007) or sex (Williams *et al.*, 2005) are considered. The latter authors, for example, report that the results suggest that there is a general willingness to sacrifice health benefits to target those with worst health prospects but there was considerable heterogeneity: in all questions there were responses implying no targeting at all as well as responses in which targeting would lead to lower health for all. Also, a result that emerged in this study is that preferences are sensitive to what inequalities exist and which groups are at stake. There were stronger preferences for reducing inequalities in life expectancy than in long-term illness and it also appears that inequalities by social class raise greater concerns than inequalities by sex.

Individuals' preferences regarding the role of age in health care resource allocation has been explored in the empirical literature as well. There is evidence to suggest that health gains to the young are weighted more highly than those to the old (Williams, 1988; Bowling, 1996; Neuberger *et al.*; 1998; Cookson and Dolan, 1999; Mossialos and King, 1999; Ratcliffe, 2000; Wilmot and Ratcliffe, 2002; Dolan and Shaw, 2004; Dolan and Tsuchiya, 2005). However, there are different forms of ageism (Tsuchiya *et al.*, 2003) and, as Dolan and Shaw (2003) note it is often difficult to tell how much of the preference for the young is due to the benefits to the young being greater (or being perceived to be greater) and how much is due to the young having lived less time. While the former argument

conforms to health maximisation, the latter is closer to the 'fair innings' argument. In one study, the authors nonetheless kept health gain constant across groups and still found that respondents were willing to give priority to the young: 40-year-olds always get a higher ranking than 60-year-olds (Dolan and Tsuchiya, 2005). Despite the support for the young over the old, Wilmot and Ratcliffe (2002) say that individuals were willing to discriminate between the ages of 5 and 80 but they did not discriminate between the ages of 40 and 60. Dolan and Shaw (2004) too report that respondents tended to focus on both ends of the age spectrum (20 and 70 year old). Anand and Wailoo (2000) say that by far the largest group of respondents in their study is the one that looks like what they termed as 'equal righters', corresponding to those individuals who preferred to allocate funds equally between diseases affecting younger and older groups.

In line with the distinction, introduced in section 2.3, between health achievement and opportunity to achieve good health, empirical studies have explored whether or not the public gives lower priority to supposedly self-inflicted diseases compared to 'natural' diseases. Empirical results show that the public is not totally neutral to the cause of illness, when it comes to allocate scarce resources (Bowie *et al.*, 1995; Bowling, 1996; Neuberger *et al.*, 1998; Mossialos and King, 1999; Anand and Wailoo, 2000; Ratcliffe, 2000; Williams *et al.*, 2005; Dolan and Shaw, 2003). Dolan *et al.* (1999) nonetheless report that the majority of respondents (above 60%, both before and after discussion/deliberation) expressed no priority regarding smokers, heavy drinkers, people with unhealthy diets, those who rarely exercise and illegal drug users. Dolan and Shaw (2004) in turn note that the issue of health related behaviour in general terms gave rise to discussion and dissent, with some respondents being very in favour of discriminating against those with self-inflicted illness, and others very against such discrimination. Nonetheless, once the discussion was put in the context of donor transplantation, when having to make a choice between two individuals, respondents often chose that person whose kidney failure is not self-inflicted.

Need principles, discussed in section 2.3, have also been addressed by empirical studies that have in this case explored the public supports for resource allocation based on ill health and on the rule of rescue. In this regard, there is evidence that people wish to give priority to the treatment of life threatening illnesses' (Bowling, 1996; Cookson and Dolan, 1999; Dolan and Shaw, 2003;

Dolan and Shaw, 2004). The severity of illness has also been studied in terms of trade-offs *vis-à-vis* other criteria. Available evidence indicates that people are willing to sacrifice health gain or quality of life in order to give priority to those with the worst no-treatment profile (Nord, 1993; Ubel *et al.*, 1998; Cookson and Dolan, 1999). Dolan and Tsuchiya (2005) oppose preferences for worse future health prospects to preferences for the young over the old, and found that younger groups are always chosen over older ones, being future health and years without treatment non-significant in terms of explaining peoples' preferences.

### *Preferences related to procedural justice*

There is less evidence about people's support for procedural characteristics compared to distributional concerns, in the context of health care resource allocation. Of the six procedural characteristics identified above (voice, neutrality, consistency, accuracy, reversibility, and transparency), voice is the dimension that has been studied more often.

Available evidence suggests that people believe that voice or involvement is important (Anand and Wailoo, 2000; Litva *et al.*, 2002; Wiseman *et al.*, 2003; Wailoo and Anand, 2005; Dolan *et al.*, 2007). However, there are some *nuances* depending on the level of involvement, on the type of decision at stake and on the level of government. For example, it seems that people wish to be consulted about their views but they are less willing to set priorities, which is seen more as a doctors' responsibility (Dolan *et al.*, 1999; Coast, 2001; Litva *et al.*, 2002). The evidence also indicates that the support for public involvement is greater at higher levels of decision making (Litva *et al.*, 2002; Wailoo and Anand, 2005). Moreover, Dolan *et al.* (2007) report that voice was seen both as a right (associated with a notion of citizenship or democracy) as well as an instrumental mechanism 'to identify needs and to reflect public preferences'.

Based on available evidence, all the other five characteristics of procedural justice are also important to people (Tsuchiya *et al.*, 2005; Wailoo and Anand, 2005; Dolan *et al.*, 2007). In Dolan *et al.* (2007) consistency does in fact rank higher than voice. The authors say that the survey questions were set at a national level in the context of an institution that has the elimination of regional variation

as one of its primary aims. The authors further say that discussions about consistency tended to focus on regional rather than temporal issues. Wailoo and Anand (2005) however report that the proportion of respondents that felt that consistency was not important was in excess of 20% at the government level compared to the other two levels despite that the problematic of variations in health care provision among regions (known in England as ‘postcode lottery’) is most relevant to the governmental decision making level. Other results are also at odds with aversion to geographic variation in resource allocation decisions. For instance, Bowling (1996) says that when asked whether ‘the government should issue guidelines to doctors about when not to use life-saving medical treatment/technology’, 77% of respondents disagree or strongly disagree, despite that this sort of guidelines could ensure equal provision across the territory. Moreover, regarding who should set priorities, only 3% (61 out of 1997) of respondents said “politicians and the government at national level”.

#### *Preferences over attributes of health care delivery*

The type of preferences addressed here differs from the previous two groups in the sense that it explicitly looks simply at what individuals prefer for themselves, on whatever basis. In our search, we found studies only for the case of out-of-hours care. The available evidence suggests that people are sensitive to aspects of delivery such as the ‘doctor’s manner’ (whether the doctor takes time to listen), time to making initial contact, time waiting for advice or treatment, type of contact (by telephone or in person; home visit or seeing an accident and emergency doctor, etc.), profession of person providing initial advice (trained nurse or doctor), and information about expected waiting time (Morgan *et al.*, 2003; Scott *et al.*, 2003; Gerard *et al.*, 2006). The latter authors also estimated some trade-offs (for example, respondents are willing to wait 58 minutes longer for advice or treatment to be seen by a doctor than a nurse and 30 minutes longer to be seen in person rather than spoken to by telephone).

## 2.7. Main findings and gaps in the literature

In this chapter, we have primarily focused on the impact of decentralisation on efficiency, from the perspective of economics and health economics. Regarding production efficiency, the theoretical literature has provided arguments both in favour and against decentralisation. Examples of the former are greater knowledge at the local level regarding local circumstances affecting production costs and greater innovation at the local level; examples of the latter are the loss of economies of scale and lower technical knowledge at the local level. Concerning the impact of decentralisation on allocative efficiency, an argument pro decentralisation is related to a better match between service provision and preferences; an argument against decentralisation is the existence of externalities. In conclusion, the theoretical literature is ambiguous regarding the net effect of decentralisation on efficiency. Similarly, the empirical evidence does not offer us conclusive results.

In the theoretical literature, a better matching between service provision and (local) preferences is achieved mainly via two models: the *principal-agent* approach and the public choice model. In the former model, the critical issue is the existence of information asymmetry between the central and local levels of government, in which it is assumed that local decision makers are better informed about local preferences than central decision makers; in the latter model, the critical point is the existence of competition among local decision makers for mobile taxpayers, in which a better match between service provision and preferences is seen as a way of attracting residents. The theoretical literature developed for the context of health care has relied not only but mainly on the issue of information asymmetry.

In both the *principal-agent* approach and the public choice model, preferences play an important role in the positive link established between decentralisation and allocative efficiency but what leads to this positive result is information asymmetry or competition (among local decision makers) and not preferences themselves. Thus, a question that might be raised here is: are the nature and content of preferences relevant in this type of analysis? In other words, might preferences *in themselves* influence the impact of decentralisation on allocative efficiency? This is the central question addressed in the theoretical

discussion presented in the next chapter.

In fact, the nature and content of preferences have not been explicitly discussed in the literature about decentralisation; however, if we cross this literature with the literature about health care resource allocation, new insights may emerge, helping to answer the previous question. Basically, the literature about decentralisation has viewed preferences under a traditional welfare economics perspective, according to which individuals are rational, selfish and are the best judges of their own welfare. However, once we look at the (theoretical and empirical) literature dealing with health care priority setting, we conclude that preferences might be many different things, depending on whom or what individuals are asked to think about, and that many other arguments besides self-regarding utility might affect the preferred social state and hence the optimal allocation of resources.

In this latter literature, it has also been acknowledged that there might be trade-offs between different arguments affecting social welfare. Contrarily, in the literature on decentralisation, trade-offs between different objectives of health care resource allocation remain a topic under-explored. This seems to be an area in which one stream of literature has something to offer to the other stream of literature. Within health economics there are also some trade-offs that have so far received little attention, namely, trade-offs involving inequality (in health or in health care provision) across geographic areas. These trade-offs might acquire particular relevance in contexts of decentralisation given that the latter have been associated with geographic variations in provision. In chapter three we consider the possibility of decentralisation creating trade-offs involving inequality across regions both in terms of health and health care provision. In chapter four, we also explore at an empirical level the trade-off between health maximisation and equality of treatment across geographic areas. In chapter five, we further argue that this is a particularly pertinent issue (and lack in empirical research) in the case of the English NHS given the lively debate about variations in provision across areas that has taken place in England.

Another sort of trade-offs less explored is that involving outcomes, on the one side, and procedures, on the other. Again this might be relevant in contexts of decentralisation given the link established in the theoretical literature between decentralisation and the procedural characteristic of voice or participation. In fact,

in the literature about decentralisation, participation has been presented as a positive outcome and it has been seen as a *good* in itself but the issue of people's preferences for participation has not been discussed; on the other hand, the literature about health care priority setting has to some extent addressed the theme of public participation or involvement and it has been acknowledged that preferences for involvement might be greater or lower and might depend on several factors but no association between participation and decentralisation has been made. Thus, there seems to be here an unexplored path between decentralisation and allocative efficiency in health care, via preferences. In chapter three (section 3.5) we also use this perspective to address the question posed above.

Finally, another important finding/gap is that, regardless of which model is used, variation in public service provision (resulting from decentralisation) increases allocative efficiency because it is assumed that preferences vary across jurisdictions, which is ultimately an empirical matter. Despite the pertinence of this question, empirical evidence about variation in preferences across regions, in the context of health care resource allocation, is almost non-existent: in one case, Neuberger *et al.* (1998), the issue of geographical heterogeneity is marginally mentioned (in this study, there were no significant differences among the responses of the general public according to location in the United Kingdom); and, for the case of preferences over health states, while Sculpher and Gafni (2001) and Birch and Gafni (2002) argue that there is variation in preferences and even question the valuation of health gains, in England and Wales, based on the values of the general population, Robinson and Parkin (2002) argue that evidence indicates the existence of variability *per se* but not the existence of systematic variation. The main objective of the empirical work reported in chapter four is precisely to test the assumption of variation in preferences across (two) geographic areas.

## **CHAPTER 3. Might preferences in themselves influence the impact of decentralisation on allocative efficiency? A theoretical discussion**

### **3.1. Introduction**

In the final section of chapter two, we noted that the concept of preferences has played a relevant role in economic analyses of decentralisation but only indirectly. We further noted that this concept has been interpreted in these analyses as consumers' preferences. However, a much broader interpretation of the concept has been proposed in the health economics literature. Thus, the main objective of the current chapter is to discuss whether or not preferences, in this broader sense, might in themselves influence the impact of decentralisation on allocative efficiency, in terms of health care resource allocation.

At the onset we define a framework, grounded on the previous review of literature. Within this framework, we identify the conditions under which decentralised resource allocation decisions improve local populations' welfare, or not, compared to centralised decisions (the *status quo*). The analyses in the various sections make use of the diagrammatic language to better illustrate our points. More formal approaches are not adopted since no particular functional forms are assumed for social welfare functions as well as for health production functions.

In section 3.2, we thus present the framework within which the subsequent discussions are developed. The main assumptions are identified and supported with the findings of the previous chapter. In this section, we also suggest that our framework can be seen as an adaptation of the *principal-agent* approach. This adapted *principal-agent* relationship is later used in chapter five to distinguish between variation in provision derived from variation in preferences and variation in provision derived from variation in technical information.

The discussion starts in section 3.3 with the impact of information asymmetry on the outcomes produced by decentralisation in terms of efficiency, stressing the trade-off between production efficiency and allocative efficiency (although the former type of efficiency is embodied in the latter, the term



allocative efficiency is used to account for the popular outcome of decentralisation concerning a better match between the provision of goods and services and preferences).

In section 3.4, preferences for equal provision of health care across jurisdictions are considered, in which decentralisation might create a trade-off between this objective of resource allocation, on one hand, and production efficiency and other concerns embodied in the preferred health distribution, on the other hand.

Section 3.5 focuses on preferences for participation (or preferences for having voice) in decision making. In this case, the trade-off potentially generated by decentralisation involves outcomes, on the one hand, and procedures, on the other. This happens because decentralisation increases participation or voice, which is a procedural characteristic that has inherent value to people regardless of outcomes but it comes at a cost in terms of health gain foregone, that is, given the assumptions identified below in section 3.2, decentralisation leads to production inefficiency.

In section 3.6, the impact of decentralisation on local populations' welfare is discussed making use of the concept of externality effects, being the latter associated with distributional concerns at the interjurisdictional level. The trade-off that emerges with decentralisation is in this case between amounts of resources allocated to health care that better reflect local preferences, on the one hand, and a health distribution between jurisdictions that is farther from the preferred distribution, on the other.

Preferences for health states are the topic of discussion in section 3.7. The main issue under consideration is in this case the use of local, rather than national, preferences to set the distribution of a common pool of resources among jurisdictions.

Based on the discussion developed throughout the chapter, the main conclusion is that preferences might be seen as a basis to determine the optimality or otherwise of decentralised resource allocation decisions, compared to centralisation. However, this result is based on some assumptions; thus, the final section of the chapter (section 3.8) presents some limitations of the current analysis. Examples of these limitations are those stemming from the assumptions of: benevolent decision-makers, information asymmetry, variation in preferences,

absence of movements of people between areas, positive link between decentralisation and participation and exogenous preferences. The implications of relaxing these assumptions are discussed and several other issues that remain to future research are also identified. The assumption of variation in preferences is actually tested empirically in chapter four.

### **3.2. The basic framework and assumptions**

In this section, we describe the overall framework that supports our discussion in the current chapter. The main assumptions adopted here are based on the review of literature presented in chapter two.

In our framework, there is a central decision-maker (that might be viewed as the central government or the department of health) and there are two jurisdictions,  $A$  and  $B$ . In some cases it is sufficient to analyse the case of a single jurisdiction, which we represent by ‘jurisdiction  $j$ ’. By jurisdiction we mean a sub-national group of people governed by a decision-making body that is a budget holder and that enjoys some autonomy in terms of health care resource allocation. In some of the sections below, a representative resident of each jurisdiction is considered; in others, we consider that there are two residents,  $X$  and  $Y$ , in each jurisdiction.

The total amount of resources available to health care is fixed. Local budgets are wholly funded from central transfers (no funds are raised by local budget holders, whether through taxes or user fees). Thus, we discuss decentralisation applied to resource allocation decisions but not to revenue raising powers (this perspective fits the scenarios in most Western European countries, for instance, where there is a generalisation of centralised taxation). In two sections, we actually allow local budgets to vary according to preferences but the assumption that local budgets are formed by central transfers still holds.

Resource allocation decisions might concern the distribution of resources among alternative health care interventions. Health care interventions may represent high technology curative interventions like heart and lung transplantation; or medicines (both for preventive or curative purposes); or

screening programmes (e.g. breast cancer screening). When we admit the existence of two residents in each jurisdiction, another decision comes into play, which is the allocation of resources between individuals  $X$  and  $Y$ , within each jurisdiction.

We assume that: (i) health gain is null in the absence of health care; (ii) health care causes no harm to individuals; (iii) health gain increases with health care; and, (iv) there is diminishing marginal health productivity (as more health care resources are allocated to  $j$ 's resident, the additional health gain obtained by  $j$ 's resident becomes smaller)<sup>15</sup>. For expository purposes, we assume that the maximum health gain, for a given amount of resources, is the same across jurisdictions. We also assume that there is no health externalities, meaning that the amount of health gain obtained in one jurisdiction is independent from the health gain obtained in the other jurisdiction.

The constraints faced by decision-makers concern the limited amount of available resources and the available technology (that defines the opportunity set of alternative uses of resources). We further assume that central transfers received by each jurisdiction under decentralisation correspond to the amount of resources that the central decision-maker would devote to the respective local populations, under centralised decision-making (that is, the amount of funds does not change as a result of decentralisation).

The concept of preferences should be interpreted in this chapter in a general sense, in line with the suggestions found in the health economics literature and reported in the previous chapter. That is, they might be personal preferences – when individuals are thinking about themselves, including (welfarist) altruistic sentiments; they might be social preferences – when individuals are thinking about the whole society (each individual might think only about others or he might think about the whole population including himself); preferences might focus on general principles of (distributive or procedural) justice, on groups of people, on specific health care interventions, or on attributes of health care delivery; they might even be the aggregate of individual views or some sort of collective preferences.

---

<sup>15</sup> These assumptions, leading to upward sloping and concave health production functions, are very common in the health economics literature (e.g., Wagstaff, 1991; Wagstaff and Culyer, 1993; Dolan and Olsen, 2001; Hauck *et al.*, 2002).

We assume that preferences, relevant for resource allocation decisions, are heterogeneous across jurisdictions. In chapter two, we saw that diversity in preferences is a central assumption in classical theory of fiscal decentralisation (e.g. Oates, 1972, 1999; Tiebout, 1956; Musgrave 1959). Other analyses that deal with preferences have relied on, or at least acknowledged the possibility of, heterogeneity in preferences (e.g. Hurley *et al.*, 1995; Gilbert and Picard, 1996). Although recognising that the concept of preferences adopted in our discussion is potentially broader than the concept that has been adopted in the literature about decentralisation, at this stage we follow without questioning the traditional assumption that diversity exists.

In our discussion, unless otherwise stated, we assume that there is imperfect information regarding preferences and that local authorities are better informed about these preferences (in their respective jurisdictions) compared to the central authority. This assumption goes along with other authors' positions (for instance, Oates, 1972, 1999; Hurley *et al.*, 1995; Gilbert and Picard, 1996; Faguet, 2004)<sup>16</sup>. In chapter two, we saw that this argument has been used to justify a positive impact of decentralisation on allocative efficiency.

On the other hand, we assume that the centre has superior technical information compared to local authorities. This means that X-efficiency or production efficiency (as defined in chapter two – 'being on the health possibility frontier') is greater with centralised than with decentralised allocation of resources. Hurley *et al.*'s (1995) discussion supports this assumption and other authors have assumed (Faguet, 2004) or considered the possibility of (Robalino *et al.*, 2001) a similar cost advantage of the central government compared to local authorities. In Norway, the lack of capacity of hospitals to manage and organise services efficiently has been put forward to partly justify the 2002 reforms of recentralisation (Mosca, 2006; Magnussen *et al.*, 2007). In England, the lack of personnel with information handling and analytical skills to transform 'loads of data' into 'usable information', within Primary Care Trusts, has also been recognised (Bate *et al.*, 2007).

Another argument, discussed in the theoretical literature, that could be used to explain this negative association between decentralisation and production

---

<sup>16</sup> See chapter two.

efficiency is related to economies of scale<sup>17</sup>. Nonetheless, there are also arguments that point to the opposite result. For example, local authorities are said to know better than the central government about local circumstances affecting production costs (Oates, 1972, 1999; Hurley *et al.*, 1995; Gilbert and Picard, 1996). The latter authors use the term of ‘geographical proximity effect’ to define the cost advantage of local authorities. But these authors also say that there might be a ‘learning effect’ about costs, which favours centralised allocations.

Additionally, it could be argued that the centre is better ‘equipped’ to elicit and meet preferences. It is therefore admissible that information asymmetry can go either way. On the whole, taking into account both the theoretical literature and empirical evidence, we concluded in the previous chapter that the impact of decentralisation on allocative efficiency is ambiguous and that the same occurs in regards to the impact of decentralisation on production efficiency. Our interest lies on the hypothesis of both impacts diverging. Here, we assume that this divergence derives from information asymmetry as previously explained but we must acknowledge that, given the current state of knowledge, different assumptions would be acceptable as well.

We make use of the theoretical assumption, mentioned in the previous chapter, that there is a positive relationship between decentralisation and participation.

Centralisation is taken as the *status quo* against which decentralisation is evaluated, in terms of its impact on social welfare. This has been the preponderant perspective adopted in the literature about decentralisation and it seems to suit the contexts of most Western European countries (and less developed countries as well), in which decisions about health care resource allocation have remained a responsibility of central governments; therefore, in general, the debates have been mostly about the pros and cons of the changes brought by decentralisation always compared to the *status quo* of centralisation.

Evaluating the impact of decentralisation on social welfare requires the definition of the desideratum and a rule of aggregation of the individuals’ levels of the desideratum. In chapter two, we said that in terms of the former there is a main distinction between a welfarist perspective (that takes the individuals’ utility

---

<sup>17</sup> In our discussion we focus only on the informational argument though.

as the desideratum) and a non-welfarist perspective according to which welfare might depend on other attributes beyond utility, like health.

Regarding the aggregation rule, Tsuchiya and Miyamoto (2007) note that the simplest aggregation rule is to add up the changes in the desideratum across individuals without any weights (or equivalently, with uniform weights) so that the outcome with the largest total is recognised as the best outcome. This rule follows the Benthamite tradition that everybody counts for one and nobody counts for more than one.

But the aggregation rule might also incorporate concerns for the distribution of the desideratum. Tsuchiya and Miyamoto (2007) say that under non-welfarism, these concerns can be derived from the informed citizen<sup>18</sup> or policy makers. Other forms of aggregation might involve some sort of bargaining or contract among individuals (Dolan *et al.*, 2003).

In health economics, the desideratum has been perceived to be health *per se* precisely because it is valued by the public at large or the relevant decision makers (Tsuchiya and Miyamoto, 2006). At the same time, the role of the public concerning the relevant aggregation rule has also been increasingly recognised within the non-welfarist approach, adopted in health economics. It is now accepted that the preferences of the general public have a role to play in terms of decisions about how to allocate resources in health care (Dolan and Tsuchiya, 2007). It is in this context that we discuss the impact of decentralisation on social welfare by reference to preferences. That is, in our discussion, decentralisation is said to improve allocative efficiency if it leads to a better match, compared to centralisation, between what is achieved with resource allocation and what are the preferences (as defined above) regarding this same allocation.

Importantly, we assume that decision-makers, both at the central and local levels, are benevolent planners. Thus, the objective of decision-makers is to maximise the social welfare of local populations.

Figure 2 diagrammatically illustrates our approach, which might be interpreted as an adaptation of the *principal-agent* framework. In the upper half of diagram, we represent the traditional focus on the vertical relationship between the central and local authorities, where the latter are explicitly seen as *agents* of

---

<sup>18</sup> See chapter two for the definition of this concept.

the former. In this case, there is decentralisation of decision-making (arrows identified by the letter 'A') but it is the central authority that defines the targets for resource allocation and the intention is to have *agents* that seek to implement these objectives, hence, the emphasis on mechanisms of incentives and sanctions that are likely to generate this result (arrows identified by the letter 'B'). The dashed arrows "C" in the upper half of Figure 2 mean that the population's welfare might be implicit in the set of objectives defined by the central authority, i.e., it might be implicit that the central authority acts as an *agent* of the population (the ultimate *principal*).

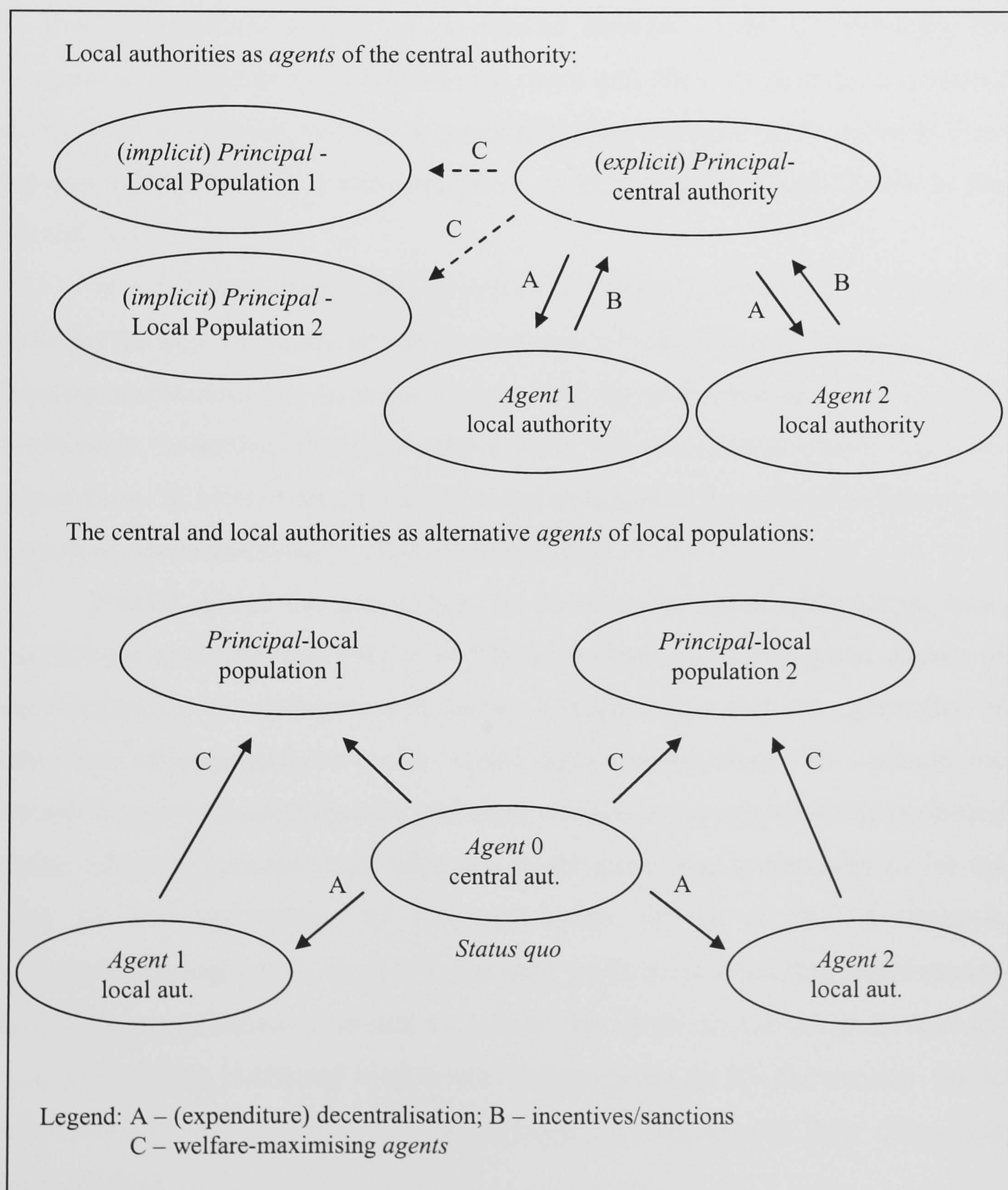


Figure 2. An adaptation of the principal-agent approach

In the bottom half of Figure 2, the analysis shifts its focus, presenting the central and local authorities as explicit alternative *agents* deciding on behalf of the respective populations (note that each local authority acts as the *agent* of the respective local population, while the central authority acts as the *agent* of all local populations). We assume that decision-makers are benevolent and therefore seek to maximise the (respective) population's welfare. The idea is to compare, for each local population, the relative performance of the central and local authorities, in meeting its preferences. For local population 1, for example, if 'Agent 1' is expected to produce greater welfare for this population than 'Agent 0', then it is optimal to rely on the former. Because 'Agent 0' represents the central decision-maker and is taken as the *status quo*, then we say that it is optimal to decentralise decision making. A major difference between the two perspectives represented in Figure 2 is the assumption of benevolent decision-makers in the second case.

In the adapted framework described in Figure 2, information asymmetry between the alternative *agents* might determine which of them performs better in meeting preferences but it might also be that, beyond information asymmetry, preferences themselves determine which agent achieves greater welfare for local populations. In other words, it might be that preferences themselves influence the impact of decentralisation on allocative efficiency.

Finally, given the assumption of diversity in preferences across local populations and admitting the possibility of preferences affecting the impact of decentralisation on allocative efficiency, it might occur that the optimality or otherwise of decentralisation also varies across jurisdictions. To evaluate the overall impact of decentralisation on social welfare, an aggregation rule is needed at the interjurisdictional level. This rule might stem from preferences as for the case of the aggregation of the individuals' levels of the desideratum (intrajurisdictional level), though in this case preferences regarding distributional issues among jurisdictions should stem from the whole population given that the whole population is affected by this rule of aggregation. In this dissertation, we do not however dwell on the issue of aggregation rules at this level (involving jurisdictions).

In summary, the key assumptions for the following discussions are: decision-makers are benevolent and their objective is to maximise the welfare of



their respective populations; central authorities are better informed than local authorities in terms of technical issues while the opposite occurs for the case of public (local) preferences; participation increases with decentralisation; and there is diversity, across jurisdictions, in preferences.

### **3.3. The trade-off between information regarding technical knowledge and information regarding preferences**

In this section, the discussion focuses only on the two types of information asymmetry assumed above and analyses the impact of decentralisation on allocative efficiency under these circumstances.

We start by considering a jurisdiction  $j$ , with two residents,  $X$  and  $Y$ , and whose health possibility frontier,  $f_{jY}f_{jX}$ , is shown in Figure 3 (note that the origin does not correspond to zero health; it represents instead the endowment point as defined in Figure 1 in chapter two). Resource allocation decisions involve the allocation of resources among alternative health care interventions and among  $X$  and  $Y$ .

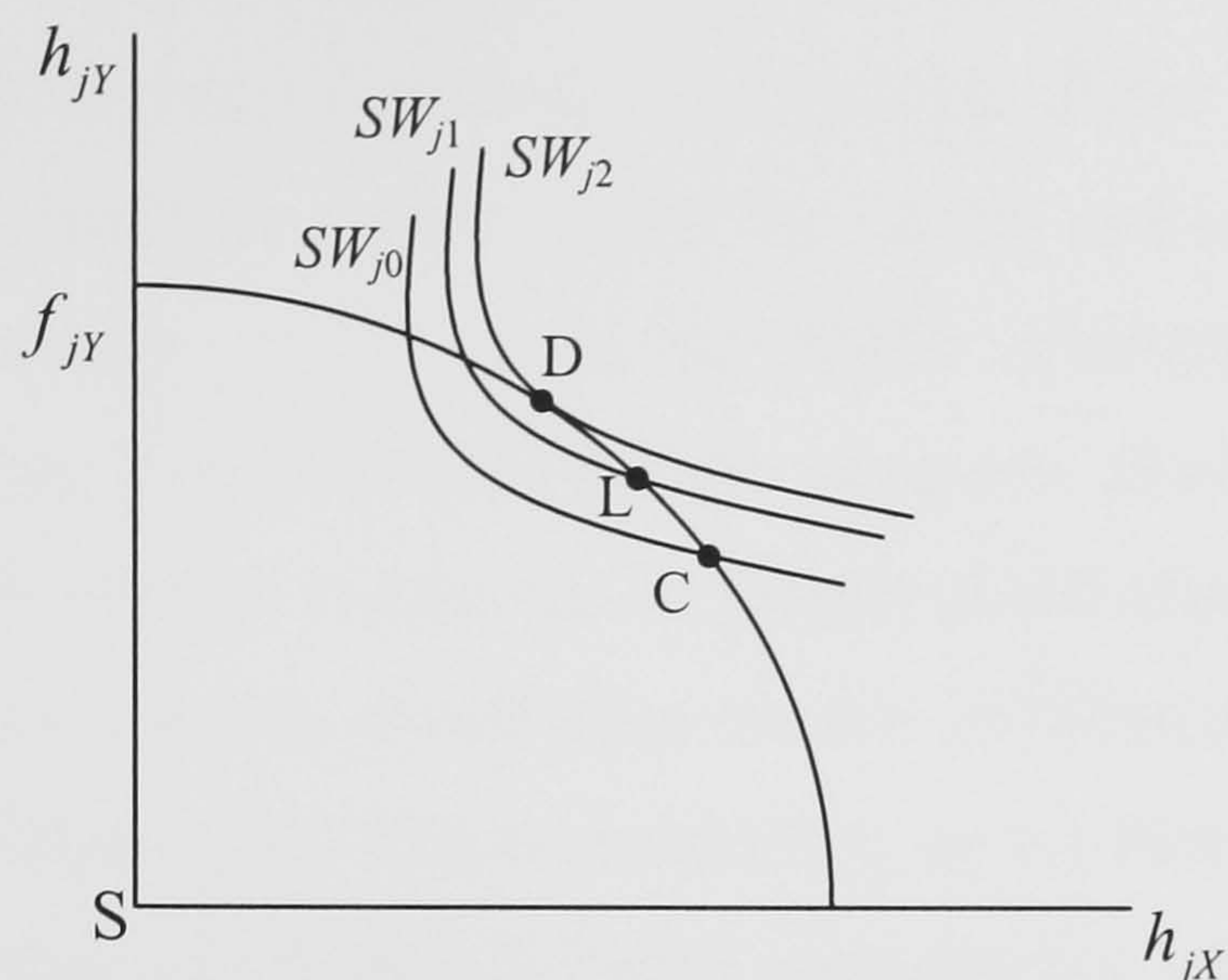
Let us consider for the moment imperfect information (and hence information asymmetry between levels of decision making) only for the case of preferences. This means that, given an allocation of resources between  $X$  and  $Y$ , the final distribution of health lies always on the frontier. Assume that preferences in jurisdiction  $j$  are such that the preferred distribution is that represented by point  $D$ <sup>19</sup>, in Figure 3. The convex indifference curves shown in this figure further indicate that we assume that the public is willing to accept deviations from the preferred distribution,  $D$ , if in return jurisdiction  $j$ 's total health ( $h_{jY} + h_{jX}$ ) also increases but the shape of indifference curves is not a particularly relevant issue for the analysis.

Suppose that point  $L$ , in Figure 3, gives the distribution reached by decentralised resource allocation, while point  $C$  represents the distribution achieved by centralised allocation. The distance between  $D$  and  $L$  is lower than

---

<sup>19</sup> The preferred distribution might result from a variety of factors discussed in the previous chapter. In chapter two we also saw that empirical evidence suggests that people are willing to trade-off different objectives, hence, the preferred distribution might itself derive from a combination of factors.

the distance between D and C due to our assumption about information asymmetry regarding preferences.



**Figure 3. Gain in social welfare due to information asymmetry about local preferences**

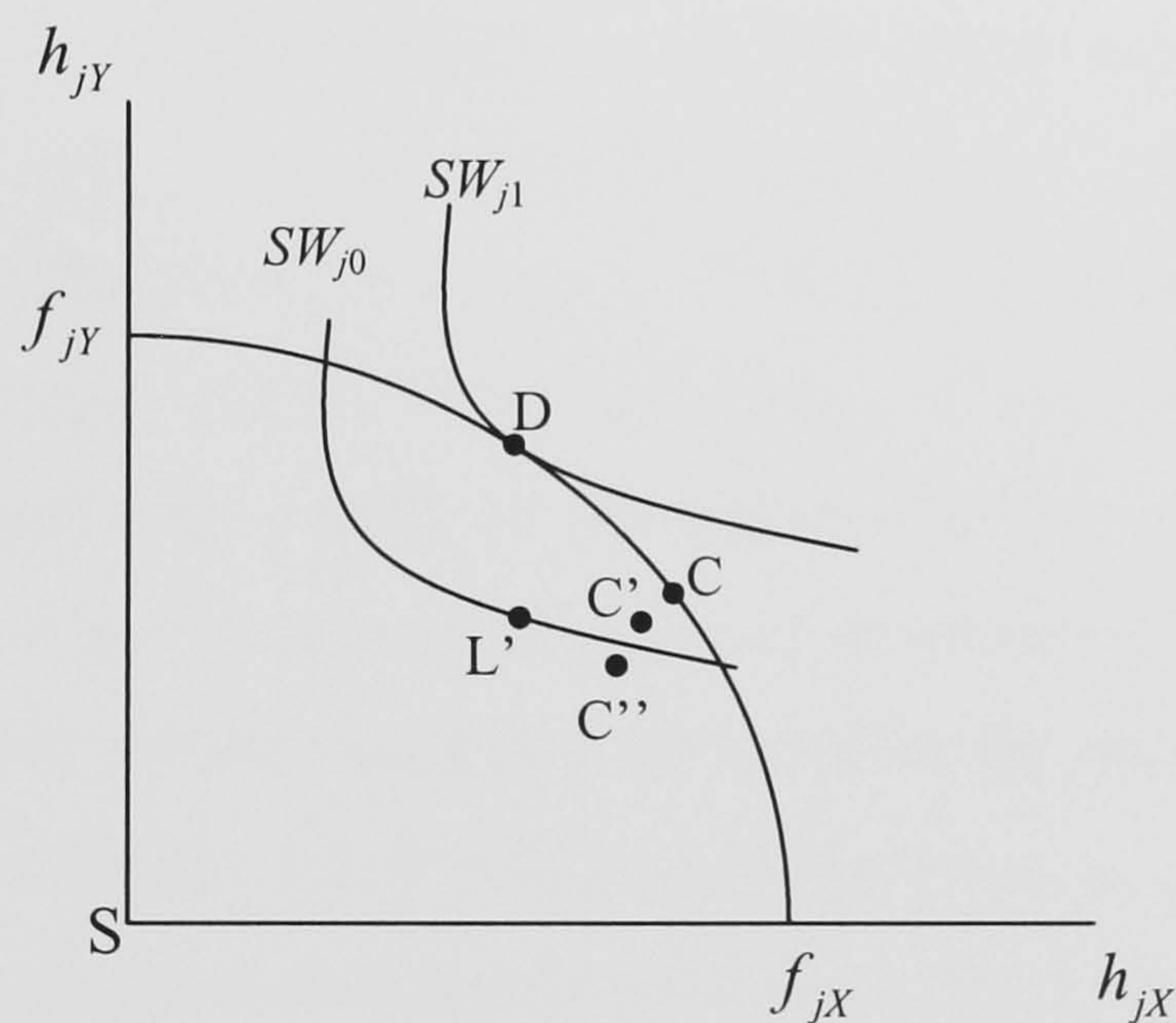
The difference  $SW_{j2}-SW_{j0}$  is the loss of social welfare due to imperfect information under centralisation and the difference  $SW_{j2}-SW_{j1}$  is the loss of social welfare due to imperfect information under decentralisation. Therefore, in this case, decentralisation increases allocative efficiency leading to an improvement in  $j$ 's social welfare represented by the difference  $SW_{j1}-SW_{j0}$ .

Let us now consider imperfect information regarding both preferences and technical issues. In this case, the former is likely to lead to deviations from point D (the preferred distribution) and given a distribution of resources between  $X$  and  $Y$ , imperfect information regarding technical matters is likely to lead to deviations from the frontier.

Due to our assumption about information asymmetry in terms of technical knowledge, deviations from the frontier are lower under centralised than decentralised allocations. Decentralisation will still have a positive impact on allocative efficiency if greater deviations from the frontier are outweighed by lower deviations from the preferred distribution. Decentralisation creates a trade-off between the two sources of welfare: on one hand it approximates the health distribution to the preferred distribution but on the other hand it implies greater losses in social welfare due to health gain foregone caused by greater production inefficiency.

Admit that point  $L'$  in Figure 4 represents the health distribution effectively achieved in jurisdiction  $j$ , in a context of decentralised allocation.

$SW_{j1} - SW_{j0}$  is in this case the total loss of social welfare in jurisdiction  $j$ , under decentralisation, due to imperfect information. Whether or not decentralisation improves  $j$ 's social welfare depends on the *status quo*, that is, it depends on the health distribution achieved by centralised allocation. If the latter is represented by e.g. C' then  $j$ 's population is better-off if the central authority continues to act as its *agent*. Contrarily, if C'' gives the health distribution obtained with centralised allocation, then decentralisation increases allocative efficiency. It would be indifferent to have centralised or decentralised allocations if the *status quo* corresponded to a health distribution on the indifference curve associated with the level of welfare  $SW_{j0}$ . Diagrammatically, as the distance between C and D expands, the point of indifference between centralisation and decentralisation is compatible with an increasing distance between D and L' (larger welfare losses, under centralisation, due to imperfect information about preferences allow larger welfare losses, under decentralisation, due to production inefficiency).



**Figure 4. Change in social welfare due to information asymmetry about preferences and technical knowledge**

If the *status quo* (centralised allocations) in jurisdiction  $A$  corresponds to a health distribution on the area filled with horizontal lines plus the area filled with crossed lines, then  $A$ 's local population is better-off if the central decision-maker continues to act as its *agent*. On the contrary, if the *status quo* is defined by a health distribution off this area, then decentralisation is optimal in jurisdiction  $A$ . The same happens for the case of jurisdiction  $B$  and the area filled with vertical lines plus the area filled with crossed lines.

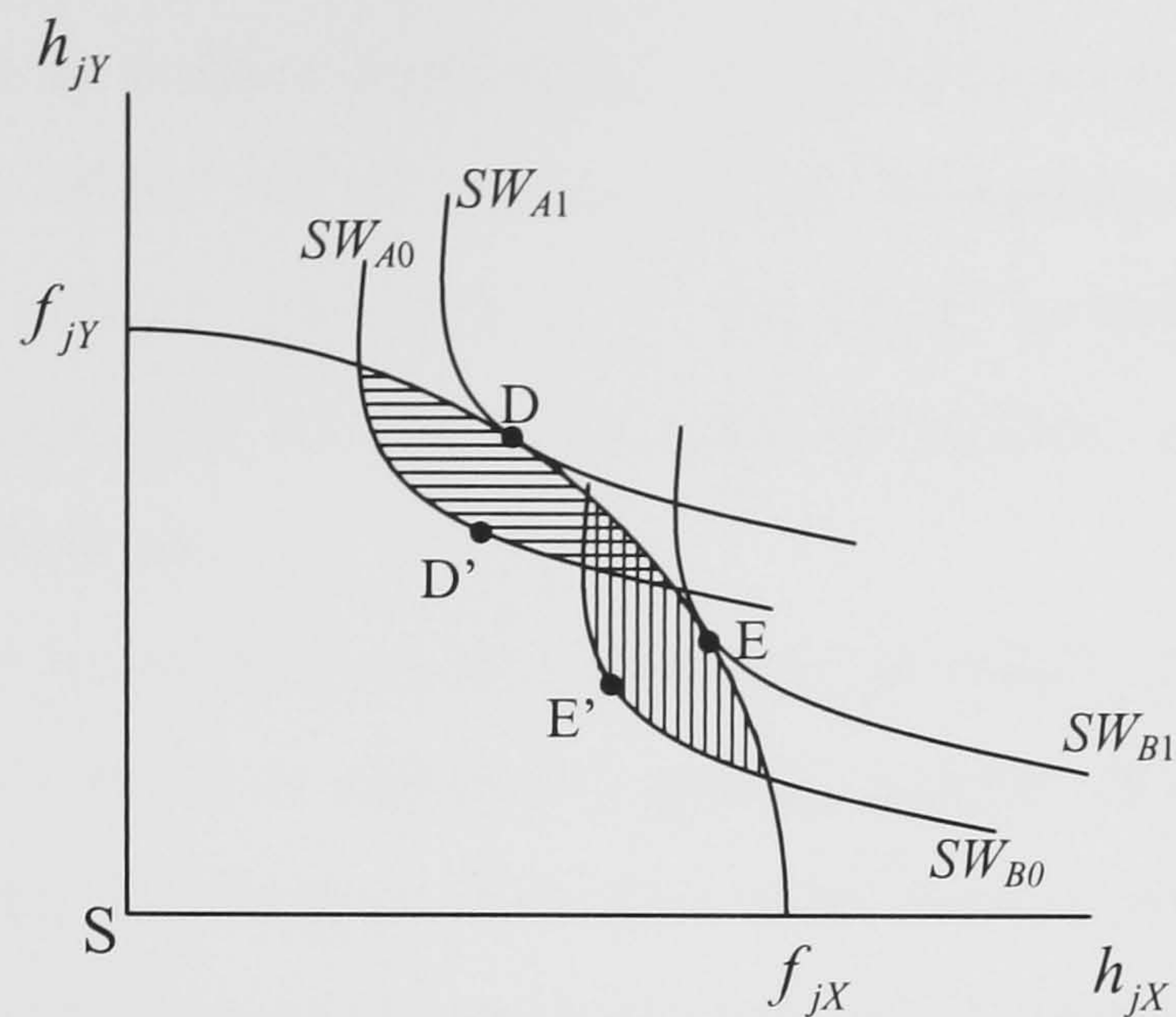


Figure 5. *Changes in social welfare given information asymmetry and variation in preferences*

Decentralisation is non-optimal simultaneously for jurisdictions  $A$  and  $B$ , if their health distributions associated with centralisation are anywhere on the filled sections of Figure 5. Decentralisation is welfare improving simultaneously for both jurisdictions only if their respective *status quo* correspond to health distributions on the white area, in Figure 5, defined by the axes, the health frontier and social indifference curves  $SW_{A0}$  and  $SW_{B0}$ . There is the possibility of decentralisation improving welfare in one jurisdiction but not in the other. For example, if the *status quo* in both jurisdictions corresponds to a health distribution on the area filled with vertical lines, then decentralisation increases social welfare in jurisdiction  $A$  but lowers social welfare in jurisdiction  $B$ . The opposite occurs for the case of the area filled with horizontal lines. In these latter situations, unless decentralisation is differently implemented across jurisdictions, depending on its effect on allocative efficiency within each jurisdiction, an aggregation rule is necessary to add up the changes in the jurisdictions' levels of welfare to reach a conclusion on the overall optimality of decentralisation compared to centralisation.

In the contexts described in Figures 3 to 5, if, for any distribution of resources between  $X$  and  $Y$ , the only thing that mattered to people was total health (that is, if the only thing that mattered was production efficiency), centralisation would be preferred to decentralisation, given that the former involves a lower waste of resources due to imperfect technical information. In this sense, by

including or not distributional concerns, local preferences determine whether or not decentralisation is welfare improving, compared to centralisation. However, this is an indirect effect via the argument of information asymmetry about preferences; with perfect information, it would be possible for the central decision-maker to allocate resources in order to achieve different targets for different local populations.

There might be an intermediate scenario, in which the central and local decision-makers do not act as alternative *agents*; instead, they share information and decision-making. Information may flow from local to central authorities or the other way around. One of the relevant aspects of information identified by Hurley *et al.*'s (1995) is precisely its communicability. This characteristic might help to determine from where to where information should flow. Following these authors, information about preferences is more difficult to transmit without loss of meaning than technical knowledge. Consequently, the intermediate scenario could be conceptualised as the central decision-maker deciding about technical matters, leaving to the local sphere of responsibility decisions about the allocation of resources among (independent)<sup>20</sup> interventions or groups of individuals.

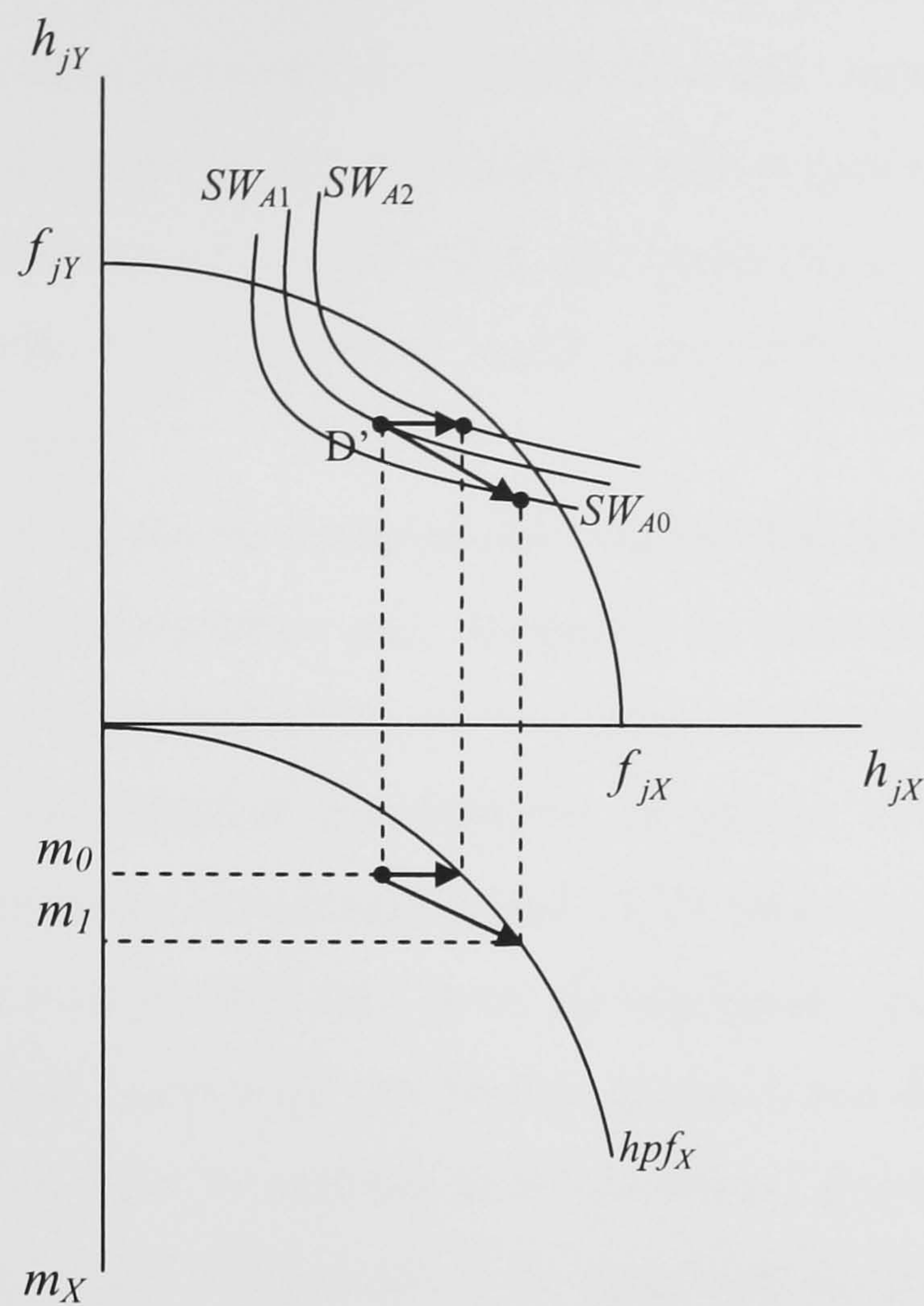
A possible way of combining the information held at both levels of decision making is to incorporate (central) technical knowledge into local decisions through centrally defined clinical guidelines for the management of specific clinical conditions. Suppose, for example, that the bottom half of Figure 6 shows the health production function for the case of individual  $X$ , living in  $A$  ( $m_x$  is the amount of resources allocated to this representative individual).

Clinical guidelines that help local authorities to move towards health production functions, given their current level of spending, always increase social welfare. For example, suppose that, in Figure 6,  $m_0$  corresponds to the amount of resources underlying the health distribution  $D'$ , obtained with decentralised allocations. Clinical guidelines that lead to an increase of the health gain obtained by  $X$ , given  $m_0$  resources, can generate an increase of social welfare up to  $SW_{A2} - SW_{A1}$  (as represented in the upper quadrant in Figure 6), compared to decentralised allocation in the absence of central guidelines. Nonetheless, results might be different if central clinical guidelines are cost-increasing compared to the

---

<sup>20</sup>For each level of resources, deciding among mutually exclusive intervention is above all a technical issue.

solution that emerges with decentralised decision making.



**Figure 6. 'Resource constrained' clinical guidelines as a mechanism to pass technical knowledge into local decisions**

Suppose that  $m_1$ , in Figure 6 is the amount of resources necessary to implement central guidelines for the management of the clinical condition represented by individual  $X$ . Compared to decentralised allocations without central guidelines, the movement towards the health production function, still represents an improvement in terms of  $X$ -efficiency. The question is that, if central guidelines themselves determine the amount of resources to be allocated to individual  $X$ , then, there is no decision resting on the local authority. This means that local knowledge about preferences is not incorporated into decisions. In Figure 6, being on  $X$ 's health production function, given the amount of resources  $m_1$ , actually entails lower social welfare in jurisdiction  $A$ , compared to what would be achieved by decentralised decisions. The difference  $SW_{A1}-SW_{A0}$ , represents (by assumption) the loss in social welfare created by a cost-increasing clinical guideline. Note that if the clinical guideline sets a greater amount of resources to individual  $X$ , compared to  $m_0$ , then, individual  $Y$  gets fewer resources compared to the amount that he or she receives in the absence of the guideline.

To conclude, if there is information asymmetry regarding technical knowledge as well as local preferences regarding health distribution, then, whether decentralisation increases social welfare or not, depends on the trade-off created by decentralisation: on the one hand, it leads to greater social welfare due to a better match between service provision and preferences; on the other, it leads to lower social welfare due to lower health gain obtained with each possible distribution of resources.

There is nonetheless an intermediate solution that increases social welfare compared to total centralisation and compared to total decentralisation. This solution entails the decentralisation of decisions concerning the allocation of resources between independent interventions or groups of individuals (in our framework, represented by individuals  $X$  and  $Y$ ), in which local knowledge about preferences can be incorporated into these decentralised decisions. At the same time, central technical knowledge can be incorporated into decisions via clinical guidelines, indicating how to manage each condition, given different levels of resources. The role of the centre here is to help local decision-makers to reach health production functions.

Under the intermediate solution, the recommendation of a specific medicine, for instance, applicable to a particular condition, might be conditioned by the amount of resources made available for this condition. We could think of these as being ‘resource-constrained’ clinical guidelines. For instance, in Figure 6, moving from  $m_0$  to  $m_1$  might allow, say, providing a new drug to  $X$ . Moreover, given the level of resources  $m_1$ , it is cost-effective to provide the new drug to individual  $X$ . But whether or not the new drug should be provided depends on whether or not it is optimal to allocate more resources to  $X$  at the expense of fewer resources allocated to  $Y$ . In this intermediate scenario, if local decision-makers decide within ‘resource-constrained’ central guidelines, there is the guarantee that variations in health care provision do not stem from differences in terms of local availability of technical information and/or local capacity to transform it into usable technical knowledge; differences derive from different preferences for health distribution. And, from the perspective of maximising social welfare, these differences are optimal.

In Figure 6, we represented the case of central guidelines affecting only individual  $X$ . When there are central guidelines that are cost-increasing for both

individuals, compared to the health distribution  $D'$ , then, it is clearly not possible to meet both guidelines, within fixed budgets. If it is not possible to meet both guidelines, anyway, there might be room for decentralised decisions (in the sense of which guideline to left partly unmet). The situation represented in Figure 6, by the movement from  $m_0$  to  $m_1$ , is still valid when there are central guidelines for some conditions but not for others. Although this result is not intended, those conditions for which guidelines are available end up being given greater priority over other conditions. For example, in England, a distinction has already been made, in terms of priority for resource allocation, between 'NICE approved technologies' and 'non-NICE technology' (Davies and Littlejohns, 2002)<sup>21</sup>.

### 3.4. Preferences for equality of health care provision between jurisdictions

In this section, we focus on equality of health care provision across jurisdictions, assuming that this issue is relevant to local populations' evaluation of alternative social states. Recall from chapter two that it has been acknowledged that inequalities in health care *per se* might have relevance to social justice and to health equity (Sen, 2002). Also, in England, NICE has been associated with equal provision of health care across jurisdictions as it was created partly to put an end on the phenomenon that became known as 'postcode lottery'<sup>22</sup>.

Given our assumption that the health frontier is the same for jurisdictions  $A$  and  $B$ , an equal allocation of resources (to individuals  $X$  and  $Y$ ) across jurisdictions also leads to an equal health distribution (between  $X$  and  $Y$ ) across jurisdictions. Thus, in this section, the latter objective should be seen simply as a proxy for equal health care provision.

In Figure 7, points  $D$  and  $E$  represent (as in Figure 5) the preferred health

---

<sup>21</sup> NICE produces guidance in three areas of health: health technologies (guidance on the use of new and existing medicines, treatments and procedures within the NHS); clinical practice (guidance on the appropriate treatment and care of people with specific diseases and conditions within the NHS); and the recently extended area of public health. The status of NICE guidance is defined as advice which should be fully taken into account by clinicians and NHS organisations but in January 2002 the UK government announced a statutory obligation for the NHS in England to provide funding for treatments and drugs recommended by NICE as a part of its technology appraisals work programme.

<sup>22</sup> Depending on the area of residence, people would get certain expenses reimbursed or not.



distributions, by  $A$ 's and  $B$ 's local populations, respectively. However, we assume that people (in both jurisdictions) are willing to give up, to some extent, of this preferred distribution in order to have equal health care provision across jurisdictions. Or, given the preferred health distribution, people are willing to give up of some health gain in order to obtain in return equality of provision between the two jurisdictions. Suppose then that the dashed indifference curves, in Figure 7, represent the trade-off that people are willing to make between equality of health provision and other considerations affecting social welfare<sup>23</sup>.

For the case of jurisdiction  $A$ , for instance, any distribution of health, involving equality of health care provision, located between the indifference curves corresponding to  $SW_{A1}$  and  $SW_{A0}$  is always preferred to any other health distribution. This implies that even when compared to point  $D$ , a distribution inside the frontier is preferred provided that it entails equal provision across jurisdictions. In the case of jurisdiction  $B$ , any distribution of health, involving equality of health care provision located between the indifference curves corresponding to  $SW_{B1}$  and  $SW_{B0}$  is always preferred to any other health distribution. In the shaded area, any distribution of health, involving equality of health care provision is preferred to any other health distribution simultaneously by both local populations.

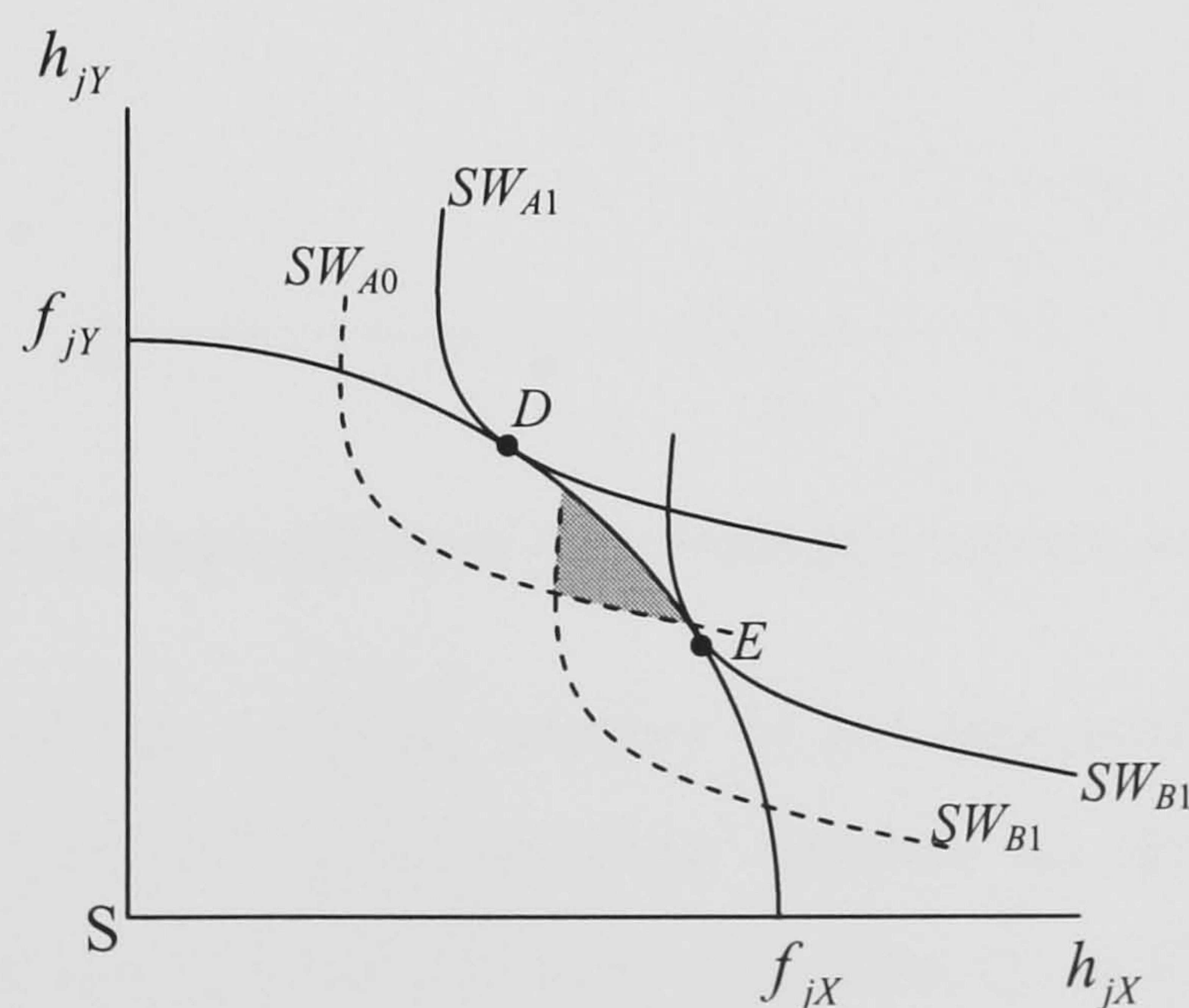


Figure 7. Preferences for equal health care provision

Figure 8 is adapted from Figure 7. Suppose that points  $D'$  and  $E'$  represent

<sup>23</sup> In chapter four, we present a questionnaire in which questions 3 and 4 are examples of how a trade-off of this kind might be operationalised in practice.

the health distributions effectively achieved under decentralised allocation (note that deviations from D and E stem from imperfect information about the populations' preferred distributions as well as from imperfect technical information that causes deviations from the frontier).

Point C represents the preferred health distribution taking into account the whole population (from both jurisdictions); it therefore lies between D and E. Given our assumption that both local populations value equal treatment across jurisdictions, the central authority takes these preferences into account and therefore allocates health care resources (between X and Y) in the same way in both jurisdictions. The reference health distribution should consequently reflect the preferences of both local populations. Assume then that point C' shows the health distribution obtained with centralised decision making (hence it gives the *status quo*), in which point C' is based on estimates of point C (deviations between the two points are also explained by imperfect technical information).

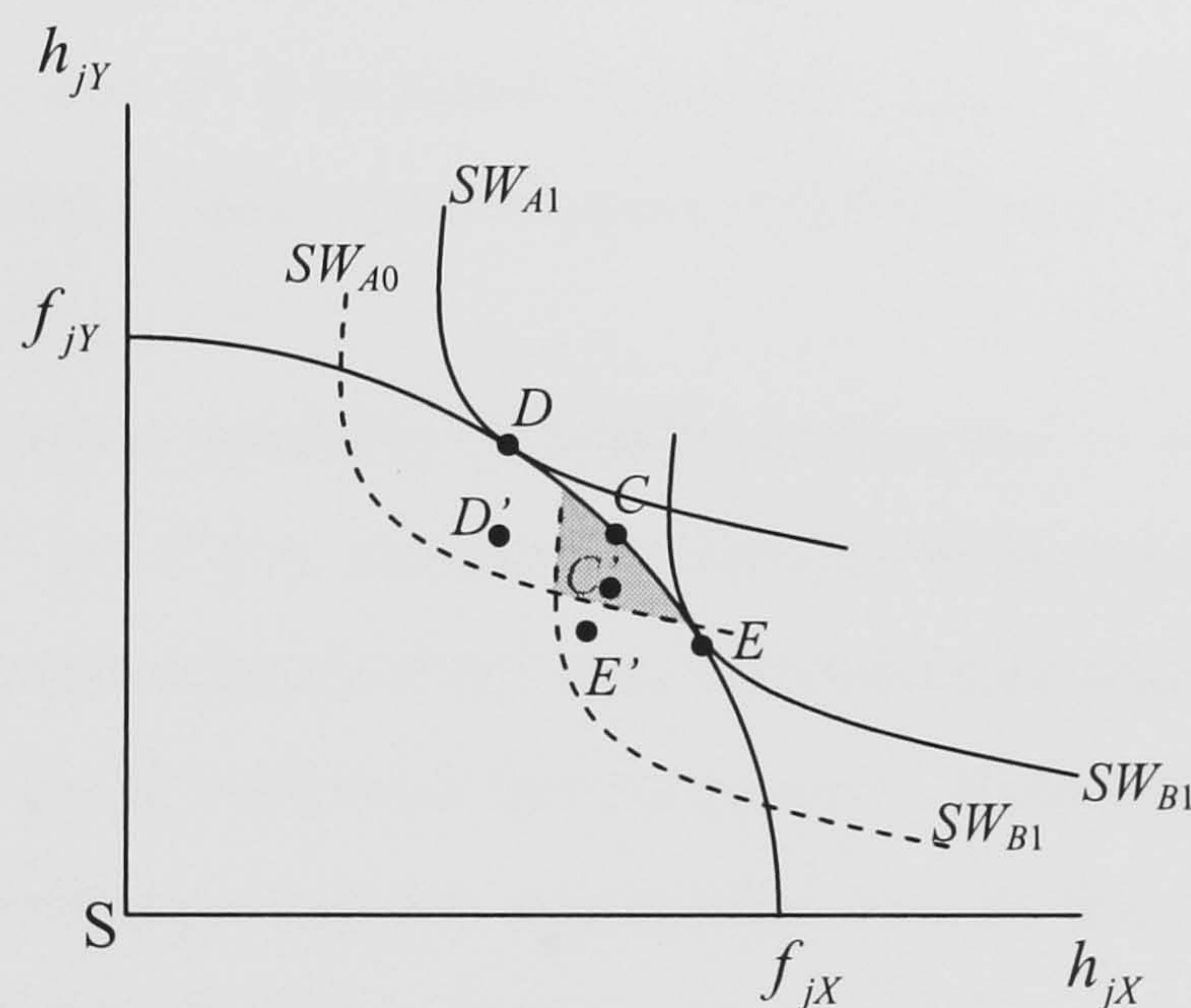


Figure 8. *Convergent effects of decentralisation across jurisdictions*

In the face of this scenario, whether or not decentralisation of resource allocation is optimal for local populations depends on the welfare changes generated by the movements from C' to D' and from C' to E'. In this particular example, both local populations would be worse-off with decentralised allocations. Let us analyse first the case of jurisdiction A. We said above that, given preferences for equal health care provision, any distribution (involving equal provision) between the indifference curves corresponding to  $SW_{A1}$  and  $SW_{A0}$  is always preferred to any other health distribution entailing unequal provision,

thus, C' is preferred to D', meaning that decentralisation, i.e. the movement from C' to D', makes A's local population worse-off. However, if the social indifference curves (for jurisdiction A) passing through points D' and C' were depicted in Figure 8, it could be seen that D' belongs to a higher indifference curve than C'. Thus, what makes decentralisation non-optimal for jurisdiction A is preferences for equal provision themselves<sup>24</sup>. In terms of the framework described in Figure 2, one might say that A's local population is better-off if the central decision-maker continues to act as its *agent* not because he possesses greater information than local decision-makers but because changing the *agent* negatively affects the achievement of the goal of equal health care. Thus, it is the local population's objectives for resource allocation that determine which *agent* represents higher local welfare.

In the case of jurisdiction B, preferences for equal health care also make point C' preferred to point E', meaning that as for the case of jurisdiction A, decentralisation makes B's local population worse-off. In the particular example shown in Figure 8, point C' is preferred to point E' anyway (from the perspective of jurisdiction B, point C' belongs to a higher indifference curve, not shown in the diagram, than point E').

Figure 9 shows a situation in which preferences in A are such that the movement from C' to D' (i.e. decentralisation) decreases social welfare in this jurisdiction, while preferences in B are such that decentralisation (movement from C' to E') increases social welfare in this jurisdiction. Here, there are winners and losers with decentralisation and an aggregation rule is needed to evaluate the overall impact of decentralisation on efficiency.

Note that, in Figure 9 C' is off the shaded area hence the potential for conflict between jurisdictions regarding the impact of decentralisation on their levels of welfare. The trade-offs that people are willing to make between equal provision and other objectives across jurisdictions might also vary between jurisdictions. The shaded area will be larger or narrower depending on this trade-off. One can see the shaded area as a measure of the relative strength of preferences for equal health care provision: the greater the relative strength of

---

<sup>24</sup> It does not mean that under decentralisation equal provision is never achieved; it could be obtained with e.g. an agreement between local decision-makers, however, in the current framework it is not optimal to change the *status quo*.

these preferences in both jurisdictions, compared to other preferences, the larger the shaded area.

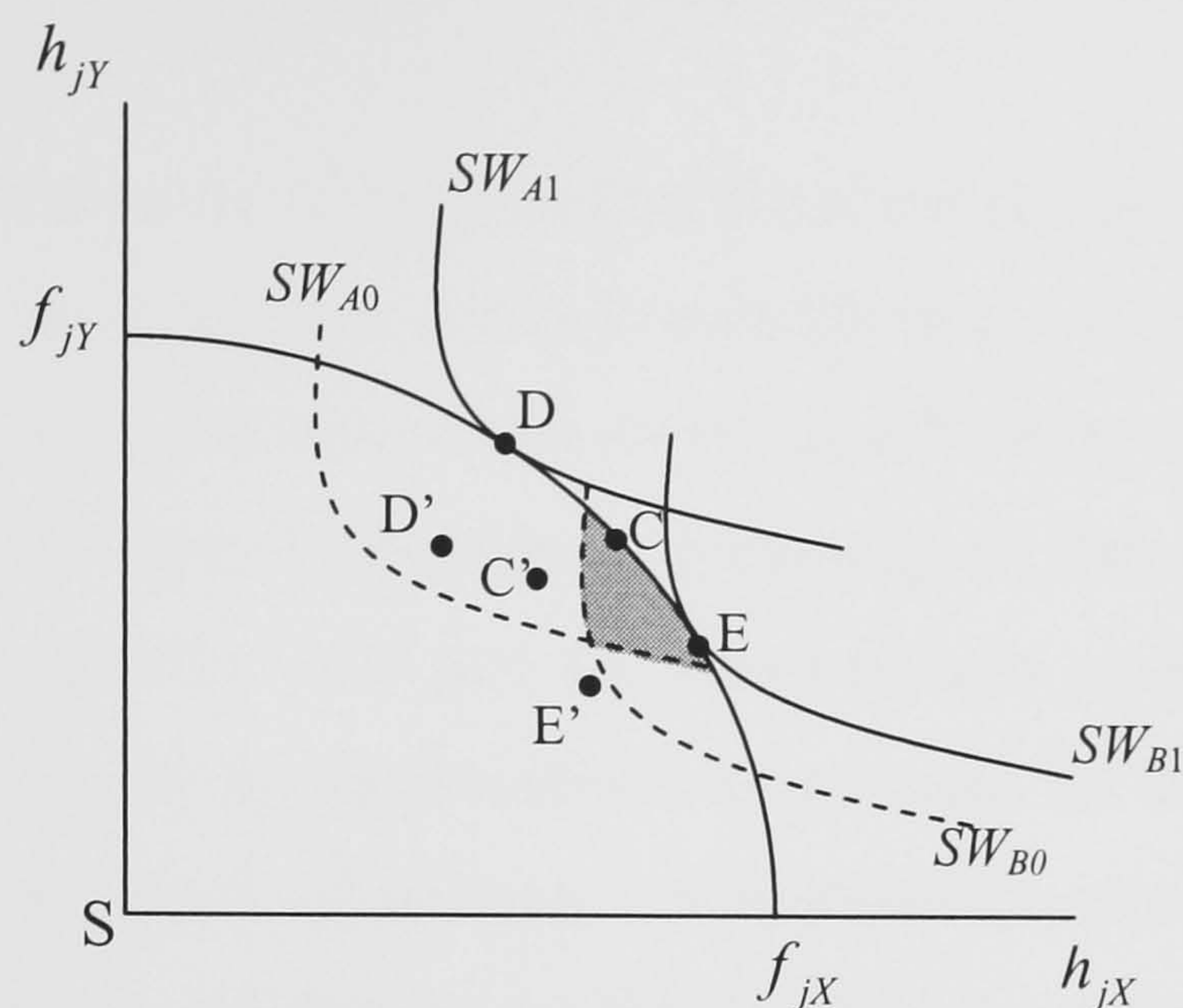


Figure 9. *Divergent effects of decentralisation across jurisdictions*

As the dashed indifference curves approximate to the frontier, the role played by preferences for equal provision in the evaluation of welfare changes produced by decentralisation becomes smaller and the analysis shifts its focus to information asymmetry as discussed in section 3.3. Further note that, while in section 3.3, for any given health distribution, points on the frontier generate greater welfare than points inside the frontier, in the current section, points inside the frontier might be preferred to points on the frontier (for the same health distribution) if the former correspond to equality of provision across jurisdictions and the latter do not.

### 3.5. Preferences for voice or participation in decision making

This section focuses on preferences over procedures, in particular, on preferences for participation in decision making or voice. Both the theoretical and empirical literature, reviewed in chapter two, suggest that participation (or voice) is valued instrumentally as well as inherently. Here, the discussion revolves around the latter aspect of the procedural characteristic of voice. Also, recall from section 3.2 that, based on the review of literature, we assume that there is a

positive relationship between decentralisation and participation.

To show that preferences for participation might affect the optimal level of decentralisation, we adapt Inman and Rubinfeld's (1997b) approach to our framework.

Inman and Rubinfeld (1997b) extend the economic approach, focusing on economic welfare, to encompass political objectives as well. They say that the twin values of political participation and economic efficiency are rarely integrated within a common conceptual framework, proposing to this effect an approach different from the suggestion in Figure 10. These authors propose the 'federalism frontier' which is formed by combinations of an index of economic efficiency (vertical axis) and an index of political participation (horizontal axis). In their framework, the former is affected by decentralisation due to spillover effects, due to economies of scale and because decentralisation 'increases the likelihood that services provided and regulations promulgated will be consistent with the desires of each member of the population' (p.1219). Regarding the index of political participation, it comprises a metric of political influence (measuring the effects of individual participation on political outcomes) and a metric of political effort (measuring the inherent value of participation, for example, by the hours given by the individual to a political activity).

Inman and Rubinfeld (1997b) consider that as decentralisation increases: initially, economic efficiency and political participation rise; then, the frontier has a downward sloping segment with economic efficiency diminishing (due to e.g. spillover effects becoming more important) and participation increasing; finally, the frontier is formed by a segment with a positive slope in which both economic efficiency and participation decrease. The two anchor points for the 'federalism frontier' are full centralisation (one national government) and full decentralisation (no government at all). The relevant range of the frontier is that in which there is a trade-off between efficiency and participation. The authors say that the institutions should be chosen in order to get us to the relevant range of the frontier; once on this segment, institutional choices should adjust in order to reflect the society's preferred mix of efficiency and participation along the frontier. Although the focus of the authors is different from the focus adopted in the current chapter, one might see that in their analysis preferences can directly influence the optimal level of decentralisation (note that Inman and Rubinfeld,

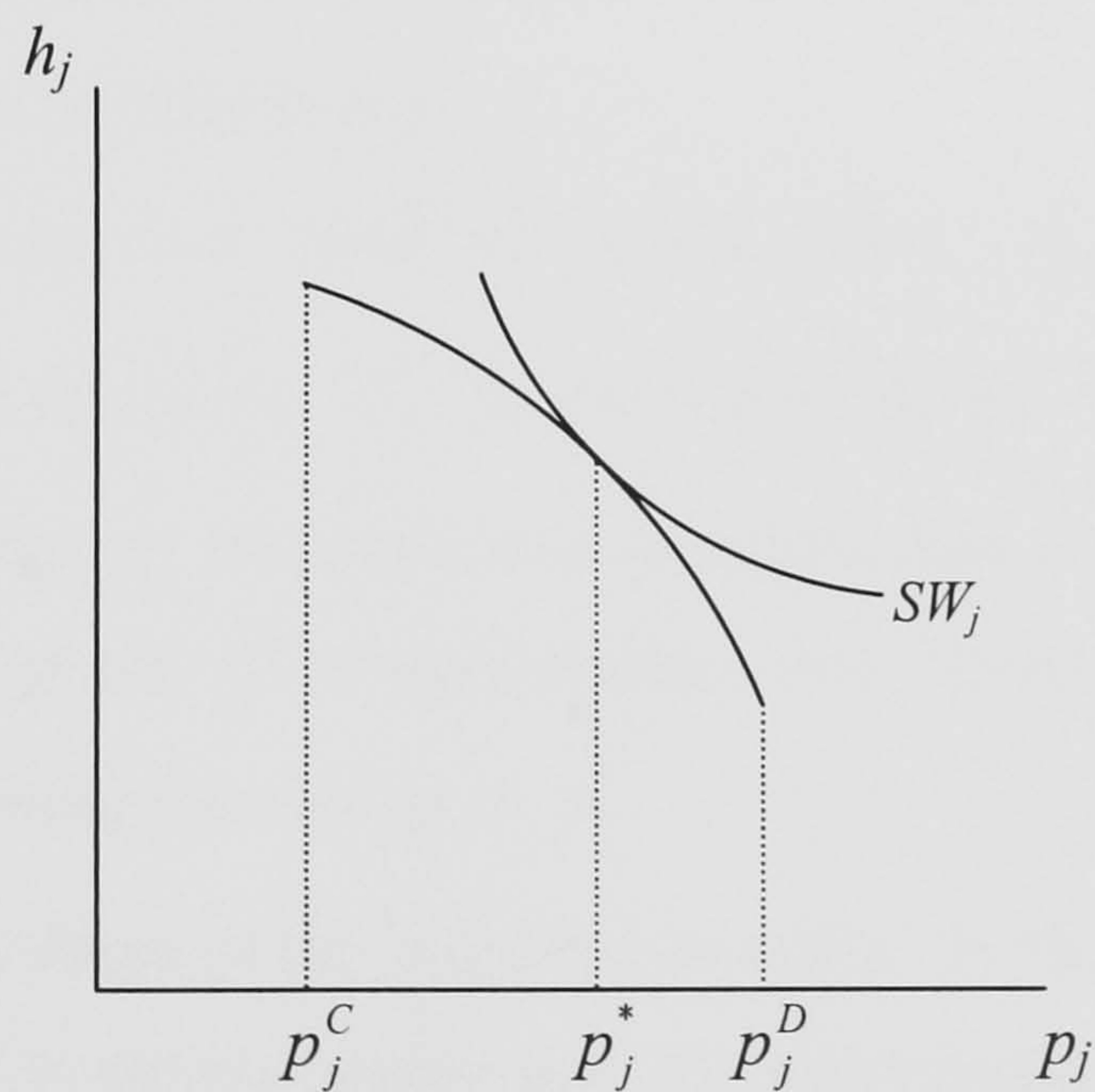
1997b, assume that participation is a continuous variable as well as decentralisation).

To adapt these authors' analysis to our framework, we define the variable 'participation' (in jurisdiction  $j$ ),  $p_j$ , as a function of decentralisation in

jurisdiction  $j$ ,  $d_j$ , being:  $p_j = p(d_j)$  and  $\frac{\partial p(d_j)}{\partial d_j} > 0$ .  $d_j$  increases as decision-

making power regarding the allocation of resources among alternative health care interventions is successively transferred from the centre to the local level (central grants also vary in this case increasing in line with augmented responsibility for resource allocation).

Further consider that preferences in jurisdiction  $j$  are such that social welfare increases with total local health (for simplicity, assume that uniform weights apply to all residents) as well as with participation in decision making regarding resource allocation in health care (we are still assuming that participation, or voice, is valued *per se*, regardless of the impact that it may have on outcomes).



**Figure 10. The impact of preferences for participation or voice on the optimal level of participation / decentralisation**

In Figure 10,  $p_j^C$  gives the level of participation that occurs with total centralisation of resource allocation decisions and  $p_j^D$  shows the level of participation that emerges with total decentralisation (i.e. all resource allocations are made by the local government).

Total health decreases with decentralisation due to information asymmetry concerning technical knowledge: more decentralisation means lower production efficiency, hence, higher levels of participation are associated with lower levels of total health. We thus have a health-participation possibility frontier with a negative slope (one might say that participation has an opportunity cost in terms of health gain foregone).

The indifference curve shown in Figure 10 has a negative slope as well, indicating that both participation and health matter for social welfare in jurisdiction  $j$ . So, if there is a decrease in participation, social welfare is kept unchanged only if total health increases, to compensate for the former welfare loss, and vice-versa. The optimal level of participation,  $p_j^*$ , occurs when an indifference curve is tangential to the health-participation possibility frontier. I.e., when the slope of the latter is the same as the slope of the indifference curve:

$$\frac{\partial h_j}{\partial p_j} = -\frac{\partial SW_j}{\partial p_j} \div \frac{\partial SW_j}{\partial h_j} .$$

In other words, the optimal level of participation is achieved when the marginal cost of participation, in terms of health foregone, is equal to the marginal rate of substitution (the rate at which society is willing to substitute health for participation).

Given the optimal level of participation,  $p_j^*$ , the optimal level of decentralisation of decisions,  $d_j^*$ , is derived from the relationship  $p_j = p(d_j)$ . The amount of resources and particular decisions that are transferred to the local decision-maker's sphere of responsibilities are those underlying the health-participation possibility frontier up to  $p_j^*$ .

The convex shape of the indifference curve, in Figure 10, is not relevant to develop our point. It simply means that the more health society has, it requires greater additional amounts of health to give up of equal additional amounts of participation. But we could obtain the same  $p_j^*$  with a linear indifference curve, provided that  $p_j^*$  corresponded to the point of tangency between the indifference curve and the health-participation possibility frontier. A linear indifference curve, for instance, implies that the rate at which society is willing to trade-off participation against health effectiveness is the same for all values of  $p_j$  and  $h_j$ .

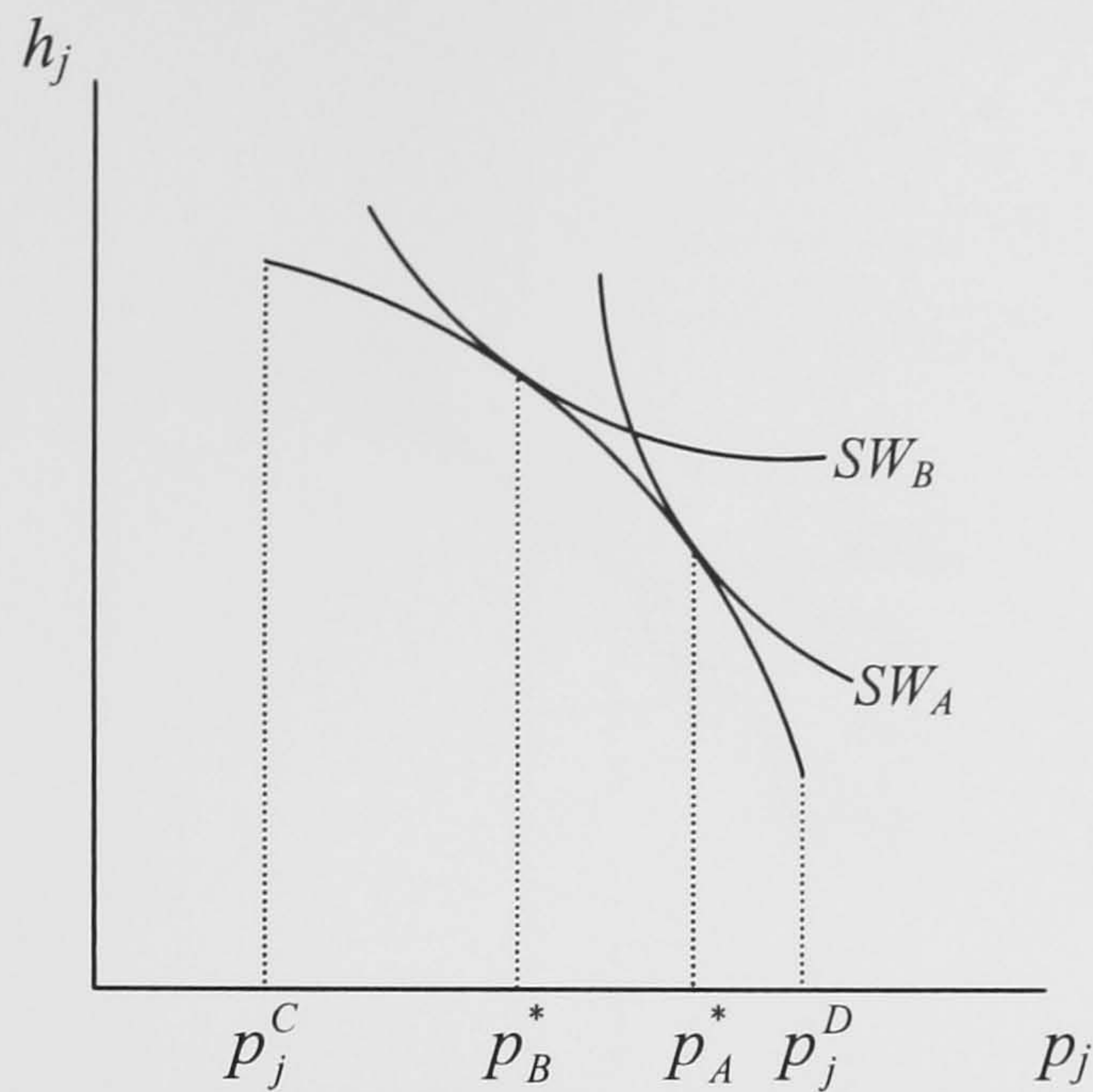


Figure 11. *Different optimal levels of participation / decentralisation as the result of different preferences*

In Figure 11, two indifference curves are represented, for jurisdictions  $A$  and  $B$ , with different slopes, their purpose being to show the effect of different preferences for participation *versus* health on the optimal level of decentralisation. The health-participation possibility frontier is assumed to be the same in both jurisdictions. The representative indifference curve of jurisdiction  $B$  (identified by the subscript  $B$ ) is flatter than the representative indifference curve of jurisdiction  $A$  (identified by the subscript  $A$ ). Thus, the same amount of participation is traded-off against a greater amount of health for the case of jurisdiction  $A$  than for the case of jurisdiction  $B$ . This implies that participation is more valuable, compared to health, to  $A$ 's residents than to  $B$ 's residents. So, with everything else the same, different preferences regarding participation *versus* health, lead to different optimal levels of participation, where  $p_A^* > p_B^*$ . Consequently, the optimal level of decentralisation also differs between these two cases:  $d_A^* > d_B^*$ .

Figure 12 shows the loss in social welfare resulting from the implementation of a unique level of decentralisation in both jurisdictions, in the presence of differences in preferences. Suppose for instance that  $d_j = \bar{d}$  for  $j=A,B$ , such that the level of participation is  $\bar{p} = p(\bar{d})$  for both jurisdictions, where  $p_B^* < \bar{p} < p_A^*$ . In this situation, the loss of social welfare in  $A$  is given, in Figure 13, by the difference  $SW_{A1} - SW_{A0}$  and the loss of social welfare in  $B$  is given by  $SW_{B1} - SW_{B0}$ , compared to the case in which different levels of



decentralisation are implemented.

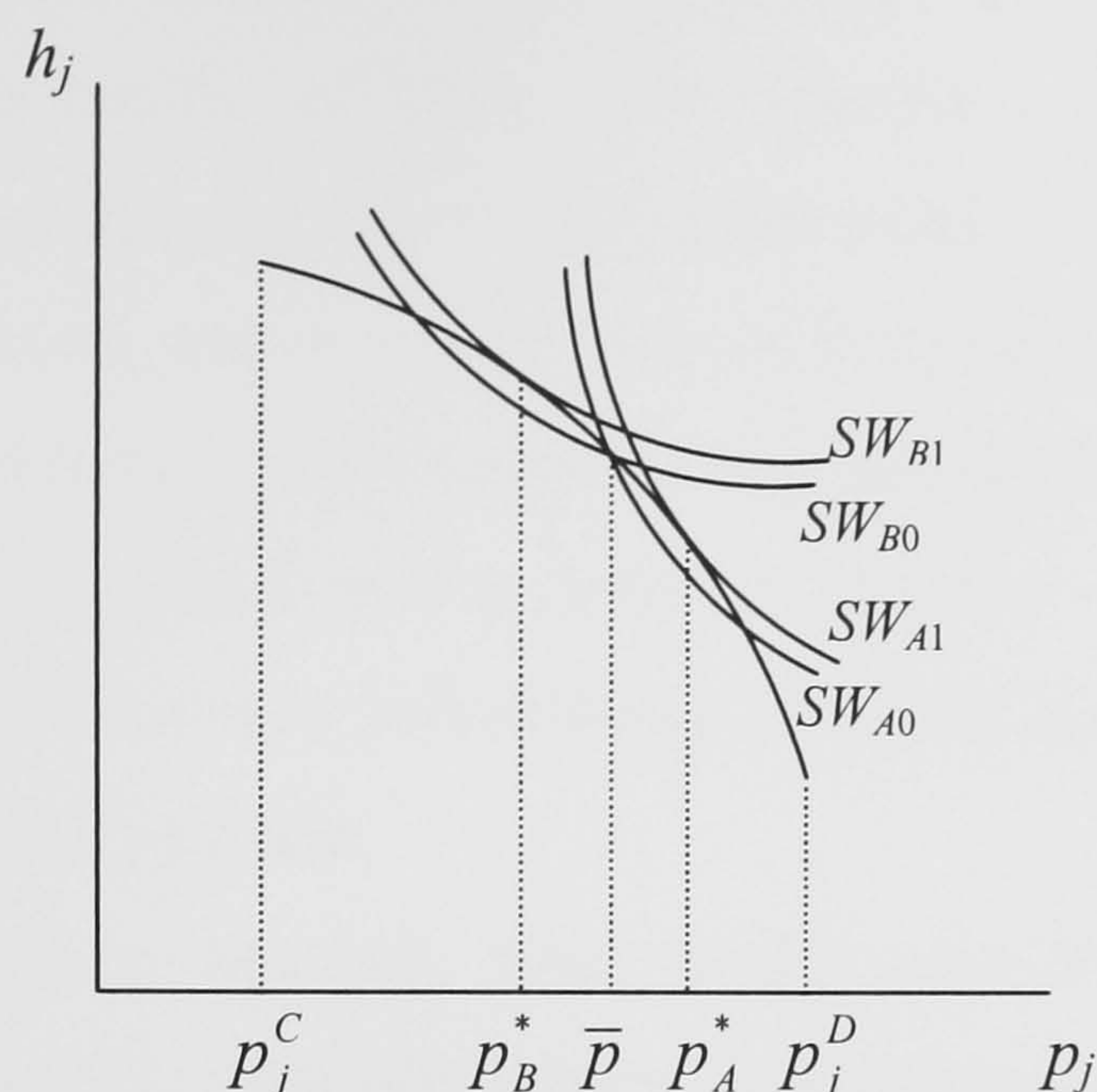


Figure 12. *Loss of social welfare created by a uniform level of decentralisation*

Finally, note that participation can impose significant opportunity (time) costs on people and significant resources might be required to promote and ensure participation. Time costs reduce the inherent benefit of participation for people though we might think about some monetary compensation and focus only on monetary costs. In this case, what might happen is that greater participation leads to lower health simply because resources are deviated from the production of health to sponsor participation (this effect reinforces the effect of information asymmetry concerning technical knowledge). Additionally, in the specific case of health care, it has been reported that denying care to others causes distress (Coast, 2001); consequently, ‘disutility of denial’ might reduce the positive impact that participation has on welfare.

### 3.6. Preferences affecting the size of local health care budgets: preferences for health *versus* non-health-related ‘enjoyment’

In the literature about decentralisation, the term externalities has been applied in cases in which the provision of services or goods within a jurisdiction benefit (or impose costs on) residents of other jurisdictions. In contexts of

decentralisation, these costs and benefits might be under-estimated by local decision-makers, thus, externalities have been identified as an argument against decentralisation. In health, an often cited example of externalities is children immunisation. In our discussion we nonetheless assume that health care provision, within each jurisdiction, has an impact only on local residents' health. Despite this assumption, local health care provision in one jurisdiction might create spillover effects on the other jurisdiction's level of welfare via preferences for the distribution of health between jurisdictions. It is this kind of externalities that is addressed in the current section.

In the previous sections, local health care budgets were given and preferences affected only their distribution within jurisdictions. In this section we focus on a case in which preferences affect the size of local budgets. However, while in the literature about decentralisation, varying local (health care) budgets have been associated mainly with variation in local taxes, here we continue to assume that local budgets are formed by central transfers only (this assumption guarantees that differences in the sizes of local health care budgets stem from different preferences across jurisdictions and not from differences in their levels of wealth; the latter might in fact lead to differences in local taxation even for constant preferences). Suppose then that the decision being considered for decentralisation is the allocation of untied resources that can be allocated to health care or to services other than health care. Here, we consider that resources are used to fund health care (generating health) or, alternatively, they are used to fund 'entertaining events' like concerts (generating 'enjoyment' non-related to health).

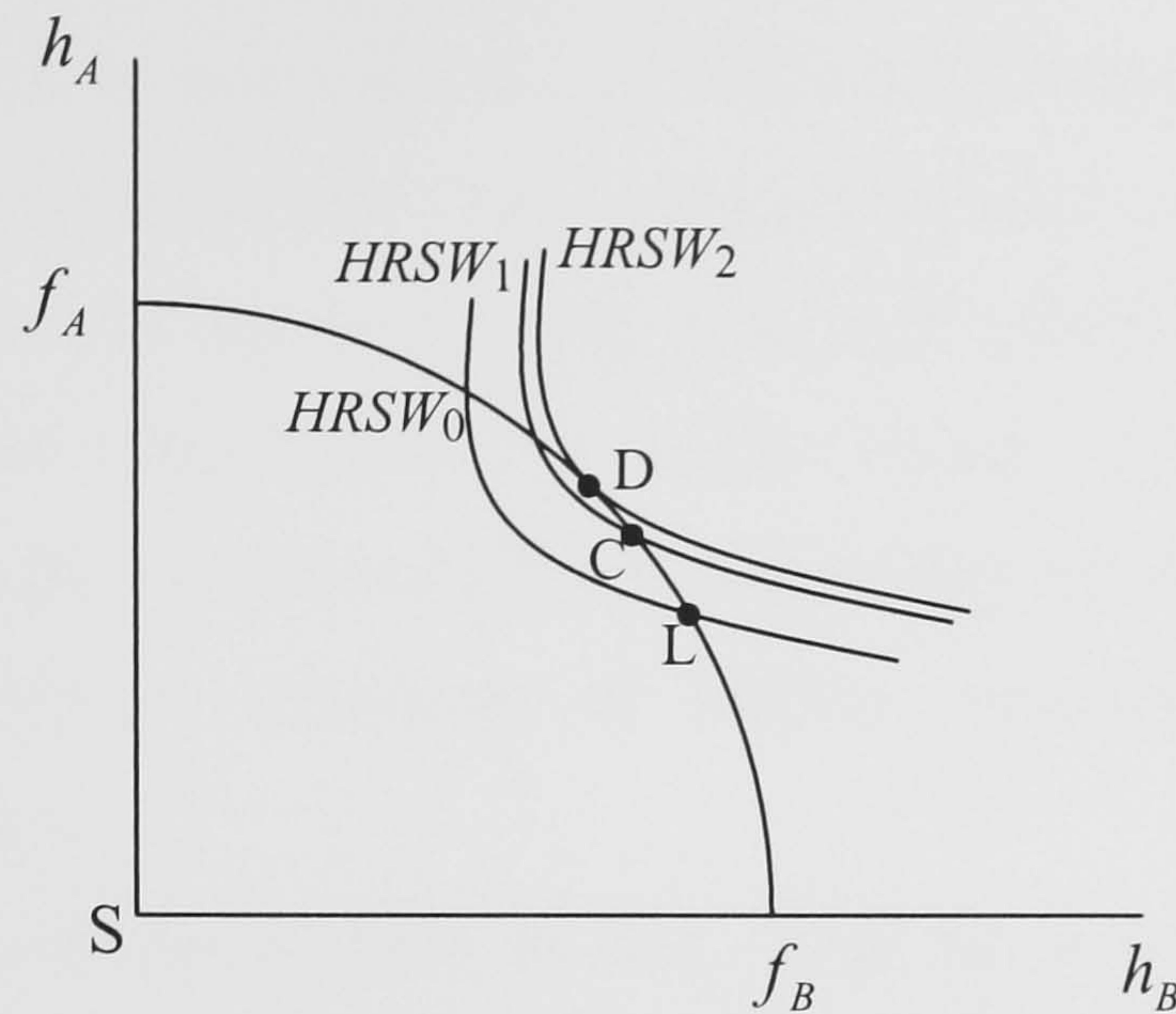
The example of 'entertaining events' is merely illustrative and it was deliberately chosen given that health has many determinants besides health care, thus, posing the question as one of distributing resources between health care and social policies such as education or housing would bring accrued difficulties to the analysis due to the impact that these policies are likely to have on health. Obviously, with a sufficiently flexible definition of health, in the limit everything affects health. So, people might feel better by attending a concert and this might be seen as a positive impact on their health (on the other hand, being healthier allows people to better enjoy the concert). However, we assume (without imposing too many restrictions, we think) that health care and 'entertaining events' have independent effects on social welfare.

We thus assume that, given preferences, social welfare within jurisdiction  $j$  is affected by total health obtained by  $j$ 's local population, by the local level of 'enjoyment' non-related to health and by the distribution of health between jurisdictions  $A$  and  $B$ . Allocating more resources to health care increases welfare via increased health. On the other hand, allocating more resources to 'entertaining events' increases welfare via increased levels of 'enjoyment'. Additionally, changes in health within jurisdictions affect the distribution of health between them and consequently affect their levels of welfare (the spill over effects previously mentioned).

We continue to assume that there is variation in preferences (in this case for health *versus* 'enjoyment') as well as information asymmetry as described in section 3.2. In the current section, one *agent* (the central decision-maker) represents greater production efficiency plus the possibility of internalising the spillover effects (that is, the central decision-maker has the capacity of affecting health distribution between jurisdictions by determining the amounts of resources allocated to health care in both jurisdictions). The other *agent* (the local decision-maker) in turn represents a distribution of resources between health care and 'entertainment' more in line with preferences (for health *versus* 'enjoyment'). Which *agent* produces higher social welfare depends on preferences in terms of the trade-off between all arguments affecting social welfare. Changing the *agent* (in the *status quo* the *agent* is the central government) might therefore have a positive or a negative impact on social welfare depending on preferences. That is, public preferences condition the impact that decentralisation has on allocative efficiency.

Figure 13 illustrates the loss in (health-related) social welfare due to externalities effects associated with preferences. To simplify the exposition we limit the analysis to points on the frontier (thus the advantage of the centre in terms of production efficiency is not considered). Moreover, we admit that decentralisation does not affect the health frontier, implying that on the whole the amount of resources allocated to health care is the same under both scenarios; what changes is the distribution of this amount between jurisdictions  $A$  and  $B$ . The health frontier for jurisdictions  $A$  and  $B$ ,  $f_A f_B$ , gives the maximum amount of health that one jurisdiction can get, given: the health of the other jurisdiction; the total amount of resources available to health care; and technology. The social

indifference curves shown give the level of health-related social welfare (Figure 13 does not incorporate ‘enjoyment’).



**Figure 13. Loss of health-related welfare as a negative 'externality effect' of local decisions**

Suppose that point D in Figure 13 represents the preferred health distribution (between *A* and *B*) by both local populations (this is in fact another simplifying assumption<sup>25</sup>). Point C represents the distribution achieved with centralised allocation and it is based on (central) estimates of point D. Under decentralised decision making, each local decision-maker is unable to manipulate the distribution of health between jurisdictions, hence, when determining the size of local health care budgets only (local) preferences for health *versus* ‘enjoyment’ are considered. Suppose that preferences for health relative to ‘enjoyment’ are weaker in jurisdiction *A* than in *B*, leading to the distribution represented by L. Weaker preferences in *A* imply that in *A* fewer resources are devoted to health care compared to *B* (hence, *A*’s health associated with point L is lower than *A*’s health associated with point C and the opposite happens for *B*’s health).

The difference  $HRSW_1 - HRSW_0$  shows the loss in health-related social welfare caused by decentralisation (movement from C to L) in each jurisdiction. This loss occurs because, by setting local health care budgets in accordance to local preferences for health *versus* ‘enjoyment’, local decision-makers moved the health distribution (between jurisdictions) farther away from the preferred

<sup>25</sup> But it could be easily relaxed by introducing two different preferred distributions on the frontier against which losses in welfare would be measured, using a different reference point for each jurisdiction.

distribution (point D). This movement can be seen as a negative externality effect of local decisions.

Although there might be negative externality effects in terms of health-related social welfare, to assess the overall impact of decentralisation on allocative efficiency, it is necessary to take into its impact on local social welfare due to a distribution of resources between health care and ‘entertaining events’, which better reflects local preferences for health *versus* ‘enjoyment’. Therefore, decentralisation might be optimal or not, depending on how local populations trade-off these different arguments of welfare, involving health and other dimensions of welfare.

The main conclusion here is that, even in a context of information asymmetry regarding preferences, decentralisation might not be optimal due to externality effects of local decisions. Given that these externality effects are associated with preferences, then, here too we conclude that preferences might themselves be seen as a basis to determine the optimality or otherwise of decentralisation.

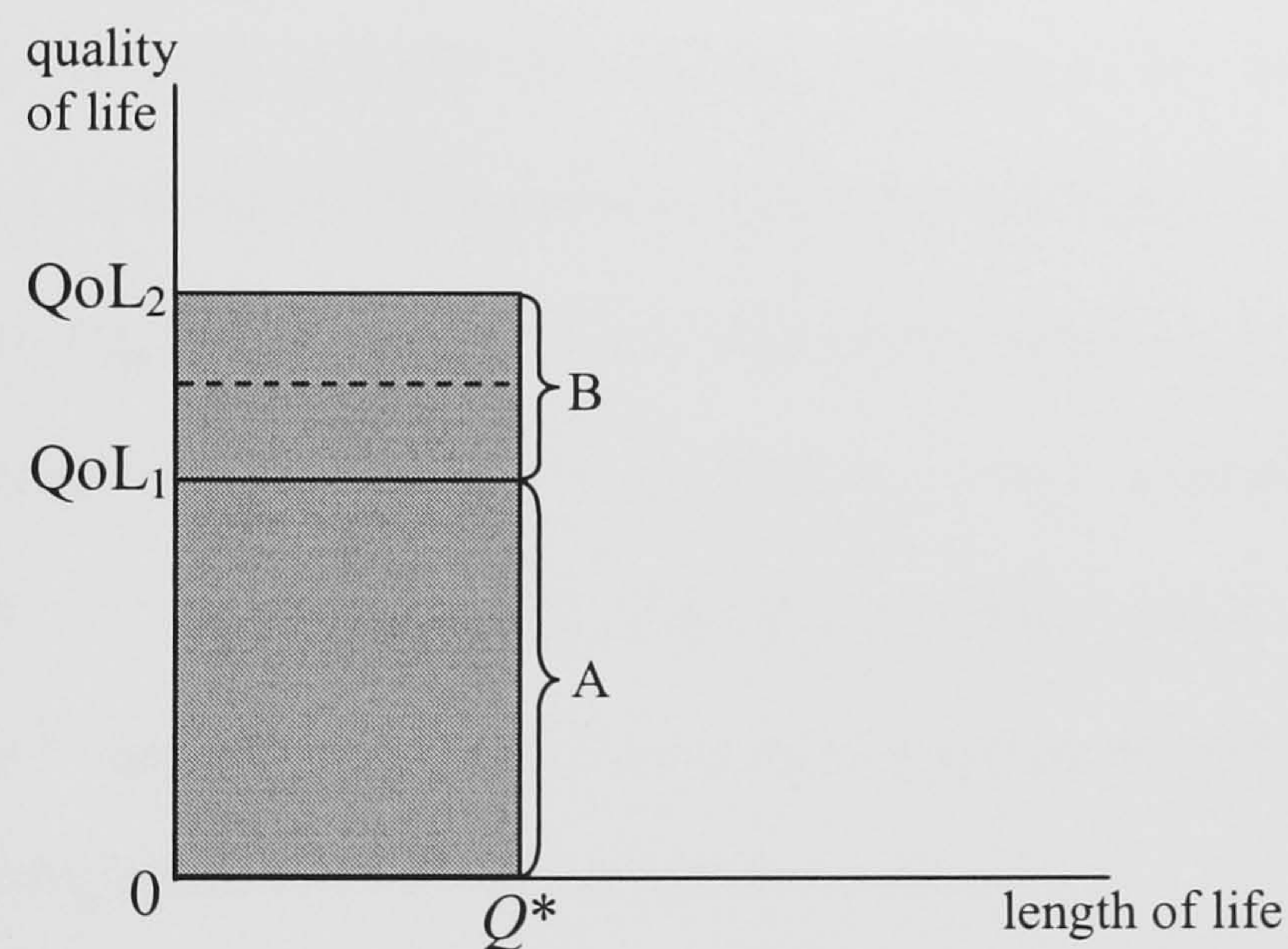
### **3.7. Is it optimal to allocate resources between jurisdictions based on local preferences?**

We have hitherto considered that local budgets are given and that preferences affect only the distribution of resources within jurisdictions. Even in the prior section, in which local health care budgets are allowed to vary in response to preferences, the whole amount of resources available in each jurisdiction to allocate between health care and ‘entertainment events’ is fixed and independent of preferences. That is, up to this point of the discussion we considered that (different) local preferences affect resource allocation only after local budgets have been set by the central government. In this section, we thus discuss the impact of using different local preferences to set local health care budgets.

The focus here is on a particular type of preferences: preferences over

health states, already presented in chapter two. Therefore, a distinction is made between quantity and quality effects of health care interventions. We assume that there is diversity across jurisdictions regarding preferences that affect health outcome measurement. There are then two alternatives to measure health gain, when setting local budgets: to use the preferences of the whole population (represented in this case by average preferences) or local preferences.

The amount of health gain obtained by  $j$ 's representative resident, when delivered  $m_{jt}$  resources,  $e(m_{jt})$ , might be decomposed into a quantity effect and a quality effect:  $e(m_{jt}) = L \cdot Q(m_{jt})$ .  $Q$  is the number of time units (e.g. days, months, years) obtained by  $j$ 's representative resident with  $m_{jt}$ , and  $L$  is the change in the weight attached to the health state achieved with  $m_{jt}$  compared to the health state that would be achieved without  $m_{jt}$ .



**Figure 14. The impact of different preferences about health states on health gain measurement**

Figure 14 shows the impact of different preferences over health states on health gain measurement. The horizontal axis measures length of life and the vertical axis measures quality of life. Total health gain is therefore given, in Figure 14, by an area whose width corresponds to changes in the length of life and whose height corresponds to changes in quality of life. Suppose then that a given intervention affects both length and quality of life, being the quantity effect equal to  $Q^*$ . If this quantity effect is weighted by the quality of life score  $QoL_1$ , then the total health gain obtained with this intervention is given by the area A.

Differently, if the quantity effect is weighted by the quality of life score  $QoL_2$ , then the total health gain obtained with this intervention is given by the area  $A+B$  (the health gain obtained with the intervention is thus greater for the case of  $QoL_2$  than  $QoL_1$ , that is, depending on preferences for health states, the health gain obtained with a given intervention will be greater or lower). The dashed line in Figure 14 indicates in the vertical axis the average value of  $QoL$ , given  $QoL_1$  and  $QoL_2$  (using the average value, the total health gain is given by the area  $A + \frac{B}{2}$ ).

Assume for the moment that there is only one health care intervention,  $t$  and that the objective of resource allocation is to maximise total health. The decision at stake here is the allocation of resources between jurisdictions  $A$  and  $B$  and the central decision-maker might use average or, instead, local quality of life scores, when distributing resource between jurisdictions.

To simplify the exposition, assume that, for intervention  $t$ , the quantity effect is the same for both jurisdictions:  $Q(m_{At}) = Q(m_{Bt})$ , for  $m_{At} = m_{Bt}$ . Let us further assume that costs are also the same across jurisdictions:  $C(m_{At}) = C(m_{Bt})$ , for  $m_{At} = m_{Bt}$ . To analyse the case of heterogeneous preferences used to derive quality of life scores, assume that  $L_A \neq L_B$ , for  $m_{At} = m_{Bt}$ . Assume, for instance, that  $L_A < L_B$ , for  $m_{At} = m_{Bt}$ , meaning that  $A$ 's residents attach lower value to what intervention  $t$  can do to their health status, compared to  $B$ 's residents.  $\bar{L}$ , representing average preferences, lies between  $L_A$  and  $L_B$ .

The central decision-maker has to allocate a fixed budget,  $M$ , between  $A$  and  $B$ . Welfare is maximised when health gain is maximised. Given the assumption of diminishing marginal health productivity (as  $m_{jt}$  increases the additional quantity effect becomes smaller) plus the assumptions that  $Q(m_{At}) = Q(m_{Bt})$  and  $C(m_{At}) = C(m_{Bt})$ , for  $m_{At} = m_{Bt}$ , if the central decision-maker uses average preferences to measure health gain, then  $e(m_{At}) = \bar{L} \cdot Q(m_{At}) = e(m_{Bt}) = \bar{L} \cdot Q(m_{Bt})$ , for  $m_{At} = m_{Bt}$ . So, welfare is maximised when each jurisdiction receives one half of the total budget. Let us identify these amounts with the superscript  $\bar{L}$ , to denote that they derive from average preferences:  $m_{At}^{\bar{L}} = m_{Bt}^{\bar{L}} = \frac{1}{2} M$ .

What happens if the central decision-maker uses  $L_A$  and  $L_B$  to allocate  $M$  between  $A$  and  $B$ ? In this case, because  $e(m_{At}) = L_A \cdot Q(m_{At}) < e(m_{Bt}) = L_B \cdot Q(m_{Bt})$ , for  $m_{At} = m_{Bt}$ , health maximisation would lead to a local budget lower in jurisdiction  $A$  than in  $B$ :  $M_A < \frac{1}{2}M$  and  $M_B > \frac{1}{2}M$ . This means that the amount of resources allocated to jurisdiction  $A$ , based on local preferences,  $m_{At}^{Local-L}$ , is lower than the amount of resources allocated to jurisdiction  $B$ ,  $m_{Bt}^{Local-L}$ , in which the latter is based on local preferences too.

Compared to a resource allocation, between jurisdictions, based on average preferences, an allocation based on local preferences makes  $B$ 's residents better-off but  $A$ 's residents become worse-off. Although the quantity effect is the same in both jurisdictions, if the distribution of resources between jurisdictions is based on (different) local preferences, one jurisdiction gets more resources than the other. The use of local preferences to allocate a common pool of resources might therefore pose some problems.

Let us now consider that there are two health care interventions,  $t=1,2$ . This means that there are two types of decisions: the allocation of resources between jurisdictions, involving different preferences for quality of life; and the allocation of resources between interventions, within each jurisdiction, involving common preferences for quality of life scores. The objective now is to discuss the implication of using average or local preferences, within jurisdictions and after local budgets have been set.

Continue to assume that  $Q(m_{At}) = Q(m_{Bt})$  and  $C(m_{At}) = C(m_{Bt})$ , for  $m_{At} = m_{Bt}$  and  $t=1,2$ . Let us further assume that quantity effects are also the same across interventions and that quality effects, based on average preferences, are also equal across interventions:  $\bar{L}_1 = \bar{L}_2$  for  $m_1 = m_2$ . Differences apply to local preferences for intervention 1 *versus* intervention 2. Suppose that  $L_{A1} < L_{B1}$ , for  $m_{A1} = m_{B1}$ , while  $L_{A2} > L_{B2}$  for  $m_{A2} = m_{B2}$ . That is, the change in quality of life produced by intervention 1 is valued more in  $B$  than in  $A$ , while the opposite is verified for the case of intervention 2.

Consider that local budgets have been determined based on average preferences  $\bar{L}_1$  and  $\bar{L}_2$  (for interventions 1 and 2, respectively). Given our



assumptions about quantity effects and costs, maximisation of social welfare (i.e. maximisation of health), implies that each jurisdiction gets half the total budget:

$$M_A = M_B = \frac{1}{2}M.$$

Once local budgets have been set, if they are allocated between interventions 1 and 2 based on average preferences,  $\bar{L}_1$  and  $\bar{L}_2$ , each intervention is allocated  $\frac{1}{4}M$  resources, in each jurisdiction (recall that we are assuming that quantity effects are also the same across interventions). Let us represent the amount of resources allocated to intervention  $t$ , in jurisdiction  $j$ , when it is based on average preferences, by  $m_{jt}^{\bar{L}_t}$ .

Now, because residents in  $A$  value more the change in quality of life produced by intervention 2, compared to intervention 1, if the allocation of  $M_A = \frac{1}{2}M$  is based on local preferences,  $L_{A1}$  and  $L_{A2}$ , more resources are allocated to intervention 2 than to intervention 1. Let us represent the allocation based on  $A$ 's local preferences by  $m_{At}^{Local-L}$ , for  $t=1,2$ , where,  $m_{A1}^{Local-L} < \frac{1}{4}M$  and  $m_{A2}^{Local-L} > \frac{1}{4}M$ . In turn, in jurisdiction  $B$ , if the allocation of  $M_B = \frac{1}{2}M$  is based on local preferences, it is intervention 1 that is allocated a greater amount of resources than intervention 2:  $m_{B1}^{Local-L} > \frac{1}{4}M$  and  $m_{B2}^{Local-L} < \frac{1}{4}M$ .

Given local budgets, moving from an allocation between interventions based on average preferences to an allocation based on local preferences, increases social welfare in both jurisdictions:  $\sum_{t=1,2} L_{At} \cdot Q(m_{At}^{Local-L}) > \sum_{t=1,2} L_{At} \cdot Q(m_{At}^{\bar{L}_t})$

$$\text{and } \sum_{t=1,2} L_{Bt} \cdot Q(m_{Bt}^{Local-L}) > \sum_{t=1,2} L_{Bt} \cdot Q(m_{Bt}^{\bar{L}_t}).$$

While distributing resources between jurisdictions based on local preferences raises some problems, once local budgets have been set, it is optimal from the perspective of local welfare to allocate local budgets based on local preferences (obviously, this result is valid assuming, as we are, that local welfare is maximised when local health is also maximised; if people value the use of equal scores across jurisdictions, the results might change as in section 3.4).

### 3.8. Discussion

Preferences have played an important role in the literature about decentralisation concerning its impact on allocative efficiency but only indirectly. In this chapter we sought to show that preferences themselves might influence the impact of decentralisation on social welfare. Without being exhaustive, we focused on particular situations in which preferences contributed to the evaluation of the optimality of decentralised allocation of resources compared to the *status quo* of centralised allocations. So, the answer to the question posed in the beginning of the chapter is ‘yes’.

The traditional perspective of the literature about decentralisation has adopted a welfarist approach, interpreting the concept of preferences as consumers’ preferences. Thus, to reach our main result (preferences themselves might affect the optimality or otherwise of decentralisation), it was crucial to bridge the literature on decentralisation with the literature about health care priority setting. Within the latter literature it has been acknowledged that the concept of preferences might be interpreted in different ways, including non-welfarist interpretations according to which preferences might reflect individuals’ concerns for others.

Many of the preferences discussed throughout this chapter can in fact be thought as society regarding preferences but in our assumptions (section 3.2) we have not imposed any constraint in this regard and some of the situations considered are actually closer to the concept of personal rather than social preferences (for example, preferences for non-health-related ‘enjoyment’ considered in section 3.6 and preferences for health states considered in section 3.7).

Some issues included in the current chapter have been addressed in the literature about decentralisation but no link has been established between them and preferences (which is understandable given the focus on consumers’ preferences). For instance, the issue of equal provision of goods and services across jurisdictions (discussed in section 3.4) has been addressed in the literature about decentralisation but it has been presented as a political/constitutional constraint that prevent central governments from introducing local differentiation

in service provision (e.g. Oates, 1999). Also, the sort of externality effects, suggested and discussed in section 3.6, is actually recognised, though only implicitly, in Petretto (2000)<sup>26</sup>. This author argues that, in the case of the social evaluation of policies, spill over effects of public expenditure should be corrected by the considerations of redistributive aims. The concept of spill over effects is used by Petretto (2000) in the welfarist sense referring to the willingness to pay by non-residents benefiting from health services provided in jurisdictions other than that where they live. Regarding the impact of redistributive aims on the social value of spill over effects, the author does not associate it with preferences however, if he recognises that the public expenditure in one region might be socially valued because it affects other ‘deserving’ regions, then this is the recognition (though the author himself does not follow this route) of the existence of spillover effects associated with distributional considerations.

In the literature about decentralisation, equal provision across jurisdictions and increased allocative efficiency by better matching service provision to heterogeneous local preferences have been addressed within separate frameworks (the former aspect as been associated with a non-welfarist perspective, in which equal provision is seen as a constraint deriving from e.g. the Constitution; the latter has been addressed within a welfarist perspective, in which the levels of utility enjoyed by individuals are used to evaluate alternative social states). In this chapter, the adoption of a broad interpretation of preferences allowed us to reconcile these dichotomous objectives (equal provision *versus* variation in provision) within a single framework. In fact, both objectives might be linked to preferences and then the optimal allocation is dictated by the trade-offs between them that people are willing to make.

Within our framework, decentralisation of resource allocation itself generates some trade-offs between objectives. Consequently, whether or not decentralisation is optimal, when compared to centralisation, depends on the capacity of local decision-makers to generate trade-offs closer to what the public prefers. The approach suggested in Figure 2 helps to understand the association between decentralisation and different objectives of resource allocation. There the focus is on the central and local decision-makers seen as alternative *agents* acting

---

<sup>26</sup> See review in chapter two.

on behalf of local populations. Depending on the goals of resource allocation envisaged by local populations and given the different capabilities possessed by *agents*, local populations will be better-off with one or the other *agent*. Decentralisation is said to improve social welfare, compared to centralisation, if local populations are better-off when local decision-makers assume the role of their *agents*.

Naturally, our conclusions are valid under the assumptions described in section 3.2. A major assumption is that decision-makers are benevolent, which has allowed us to focus the discussion on the relative capacities of *agents* to reach resource allocation objectives, ignoring the possibility of the introduction of an additional tier of government bringing into the analysis additional self-regarding objectives of decision-makers (Smith *et al.*, 1997). In Bossert's (1998) 'decision space'<sup>27</sup>, for instance, the objectives of resource allocation are set by the central government, the *principal*. His approach is therefore closer to the relationship described in the upper-half of the diagram presented in Figure 2. One must say nonetheless that, while his view takes into account that local authorities, the *agents*, may pursue self-regarding interests, it does not take into account that *agents* (local decision-makers) may possess superior knowledge about preferences than the (intermediate) *principal* (the central government).

Another important assumption concerns the existence of information asymmetry between the central and local decision-makers regarding two types of information: technical information and information about preferences. The assumption that each decision-maker has an advantage over the other regarding one of the two types of information is fundamental to explain the trade-off created by decentralisation between production efficiency and the match between service provision and preferences. The direction of the information asymmetry, regarding each type of information, is not however determinant to conclude that preferences might affect the impact of decentralisation on allocative efficiency. That is, we could reach the same conclusion assuming that the central decision-maker is better informed about preferences than local authorities while assuming the opposite for the case of technical information.

Variation in preferences across jurisdictions is an important assumption

---

<sup>27</sup> See chapter two.

too in the current chapter: even with wide variation in preferences, there will be no systematic way to address these differences in preferences by varying local health services provided if people are located randomly across jurisdictions (only in the case of preferences for participation in decision-making, the impact of preferences on the optimal level of decentralisation does not depend on the existence of variation across jurisdictions). Although variation in preferences is a common assumption in discussions about decentralisation, in chapter two it was said that there is a lack of empirical evidence supporting this assumption. Oates (1999, p.1122) himself says, in respect to consumers' preferences, that the 'spatial pattern of consumption of certain goods and services like education and health is open to some debate'. The main objective of the empirical work reported in chapter four is precisely to test the assumption of geographic diversity in preferences. For the particular samples and preferences used in this work, the hypothesis of variation is actually not corroborated.

The discussion in section 3.5 is valid under the assumption that participation increases as decentralisation increases but empirical evidence in this respect is scant too. Additionally, in Inman and Rubinfeld's (1997b) analysis (on which Figures 10 to 12 are based), decentralisation increases with the number of local governments, thus, the link between decentralisation and participation is clearer in their discussion than in ours because more local governments mean smaller groups of people governed by a single entity and hence greater possibility of participation.

The discussion developed in section 3.6 is based on the assumption that the public is sensitive to inequality in health across jurisdictions nonetheless the empirical evidence about this issue is also scarce. Still, *a priori* it seems that the geographical dimension of inequality is relevant for policy making. In England, for example, in February 2001, one of the two health inequality targets announced by the Secretary of State for Health at the time, Alan Milburn, was "starting with Health Authorities, by 2010 to reduce by at least 10% the gap between the quintile of areas with the lowest life expectancy at birth and the population as a whole" (Department of Health, 2001a). The results of the empirical work presented in chapter four show that people do in fact value the geographical dimension of (in)equality in health and health care. The majority is however sensitive to opportunity costs of equality in terms of health gain foregone, meaning that they

are willing to give up of health gain to keep equality but only up to a certain level.

In section 3.6, we did not however consider movements of people between jurisdictions but these movements might themselves alter health distribution between areas. In this case, if resource allocation between areas is not a feasible instrument to impact upon health distribution, then the advantage of centralised allocation in terms of internalising the spill over effects disappears and the discussion falls in the scenario adopted in section 3.3.

Here, as in the mainstream economic theory, we have assumed that preferences are exogenous and stable. This assumption has nonetheless been challenged (e.g. Sunstein, 1993; Bowles, 1998; Palacios-Huerta and Santos, 2004). Sunstein (1993, p. 221), for instance, says for the case of environmental law that ‘policymakers cannot simply identify preferences and try to satisfy them, since preferences are influenced by law, and since there is no way to identify the preferences that would exist in the absence of law’.

In the current chapter, the consequence of endogenous preferences is that the optimal setting (centralised or decentralised decision making) cannot be identified based on preferences given that the latter are in turn affected by the chosen setting. So, for example, suppose that, given preferences for participation, we conclude that it is not optimal to decentralise resource allocation decisions. However, people might attach little value to participation because they have not been given the opportunity to participate. It has in fact been said that ‘when people are induced to engage in an activity with little or no extrinsic reward, they come to value the activity more highly, that is, they come to believe that their actions were intrinsically motivated’ (Bowles, 1998, pp. 90-91). Participation might even be viewed as a way of enhancing the credibility of, and getting support to, decisions that would otherwise be implemented (even more so as expectations and demand for health services have steadily increased). So, in this case participation might be deliberately used by decision-makers to lower expectations and possibly affect preferences regarding health care priority setting (Bagott (2005), for example, says that Patient and Public Involvement Forums in England - see chapter five – could be seen as a way of controlling the demand for services). Nonetheless, in a scale measuring the degree of participation in decision-making, this sort of participation is rated low and it has actually been termed as ‘manipulation’ (Litva *et al.*, 2002). Another example is the introduction

of guidelines, as in Figure 6, which might itself shape preferences regarding the allocation of resources to specific groups of people or to particular health care services.

Sunstein (1993) says that it has been shown that people place a higher value on rights or goods that they currently hold than they place on the same goods in the hands of others. The author provides some explanations namely the existence of a *status quo* bias (losses are weighed more heavily than equivalent gains) and adaptive preferences (people adapt their conduct or even their desires to what has been available). Sunstein (1993) distinguishes however private preferences from what he calls ‘collective judgements’, saying that the choices that people make as political participants are different from those they make as consumers. Moreover, he says that the collective character of politics might overcome the problem of adaptive preferences, further suggesting that private preferences are an inadequate basis for (environmental) policy.

This suggestion raises some normative questions to future research on the nature of preferences used to inform policy making. Sunstein (1993) actually refers the possibility of shifting the analysis to normative issues, that is, he says that given the possibility of endogenous preferences, it may be important, when setting policy, to make some choice about the sorts of preferences that ought to be encouraged, rather than to act as if preferences can be kept constant.

An important question that remains to future research concerns the rule of aggregation of the desideratum at the suprajurisdictional level, that is, how to evaluate changes in overall social welfare given different preferences across jurisdictions (in Figure 5, for instance, decentralisation might represent an improvement relative to centralisation, or not, depending on which indifference curves, *A*’s or *B*’s curves, are used to evaluate welfare changes). This issue has been avoided either by not allowing the possibility of inconsistent objectives across jurisdictions (e.g. Gordon, 1983) or by assuming that there are two alternative, not necessarily convergent, perspectives (regional and national) to evaluate welfare changes (e.g. Petretto, 2000) but which alternative should be used to evaluate the impact of decentralisation on allocative efficiency is still an open question.

In terms of empirical work, there are many topics that have remained under-explored. The question of variation in preferences and distributional

concerns involving geographic areas are addressed in the next chapter but other questions remain to future research, namely the operationalisation and elicitation of several trade-offs considered in the current dissertation (e.g trade-offs involving outcomes and procedures).



## **CHAPTER 4. Do preferences vary across regions? An empirical study applied to two Portuguese municipalities**

### **4.1. Introduction**

Following the review of literature in chapter two, we stressed that the assumption of variation in preferences across jurisdictions is crucial to the positive relationship established, in the theoretical literature of economics and health economics, between decentralisation and allocative efficiency. This same assumption is present (with few exceptions) throughout the theoretical analysis that we have developed in the previous chapter. However, despite the relevance of this assumption in economic analyses in contexts of decentralisation, in chapter two we also said that empirical evidence in this area is to the extent of our knowledge very scarce. This chapter reports on a small-scale empirical study that sought to approach this issue of geographic variation in preferences. Thus, the main objective of this chapter is to assess whether or not there are differences in preferences (some of which are addressed in the theoretical chapter), in this case, between two independent samples drawn from two distinct Portuguese municipalities.

To pursue this objective we developed and administered the same questionnaire to both samples. The selection of regions from which to draw the samples was partly by practical convenience and partly judgemental (as explained in section 4.3.2). Given the framing of questions, and making use of the taxonomy proposed in the literature (reviewed in chapter two), one might say that two types of preferences are elicited: personal preferences that ask people about what they prefer for themselves, on whatever basis, and social preferences that ask people about what ought to be the distribution of resources affecting other people.

Social preferences considered in the current study can be related to maximising as well as egalitarian principles of social justice (see chapter two, section 2.3). Trade-offs between these two principles are elicited using person trade-off questions. Several trade-offs involving health maximisation and inequality of different kinds have been elicited, however, based on the results of

the literature search reported in chapter two, the trade-off between health maximisation and equality of treatment (equality of health gains) between regions, which is addressed in the current study, seems to have received little attention in the empirical (health economics) literature. Preferences regarding the geographical dimension of (in)equality are of particular importance in contexts of decentralisation given that the latter have been associated with variations in service provision.

The chapter is organised as follows: in section 4.2, we introduce in brief the Portuguese context, in which the empirical study was carried out; section 4.3 describes the methods, namely the design of the questionnaire, the selection of respondents and the statistical methods used to analyse the data; section 4.4 presents the main results that emerged from the data; and section 4.5 closes this chapter with a discussion concerning the results and limitations of the current study.

## **4.2. The Portuguese context**

In this section, we outline some aspects that characterise the health care sector in Portugal. We start by presenting some demographic and health indicators of the country (sub-section 4.2.1). The focus then shifts to the health care system (sub-section 4.2.2), covering a bit of history that explains to some extent the current state of affairs as well as organisational features, aspects related to finance and delivery, and, finally the issue of decentralisation. This section is mainly based on Bentes *et al.* (2004) and some OECD health data (2006) are also used.

### **4.2.1. Some demographic and health indicators**

Portugal is part of the Iberian Peninsula in the southwest of Europe. Two archipelagos (Azores and Madeira) in the Atlantic Ocean are also part of Portugal. According to the 2001 census, the total resident population of Portugal was 10.355.824, being 15,5% of the population 65 years or older and about 65,6% of the

population resided in urban areas. The two main metropolitan areas are greater Lisbon and greater Porto.

In 2000, GDP per capita was €11.288, a figure that rose from €5.135 just in 1990. Despite the economic growth, Portugal has one of the highest levels of income inequality in Europe. Portugal has developed an increasingly service-based economy. The three main sectors of the labour force in 1999 were services (60%), industry (30%) and agriculture (10%). The main industries were textiles and footwear, wood pulp, paper and cork, metalworking, oil refining, chemicals, fish canning, wine, and tourism.

Total health care expenditure in Portugal has risen steadily from 3% in 1970 to 10,1% of GDP, in 2004 (OECD health data). In 2004 the average life expectancy at birth in Portugal was 77,4 years (80,5 years for women and 74,2 for men). The infant mortality rate decreased fivefold between 1970 and 1990. In 2005 this rate was 4,0 per 1000, being now lower than the European Union average.

Figures for 2004 show that diseases of the circulatory system accounted for 39% of all deaths and cancers for 20% of all deaths. Cerebrovascular diseases account for almost half of the deaths associated with diseases of the circulatory system. Gastro-intestinal tumours are the most frequent type of cancer in both men and women. Approximately a quarter of premature mortality in men comes from traffic accidents, while among women the highest proportion of potential years of life lost is caused by cancer (25.9%).

#### 4.2.2. The health care system

One might say that Portugal has a complex health care system, greatly explained by historical factors. Up until 1979 and the establishment of the National Health Service, the Portuguese state had traditionally left the responsibility for paying for health care to the individual patient and his or her family. Care of the poor was the responsibility of charity hospitals and out-of-hospital care remained the responsibility of the Department of Social Welfare. Civil servants were the exception as well as maternal and child health care, and some interventions in the control of infectious diseases and mental health.

After the revolution of 1974, a process of health services “nationalisation” began which culminated into the establishment of the NHS in 1979. First, in 1974

district and central hospitals owned by the religious charities were taken over by the government. Local hospitals followed in 1975 and were integrated with existing health services. Finally in 1977, over 2000 medical units or health posts situated throughout the country were taken over by the government.

The principle of the citizen's right to health was embodied in the Portuguese constitution of 1976 and was to be delivered through "a universal, comprehensive and free of charge National Health Service" but only in 1979 the law establishing the NHS was enacted. Despite the development of a unified publicly financed and provided health care system and the incorporation of most of the health facilities previously operated by the social welfare and religious charities, some aspects of the pre- 1970s system persisted. In particular, the health subsystems (*Subsistemas*) continued to cover a variety of public and private employees. These schemes offered better services and greater choice of provider than would be available under the NHS. In addition to the coverage provided by the NHS, about 25% of the population is covered by the health subsystems. Along with the subsystems, a strong private sector remains an important part of the Portuguese health care system. In many cases, the medical professionals simultaneously work in the public and private sectors. This situation too has its roots in historical events, namely in the resistance offered by the medical profession to the implementation of a National Health Service.

At the top of the hierarchy of the public health care sector is the Ministry of Health, with the Minister of Health himself embracing much of the core function of the Ministry (regulation, planning and management of the NHS). The Ministry of Health is made up of four directorates and seven institutes and there are also four vertical programmes run by national bodies attached to the Ministry of Health (there are however some reforms currently under way).

The 1990 Law on the Fundamental Principles of Health (Law n°47/90 - 24<sup>th</sup> August) introduced new principles for the organization and functioning of the health system, namely the decentralisation of management at the regional level. As a consequence, in 1993, five Regional Health Administrations were established: North; Centre; Lisbon and Tagus Valley; Alentejo; and Algarve (see Appendix A). In each region a health administration board, accountable to the Minister of Health, manages the NHS. Their management responsibilities are a mix of strategic management of population health, supervision and control of hospitals

and centralized direct management responsibilities for primary care/NHS health centres. Each region is further divided into eighteen sub-regions (corresponding to the continental eighteen districts). Here too some organisational changes are to be introduced.

Concerning public hospitals, legal and formal changes have occurred, with the introduction of an entrepreneurial management style, affecting the composition of boards and with most members appointed by the Minister of Health. Changes in management have also led to greater flexibility and accountability in resource use as well as the progressive change of the NHS personnel status from salaried civil servants to private employees working under negotiable contracts. We must note that there has been a series of experiences regarding the management of hospitals but fewer have been the lessons and consequences assumed, that is, experiences have neither been expanded nor abandoned.

Primary health care centres are directly under the managerial control of the Regional Health Authorities through sub-regional coordinators. In their case too, there have been some experiences concerning their management and one of the most recent reforms has created the figure of executive director that for the first time does not have to be a doctor. The creation of Family Health Units is another aspect of the recent reforms (they are multidisciplinary teams involving doctors, nurses and administrative staff, being responsible for the provision of primary care and some specialised care to patients registered in their respective lists).

Some of the health subsystems, for which membership is based on occupational category, are run by trade unions and managed by boards of elected members. The largest fund (ADSE- for civil servants) is controlled by the Ministry of Finance. It covers 15% of the population, corresponding to 60% of all subsystem members and includes amongst its members all employees of the NHS.

The private sector has performed a supplementary role to the NHS and mainly includes private practitioners, *Misericórdias* and private hospitals and clinics. Overall the private sector accounts for 32% of all medical consultations, being the majority specialist consultations (Bentes *et al.*, 2004).

The health care system benefits from both public and private sources of financing. The National Health Service is funded out of general taxation, though there are some flat rate charges namely for consultations (primary care and hospital outpatient visits), emergency visits, home visits and diagnostic tests.

Recently, flat rates were introduced for the case of in-patient care and elective surgery. Certain groups (e.g. pregnant women, children, low income patients) are exempted from these charges. Hospitals have other sources of revenue such as donations and payments from renting premises and equipment. In total, sources other than general taxation account for about 8% of the NHS revenue (Bentes *et al.*, 2004).

The health subsystems are financed through employer and employee contributions, with the largest part paid by the employer. However, in the case of civil servants, the employer is the State, so this too should be considered public financing.

An important source of financing of health care is out-of-pocket payments. This category of financing has consistently accounted for over 20% of total health expenditure over the last ten years, being among the highest in Europe. The majority of payments are for drugs and therapeutic products. Voluntary health insurance is marginal.

Delivery of health care is also based on a mix of public and private providers. Primary health care in the public sector is mostly delivered through publicly funded and managed health centres, by GPs/family doctors and primary care nurses. Some health centres also provide a limited range of specialized care (for example, paediatrics, gynaecology and obstetrics). Patients must register with a GP, and can choose among the available clinicians within a geographical area. In theory there is no direct access to secondary care and GPs are supposed to act as gatekeepers. However, people tend to by pass the referral process by going directly to emergency care services. Alternatively, they go to the private sector (particularly for the case of specialist visits in which there are longer delays in public hospitals).

Portugal also has a large independent private sector which provides diagnostic and therapeutic services to NHS beneficiaries under contracts called *convenções*. Patients can choose from any of the providers who appear on the contracts.

There are very few NHS dentists; as a consequence of this shortage in public supply, people use the private sector for oral care. In the case of members of health subsystems, they have easier access to this form of care, either by paying directly and being partially reimbursed afterwards or by paying a lower fee in advance by visiting a dentist with a contract with the subsystems.

Secondary and tertiary care is mainly provided in hospitals. In 1999, Portugal had 205 hospitals, 110 public and 84 private. Most hospital services are directly provided by the NHS, though some non-clinical services have been outsourced. There are four categories of hospitals: central hospitals (provide highly specialized services with advanced technology and specialist human resources); specialized hospitals (provide a broad range of specialized services); district hospitals (located in the main administrative district and provide a range of specialist services); and district level-one hospitals (only provide internal medicine, surgery and one or two other basic specialties).

Health resources are concentrated in the capital, Lisbon, and along the coast, between Lisbon and Porto. The regions of Algarve and Alentejo account only for about 10% of public hospitals, though they are also less populated areas. The number of NHS beds per 1000 people ranges from 1,87 in Algarve to 2,68 in the Centre.

The NHS budget is set annually by the Ministry of Finance, based on historical spending and the plans put forward by the Ministry of Health. Capital and current expenditure are separated; being the former under the direct control of the Ministry of Health. Concerning current expenditure, the Ministry of Health allocates a budget to each Regional Health Administration for the provision of health care to a geographically defined population. Their autonomy has nonetheless been limited to primary care. Regional budgets for primary care are set according to historical expense and, since 1998, capitation, adjusted by age, sex and a disease burden index (computed according to the regional prevalence of selected health indicators).

Despite that decentralisation is formally addressed within the Law of Fundamental Principles of Health (Law n<sup>o</sup>47/90 - 24<sup>th</sup> August), where it is stated that the NHS is managed at the regional level, Bentes *et al.* (2004) note that, in practice, responsibility for planning and resource allocation has remained highly centralised. Campos (2004) argues in the same line, saying that the public sector, including health services, remains highly centralised. In fact and as previously mentioned, Regional Health Administrations' autonomy has been limited to primary care, being hospital budgets still defined and allocated at the central level.

In the case of hospitals, one might say that some degree of operational decentralisation has been achieved via the introduction of some of the principles of the New Public Management approach in the NHS.

A further measure potentially leading to decentralisation and public involvement was adopted, in 1997, with the creation of regional contracting agencies (one in each of the five Regional Health Administrations). These agencies were expected to increase cost awareness and to provide incentives for efficiency. The impact of contracting agencies was however marginal and as a result of governmental changes, in 1999 and 2001, agencies were struck by a lack of strategic guidelines and uncertainty about their role, coming to an effective end. The concept has been reintroduced in the political agenda but no practical changes have occurred so far. Another aim envisaged by the establishment of the contracting agencies was the promotion of public involvement in health care decision making. However, at the moment, talking about an increase in public participation in decision making basically means that formal mechanisms for ‘consumers’ to make complaints have now been arranged in a generalised and systematic way.

### **4.3. Methods**

In this section, we describe the methods used to collect and analyse the data, including: (i) the design and administration of the questionnaire as well as the designation and codification of the variables derived from the questionnaire (sub-section 4.3.1); (ii) the selection of samples (sub-section 4.3.2); and (iii) the identification of statistical methods used to analyse the data collected (sub-section 4.3.3).

#### **4.3.1. The questionnaire**

Being the main objective of the empirical work to test for diversity in preferences, we had to select some preferences to start with. We decided to elicit people’s preferences regarding the criteria of health maximisation and equality of treatment (across geographic regions) as well as the trade-off between them. The main reasons to choose in this way were two-fold: firstly, we saw in chapter two that the geographical dimension of (in)equality has not been addressed in empirical works that have elicited preferences in the context of health care



resource allocation, thus, this seemed to be an opportunity to contribute to current knowledge by looking at people's views about the issue, namely, whether or not (in)equality across regions matters and whether or not people are willing to make trade-offs; and, secondly, this sort of trade-offs is among the ones that were explored in the previous chapter (section 3.4), thus, one could use the empirical findings to test the pertinence of the theoretical discussion.

In addition, we hypothesised that there could be different results concerning geographic diversity in preferences depending on the nature (personal or social) of preferences. A secondary objective of the questionnaire was therefore to elicit personal preferences, choosing for this effect preferences over attributes of health care delivery. In one case, the doctor's gender is the attribute at stake (which might be more relevant, for example, in conservative communities than in liberal communities). In the other case, as explained below, the aim is to identify thresholds below which people do not seek care (or, a particular form of care). We found this idea of thresholds in Rice and Smith (2001). These authors suggest that there may be important areawide cultural influences on the use of health services that might, for example, influence a threshold of ill health below which individuals choose not to seek medical intervention.

In designing the questionnaire, we have followed common wisdom (Hill and Hill, 2002; Saunders *et al.*, 2003). Examples of aspects that should be taken into account are the wording, lay out and length of questions, the size of letters, coding issues and the use of an odd or even number of points in rating scales. In some cases, one has to weigh advantages and disadvantages of each solution. For instance, in our questionnaire and particularly in questions 3 and 4, the reader has to go through the description of the hypothetical situations before answering the questions. Longer questions might disincentive the completion of questionnaires; however, we thought that it was important to make the respondent aware of what was at stake, trying in addition to control for some of the factors that could affect responses. The two stages and the graphic presentations of questions 3 and 4 were inspired by Williams *et al.* (2005). We also carried out a small pilot test (with a Portuguese version of the questionnaire), being some of the suggestions of participants included in the final version of the questionnaire.

Considering that the questionnaire was aimed at Portuguese people, we have given extra attention to the Portuguese version (Appendix C) in terms of the

specific words put into the questionnaire (namely, we have avoided the use of economic ‘jargons’ like health maximisation). As suggested by participants in the pilot study, we omitted the codes attributed to the different responses in order to avoid any confusion with numbers (these codes are shown only in the English version - Appendix B).

The questionnaire comprises 6 questions, which can be classified in two groups. The first group (questions 1 to 4) is intended to elicit preferences from a societal perspective, in which individuals are asked about the allocation of resources affecting other people<sup>28</sup>. Moreover, the criteria in question, health maximisation and (geographic) equality of treatment, have echoes in social justice. Questions 1 and 2 simply ask respondents about their support for the two previous criteria; while questions 3 and 4 evaluate the extent to which people are willing to trade-off one principle against the other. Both questions 3 and 4 apply to hypothetical situations but in the latter the hypothetical situation involves actual Portuguese regions.

The second group of questions (formed by questions 5 and 6) is intended to elicit personal preferences, asking individuals about what they prefer for themselves, on whatever basis.

The last part of the questionnaire asks respondents about some personal data, for statistical purposes only (to assess the representativeness of samples). Given this overview of the questionnaire, we will now go through each question.

Question 1 asks respondents about their support for the inclusion of expected outcome (health gain) in the set of criteria that might be used to guide resource allocation. A brief paragraph explains that not all health care interventions have the same impact on individuals’ health and that, depending on the situation, what might be at stake is a health improvement or a lower deterioration in health. There is not an explicit reference to health maximisation as this seems to be a more complex concept. A Likert scale was used to measure the level of respondents’ agreement or disagreement to the above criterion<sup>29</sup>. Answers

---

<sup>28</sup> Actually, questions 1 and 2 can be thought as an example of *ex-ante* ‘socially inclusive personal preferences’ (Dolan *et al.*, 2003), though respondents are not given any information about the probability of them or others becoming ill (see chapter two).

<sup>29</sup> Rating or scale questions are often used to collect opinion data. The most common approach is the Likert-style rating scale in which respondents are asked how strongly they agree or disagree with a statement or series of statements, usually on a four-, five-, six- or seven-point rating scale (Saunders *et al.*, 2003, p.296).

could range from 1 (strongly disagree) to 5 (strongly agree), in which 3 represents a neutral position.

It is important to note that we mention in the question that health gain is a potential criterion, among others. Consequently, respondents see themselves in a comfortable standpoint given that they are not forced to choose a single criterion over potential contenders nor are they asked to establish any ranking. This might result in high support to health maximisation. Nonetheless, presenting health maximisation as the only criterion would not serve our objectives (given that the aim was to ask respondents to make trade-offs between two potentially relevant criteria). Moreover, difficult decisions in real world situations are not so much about which principles matter as they are about how much weight should be given to each one; it is therefore reasonable to expect high levels of support to each principle, when considered in isolation.

In question 2, we start by briefly explaining that resource allocation involves a geographical dimension as well and asks respondents whether (and to what extent) or not they agree that geographic distribution of resources should be taken into account in health care resource allocation. We use a Likert scale as in question 1. We do not explicitly mention geographic equality of treatment and again this is presented as one among other possible criteria.

Despite the fact that the framing of questions 1 and 2 is likely to lead to very similar responses across respondents, we think that these two questions needed to be set prior to questions 3 and 4. Firstly, they serve to introduce and explain to respondents the issues of capacity to benefit from health care and geographic distribution of resources. And, secondly, when respondents are asked to make trade-offs between two principles, in question 3 and 4, they were already given the opportunity to think about the relevance of each one.

Two variables, HMAX and HGEO, are defined to represent responses to questions 1 and 2, respectively (see Table 1 below). These variables directly follow from the questionnaire and can therefore take the values of 1 to 5.

In question 3, respondents are faced with a hypothetical situation, in which there is a limited quantity of a given vaccine. This vaccine protects children against a disease that is not contagious but is lethal. Thus, each episode of illness avoided is a children's life saved.

Respondents are asked to imagine themselves in a position in which they

must choose whom to immunise. The situation described involves a physical constraint that cannot be overcome by diverting more financial resources to health care. This means that respondents cannot consciously or unconsciously adopt a defensive attitude, claiming or thinking that rationing is not inevitable<sup>30</sup>.

In order to deal with the issue of geographic distribution of resources, we hypothesise a country divided in two regions: Region 1 and Region 2<sup>31</sup>. Respondents must choose between two scenarios, A and B. In the former case, all doses of the vaccine are distributed among children living in Region 1; in the latter case, each region receives half of the doses. Because it is assumed that half of the children live in Region 1 and half of the children live in Region 2, Scenario A is not chosen over Scenario B on the grounds that Region 1 has more children than Region 2.

In the first stage, both scenarios produce the same total health gain (100 episodes of illness avoided) but, while in Scenario A, this gain is concentrated in Region 1, in Scenario B, the total health gain is equally divided between the two regions. From the perspective of health maximisation, respondents should be indifferent between the two scenarios. However, unless people disagree with the inclusion of geographic distribution of resources in the set of guiding criteria of resource allocation, we expect Scenario B to be chosen over Scenario A.

Further note that Scenario B represents geographic equality of treatment in the sense that the relation between the total number of doses of the vaccine and the total number of children is the same in both regions. However, in our questionnaire, geographic equality of treatment also results in equal distribution of health gain between the two regions. This need not be the case (we could simply compare the total health gain obtained in both Scenarios and ignore how it is distributed between regions) but it greatly simplifies the questionnaire. It is nonetheless not possible to disentangle both effects of geographic equality of treatment and geographic equality of health gain. Still, the description of Scenario B emphasises precisely, and brings the attention of respondents to, the issue of equality of treatment.

---

<sup>30</sup> In a cross-country study in which respondents were given two choices for managing health services- 'unlimited funding' or 'limits should be set – some treatments given priority' – in almost all countries more than 50% of respondents chose the former option (Mossialos and King, 1999). When given the possibility, people tend to avoid hard decisions (Dolan and Cookson, 2000).

<sup>31</sup> We had initially used the designations 'North' and 'South', however, after the pilot study we changed it to Regions 1 and 2 to avoid any association with the North and South of Portugal.

If Scenario B is chosen in the first part of question 3, then the respondent should continue on the following page. In the second part of question 3, it is said that Scenario B actually produces a total health gain lower than previously stated. There are five hypotheses and while the health gain is kept constant under Scenario A, the total health gain associated with Scenario B is progressively reduced (being health gain always equally distributed between regions). It is expected that at some point respondents change their choice from Scenario B to Scenario A. The exact point where this change occurs depends on how people trade-off geographic equality of treatment against health maximisation. The more people insist in choosing Scenario B over Scenario A, the greater the relative value that they attach to geographic equality of treatment compared to health maximisation. Note that what is at stake is a greater or lower total health improvement. Therefore, when we talk about loss (or sacrifice), this is not an effective loss; what it means is giving up of a given health improvement.

When the total health gain produced under Scenario B is lower than the total health gain produced under Scenario A, it means that meeting geographic equality of treatment has an opportunity cost in terms of total health gain. In our case, this cost can be measured in terms of number of persons (children). Actually, in the second part of question 3, choosing one Scenario over the other always involves some sacrifice. If Scenario B is chosen, the same number of children from both regions will avoid the disease but globally fewer children will avoid it. The sacrifice can be measured in terms of a given number of children living in Region 1. On the other hand, choosing Scenario A means that in total more children will avoid the disease but now no such gain is verified in Region 2. Here, the sacrifice can be measured in terms of a given number of children living in Region 2.

This kind of approach is known as the person trade-off (PTO), in which subjects are asked to decide what sacrifices they are prepared to make in the lives of some people in order to provide health benefits to some other people (Nord, 1995). The person trade-off technique has been proposed as a way of ascertaining the relative values that respondents attach to criteria relevant to social decisions. Unlike methods such as the standard gamble and the time trade-off in which individuals are asked to value health states for themselves, the person trade-off is proposed as a technique more appropriate to use in a resource allocation context

in which respondents are asked to distribute resources between groups of other people (Nord, 1995; Nord *et al.*, 1999).

Green (2001) presents the main pros and cons of the person trade-off technique and based on empirical evidence regarding its practicality, reliability and empirical validity, he makes an overall evaluation of the PTO. According to the author, the main advantages of the PTO are its clarity of meaning and directness of measurement. Moreover, given that allocating scarce resources in health care essentially means dealing with person trade-offs (Nord, 1995), it is said that PTO asks the right question (Prades, 1997).

Regarding the limitations, Green (2001) concludes that available evidence on the acceptability and feasibility of the PTO technique is inconclusive and that empirical evidence on its reliability is virtually non-existent. Green (2001) says that it would seem that respondents often require greater explanation of the task than can be offered through a written survey instrument. Nord (1995) also notes that other factors such as the arguments used in the questions, the start-point, the numbers in pairwise comparisons and the decision context, may affect PTO responses.

Green (2001) nonetheless concludes that the limitations of the person trade-off technique should not deter its further development, being PTO an appealing option for the elicitation of societal preferences and offering a means of bringing into the open the real trade-offs that decision-makers are frequently faced with (pp. 239-240). Moreover, the author argues that other alternative methods have their own weaknesses and that “at the present time all techniques for the measurement of societal values in health care are in their early stages of development” (p.240). Some posterior works (e.g., Dolan and Tsuchiya, 2003; Mansley and Elbasha, 2003; Schwarzinger *et al.*, 2004) have looked into methodological aspects of the PTO, underlining its limitations without nonetheless defending its abandonment.

Based on question 3, the variable TOFF3 is defined as the number of times respondents select Scenario B over Scenario A in this question. TOFF3 can therefore assume values from 0 to 5, showing how far respondents are willing to go in terms of giving up of additional health improvements in order to keep geographic equality of treatment; the greater the value of TOFF3, the greater the willingness to forego health gain. Naturally, the hypotheses in which Scenario B

is chosen (before changing to Scenario A) should be consecutive; otherwise answers should be ignored and eliminated from the analysis. In the case of respondents rightfully skipping this part because Scenario A was chosen in page 3 of the questionnaire, then TOFF3 is attributed the code '-1' (see Table 1). Contrarily, if respondents skip this part, despite having chosen Scenario B in the first part, then TOFF3 is registered as a missing value (999).

If the respondent selects Scenario A right from the start (implying that  $TOFF3=0$ ), this means that, although he or she is not indifferent to geographic equality of treatment (in the sense that Scenario B was chosen in the first part of the question though, one must acknowledge, we force the respondent to select one out of the two scenarios), he or she is not willing to forego any potential health gain in order to keep in return geographic equality of treatment. More specifically, the respondent is not prepared to give up of 10% of the total health gain. Still, we must note that Scenario B could eventually be chosen by the respondent for trade-offs smaller than that considered in the first hypothesis.

In hypothesis 1, choosing Scenario B (90 children) implies the acceptance of a 10% loss in total health benefit, compared to Scenario A (100 children). In hypothesis 2, this loss raises to 30% (70 children vis-à-vis 100 children); in hypothesis 3, the loss is 50% (50 children vis-à-vis 100 children); in hypothesis 4, the loss is 80% (20 children vis-à-vis 100 children); finally, in hypothesis 5, the loss is 90% (10 children vis-à-vis 100 children). If a respondent chooses Scenario B up to, say, hypothesis 3 and then selects Scenario A in the hypotheses 4 and 5, then  $TOFF3=3$ , meaning that this respondent is prepared to forego 50% of the potential health gain in order to keep geographic equality of treatment. But it also means at the same time that, according to this respondent's preferences and given the set of options presented in the questionnaire, the maximum acceptable opportunity cost of geographic equality of treatment is 50% of the total potential health gain.

Question 4 is very similar to question 3; its objective being to elicit trade-offs between geographic equality of treatment and health maximisation, using the person trade-off as well. It applies to a hypothetical situation but now actual Portuguese regions are involved and the introduction of the question uses real 2006 data. Also, we use an example that is quite feasible given the strategy of the Ministry of Health in recent years. In a way, this question introduces a further

difficulty for respondents, in the sense that they are now led to think about actual regions, including that where they live. It is still a hypothetical situation but we reduce a bit the level of abstraction. And this is not an irrelevant issue when it comes to preferences for guiding principles of resource allocation (think for example about the rule of rescue, reviewed in chapter two). Some authors have recommended the use of realistic examples in surveys to study on a less abstract level support for different criteria to be employed in resource allocation (Cuadras-Morató *et al.*, 2001; Stolk *et al.*, 2005). Though it may also be argued that introducing too many details (instead of having questions described in a vague fashion) may lead respondents to focus on considerations of feasibility (Tsuchiya and Dolan, 2007).

Respondents are asked to decide about the allocation of extra resources among three regions of Portugal. The objective is to reduce the number of people waiting for surgery beyond the clinically acceptable time. We inform respondents that the number of people in waiting lists for surgery and the average waiting time are the same in all three regions (the actual numbers are not exactly the same but they are very similar). Thus, from this perspective, all regions might be considered equally entitled to the extra resources in question.

Respondents must choose between two alternatives, A and B. In the former case, the total amount of resources is concentrated on one region (Lisbon and Tagus Valley) and, in the latter case, resources are equally divided among the three regions (North, Centre and Lisbon and Tagus Valley). Nothing is said about the benefits of having a surgery. However, because there are no particular reasons to think that the characteristics of people waiting for a hip replacement or cataract surgery (the examples used in the questionnaire) significantly differ from one region to another, with some confidence we can say that responses were based on the assumption that the average benefit per surgery is constant across regions and therefore the relevant information to decide about the allocation of resources is the number of surgeries. For this reason, in what follows, we also consider that total health gain is proportional to the total number of surgeries.

In the first part of question 4, respondents must choose between allocating all resources to Lisbon and Tagus Valley, benefiting 10.000 people (Alternative A) and allocating resources to all three regions, benefiting 3.300 people in each region (Alternative B). Here, right from the beginning there is a trade-off: 10.000



surgeries concentrated on one region compares to 9.900 surgeries equally divided among three regions (in this case, selecting Alternative B represents a loss of 1% of the total potential health gain). Choosing the alternative that produces greater benefit implies a sacrifice born by people living in the North and Centre. Thus, respondents are faced with potential sacrifices affecting the regions where they reside<sup>32</sup>.

For those respondents selecting Alternative B, in the first part of question 4, the questionnaire continues on the following page, where the total benefit under Alternative A is kept constant and the total benefit under Alternative B is progressively reduced. Whether and when people change their choice from Alternative B to Alternative A, again depends on how people trade-off geographic equality of treatment against health maximisation. In proportional terms, the losses regarding the potential health improvements foregone, in the five hypotheses considered in question 4, are about the same as in question 3.

The variable TOFF4 is defined as the number of times that Alternative B is chosen in the second part of question 4. The interpretation and codification of TOFF4 (see Table 1) are *mutatis mutandis* as those of TOFF3.

Dolan *et al.* (2002) demonstrate how the stated preferences of a sample of the general public can be used to estimate the parameters of a social welfare function (SWF) in the domain of health benefits. Therefore, responses to questions 3 and 4 are also used to derive these parameters.

Throughout the previous chapter we have not assumed any particular functional form for the social welfare function and there are in fact a number of functional forms that it can take (see e.g., Wagsaff, 1991 and Dolan, 1998). Dolan *et al.* (2002) adopt an additive SWF with a constant elasticity of substitution:

$$W = \left[ \alpha H_a^{-r} + \beta H_b^{-r} \right]^{\frac{1}{r}}, \quad H_a, H_b \geq 0, \quad \alpha + \beta = 1, \quad r \geq -1, \quad r \neq 0$$

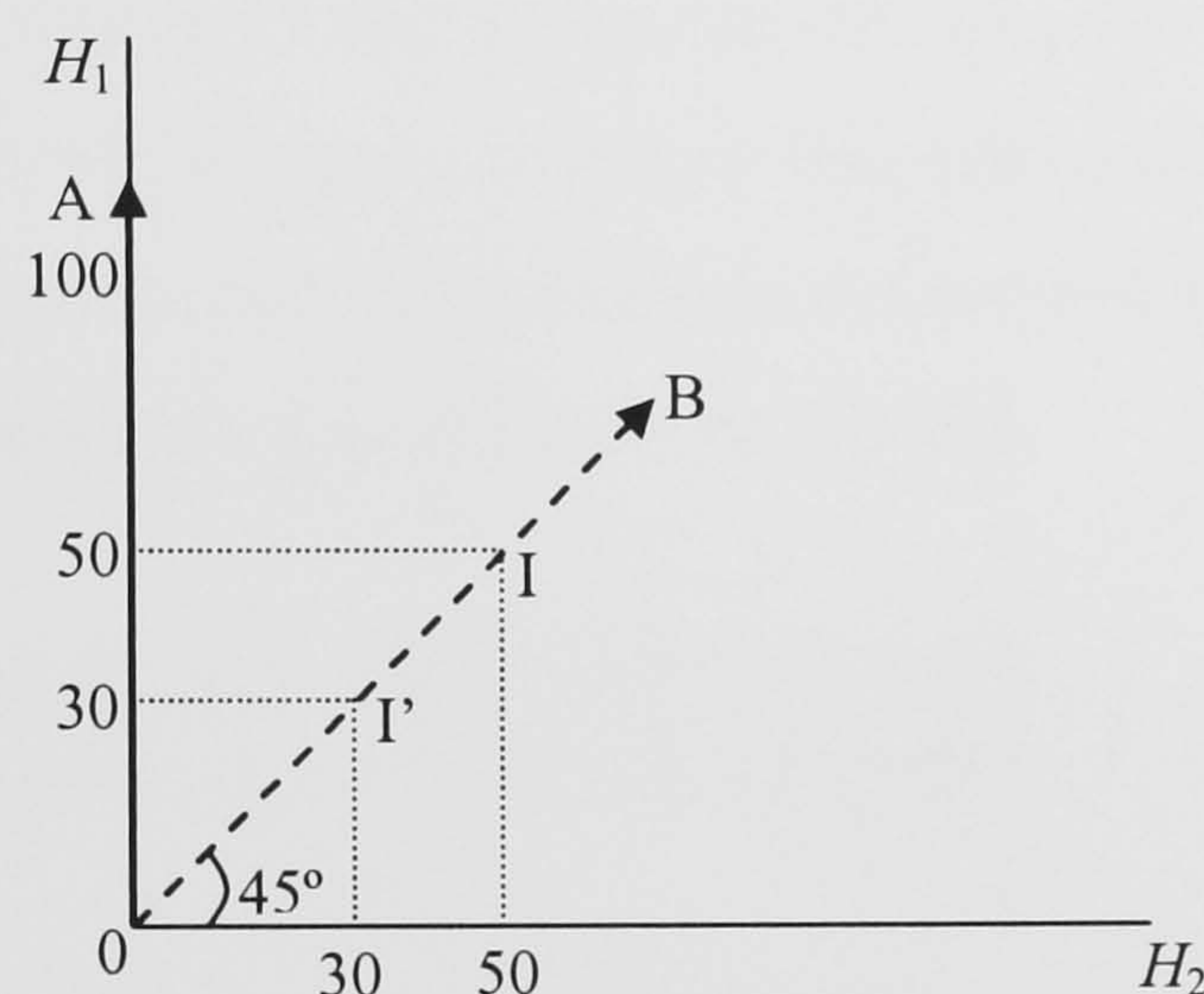
where  $W$  is the level of overall population health and  $H_a$  and  $H_b$  are the levels of health of groups of equal size. In our case,  $W$  represents the overall population health gain (given that we do not provide information about health endowments) and the two groups ( $a$  and  $b$ ) are defined by residence. Dolan *et al.* (2002) assume that  $\alpha$  and  $\beta$  (representing the relative value attached to health gain

---

<sup>32</sup> Bragança is located in the North and Coimbra is located in the Centre, as shown in the Appendix A.

going to each group) are both equal to 0,5 (that is, both groups are considered equally deserving of any health gain). The parameter  $r$  measures the degree of aversion to inequality: if  $r=-1$ , then social welfare is a function of total health and indifference curves will be straight lines with a gradient of -1; if  $r>-1$  there is aversion to inequality and indifference curves become convex to the origin; in the extreme,  $r$  approaches  $\infty$  and indifference curves become L-shaped. In our case,  $r$  should be thought as aversion to inequality in the distribution of health gain.

To calculate the parameter  $r$  we need first to determine the point of indifference between the two alternative programmes that respondents are faced with.



**Figure 15. Deriving points of indifference from trade-off questions**

Figure 15 shows the case of question 3 ( $H_1$  and  $H_2$  represent the health gain going to *Region 1* and *Region 2*, respectively). There are two alternatives to allocate resources: in all hypotheses, Scenario A allows 100 children from *Region 1* and none from *Region 2* to avoid the disease (this corresponds in Figure 15 to a movement in the vertical axis, from 0 to 100); in turn, Scenario B always produces equal benefit for both regions, thus, it corresponds in Figure 15 to movements along the  $45^\circ$  line through the origin. The aim is then to determine what is the minimum benefit that Scenario B would have to provide to each region in order to be considered equally as valuable as Scenario A. Note that the latter represents maximisation of health gain, while the former represents geographic equality in the distribution of total health gain.

In the first part of question 3, respondents are asked to choose between two alternatives represented in Figure 15 by points A and I. Respondents were not actually given the opportunity to state that they were indifferent between the two

Scenarios but if points A and I were considered equally valuable, then we would have an indifference curve with a gradient of -1.

As in Dolan *et al.* (2002), for those respondents that initially chose Scenario B and then switched at some point to Scenario A, we consider that their point of indifference is half-way between the last point at which they chose Scenario B and the first point at which they chose Scenario A. For instance, if a respondent chose Scenario B in hypotheses 1 and 2 and then switched to Scenario A in hypotheses 3,4 and 5 (TOFF3=2), he is willing to forego between 30 and 50 children, the midpoint of which is 40. Thus, this respondent is indifferent between 100 children concentrated in *Region 1* and 60 children equally distributed between the two regions. In Figure 15 this means that the respondent is indifferent between points A and I' (the indifference curve now becomes convex).

Once indifference between alternatives has been established, the value of  $r$  can be calculated. According to Dolan *et al.* (2002):

$$r \approx \frac{\log((H_a(A) - H_a(B))/(H_b(B) - H_b(A)))}{\log((H_a(A) + H_a(B))/(H_b(A) + H_b(B)))} - 1$$

The authors further note that the precise trade-offs made by those who never switch between alternatives is indeterminate.

For the example above (TOFF3=2):

$$r \approx \frac{\log((H_1(A) - H_1(I'))/(H_2(I') - H_2(A)))}{\log((H_1(A) + H_1(I'))/(H_2(A) + H_2(I')))} - 1 = \frac{\log((100 - 30)/(30 - 0))}{\log((100 + 30)/(30 + 0))} - 1$$

Another parameter that can be derived from responses to questions 3 and 4 is the implied trade-off between health equality and health maximisation. It is calculated as the difference between average health and the 'equally-distributed equivalent health' (the overall population health that, if distributed equally across the population, is as good as a given unequal distribution - Dolan *et al.*, 2002, p.17). So, for the case of TOFF3=2, the respondent is indifferent between the point where 100 children from *Region 1* and none from *Region 2* avoid the disease and the point where 30 children from both regions avoid the disease. This means that he would be willing to trade-off up to 20 children of the average number of children avoiding the disease of these two regions if the total number were

distributed equally between them.

In question 5, we hypothesise a screening programme for colon cancer, in which tests are offered free of cost within the National Health Service. The success of screening programmes depends a great deal on people voluntarily coming forward to take tests, which in turn might be affected by some attributes of health care delivery. In this case, the attribute is the doctor's gender (the case of colon cancer was deliberately chosen given that it might be sensitive for both men and women).

What we ask respondents is whether or not the possibility of choosing the gender of the professional performing the test has an impact on their decision to take the test. We might expect that in more conservative regions or regions where religion plays a greater role in people's lives, individuals will be more sensitive to this issue. It might also be that once individuals come forward to take the test, they do not feel very strong about their preferences up to the point of exercising their choice regarding the doctor's gender, hence the second part of question 5 (see Appendix B). Two variables are defined to address the ('Yes' or 'No') responses to question 5: DOCGEND and CHOICE (see Table 1).

The last question of the questionnaire seeks to identify preferences for different forms of action, in the face of given health symptoms. As in the previous question, these are personal preferences (we ask individuals about what they prefer for themselves), which might affect the demand for health care. There are five situations, representing successively greater deterioration in the individual's health. For each situation, there are four alternative forms of action, representing preferences that require successively greater effort from the individual, in the process of using health care resources (see Appendix B). One might think about two main thresholds: one below which individuals prefer not to seek any form of care and another below which individuals prefer to seek forms of care (if any) other than visiting a doctor. The latter case might have a direct impact on the utilisation of public health care resources.

We derive two variables from question 6, THRESH61 and THRESH62, representing the two thresholds above (see Table 1). The two thresholds might be different or not. That is, it might be that people start by doing, for example, self medication, or, it might be that people do nothing until they visit the doctor. Let us consider an example, in which an individual answers to question 6 as follows:

1 - 'I do nothing'; 2 - 'I do self-medication'; 3, 4 and 5 - 'I visit the doctor'. Here, we would have THRESH61=2 (given that the first situation involving some form of action is that described in hypothesis 2) and THRESH62=3 (given that the first time that the option 'I visit the doctor' is selected is in hypothesis 3). Further note that, because the five hypotheses represent successively greater deterioration in the individual's health, once respondents choose 'I visit the doctor', we should expect to see this same answer in the remainder situations.

In the final part of the questionnaire respondents are asked to indicate their gender, age and level of education (corresponding to the last three variables in Table 1, which are only used to characterise and compare our samples with their respective local populations). The variable RESID, in Table 1, represents the region of residence, which is either Coimbra or Bragança<sup>33</sup>.

The questionnaire was self-completed by respondents and we adopted a delivery and collection method of administration. No covering letter accompanied the survey form, thus, the first page of the questionnaire provided some information about the aim, nature and relevance of the questionnaire. Contact details were also provided in the front page. Respondents were told in advance that any query could be addressed to us, using these contacts and that at the time of collection there would be a further opportunity to clarify any remaining queries.

The method of delivery and collection offers the advantage of ensuring a high rate of response within a short period of time, still, not all questionnaires were returned and several follow-up personal contacts were necessary to reach our targets regarding the size of samples and to meet what we had defined as the deadline for collection.

---

<sup>33</sup> This information was inserted in the first page of the questionnaire, as completed questionnaires were collected (we used the small box on the top of the front page to insert a survey form identifier, composed by the letters C or B, indicating the region of residence, and a number).

**Table 1. Variables designation and codification**

Question*	Variable designation	Codification
—	RESID	=1 if respondent resides in Coimbra; =2 if respondent resides in Bragança
1	HMAX	=1 if strongly disagree; =2 if disagree; =3 if neutral; =4 if agree; =5 if strongly agree
2	HGEO	=1 if strongly disagree; =2 if disagree; =3 if neutral; =4 if agree; =5 if strongly agree
3	QUEST3	=1 if choose Scenario A in page 3; =2 if choose Scenario B in page 3
3	TOFF3	If QUEST3=1 then TOFF3=-1; Otherwise, TOFF3 = number of times that Scenario B is selected in page 4 (0 to 5)
4	QUEST4	=1 if choose Alternative A in page 5; =2 if choose Alternative B in page 5
4	TOFF4	If QUEST4=1 then TOFF4=-1; Otherwise, TOFF4 = number of times that Alternative B is selected in page 6 (0 to 5)
5	DOCGEND	=1 if respondent answers 'YES' in the first part; =2 if respondent answers 'NO' in the first part
5	CHOICE	=1 if respondent says that would exercise choice =2 if respondent says that would be indifferent
6	THRESH61	= number of first hypothesis involving some form of action: ranges from 1 to 5
6	THRESH62	= number of first hypothesis involving a visit to the doctor: ranges from 1 to 5
—	GENDER	=1 if respondent is male; =2 if respondent is female
—	AGE	=1 if 15 to 24 years; =2 if 25 to 44 years; =3 if 45 to 64 years; = 4 if ≥65 years
—	LEVELED	= 1 if no degree; =2 if 1 <sup>st</sup> cycle or 2 <sup>nd</sup> cycle or 3 <sup>rd</sup> cycle; =3 if secondary; =4 if medium or superior; =5 if attending (any degree)

\* See Appendix B

#### 4.3.2. The respondents

This study was carried out in two Portuguese municipalities, Coimbra and

Bragança, in March-April 2007. The selection of respondents (described below) ensures that the two samples are independent; being each one composed by 35 members of the general public.

Although Portugal is a small country, there are some differences across regions. People living in the north tend to be more conservative when compared to people living further in the south. There are also apparent differences between coastal and inland areas. These differences emerge when there are, for example, general elections. They also clearly emerged, in February 2007, when a referendum about the de-penalisation of voluntary abortion up to 10 weeks of pregnancy, was carried out in Portugal. The referendum question was of the type ‘Yes’ (agree with de-penalisation) or ‘No’. In regions such as Braga, in the north, 59% of people voted ‘No’, while 41% voted ‘Yes’; in Beja, in the south (Alentejo), 84% of people voted ‘Yes’, while only 16% of people voted ‘No’.

In this context, the choice of the two regions, from which to draw the samples, was partly oriented by practical convenience and partly judgemental. That is, we started by considering a series of regions that for several reasons are more accessible to us and then, from this set, we chose two that reflected most closely the division of the country in terms of north *versus* south and coastal *versus* inland areas<sup>34</sup>.

We have consequently a sample drawn from a municipality (Coimbra) located in the most populated coastal line in the centre of the country and another drawn from a municipality (Bragança) located in a northern and inland area of the country – see Appendix A. Out of curiosity, the results of the referendum (above mentioned) in Coimbra and Bragança, were as follows: in the former case, 63% of people voted ‘Yes’ and 37% voted ‘No’; in the latter case, 59% of people voted ‘No’ and 41% voted ‘Yes’.

Considering that this study was conceived from the start as a small-scale study, we followed existing rules of thumb regarding the size of samples (Saunders *et al.*, 2003), aiming for a minimum of 30 respondents in each sample. We used the method of quota sampling to define the composition of samples. The characteristics used to set the quotas were gender, age and level of education. We

---

<sup>34</sup> Besides, taking into account that this is an exploratory study, the choice of regions did not seem to be of critical importance.

defined four age bands: 15-24; 25-44; 45-64; and 65 or older. Concerning education, we used 5 categories: no degree successfully completed; basic (which in turn includes first cycle – four years of schooling in total; second cycle – six years of schooling; and third cycle – 9 years of schooling); secondary; superior; and currently attending (any degree). We considered using income instead of level of education; however, Portuguese people are usually not receptive to questions related to income.

Based on the composition of local populations and given the projected size of samples, we defined quotas and selected respondents according to their gender, age and, whenever possible, according to their level of education. The groups harder to reach were older people and people with lower or no degree of education successfully completed. We distributed 40 survey forms in Bragança and 40 in Coimbra. Overall, it was not possible to collect five questionnaires; three were eliminated due to inconsistencies (reasons for exclusion are explained in sub-section 4.3.1); and two were deliberately not included to keep samples with equal sizes.

Tables 2 and 3 show the characteristics of our respondents, in each sample, and how they compare with the respective local population (population data refers to the municipality level). It can be seen that in both cases the samples are definitely better educated and slightly younger than would have been the case had the samples been representative of the wider population. In particular, people with 65 years or older are under-represented (with the age band 15-24 being over-represented) as well as people with no degree of education successfully completed and people that completed only the basic level (in this case, the group that ended up over-represented concerns individuals with superior education).



Table 2. Respondent characteristics in sample 1 - Coimbra (NI=35)

	Gender	Age														
		TOTAL				15-24			25-44			45-64		65+		
		N	(%)	Population Data <sup>a</sup> (%)	Sample 1 N	(%)	Population Data <sup>a</sup> (%)	Sample 1 N	(%)	Population Data <sup>a</sup> (%)	Sample 1 N	(%)	Population Data <sup>a</sup> (%)	Sample 1 N	(%)	Population Data <sup>a</sup> (%)
	Male/Female	35	100	100	8	22,86	16,96	12	34,28	34,54	11	31,43	29,31	4	11,43	19,19
	Male	16	45,71	46	4	11,43	8,44	6	17,14	16,58	4	11,43	13,61	2	5,71	7,68
	Female	19	54,29	54	4	11,43	8,52	6	17,14	17,96	7	20,00	15,70	2	5,71	11,51
<b>Education</b>																
None	Male/Female			6,47			0,06			0,27			1,25			4,89
	Male			1,60			0,03			0,13			0,34			1,10
	Female			4,87			0,03			0,14			0,91			3,79
Basic	Male/Female	9	25,72	44,30			2,31	3	8,57	13,04	4	11,43	17,95	2	5,71	11,01
	Male	4	11,43	21,13			1,48	2	5,71	6,65	1	2,86	8,16	1	2,86	4,85
	Female	5	14,29	23,17			0,83	1	2,86	6,39	3	8,57	9,79	1	2,86	6,16
Secondary	Male/Female	4	11,43	13,95			1,84			7,57	4	11,43	3,28			1,24
	Male	2	5,71	7,52			1,00			3,99	2	5,71	1,84			0,67
	Female	2	5,71	6,43			0,84			3,58	2	5,71	1,44			0,57
Superior	Male/Female	14	40,00	18,32			0,58	9	25,72	9,30	3	8,57	6,49	2	5,71	1,97
	Male	6	17,14	7,93			0,17	4	11,43	3,65	1	2,86	3,09	1	2,86	1,03
	Female	8	22,86	10,39			0,41	5	14,29	5,65	2	5,71	3,40	1	2,86	0,94
Attending	Male/Female	8	22,86	16,96	8	22,86	12,17			4,36			0,35			0,08
	Male	4	11,43	8,12	4	11,43	5,75			2,16			0,18			0,03
	Female	4	11,43	8,84	4	11,43	6,42			2,20			0,17			0,05
<b>TOTAL</b>	Male/Female	35	100	100	8	22,86	16,96	12	34,28	34,54	11	31,43	29,31	4	11,43	19,19

a- Population data (for the municipality of Coimbra) come from 2001 Portuguese Census

Table 3. Respondent characteristics in sample 2 - Bragança (N2=35)

	Age																							
	TOTAL			15-24			25-44			45-64			65+											
	Sample 2	Population Data <sup>a</sup>	Gender	N	(%)	Sample 2	Population Data <sup>a</sup>	Gender	N	(%)	Sample 2	Population Data <sup>a</sup>	Gender	N	(%)	Sample 2	Population Data <sup>a</sup>	Gender	N	(%)	Sample 2	Population Data <sup>a</sup>		
None	Male/Female	100	Male/Female	35	100	16,84	12,38	Male/Female	13	37,14	31,68	2,46	Male/Female	4	11,42	28,80	9,1	Male/Female	10	28,57	28,80	4	11,42	22,68
	Male	48	Male	16	45,71	8,32	4,55	Male	6	17,14	15,67	0,89	Male	2	5,71	14,02	3,23	Male	4	11,43	14,02	2	5,71	10
	Female	52	Female	19	54,29	8,52	7,83	Female	7	20,00	16,01	1,57	Female	6	17,14	14,78	5,87	Female	6	17,14	14,78	2	5,71	12,68
Basic	Male/Female	31,45	Male/Female	5	14,29	3,28	52,11	Male/Female	5	14,29	15,69	20,69	Male/Female	1	2,86	20,69	12,56	Male/Female	5	14,29	20,69	1	2,86	12,56
	Male	27,73	Male	6	17,16	2,05	27,73	Male	2	5,71	8,90	10,50	Male	3	8,57	10,50	6,27	Male	3	8,57	10,50	1	2,86	6,27
	Female	24,38	Female	6	17,16	1,23	24,38	Female	3	8,57	6,79	10,19	Female	2	5,71	10,19	6,19	Female	2	5,71	10,19	1	2,86	6,19
Secondary	Male/Female	10,21	Male/Female	3	8,57	1,83	10,21	Male/Female	1	2,86	6,04	2,02	Male/Female	2	5,71	2,02	0,34	Male/Female	1	2,86	2,02	2	5,71	0,34
	Male	5,46	Male	2	5,71	1,06	5,46	Male	1	2,86	2,99	1,23	Male	2	5,71	1,23	0,20	Male	1	2,86	1,23	2	5,71	0,20
	Female	4,75	Female	1	2,83	0,77	4,75	Female	1	2,86	3,05	0,79	Female	1	2,86	0,79	0,14	Female	1	2,86	0,79	1	2,86	0,14
Superior	Male/Female	10,39	Male/Female	13	37,15	0,54	10,39	Male/Female	8	22,86	5,92	3,19	Male/Female	1	2,86	3,19	0,73	Male/Female	4	11,43	3,19	1	2,86	0,73
	Male	3,73	Male	5	14,29	0,16	3,73	Male	4	11,43	2,02	1,26	Male	1	2,86	1,26	0,29	Male	1	2,86	1,26	1	2,86	0,29
	Female	6,66	Female	8	22,86	0,38	6,66	Female	4	11,43	3,90	1,93	Female	3	8,57	1,93	0,44	Female	4	11,43	1,93	1	2,86	0,44
Attending	Male/Female	14,91	Male/Female	8	22,86	10,97	14,91	Male/Female	8	22,86	3,43	0,44	Male/Female	1	2,86	0,44	0,05	Male/Female	8	22,86	3,43	1	2,86	0,05
	Male	6,52	Male	4	11,43	4,90	6,52	Male	4	11,43	1,44	0,15	Male	1	2,86	0,15	0,01	Male	4	11,43	1,44	1	2,86	0,01
	Female	8,39	Female	4	11,43	6,07	8,39	Female	4	11,43	1,99	0,29	Female	3	8,57	0,29	0,04	Female	4	11,43	1,99	1	2,86	0,04
TOTAL	100	TOTAL	35	100	16,84	100	TOTAL	35	100	31,68	28,80	TOTAL	35	100	28,80	22,68	TOTAL	35	100	28,80	4	11,42	22,68	

a- Population data (for the municipality of Bragança) come from 2001 Portuguese Census

### 4.3.3. The analysis

In this sub-section we describe the methods used to analyse the data. Two kinds of results are of interest: (i) the people's views regarding the issues addressed in the questionnaire; and above all (ii) the differences between samples in terms of people's views. Descriptive statistics and graphical methods are thus firstly used to explore the data, giving an account of respondents' preferences and an initial overview of differences between samples. Then these differences are analysed using inductive statistical methods.

We begin by presenting some (comparative) descriptive statistics, showing frequency tables, for all variables, calculating afterwards some summary statistics to describe the empirical distribution of each variable. Regarding measures of central tendency, we calculate the mean for HMAX and HGEO. Siegel and Castellan (1988, p.28) argue that in matters of opinion categories such as 'agree', 'ambivalent' and 'disagree' may be thought to fall on a continuum reflecting strength of agreement/disagreement. Thus, the calculation of average values for the case of HMAX and HGEO is meaningful. We also calculate the mean for the implied parameter  $r$  and the implied trade-off as well as the median for TOFF3, TOFF4, THRESH61 and THRESH62.

In addition to descriptive statistics, we conducted some statistical tests (*STATA ver. 9.0*) to evaluate statistical hypotheses formulated in order to assess the cross-sample variation in preferences (as reflected by the values of the variables defined in Table 1). In all tests, the null hypothesis is that both samples have been drawn from the same population (or, from populations with the same mean, median or distribution, depending on the test). In each case, the statistical hypothesis to be tested is:

$H_0$ : There are no differences in preferences across samples

$H_1$ : There are differences in preferences across samples

Consequently, to conclude that there are differences in preferences between samples, the null hypothesis should be rejected. Tests are all two-tailed given that we are interested in testing for differences, irrespective of the relation between values obtained in both samples.

The choice of the specific test to apply in each situation was based mainly

on Siegel and Castellan (1988) and on Saunders *et al.* (2003). To test whether or not differences between the two (independent) samples, regarding the mean of HMAX, HGEO, implied parameter  $r$  and implied trade-off are statistically significant, we use the Student's  $t$ -test. Despite that the  $t$ -test assumes that the data is normally distributed, it is argued that this can be ignored without too many problems even with sample sizes of less than 30 (Saunders *et al.*, 2003, p.361). The same applies to the assumption that the data for the two groups have the same variance, provided that the two samples are of similar sizes. Regarding the variables TOFF3, TOFF4, THRESH61 and THRESH62, we test for differences between samples using the Wilcoxon-Mann-Whitney test (in the case of obtaining different medians).

In addition to the previous tests, we used the Kolmogorov-Smirnov two-sample test, which is a non-parametric and distribution free statistical test, to check for differences, across samples, in the distributions of most variables.

Concerning the variables DOCGEND and CHOICE, we use the Fisher Exact test for 2x2 tables to examine the significance of the association between each of these variables and the variable RESID. To conclude that there is a statistically significant association between DOCGEND and CHOICE, on one hand, and RESID, on the other hand (i.e., to conclude that residence matters when determining preferences), the null hypothesis should be rejected as in the other tests.

The variables QUEST3 and QUEST4 are also eligible to apply the Fisher Exact test for 2x2 tables; however the results obtained in descriptive statistics decisively dispense further tests because there is basically a perfect match between responses in one sample and responses in the other sample.

The final test used is the Wilcoxon matched-pairs signed-ranks test to test within-respondent differences; more specifically, it is used to check whether or not the distribution of responses given to question 3 (TOFF3) and question 4 (TOFF4) are significantly different. The same statistical method is used to test for differences between THRESH61 and THRESH62.

#### 4.4. Results

The principle of health maximisation gathers the support of basically all respondents. This result is not surprising given (as previously explained) that this principle was presented as one among other possible alternatives. The average value of HMAX (see Table 4) is 4,5 for respondents from Coimbra and 4,2 for respondents from Bragança, showing that opinions lie between agree and strongly agree.

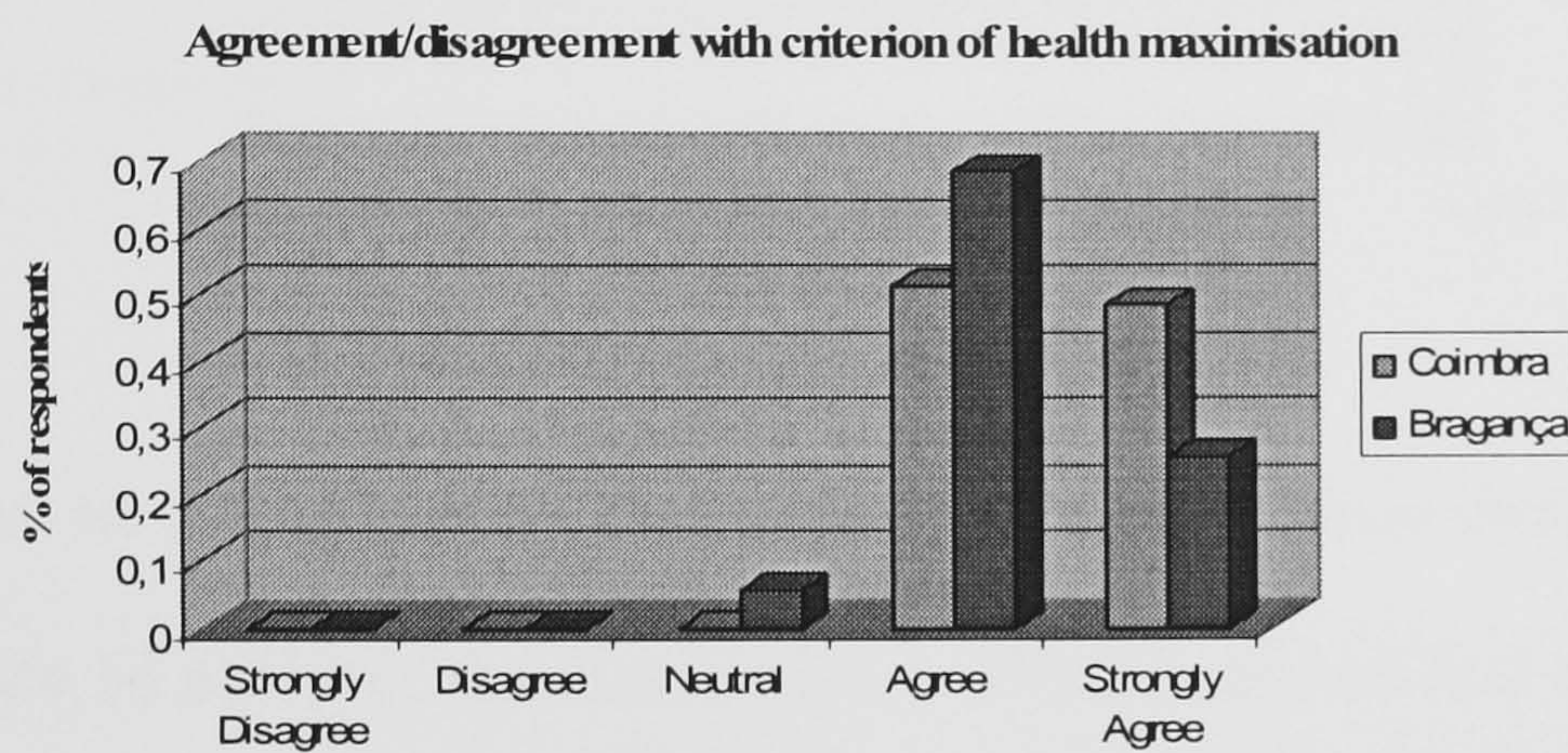


Figure 16. *Bar chart of HMAX by region of residence*

Figure 16 further shows that the difference between means derives from the fact that respondents in Coimbra are equally distributed between the category ‘agree’ and ‘strongly agree’ (18 and 17 respondents, respectively), while in Bragança there is a greater concentration on the category ‘agree’ compared to ‘strongly agree’ (9 *versus* 24 respondents).

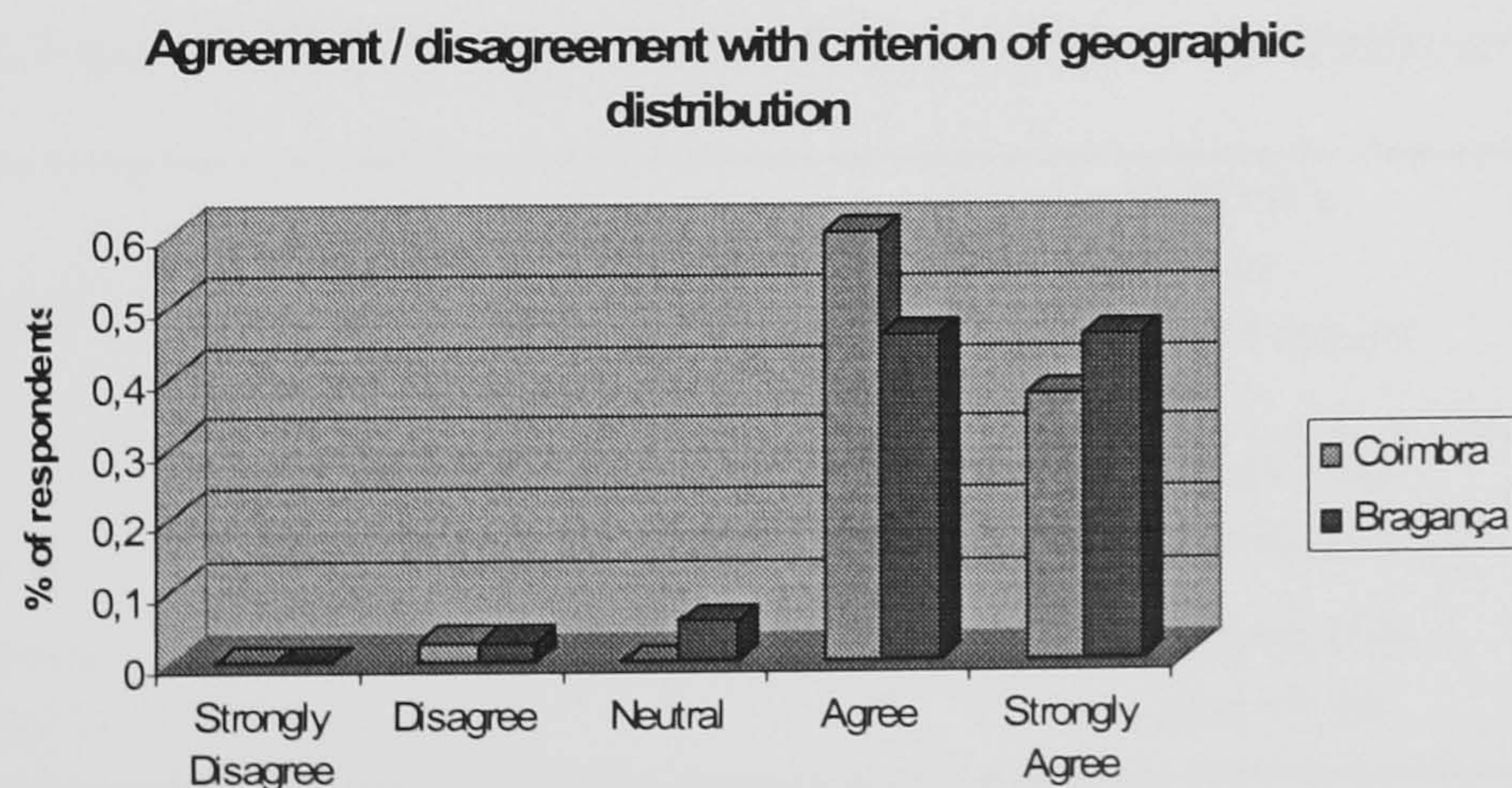


Figure 17. *Bar chart of HGEO by region of residence*

Regarding the variable HGEO, again opinions lie between agree and strongly agree, with an average value of 4,3 (Table 4) for both regions. We now have an equal percentage of respondents in the categories ‘agree’ and ‘strongly

agree' in Bragança, while in Coimbra there is a greater concentration on the category 'agree' (Figure 17). The means are still the same given that in Bragança some respondents are neutral regarding the geographic distribution of resources, therefore lowering the average in this group.

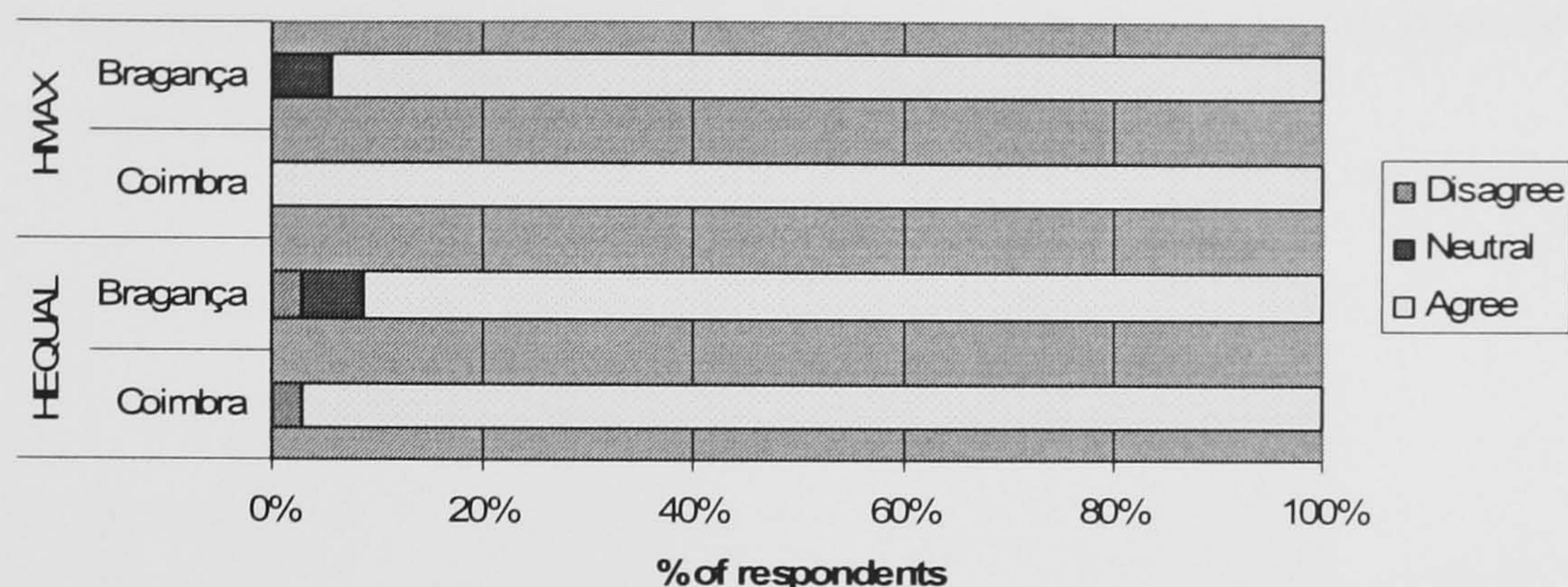


Figure 18. *HMAX and HGEO by region of residence, with three categories*

Figure 18 analyses the distribution of HMAX and HGEO considering only three categories: disagree, neutral and agree. The most noticeable result is the strong support to both criteria as potential guiding principles of health care resource allocation. Curiously, in both samples there was one respondent that manifested a negative opinion regarding any role played by geographic distribution of resources. Also, for the case of Bragança, there is an equal percentage of respondents that are neutral regarding the two criteria (we note that the respondents are not the same).

Table 4. *HMAX and HGEO by region of residence: summary statistics and hypothesis test results*

Var.	Region	Mean	S.D.	Max.	Min.	Student's <i>t</i> -test H <sub>0</sub> : equal means	Kolmogorov- Smirnov test H <sub>0</sub> : equal distributions
HMAX	Coimbra	4,50	0,508	5	4	<i>t</i> -value =2,40 p-value=0,019	P-value=0,281
	Bragança	4,20	0,531	5	3		
HGEO	Coimbra	4,31	0,631	5	2	<i>t</i> -value =0,176 p-value=0,86	P-value=0,999
	Bragança	4,34	0,725	5	2		

The results reported in Table 4 suggest that overall the differences between the two samples are not statistically significant. The strongest result pointing to the existence of differences regards the Student's *t*-test for HMAX (different means) but the Kolmogorov-Smirnov test suggests that the distributions of

HMAX in both samples are not statistically different.

In terms of question 3 (variable QUEST3), as expected, when faced with two situations involving equal total health gain, all respondents choose the scenario ensuring geographic equality of treatment (Figure 19). Probably, those that are neutral towards the criterion of geographic distribution of resources are indifferent between the two scenarios, however, once forced to choose between them, they select the second option.

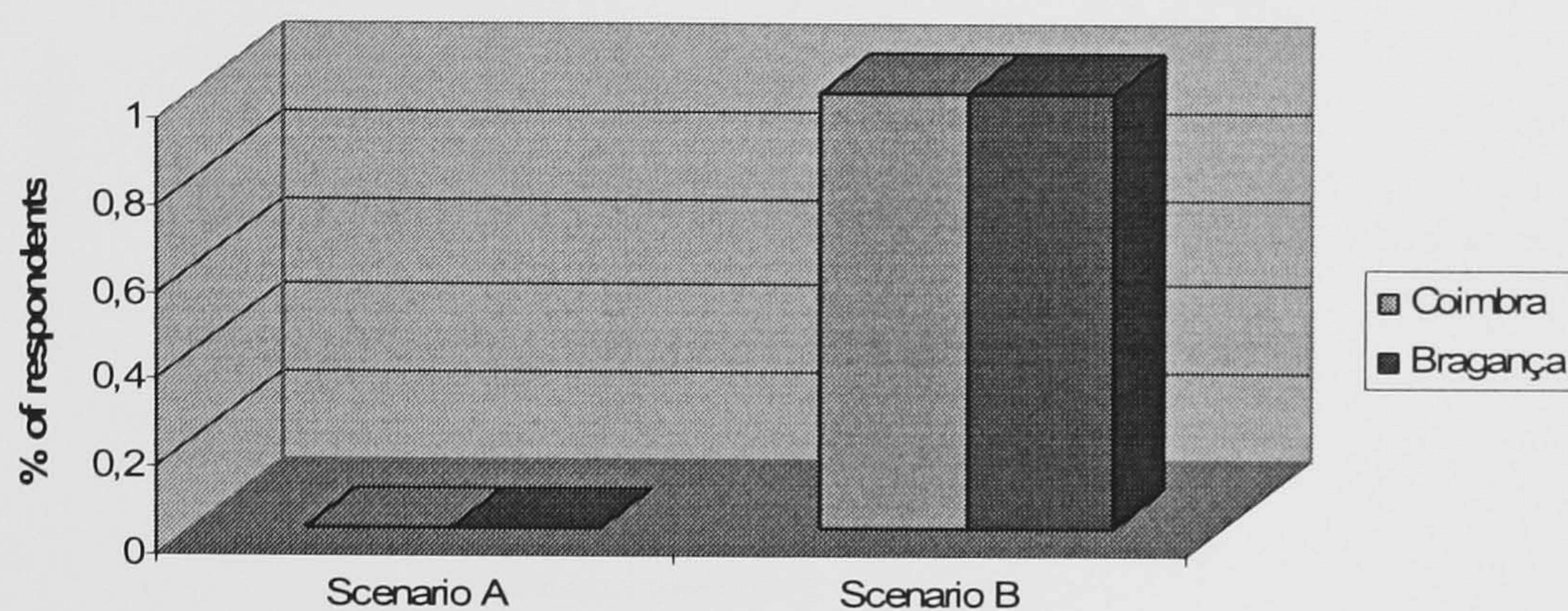


Figure 19. Bar chart of QUEST3 by region of residence

For the case of question 4 (variable QUEST4), a few respondents select Alternative A (see figure 20) given that geographic equality of treatment (Alternative B) is obtained here at the cost of 1% of total health gain foregone. Still, there is some inconsistency in responses. That is, following our previous argumentation, one would expect that those selecting Alternative A, in question 4, coincide with those manifesting a lower support to the criteria of geographic distribution of resources, in question 2. However, this does not happen.

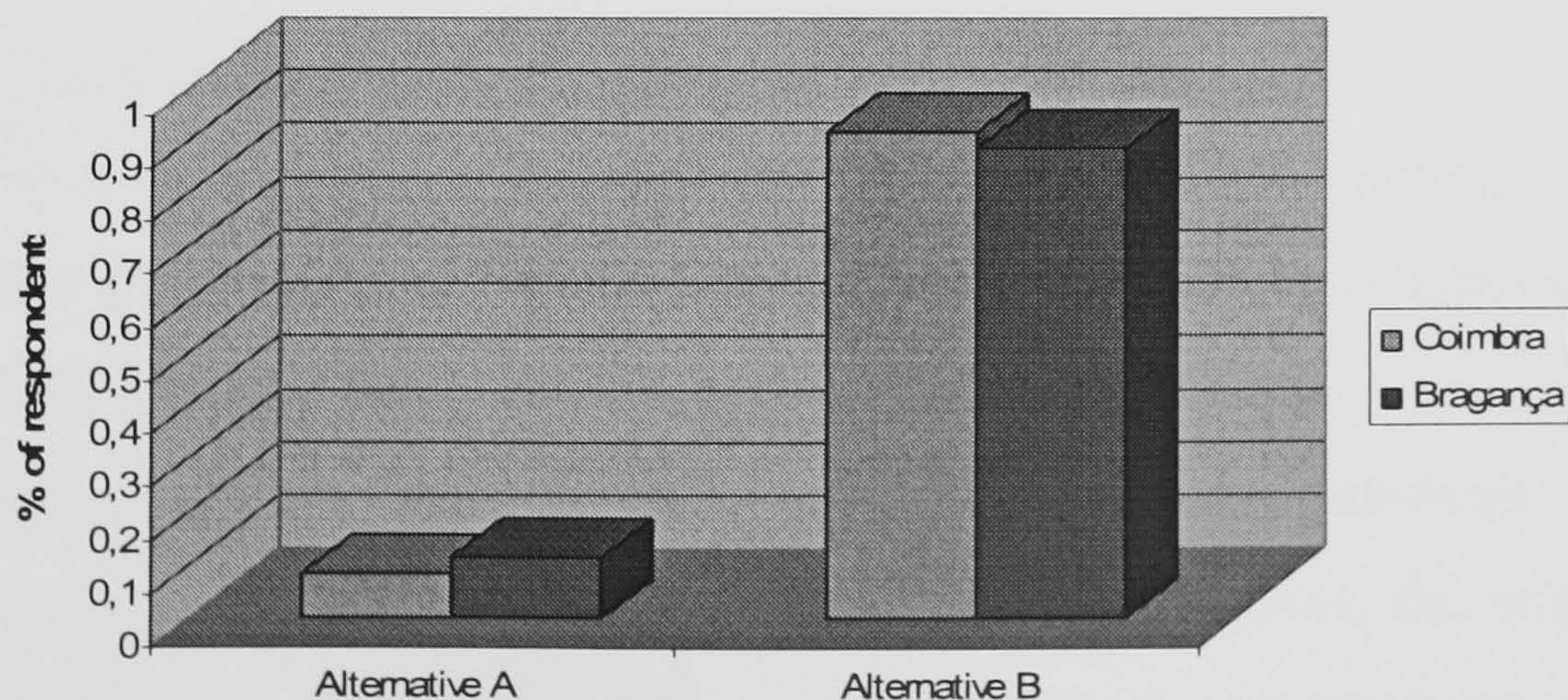


Figure 20. Bar chart of QUEST4 by region of residence

Regarding the trade-off between health maximisation and geographic equality of treatment, for the case of question 3 (TOFF3), Table 5 shows the location of the median respondent. In both samples, the median corresponds to TOFF3=1 (which in turn corresponds to 10% of total health gain foregone – see sub-section 4.3.1). Half of the respondents is either not making any trade-off between total health gain and geographic equality of treatment or is willing to forego a maximum of 10% of the total health gain to keep equality of treatment.

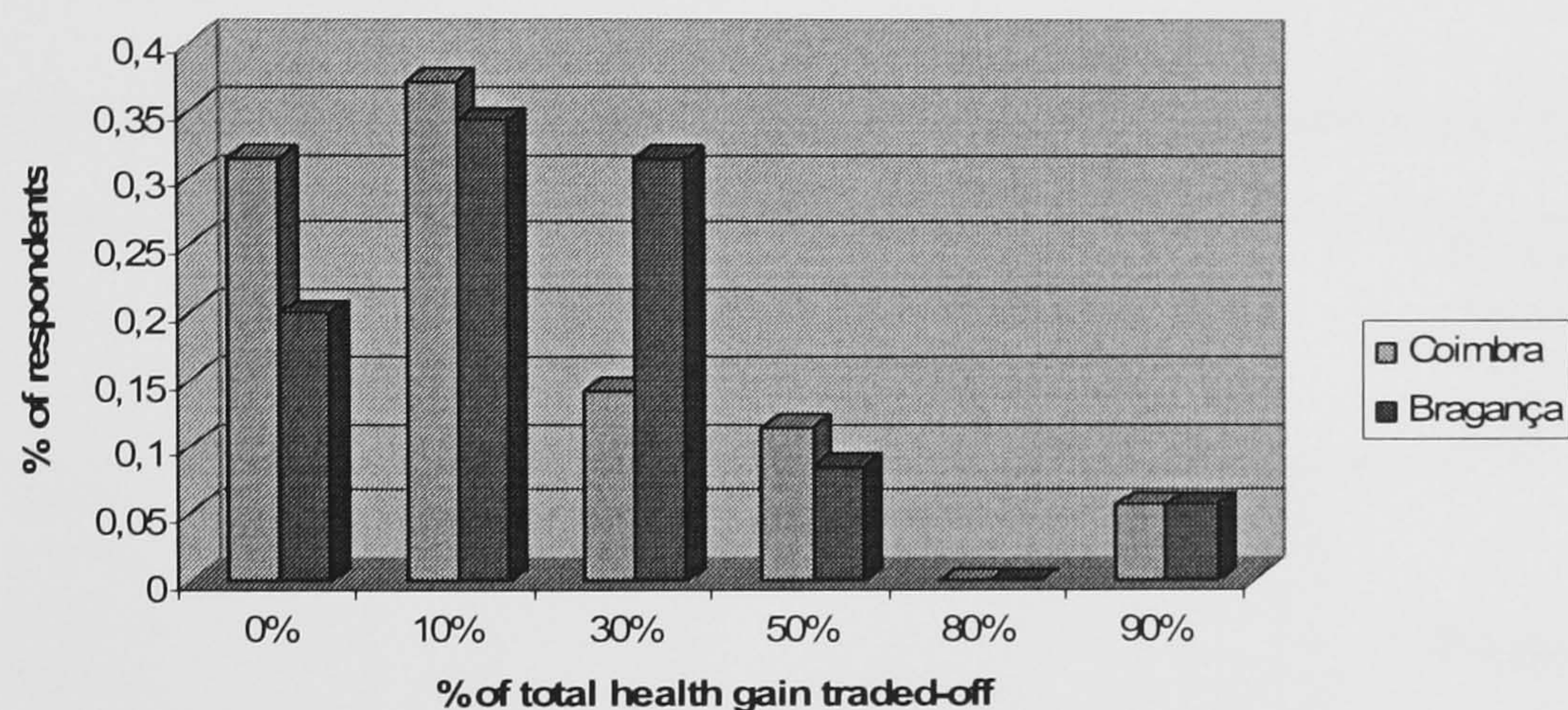


Figure 21. Bar chart of TOFF3 by region of residence

In Figure 21, we can further see that the mode corresponds to TOFF3=1 (10% of total health gain foregone) in both samples. There is also a greater percentage of respondents, in Coimbra, not making any trade-off, than in Bragança. On the other hand, there is a greater percentage of respondents, in Bragança, willing to forego 30% of total health gain compared to Coimbra. In both cases there are respondents (about 15%) willing to forego 50% of total health gain and in both samples two respondents always choose Scenario B (we suspect that the question was misunderstood; that is, given that no one disagrees with health maximisation, it is surprising to obtain this latter result).

Despite the differences suggested by Figure 21, the Kolmogorov-Smirnov test does not allow us to reject the null hypothesis of equal distributions of the variable TOFF3 (Table 5).

Concerning the variable TOFF4, because some respondents selected Alternative A, in question 4 (see bar chart of QUEST4, above), this means that TOFF4 ranges from -1 to 5. Consequently, TOFF4=-1 represents no trade-off between criteria. If TOFF4=0, which applies to respondents that choose Alternative B in the first part of question 4 but then in the second part change to



Alternative A right in the first hypothesis, then respondents are willing to forego a maximum of 1% of total health gain, to keep equality of treatment. In Table 5, one can see that the median respondent, in Coimbra, falls in the category TOFF4=1 (10% of health gain foregone) and the median respondent in Bragança falls in the category TOFF4=2 (25% of health gain foregone). Although seemingly different, the medians of both samples are not statistically different, as indicated by the Wilcoxon-Mann-Whitney test (Table 5).

**Table 5. TOFF3 and TOFF4 by region of residence: summary statistics and hypothesis test results**

Var.		Coimbra	Bragança	Wilcoxon-Mann-Whitney test H <sub>0</sub> :equal medians	Kolmogorov-Smirnov test H <sub>0</sub> : equal distributions
TOFF3	0(0%)	11	7		P-value=0,967
	1(10%)	13	12		
	2(30%)	5	11		
	3(50%)	4	3		
	4(80%)	0	0		
	5(90%)	2	2		
TOFF4	-1(0%)	3	4	P-value= 0,322848 (approximate value)	P-value=0,640
	0(1%)	4	2		
	1(10%)	11	5		
	2(25%)	9	11		
	3(55%)	2	8		
	4(76%)	1	0		
	5(91%)	5	5		

Median respondent is in shaded cell;

Percentages in brackets correspond to total health gain foregone

Comparing the distribution of TOFF4 for both samples (Figure 22), the mode varies from one sample to the other (in Coimbra the highest frequency is registered for TOFF4=1 and in Bragança the highest frequency occurs for TOFF4=2). There is in addition a greater percentage of respondents, in Bragança, willing to sacrifice 55% of total health gain in order to keep equality of treatment than in Coimbra. As for the case of TOFF3, the Kolmogorov-Smirnov test concludes that distributions of TOFF4 in both regions are not statistically different.

In TOFF4 there are also some respondents that always choose Alternative B, meaning that they are prepared to forego 91% of total health gain to obtain in return equality of treatment. Previously, we said that this kind of result was inconsistent with answers to question 1. The puzzle here is that not only there are more respondents choosing in this way in TOFF4 than in TOFF3 but also they do

not coincide with those respondents that did the same in TOFF3. It looks as if respondents do understand in fact what they are doing. Still, we should consider the possibility of these ‘outliers’ being generated by a different process (misunderstanding of the questions or they might even be ‘protest’ votes who refuse to make any trade-off with geographical equality). Thus, we checked whether it would make a difference to the results if these outliers were excluded from the analysis. However, the medians are the same as in table 5 and the Kolmogorov-Smirnov test still suggests that distributions in both regions are not statistically different (the K-S test generated a P-value equal to 0,957 and a P-value equal to 0,537, for TOFF3 and TOFF4, respectively).

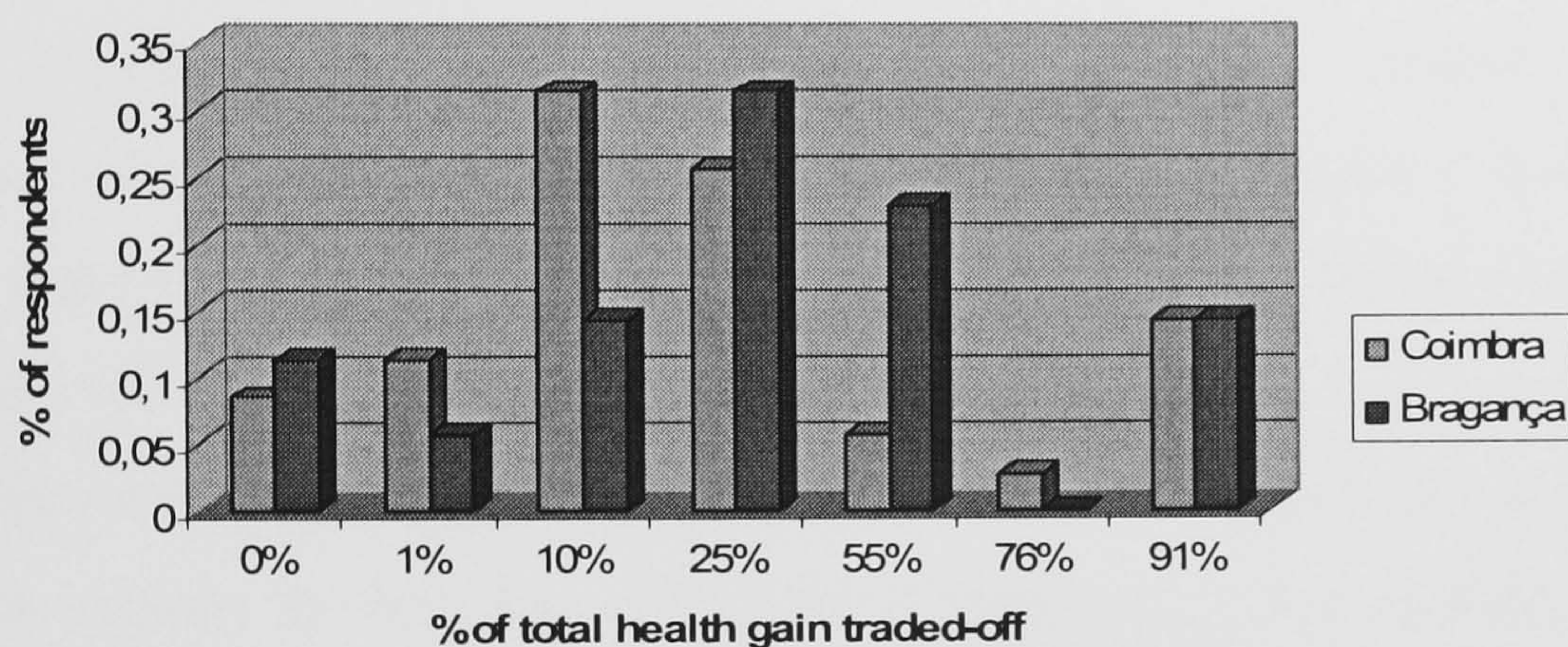


Figure 22. Bar chart of TOFF4 by region of residence

Table 6 shows the value of the parameter  $r$  and the implied trade-off for both questions 3 and 4. This parameter is indeterminate for those who never switch between alternatives (corresponding to TOFF3=5 and TOFF4=5), which also implies that no trade-off is made, hence the empty cell in column 4. A possibility would be to assume, as Dolan *et al.* (2002), that these respondents are concerned only with equality and therefore to consider that  $r$  approaches infinity (but as the authors stress, one cannot be sure).

In the case of TOFF4, we calculated  $r$  taking into account only two regions (the total health gain under Alternative A was therefore adjusted to  $\frac{2}{3} \times 10000$ ).

Also, we assumed that the respondents that chose Alternative A in the first part of question 4 (TOFF4=-1) would have chosen Alternative B if the two alternatives presented the same total health gain being this gain equally distributed between

regions in the case of Alternative B (as it happened in question 3).

**Table 6. Implied  $r$  and implied trade-off for TOFF3 and TOFF4**

Var.		Implied $r$	Implied trade-off
TOFF3	0	-0,91	-2,5
	1	-0,67	-10
	2	-0,42	-20
	3	-0,18	-32,5
	4	-0,05	-42,5
	5	Indet.	
TOFF4	-1	-0,99	-16,5
	0	-0,9	-183
	1	-0,71	-583
	2	-0,42	-1333
	3	-0,18	-2183
	4	-0,06	-2783
	5	Indet.	

In question 3, the median respondent in both samples is indifferent between a health gain of 100 children concentrated in *Region 1* and a health gain of 80 children equally divided by the two regions. In question 4 and for case of sample 1 (Coimbra), the median respondent is indifferent between a total of 6666 surgeries concentrated in LTV and 5500 surgeries equally divided by LTV and another region; in the case of sample 2 (Bragança), the median respondent is indifferent between a total of 6666 surgeries concentrated in LTV and 4000 surgeries equally divided by LTV and another region.

In question 3, the implied trade-off for the median respondent in both samples is -10 children and in question 4 it is -583 surgeries for sample 1 and -1333 surgeries for sample 2.

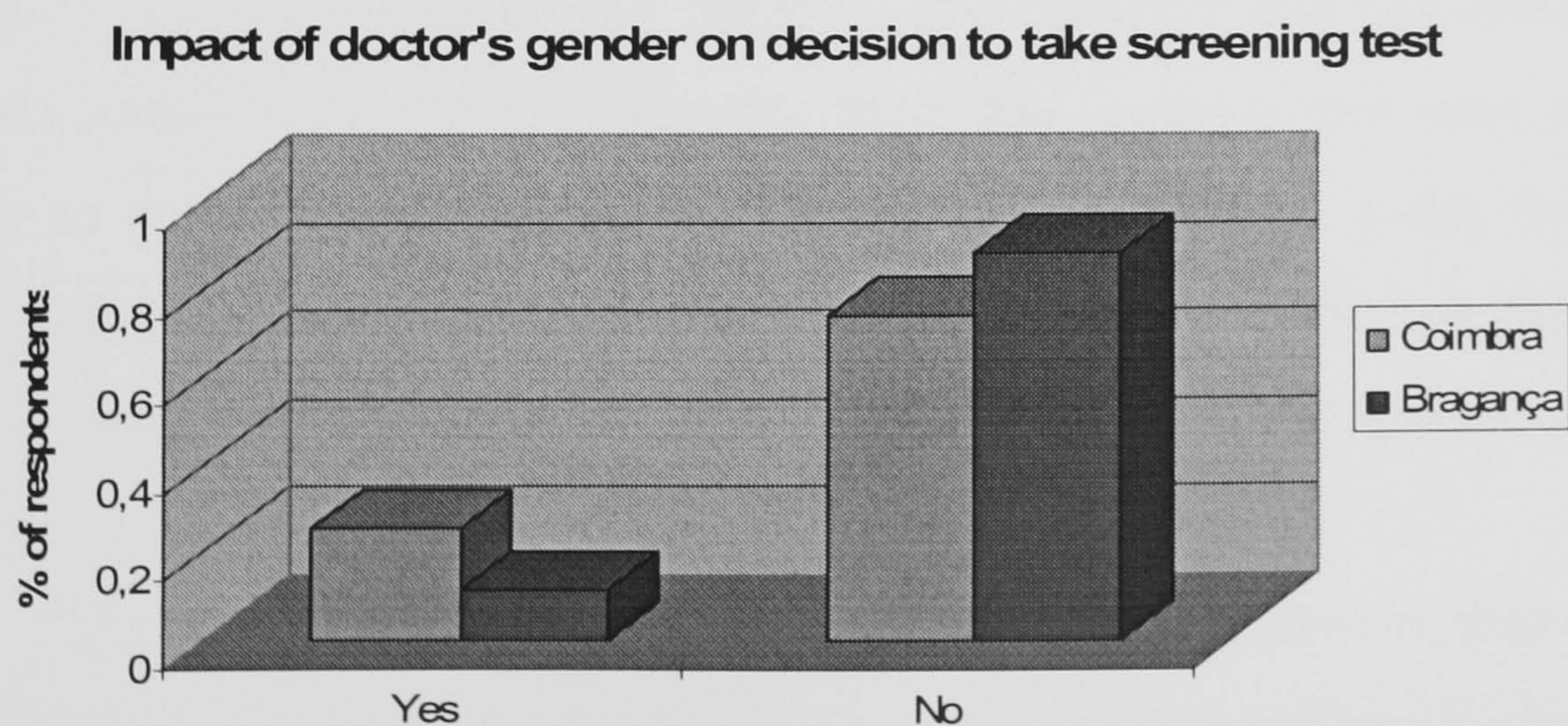
Regarding the parameter  $r$ , the results shown in Table 6 indicate that there is aversion to inequality in the distribution of total health gain between regions (as  $r > -1$ ) though the values are relatively low. Note that the parameter  $r$  (see section 4.3.1) depends not only on the trade-offs made (the numerator in the formula presented in section 4.3.1) but also on the level of inequality (considered in the denominator). In our questionnaire, one of the alternatives produces zero health gain for one or some regions, which implies a high level of inequality under this alternative. Thus, the impact of large trade-offs is mitigated by large levels of inequality tolerated.

**Table 7. Implied  $r$  and implied trade-off by region of residence: mean values and hypothesis test results**

		Region	Mean	Student's $t$ -test $H_0$ : equal means
Implied $r$	TOFF3	Coimbra	-0,653	t-value =0,907 p-value=0,368
		Bragança	-0,601	
	TOFF4	Coimbra	-0,619	t-value =1,29 p-value=0,20
		Bragança	-0,530	
Implied trade-off	TOFF3	Coimbra	-11,7	t-value =0,654 p-value=0,52
		Bragança	-13,3	
	TOFF4	Coimbra	-878	t-value =1,34 p-value=0,19
		Bragança	-1129	

With the parameterisation of the trade-offs it was possible to calculate mean values for both samples and to use the Student's  $t$ -test. This test is more sensitive than the Kolmogorov-Smirnov test; nonetheless the results (shown in Table 7) suggest that differences between samples are still not significant.

When it comes to the variable DOCGEND, Figure 23 shows that most respondents, in both samples, say that the possibility of choosing the doctor's gender does not affect their decision to take a screening test. Eventually, this might be a sign of a reasonable public awareness regarding the kind of disease in question (colon cancer); as pointed in sub-section 4.2.1, gastro-intestinal tumours are the most frequent type of cancer in both men and women, in Portugal. There is a slight difference in responses between samples, with more respondents in Coimbra saying 'yes' than in Bragança but, as expected (based on Figure 23), the Fisher exact test rejects the existence of a significant association between answers to this question and region of residence (Table 8).

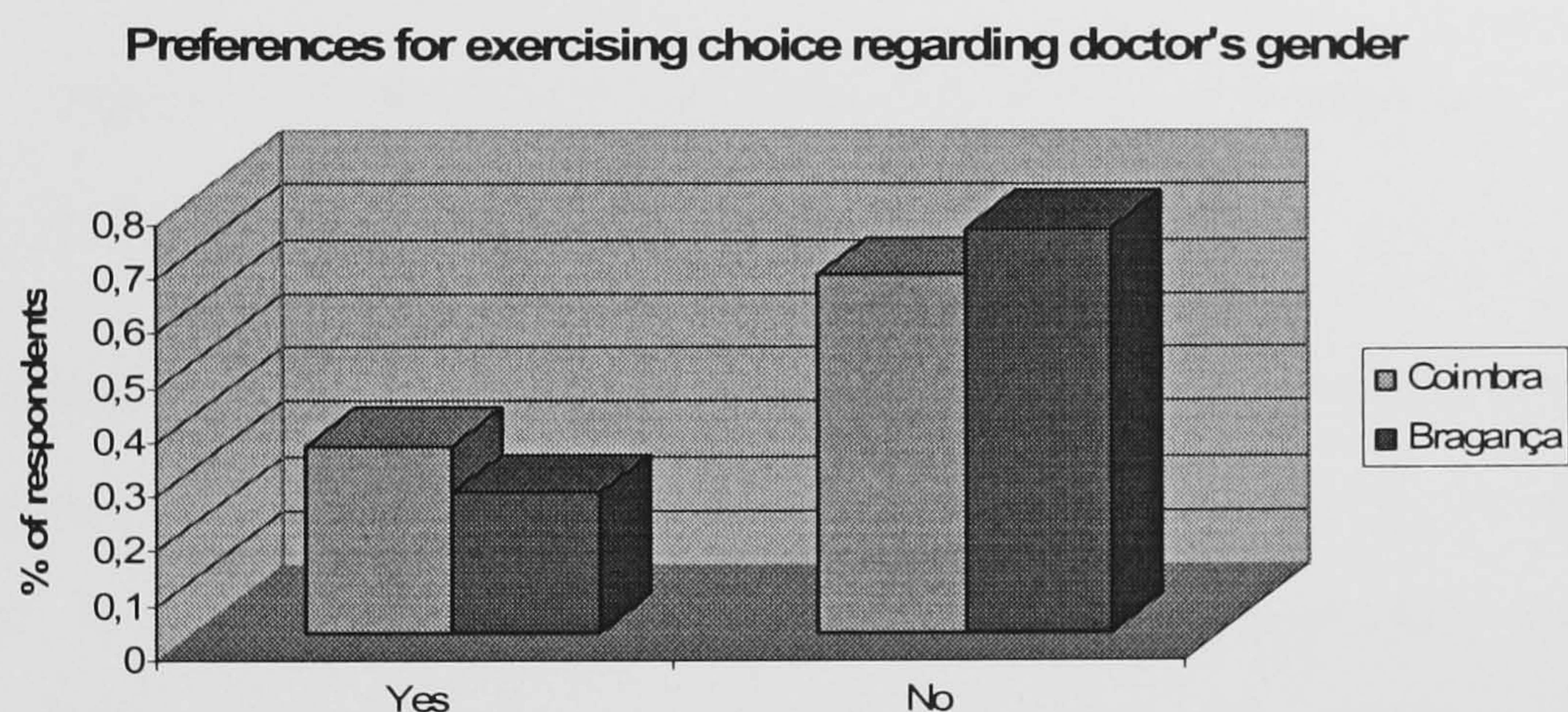


**Figure 23. Bar chart of DOCGEND by region of residence**

**Table 8. Fisher exact test for DOCGEND and CHOICE**

Variable	Fisher exact test
	H <sub>0</sub> : random assignment of cases
DOCGEND	p-value=0,218
CHOICE	p-value=0,602

Concerning the variable CHOICE, here too most respondents in both samples say that, once in the health centre, they would be indifferent regarding the gender of the doctor performing the test (Figure 24). Again, there is a slightly greater percentage of respondents in Coimbra saying that would exercise choice, compared to Bragança but the Fisher exact test rejects any association between CHOICE and residence (Table 8). Comparing the distributions of DOCGEND and CHOICE, the percentage of respondents answering 'yes' is higher for the latter variable, in both samples, contrarily to what we have conjectured in sub-section 4.3.1.

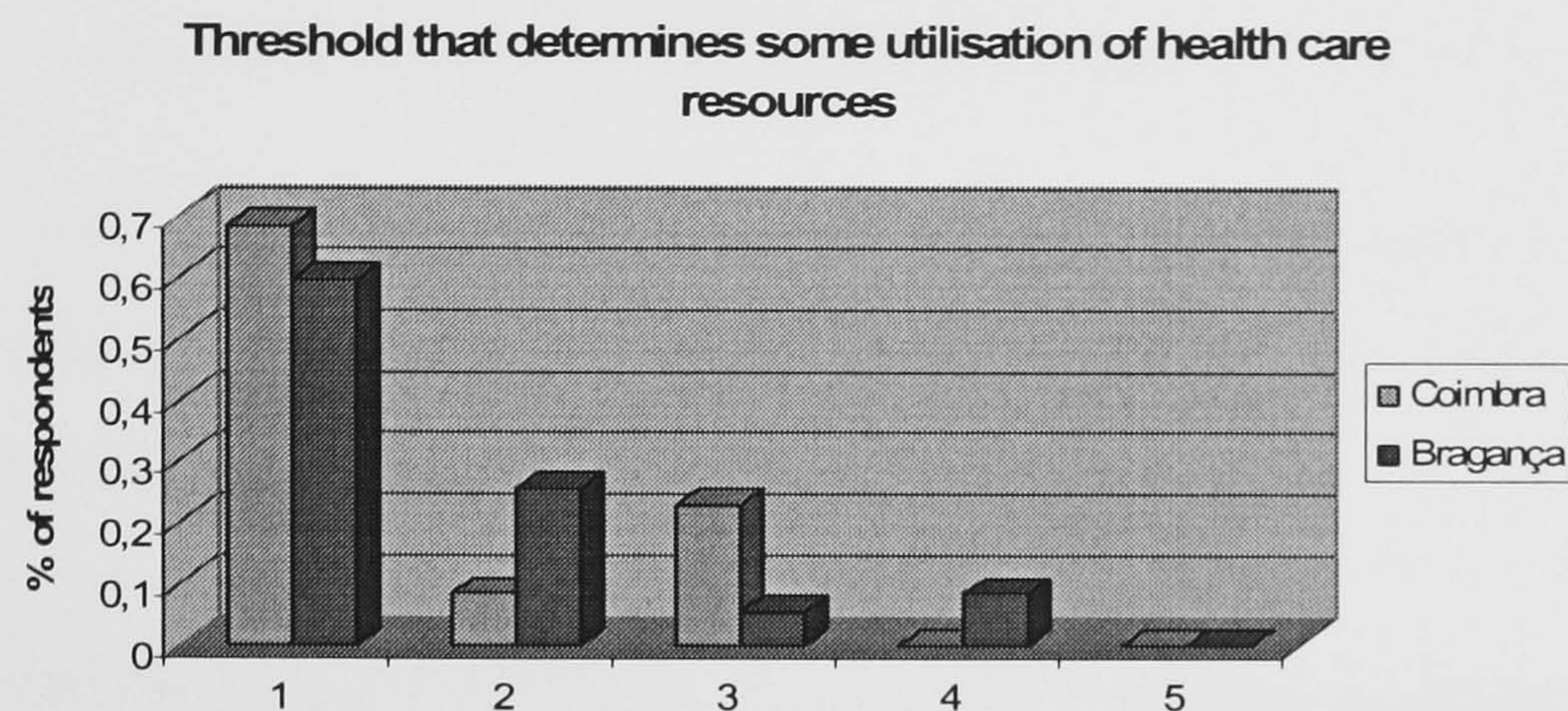


**Figure 24. Bar chart of CHOICE by region of residence**

In terms of the variable THRESH61 the median respondent falls in category 1, in both samples (see Table 9). This category also represents the mode in both samples, as seen in Figure 25. What this means is that most respondents prefer to utilise some form of health care resources right in the first situation hypothesised in question 6 (see appendix B). The Kolmogorov-Smirnov test, whose result is shown in Table 9, suggests that the distributions for the two regions, in Figure 25, are not statistically different.

The variable THRESH62 corresponds to the situation in which a visit to the doctor is mentioned for the first time as the preferred form of action. In both regions, the second hypothetical situation is the threshold below which individuals do not seek personalised medical care for more respondents (in Figure 26, one can

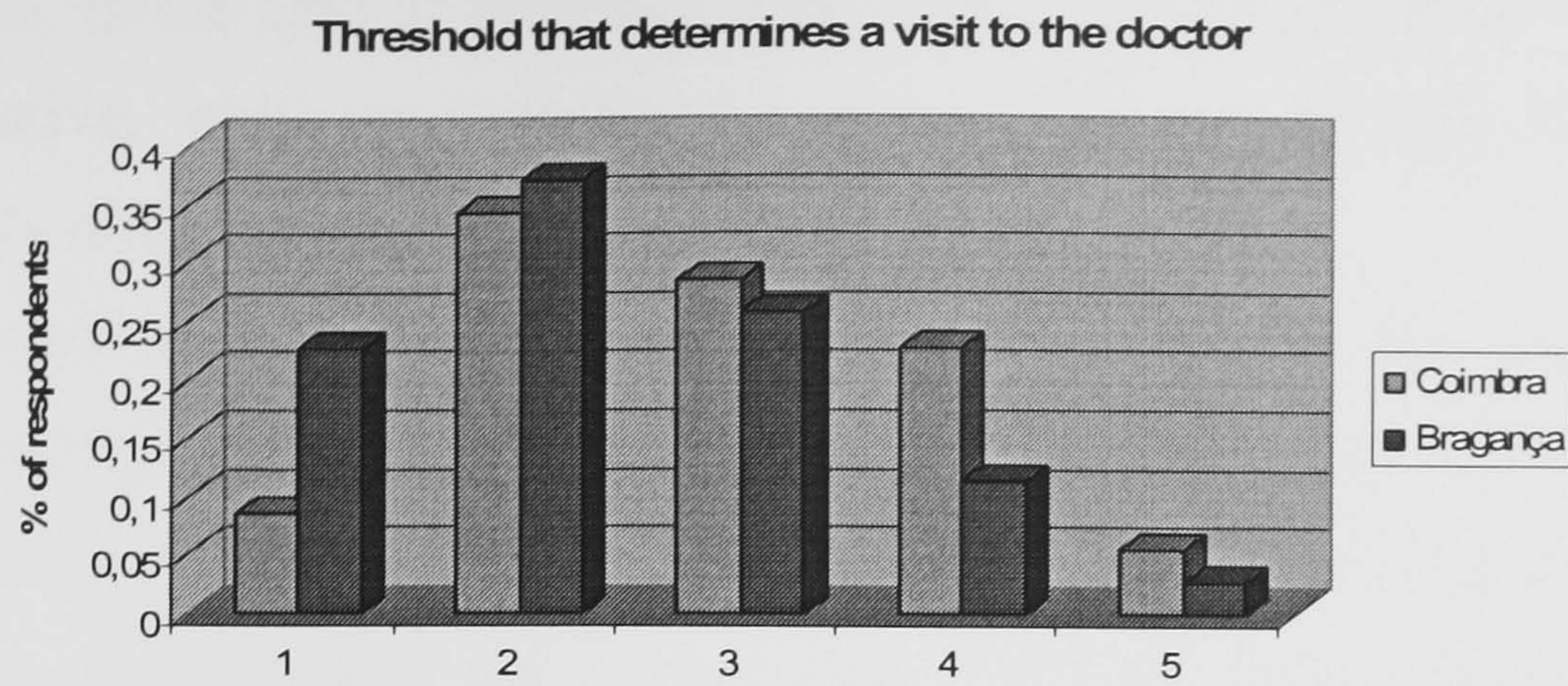
easily verify that the mode corresponds to THRESH62=2). The medians (Table 9) are 2 and 3 for Bragança and Coimbra, respectively. There is a noticeable difference between samples, regarding the category 1, which corresponds to the THRESH62 for about 20% of respondents in Bragança, while in Coimbra this figure is about 8%. These results suggest that people in Bragança prefer to seek personalised medical care sooner (in terms of levels of severity of illness) than in Coimbra<sup>35</sup>.



**Figure 25. Bar chart of THRESH61 by region of residence**

Statistical tests regarding THRESH62 (Table 9) suggest nonetheless that medians are not statistically different (at least at the 5% level) as well as the distributions. We should also note that we did not obtain any inconsistent response to this question, that is, after choosing once the option ‘I visit the doctor’, all respondents chose this option in the remainder situations.

<sup>35</sup> In Bragança, the median respondent falls in the category ‘feeling strong pain frequently’, while in Coimbra the median respondent falls in the category ‘feeling pain frequently and having sleeping difficulties’. So, the threshold corresponding to the median respondent is lower in Bragança than in Coimbra. This result is somehow unexpected; up to a dozen years ago, Bragança was relatively isolated and the primary sector still played (and plays) an important role in the region’s economy. Thus, the dominant idea was that people’s lives were harsh under these conditions. One would expect that, given this background, individuals would be more prepared and willing to endure difficult situations, including illness. Naturally, although our interest lies on preferences, it is not possible to control for a variety of factors that could affect responses. For example, it seems reasonable to expect that a greater availability of doctors leads to higher demand and hence affect expressed preferences for this kind of health care. However, our results again do not confirm this expectation given that it is in Coimbra that there is a greater concentration of resources, compared to Bragança (in 2003, the number of physicians per 1000 residents was 20,5 in the former compared to only 2,7 in the latter – cf. Portuguese National Statistics Office <http://www.ine.pt>).



**Figure 26. Bar chart of THRESH62 by region of residence**

A final set of statistical tests compares the distributions of different variables within each sample; in other words, we test for within-respondent differences, in each sample.

**Table 9. THRESH61 and THRESH62 by region of residence: summary statistics and hypothesis test results**

Variable	Coimbra	Bragança	Wilcoxon-Mann-Whitney test H <sub>0</sub> : equal medians	Kolmogorov-Smirnov test H <sub>0</sub> : equal distributions
THRESH61	1	24		P-value=0,999
	2	3		
	3	8		
	4	0		
	5	0		
THRESH62	1	3	P-value=0.089034	P-value=0,640
	2	12		
	3	10		
	4	8		
	5	2		

Median respondent is in shaded cell

We use the Wilcoxon matched-pairs signed-ranks test to test firstly whether or not respondents make different trade-offs in questions 3 and 4, though we should say that questions are not framed in the same way. Thus, the interpretation of this test is for this reason limited. Still, the results (Table 10) suggest that answers are statistically different for the case of Bragança while this does not happen for the case of Coimbra.

Further note that there are movements in both directions (there are positive and negative differences), that is, there are respondents making lower trade-offs in question 4 compared to question 3 (positive difference) but there are also people making higher trade-offs in question 4 compared to question 3 (negative

difference). In the case of Bragança, positive differences are given more weight than negative, while in Coimbra the opposite is true (though they are not statistically different).

People might make higher trade-offs in question 4 than in question 3 (corresponding to a negative difference between TOFF3 and TOFF4) because now choosing the option that generates greater total health gain benefits residents in a region other than that where respondents live and implies a sacrifice born by residents in their own regions of residence. On the other hand, people might make lower trade-offs in question 4 than in question 3 (corresponding to a positive difference between TOFF3 and TOFF4) because they are more aware of opportunity costs (in terms of health gain foregone) of geographic equality of treatment. This latter suggestion is based on the idea that reducing the level of abstraction of questions makes people more conscious of the consequences associated to each decision.

**Table 10. Tests for within-respondent differences in each sample**

Variables	Region	Wilcoxon matched-pairs signed-ranks test H <sub>0</sub> : equal medians and distributions
TOFF3 vs TOFF4	Coimbra	W+=82; W-=194; N=23; p-value $\leq 0,0914$
	Bragança	W+=224; W-=133,5; N=18; p-value $\leq 0,0342$
THRESH61 vs THRESH62	Coimbra	W+=0; W-=435; N=29; p-value $\leq 0,000$
	Bragança	W+=0; W-=171; N=18; p-value $\leq 0,000$

Finally, we use the Wilcoxon matched-pairs signed-ranks test to test whether THRESH61 is statistically different from THRESH62 or not. Here there is no limitation in the interpretation of the test given that these two variables are derived from the same question of the questionnaire. In this case, positive differences are null by definition. That is, THRESH61 corresponds to the first situation involving some form of action. Because visiting the doctor (THRESH62) is already an alternative form of action, then, THRESH62 is in the minimum equal to THRESH61. When respondents prefer to start with a form of action other than visiting the doctor, then THRESH61 is lower than THRESH62 and the



difference between the two will be negative.

The results in Table 10 suggest that THRESH62 is statistically different from THRESH61 in both samples. There are however some differences, which in fact have already been spotted to some extent in Figure 26. While in Bragança 17 out of 35 respondents prefer to start by visiting the doctor, in Coimbra, only 6 out of 35 respondents express this preference (in Table 10, N is the number of respondents that answer differently in the two questions). Also, the gap between taking some form of action and visiting the doctor is larger in Coimbra than in Bragança.

#### **4.5. Discussion**

In the study reported here, two independent samples drawn from two distinct (Portuguese) municipalities were given the same questionnaire to allow us to test the hypothesis of variation in preferences across regions.

We elicited preferences regarding the criteria of health maximisation and geographic equality of treatment. To measure the trade-off between these two principles, we asked respondents to choose between two hypothetical alternative programmes: one involving a fixed health gain concentrated in one region; and the other involving successively lower total health gain, equally distributed between two regions. Although the latter alternative is characterised by an equal distribution of health gain, when introducing questions, we have focused on equal treatment across regions. For simplicity and given that questionnaires were to be self-administered, we have considered that equal treatment would result in equal health gain. Trade-offs were elicited in two questions; one involving hypothetical regions; and the other making use of actual Portuguese regions.

Respondents were also asked about personal preferences. In one case, respondents were asked if the possibility of choosing the gender of the doctor performing a screening test, for the case of colon cancer, would affect their decision to take the test. In the other case, we addressed the issue of preferences for different forms of health care, depending on the perceived severity of illness.

Concerning the main objective of this study, the statistical tests showed

that basically for all variables the null hypothesis, that there are no differences in preferences across samples, cannot be rejected. This is true for the Student's *t*-test, Wilcoxon-Mann Whitney and Kolmogorov-Smirnov tests, meaning that means, medians and distributions of variables in both samples are not statistically different. The Fisher exact test also shows that personal preferences for two attributes of health care delivery do not vary between samples. The strongest result concerns the variable HMAX, for which the Student *t*-test suggests the rejection of the null hypothesis at the 5% level. In this case, the support for the criterion of health maximisation is stronger in Coimbra than in Bragança but mean values are actually closer to each other (4,5 and 4,2, respectively). Regarding the threshold indicating a visit to the doctor, our results produced different medians but the Wilcoxon-Mann Whitney test suggests that the null hypothesis can be rejected only at the 10% level.

The null hypothesis of absence of diversity could not be rejected in both cases of social and personal preferences. Thus, it seems that the nature of preferences is not decisive for the existence or not of diversity in preferences. Also, some changes in responses were registered when respondents were led to think about actual Portuguese regions instead of hypothetical regions (the example used - reducing waiting lists for hip replacement or cataract surgery - is very close to recent measures adopted by the Ministry of Health, in Portugal, as well); median values actually became different between samples but the Wilcoxon-Mann Whitney test suggests that the null hypothesis cannot be rejected in the same. The difference between mean values of *r* is also greater for TOFF4 than for TOFF3 but again the Student's *t*-test suggests that the null hypothesis cannot be rejected in both situations.

Thus, the main conclusion is that differences between samples are not sharp and globally they are not statistically significant; our results do not consequently corroborate the hypothesis of variation in preferences across regions. Accepting these results means that the theoretical discussion about the impact of decentralisation on allocative efficiency should be revisited and greater attention should be paid to the role that the assumption of diversity in preferences has played.

In terms of the results obtained for preferences weighing health maximisation against geographic equality of treatment, respondents agree or

strongly agree that resource allocation should take into account both principles. Given a fixed total health gain, all respondents in both samples choose the alternative that equally distributes this health gain between two regions instead of choosing the alternative that concentrates this health gain in one of the two regions.

Once respondents are asked to trade-off health maximisation and equality of treatment (equal distribution of health gain), between 70 and 80 per cent of them are prepared to make trade-offs between the two criteria. Of those making trade-offs, most people are willing to forego between 10 and 30 per cent of total health gain to keep geographic equality of treatment in return. For sample 1 (Coimbra), we obtained a median willingness to sacrifice health maximisation on behalf of geographic equality of treatment at 10% of total health gain in both questions 3 and 4. For the case of sample 2 (Bragança), we found a median willingness to sacrifice health maximisation on behalf of geographic equality of treatment at 10% of total health gain, in question 3 and at 25% of total health gain, in question 4.

The results further show that, in question 3, the median respondent is willing to give up of 10 children of the average number of children avoiding the disease in the two regions if the total number were distributed equally between them, corresponding to an  $r$  equal to -0,67; in question 4, the median respondent in sample 1 is willing to give up of 583 people of the average number of people having a surgery in LTV and another region if the total number were equally distributed between them (corresponding to an  $r$  equal to -0,71), while in sample 2 this figure raises to 1333 people (corresponding to an  $r$  equal to -0,42).

Taking into account both samples and questions, the mean value of  $r$  ranges between -0,53 and -0,65 suggesting the existence of aversion to inequality in the geographic distribution of total health gain.

Despite that the majority of respondents is willing to make trade-offs, there are some of them not willing to sacrifice one principle on behalf of the other. That is, some are not prepared to give up of any amount of health gain (not even 1%) to keep equality of treatment, while there are other respondents not prepared to give up of equality of treatment even if this is obtained at a high cost in terms of health gain foregone. The former group ranges between 8 (TOFF4) and 30 (TOFF3) per cent of respondents, in Coimbra, while in Bragança it ranges between 11 (TOFF4)

and 20 (TOFF3) per cent of respondents. The groups of respondents not willing to give up of equality of treatment are less expressive, ranging between 6 (TOFF 3) and 14 (TOFF4) per cent of respondents, in both samples.

The results suggest that the geographical dimension of (in)equality in treatment matters to people and that a maximum opportunity cost of equality, in terms of health gain foregone, is likely to exist. Hence, the empirical findings reported here suggest that analyses of the sort developed in the previous chapter, in which equality of treatment is traded-off against other criteria are pertinent and might be operationalised.

To conclude this chapter, one must say that the main conclusion of the current empirical work, i.e. that there is absence of diversity in preferences across geographic areas, applies for the case of the specific regions and for the particular topics approached in this study. Additionally, this study has its own limitations like the method of administration of the questionnaire: it has been acknowledged that, in the case of PTO questions, respondents often require greater explanation of the task than can be offered through a written survey instrument (Green, 2001). Still, given the dimension and exploratory nature of our study, a delivery and collection method of administration seemed acceptable. Plus, when designing the questionnaire we have taken into consideration the method of administration, hence the introduction and explanation of themes in each question. Moreover, at the time of collection some respondents mentioned their difficulty in completing the questionnaire not because they felt difficulties in understanding the questions; instead they felt that questions were difficult in themselves but, as Green (2001) comments about the PTO, it may be that difficult choices are necessary to establish true preferences. In future works, however, the use of interviews or discussion groups should be considered. Plus, with these latter methods of administration, other principles and trade-offs could be introduced in the questionnaire without representing too many additional difficulties. Questions could also be refined in order to make more explicit the difference between geographic equality of health gain and geographic equality of treatment.

The sizes of samples might also be considered a limitation of the study. therefore, in the future, samples should be larger (generating more powerful tests) and should include some fringes of the population that were excluded from our study. We must nonetheless acknowledge that this will be a challenge considering

that we are talking about people with no degree of education or at best people with 9 years of schooling in the maximum. In addition, these individuals are on average older than others better educated. Thus, more inclusive studies should be designed taking into account the specific characteristics of the respondents.

Beyond these limitations, several other issues remain to future research. For the particular context of Portugal, other regions should be included in analyses, namely southern regions. Also, it might be interesting, from the perspective of testing for differences between geographically defined samples, to group regions in terms of mainly rural and mainly urban areas. In addition, future research might look into the impact of the socio-demographic composition of samples on expressed preferences in order to test whether eventual differences between samples stem from their different compositions (Dolan and Tsuchiya (2007), for example, use binary logistic regressions to test whether age, gender, education and insurance status influence willingness to trade-off health gain against equality of health).

Moreover, the trade-off between total health gain and geographic equality of treatment might be defined in terms of total health gain at the local level *versus* equal availability of services across regions. Another trade-off that seems to be worth exploring is that between the allocation of resources at the local level based on local preferences *versus* uniform provision throughout the country. Actually, in chapter two it is said that there has been a tendency to see variations in provision as an undesirable outcome of decentralisation but in our study, although respondents value equality between regions up to the point of being willing to give up of some health gain, the vast majority does not support this equality at all costs.

Addressing methodological issues was not a particular objective of the current work; nonetheless, the results generated by the Wilcoxon matched-pairs signed-ranks test for TOFF3 *versus* TOFF4 suggest that the level of abstraction of (PTO and other) questions might be a relevant aspect to consider in future research about methodologies.

## **CHAPTER 5. A discussion of ambiguities and confusions affecting the debate about decentralisation in the English NHS**

### **5.1. Introduction**

The literature and the debate about decentralisation have not been always clear and unequivocal about the meaning of concepts like preferences. In particular, we saw in chapter two that, in the context of health care resource allocation, the concept of preferences might reach beyond its narrowest definition of consumers' preferences. It is therefore important to point where these ambiguities arise and further clarifications are required.

In chapters two and three, we have distinguished between information regarding preferences and information regarding technical matters. Because variations in both forms of information might lead to variations in health care provision, there might be some confusion between the two sources of variations in provision. It is consequently pertinent to clarify this issue, which can be done making use of the framework proposed in chapter three (section 3.2).

It has already been noted in this dissertation that participation or voice might be inherently valued, directly contributing to social welfare (procedural approach) and that it might also be seen as a means of matching outcomes to the preferences of the population (consequentialist approach). In practice, references to both approaches have been used by authorities as well as by the public. Some ambiguities might however arise when, for instance, the promotion of participation as an instrument to bring public values into decision making is not accompanied by greater local discretion to accommodate such values.

The objective of the current chapter is therefore to identify some ambiguities and confusions in the context of decentralisation in health care decision making in the English National Health Service (NHS). The English context was chosen as the reference for the current discussion because decentralisation in health care is one of the current policy themes in England, reflected in the political discourse of 'shifting the balance of power'. Among other features, this reform contemplates a form of decentralisation of expenditures:

allocations were made directly to locally based Primary Care Trusts (PCTs). for the first time in 2003/04. Discussions about participation in decision making are also in line with recent changes in the English NHS, in which a 'patient and public involvement' movement was initiated and stimulated by the central government itself.

In what follows, the discussion applies at some points to interpretations that might simply represent personal opinions of some authors. However, because our objective is to identify ambiguities and confusions in the debate, the message of those that might in fact act as opinion makers is important too. Moreover, the whole chapter is intended to be illustrative rather than exhaustive; suggestive rather than conclusive.

We start by introducing, in section 5.2, the English NHS. This section is relevant to understand the geographical organisation of the NHS, including Primary Care Trusts, as well as the role of institutions like NICE and recent tendencies in health care policy. Because our discussion is centred on issues that might be related in one way or another with the debate about decentralisation, section 5.3 gives a summarised account of decentralisation in the NHS. Section 5.4 then stresses the ambiguity that has characterised the use of the concept of preferences in the context of the English NHS. In section 5.5, the main argument is that variations in service provision in England have been dealt with mainly as variations in technical information, whereas they can derive not only but also from variations in health care priorities. In section 5.6, the discussion focuses on ambiguities associated with public participation in decision making. Finally, section 5.7 presents some suggestions and directions to future analyses.

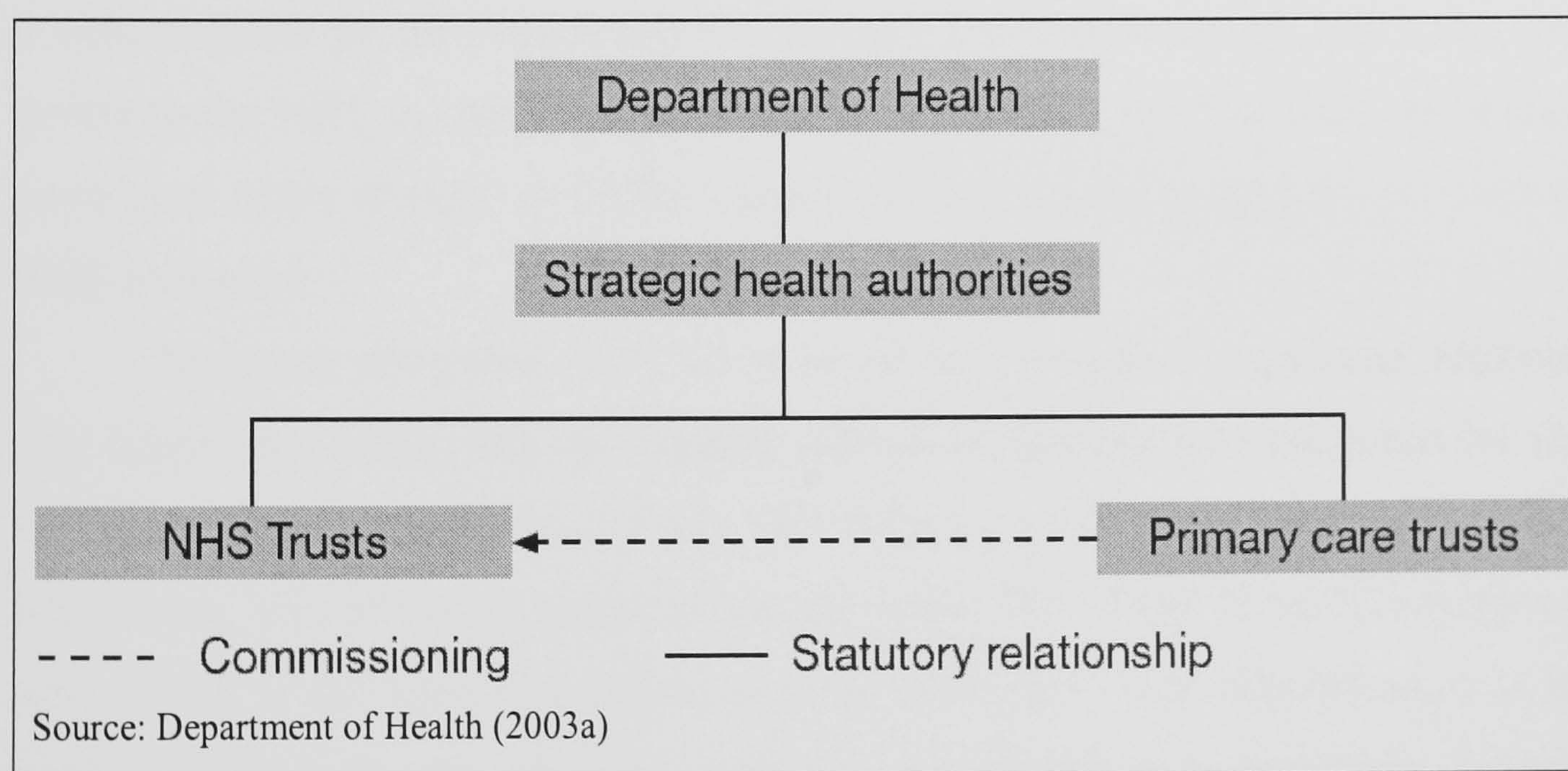
## **5.2. Brief characterisation of the NHS**

The objectives of the NHS, created in 1948, are characterised by the general principles of universality and equality of access. The NHS is funded out of general taxation and services are free at point of delivery (with some exceptions related to dental care, prescriptions and ophthalmology). The NHS expenditure, in England, is about 6 per cent of GDP (figure for 2003/04).

The NHS is one of the biggest employers in England. General practitioners are independent workers and have a contractual link with the NHS. They act as gatekeepers controlling therefore the access to secondary care.

The NHS has been the target of successive reforms and is in continuous evolution. As Greener and Powell (2003, p. 43) put it “the health arena has undergone reform after reform through the 1980s and 1990s, from the introduction of general management, to resource management and clinical audit, to the introduction of internal markets (...) to evidence-based medicine and the restructuring of the internal market by New Labour in 1997. Since then, the pace has, if anything, picked up further with the introduction of the performance assessment framework, the reform of regional health authorities, the introduction of the new bodies to ‘modernize’ health care (...)”.

There have been some changes in the NHS structure. The main feature of change has been giving locally based Primary Care Trusts the role of running the NHS in their areas. All existing health authorities were consequently abolished and new ones were created.



**Figure 27. The NHS in England**

27 gives a general perspective of the organisation of the NHS. The Department of Health (DoH) is the government department responsible for delivering health and social care services to the English population. Its functions include, among others, securing management and accountability of the overall health and social care system and the overall regulation and inspection of the NHS.



Eight previous Regional Offices were abolished while four new Directors of Health and Social Care have been appointed. The latter work directly with the NHS and performance manage strategic health authorities (StHA).

100 former health authorities were substituted by 28 strategic health authorities and then (on July 2006) these were in turn replaced by 10 strategic health authorities (the *Statutory Instrument* 2006 No. 1408 gives an account of the strategic health authorities abolished and sets out which areas are covered by each of the 10 new authorities). StHAs are presented by the Department of Health as a key link between the Department and the NHS. They are expected to ensure that national priorities are integrated into plans for the local health service.

Primary Care Trusts are run by GPs, nurses, other health and social care professionals, and representatives of patients and the community. PCTs provide some services themselves and are responsible for commissioning other services (a role previously carried out by the former health authorities)<sup>36</sup>. Commissioning might include services provided by hospitals, dentists, mental health care, Walk-In Centres, NHS Direct, patient transport (including accident and emergency), etc. Primary Care Trusts have been presented as the cornerstone of the NHS as well as a central piece to the Government's strategy for decentralising and devolving power in the NHS to local communities. In 2003/04, for the first time allocations have been made directly to PCTs, controlling now about control 80 per cent of NHS resources.

Resource allocation to PCTs is based on a weighted capitation formula. The basic idea underlying the current allocation formula was proposed by the Resource Allocation Working Party (RAWP), in 1976, in which the weighted population, WP, of a local authority was given by:  $WP = POP * (1+a) * (1+n) * (1+c)$ , where POP is the local population,  $a$  is the authority's age adjustment,  $n$  is its needs adjustment (over and above age) and  $c$  its relative cost adjustment. Several revisions occurred in the meantime, in particular in terms of the needs adjustment factor, but the core idea remains the same. The current formula (see e.g. Department of Health, 2003b) is the result of a process coordinated by the Advisory Committee on Resource Allocation, set up in 1997. It has been firstly used for the 2003/04 to 2005/06 PCTs' allocations round.

---

<sup>36</sup> PCTs are expected to increasingly become commissioning-led organisations with their role in provision reduced to a minimum (Department of Health, 2005).

Special health authorities have been set up to provide a national service to the NHS or to the public. They are independent, but can be subject to ministerial direction like other NHS bodies. Some examples of these special health authorities are: National Blood Service, Health Development Agency, National Institute for Clinical Excellence (NICE)<sup>37</sup>, among others.

NICE is a special health authority for England and Wales, launched on 1st April 1999. An important reason behind the creation of NICE was the highly criticised (particularly by the press) geographic variations in service provision - the so-called 'postcode prescribing' or 'postcode lottery'<sup>38</sup>. Precisely for its role in terms of equal provision of health care (see section 5.5), it seems opportune to provide a bit more of information about NICE.

NICE produces guidance in three areas of health: health technologies (guidance on the use of new and existing medicines, treatments and procedures within the NHS); clinical practice (guidance on the appropriate treatment and care of people with specific diseases and conditions within the NHS); and the recently extended area of public health. The Commission for Health Improvement is responsible for monitoring the implementation of NICE recommendations. The Health Select Committee (NICE, 2003, paragraph n) also recommends that strategic health authorities include the implementation of NICE guidance as part of their regular monitoring of PCTs and acute Trusts.

An important attribution of NICE is technology appraisals. These are recommendations on the use of new and existing technologies (e.g. medicines, medical devices, diagnostic techniques, surgical procedures and health promotion activities) within the NHS. Recommendations issued by NICE might be that the technology is a cost-effective use of NHS resources in general, or for specific conditions or defined subgroups of patients. Recommendations might also be to reject the technology for the NHS or demand additional research before a final decision is made. The status of NICE guidance is defined as advice which should be fully taken into account by clinicians and NHS organisations, though it does not override the individual responsibility of health professionals to make decisions appropriate to the circumstances of the individual patient. In January 2002 the

---

<sup>37</sup> On 1st April 2005, NICE joined with the Health Development Agency to become the new National Institute for Health and Clinical Excellence (still known as NICE).

<sup>38</sup> The term 'postcode lottery' was used to represent situations in which, depending on where people lived, some drugs or treatments would be provided, or not, within the NHS.

Government announced a statutory obligation for the NHS in England to provide funding for treatments and drugs recommended by NICE as a part of its technology appraisals work programme.

Hospital Trusts are found in most large towns and cities, and usually offer a general range of services. Some Trusts also act as regional or national centres of expertise for more specialised care, while some are attached to universities and help to train health professionals. Trusts can also provide services in the community, for example, through health centres, clinics or in people's homes. Except in the case of emergencies, hospital treatment is arranged through GPs (via the so-called referrals). Together, NHS Trusts employ the majority of the NHS workforce.

Care Trusts are NHS bodies, which work in both health and social care. They can be established where NHS organisations and local authorities agree to work together and their actual functions are determined by this partnership. The role for the Care Trust is usually where it is felt that closer integration between health and social care is needed or would be beneficial at a local level. At the moment there are only a small number of Care Trusts in development.

NHS Direct opened in March 1998 and offers free 24-hour advice about personal health care. NHS Direct nurses aim to provide callers with the advice and reassurance they need to care for themselves at home, or, if they need further help, to direct them quickly to the right service.

Regarding some targets defined by the Government, cancer, heart disease and stroke, accidents and suicide constitute the four priority areas for intervention identified in Department of Health (1999). The targets set for 2010 are: (i) to reduce deaths from cancers at least a fifth in people under the age of 75 years; (ii) to reduce the death rate from heart disease, stroke and related illnesses by at least two fifths in people under the age of 75 years; (iii) to reduce the death from accidents by at least a fifth and the rate of serious injury at least a tenth and (iv) to reduce deaths by suicide and undetermined injury by at least a fifth.

In February 2001, the Secretary of State for Health at the time, Alan Milburn, announced two health inequality targets: (i) "starting with children under one year, by 2010 to reduce by at least 10% the gap in mortality between manual groups and the population as a whole" and (ii) "starting with Health Authorities,

by 2010 to reduce by at least 10% the gap between the quintile of areas with the lowest life expectancy at birth and the population as a whole” (Department of Health, 2001a). Regarding waiting times, by 2008 there should be maximum waiting times of three months for outpatients and inpatients (Budget Report 2003)

### **5.3. A summarised account of decentralisation in the NHS: from Conservatives (1979-1997) to Labour<sup>39</sup>**

The rhetoric of decentralisation does not always correspond to effective devolution of decision making power. Regarding the NHS, in particular, Peckham *et al.* (2005) argue that many British governments have claimed that they wish to decentralise the NHS; very few claimed to want to centralise the NHS. However, based on their review of literature, these authors note that the existing accounts of decentralisation in the NHS are unclear and that some of the conclusions are conflicting.

According to Peckham *et al.* (2005), it appears that the NHS moved in two different directions at once, during the Conservative period of office. For example, they say that most commentators agree that the move from Regional Health Authorities to Regional Offices of the NHS Executive was centralist as regional staff became classified as national ‘civil servants’ rather than ‘local’ NHS personnel. Also, the system of performance reviews designed to monitor progress towards very specific targets was associated with a tighter system of control. Movements that have been classified as decentralist are, for example, devolution of actual purchasing budgets and of local pay.

Although opposite movements towards more centralisation or more decentralisation have been associated with different initiatives or reforms, it is common to find, in Peckham *et al.*’s (2005) report, divergences among authors regarding their interpretation of any single reform, in terms of decentralisation *versus* centralisation.

The philosophy of the new public management (introduced into the NHS

---

<sup>39</sup> Section 5.3 is mainly based on Peckham *et al.* (2005).

in the 1980s and greatly motivated by the *Griffiths* Report) is seen as decentralising. General Managers are meant to be responsive to consumers and in principle should be allowed autonomy to achieve established objectives. Equally, the idea of the purchaser/provider split (suggested by the White Paper *Working for Patients*, released in 1989 by the Department of Health), with decentralised institutions of self-governing NHS Trusts and General Practitioner Fund Holders, is identified as decentralist. However, centrally determined targets and performance management have increased centralisation and created lines of command between the executive and the staff on the field. Exworthy (1994) argues that although decentralisation is associated with managerialism, local managers manage within closely defined central terms. Paton (1993, 1998), for example, talks about decentralisation of pay bargaining and centralisation of objectives in the NHS. This author argues that in some cases responsibilities were devolved but not power and that despite the operational decentralisation, strategic control was kept at the central level.

The opposite movements towards more centralisation, or, inversely, towards more decentralisation, have been reported for the period that began in 1997 as well, when Labour came to power: “while the rhetoric of decentralization is frequently invoked, the practices in areas like health, education and social services are frequently centralizing” (Ross and Tomaney, 2001, p.267); “while the centre is committed to the rhetoric of devolved implementation, it seems frequently unwilling to trust the front line with discretionary powers or has overlaid it with potentially contradictory initiatives of its own” (Gray and Jenkins, 2001, p.216).

Peckham *et al.* (2005) conclude that, on the whole, the period between 1997 and 2000 is attributed centralising tendencies, justified by the need to correct either organisational failures or health inequalities. There was strong central pressure to reduce waiting times (a Labour manifesto commitment) and the *clinical governance* strategy, seeking the improvement of quality of care and reduction of variations in services led to the creation of centralising organisations such as NICE and CHI (Commission for Health Improvement, seconded by the Commission for Healthcare Audit and Inspection, later renamed the Healthcare Commission - HC). Centralising are also the National Service Frameworks (NSFs), launched in April 1998 within the spirit of *clinical governance*, with the

objectives of setting national standards and defining service models for a particular service or care group.

Regarding the NHS Plan, in 2000, Ham (2004) classifies it as having centralising tendencies. Although allowing some potential for greater autonomy for high performing organisations (with performance assessment still depending on the centre), the NHS Plan represents, according to this author, a new delivery model for the NHS, putting in place arrangements for the inspection and performance measurement of health organisations that are ‘strong centralising’.

The policy document *Shifting the Balance of Power*, released in 2001 (Department of Health, 2001b), represents a movement towards a more local approach to decision making. Primary Care Trusts are presented as the cornerstone of the NHS and it is announced the devolution of about 80% of the NHS budget to PCTs, from 2004 onwards. In the case of good performance, there exists the opportunity for an NHS Trust to become a Foundation Trust, earning more autonomy by, for instance, retaining revenues from land sales and determining their own investment plans. The spirit continues to be one of decentralising decisions over means of meeting health targets set by the centre.

Oliver (2005, p. S79) says that in encouraging greater local autonomy through PCTs and Foundation Trusts, and yet introducing more extensive national regulation through NICE, the NSFs and the Healthcare Commission, the Labour Government has developed ‘seemingly contradictory policy instruments though it is also possible that the Healthcare Commission will ultimately help to generalise transferable local innovation and good practice’.

These policy reforms (PCTs, on one hand; NICE, NSFs and HC, on the other) need not to be seen necessarily as contradictory policy instruments; they might represent decentralisation of some decisions and centralisation of other decisions. Peckham *et al.* (2005) classify the former movements as input decentralisation and the latter as output centralisation (see ‘arrows framework’ in section 2.2 – chapter two). Though, we must say that NICE’s recommendations, for example, can also be viewed as input centralisation (as they affect resource allocation to different health care interventions or groups of patients).

#### 5.4. The ambiguous use of the concept of preferences in the context of the NHS

In chapter two (section 2.5) we saw that preferences might be classified mainly taking into account ‘whom respondents are asked to think about’ as well as ‘what respondents are asked to think about’. In broad terms, the former classification distinguishes above all between consumers’ preferences and citizens’ preferences. Consumers’ preferences can further be elicited from actual patients or from non-patients. Concerning ‘what respondents are asked to think about’, different things might be at stake, namely health states, the criteria that ought to be used to guide resource allocation, priorities among health care services, and attributes of health care delivery.

Both consumer and citizen perspectives might be associated with the NHS but the distinction between the two has not been always clear. Callaghan and Wistow (2006) say that the co-existence of those two approaches is a source of ambiguity about the status of those involved. The use of the concept of preferences in practice is actually closely related to the role accorded to public involvement in health care decision making<sup>40</sup>.

Rowe and Shepherd (2002, p.276) note that, prior to the 1990s, the pattern and delivery of health services were largely determined by the medical profession, “with the ‘public interest’ reduced to the sum of clinical judgements”. With the New Public Management, the emphasis became on making services more responsive to the public, more precisely, to consumers’ preferences. Milewa (2004) argues that the involvement agenda has clearly its origins in the work of Conservative administrations. Initiatives such as *Working for Patients*, *Local Voices* and *Patient’s Charter* (White Papers released in the late 1980s and early 1990s) led way to the institutionalisation of mechanisms to involve the public<sup>41</sup>. However, at this stage, although the importance of listening to local communities was advocated, in practice the focus was primarily on developing methods of

---

<sup>40</sup> While the literature about decentralisation uses mainly the term ‘participation’ and the literature about procedural justice uses mainly the term ‘voice’ (as noted in chapter two), at the policy level, the term most commonly used to represent participation is ‘involvement’.

<sup>41</sup> Prior to the New Public Management, Community Health Councils (which came to be abolished in 2003) were founded in 1974, being classified as bodies to represent the interests of the public. Baggott (2005) however says that although these bodies were able to contest major changes, they were not systematically consulted on matters of strategic planning, service design or delivery.

consultation about (consumer) satisfaction with existing services (Milewa, 2004).

The initial philosophy of partnership and responsiveness thus focused above all on patients and their involvement on decisions affecting their own health. Nonetheless, acknowledging the diversity of roles that individuals may play, ranging from patients to carers or members of the public, and that these roles can produce divergent perspectives, namely between the perspectives of patients and non-patients, the Kennedy Inquiry into paediatric heart surgery in Bristol (*Bristol Royal Infirmary Inquiry* 2001) urged the inclusion of the wider public, besides patients, in the decision making process. This can be seen as an attempt to expand the concept of preferences though it is still not clear the distinction between consumers' (non-patients) and citizens' preferences.

The Health and Social Care Act 2001 (section 11) brought some changes at a more formal level, by placing a statutory duty (commencing on 1 January 2003) on NHS Trusts, Primary Care Trusts and strategic health authorities to make arrangements to involve and consult patients and the public in service planning and operation, and in the development of proposals for changes. To assist NHS organisations in meeting this legal duty, the Commission for Patient and Public Involvement in Health was established in 2003 as an independent, non-departmental public body.

New advisory groups, Patient and Public Involvement Forums (PPI Forums), were established in 2003 to parallel the now over 570 NHS primary and secondary health care Trusts. Forums are independent of the bodies to which they relate and they will be funded, supported and performance managed by the Commission for Patient and Public Involvement in Health. PPI Forums comprise patients, carers and members of voluntary organisations that represent the interests of patients or carers (over 5000 people are currently members of existing Forums). According to the leaflet that publicised PPI Forums, their objectives include helping the public to feed its views about health services into the NHS and to make recommendations to NHS managers about changes to services. PPI Forums are also advertised as one way to influence and change local health services. It is nonetheless unclear if these views or recommendations are to reflect what individuals prefer for them or for others and it is not also clear what is meant by e.g. 'changes to services'. That is, are these changes mainly about attributes of health care delivery of the type reviewed in chapter two? Or, are these changes



about more fundamental things such as the elimination of services and inclusion of new ones?

In the literature about decentralisation, participation is associated with decision making at the local level. In England, PPI Forums are in fact associated with local organisations. But input might occur at a more central level, being the NICE Citizens Council, set up in 2001, an example of this situation. The Citizens Council is a form of citizens' jury, with its 30 members drawn from different age groups, social circumstances, ethnic backgrounds and abilities. The Council considers topics relating to social, ethical or moral questions which arise in the Institute's work and that are referred to it by the Institute's Board. When NICE Citizens Council was created, the Chief Executive of NICE expressed his expectations towards the Council, saying: "We expect this Council to provide NICE with advice that reflects the public's perspective on what are often challenging moral and ethical issues" (NICE, 2001). Here, the implicit concept of preferences seems closer to the definition of social preferences (compared to personal preferences). As we noted in chapter two, in some cases the 'whom' might depend on the 'what', that is, thinking about e.g. the role that age should play in resource allocation (NICE, 2005) forces people to think about others (the young; the elderly). Therefore, the concept of social preferences associated with the Citizens Council might itself derive from the topics under deliberation. In chapter two, we further noted that the views of the NICE Citizens Council can be thought as an example of collective preferences. Thus, another distinction that applies to the English NHS is that between individual *versus* collective preferences.

Starting with a situation in which the medical profession was predominant, there has been an evolution that represented in a first phase the consideration of patients' (consumers') views regarding health services; later, the public in general, not necessarily patients, was also brought into the discourse at the policy level, though this emphasis in itself does not seem sufficient to clarify the status of the public in terms of consumers *versus* citizens. Nonetheless, as the views of the public are sought in regards to 'moral and ethical issues' (e.g. NICE, 2005), one might say that individuals are led to think more as citizens and less as consumers. On the contrary, the focus on patient choice fits better a consumer perspective. Baggott (2005), for instance, says that in the NHS it appears that 'choice rather

than voice is being prioritised as a mechanism of securing more responsive services'. 'Choice' here means that patients should be able to visit primary care centres and, where they need additional treatment, choose from a list of service providers and book care at the location and time of their choosing. These are essentially attributes of health care delivery; again the 'what' might influence the 'whom'. But this continues to be an implicit rather than explicit issue in policy documents.

Thus, in the future it would be useful if some guidance were provided in terms of the perspective that should be adopted when addressing public views. Additionally, we already referred that what respondents should be asked to think about has not always been clear as well. The role attributed to PCTs is another example of this ambiguity. In the Department of Health (<http://www.doh.gov.uk/>), PCTs are presented as local organisations, being therefore best positioned to 'understand the needs of their community'. In Department of Health (2005), we can find the expressions 'greater variety of service offerings and responsiveness to patient needs' and 'wider variety of local services and models of provision in response to patient needs'. There is here a strong parallelism with the literature about decentralisation (in which decentralisation is associated with diversity in provision according to heterogeneous local preferences). But what is meant by 'needs of their community' is open to question.

### **5.5. Variations in service provision: the confusion between variations in preferences and variations in technical information**

Variations in health care provision, across geographic areas, have been at the heart of discussions about decentralisation in health care because it has been pointed simultaneously as an advantage and a disadvantage of decentralisation (e.g. Saltman and Bankauskaite, 2004; Levaggi and Smith, 2005 – see chapter two for more details). In England, geographic variation in health care provision has also been at the centre of attention and it has led, at least in part, to the creation of the National Institute for Clinical Excellence (see section 5.2 above). Variation in provision might be considered *per se* desirable or not but given that there seems to

exist some confusion regarding its origin, it is important to clarify this matter and distinguish between the two sources of variation. That is, health care provision might vary across regions because local populations have different priorities in health care, or, it might vary because information about technical issues varies among local decision-makers.

This distinction might be explained using the framework proposed in section 3.2 (Figure 2). When decision making is transferred from the central authority to local authorities, this might give rise to differences in health care provision between local populations because local authorities are better informed about local preferences and in the case of these preferences varying between populations, the provision of health care will accordingly vary. This corresponds to the perspective adopted in chapter three. We also considered that technical information possessed by authorities might vary between the centre, on one hand, and localities, on the other hand. But technical information might vary between local authorities as well. Consequently, when transferring decision making from the central authority to local authorities, this might give rise to differences in health care provision due to differences in technical knowledge possessed by local authorities. How variations are viewed might well depend on their origin. For instance, people might accept variations stemming from differences in preferences, while variations stemming from differences in technical knowledge are less tolerated. Or, they might simply care about variations irrespective of their sources. But it is pertinent to be clear about this issue and to point the confusion that there seems to exist regarding variations in the NHS.

The principle of equal access has been given extreme importance in the context of the English NHS, which is reflected namely in the 1997 Labour manifesto ('access to NHS will be based on need and need alone'), cf. Oliver (2005). This author argues that securing equal access for equal need remains the overriding objective of the system that ought not to be sacrificed in order to meet other 'secondary concerns' such as improving health outcomes and patient satisfaction, reducing waiting times and health inequalities. Sassi *et al.* (2001), in turn, give two examples (cervical cancer screening and sickle cell disease screening) in which equal provision of opportunities to receive care have been pursued at a high price, in terms of small health gains.

Because health care is funded out of general taxation and is, with very few exceptions, free at the point of delivery, ability to pay is eliminated as a barrier to equal provision. There is also a geographic dimension to resource allocation in England, with more than 75% of NHS resources being allocated directly from the Department of Health to PCTs. At this level, equal access for equal need is to be achieved through the weighted capitation formula used in resource allocation to PCTs, so that PCTs in greater need get proportionately more funds (see section 5.2). However, as noted by e.g. Smith (2003) and Oliver (2005), even assuming that the formula in question adequately captures variation in needs, it can only secure the opportunity for equal provision for equal need; it cannot by itself to secure equal provision for equal need.

Additionally, and as noted above, NICE (see section 5.2) is associated with equal provision of health care across regions as it was created partly to put an end on the 'postcode lottery'. Rawlins and Culyer (2004) note that NICE recommendations are intended to apply across the whole England (and Wales), regardless of where people live or work. They further argue that local variations in cost ought not to result in variations in availability of health care. NSFs should also lead to more consistent standards of treatment. Thus, the role of NSFs and NICE might be seen as ensuring that the equal availability of resources for equal need is translated into equal availability of specific technologies and standards of treatment. Equal provision of health care across geographic areas should therefore result from the combination of the needs-based resource allocation system (which already includes a relative cost adjustment) with *clinical governance*.

Although equal availability of technologies and standards of treatment should derive from NICE guidance, there is a difference between funding decisions and decisions about clinical and cost-effectiveness, as stressed by NICE (NICE, 2003, paragraph y). Rawlins and Culyer (2004) too reiterate that NICE does not take affordability when making judgments about cost effectiveness. They say that this is a matter for the government when deciding the annual budget for the NHS; the authors add that it is NICE's job to judge whether something ought to be purchased from within the resources made available to the NHS. The question is that, when the implementation of guidance requires additional resources, this in itself is already affecting health care priorities, which is not in essence a technical matter. In Figure 6 (chapter three), cost-increasing

recommendations are illustrated by the movement from  $m_0$  to  $m_1$ .

In the case of the English NHS, NICE and NSF recommendations have been cost-increasing (Oliver, 2005) and NICE has said ‘yes’ more often than it has said ‘no’ (Devlin and Parkin, 2004). What might be happening here is that the elimination of variations in technical knowledge across localities is simultaneously eliminating variations due to different funding priorities. The Health Select Committee acknowledges this situation, stating that “in making the implementation of NICE Health Technology Appraisals mandatory in a healthcare system which operates within fixed budgets, there is the potential to give the provision of certain, NICE-approved treatments priority over other” (NICE, 2003, paragraph o). Others have drawn attention on this issue in the early years of NICE (e.g. Hutton and Maynard, 2000; Sculpher *et al.*, 2001; Birch and Gafni, 2002). These suspicions are confirmed to some extent by Davies and Littlejohns’ (2002) research. These authors, based on a questionnaire returned by 92 Directors of Public Health about their views of NICE’s technology appraisals, report that most respondents (76%) agreed that NICE was succeeding in making the NHS set aside resources locally to provide NICE approved technologies, however, nearly all (95%) agreed that this was causing difficulty locally for the implementation of ‘non-NICE’ technology.

Justifications for (and efforts to eliminate) variations in provision have focused primarily on technical decisions but, in a resource constrained environment, variations might well reflect different priorities in health care. Therefore, it would be important to distinguish the two sources of variation and make this issue clearer to the general public. There is some evidence that the latter’s attention is also focused on technical matters. For example, in Dolan *et al.* (2007), it is said that geographic consistency is important to people and that respondents expressed a greater willingness to accept variations in outcomes if authorities used “the same way to come to the conclusion of which treatment to use”. But because different treatments usually involve different costs, the question might not be ‘which treatment to use’ but instead ‘which treatments should be funded’.

The debate preceding and following NICE’s guidance on infertility treatment is an example of this confusion between the different sources of

variations in provision<sup>42</sup>. When the (at the time) health secretary Alan Milburn announced that NICE was to provide national guidance on infertility treatment, some health authorities were spending up to £2m on this type of treatment whereas others were spending nothing at all (cf. Kmietowicz, 2000). Hence the ‘postcode lottery’ associated with this kind of treatment. Clare Brown, executive director of the Charity Child and president of the National Infertility Awareness Campaign, said in respect to the government’s instructions to NICE: “This clearly indicates that the government intends to make sure that infertility treatment should be funded by the NHS” (Kmietowicz, 2000). This statement illustrates how ending ‘postcode lottery’ might be in some cases almost automatically interpreted as providing in all authorities what is being provided in only some of them. Nonetheless, in theory, NICE’s guidance could result in no funding of infertility treatment in the NHS. That is, the elimination of variations could in principle be achieved via no provision at all.

NICE’s guidance on this topic was issued in 2004 and a key recommendation is the provision of up to three free cycles of in vitro fertilisation (IVF) for couples in which the woman is aged between 23 and 39 years at the time of treatment and who have an identified cause for their fertility problems or who have been unable to conceive for three years (NICE, 2004). If the guidance was fully implemented, the demand for infertility treatment was estimated to increase by 80% at a projected additional cost to the NHS of £85m in England and Wales (cf. White, 2004) and the Department of Health directed the health service to move gradually towards this target. The Department of Health guidelines say that the NHS in England must offer couples at least one cycle of IVF treatment and that those without children living at home should receive priority (Shannon, 2005).

Rawlins and Culyer (2004) say that if the government judges a particular intervention unaffordable for the NHS, even though NICE had judged it cost effective, it might respond in two ways: it might formally advise the NHS to ignore NICE’s advice; or, alternatively, ministers might instruct NICE to take account of advice from ministers on available resources. According to Rawlins

---

<sup>42</sup> Although infertility treatment is not seen by the public as a high priority service (Bowling, 1996; Schickle, 1997), we use this example because it illustrates at once the various points under discussion.

and Culyer (2004), neither of these responses had been proposed or threatened. In the case of infertility treatment, the Government chose a gradual implementation of NICE's advice but the health secretary at the time guidance was issued, John Reid, did not mention how long full implementation would take and he did not mention additional funding either (cf. White, 2004).

When draft guidelines were issued for consultation, the largest part consisted, as Ashcroft (2003) notes, in a series of evidence based recommendations about particular techniques and protocols used in the management of infertility. However, most media attention was paid to the proposed recommendation about the level of service to be provided to couples experiencing difficulties in conceiving. But most importantly, Ashcroft (2003) notes that the clinical criteria from NICE for the provision of in vitro fertilisation are framed in terms of the effectiveness and likely success of treatment for men and women of different age and diagnostic groups on the basis of the available data. He stresses nonetheless that what arguably explained most of the variation in service provision was social values. Ashcroft (2003) continues: "The debate about publicly funding assisted conception is in essence a debate about values and priorities rather than a debate about what works. Under a publicly funded NHS it is a political decision, not a clinical one, about how far we as a society want to provide fertility services and to whom. (...) Local priority means setting local determinations of relative importance and local judgments of value. Differences should be allowed to flourish so that people can make personal and political choices. (...) As long as people have reliable information (including reliable technical information such as that provided by NICE) and decision makers are accountable for the reasonableness of their decisions, why should there be national policy?"

The testimony of Brambleby (2001), a consultant in public health, corroborates the view that variations in priorities did exist before guidance. He describes the situation in the Norfolk (former) health authority as follows: "In east Norfolk last year we lengthened screening intervals for cervical cancer from 36 months to 54 months, at a small but tangible health loss. But we are now able to sustain our substantial in vitro fertilisation programme as well as allow sterilisations on the NHS (...) this has been supported at public consultation. A neighbouring health authority claims that in vitro fertilisation and sterilisations are

unaffordable and continues its 36 month cervical screening cycle”.

The uneasiness, in England, to deal with unequal provision seems to apply to various public services and not only to health care. Burns (2000, p.967) talks about a ‘deep-seated’ difficulty that has emerged with local democracy: even those councils with a strong ideological commitment have found it difficult to accept different levels and patterns of service in different neighbourhoods. “It is almost as if they had not realised that the very local responsiveness for which they argued, produces the diversity that they are uneasy about”.

Being the geographic variation in health care provision a very sensitive question in England, the debate would only benefit, we think, if the different sources of variation were clearly distinguished and the public’s view towards geographic variation was open and directly addressed, namely in empirical terms.

## **5.6. Ambiguities regarding the role attached to participation in the NHS**

We saw in chapter two that increased participation in decision making is an outcome attributed to decentralisation, in the respective literature. In England, at the policy level, there has been an association between participation in the NHS affairs and local organisations as well. For instance, the statutory duty to involve patients and the public is placed specifically on local organisations (Health and Social Care Act 2001 - section 11). Also, one of the criteria to evaluate proposals regarding the re-structuring of PCTs is their ability to ‘improve public involvement’ (Department of Health, 2005). As previously reported in chapter two, some authors have expressed doubts that a positive link between decentralisation in the NHS and participation exists and pointed the lack of empirical evidence in this respect (Peckham *et al.* 2005; Allen, 2006). Additionally, despite the focus of authorities on local organisations, there is some evidence that suggests that participation is seen by the public as more relevant at higher levels of decision making (Litva *et al.*, 2002; Wailoo and Anand, 2005).

Concerning the various roles attached to participation, we know from chapters two and three that in theory it can be seen both as an end (procedural view) and as a means (consequentialist view). In addition, there are two, not



necessarily coincident, perspectives to consider in practice: the perspectives of the public and of the authorities.

Regarding the general public's views towards participation, much of the empirical evidence produced refers to the UK or English context. Consequently, the main results presented in chapter two already give an indication about the English public's approach to participation. That is, participation is valued by the general public (e.g. Anand and Wailoo, 2000; Litva *et al.*, 2002; Wailoo and Anand, 2005; Dolan *et al.*, 2007), both as an end and as an instrumental mechanism to reach certain outcomes (Dolan *et al.*, 2007). Nonetheless, in the eyes of the general public, doctors are in the vast majority of cases ranked first when it comes to determine who should set priorities (e.g., Bowling, 1996; Mossialos and King, 1999; Dolan *et al.*, 1999; Coast, 2001; Litva *et al.*, 2002). Still, there seems to exist some ambiguities in this regard as well. For example, in Anand and Wailoo (2000), respondents say that a health authority should follow public opinion in a case where a treatment has a very low expected health gain and therefore is not supported by doctors, but which the public has said should be provided. It seems that people are willing to surpass doctors' view when it means providing health care that would not otherwise be provided.

Furthermore, although surveys of public views suggest that people value participation, Milewa *et al.* (2002), for instance, reporting on a study of 167 Primary Care Groups (PCGs), say that only 8.7% of PCGs give pressures from the public as a reason to involve it in decision making. Rowe and Shepherd (2002), based on their study of PCGs too, identify the existence of divergence between local and national decision-makers regarding the extent to which they considered there to be a demand for involvement from the public. The Department of Health (2004b, p.3) claims however that the low public awareness of the opportunities for involvement in the NHS does not reflect a lack of willingness to become involved. The latter perspective might in fact reflect a view that preferences for participation are to some extent endogenous (see section 3.8, in chapter three, about endogenous preferences).

English authorities have also attributed instrumental value to participation. One of the stated objectives of PPI Forums is precisely to help the public to feed its views about health services into the NHS. There are suggestions that the instrumental approach is also embedded in local organisations. For instance,

Rowe and Shepherd (2002), after surveying Primary Care Groups, conclude that the instrumental approach to public involvement is an area of consensus between national and local decision-makers. In addition, two other research projects (Department of Health, 2004b, p. 21) came to the conclusion that the instrumental approach to participation is embodied in PCTs' views regarding public involvement.

But here too things seem to be non-linear. Milewa *et al.* (2002), for example, report that only 32% of the PCGs included in their study claimed that changes had taken place due to public and patient involvement. From these, only 29.3% could cite examples in detail sufficient to be recorded. Similarly, Alborz *et al.* (2002) report that few chairs of their sample of PCG/Ts felt that consultation had much impact on decisions taken by their respective PCG/Ts (e.g., 77% said that consultation had little or no impact on decisions relating to *clinical governance*).

There is some ambiguity in the central authorities' perspective. That is, on one hand, the instrumental value of participation is stressed; on the other hand, this participation is to occur at the local level but local organisations (e.g. PCTs) are greatly constrained by central regulations such as NICE guidance and NSFs. For example, Bate *et al.* (2007) say that national drivers were described by participants in their study (decision-makers from six PCTs) as 'the must-do's' and "it was suggested throughout the interviews that the PCTs tended to focus on reacting to and directing/allocating resources towards the national drivers first and foremost". Thus, the flexibility of local organisations to be responsive to preferences seems limited.

In fact, this might partly explain why the pressure to involve the public has come mainly from the centre. That is, lower propensity at the local level to promote participation might be a sign of little decentralisation of effective decision making power. This is basically what Greener and Powell (2003) suggest. These authors, based on a questionnaire sent in 1999 to 121 health authorities in England and Wales, suggest that health authorities' attitude towards public consultation is itself conditioned by local discretion: in essence, if health authorities have few resources to use at the local level, consulting widely their local populations is seen as a waste of time; on the other hand, those authorities that consult widely with the public are either paying less attention to national

guidelines than other authorities or have already met these guidelines and are relatively resource-rich. Milewa (2004, p.242) too refers that despite the role of local 'consumers', under the quasi-market, in advising health authorities on the type of health services that should be commissioned, "the limited upheaval in local health services and professional-managerial discretion in acting upon such input has limited the change that could be attributed to public and patient involvement".

Another explanation for the limited impact of public views on policies might be of course that public involvement is seen as a way of giving pressure groups an opportunity to use preferences merely as advocacy. In this line, Rowe and Shepherd (2002, p. 279) say that user-group representatives have been labelled as 'activists' and that their views have been dismissed as not being typical of 'normal users'. Also, in Department of Health (2004b, p. 20), it is reported that some professional Primary Care Trust members are less certain about the value of public involvement, fearing that the people selected may not provide a representative viewpoint, or that they and their voluntary organisation representatives may push a self-interested agenda. Callaghan and Wistow (2006, p.2291) too talk about "manipulation by the better organised and more powerful, whether these be professionals, managers or, in some cases, user groups".

Although different approaches to participation are acceptable and have been inclusively sustained at the theoretical level, it seems that there is some discrepancy between the objectives (particularly, instrumental objectives) of public involvement announced by the government and what local organisations like PCTs can effectively achieve.

## **5.7. Concluding comments**

Some ambiguities and confusions seem to exist in the debate direct and indirectly related to decentralisation in the NHS. In the future, more attention should be given to these issues in order to make the discourse and policy guidance clearer. Regarding the question of variations in health care provision, and although geographic inequalities in health have been considered in policy

objectives (Department of Health, 2001a), there is a lack of empirical evidence focusing on people's preferences about equality in provision across geographic areas. In fact, much of the empirical evidence about people's preferences, in the context of health care resource allocation, produced thus far has used samples from England or UK but studies of the sort reported in chapter four seem to be lacking in this scenario.

There is also some ambiguity regarding the instrumental role attributed to public involvement given the centrally driven constraints that PCTs are faced with. Here we must however note that greater flexibility at the local level to meet local preferences has to be weighed against poorer accountability in the sense that, as noted by Saltman and Bankauskaite (2004), a basic principle of public finance is that spending and taxing authority should be located in the same agency to facilitate accountability. Given that tax decentralisation is unlikely in England, as in most Western European countries, any instrumental approach to public participation might have to be thought in restrictive terms.

Finally, participation is seen as being inherently good (the arguments supporting this perspective are reviewed in chapter two); Harrison and Mort (1998, p.66) say that being in favour of better public consultation or more user involvement is "rather like being against sin: at a rhetorical level, it is hard to find disagreement". However, as we acknowledged in our discussion in chapter three, participation might have relevant opportunity costs; at some point, costs might even outweigh benefits. Thus, more complete analysis should take those into account and the debate at the policy level should be more open about benefits, costs and limits to participation.

## CHAPTER 6. Conclusion

The objective of this concluding chapter is to summarise and give an overview of the main results, limitations and directions to future research, that have already been identified throughout this dissertation.

The main contributions of the current work are, at the theoretical level, the suggestion of a new role that might be attributed to preferences in the evaluation of decentralisation in terms of its impact on social welfare compared to centralisation; at the empirical level, we addressed two topics that have been given little attention, in the empirical literature: the hypothesis of geographic variation in preferences was tested and preferences regarding a trade-off involving a geographical dimension of inequality were elicited.

Within our framework, central and local decision-makers are seen as alternative *agents* acting on behalf of local populations. Given the different capabilities possessed by *agents*, decentralisation of resource allocation generates some trade-offs between objectives. Depending on the goals of resource allocation envisaged by local populations, that is, depending on the trade-offs that they are willing to make, local populations will be better-off with one or the other *agent*. Therefore, we conclude that not only information asymmetry regarding preferences or competition among decision-makers (the usual economic approaches) matter but the specific preferences held by the public might also determine in themselves whether or not decentralisation is optimal, when compared to centralisation.

Concerning the empirical work, the principal conclusion is that the results do not corroborate the hypothesis of geographic variation in preferences. Basically for all variables, the null hypothesis, that there are no differences in preferences across samples, cannot be rejected. This is true for the Student's *t*-test, Kolmogorov-Smirnov test, the Fisher exact test, and the Wilcoxon-Mann Whitney test. Only in the case of health maximisation the Student *t*-test suggests the rejection of the null hypothesis at the 5% level but mean values (representing the support to this principle on a scale ranging from 1 to 5) are similar in both samples (4,5 and 4,2). Moreover, the null hypothesis could not be rejected in both cases of social and personal preferences. Thus, it seems that the nature of

preferences is not decisive for the existence or not of diversity in preferences.

Based on these results, the theoretical discussion about the impact of decentralisation on allocative efficiency should be revisited and greater attention should be paid to the role that the assumption of geographic diversity in preferences has played.

Regarding the variables that measure the public support to the principles of health maximisation and geographic equality of treatment as well as the trade-off between them, respondents agree or strongly agree that resource allocation should take into account both principles. Once respondents are asked to trade-off both principles, between 70 and 80 per cent of them are prepared to make trade-offs between the two criteria. Of those making trade-offs, most people are willing to forego between 10 and 30 per cent of total health gain to keep geographic equality of treatment in return. The median willingness to sacrifice health maximisation on behalf of geographic equality of treatment ranges from 10% to 25% of total health gain, depending on samples and questions (two person trade-off questions were included in the questionnaire). Also, the median respondent is willing to trade-off 10 children of the average number of children avoiding the disease in the two regions (in the case of question 3) and 583 (or 1333, depending on sample) people of the average number of people having a surgery in two regions (in the case of question 4). Taking into account both samples and questions, the mean value of  $r$  ranges between -0,53 and -0,65 suggesting the existence of aversion to inequality in the geographic distribution of total health gain. There are some respondents not willing to sacrifice one principle on behalf of the other but they form a minority.

These results thus suggest that the geographical dimension of (in)equality in treatment matters to people and that a maximum opportunity cost of equality, in terms of health gain foregone, is likely to exist. Hence, the empirical findings support the relevance of the discussion developed in chapter three in which this kind of trade-off is addressed.

Overall, the results obtained in this dissertation imply that strategies of decentralisation might have to focus on outcomes other than a better matching between provision and preferences and, given the possibility of decentralisation causing variations in provision, it should be taken into account that this is likely to have a negative impact on social welfare but it should also be noted that the results suggest that people are not willing to pursue equality in provision at all

costs in terms of health gain foregone. The conclusions are nonetheless valid under some assumptions that were made namely in the theoretical discussion and the empirical study also has some limitations.

We have already mentioned the implications of relaxing the hypothesis of geographic variation in preferences. Another assumption is that central and local authorities are benevolent decision-makers however it has been acknowledged, e.g. within the usual *principal-agent* approach, that decision-makers might have self-regarding interests. Therefore, potential gains derived from superior information about local preferences might not be materialised because local decision-makers simply seek to achieve their own objectives. It might also be that greater public participation motivated by decentralisation gives pressure groups the opportunity to pursue their own agendas (as mentioned in chapter five, for the context of the English NHS). In any case, local populations are not necessarily better-off by having decisions made at the local level.

Information asymmetry is an important assumption in our framework, in particular, the assumption that each decision-maker has an advantage over the other regarding one of the two types of information. Still, we could reach the conclusion that preferences might determine the impact of decentralisation on allocative efficiency, in the same, assuming that information asymmetry goes the other way around (i.e. assuming that the central decision-maker is better informed about preferences than local authorities while the opposite is true for technical information).

Part of the theoretical discussion makes use of the assumption that there is a positive link between decentralisation and participation in decision-making. This assumption is in fact widely accepted in the literature about decentralisation. This link has nonetheless been questioned for the case of health care, namely for the case of the English NHS. Moreover, we have not considered the opportunity costs of participation that are likely to exist for participants themselves (e.g. time costs) and in terms of diverting resources of health care budgets from health production to sponsor participation activities. Determining health care priorities might even cause what has been termed as 'disutility of denial'. These are factors that tend to lower the positive and direct impact of participation on welfare and which might be considered in future analyses.

Throughout our discussion we have assumed, as it has been done in the

mainstream economic theory, that preferences are exogenous. Nevertheless, in the face of the criticisms that this assumption has stimulated, the assumption of exogenous preferences might represent a limitation of the current work. Relaxing this assumption has a relevant impact on our conclusions given that, in the case of endogenous preferences, the optimal setting (centralised or decentralised decision making) cannot be identified based on preferences given that the latter are in turn affected by the chosen setting. For example, if people attach little inherent value to participation because under centralised decision making they have not been given the opportunity to participate, then one cannot say that decentralisation is not optimal compared to centralisation because people do not value participation.

When discussing externality effects associated with concerns for geographic inequality in health, we did not consider the possibility of movements of people between jurisdictions. Thus, if this happens, resource allocation between areas might not be a feasible instrument to impact upon health distribution and then the advantage of centralised allocation in terms of internalising the spillover effects disappears and the discussion in section 3.6 falls in the scenario adopted in section 3.3 (that is, the argument to decentralise stops being the content of preferences and becomes the usual argument of information asymmetry about them).

Regarding the empirical results, in addition to the limitations related to the specific regions from which samples were drawn and the specific topics included in the questionnaire, the method of administration might be viewed as a limitation because it has been acknowledged that, in the case of person trade-off questions, respondents often require greater explanation of the task than can be offered through a written survey instrument. The sizes of samples might also be considered a limitation for the generalisation of results.

Several theoretical and empirical topics remain to future research. For instance, it has been suggested that, in the case of endogenous preferences, the nature of preferences (in terms of personal *versus* social preferences) might be a relevant issue to determine which preferences should be used to inform policy. Moreover, it has been suggested that the focus of discussion might even shift to the question of which preferences ought to be encouraged. Thus, in the future, normative questions of this kind might have to be addressed.

An important question that remains to future research concerns the rule of



aggregation of the desideratum at the suprajurisdictional level, that is, how to evaluate changes in overall social welfare given different preferences across jurisdictions? It is of course possible to evaluate changes in social welfare from a regional perspective and from a national perspective. But in the case of conflict between the two, how can we reach a conclusion on the overall impact of decentralisation?

In terms of the empirical work, considering the Portuguese context in which this work was initiated, future studies should include other regions, namely southern regions as well as some fringes of the population that were excluded from our study. Methods of administration such as structured interviews or discussion groups might be considered, in particular for older people and people with no degree (or with a low degree) of education. It might also be interesting, from the perspective of testing for differences between geographically defined samples, to group regions in terms of mainly rural and mainly urban areas. In addition, future research might look into the impact of the socio-demographic composition of samples on expressed preferences in order to test whether eventual differences between samples stem from their different compositions.

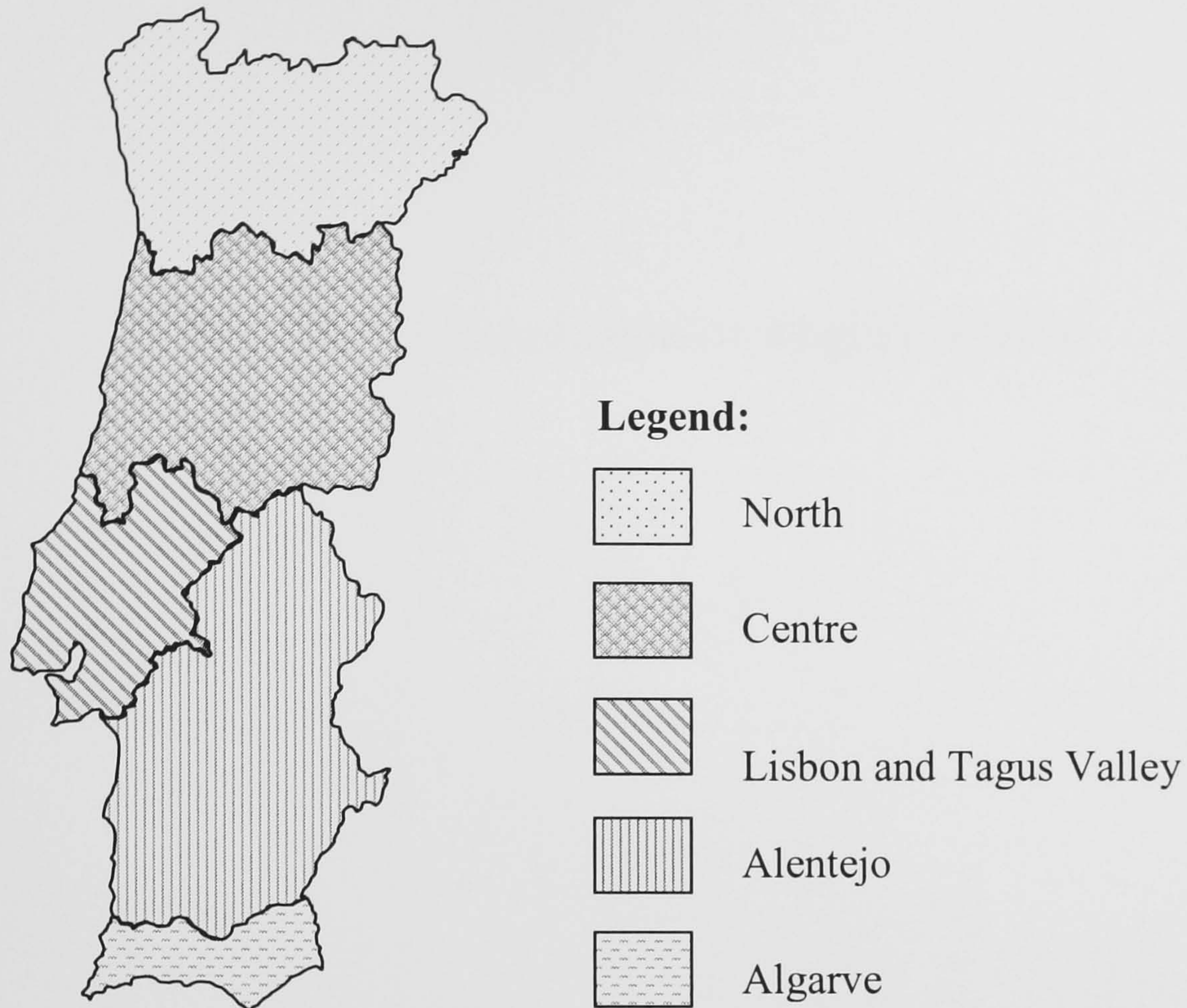
Generally, and beyond the Portuguese context, questions could be refined in order to make more explicit the difference between geographic equality of health gain and geographic equality of treatment (studies might inclusively investigate whether or not this distinction has a significant impact on results). Also, a trade-off that seems to be worth exploring is that between the allocation of resources at the local level based on local preferences *versus* uniform provision throughout the country. The latter seems to be a particularly interesting issue to explore in the English context given the ambiguity that seems to exist there between the focus on 'localism', on one hand, and the efforts to eliminate variations in provision, on the other hand. In fact, much of the empirical evidence about people's preferences, in the context of health care resource allocation, produced thus far has used samples from England or UK but studies of the sort reported in chapter four, in which the geographic dimension of (in)equality is explored, have not been implemented in England. Therefore, future analyses should take into account that there is this gap in the empirical literature. Yet another type of trade-offs that remains unexplored in the empirical literature is that involving outcomes, on the one side, and procedures, on the other. For instance, it would be interesting to see if; what; and

how much, are people willing to give up in order to increase their participation in decision making.

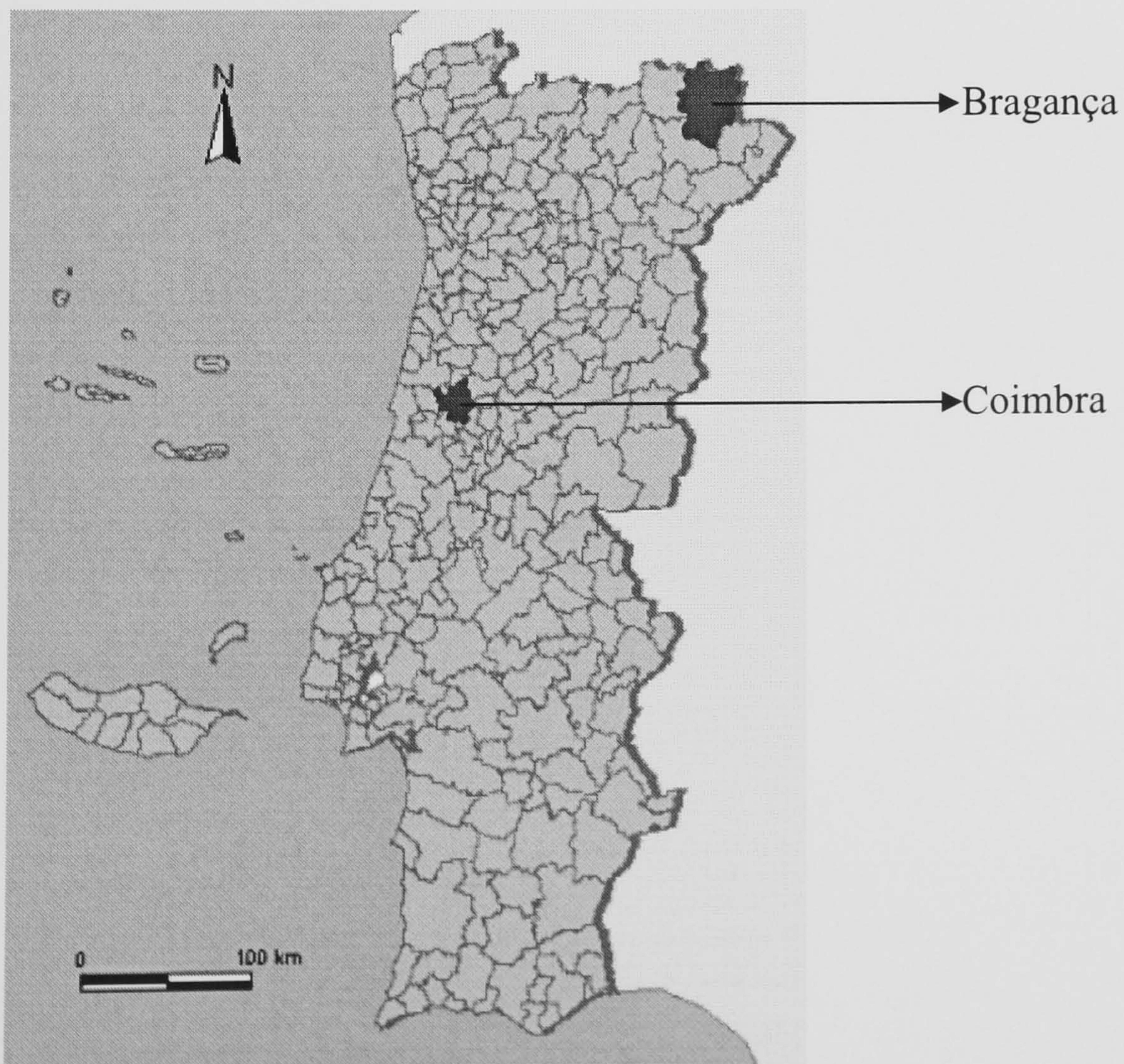
Decentralisation is a current theme in the political agendas of different countries. Looking at the case of the English NHS, we suggest that there is a lack of clarity in the use of concepts like preferences and that other ambiguities and confusions have characterised the debate about (and related to) decentralisation. Although our discussion has focused on the English case, the same is likely to be happening in other countries as well. Thus, researchers might have a role to play here in terms of identifying these ambiguities and point the need for clearer policy guidance.

## APPENDIX A – Maps of Portugal

Regional Health Administrations:



The Municipalities (Concelhos) of Bragança and Coimbra:



## **APPENDIX B – Questionnaire: English version**

## Questionnaire

### How ought health care resources to be distributed?

- This questionnaire is aimed at the general public.
- Thanks to scarcity, decisions about the distribution of health care resources are inescapable – It is important for decision makers to know what citizens would like from their National Health Service.
- There are no right or wrong answers; the researchers are only interested in **your opinion**.
- Answers are **anonymous**.
- There are 6 questions which will take about **10-15 minutes** to complete

Thank you very much for your co-operation.

If you have any queries, please contact:

Carlota Quintal

Faculdade de Economia da Universidade de Coimbra

Av. Dias da Silva, 165, 3004-512 Coimbra

Tel: 239790580

Email: [qcarlota@fe.uc.pt](mailto:qcarlota@fe.uc.pt)

**In the following two questions, we are asking you about your opinion regarding principles or criteria that might guide resource allocation.**

For each of the following two questions please circle ○ the number that matches your view most closely:

**Question 1**

There are some health care interventions (for example, medicines and medical procedures) that have a greater impact on the individuals' health state than others. In some cases, this impact means improving the individuals' health; in other cases, it means narrowing the deterioration of the individuals' health.

Do you agree that decisions about the distribution of resources available to health care ought to take into account, among other possible criteria, the impact that these resources have on the individuals' health state as well?

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree (neutral)	Agree	Strongly Agree

**Question 2**

Resource allocation in health care involves the distribution of resources among different interventions but it also has a geographical dimension, involving the allocation of resources among different regions of the country.

Do you agree that decisions about the distribution of resources available to health care ought to take into account, among other possible criteria, how these resources are distributed across different regions as well?

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree (neutral)	Agree	Strongly Agree

In the following two questions, we are asking you about your opinion regarding the relative importance of the two principles mentioned in questions 1 and 2.

**Question 3**

Imagine that following a scientific discovery, a new vaccine for children becomes available. This vaccine is capable of completely immunising (protecting) children against a given disease, saving their lives. The disease is **not contagious** but is lethal.

Without the vaccine, some children will become ill and consequently die; others, even without the vaccine might not become ill.

There is however a limited quantity of this vaccine for each country, which is not enough to immunise all children. A decision must be made regarding which children to immunise.

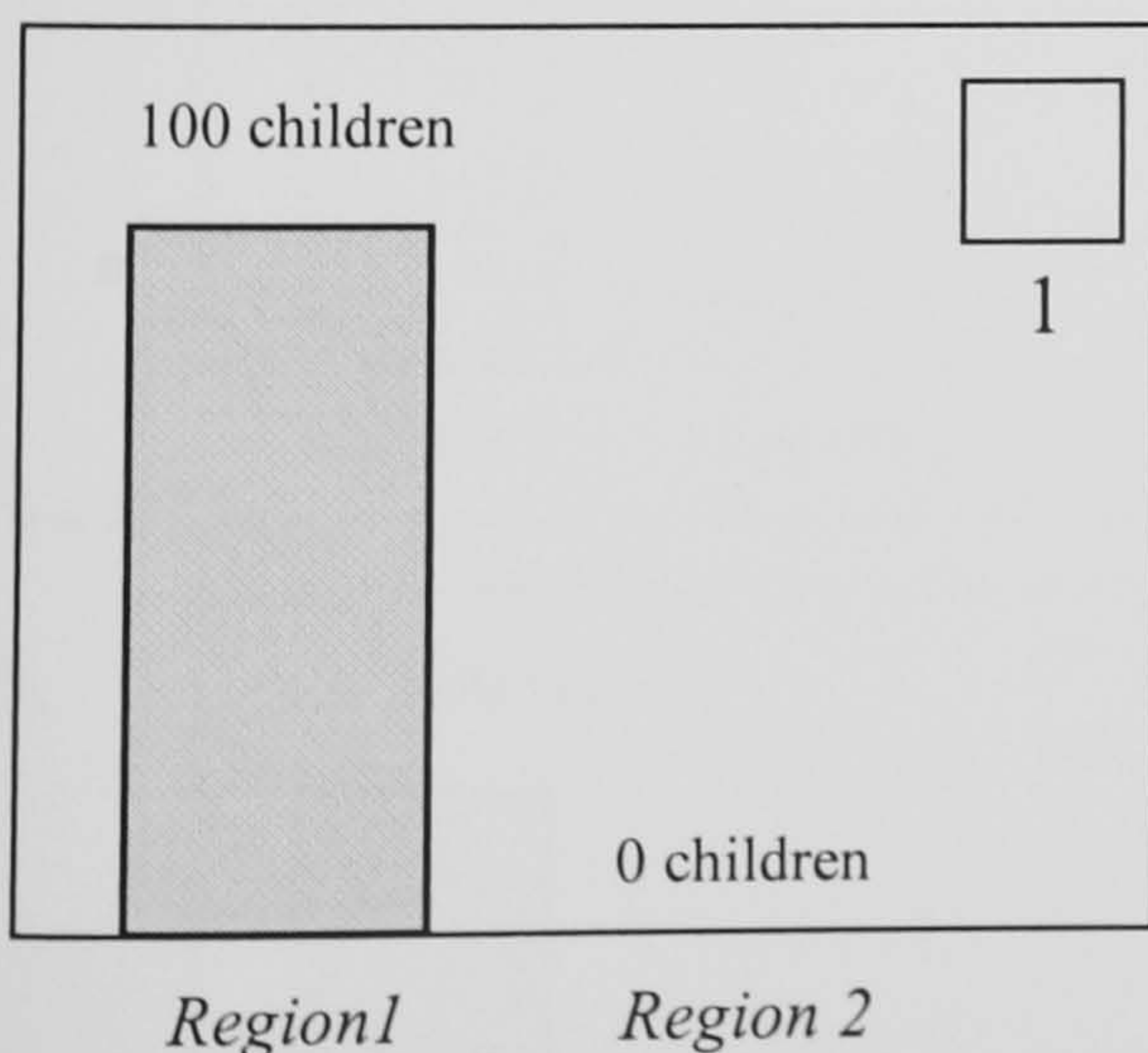
Imagine then that the country is divided in two regions, *Region 1* and *Region 2*, where half of the children lives in *Region 1* and the other half of the children lives in *Region 2*.

You must choose between distributing all doses of the vaccine among the health centres in *Region 1* (Scenario A) or distributing half of the doses among health centres in *Region 1* and half among health centres in *Region 2* (Scenario B). Both scenarios involve the same total reduction in episodes of illness but differ in terms of the geographic distribution of the episodes of illness that are avoided.

Please, indicate which scenario A or B you would choose, crossing  the corresponding box:

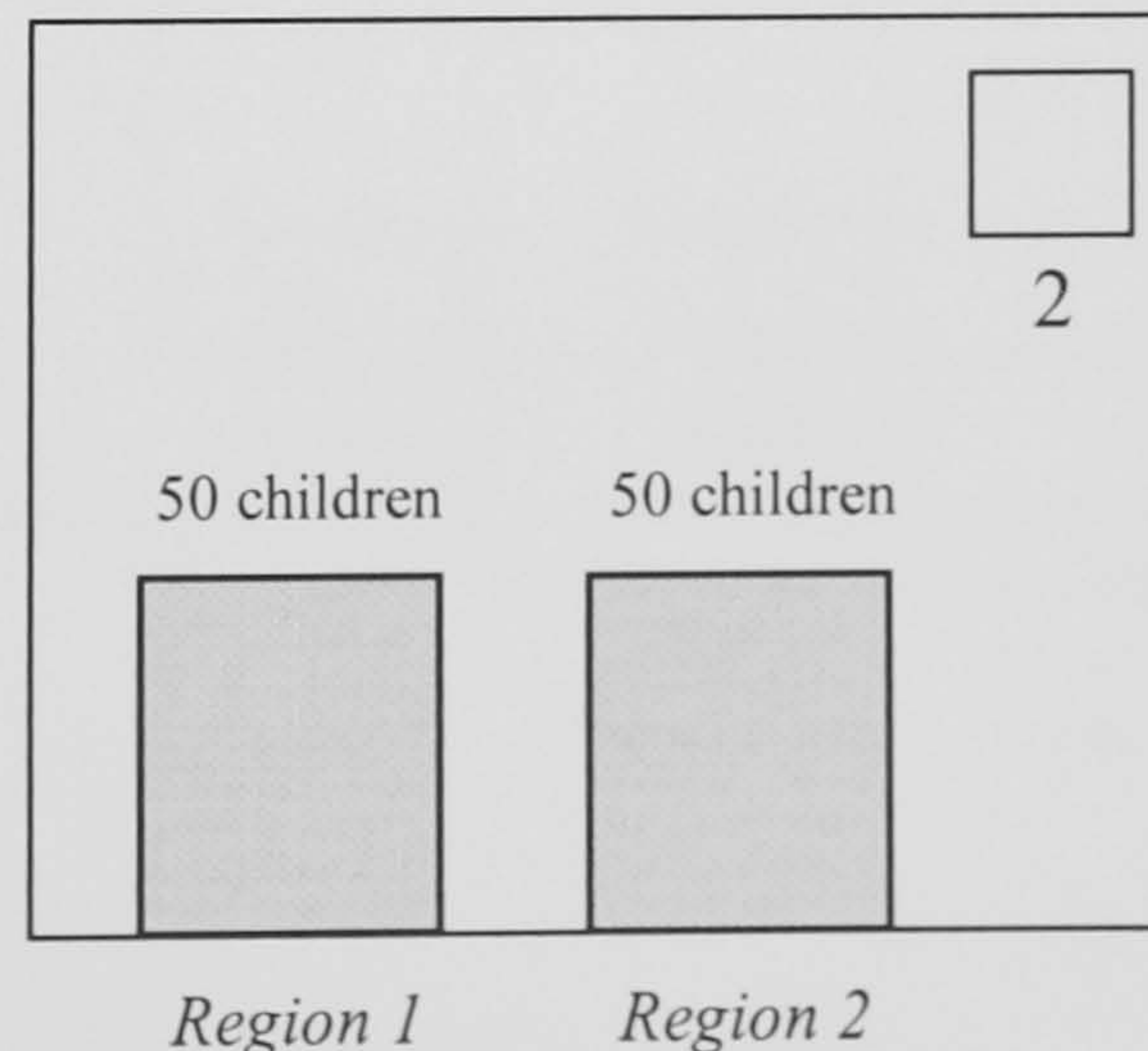
**Scenario A**

Number of children who will avoid the disease



**Scenario B**

Number of children who will avoid the disease



or

**ATTENTION:** If you chose **Scenario A** please turn to **question 4** in page **5**.

If you chose **Scenario B**, please continue on the following page (p.4)

Suppose now that targeting two regions (Scenario B) is less effective than concentrating on one region (Scenario A).

For each of the five choices bellow, please cross  one box, indicating whether you would still choose Scenario B, or whether you would now choose Scenario A:

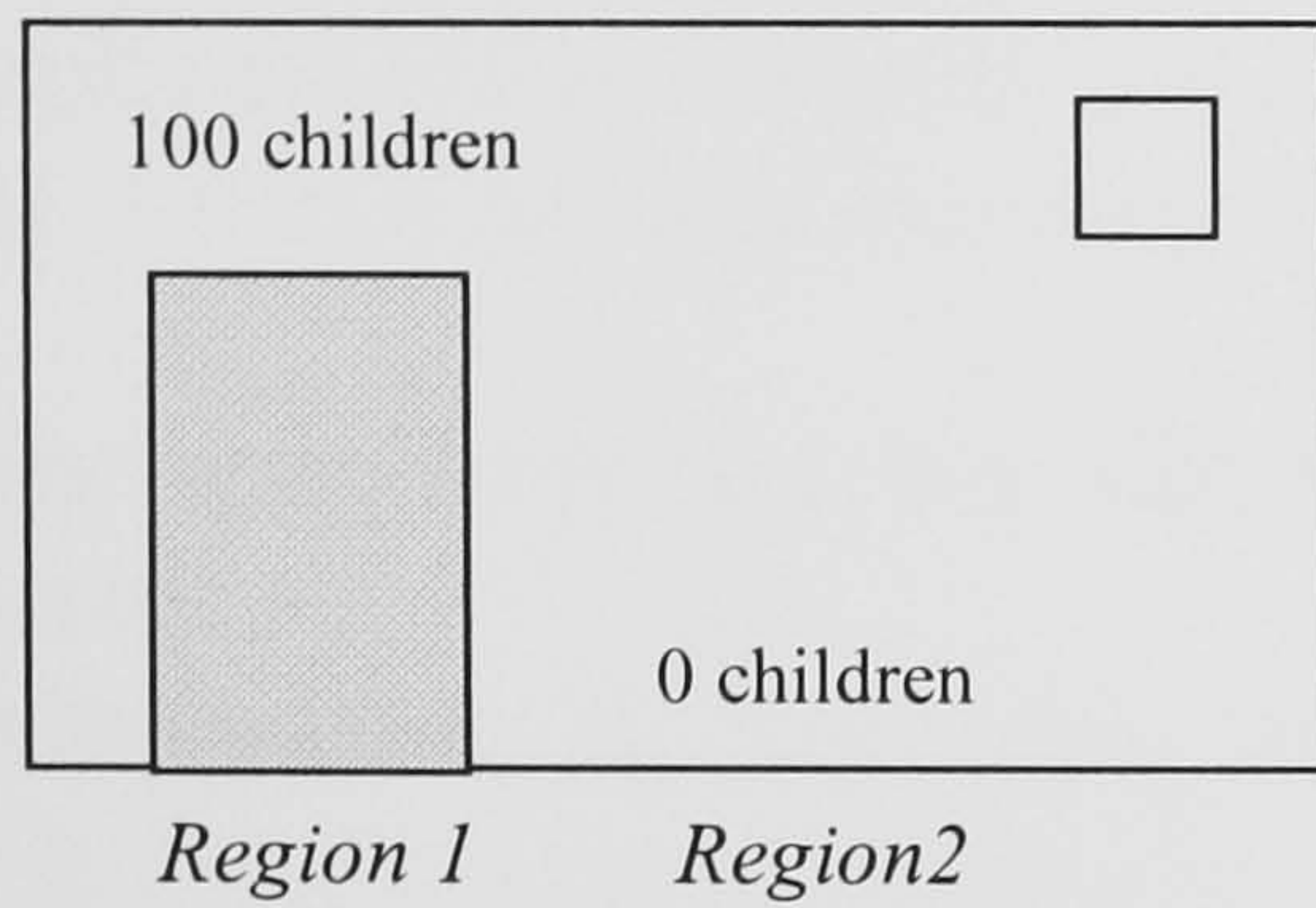
**Scenario A**

**Scenario B**

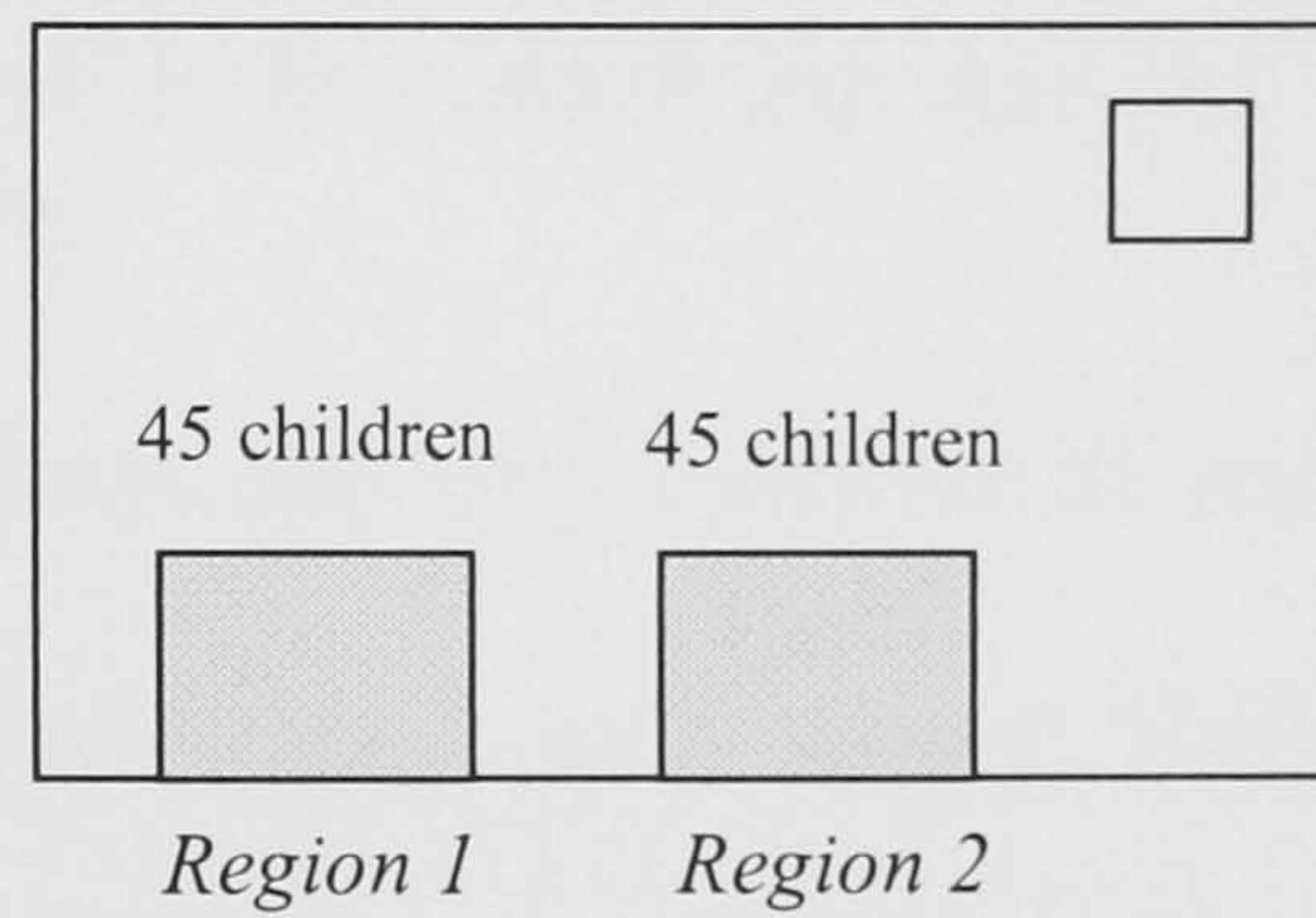
Number of children who will avoid the disease

Number of children who will avoid the disease

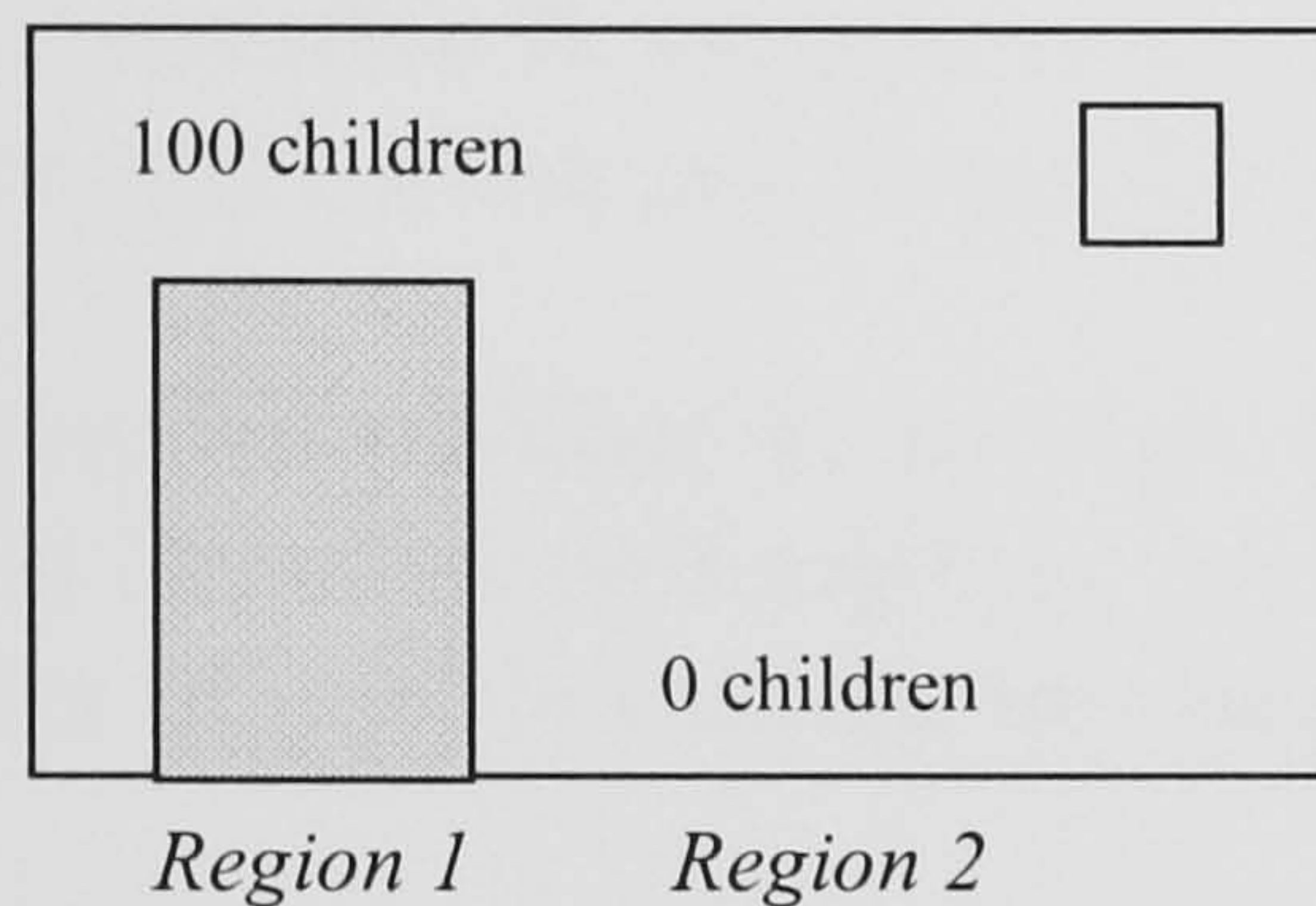
1.



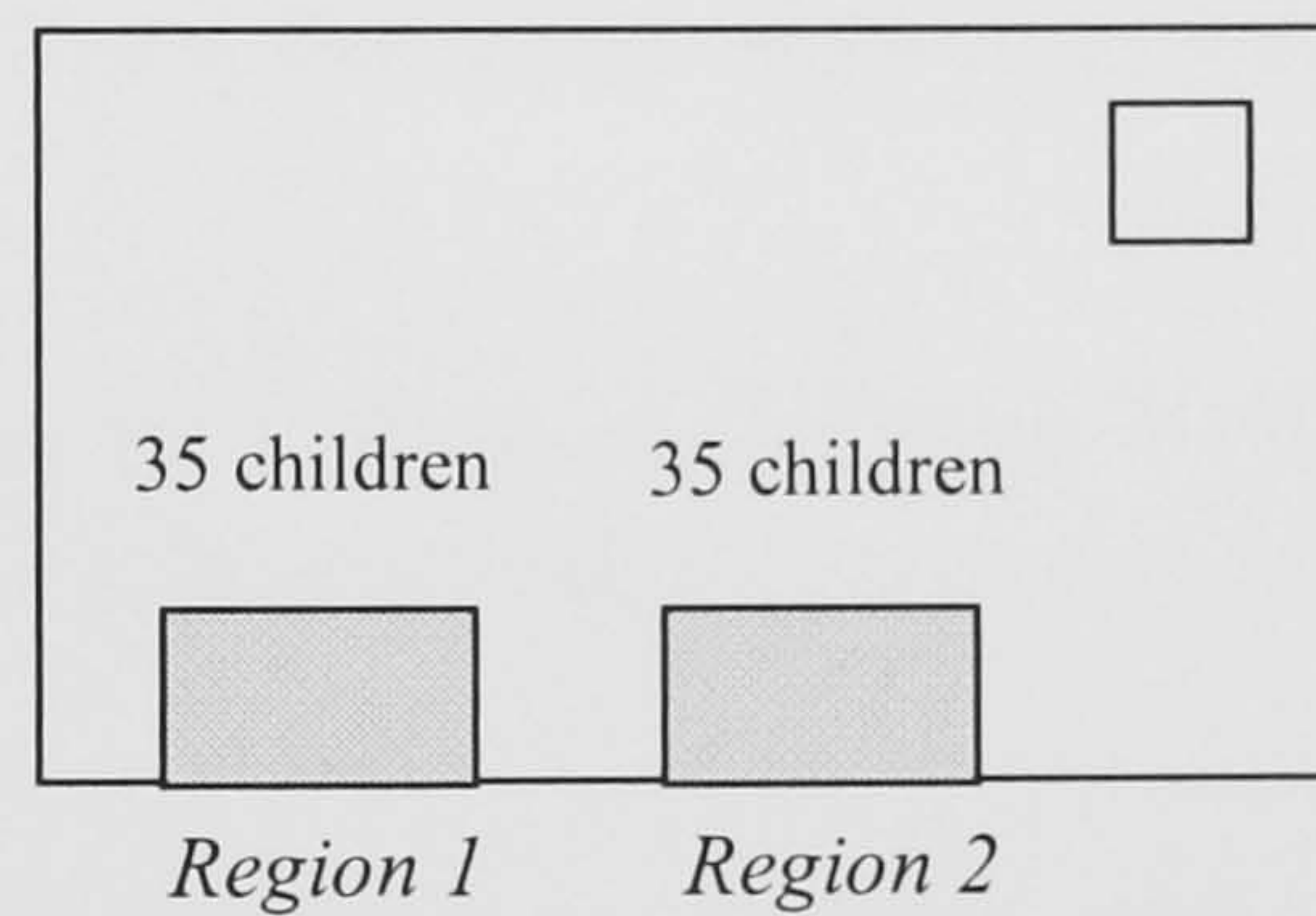
or



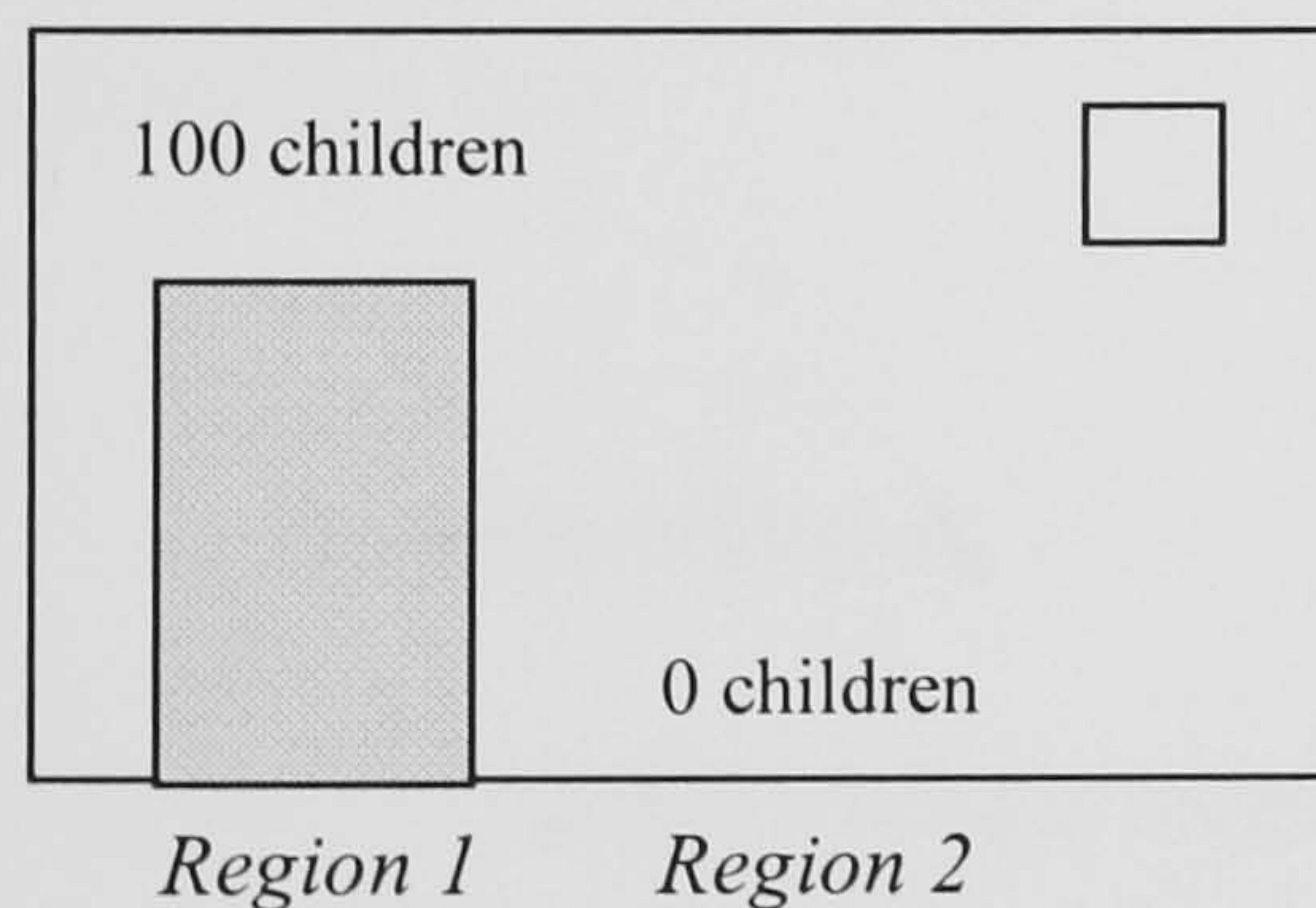
2.



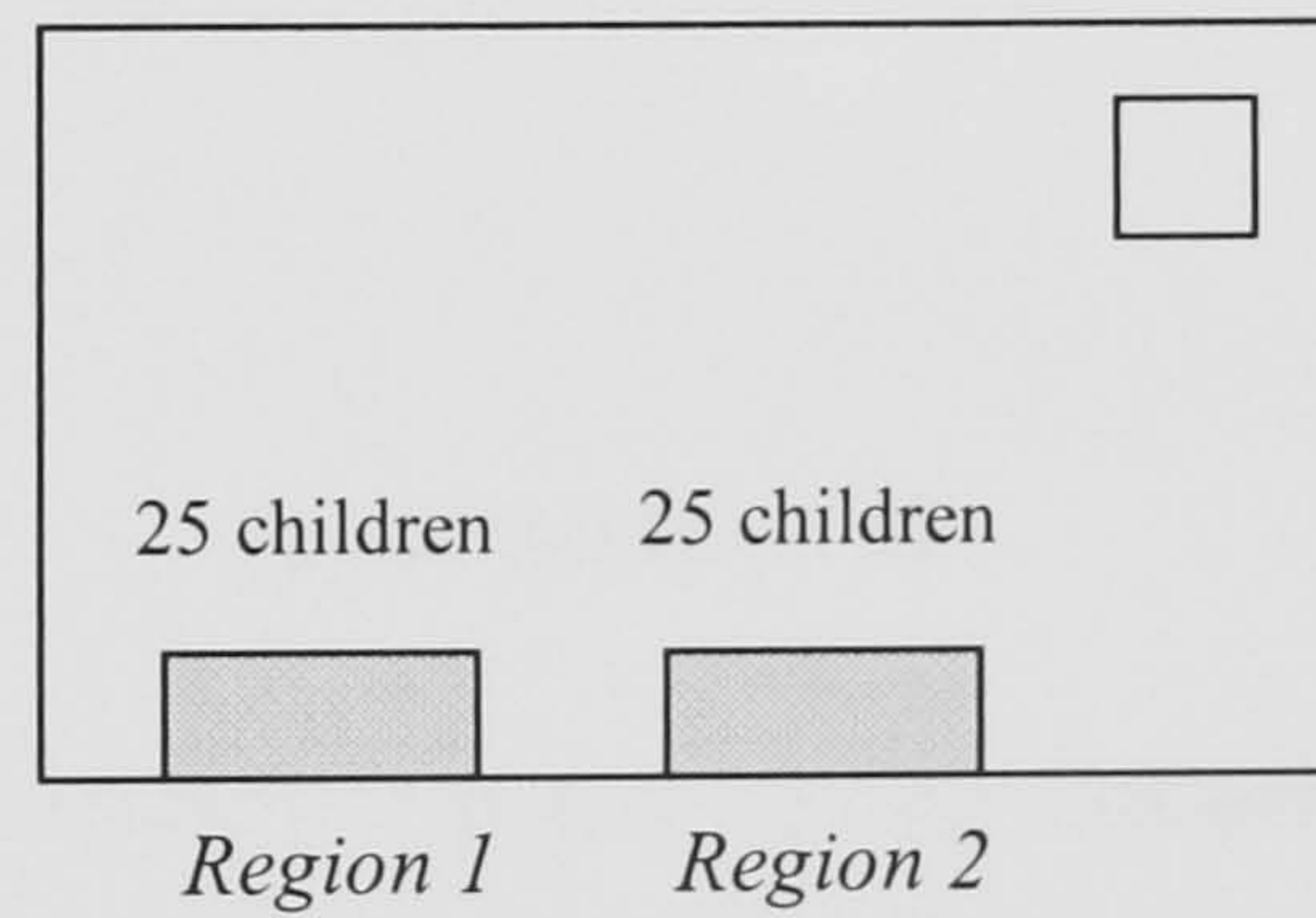
or



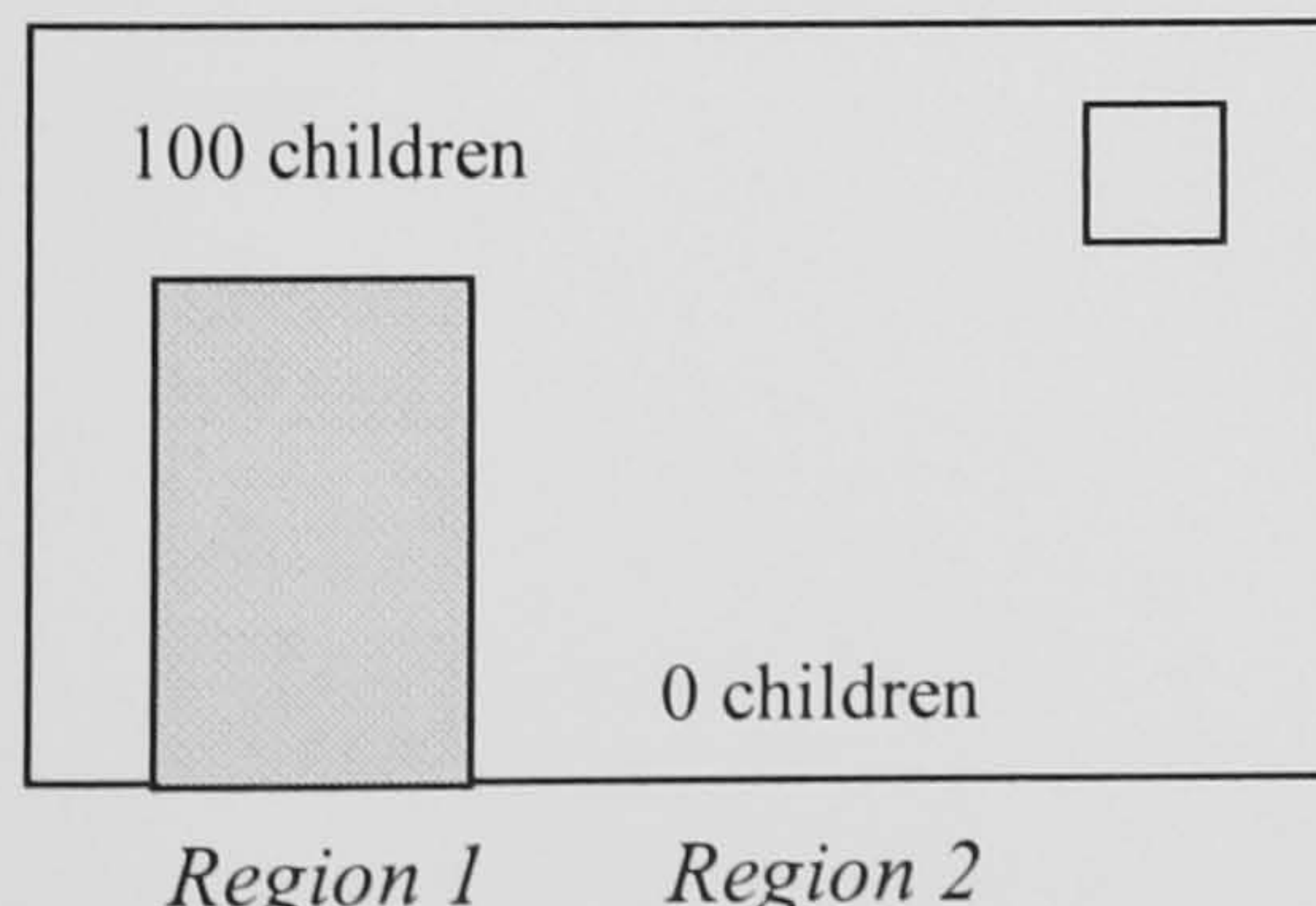
3.



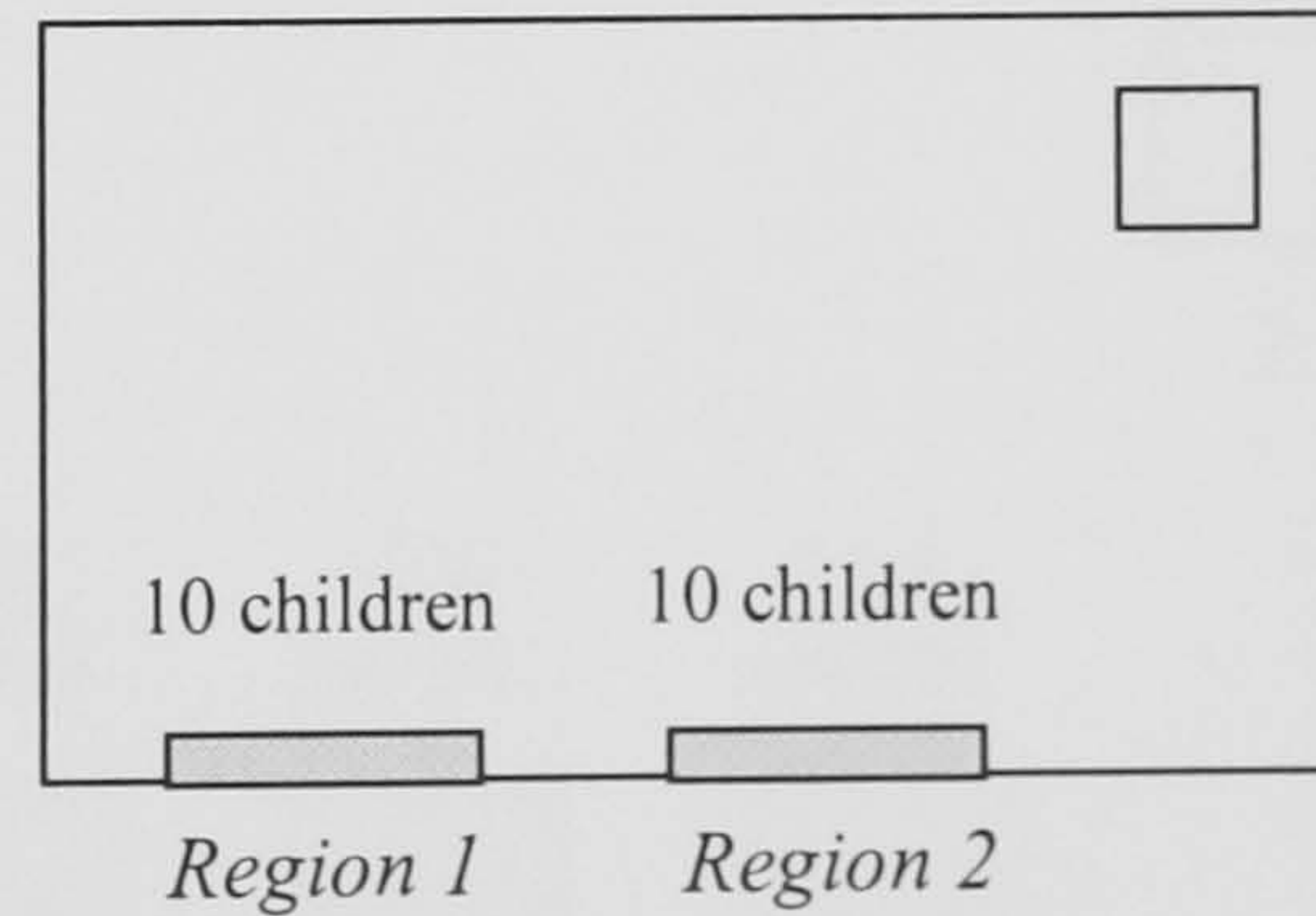
or



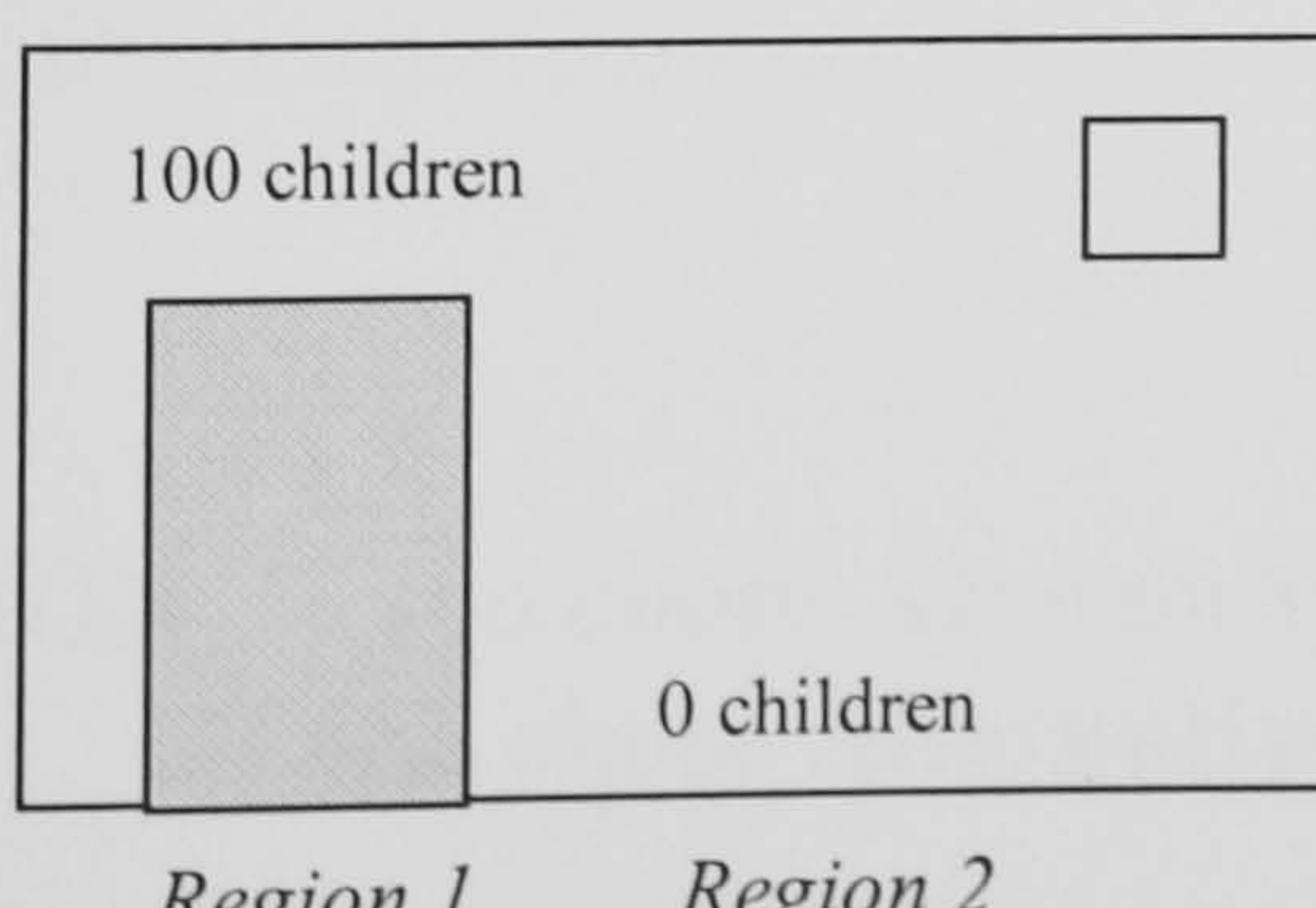
4.



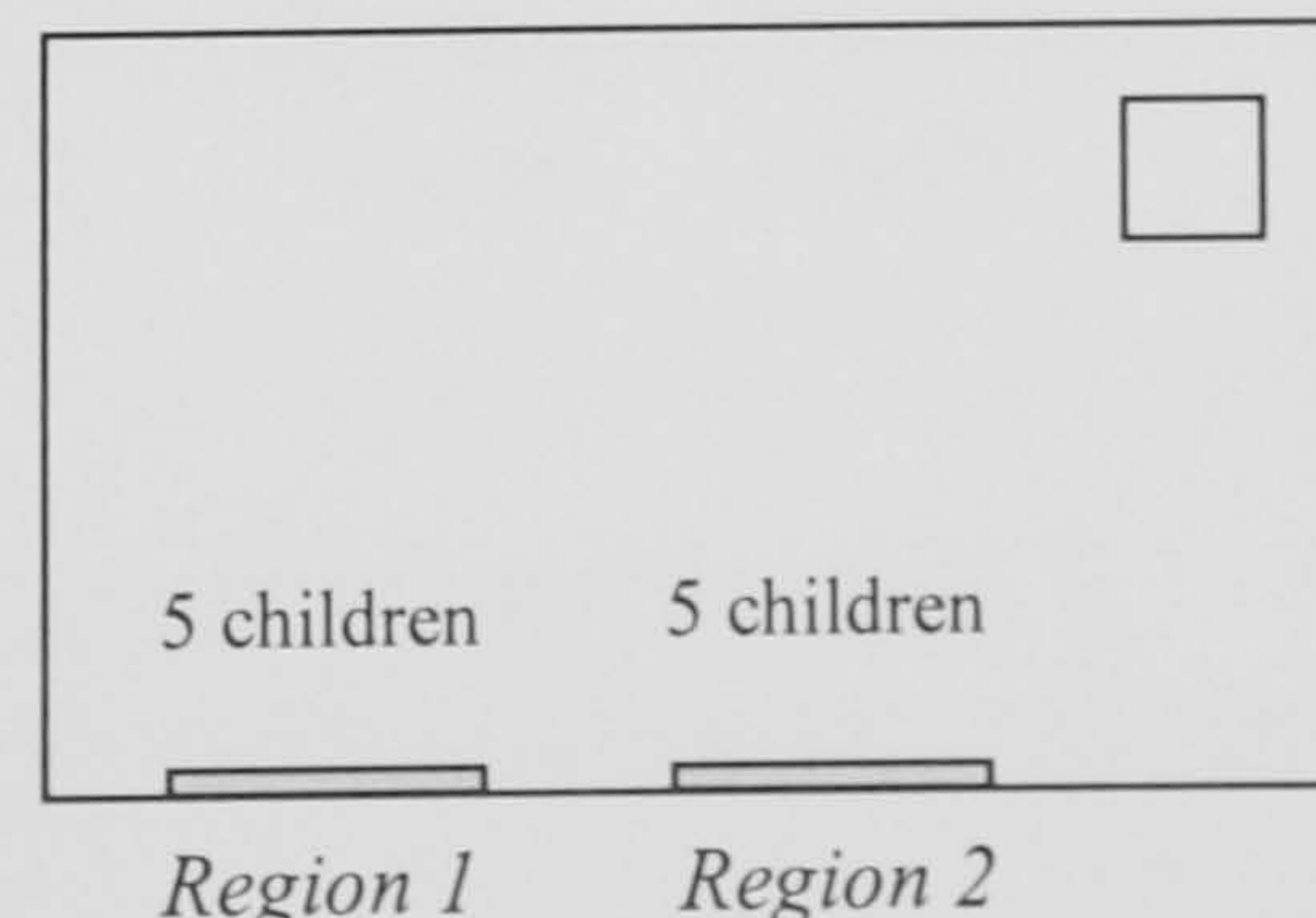
or



5.



or





Question 4

Based on 2006 data, the average waiting time for a surgery, in Portugal, in the three regions where there are more people waiting (North, Centre and Lisbon and Tagus Valley) is about 7 months and the number of people registered in waiting lists is about 70000 in each of these regions.

Suppose that the Department of Health decides to put extra money in the next year's health care budget to tackle waiting lists for a given surgery (for example, hip replacement or cataract surgery). The objective is to reduce the number of people waiting beyond the clinically acceptable time.

Imagine that there are two alternative ways to distribute this amount of extra resources:

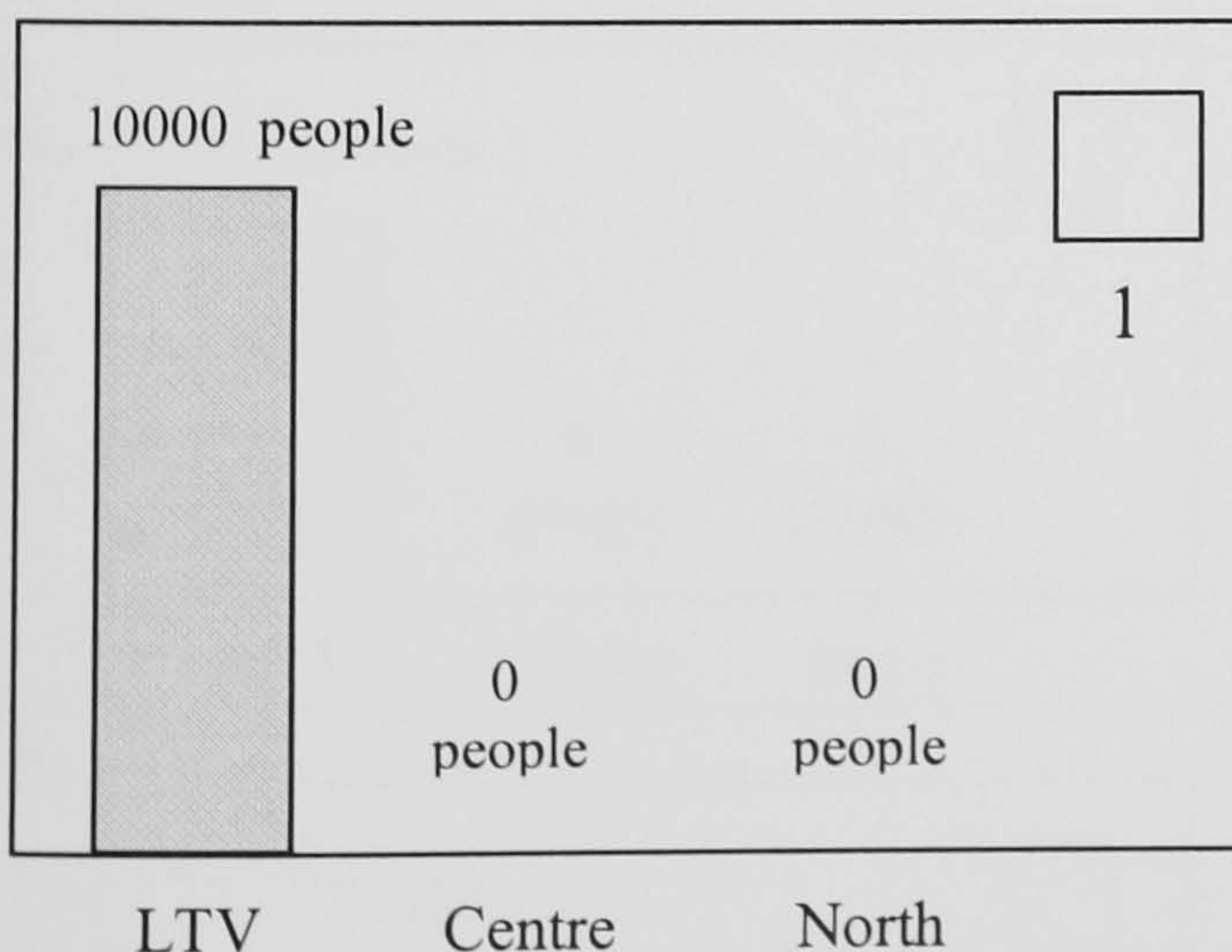
- In **Alternative A**, resources are concentrated on Lisbon and Tagus Valley (LTV);
- In **Alternative B**, resources are equally divided among the three regions (North, Centre and Lisbon and Tagus Valley)

Because the number of specialized teams is greater in Lisbon and Tagus Valley than in the other two regions, Alternative A leads to a larger reduction in the total number of people waiting for surgery.

Please, indicate which alternative, A or B, you would choose, crossing  the corresponding box:

**Alternative A**

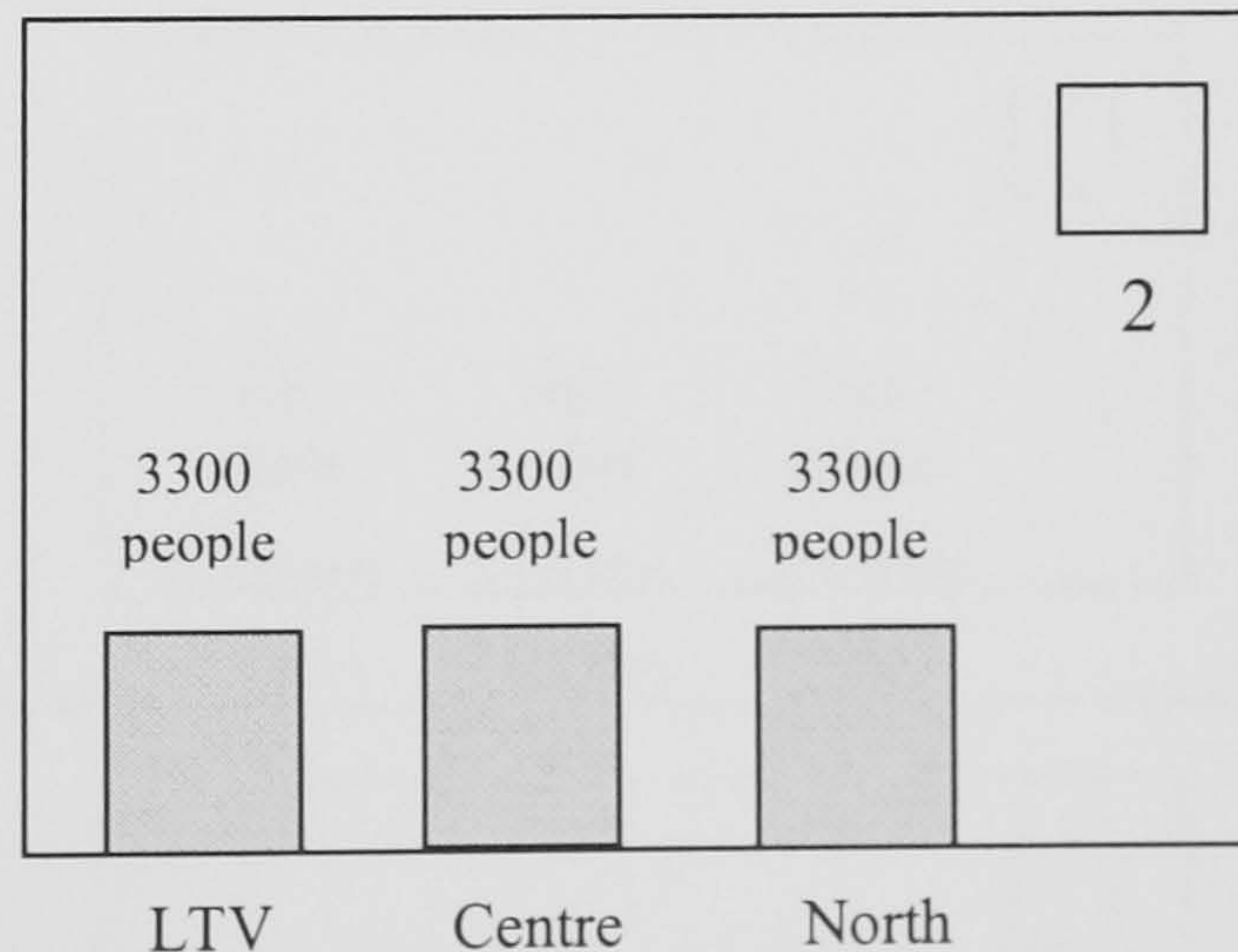
Total number of people having a surgery



or

**Alternative B**

Total number of people having a surgery



**ATTENTION:** If you chose **Alternative A**, please turn to **question 5** in page 7.

If you chose **Alternative B**, please continue on the **following** page (p.6)

Suppose now that the total number of people having a surgery under Alternative B is lower than previously stated.

For each of the five choices bellow, please cross  one box, indicating whether you would still choose Alternative B, or whether you would now choose Alternative A:

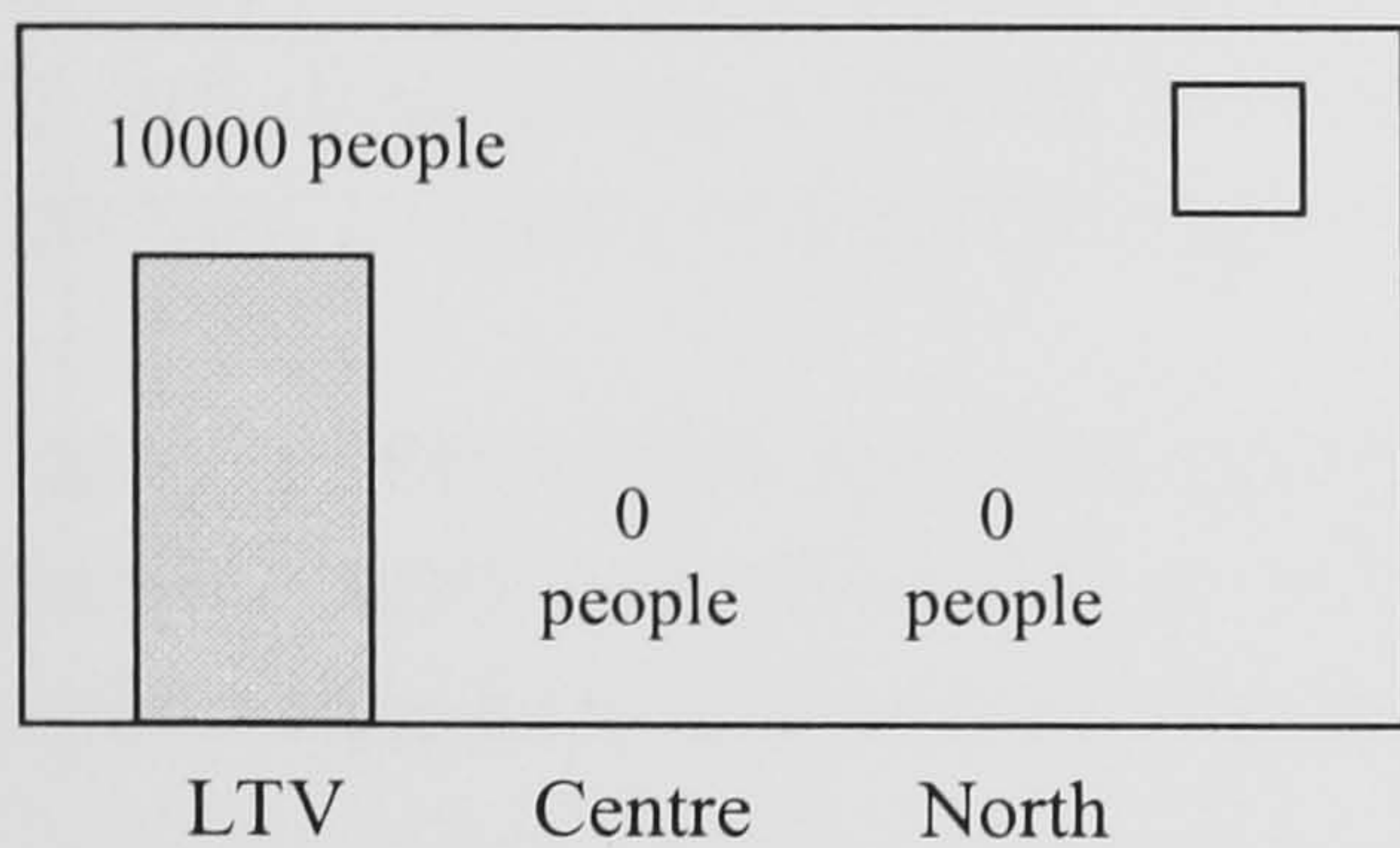
**Alternative A**

Total number of people having a surgery

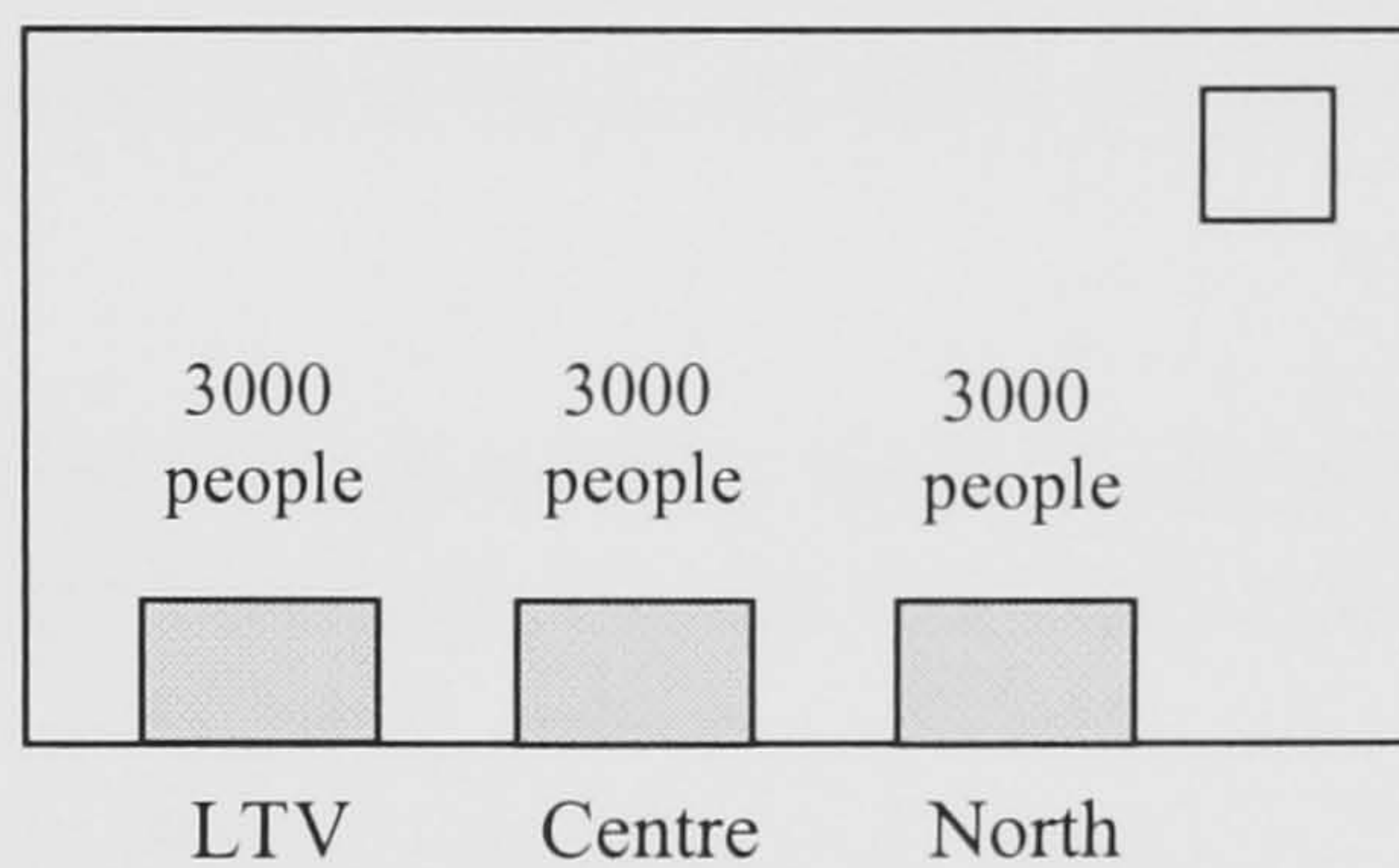
**Alternative B**

Total number of people having a surgery

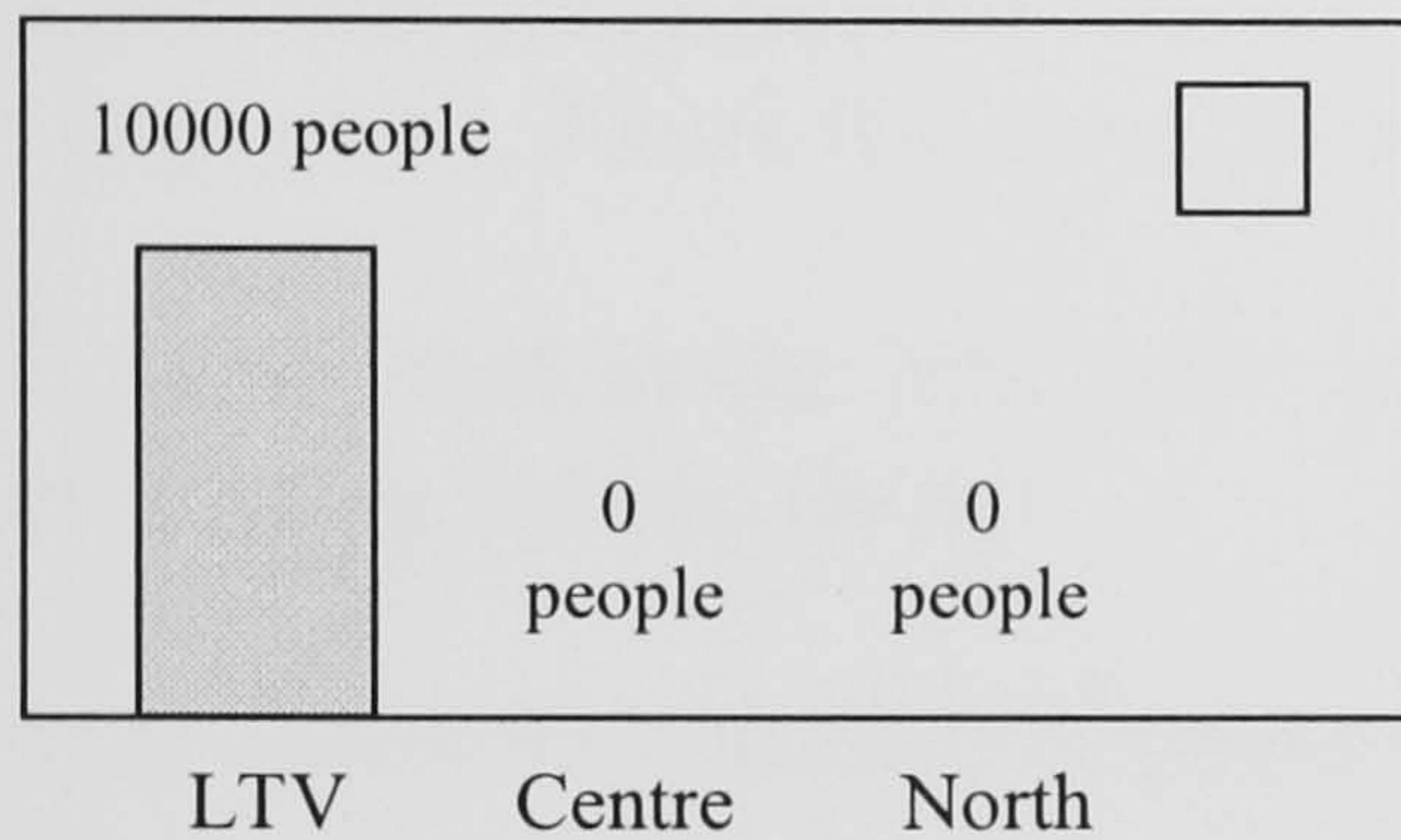
1.



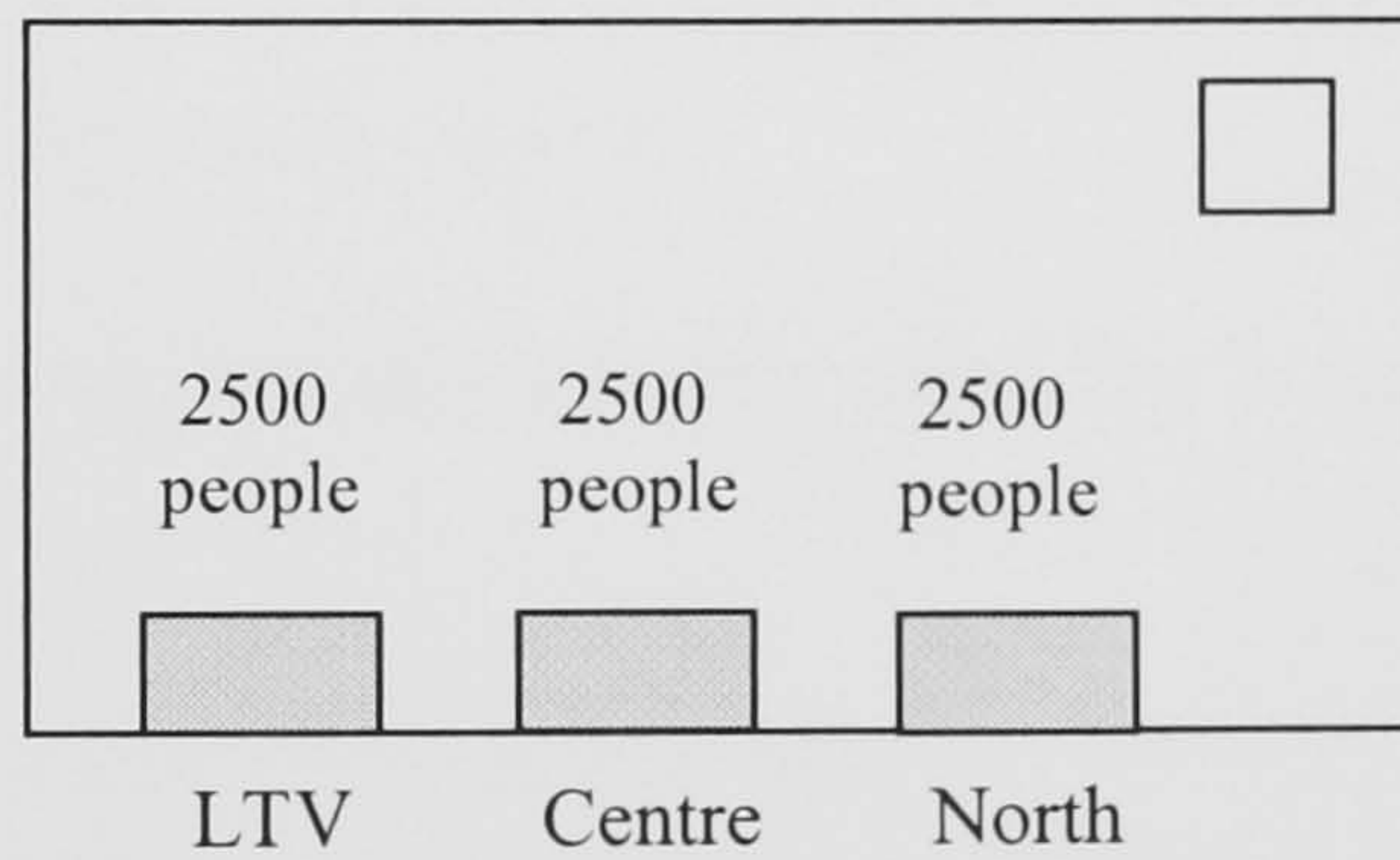
or



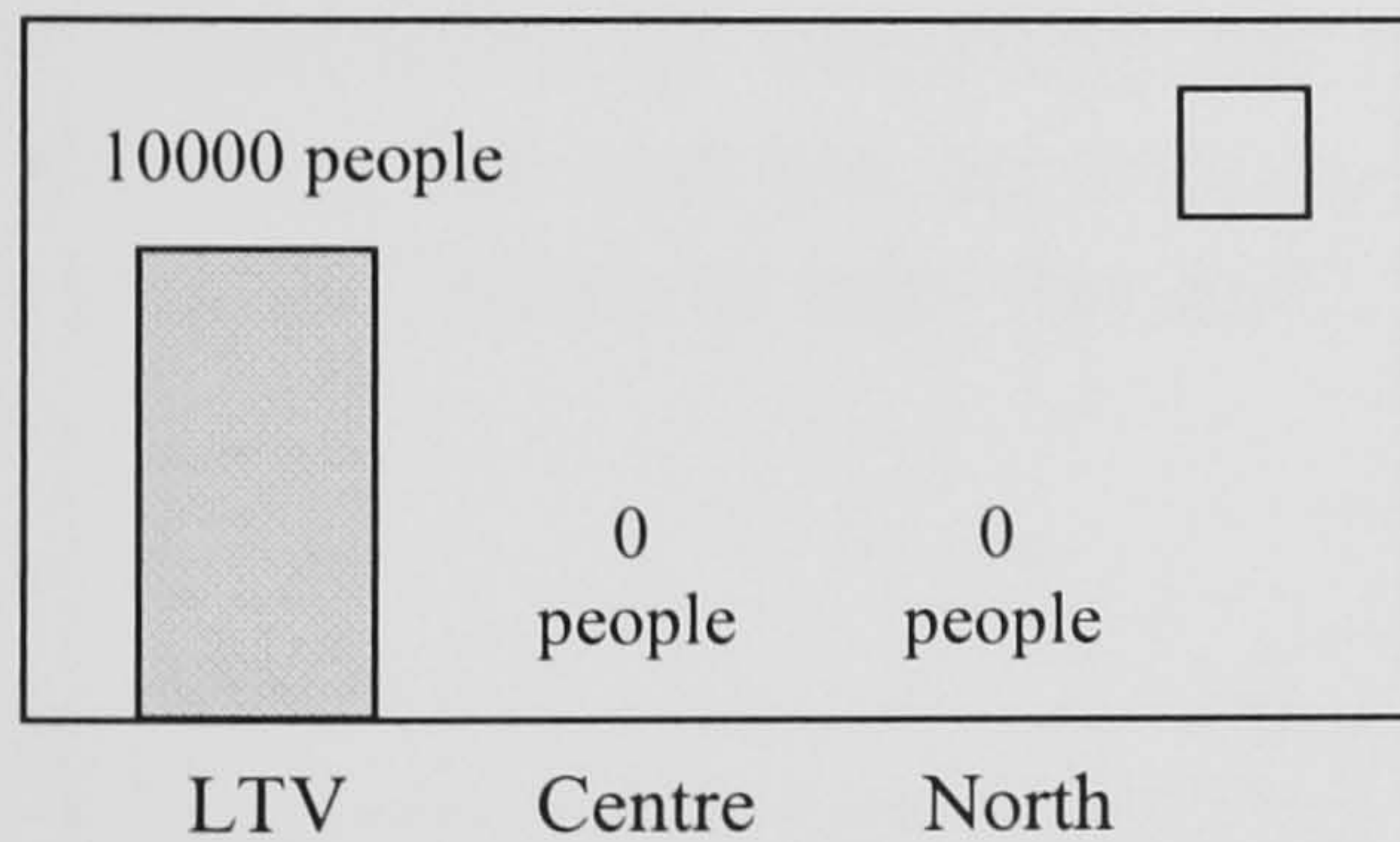
2.



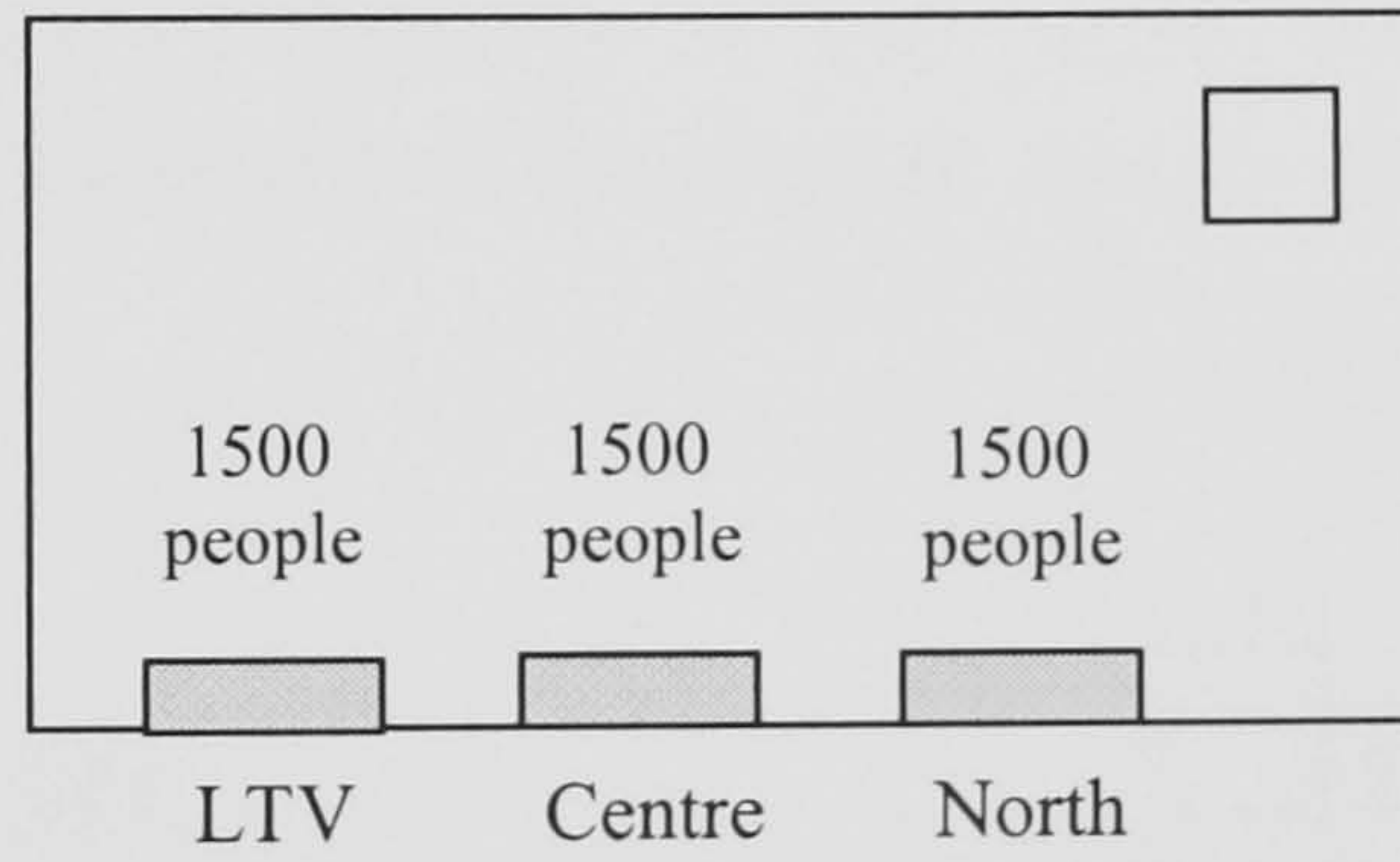
or



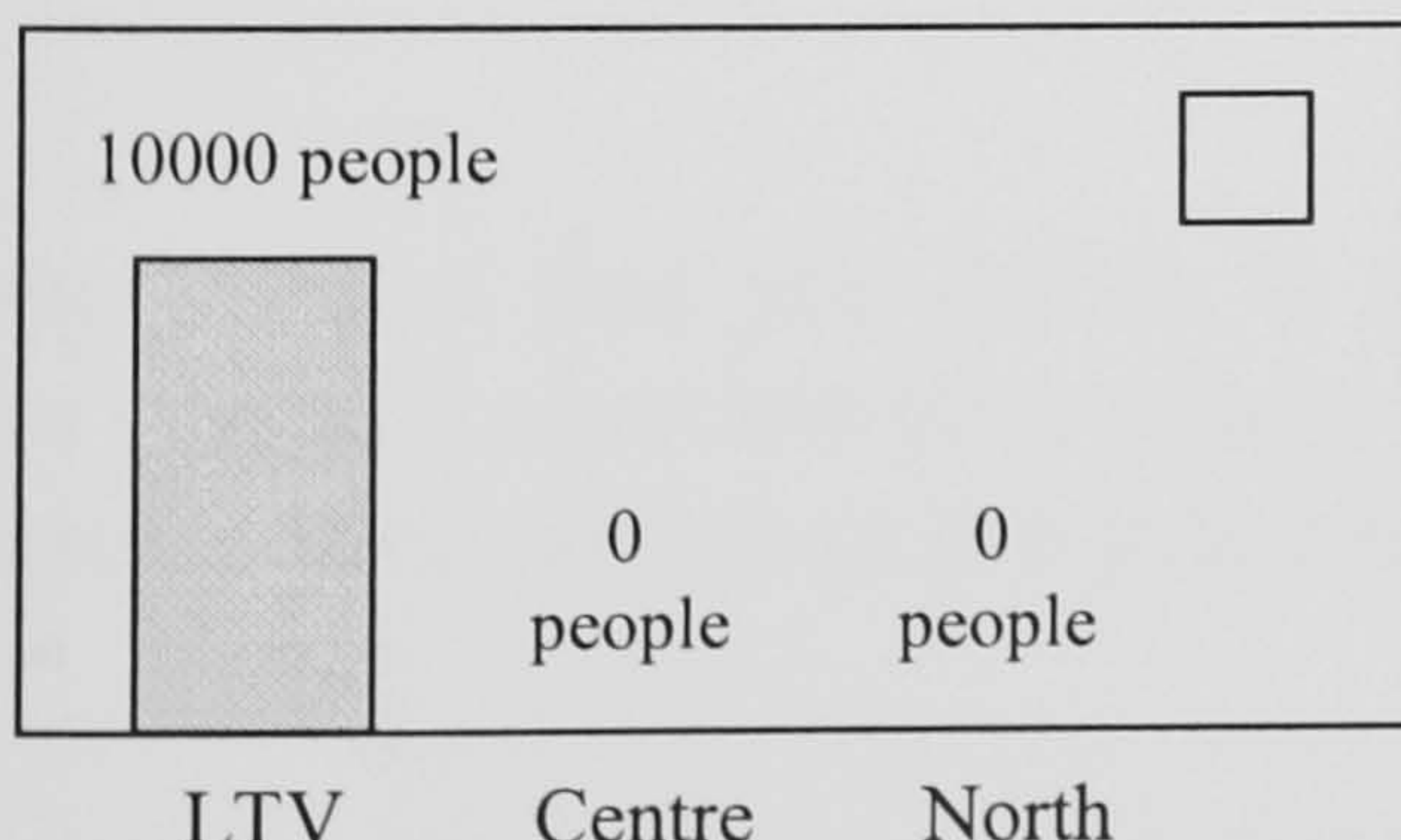
3.



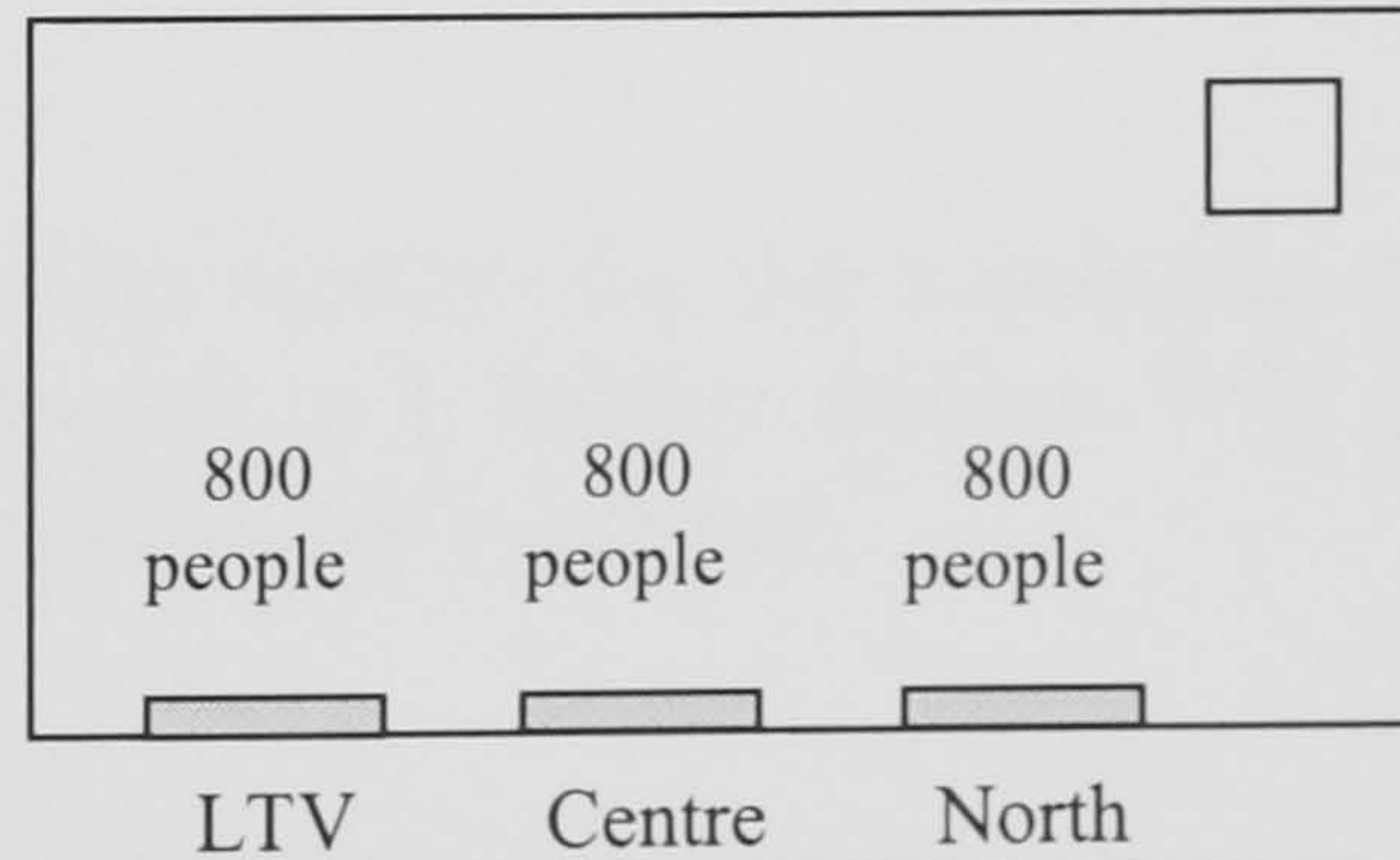
or



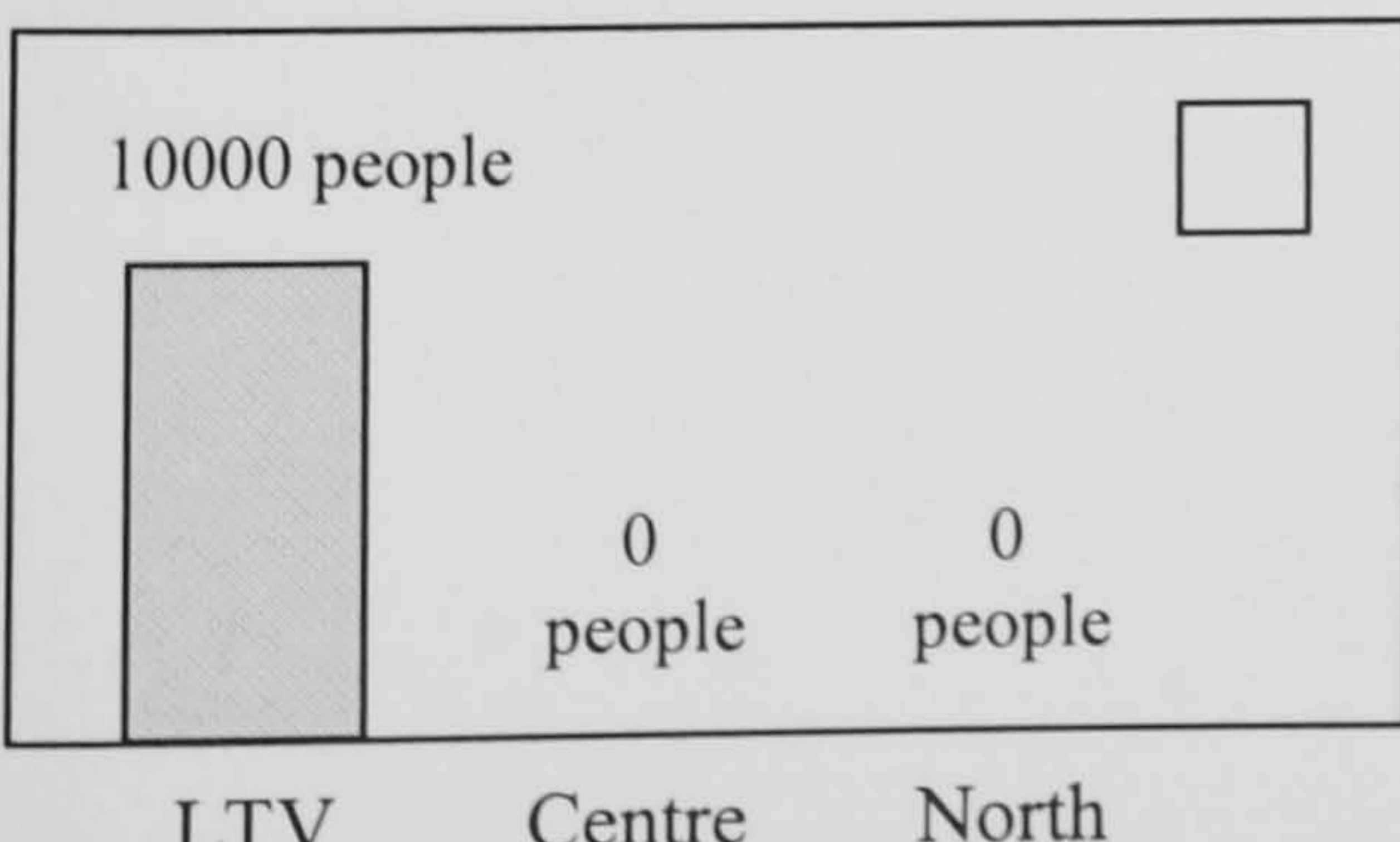
4.



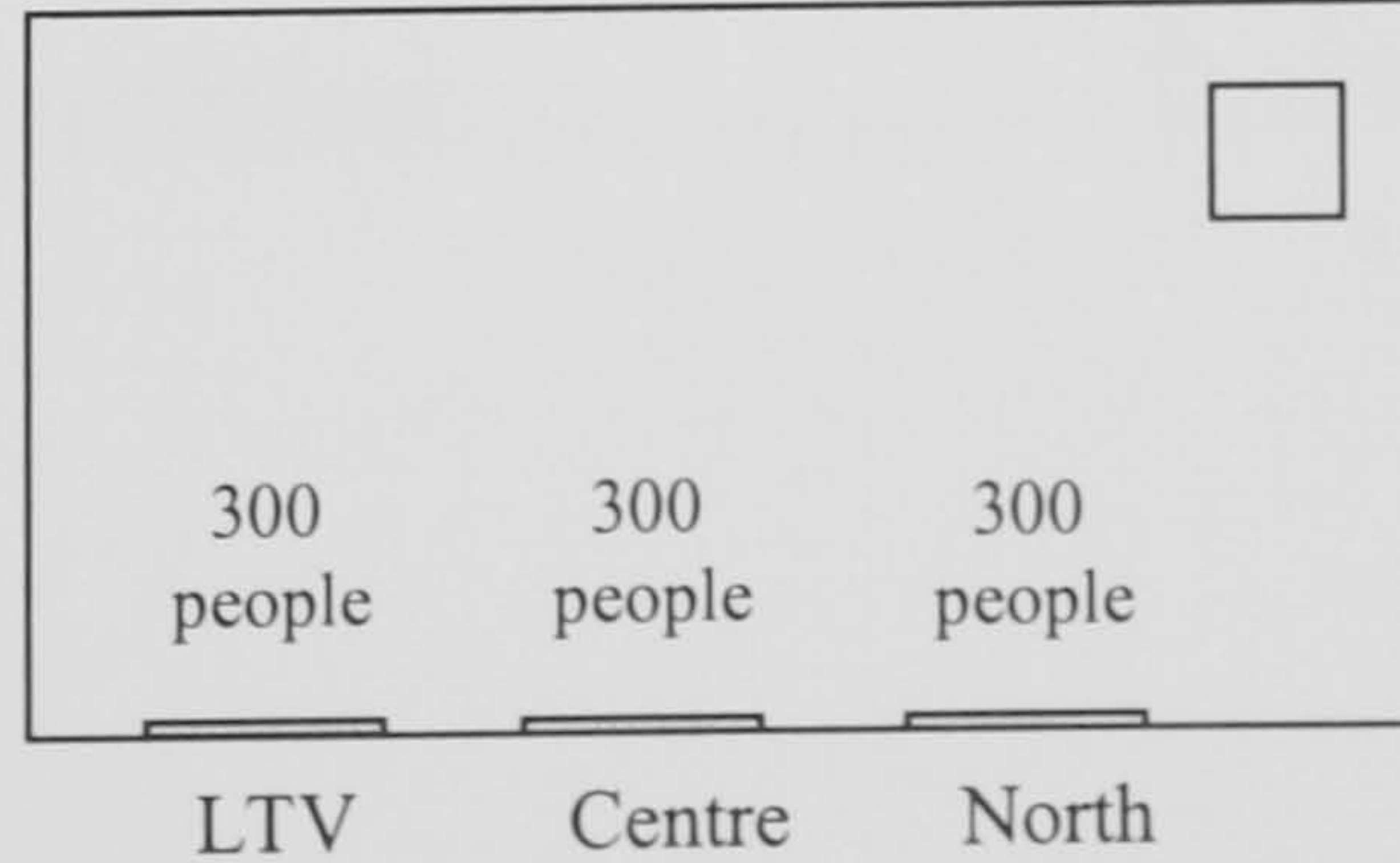
or



5.



or



In the two following questions, we are asking you about your personal (private) preferences.

**Question 5**

There are many diseases, like cancer, that have a greater probability of cure if they are diagnosed at an early stage. To diagnose diseases at an early stage it is important to carry out screening programmes.

In our country there are few systematic screening programmes. Still, even when screening tests are offered free of cost within the National Health Service, not all people come forward to take the test.

Imagine that you received a letter from your Health Centre inviting you to take a screening test. Think for example about a test for colon cancer.

It is mentioned in the letter that you may choose between being seen by a male doctor and a female doctor.

For the following questions, please cross  the corresponding box:

The possibility of choosing the gender (male or female) of the professional performing the test has an impact on your decision about whether or not to go to the Health Centre to take the test?

YES.....  1                      NO.....  2

Suppose now that you decide to go to the Health Centre for the screening test. You can still exercise your choice of a male doctor or a female doctor. Will you exercise your choice or are you indifferent to the doctor's gender?

I choose according to my preferences.....  1                      I am indifferent.....  2

## Question 6

Different people have different attitudes towards the same health problems. There are, for example, some vulgar symptoms like headache or back pain that, in most cases, do not lead people to seek personalised medical treatment.

We would like to learn about your attitude / preferences in the hypothetical situations described below. Think for instance about the cases of headache or back pain.

For each of the five hypotheses, please cross  the box that in your view matches or would match your preferred form of action most closely:

	I do nothing	I do self-medication	I ask advice to pharmacist	I visit the doctor
1. I feel strong pain occasionally	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. I feel strong pain frequently but I carry out my normal life	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. I feel strong pain frequently and although I carry out my normal life during the day, I have sleeping difficulties overnight because of pain	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. I feel strong pain frequently, I do not sleep well and sometimes I have to interrupt my daily activities because of pain	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. I feel strong pain frequently, I do not sleep well and sometimes I am even forced to miss work / school because of pain	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Finally, we are asking you about some personal data (gender, age and level of education) for statistical purposes only.**

Please, fill or cross  the corresponding box:

**1. Gender:**

Male.....  1      Female.....  2

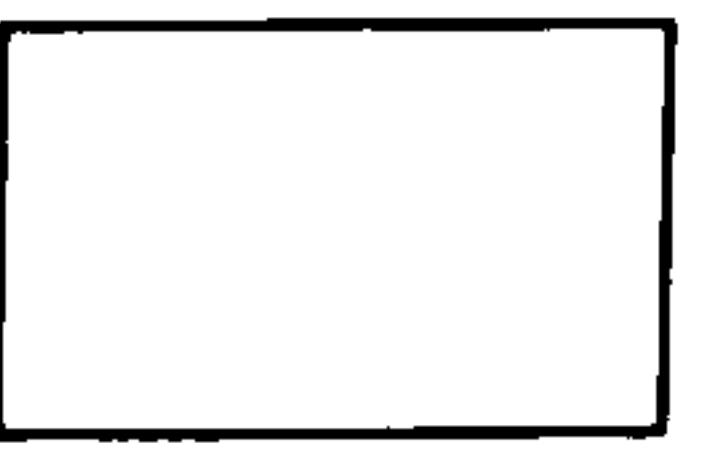
**2. Age: \_\_\_\_\_ year old.**

**3. Education - please indicate the highest degree of education that you have successfully completed. If you are still a student (any degree) please select the box "Attending":**

None.....	<input type="checkbox"/> 1	Secondary .....	<input type="checkbox"/> 3
1st Cycle .....	<input type="checkbox"/> 2	Medium .....	<input type="checkbox"/> 4
2nd Cycle .....	<input type="checkbox"/> 2	University .....	<input type="checkbox"/> 4
3rd Cycle .....	<input type="checkbox"/> 2	Attending.....	<input type="checkbox"/> 5

<p>Thank you very much for completing this questionnaire. Your contribution is valuable to our research.</p>
--

**APPENDIX C – Questionnaire: Portuguese version**



## Questionário

### Como devem os recursos do Ministério da Saúde ser distribuídos?

- Este questionário destina-se aos cidadãos em geral.
- Devido à falta de recursos, as decisões sobre como distribuir a parte destinada aos cuidados de saúde são inevitáveis - É importante que os responsáveis políticos saibam o que é que os cidadãos desejam do seu Serviço Nacional de Saúde.
- Não há respostas certas ou erradas; os investigadores estão apenas interessados na **sua opinião**.
- As respostas são de carácter **anónimo**.
- Este questionário é composto por seis questões que deverão levar cerca de **10-15 minutos** a preencher.

Muito obrigada pela sua colaboração.

Se tiver alguma dúvida sobre este questionário, por favor, contacte:

Carlota Quintal

Faculdade de Economia da Universidade de Coimbra

Av. Dias da Silva, 165, 3004-512 Coimbra

Tel: 239790580

Email: [qcarlota@fe.uc.pt](mailto:qcarlota@fe.uc.pt)

**Nas duas questões seguintes dê-nos a sua opinião sobre dois critérios que considere poderem vir a orientar a distribuição de recursos.**

Para cada uma das questões seguintes, faça um círculo ○ em redor do número que melhor reflecte a sua opinião:

**Questão 1**

Há cuidados de saúde que têm maior impacto sobre o estado de saúde das pessoas do que outros. Em alguns casos, este impacto significa melhorar o estado de saúde das pessoas; em outros casos, significa atenuar a deterioração do seu estado de saúde.

Concorda que ao distribuir os recursos disponíveis para os cuidados de saúde é importante ter em conta, entre outros possíveis critérios, também o seu impacto sobre o estado de saúde das pessoas?

1	2	3	4	5
Discordo Plenamente	Discordo	Nem concordo Nem discordo (neutro)	Concordo	Concordo Plenamente

**Questão 2**

A distribuição de recursos envolve a sua distribuição pelos diferentes cuidados mas tem igualmente uma componente geográfica, envolvendo a distribuição pelas diferentes regiões do país.

Concorda que ao distribuir os recursos disponíveis para os cuidados de saúde é importante ter em conta, entre outros possíveis critérios, também como é que esses recursos são distribuídos pelas várias regiões do país?

1	2	3	4	5
Discordo Plenamente	Discordo	Nem concordo Nem discordo (neutro)	Concordo	Concordo Plenamente



Nas duas questões seguintes dê-nos a sua opinião sobre a importância relativa dos dois critérios referidos nas questões 1 e 2.

**Questão 3**

Suponha que na sequência de uma descoberta científica, surgiu uma vacina destinada às crianças que as imuniza (protege) totalmente contra uma determinada doença. A **doença não é contagiosa** mas é fatal.

Sem a vacina, algumas crianças virão a contrair a doença e a morrer em consequência; outras, mesmo sem a vacina poderão não chegar a contrair a doença.

Há, no entanto, uma quantidade limitada da vacina para cada país e esta quantidade não é suficiente para vacinar todas as crianças. Há que tomar uma decisão sobre quem vacinar.

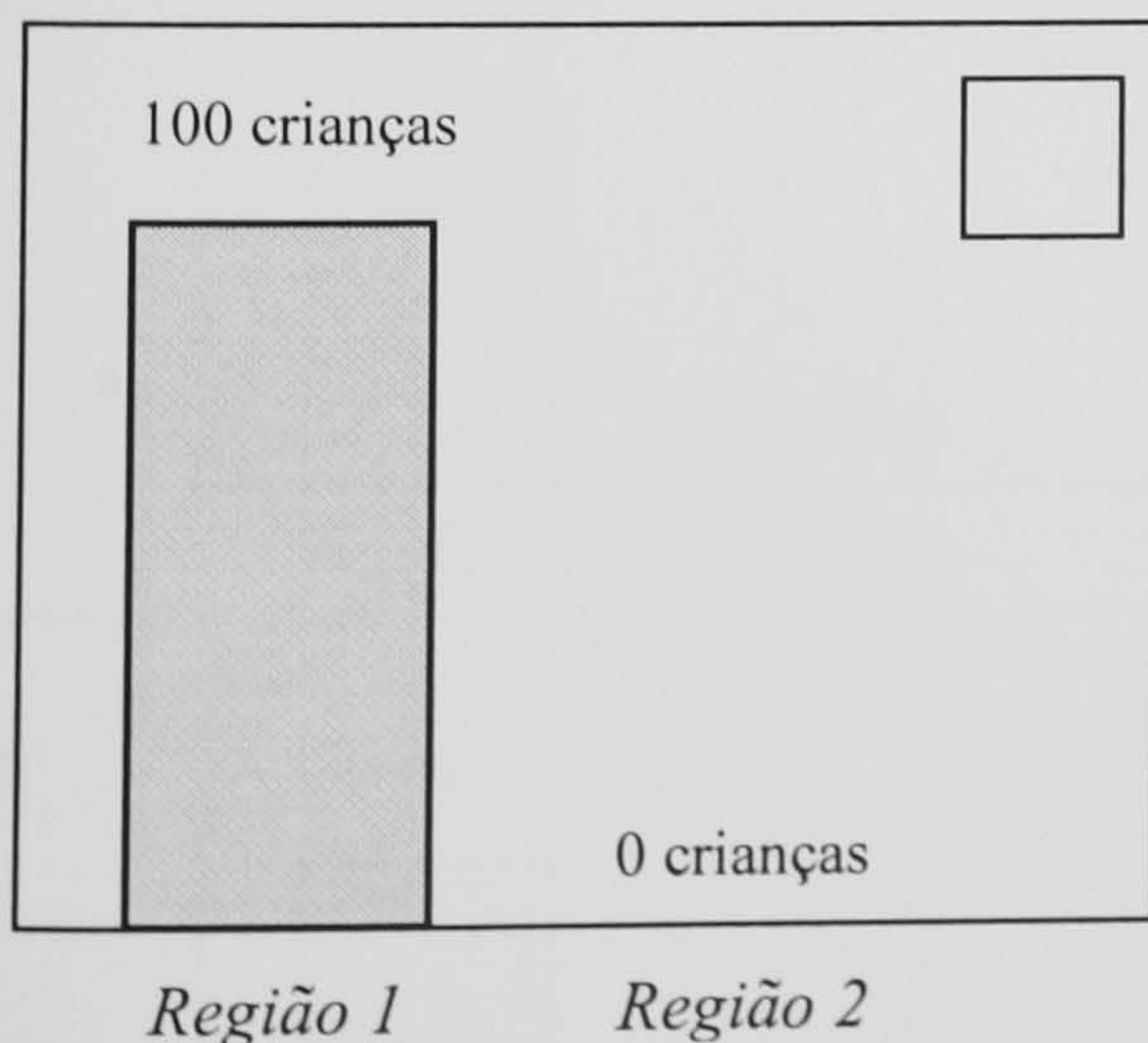
Suponha então que o país está dividido em duas regiões, *Região 1* e *Região 2*, em que metade das crianças habita na *Região 1* e a outra metade habita na *Região 2*.

Tem de escolher entre distribuir todas as doses da vacina pelos centros de saúde da *Região 1* (Cenário A) ou distribuir metade das doses pelos centros de saúde da *Região 1* e metade pelos centros de saúde da *Região 2* (Cenário B). Os dois cenários envolvem a mesma redução do número total de casos de doença mas diferem quanto à distribuição geográfica dos casos de doença evitados.

Por favor, indique qual dos cenários, A ou B, escolheria, marcando com uma cruz ☒ o quadrado correspondente:

**Cenário A**

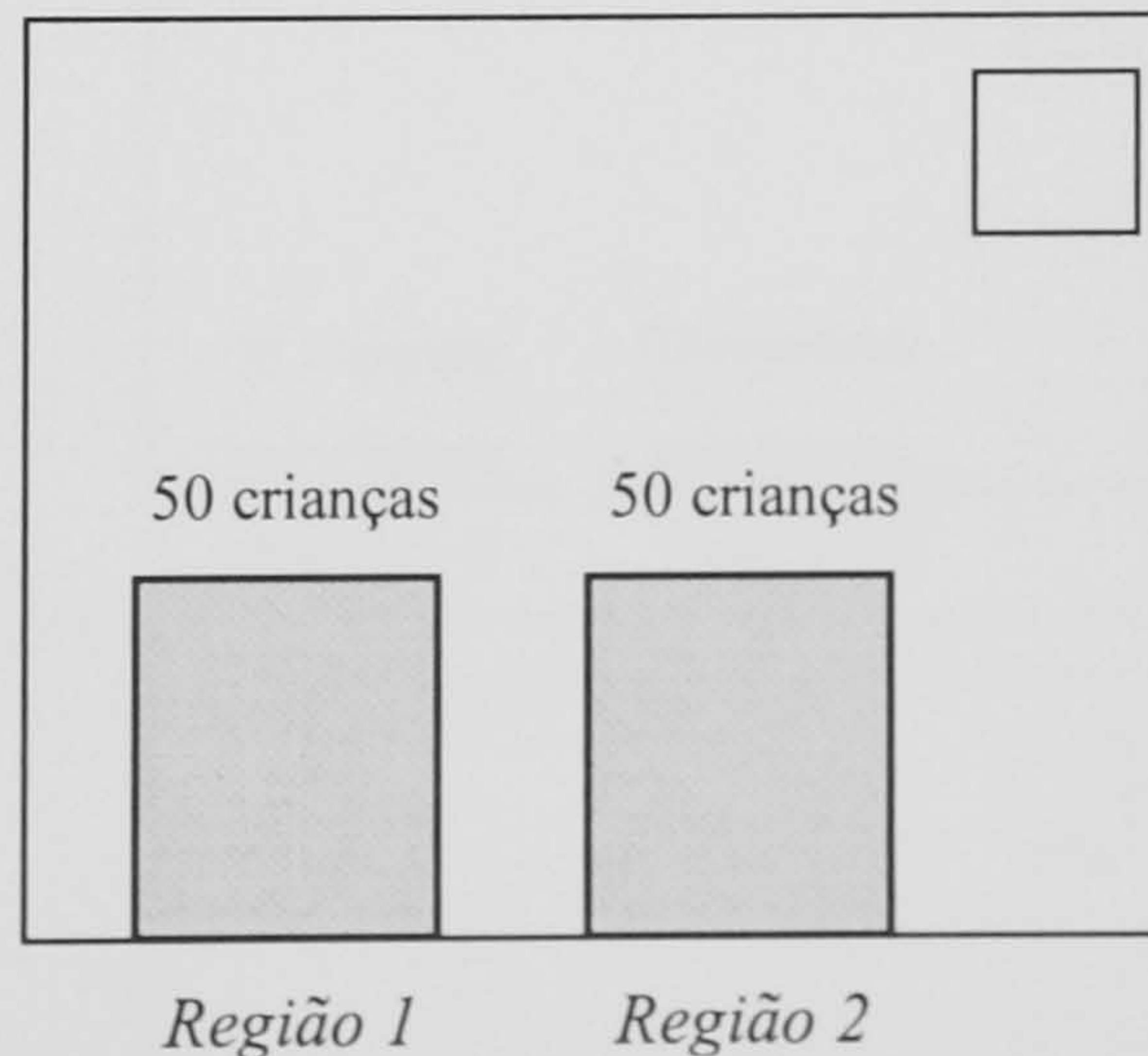
Número de crianças que irão evitar a doença



ou

**Cenário B**

Número de crianças que irão evitar a doença



**ATENÇÃO:** Se escolheu o **Cenário A**, por favor vá para a **questão 4** da página **5**  
Se escolheu o **Cenário B**, por favor continue na página **seguinte** (p.4)

Admita agora que distribuir as vacinas pelas duas regiões (Cenário B) é menos eficaz do que concentrar a sua distribuição apenas na *Região 1* (Cenário A).

Para cada uma das cinco hipóteses em baixo, assinale com uma cruz um dos quadrados, indicando se continua a preferir o Cenário B ou se agora passaria a optar pelo Cenário A:

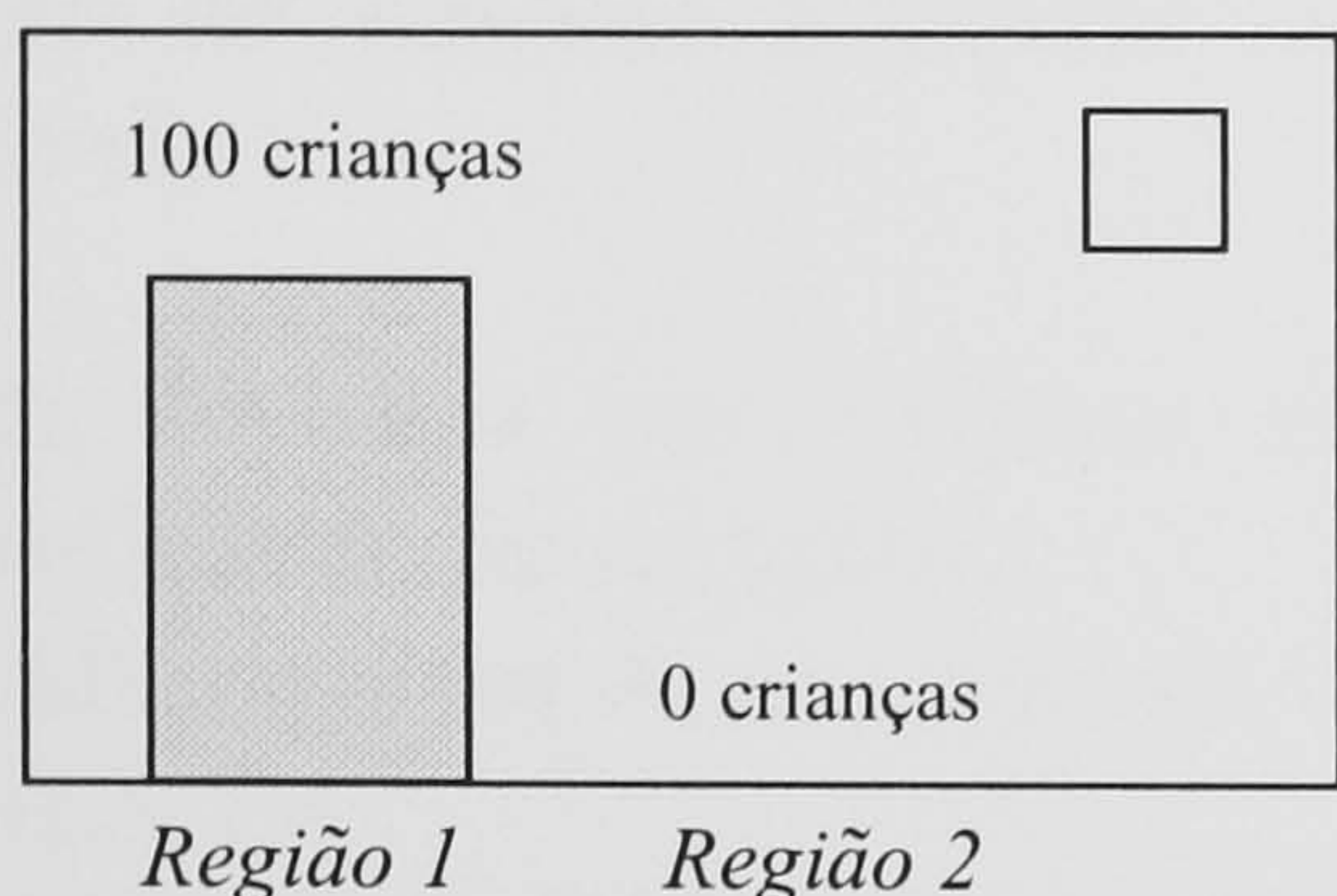
### Cenário A

Número de crianças que irão evitar a doença

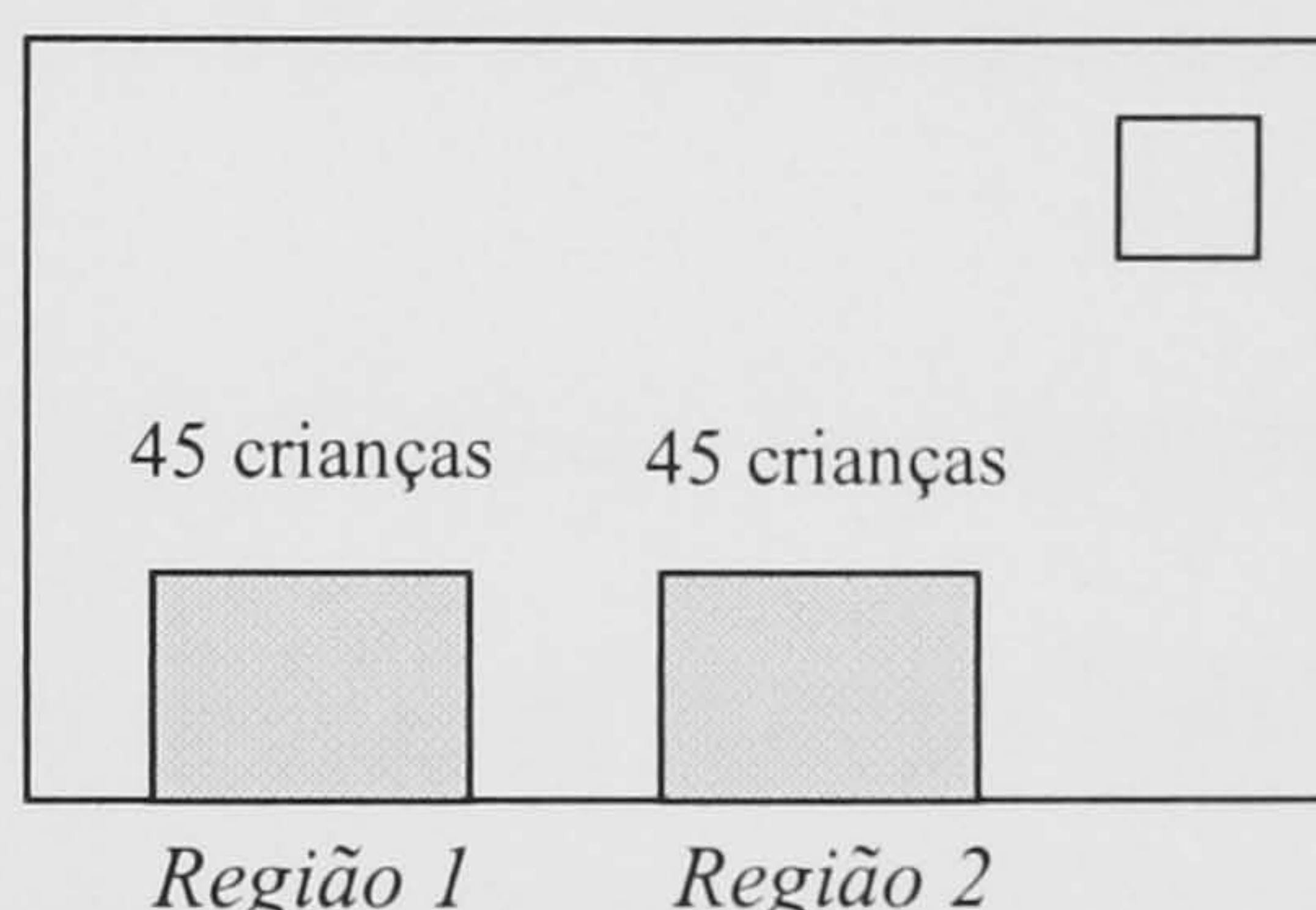
### Cenário B

Número de crianças que irão evitar a doença

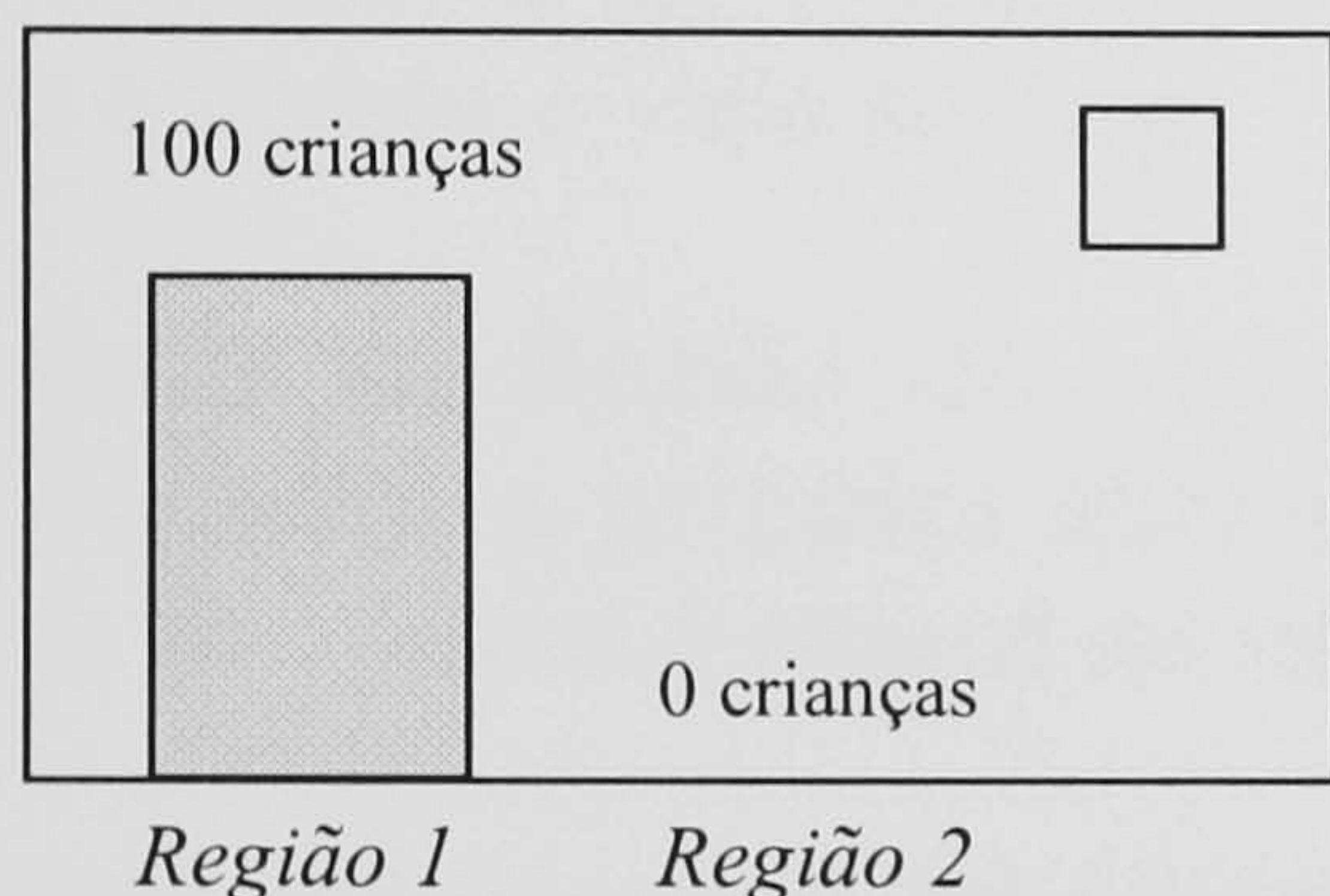
1.



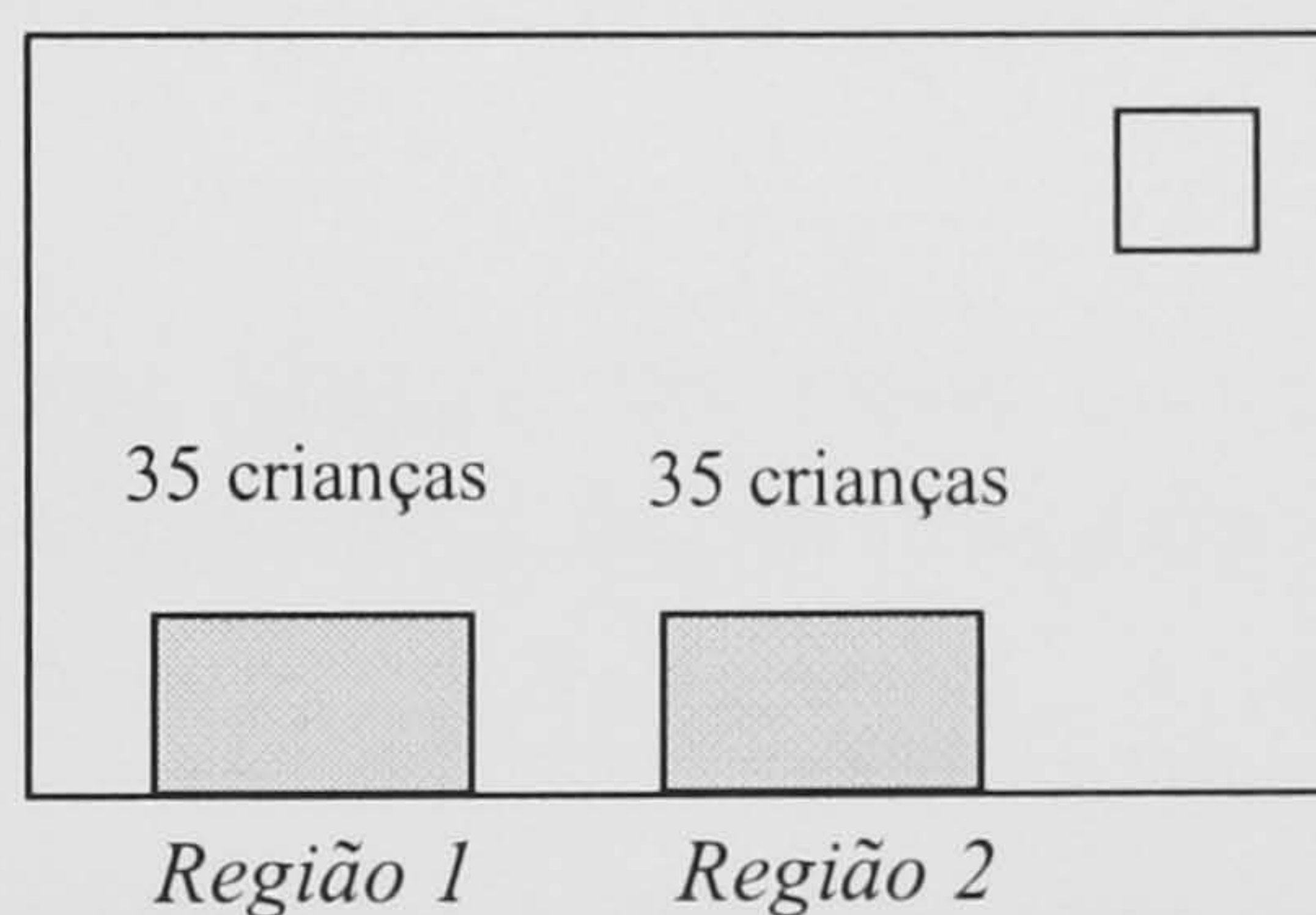
ou



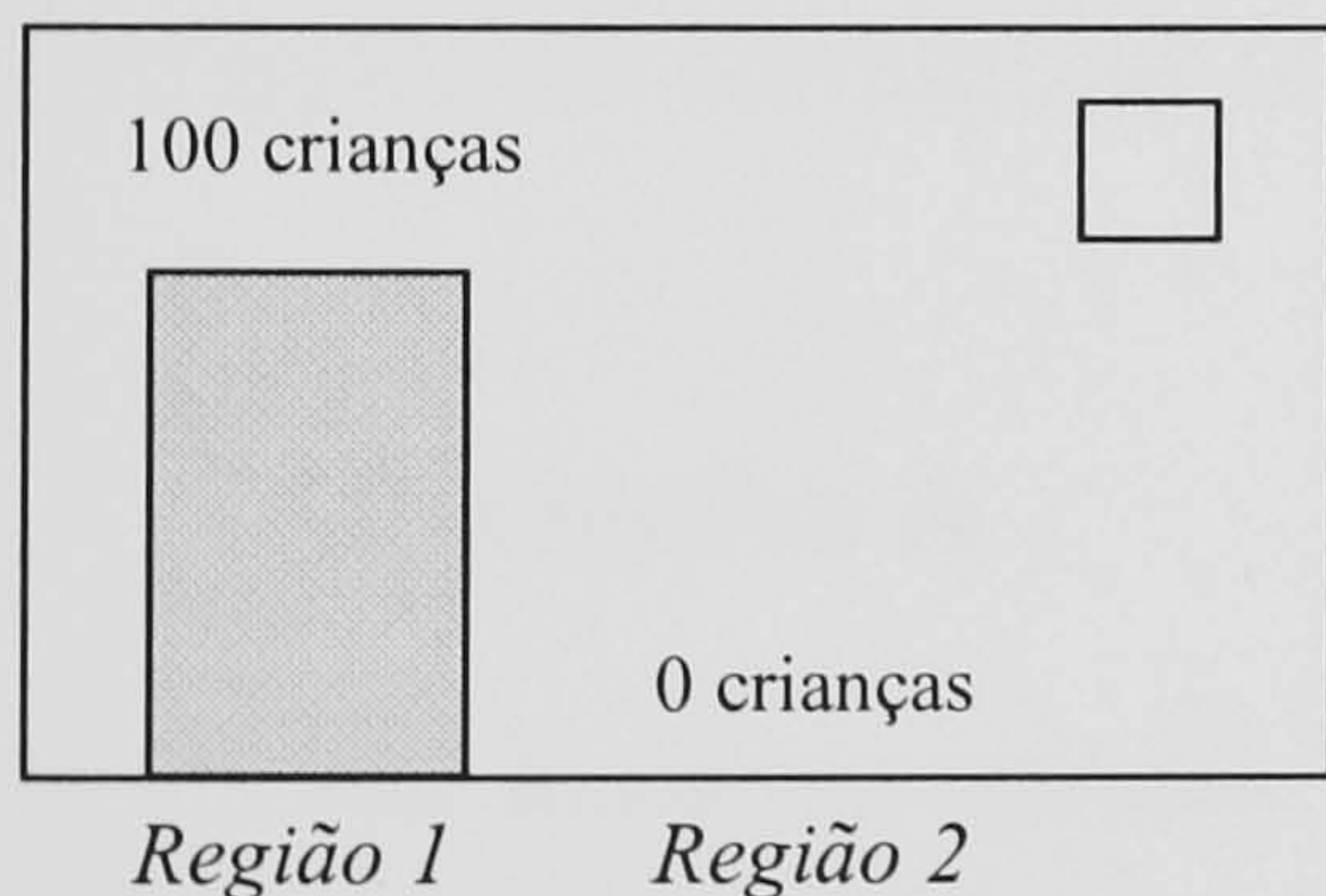
2.



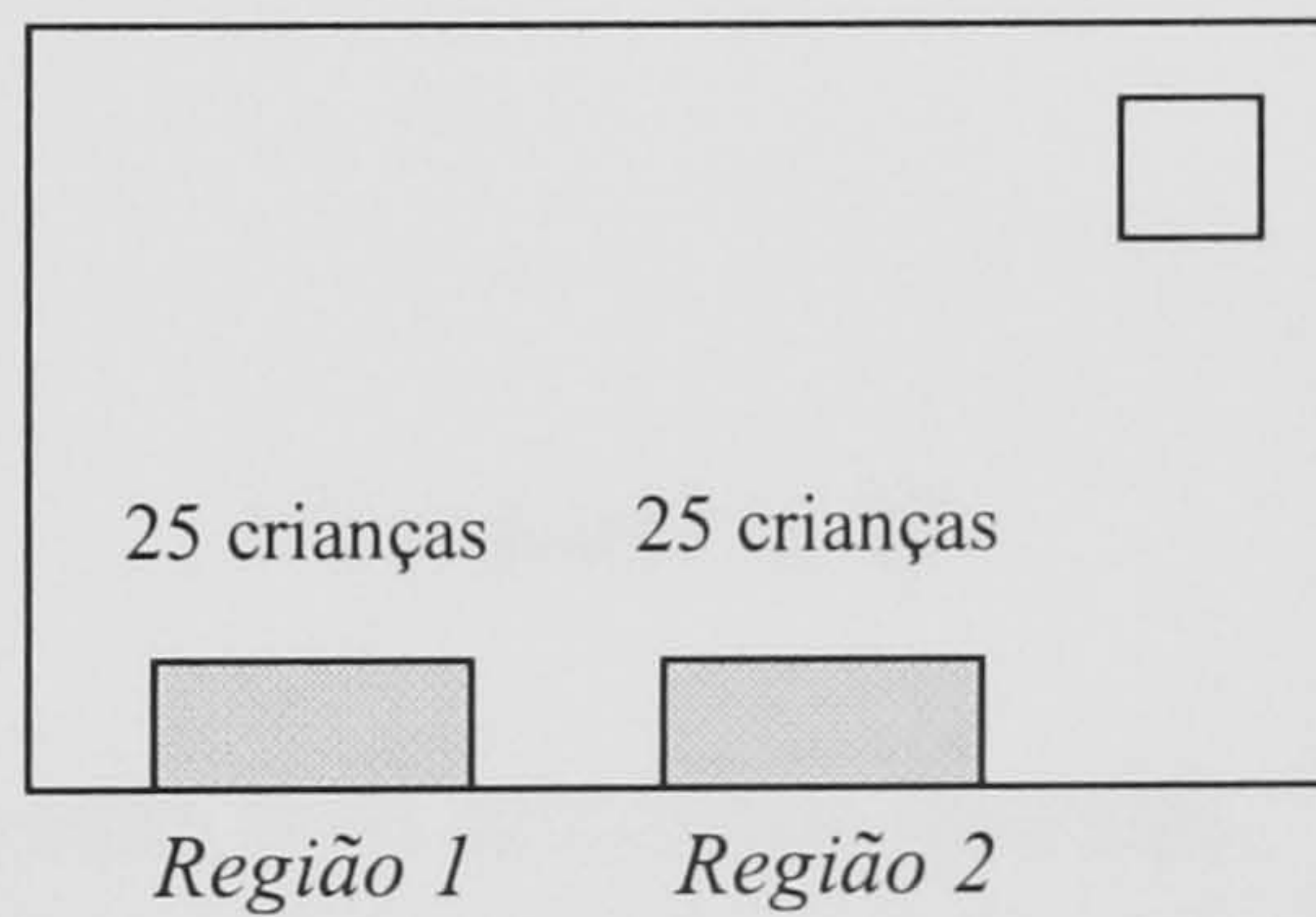
ou



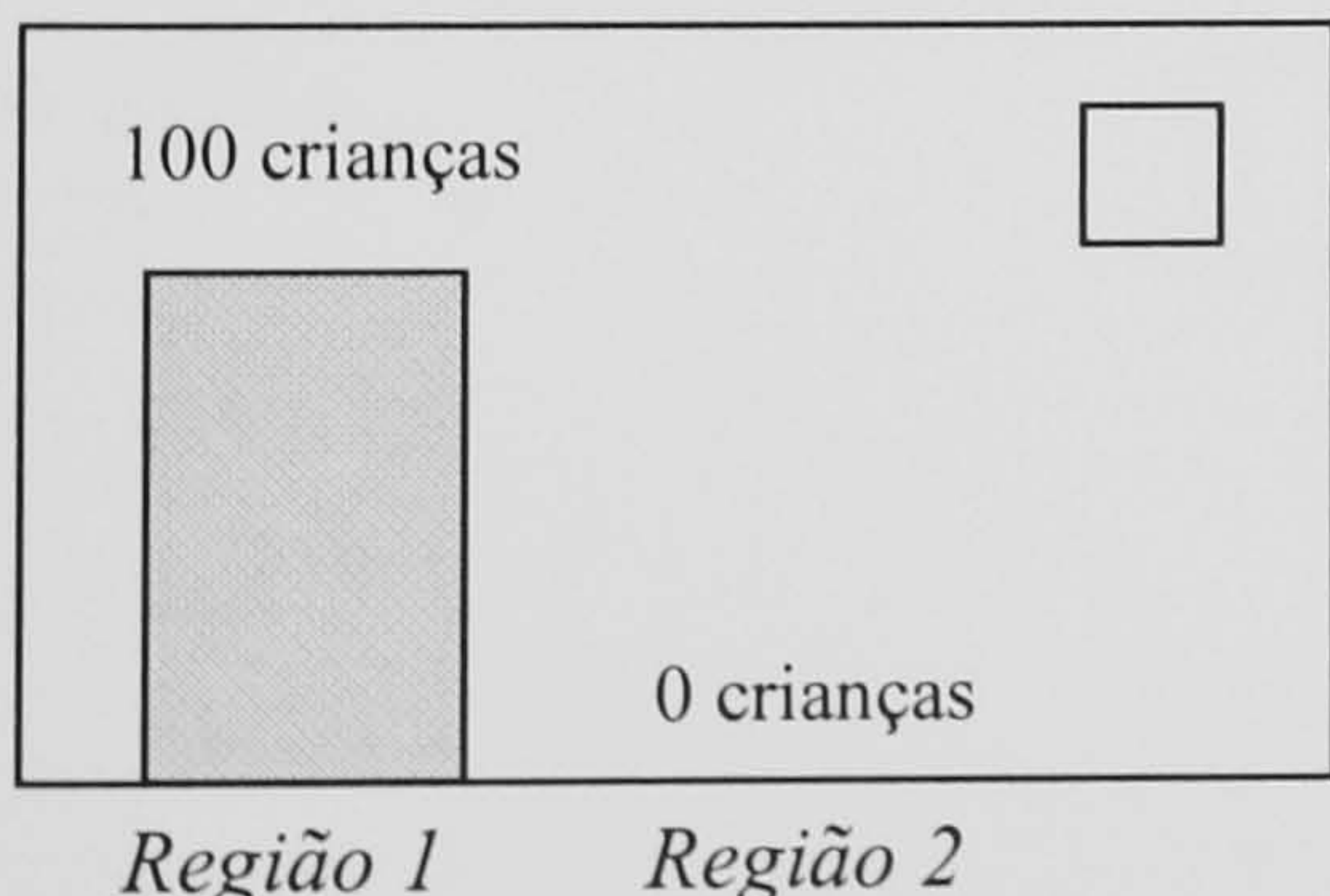
3.



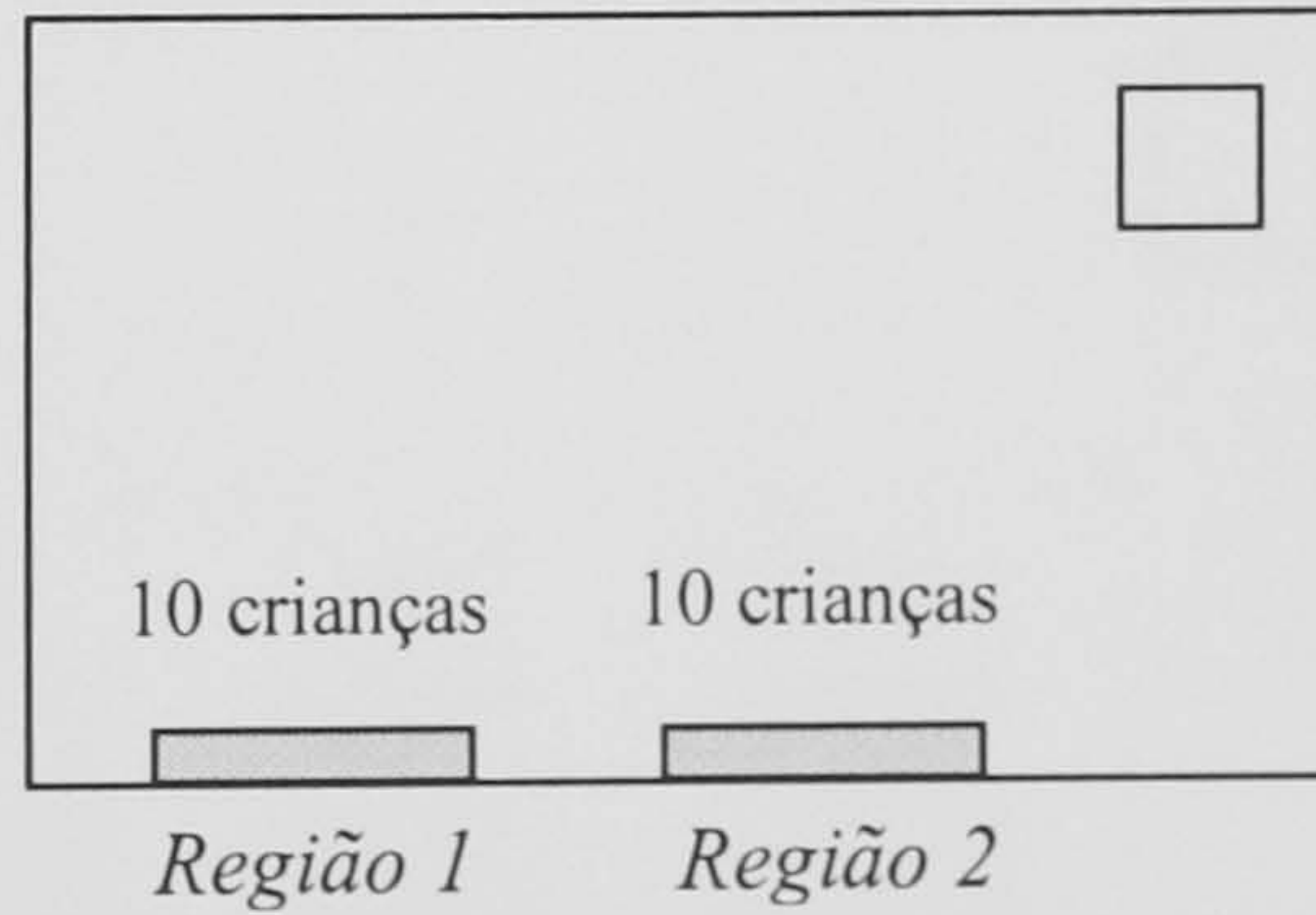
ou



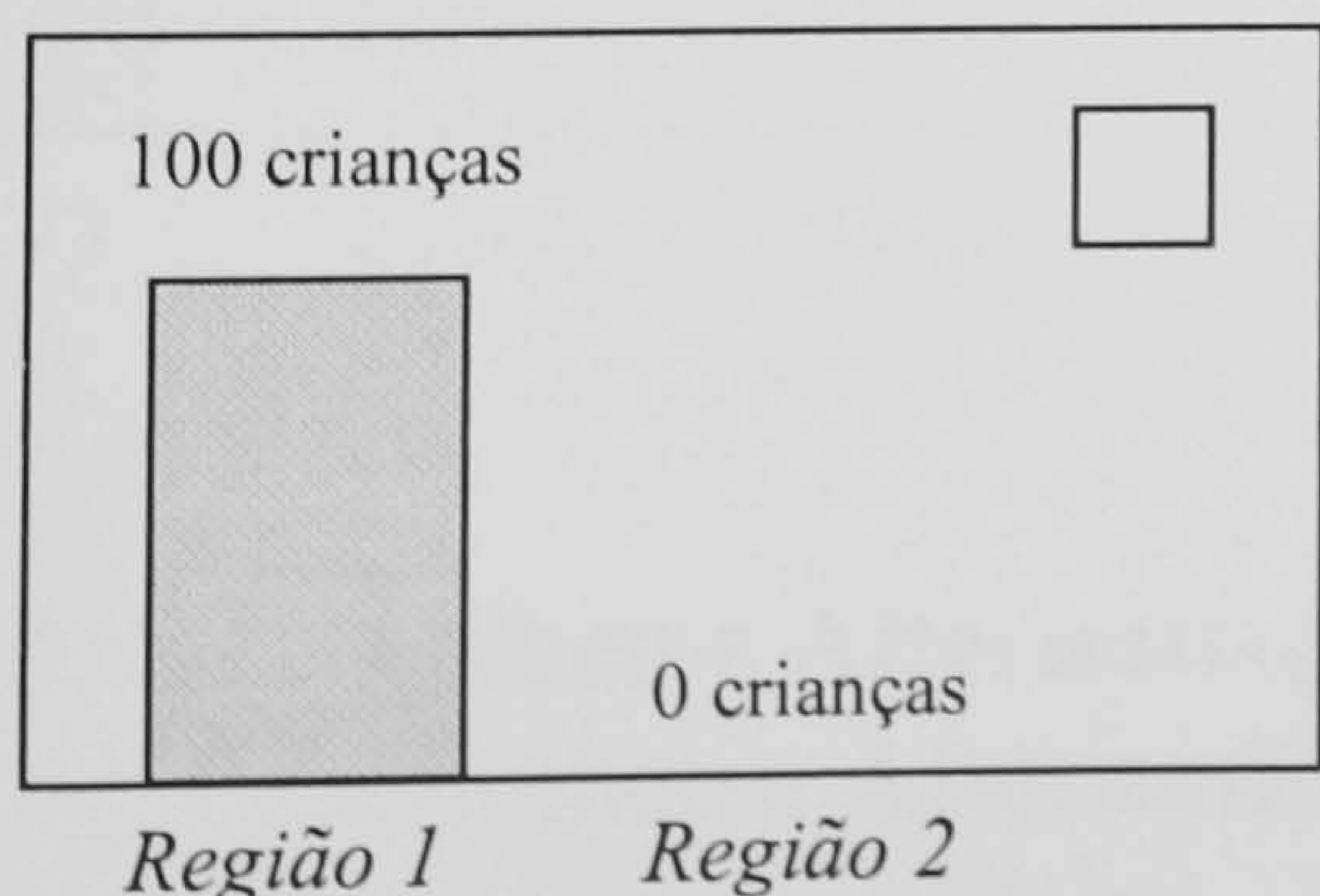
4.



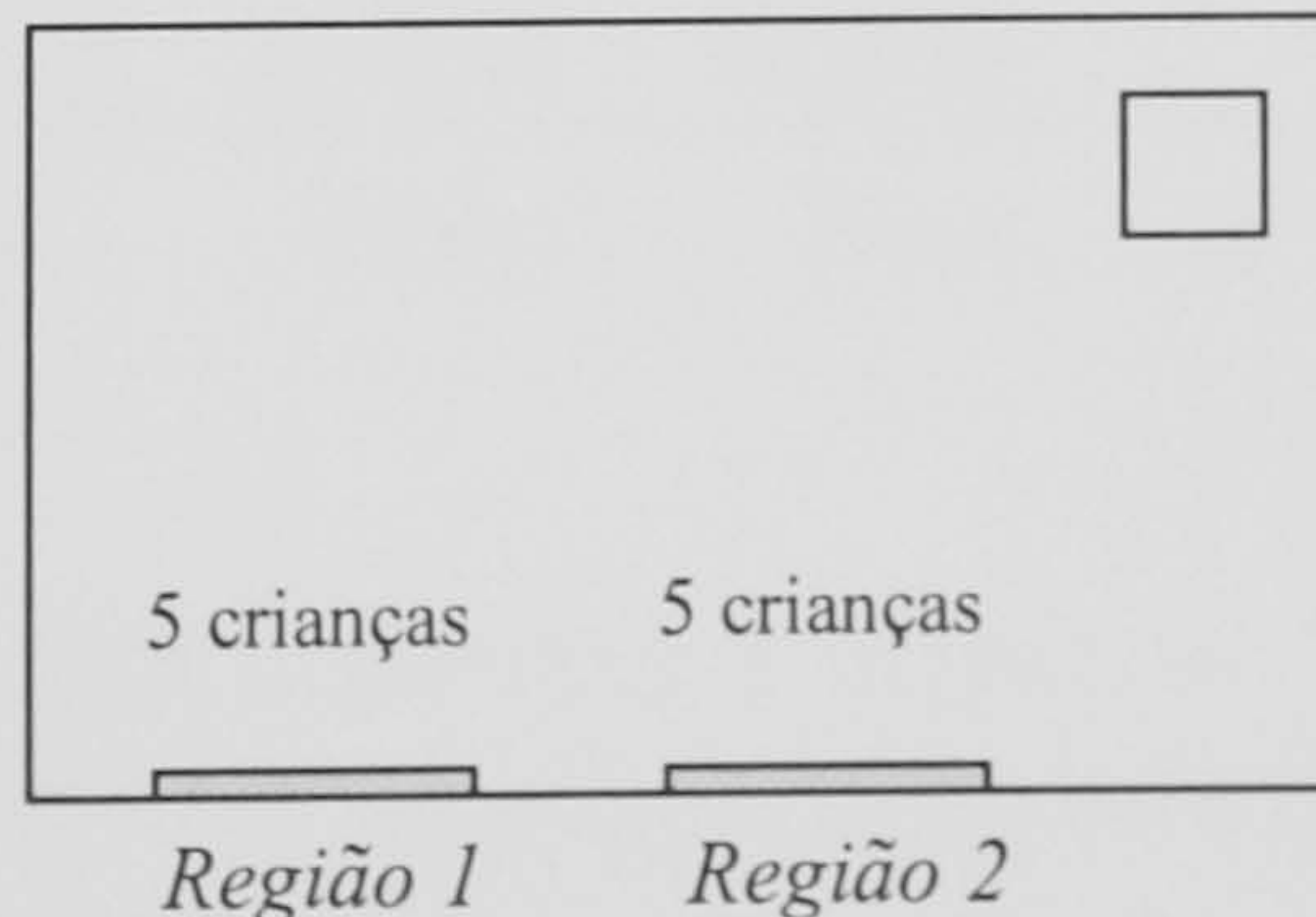
ou



5.



ou



#### Questão 4

Com base em dados do ano de 2006 sabe-se que o tempo médio de espera para uma cirurgia, em Portugal, nas três regiões com maior número de pessoas à espera (Norte, Centro e Lisboa e Vale do Tejo) é de 7 meses e que estão inscritas nas listas de espera de cada uma destas regiões cerca de 70000 pessoas.

Suponha que o Ministério da Saúde decide, como medida extraordinária, reforçar o financiamento do combate às listas de espera para determinada cirurgia (por exemplo, cataratas ou prótese da anca) no próximo ano. O objectivo é reduzir o número de pessoas à espera para além do tempo considerado clinicamente aceitável.

Assim, suponha que existem duas alternativas para distribuir este montante extraordinário de recursos:

-Na **Alternativa A**, os recursos são concentrados na Região de Lisboa e Vale do Tejo (LVT);

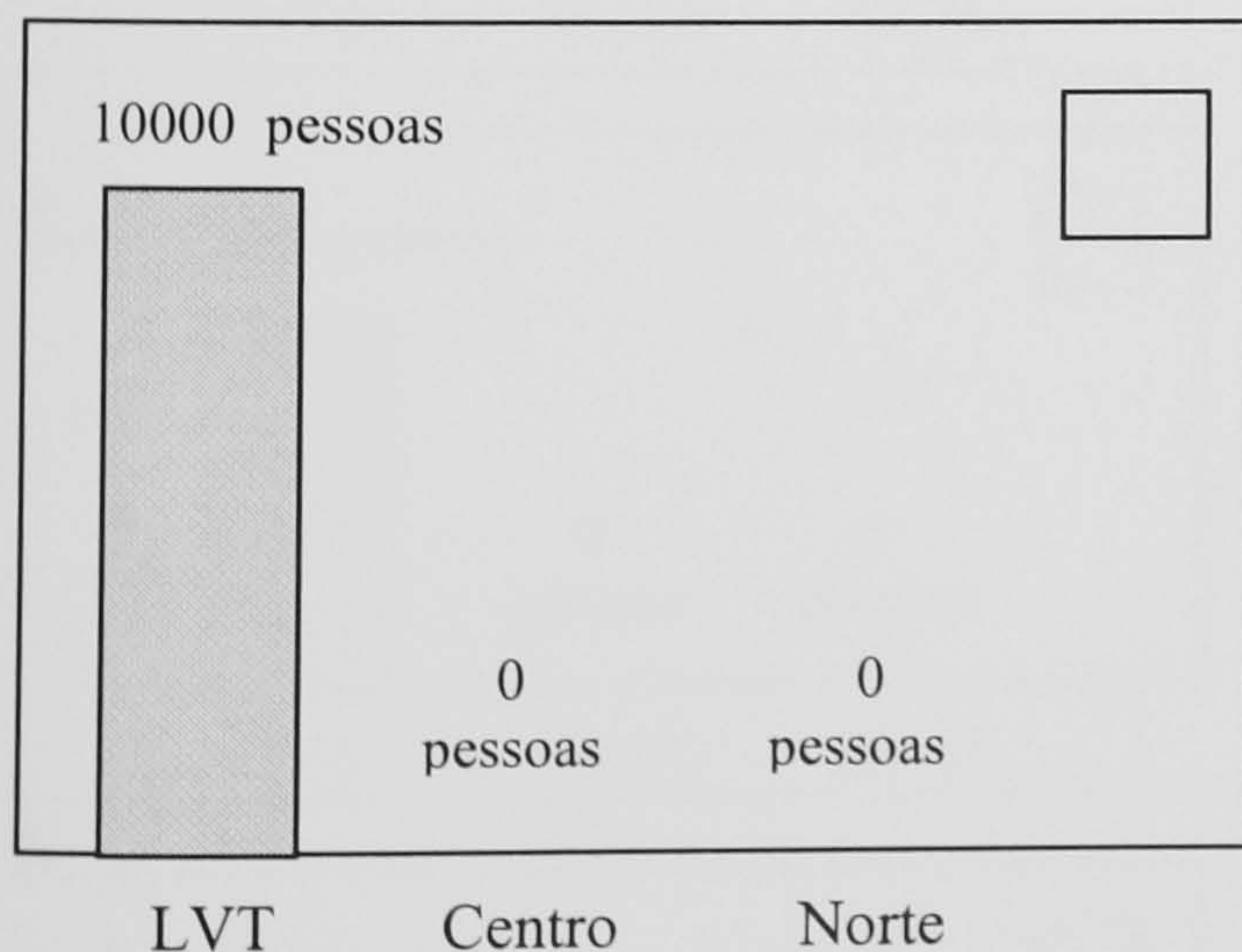
-Na **Alternativa B**, o montante é dividido igualmente pelas Regiões Norte, Centro e Lisboa e Vale do Tejo

Pelo facto da Região de Lisboa e Vale do Tejo dispor de mais equipas especializadas, a primeira alternativa permite alcançar uma maior redução do número de pessoas à espera, no total das três regiões.

Por favor, indique qual das alternativas, A ou B, escolheria, marcando com uma cruz ☒ o quadrado correspondente:

#### Alternativa A

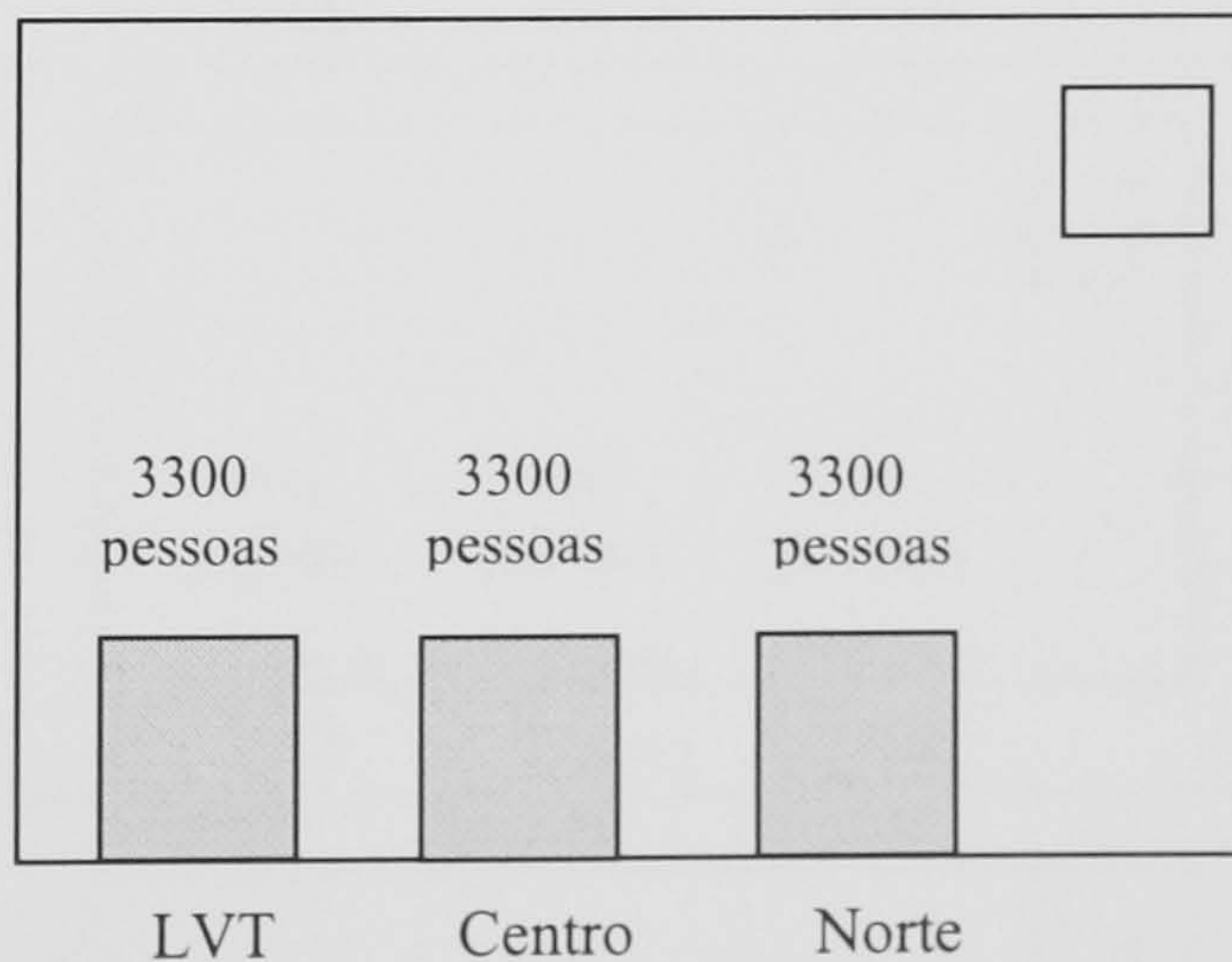
Número total de pessoas operadas



ou

#### Alternativa B

Número total de pessoas operadas



**ATENÇÃO:** Se escolheu a **Alternativa A**, por favor vá para a **questão 5** da página 7.

Se escolheu a **Alternativa B**, por favor continue na página seguinte (p.6)

Admita agora que na Alternativa B, em que o montante é dividido igualmente pelas três Regiões, o número total de pessoas operadas é inferior ao anteriormente anunciado.

Para cada uma das cinco hipóteses em baixo, assinale com uma cruz um dos quadrados, indicando se continua a preferir a Alternativa B ou se agora passaria a optar pela Alternativa A:

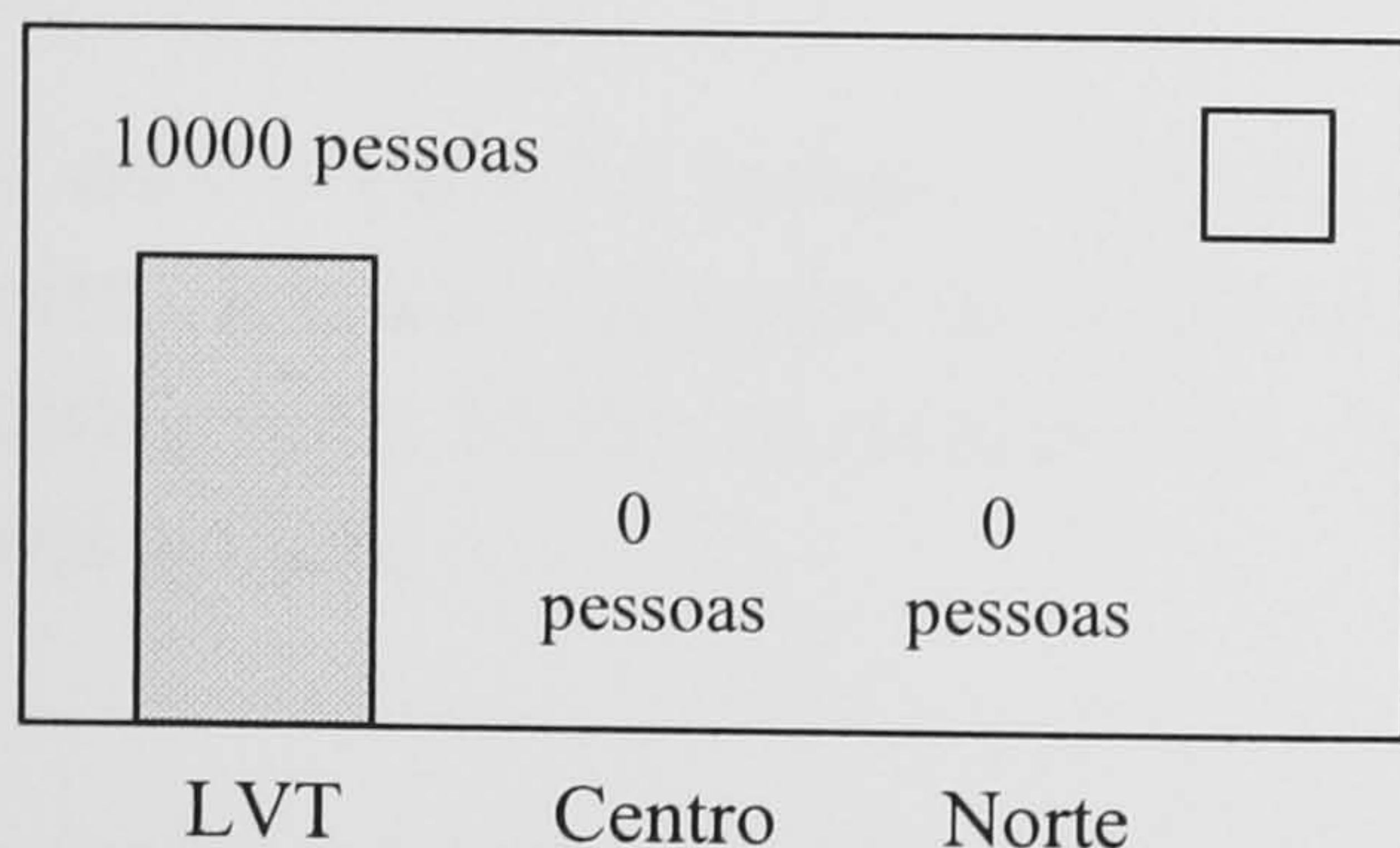
### Alternativa A

Número total de pessoas operadas

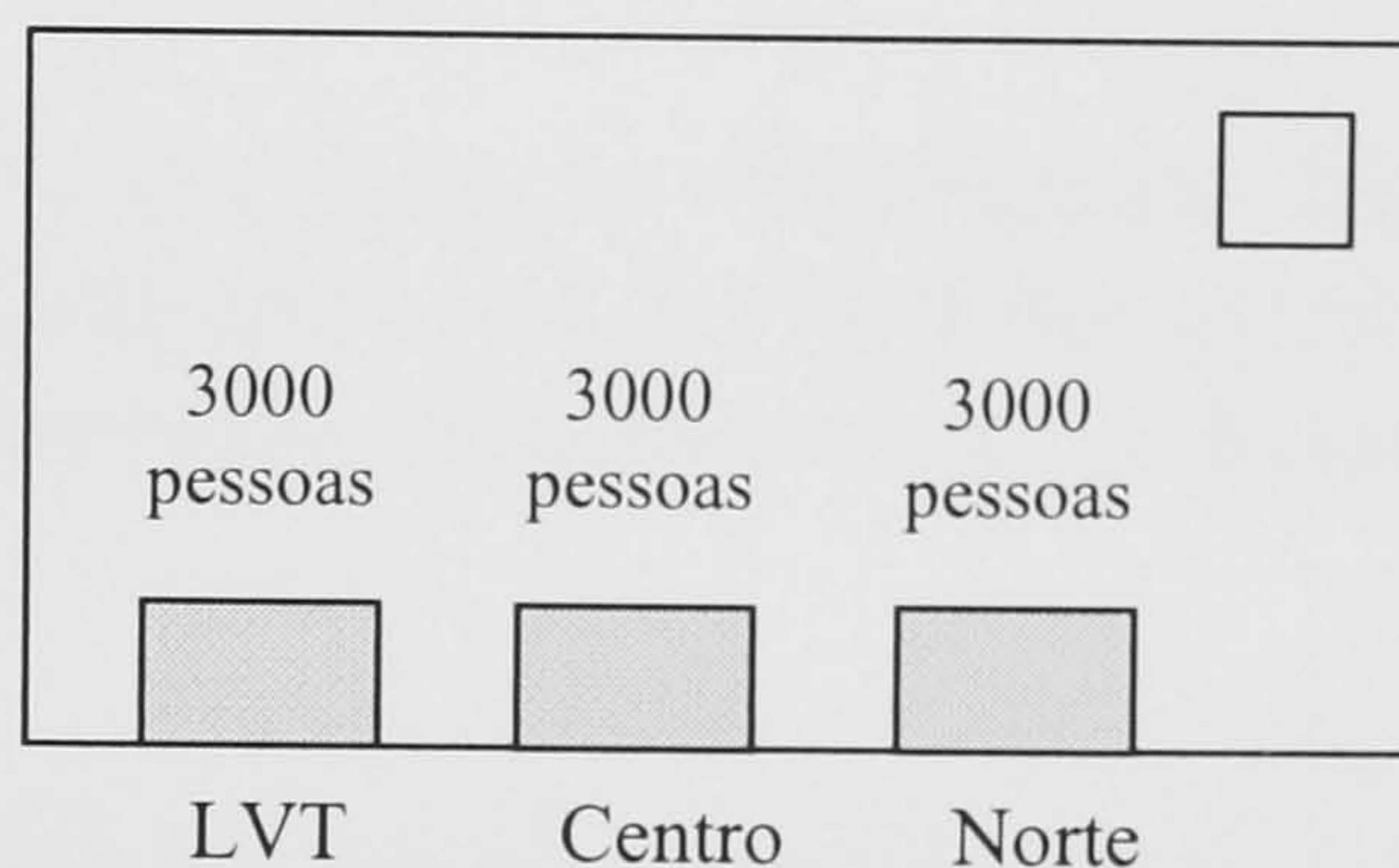
### Alternativa B

Número total de pessoas operadas

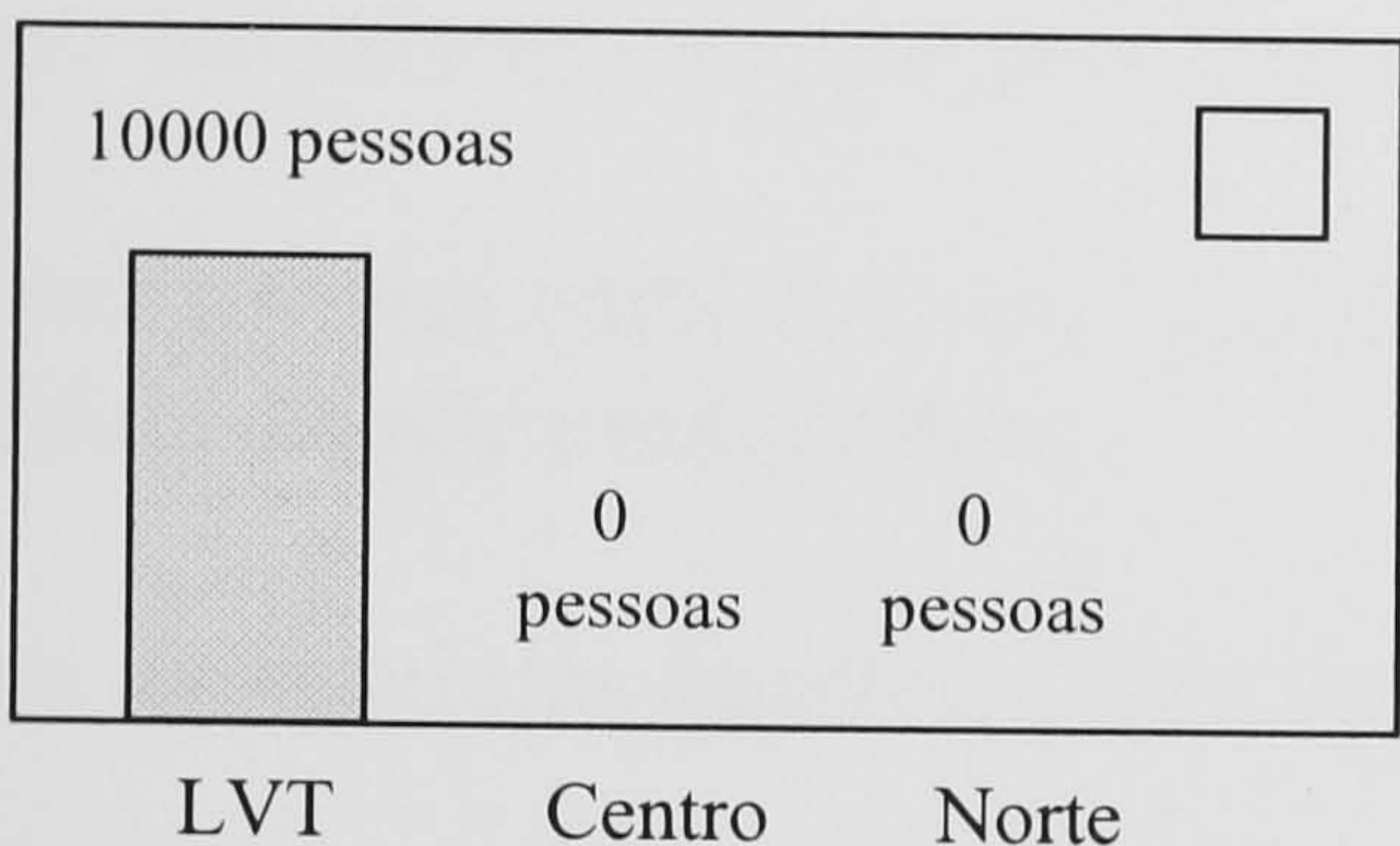
1.



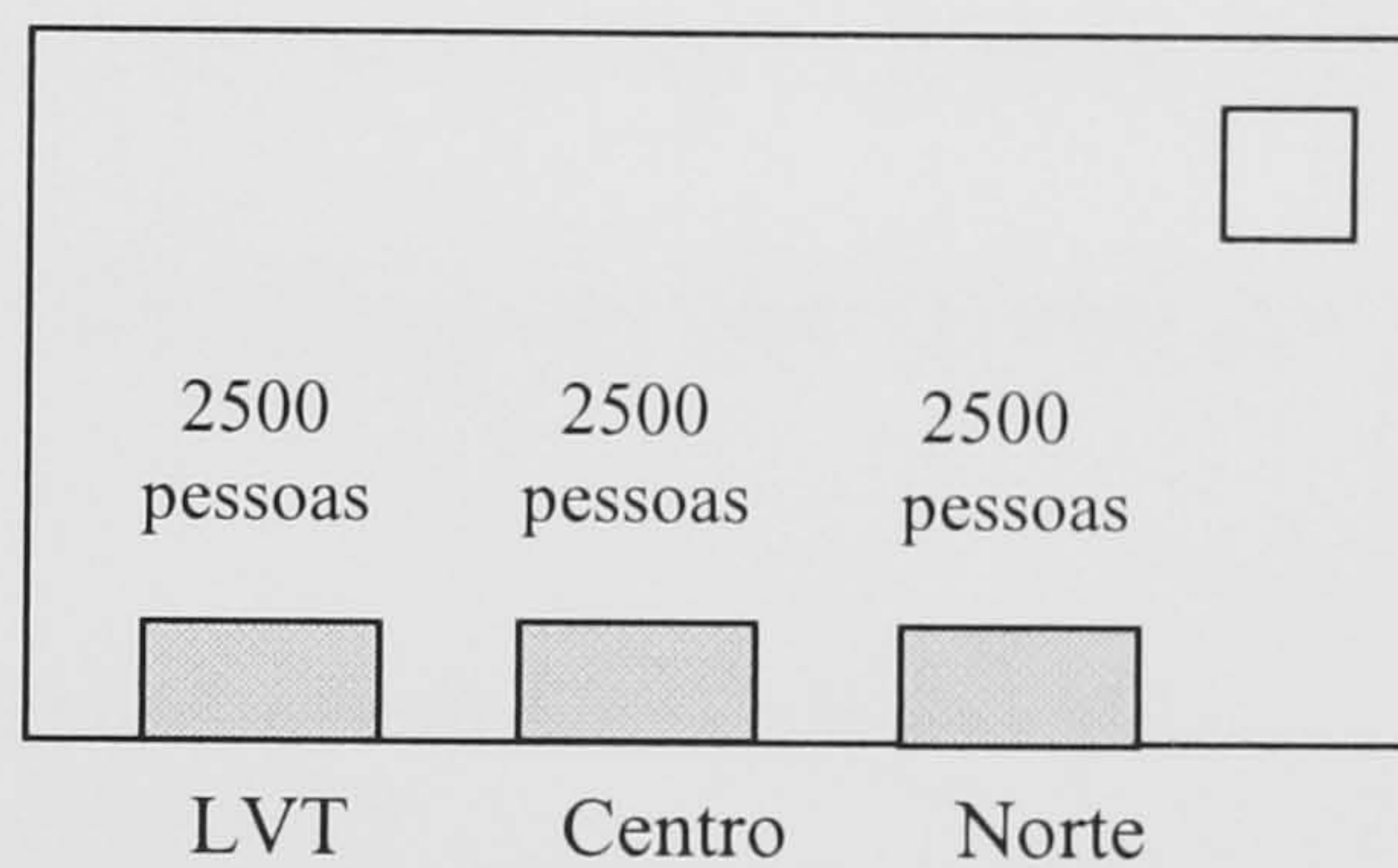
ou



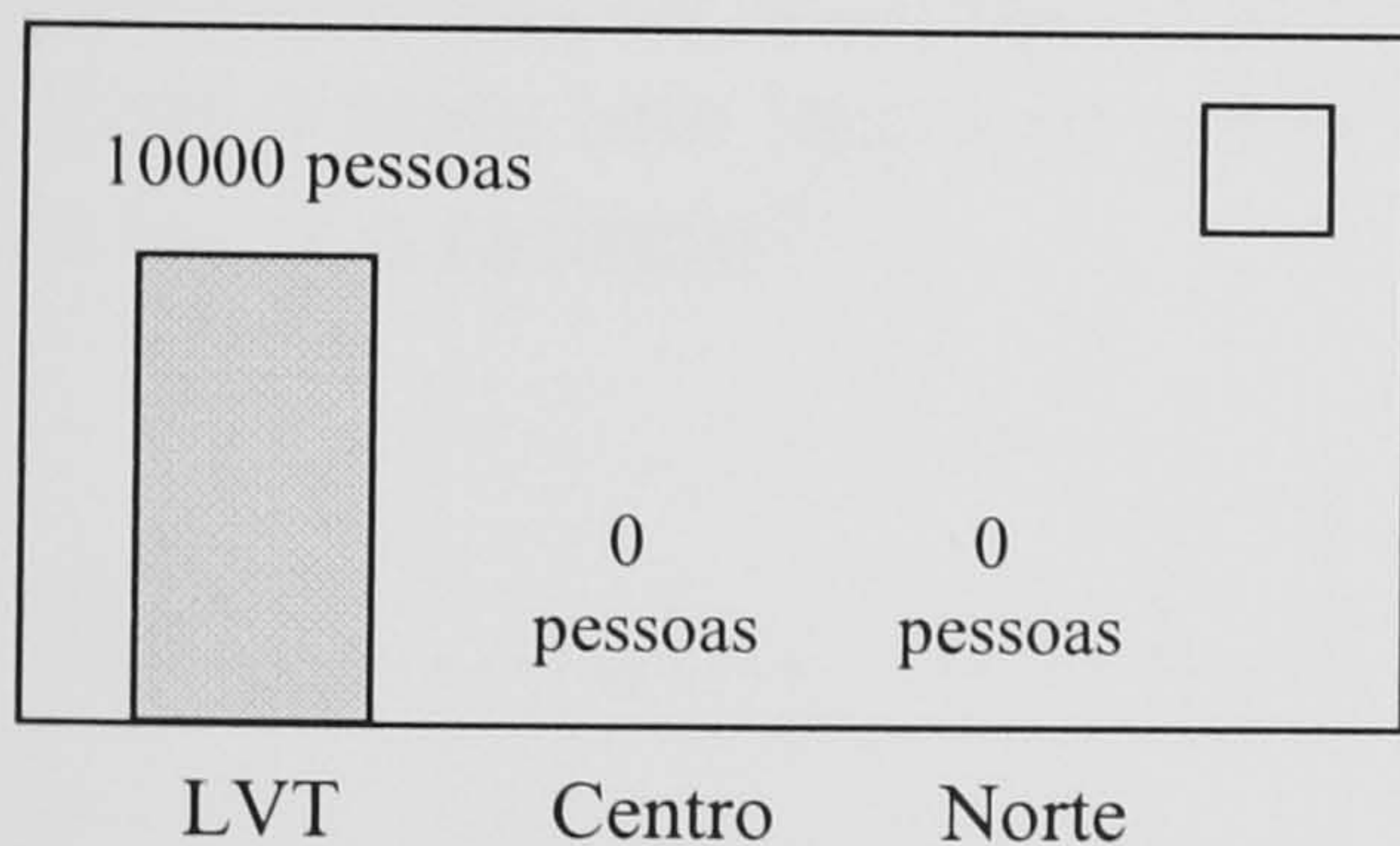
2.



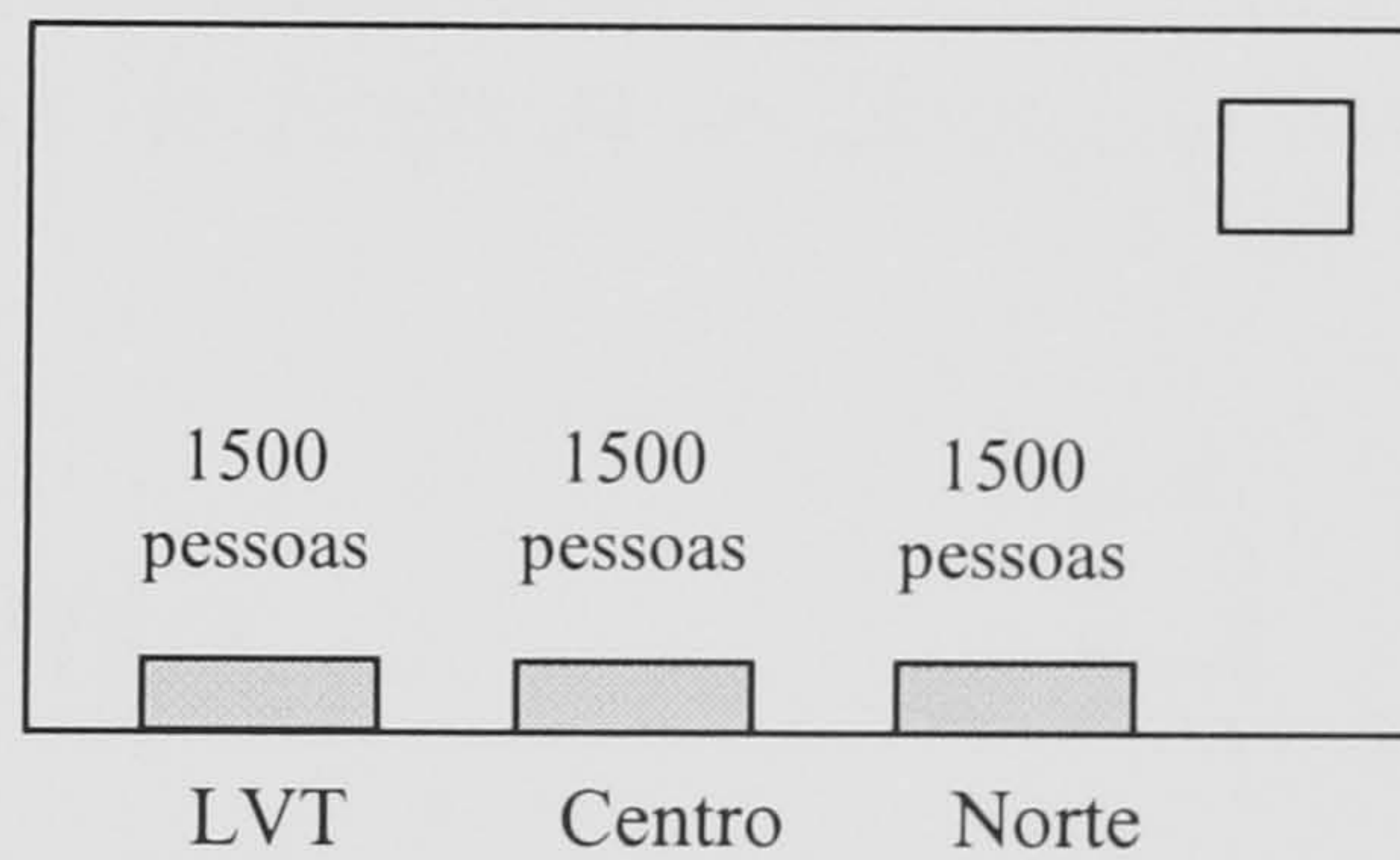
ou



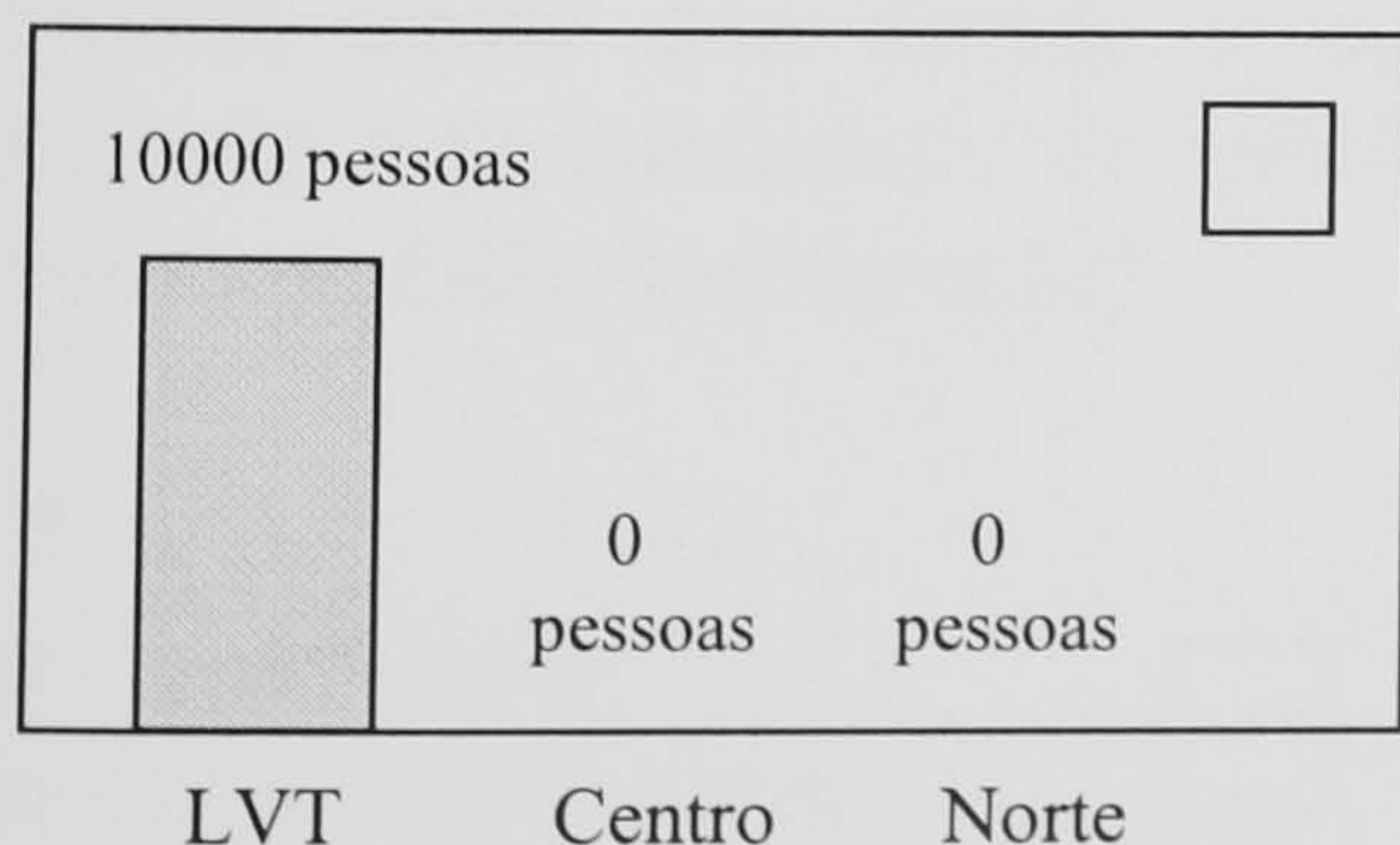
3.



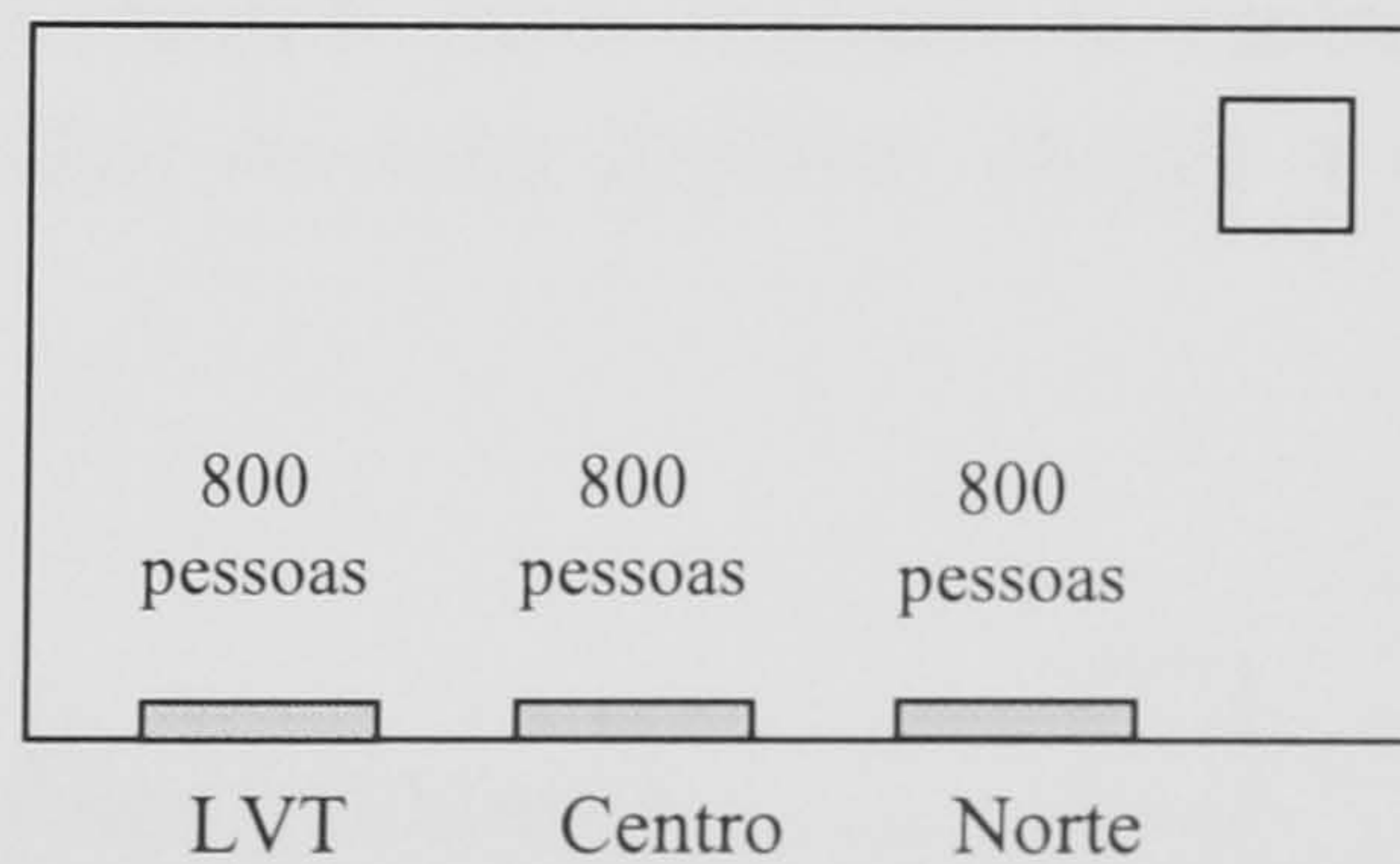
ou



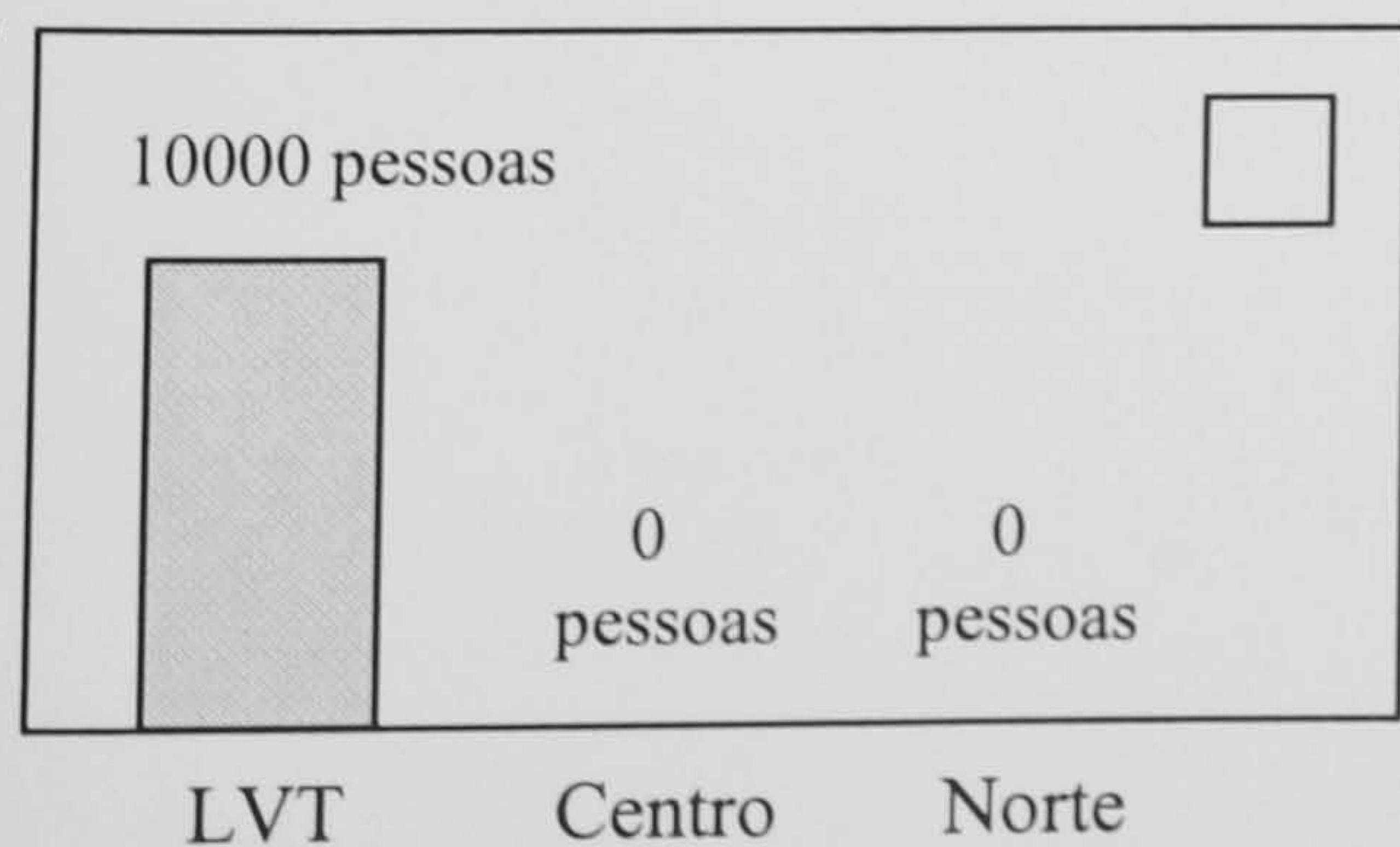
4.



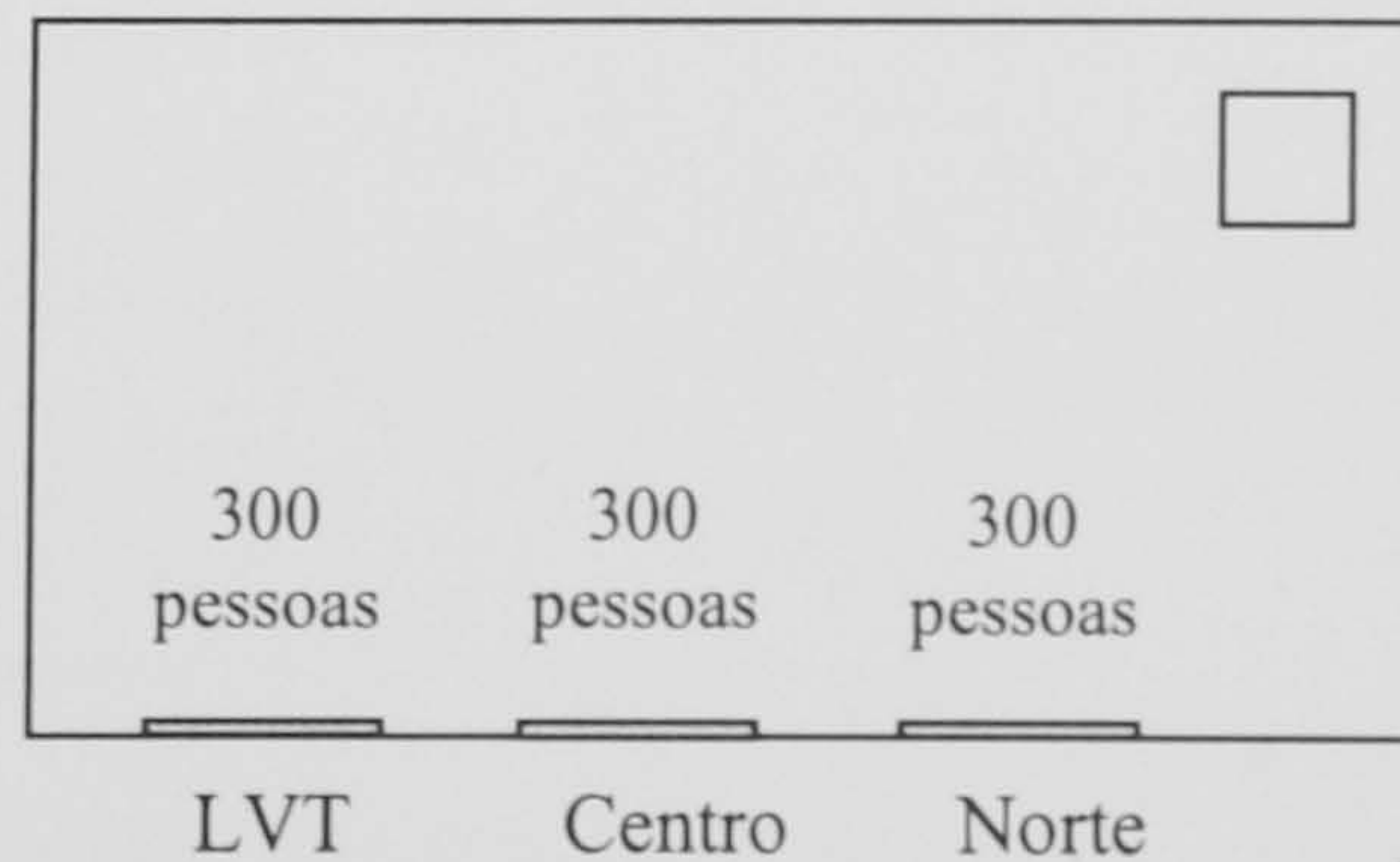
ou



5.



ou



Nas duas questões seguintes refira-nos as suas preferências pessoais.

**Questão 5**

Existem muitas doenças, como as doenças oncológicas, que ao serem detectadas precocemente têm uma maior probabilidade de cura. Para se diagnosticar uma doença no seu início é fundamental que se façam rastreios.

No nosso país há poucos rastreios sistemáticos das doenças oncológicas. Ainda assim, mesmo quando se oferece o teste gratuitamente no Serviço Nacional de Saúde, nem todos os utentes se dirigem aos serviços competentes para fazerem esses testes.

Suponha que recebeu uma carta do seu Centro de Saúde a convidá-lo(a) para fazer um rastreio oncológico. Pense por exemplo no rastreio do cancro do cólon.

Nessa mesma carta referem que lhe dão a escolher entre ser atendido(a) por um médico ou por uma médica.

Para as seguintes questões, faça uma cruz ☒ no quadrado correspondente:

A possibilidade de escolher o género (homem ou mulher) do profissional que irá realizar o teste tem impacto sobre a sua decisão de dirigir-se ao Centro de Saúde para fazer o rastreio?

SIM.....

NÃO.....

Suponha agora que decide ir ao Centro de Saúde para realizar o rastreio. Mediante a possibilidade de escolher um médico ou uma médica, exerce a sua escolha ou é-lhe indiferente?

Escolho de acordo com a minha preferência.....

É-me indiferente.....

## Questão 6

A atitude das pessoas perante os mesmos problemas de saúde nem sempre é igual. Existem, por exemplo, sintomas comuns como ‘dores de cabeça’ ou ‘dores de costas’ que em grande parte dos casos não levam as pessoas a procurar cuidados médicos personalizados.

Gostaríamos de saber qual a sua atitude / preferência perante as situações hipotéticas descritas em baixo. Pense, por exemplo, no caso das ‘dores de cabeça’ ou ‘dores de costas’.

Para cada uma das cinco hipóteses faça uma cruz ☒ no quadrado que no seu entender melhor corresponde ou corresponderia à sua forma de agir preferida.

	Não faço nada	Faço auto-medicação	Peço conselho ao farmacêutico	Vou ao médico
1. Sinto dores fortes uma vez por outra	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Sinto dores fortes frequentemente mas faço a minha vida normal	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Sinto dores fortes frequentemente e embora faça a minha vida normal durante o dia, durmo mal de noite por causa das dores	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Sinto dores fortes frequentemente, durmo mal e por vezes tenho de interromper as minhas tarefas diárias por causa das dores	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Sinto dores fortes frequentemente, durmo mal e por vezes sou mesmo forçado(a) a faltar ao trabalho / escola por causa das dores	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Por fim, gostaríamos que nos fornecesse alguns dados pessoais (sexo, idade e nível de escolaridade), apenas para fins estatísticos.**

Por favor, preencha ou faça uma cruz ☒ no quadrado correspondente:

1. Sexo:

Masculino.....

Feminino.....

2. Idade: \_\_\_\_\_ Anos

3. Nível de escolaridade: Qual o maior grau que frequentou com êxito? (Se ainda se encontra a estudar, em qualquer nível de ensino, por favor, faça uma cruz onde se lê “A frequentar”):

Nenhum.....

Ensino secundário.....

1º Ciclo (4ª classe)....

Curso médio.....

2º Ciclo (6º ano).....

Curso superior.....

3º Ciclo (9ºano).....

A frequentar.....

Muito obrigada por ter preenchido este questionário. A sua contribuição é valiosa para a nossa investigação.

## REFERENCES

- Abasolo, I. and A. Tsuchiya (2004), "Exploring social welfare functions and violation of monotonicity: an example from inequalities in health", *Journal of Health Economics* 23: 313-329.
- Akin, J.; P. Hutchinson and K. Strumpf (2005), "Decentralisation and government provision of public goods: The public health sector in Uganda", *The Journal of Development Studies* 41: 1417-1443.
- Alborz A.; D. Wilkin and K. Smith (2002), "Are primary care groups and trusts consulting local communities?", *Health & Social Care in the Community* 10: 20-28.
- Allen, P. (2006), "New localism in the English National Health Service: What is it for?", *Health Policy* 79: 244-252.
- Anand, P. (2001), "Procedural fairness in economic and social choice: Evidence from a survey of voters", *Journal of Economic Psychology* 22: 247-270.
- Anand, P. and A. Wailoo (2000), "Utilities versus Rights to Publicly Provided Goods: Arguments and Evidence from Health Care Rationing", *Economica* 67: 543-577.
- Ashcroft, R. (2003), "In vitro fertilisation for all?", *British Medical Journal* 327: 511-512.
- Baggott, R. (2005), "A funny thing happened on the way to the forum? Reforming Patient and Public Involvement in the NHS in England", *Public Administration* 83: 533-551.
- Bankauskaite, V. and R. Saltman (2006), Central issues in the decentralization debate in: R. Saltman and Bankauskaite, eds, *Decentralization in Health Care: Strategies and Outcomes*, The European Observatory on Health Systems and Policies, Open University Press.
- Bate, A.; C. Donaldson and M. Murtagh (2007), "Managing to manage healthcare resources in the English NHS? What can health economics teach? What can health economics learn?", *Health Policy*, article in press.



- Bentes, M.; C. Dias; C. Sakellarides and V. Bankauskaite (2004), *Health Care Systems in Transition – Portugal*, European Observatory on Health Systems and Policies.
- Birch, S. and A. Gafni (2002), “On being NICE in the UK: guidelines for technology appraisal for the NHS in England and Wales”, *Health Economics* 11: 185-191.
- Bossert, T. (1998), “Analyzing the decentralization of health systems in developing countries: decision space, innovation and performance”, *Social Science and Medicine* 47: 1513-1527.
- Bossert, T. and Joel Beauvais (2002), “Decentralisation of health systems in Ghana, Zambia, Uganda and the Philippines: a comparative analysis of decision space”, *Health Policy and Planning* 17: 14-31.
- Bossert, T.; M. Chitah and D. Bowser (2003), “Decentralization in Zambia: resource allocation and district performance”, *Health Policy and Planning* 18: 357-369.
- Bowie, C; A. Richardson and W. Sykes (1995), “Consulting the public about health service priorities”, *British Medical Journal* 311: 1155-1158.
- Bowles, S. (1998), “Endogenous preferences: The cultural consequences of markets and other economic institutions”, *Journal of Economic Literature* XXXVI: 75-111.
- Bowling, A. (1996), “Health care rationing: the public’s debate”, *British Medical Journal* 312: 670-674.
- Burns, D. (2000), “Can local democracy survive governance?”, *Urban Studies* 37: 963-973.
- Callaghan, G. and G. Wistow (2006), “Governance and public involvement in the British National Health Service: Understanding difficulties and developments”, *Social Science and Medicine* 63: 2289-2300.
- Campos, A. (2004), “Decentralization and privatization in Portuguese health reforms”, *Revista Portuguesa de Saúde Pública*, vol. temático.

- Cappelen, A. and O. Norheim (2006), "Responsibility, fairness and rationing in health care", *Health Policy* 76: 312-319.
- Chubb, J. (1985), "The political economy of fiscal federalism", *American Political Science Review* 79: 994-1015.
- Coast, J. (2001), "Citizens, their agents and health care rationing: an exploration study using qualitative methods", *Health Economics* 10: 159-174.
- Collins, C.; J. Araujo and J. Barbosa (2000), "Decentralising the health sector: issues in Brazil", *Health Policy* 52: 113-127.
- Cookson, R. and P. Dolan (1999), "Public views on health care rationing: a group discussion study", *Health Policy* 49: 63-74.
- Cookson, R. and P. Dolan (2000), "Principles of justice in health care rationing", *Journal of Medical Ethics* 26: 323-329.
- Cuadras-Morató, X.; J. Pinto-Prades and J. Abellán-Perpiñan (2001), "Equity considerations in health care: the relevance of claims", *Health Economics* 10: 187-205.
- Culyer, A. J. (1989a), "The normative economics of health care financing and provision", *Oxford Review of Economic Policy* 5: 34-58.
- Culyer, A. J. (1989b), "Commodities, characteristics of commodities, characteristics of people, utilities and the quality of life" in: S. Baldwin *et al.*, eds, *The quality of life: Perspectives and policies*, Routledge, London.
- Culyer, A. J. and A. Wagstaff (1993), "Equity and equality in health and health care", *Journal of Health Economics* 12: 431-457.
- Culyer, A. J. (1995), "Equality of What in Health Policy? Conflicts Between the Contenders", University of York, Centre for Health Economics, Discussion Paper 142.
- Culyer, A. J. and A. Castelli (2004), "Rationing health care in Europe – the United Kingdom" in: J.-M. von der Schulenburg and M. Blanke, eds, *Rationing of Medical Services in Europe: An empirical study*, IOS Press, Amsterdam.

Culyer, A. J. (2006), "The bogus conflict between efficiency and vertical equity", *Health Economics* 15: 1155-1158.

Davies, E. and P. Littlejohns (2002), "Views of directors of public health about NICE appraisal guidance: results from a postal survey", *Journal of Public Health Medicine* 24: 319-325.

Department of Health (1999), *Saving Lives – Our Healthier Nation*, London.

Department of Health (2001a), Press Notice 2001/0108.

Department of Health (2001b), *Shifting the balance of power within the NHS – securing delivery*, London.

Department of Health (2003a), *A guide to the NHS for members and officers of health scrutiny committees*, London.

Department of Health (2003b), *Resource Allocation: weighted capitation formula*, London.

Department of Health (2004a), *A responsive and high-quality local NHS: The primary care progress report 2004*, London.

Department of Health (2004b), *Patient and Public Involvement in Health: The Evidence for Policy Implementation*, London

Department of Health (2005), *Commissioning a Patient-Led NHS*, London.

Devlin, N. and D. Parkin (2004), "Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis", *Health Economics* 13: 437-452.

Dolan, P. (1998), "The measurement of individual utility and social welfare", *Journal of Health Economics* 17: 39-52.

Dolan, P. and C. Green (1998), "Using the person trade-off approach to examine

differences between individual and social values”, *Health Economics* 7: 307-312.

Dolan, P.; R. Cookson and B. Ferguson (1999), “Effect of discussion and deliberation on the public's views of priority setting in health care: focus group study”, *British Medical Journal* 318: 916-919.

Dolan, P. and R. Cookson (2000), “A qualitative study of the extent to which health gain matters when choosing between groups of patients”, *Health Policy* 51: 19–30.

Dolan, P. and J. Olsen (2001), “Equity in health: the importance of different health streams”, *Journal of Health Economics*, 20: 823-834.

Dolan, P. and A. Robinson (2001), “The measurement of preferences over the distribution of benefits: The importance of the reference point”, *European Economic Review* 45: 1697-1709

Dolan, P.; A. Tsuchiya; P. Smith; R. Shaw and A. Williams (2002), Determining the parameters in a social welfare function using stated preference data: an application to health. Sheffield Health Economics Group Discussion Paper Series May 2002.

Dolan P. and R. Shaw (2003), “A note on the relative importance that people attach to different factors when setting priorities in health care”, *Health Expectations* 6: 53–59.

Dolan, P. and A. Tsuchiya (2003), “The person trade-off method and the transitivity principle: an example from preferences over age weighting”, *Health Economics* 12: 505-520.

Dolan, P.; J. Olsen; P. Menzel and J. Richardson (2003), “An inquiry into the different perspectives that can be used when eliciting preferences in health”, *Health Economics* 12: 545-551.

Dolan, P and R. Shaw (2004), “A note on a discussion group study of public preferences regarding priorities in the allocation of donor kidneys”, *Health Policy* 68: 31–36.

- Dolan, P. and A. Tsuchiya (2005), “Health priorities and public preferences: the relative importance of past health experience and future health prospects”, *Journal of Health Economics* 24: 703–714.
- Dolan, P.; R. Shaw; A. Tsuchiya and A. Williams (2005), “QALY maximisation and people’s preferences: a methodological review of the literature”, *Health Economics* 14: 197-208.
- Dolan, P.; R. Edlin; A. Tsuchiya; and A. Wailoo (2007), “It ain’t what you do, it’s the way that you do it: Characteristics of procedural justice and their importance in social decision-making”, *Journal of Economic Behavior & Organization* 64: 157-170.
- Faguet, J.-P. (2004), “Does decentralization increase government responsiveness to local needs? Evidence from Bolivia”, *Journal of Public Economics* 88: 867-893.
- Fehr, E. and U. Fischbacher (2002), “Why social preferences matter – the impact of non-selfish motives on competition, cooperation and incentives”, *The Economic Journal* 112: C1-C33.
- Frey, B.; F. Oberholzer-Gee and R. Eichenberger (1996), “The old lady visits your backyard: A tale of morals and markets“, *Journal of Political Economy* 104: 1297-1313.
- Frey, B.; B. Matthias and A. Stutzer (2004), “Introducing procedural utility: Not only what, but also how matters”, *Journal of Institutional and Theoretical Economics* 160: 377-401.
- Gerard, K. *et al.* (2006), “The introduction of integrated out-of-hours arrangements in England: a discrete choice experiment of public preferences for alternative models of care”, *Health Expectations* 9: 60-69.
- Giacomini, M. *et al.* (2004), “The policy analysis of ‘values talk’: lessons from Canadian health reform”, *Health Policy* 67: 15-24.
- Gilbert, G and P. Picard (1996), “Incentives and optimal size of local jurisdictions”, *European Economic Review* 40: 19-41.
- Gordon, R. (1983), “An optimal taxation approach to fiscal federalism”. The

Quarterly Journal of Economics 98: 567-586.

Gray, A. and B. Jenkins (2000), "Government and Administration 1998-99: overcoming 'conservatism' – a job half done?", *Parliamentary Affairs* 53: 219-241.

Green, C. (2001), "On the societal value of health care: what do we know about the person trade-off technique?", *Health Economics* 10: 233-243.

Greener, I. and J. Powell (2003), "Health authorities, priority-setting and resource allocation: a study in decision-making in New Labour's NHS", *Social Policy & Administration* 37: 35-48.

Ham, C. (2004), *Health Policy in Britain*, 5th edn, Palgrave Macmillan, Basingstoke.

Harrison, S. and M. Mort (1998), "Which champions, which people? Public and user involvement in health care as a technology of legitimisation", *Social Policy & Administration* 32: 60-70.

Hill, M. and A. Hill (2002), *Investigação por questionário*, Edições Sílabo, Lisbon.

Hauck, K.; R. Shaw and P. C. Smith (2002), "Reducing avoidable inequalities in health: a new criterion for setting health care capitation", *Health Economics* 11: 667-677.

Heuer, L. (2005), "What's just about the criminal justice system? A psychological perspective", *Journal of Law and Policy* XIII: 209-228.

Homedes, N. and A. Ugalde (2002), "Why neoliberal health reforms have failed in Latin America", *Health Policy*, *article in press*, available online on 2 September.

Hurley, J.; S. Birch and J. Eyles (1995), "Geographically-decentralised planning and management in health care: some informational issues and their implications for efficiency", *Social Science and Medicine* 41: 3-11.

Hutchinson, P. and A. LaFond (2004), *Monitoring and evaluation of*

- decentralization reforms in developing country health sectors, Partners for Health Reform*plus* Project.
- Hutton, J. and A. Maynard (2000), "A NICE challenge for Health Economics", *Health Economics* 9: 89-93.
- Inman, R. and D. Rubinfeld (1997a), "Rethinking federalism", *Journal of Economic Perspectives* 11: 43-64.
- Inman, R. and D. Rubinfeld (1997b), "Making sense of the antitrust state-action doctrine: balancing political participation and economic efficiency in regulatory federalism", *Texas Law Review* 75: 1203-1299.
- Khaleghian, P. (2004), "Decentralization and public services: the case of immunization", *Social Science & Medicine* 59: 163-183.
- Kmietowicz, Z. (2000), "NICE to tackle infertility treatment", *British Medical Journal* 321: 1432.
- Koivusalo, M. (1999), "Decentralisation and equity of healthcare provision in Finland", *British Medical Journal* 318: 1198-1200.
- Le Grand, J. (1991), *Equity and choice*, Harper Collins, London.
- Levaggi, R. and P. C. Smith (2005), "Decentralization in health care: lessons from public economics" in: P. C. Smith; L. Ginelly and M. Sculpher, eds, *Health and Economics: Opportunities and Challenges*, Open University Press, Maidenhead.
- Leventhal, G. S. (1980), "What should be done with equity theory? New approaches to the study of fairness in social relationships" in: K. Gergen, M. Greenberg and R. Willis, eds, *Social exchange: Advances in theory and research*. Plenum Press, New York.
- Lind, E. and T. Tyler (1988), *The Social Psychology of Procedural Justice*. Plenum Press, New York.
- Lindaman, K. and K. Thurmaier (2002), "Beyond efficiency and economy: An examination of basic needs and fiscal decentralization", *Economic*

*Development and Cultural Change* 50: 915-934.

Lindholm, L. and M. Rosen, (1998), "On the measurement of the nation's equity adjusted health", *Health Economics* 7: 621-628.

Litva, A. *et al.* (2002), "'The public is too subjective': public involvement at different levels of health-care decision making", *Social Science & Medicine* 54: 1825-1837.

Magnussen, J.; T. Hagen and O. Kaarboe (2007), "Centralized or decentralized? A case study of Norwegian hospital reform", *Social Science & Medicine* 64: 2129-2137.

Mansley, E. and E. Elbasha (2003), "Preferences and person trade-offs: forcing consistency or inconsistency in health-related quality of life measures?", *Health Economics* 12: 187-198.

Mckie, J. and J. Richardson (2003), "The rule of rescue", *Social Science and Medicine* 56: 2407-2419.

Milewa, T.; S. Harrison; W. Ahmad and P. Tovey (2002), "Citizens' participation in primary healthcare planning: innovative citizenship practice in empirical perspective", *Critical Public Health* 12: 39-53.

Milewa, T. (2004), "Local participatory democracy in Britain's health service: innovation or fragmentation of a universal citizenship?", *Social Policy & Administration* 38: 240-252.

Mooney, G. (1998), "'Communitarian claims': As an ethical basis for allocating health care resources", *Social Science & Medicine* 47: 1171-1180.

Mooney, G. (2005), "Communitarian claims and community capabilities: furthering priority setting?", *Social Science & Medicine* 60: 247-255.

Morgan, A.; P. Shackley; M. Pickin and J. Brazier (2000), "Quantifying patient preferences for out-of-hours primary care", *Journal of Health Services Research and Policy* 5: 214-218.

Mosca, I. (2006), "Is decentralisation the real solution? A three country study",



*Health Policy* 77: 113-120.

Mossialos, E. and D. King (1999), "Citizens and rationing: analysis of a European survey", *Health Policy* 49: 75-135.

Musgrave, R. (1959), *The theory of public finance*, McGraw-Hill, New York.

Needham, D. *et al.* (2004), "Patient care seeking barriers and tuberculosis programme reform: a qualitative study", *Health Policy* 67: 93-106.

Neuberger, J. *et al.* (1998), "Assessing priorities for allocation of donor liver grafts: survey of public and clinicians", *British Medical Journal* 317: 172-175.

NICE (2001), NICE to seek advice from new Citizens Council, Press release 2001/038.

NICE (2003), NICE response to the report of the Health Select Committee following the Committee's short inquiry.

NICE (2004), *Fertility: assessment and treatment for people with fertility problems – Clinical Guideline 11*.

NICE (2005), *Social Value Judgements – Principles for the development of NICE guidance*.

Nord, E. (1993), "The trade-off between severity of illness and treatment effect in cost-value analysis of health-care", *Health Policy* 24: 227-238.

Nord, E. (1995), "The person trade-off approach to valuing health care programs", *Medical Decision Making* 15: 201- 208.

Nord, E.; J. L. Pinto; J. Richardson *et al.* (1999), "Incorporating societal concerns for fairness in numerical valuations of health programmes", *Health Economics* 8: 25-39.

Oates, W. (1972), *Fiscal federalism*, Harcourt Brace Jovanovich, New York.

- Oates, W. (1999), "An essay on fiscal federalism", *Journal of Economic Literature* 37: 1120-1149.
- Oliver, A. (2005), "The English National Health Service: 1979-2005", *Health Economics* 14: S75-S99.
- Olsen, J.; J Richardson; P. Dolan and P. Menzel (2003), "The moral relevance of personal characteristics in setting health care priorities", *Social Science and Medicine* 57: 1163-1172.
- Palacios-Huerta, I. and T. Santos (2004), "A theory of markets, institutions, and endogenous preferences", *Journal of Public Economics* 88: 601-627.
- Paton, C. (1993), "Devolution and centralism in the NHS", *Social Policy and Administration* 27: 83-108.
- Paton, C. (1998), *Competition and Planning in the NHS*, Thornes, Cheltenham.
- Peckham, S.; M. Exworthy; M. Powell and I. Greener (2005), *Decentralisation as an organisational model for health care in England*, Report to NCCSDO.
- Petretto, A. (2000), "On the cost-benefit of the regionalisation of the National Health Service", *Economics of Governance* 1: 213-232.
- Prades (1997), "Is the person trade-off a valid method for allocating health care resources?", *Health Economics* 6: 71-81.
- Prud'Homme, R. (2003), "Fiscal decentralisation in Africa: a framework for considering reform", *Public Administration and Development* 23: 17-27.
- Ratcliffe, J. (2000), "Public preferences for the allocation of donor liver grafts for transplantation", *Health Economics* 9: 137-148.
- Rawlins, M. and A. J. Culyer (2004), "National Institute for Clinical Excellence and its value judgments", *British Medical Journal* 329: 224-227.
- Rawls, J. (1971), *A Theory of Justice*, Harvard University Press, Cambridge, MA.

- Rawls, J. (1999), *A Theory of Justice*, Rev. Edition, Oxford University Press, Oxford.
- Reich, M. (2002), "Reshaping the state from above, from within, from below: implications for public health", *Social Science & Medicine* 54: 1669-1675.
- Rice, N. and P. C. Smith (2001), "Ethics and geographical equity in health care". *Journal of Medical Ethics* 27: 256-261.
- Robalino, D.; O. Picazo and A. Voetberg (2001), "Does fiscal decentralization improve health outcomes?", *Africa Technical Families – Human Development* 1, The World Bank.
- Roberts, T. *et al.* (1999), "Public involvement in health care priority setting: an economic perspective", *Health Expectations* 2: 235-244.
- Robinson, A. and D. Parkin (2002), "Recognizing diversity in public preferences: the use of preference sub-groups in cost-effectiveness analysis. A response to Sculpher and Gafni", *Health Economics* 11: 649-651.
- Ross, W. and J. Tomaney (2001), "Devolution and health policy in England", *Regional Studies* 35: 265-270.
- Rowe, R. and M. Shepherd (2002), "Public participation in the New NHS: no closer to citizen control?", *Social Policy & Administration* 36: 275-290.
- Saltman, R. and V. Bankauskaite (2004), "Implementing decentralization in European health care systems: searching for policy lessons", paper presented at the 5<sup>th</sup> European Conference on Health Economics, 8-11 September, London.
- Sassi, F.; L. Archard and J. Le Grand (2001), "Equity and economic evaluation of healthcare", *Health Technology Assessment Papers* 5.
- Saunders, M; P. Lewis and A. Thornhill (2003), *Research Methods for Business Students*, 3<sup>rd</sup> Edition, Prentice Hall, Essex.
- Schickle, D. (1997), "Public preferences for health care: prioritisation in the United Kingdom", *Bioethics* 11: 277-290.

- Schwappach, D. (2002), "Resource allocation, social values and the QALY: a review of the debate and empirical evidence", *Health Expectations* 5: 210-222.
- Schwarzinger, M. *et al.* (2004), "Lack of multiplicative transitivity in person trade-off responses", *Health Economics* 13: 171-181.
- Scott, A.; M. Watson and M. Ross (2003), "Eliciting preferences of the community for out-of-hours care provided by general practitioners: a stated preference discrete choice experiment", *Social Science & Medicine* 56: 803-814.
- Sculpher, M. and A. Gafni (2001), "Recognizing diversity in public preferences: the use of preference sub-groups in cost-effectiveness analysis", *Health Economics* 10: 317-324.
- Sculpher, M. ; M. Drummond and B. O'Brien (2001), "Effectiveness, efficiency, and NICE", *British Medical Journal* 322: 943-944.
- Seabright, P. (1996), "Accountability and decentralisation in government: An incomplete contracts model", *European Economic Review* 40: 61-89.
- Sen, A. (1977), "Social choice theory: A re-examination", *Econometrica* 45: 53-90.
- Sen, A. (1980), "Equality of what?" in: S. McMurrin, ed, *The Tanner lectures on human values*, Cambridge University Press, Cambridge.
- Sen, A. (1997), "Maximisation and the act of choice", *Econometrica* 65: 745-779.
- Sen, A and J. Foster (1997), *On economic inequality*, Enlarged Edition, Clarendon Press, Oxford.
- Sen, A. (2002), "Why health equity?", *Health Economics* 11: 659-666.
- Shannon, C. (2005), "New guide ranks success of IVF treatment centres", *British Medical Journal* 330: 1228.
- Shiell, A.; P. Hawe and J. Mour (1997), "Values and preferences are not

- necessarily the same”, *Health Economics* 6: 515-518.
- Siegel, S. and N.Castellan, Jr. (1988), *Nonparametric Statistics for the Behavioural Sciences* (2nd ed.), McGraw-Hill, New York.
- Smith, B.C., (1997), “The decentralization of health care in developing countries: organizational options”, *Public Administration and Development* 17: 399-412.
- Smith, P. C.; A. Stepan; V. Valdmanis and P. Verheyen (1997), “Principal-agent problems in health care systems: an international perspective”, *Health Policy* 41: 37-60.
- Smith, P. C. (2003), “Formula funding of public services: an economic analysis”, *Oxford Review of Economic Policy* 19: 301-322.
- Smith, R. and J. Richardson (2005), “Can we estimate the ‘social’ value of a QALY? Four core issues to resolve”, *Health Policy* 74: 77-84.
- Stolk, E.; *et al.* (2005), “Equity in health care prioritisation: An empirical inquiry into social value”, *Health Policy* 74: 343-355.
- Stutzer, A. and B. Frey (2006), “Political participation and procedural utility: An empirical study”, *European Journal of Political Research* 45: 391-418.
- Sunstein, C. (1993), “Endogenous preferences, environmental law”, *The Journal of Legal Studies* XXII: 217-254.
- Thibaut, J. and L. Walker (1978), “A theory of procedure”, *California Law Review* 66: 541-566.
- Tiebout, C. (1956), “A pure theory of local government expenditures”, *Journal of Political Economy* 64: 416-424.
- Tsuchiya, A. and A. Williams (2001), *Welfare economics and economic evaluation in: M. Drummond and A. McGuire, eds, Theory and Practice of Economic Evaluation in Health Care*, Oxford University Press.
- Tsuchiya, A.; P. Dolan and R. Shaw (2003), “Measuring people’s preferences

regarding ageism in health: some methodological issues and some fresh evidence”, *Social Science & Medicine* 57: 687–696.

Tsuchiya, A. *et al.* (2005), “Procedural justice in public healthcare resource allocation”, *Applied Health Economics & Health Policy* 4: 119-127.

Tsuchiya, A. and P. Dolan (2007), “Do NHS clinicians and members of the public share the same views about reducing inequalities in health?”, *Social Science & Medicine* 64: 2499-2503.

Tsuchiya, A. and J. Miyamoto, Chapter 20 Social choice in health and health care in: P. Anand; C. Puppe and P. Pattanaik, eds, *Handbook of Rational and Social Choice*, Oxford University Press (forthcoming).

Ubel, P. *et al.* (1998), “Public preferences for prevention versus care: what if an ounce of prevention is worth only an ounce of cure?”, *Medical Decision Making* 18: 141-148.

Ubel, P; J. Richardson and P. Menzel (2000), “Societal value, the person trade-off, and the dilemma of whose values to measure for cost-effectiveness analysis”, *Health Economics* 9: 127-136.

Varatharajan, D.; R. Thankappan and S. Jayapalan (2004), “Assessing the performance of primary health centers under decentralized government in Kerala, India”, *Health Policy and Planning* 19: 41-51.

Wagstaff, A. (1991), “QALYs and the equity-efficiency trade-off”, *Journal of Health Economics* 10: 21-41.

Wailoo, A. and P. Anand (2005), “The nature of procedural preferences for health-care rationing decisions”, *Social Science and Medicine* 60: 223-236.

White, C. (2004), “Infertile couples to be given three shots at IVF”, *British Medical Journal* 328: 482.

Williams, A. (1988), “Ethics and efficiency in the provision of health care” in M. Bell and S. Mendus, eds, *Philosophy and Medical Welfare*, Cambridge University Press, Cambridge.

- Williams, A. (1996), "QALYs and ethics: a health economist's perspective". *Social Science and Medicine* 43: 1795-1804.
- Williams, A. (1997a), "The rationing debate - rationing health care by age: the case for", *British Medical Journal* 314: 820-822.
- Williams, A. (1997b), "Intergenerational equity: an exploration of the fair innings argument", *Health Economics* 6:117-132.
- Williams, A. and R. Cookson (2000), "Equity in Health" in: A.J. Culyer and J.P. Newhouse, eds, *North Holland Handbook of Health Economics*, Elsevier.
- Williams, A. (2001), "How economics could extend the scope of ethical discourse", *Journal of Medical Ethics* 27: 251-255.
- Williams, A.; A. Tsuchiya and P. Dolan (2005), "Eliciting equity-efficiency trade-offs in health", in: P. C. Smith; L. Glinely and M. Sculpher, eds, *Health and Economics: Opportunities and Challenges*, Open University Press, Maidenhead.
- Wilmot, S and J. Ratcliffe (2002), "Principles of distributive justice used by members of the general public in the allocation of donor liver grafts for transplantation: a qualitative study", *Health Expectations* 5: 199-209.
- Wiseman, V.; G. Mooney; G. Berry and K. Tang (2003), "Involving the general public in priority setting: experiences from Australia", *Social Science & Medicine* 56: 1001-1012.