THE UNIVERSITY OF WESTERN AUSTRALIA

Business School

ECONOMICS

THE ECONOMICS OF MEASURING FISCAL DECENTRALISATION

PART III: SUBNATIONAL FISCAL INEQUALITIES: IMPLICATIONS FOR THE MEASUREMENT OF FISCAL DECENTRALISATION

By

Duc Vo
The University of Western Australia

DISCUSSION PAPER 08.15

THE ECONOMICS OF

MEASURING FISCAL DECENTRALISATION

PART III:

SUBNATIONAL FISCAL INEQUALITIES: IMPLICATIONS FOR THE MEASUREMENT OF FISCAL DECENTRALISATION[1]

By

Duc Hong Vo

UWA Business School

The University of Western Australia

DISCUSSION PAPER 08.15

[1] This paper comprises Chapters 5 and 6 of my PhD thesis, *The Economics of Measuring Fiscal Decentralisation*, The University of Western Australia, 2008. The full thesis is available as Discussion Papers 08.13 to 08.16.

CHAPTER 5
INFORMATION THEORY AND
ENTROPIC APPROACH TO AN ANALYSIS OF FISCAL INEQUALITY

5.1 Introduction

The fiscal decentralisation indices, as developed in Chapters 3 and 4, have two potentially significant limitations. *First*, each subnational government ("SNG") is implicitly treated as fiscally homogenous. In effect, per capita revenue and expenditure in each subnational region are implicitly assumed to be equal. However, SNGs typically involve large fiscal differences that may have implications for fiscal decentralisation. *Second*, but related to the first point, the structure of fiscal arrangements is ignored. SNGs are not differentiated by type – the state government level is not distinguished from the local government level. As such, the new indices developed earlier account only for the more fundamental influences on the fiscal autonomy and fiscal importance of SNGs while ignoring the impact of fiscal differences between them.

To redress these shortcomings, the background for the extension of the fiscal decentralisation index in future studies is developed, using information theory developed by Theil (1967). The main goals are to account for: (i) the distributions of state and local government revenue and expenditure shares between the regions physically defined by the border of state jurisdictions, and (ii) the distribution of state and local government revenue and expenditure shares within a physical region defined by the state-level governments. The concepts of "between-set entropy" and "within-set entropy" appear to have the potential to account for heterogeneity in fiscal shares across different levels of government.

The main purpose of the analysis of SNGs' fiscal inequality is to consider the impact of the distribution of revenue and expenditure between various levels of government on fiscal decentralisation. SNGs consist of two distinct levels: state (or province) and local levels. Fiscal inequality is examined in two different contexts: (i) geographic (subnational government units are grouped based on their geography), when one set of governments includes the state government and a number of local governments; and (ii) hierarchical (subnational government units are grouped based on their hierarchy), when one set of governments consists of all state governments and the other set includes all local governments. Fiscal inequality is relevant to fiscal decentralisation because it captures the influence of the structure of fiscal arrangements on decentralisation. When fiscal inequality of subnational governments within a geographically defined area is relatively high, the fiscal difference between the different tiers of SNGs is also relatively high, and, hence, there is a higher degree of fiscal decentralisation relative to the case where fiscal inequality is low.

In this chapter, following this introduction, Section 5.2 discusses information theory from Theil (1967), which lays the foundation for the analyses of fiscal inequality across subnational regions. Analytical framework for the analysis of fiscal inequality is established and discussed in Section 5.3. Section 5.4 discusses the compositions of revenue inequality. Concluding remarks are included in Section 5.5.

5.2 Analytical framework: information theory

The future is uncertain since no one knows for sure what will happen. This forms the role of expectation for various purposes of life. In every corner of life, with millions of events, everything has some possibility of occurring. Some events may simultaneously occur at the same time whereas some cannot. In the first case, two or more events are happening at the same time because their occurrence is just a happy coincidence. However, for the latter case, events are said to be mutually exclusive to each other. The possibility of occurrence of two events, sunny hours and dark rainy weather, at the same time is impossible. Every event is expected to happen with its own “probability”. This section sets out the principles of information theory, drawing on Theil (1967, chapters 2 and 3).

5.2.1 What is information?

I could not find my keys to open the drawers when I got back from lunch, and I have no idea where they could be. Some options may be available to answer my query. They may be at home. I may have left them in my supervisor's office when I was there before lunch to discuss my paper. They could also be left at University House where I went for lunch today. Another possibility is that the keys were left in the photocopying room where I went to collect my printing after lunch. Or, I have lost them on the way to university this morning. Each of these carries its own "probability" to be correct. It is very time-consuming and difficult to answer the question of where my keys are since there may be many other options. In this case, to answer my question, some "information" is needed.

While I am searching everywhere in my office to see where my keys are and thinking of where they could possibly be, my supervisor came in and let me know that: "I saw your keys near the printer in the photocopying room downstairs two minutes ago". With this "message", many "possible" places where my keys could be are excluded. This message confirms that my keys are in specific place. This message can be considered the definite and reliable message since it confirms accurately where my keys are. As a result, a possibility E will occur with the probability x with $0 \leq x \leq 1$ where $x = 0$ means that this possibility will not be realised and $x = 1$ means that this possibility is definitely realised.

Now, the ideas of the so-called "information content" of a definite and reliable message are discussed. If, before the message is received, I never thought that my keys could be in the photocopying room, the information content of the message is very high because it is really outside my expectation. However, if the photocopying room is one of the places I am thinking of, the information context should be lower than in the first case since, at least, it is also listed in the selected places I am searching for my keys. When x is close to 0, say, $x = 0.01$, the information content of the message is very large.

However, when x is close to one, say, $x = 0.95$, the message has provided a little information content. To formalise these ideas, let $h(x)$ be the information content of a definite and reliable message x. It is obvious that $h(x)$ will be the decreasing function of the probability x. This is because "the more unlikely the event before the message on its realisation, the larger the information content" (Theil, 1967, p.3). Among many different decreasing functions, the logarithm of the reciprocal of the probability x is widely used.

(2.1) $$h(x) = \log \frac{1}{x} = -\log x$$

The other reason for the logarithmic function to be selected among many decreasing functions is the additivity of this function in the case of independent events. Suppose that E_1 with probability x_1 and E_2 with probability x_2 are stochastically independent, their product $x_1.x_2$ is the probability that both events occur. In this case, the information content of the message which informs us that "*both events did occur*", $h(x_1, x_2)$, will be as follows:

(2.2) $$h(x_1, x_2) = \log \frac{1}{x_1.x_2} = \log \frac{1}{x_1} + \log \frac{1}{x_2} = h(x_1) + h(x_2)$$

The far right-hand side of the equation (2.2) includes the information content of the message telling us that "Event E_1 occurred", $h(x_1)$, and the information content of the other message of "Event E_2 occurred", $h(x_2)$. As a consequence, as the equation (2.2) shows, the information content of the message which informs us that "*both events* did occur" is the sum of the information content of "Event E_1 occurred" and the information content of "Event E_2 occurred". This additivity is a very convenient property of definition in equation (2.1).

5.2.2 The entropy as the information content

In light of the previous discussion, it is clear that different values of probabilities x_i of the event E_i will provide different meanings. In short, it means that the lower the probability of an event occurring, the larger the "information content" of a message.

Until the message is released, no one can predict how significant the "information content" will be as either $h(x_1), or\ h(x_2),..., or\ h(x_n)$ with different probabilities $x_1 \neq x_2 \neq ... \neq x_n$ can occur. However, the *average or expected information content* can be calculated before the message arrives, since we know the probabilities. In this sense, the expected information content of the message is just the expected value of the information content, that is, the probability weighted average of $h(x_1), h(x_2),..., h(x_n)$:

(2.3) $$H(x) = \sum_{i=1}^{n} x_i h(x_i) = \sum_{i=1}^{n} x_i \log \frac{1}{x_i} = -\sum_{i=1}^{n} x_i \log x_i$$

Since x_i is a probability for a particular event to occur, it follows that $0 \leq x_i \leq 1$ and $\log x_i$ will always be negative. As the product of $x_i \log x_i$ is always negative, $\sum_{i=1}^{n} x_i \log x_i < 0$. Therefore, the negative of this sum, $H(x)$, cannot be negative. In other words, $H(x)$ cannot be negative since it is the weighted average, with all non-negative weights $x_1, x_2,..., x_i$, of the non-negative information values $h(x_1), h(x_2),..., h(x_n)$. The measure $H(x)$ is the expected information of a distribution, which Theil calls "entropy". In addition, the value of the entropy $H(x)$ has a lower limit of zero and the upper limit of $\log n$, where n represents a number of events or possibilities, so that $0 \leq H(x) \leq \log n$. Appendix A.5.1 at the end of the chapter establishes this range for $H(x)$.

5.2.3 Entropy, uncertainty and dispersion

The measure $H(x)$, defined in equation (2.3), is known as the expected information content or the expectation of information. It is developed from the notion of the probability of occurrence of certain events. Based on the limits of this entropy, $0 \leq H(x) \leq \log n$, it is said that, prior to the presence of a message which states that A occurred, the more uncertainty there is, the larger the expected information content of the message. As a consequence, entropy $H(x)$ can also be used to measure uncertainty of an event or an outcome. When an event is certain to occur, its probability is unity. There is no uncertainty, and $H(x)=0$, the lower limit, in this case. On the other hand, for a given number of events, uncertainty is at its maximum level when all events have the same probability, $1/n$, of occurrence. This case corresponds with the upper limit of the expected information content $H(x)=\log n$. Moreover, the level of uncertainty will increase with an increase in the number of outcomes n. For example, if there are only two possible outcomes, the probability of $1/2$ for each outcome presents less uncertainty than in the case with 20 possible outcomes which carries a probability of $1/20$ to occur. In other words, the more equi-likely events that can occur, the more uncertainty there is.

In addition, the entropy $H(x)$ can also be used to measure dispersion. The variance is the most common approach to measure dispersion of the distribution. The variance of a continuous random variable with a probability distribution $f(x)$ is defined as: $\sigma^2=\int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx$, where $\mu=\int_{-\infty}^{\infty}xf(x)dx$ is the mean. In the discrete case, entropy is defined as the negative value of the expected logarithms of event probabilities: $H(x)=-\sum_{i=1}^{n}x_i\log x_i$. When x is continuous, entropy is the negative value of expectation of the logarithms of the density: $H(x)=-\int_{-\infty}^{\infty}f(x)\log f(x)dx$.

To illustrate, suppose x is normally distributed, with the mean μ and variance σ^2, so that:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1/2(x-\mu)^2}{\sigma^2}}, \text{ so that } \log f(x)=-\log\sigma\sqrt{2\pi}-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}$$

The entropy now becomes:

$$H(x) = -\int_{-\infty}^{\infty} f(x)\left\{-\log\sigma\sqrt{2\pi} - \frac{(1/2)(x-\mu)^2}{\sigma^2}\right\}dx$$

$$= \left(\log\sigma\sqrt{2\pi}\right)\int_{-\infty}^{\infty} f(x)dx \;+\; \frac{1}{2}\int_{-\infty}^{\infty}\left(\frac{x-\mu}{\sigma}\right)^2 f(x)dx$$

$$= \log\sigma\sqrt{2\pi} + \frac{1}{2} \;=\; \log\sigma + \left(\frac{1}{2} + \log\sqrt{2\pi}\right).$$

Thus, the entropy of a normal distribution is the sum of the logarithm of the standard deviation σ and a constant equal to $1/2 + \log\sqrt{2\pi}$. Since $1/2 = (1/2)\log e$, the relationship between the entropy and the variance σ^2 of the normal distribution can also be expressed as: $H(x) = (1/2)\log\left(2\pi e\sigma^2\right)$. This shows that the entropy is an increasing function of the variance in the case of the normal distribution. Even though when things are not normally distributed, however, the general idea that the entropy measures dispersion continues to hold.

In conclusion, the entropy $H(x)$ can be used to measure the expected information content, the uncertainty and the dispersion. The entropy $H(x)$ is developed based on the concepts of probability alone, so it can take both numerical values (say, $0.1,\ 0.5, ...$) and "nominal" values (say, "rich" and "poor"). It sheds light on the view that the entropy $H(x)$ is in contrast to the variance since the variance can only take the numerical values.

The concept of expected information for a direct and reliable message, known as a "direct message", as discussed in Sections 5.2 above, guarantees that, among many possibilities, with the presence of a message, at least one event, or possibility, will occur. When a message is released, we know with certainty what happened. Recalling the example of my lost keys, the message of "I saw your keys near the printer in the photocopying room downstairs two minutes ago" provided the certain conclusion that the event did occur: *My keys were found*. However, this type of direct message may not be the case for all situations in life. Further analysis on the example of my lost keys below will explore this aspect – an indirect message.

5.2.4 An indirect message: prior and posterior probabilities

Let the previous example now be revised. Instead of receiving the message from my supervisor that "I saw your keys near the printer in the photocopying room downstairs two minutes ago", I receive a different message of "I did remember I saw your keys somewhere in our school building one minute ago". After this message is released, I have general information on where my keys could be, but not direct information on the specific place where they could be found. This has information content: the chance of finding my keys somewhere in the school building becomes more probable while the chance of finding them in University House, or anywhere else outside our school building becomes less probable upon the message being released. Theil calls this type of information an *indirect message*.

When we take one possibility into consideration, an indirect message does not confirm any event but it does provide additional information regarding an event that may occur in the future. If so, then the expected information content will change. This is because, with the release of the message, some events have a higher chance of occurring and others have a lower probability of occurring, no guarantee of an event is provided with the release of the message. Similar to previous discussions, it is assumed we have n chances as $E_1, E_2, ..., E_n$ with the probabilities to occur are $x_1, x_2, ..., x_n$, respectively. These probabilities are known as *prior* probabilities since they existed *before* the message comes in. When the message comes in, these probabilities will be changed because with the presence of the message, some chances become more probable to occur and others become less probable to occur. The probabilities for these events $E_1, E_2, ..., E_n$ to occur become $y_1, y_2, ..., y_n$, respectively. These are called as *posterior* probabilities (Theil, 1967). As a result, the sum of these posterior probabilities is unity. That means:

(2.4) $$\sum_{i=1}^{n} y_i \quad = \quad 1, \quad y_i \geq 0 \quad \forall\, i = 1, 2, ..., n.$$

These posterior probabilities are also non-negative. If it turns out that one of these probabilities is one, all the others are zero, then the message becomes a direct one since this message guarantees one particular event with probability of unity occurs. Recall from equation (2.1) regarding the information content, we will then apply for the event E_i to occur with the probabilities *before* and *after* the message is released (i.e. its *prior* and *posterior* probabilities) are x_i and y_i, respectively.

"Probability ex post" is the probability of the event to occur *after* the message is released. In this case, we will not know what happens for sure with the release of the message. And the probability in this case is y_1. In addition, "probability ex ante" is the probability of the event to occur *before* the message is released, still x_i in this case. Therefore, the information content in the case of an "indirect message" is as follows:

(2.5) $$h(y_i, x_i) = \log \frac{y_i}{x_i},$$

or in words: The information received with message = $\log\left(\frac{\text{probability ex post}}{\text{probability ex ante}}\right)$.

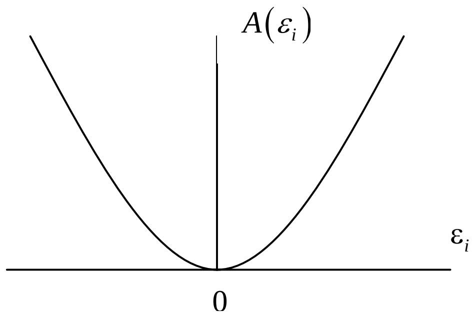
It is important to note that the message itself does not mention any possibility or event E_i in particular. This means that the presence of the message does not guarantee the occurrence of any event. Any event has its own *posterior* probability y_i to occur. In this case, the expected information of the indirect message is as follows:

(2.6) $$I(y:x) = \sum_{i=1}^{n} y_i \log \frac{y_i}{x_i}.$$

The expected information of an indirect message $I(y:x)$ transforms the prior probabilities $x_1, x_2, ..., x_n$ into the posterior probabilities $y_1, y_2, ..., y_n$. And, $I(y:x)$ is non-negative, which can be shown as follows. It is assumed that $y_i > 0$, $i = 1, 2, ..., n$ in the first instance. In addition, let us assume that there exists a small number $\varepsilon_i\ (i = 1, 2, ..., n)$, such that $\sum_{i=1}^{n} y_i \varepsilon_i = 0$. In this case, the equation $x_i = y_i(1+\varepsilon_i)$ holds, or equivalently: $x_i / y_i = 1 + \varepsilon_i$. Equation (2.6) can then be rewritten as follows:

(2.7) $$I(y:x) = -\sum_{i=1}^{n} y_i \log(1+\varepsilon_i).$$

Since $\sum_{i=1}^{n} y_i \varepsilon_i = 0$, equation (2.7) can be rewritten as $I(y:x) = \sum_{i=1}^{n} y_i \left[\varepsilon_i - \log(1+\varepsilon_i)\right]$. In proving that $I(y:x)$ is non-negative, because $y_i \geq 0$, it is only necessary to prove that $A(\varepsilon_i) = \left[\varepsilon_i - \log(1+\varepsilon_i)\right] \geq 0$. Taking the first-order derivative of $A(\varepsilon_i)$ is: $dA/d\varepsilon_i = 1 - 1/(1+\varepsilon_i) = \varepsilon_i/(1+\varepsilon_i)$. This is obvious that this derivative disappears when $\varepsilon_i = 0$. In addition, the derivative $dA/d\varepsilon_i$ is positive when $\varepsilon_i > 0$ and negative when $\varepsilon_i < 0$. However, regardless of the value of ε_i, negative or positive, the function $A(\varepsilon_i)$ is always positive as long as ε_i is a small number. For example, when $\varepsilon_i = 0.3$, $A(0.3) = \left[0.3 - \log(1+0.3)\right] = 0.0376 \geq 0$. In addition, when $\varepsilon_i = -0.3$, then $A(-0.3) = \left[-0.3 - \log(1-0.3)\right] = 0.0567 \geq 0$. It is clear that $A(\varepsilon_i) = 0$ when $\varepsilon_i = 0$, and $A(\varepsilon_i) > 0$ when $\varepsilon_i \neq 0$, and the function looks like below. In short, the function $A(\varepsilon_i) \geq 0$, so that $I(y:x) \geq 0$ and the equality sign holds when and only when each ε_i disappears, that is, when $x_i = y_i$ for all i. It means that the expected information of an indirect message disappears when all probabilities are left unchanged.



It is important to further note that, as previously discussed, in the case where $x_1 = x_2 = ... = x_n = 1/n$, the entropy is at its maximum value. That is $H(x) = \log n$. In this case, the expected information content of an indirect message $I(y:x)$ is:

(2.8) $$\log n - \sum_{i=1}^{n} y_i \log \frac{1}{y_i} = \log n - H(y).$$

Equation (2.8) tells us that, in a special case for equal prior probabilities, the expected information of an indirect message is the difference between the maximum value ($\log n$) of the entropy of the posterior probabilities, and the actual value of the entropy H(y).

In addition, the expected information of an indirect message $I(y:x)$ as in equation (2.6) can be named as the information inaccuracy. This is because the message transforms the prior probabilities (before a realisation of an event) into posterior probabilities (after a realisation of an event). The presence of the posterior probabilities reveals how allocation of occurrence among events actually took place. When the message has a zero expected information (i.e. $I(y:x)=0$), we have $x_i = y_i$, where $i=1,2,...,n$. In this case, the forecast is perfect. As a result, the higher the expected information of an indirect message is, the more inaccurate the forecast is. Further analysis and conclusion can be found in the following numerical examples.

5.2.5 A numerical example

The following example illustrates the expected information content of the indirect message. It is assumed that there are two events E_1 and E_2 which occur with probabilities x_1 and x_2, respectively. If one of these two events must happen, then $x_1 + x_2 = 1$. It is also assumed that each of these events has the same chance to occur, so that $x_1 = x_2 = 0.5$. Suppose the message S states that event E_1 occurred, then, $x_1 = 1$ and $x_2 = 0$. The statement S confirming the occurrence of event E_1 is called the direct message.

It is now assumed that the above statement S does not confirm that E_1 will occur. It only provides some more evidence that E_1 is more likely to occur. As such, prior probabilities x_1 and x_2 are transformed into the posterior probabilities y_1 and y_2. With the release of the message S, no event is confirmed to occur. As a result, it is further assumed that $y_1 = 0.7$. It means that, with the presence of the statement S, the probability for the event E_1 to occur increases from 0.5 to 0.7. Accordingly, the probability for the event E_2 to occur decreases from 0.5 to 0.3. The statement S is no longer the direct message because, with its presence, no event is guaranteed to occur. The only effect of the message S is to change the probabilities of related events. The expected information content of this statement S, as in case 4 in the following table, is:

$$I(y:x) = y_1 \log \frac{y_1}{x_1} + y_2 \log \frac{y_2}{x_2}$$
$$= 0.7 \log \frac{0.7}{0.5} + 0.3 \log \frac{0.3}{0.5} = 0.0375$$

Table 5.1 explores the changes of the expected information of an indirect message at various levels of both prior and posterior probabilities for both events E_1 and E_2. The difference between prior probabilities x_1, x_2 and posterior probabilities y_1, y_2 reveals the inaccuracy of the indirect message: the higher the difference, the more inaccurate the message is.

TABLE 5.1
THE EXPECTED INFORMATION CONTENT OF AN INDIRECT MESSAGE

Case	Probability $(\times 100)$				I (y:x) $(\times 100)$	Difference $(y_1 - x_1)$ $(\times 100)$
	Prior		Posterior			
	x_1	x_2	y_1	y_2		
1	50	50	50	50	0.00	0
2	50	50	80	20	8.37	30
3	50	50	75	25	5.68	25
4	50	50	70	30	3.57	20
5	50	50	65	35	1.98	15
6	50	50	60	40	0.87	10

For example, as in the first case, given in case 1, the indirect message plays no role because both prior and posterior probabilities for events E_1 and E_2 are unchanged. In addition, as illustrated by cases 2 and 3, when the probability of occurrence for Event E_1 increases from 0.50 to 0.80 (or 0.75), probability of occurrence of event E_2 decreases from 0.5 to 0.2 (or 0.25), the expected information of the indirect message falls from 8.37

per cent (case 2) to 5.68 per cent (case 3). As a consequence, as presented by the last two columns, the expected information content of an indirect message $I(y:x)$ decreases as the difference between the posterior and prior probabilities, $(y_i - x_i)$, decreases.

5.2.6 The expected information content

The following section explores the link between the expected information content of an indirect message with both prior and posterior probabilities, being weighted by respective posterior probabilities. Because the sums of prior or posterior probabilities are both unity, the expected information content of an indirect message could be expressed as the weighted sum of these two probabilities. From equation (2.6), the expected information content of an indirect message is the sum of n terms involving x_i and y_i. The x_i and y_i are prior and posterior probabilities of an event E_i to occur and $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i = 1$. Suppose that $y_i > x_i$ for each i, so that $y_i - x_i > 0$ for each i. This is contrary to the fact that the sum of both sets of probabilities is unity. As a result, n terms in equation (2.6) must consist of some negative terms and some positive terms so that $y_i > x_i$ for some i and $y_j < x_j$ for some j where $i \neq j$. We start with the function in logarithms $\log(y_i / x_i)$ which we express as:

(2.9) $$\log \frac{y_i}{x_i} = -\log\left[1 + \frac{x_i - y_i}{y_i}\right].$$

For convenience, let $a = (x_i - y_i)/y_i$, so that we can write equation (2.9) as $\log(y_i/x_i) = -\log(1+a)$. Function $f(a)$ can be expanded as Maclaurin series:

(2.10) $$f(a) = f(0) + \frac{f^{'}(0)}{1!}a + \frac{f^{''}(0)}{2!}a^2 + \frac{f^{'''}(0)}{3!}a^3 + \frac{f^{4}(0)}{4!}a^4 + \ldots$$

With $f(a) = -\log(1+a)$, and $a = (x_i - y_i)/y_i$, we have $f(0) = 0$, $f^{'}(0) = -1$, $f^{''}(0) = 1$, $f^{'''}(0) = -2$ and $f^{4}(0) = 6$. Using these values in equation (2.10), we then obtain:

$$(2.11)\quad -\log\left(1+\frac{x_i - y_i}{y_i}\right) = -\frac{x_i - y_i}{y_i} + \frac{1}{2}\left(\frac{x_i - y_i}{y_i}\right)^2 - \frac{1}{3}\left(\frac{x_i - y_i}{y_i}\right)^3 + \frac{1}{4}\left(\frac{x_i - y_i}{y_i}\right)^4 - \ldots$$

The above expansion converges if $(x_i - y_i)/y_i < 1$, or $x_i < 2y_i$. The first term of the right-hand side of equation (2.11) is worth considering. If we multiply it by y_i and take the sum, we have: $-\sum_{i=1}^{n} y_i\left[(x_i - y_i)/y_i\right] = -\sum_{i=1}^{n}(x_i - y_i) = 0$. The expected information content now becomes:

$$(2.12)\quad I(y:x) = \sum_{i=1}^{n} y_i \log\frac{y_i}{x_i} = \frac{1}{2}\sum_{i=1}^{n}\frac{(x_i - y_i)^2}{y_i} - \frac{1}{3}\sum_{i=1}^{n}\frac{(x_i - y_i)^3}{y_i^2} + \frac{1}{4}\sum_{i=1}^{n}\frac{(x_i - y_i)^4}{y_i^3} - \ldots$$

From these results, the expected information content of an indirect message can also be used to represent information inaccuracy because it translates the prior probability into posterior probability: the higher the differences between these two probabilities $x_i - y_i$ are, the more inaccurate the information is.

As discussed in Chapter 3, many previous attempts to measure the degree of fiscal decentralisation involve the use of some form of share of revenue/expenditure at lower-level jurisdictions in the national total. It is the claim of this chapter that such an approach completely ignores important distributional aspects of fiscal arrangements. Consider two hypothetical economies, A and B. In both economies, government spending and revenue at the national level accounts for 50 per cent of the total, so that the remaining 50 per cent is the responsibility of SNGs. In country A, there are only two large subnational governments, each with an equal share of total subnational fiscal activity (i.e. 50 per cent each); while in country B there are 100 subnational units, each accounting for 1 per cent of the 50 per cent total. It is clear that there is substantially more fiscal decentralisation in B as compared to A. However, an exclusive focus of the split of the total between the national and subnational levels would lead one to erroneously conclude that both economies exhibit the same degree of fiscal decentralisation. In other words, both the first and second moments of the distribution of revenue/expenditure are important for understanding the workings of fiscal arrangements.

The ideas of expected information of a direct message and an indirect message were originally developed by Theil in his influential book *Economics and Information Theory* in 1967. These ideas were further developed to measure the income inequality by comparing the income share with the population share of the states. These works lay a strong foundation for the development of an analytical framework of fiscal inequality which takes into account the dispersions of the revenue and expenditure of various levels of SNGs. The following section is devoted to this development.

5.3 Analytical framework for the analysis of subnational fiscal inequality

In his influential study, Theil (1967) advocated the use of entropy-based measurement for the analysis of income inequality. In this section, we apply Theil's notion of entropy, as outlined in Section 5.2, to public finances in multi-tiered governments. The analysis that follows is devoted to the development of an analytical framework which reveals SNGs' fiscal inequality in terms of revenue shares among SNGs. The same framework can be directly applied to the expenditure shares among SNGs. The notion of fiscal inequality (or fiscal dispersion) is important for fiscal theory on decentralisation because it accounts for the heterogeneity of various subnational units in terms of revenue and expenditure shares. However, it should be emphasised that fiscal inequality and fiscal equalisation are two distinct concepts, in that fiscal equalisation is not designed to redress the notion of fiscal inequality in this chapter. Specifically, the concept of fiscal inequality in this chapter relies on "money" (such as revenue and expenditure of subnational governments) as the unit of comparison, whereas the fiscal equalisation process (such as that adopted in Australia) is concerned with equalising the capacity of SNGs to provide the same "real" level of service.

It is assumed that a country has P states (the second level of government) and Q local councils (the third level of government) and each local council belongs to one state. Let $N = P + Q$ be a total number of local and state governments, the number of subnational governments (SNGs). It is further assumed that each subnational government accounts for a non-negative fraction of total subnational revenue, to be denoted by r_i which, for short, we shall refer to as the "regional revenue share". The sum of all these revenue shares is equal to unity: $\sum_{i=1}^{N} r_i = 1, \quad r_i \geq 0 \quad \forall\, i = 1,\ldots,N$. Let $\boldsymbol{r}$ denote the vector of revenue shares $r_1,..,r_N$. The entropy of revenue shares is defined as:

(3.1) $$H(\boldsymbol{r}) = \sum_{i=1}^{N} r_i \log \frac{1}{r_i}.$$

Entropy $H(\boldsymbol{r})$ can be regarded as the measure of the equality with which revenue is distributed among the SNGs. When the revenue distribution is extremely equal in that each SNG has the same revenue share (i.e., $r_i = 1/N$) and revenue entropy is at its maximum: $H(\boldsymbol{r}) = \log N$. At the other extreme, when only one SNG collects all SNGs' revenue so that others have no revenue at all (i.e., $r_i = 1$ and $r_j = 0$ for $i \neq j$), the minimum value of the entropy is achieved: $H(\boldsymbol{r}) = 0$. As a result, the range of the entropy is $0 \leq H(\boldsymbol{r}) \leq \log N$.

In the context of considering the relevance of the distribution of revenue among SNGs for its impact on fiscal decentralisation, it is appropriate to focus on revenue inequality between SNGs, rather than revenue equality. Revenue inequality is measured by deducting the revenue entropy, $H(\boldsymbol{r})$, from its maximum value, $\log N$:

(3.2) $$\log N - H(\boldsymbol{r}) = \log N - \sum_{i=1}^{N} r_i \log \frac{1}{r_i} = \sum_{i=1}^{N} r_i \log N r_i.$$

Due to the constraints on the range of the entropy $H(\boldsymbol{r})$, it is clear that the range of this measurement of revenue inequality is 0 - perfect equality (when $H(\boldsymbol{r}) = \log N$) - and $\log N$ - maximum inequality (when $H(\boldsymbol{r}) = 0$). The entropy $H(\boldsymbol{r})$ is an attractive way to measure equality as it satisfies three axioms or tests described below.

5.3.1 Axiom 1: The proportionality test

The entropy (3.1) is expressed in terms of the revenue shares of SNGs. Thus, if all revenues change proportionally, the shares do not change, and measure (3.2) remains unchanged. This invariance of revenue inequality to a proportional change is the proportionality test.

5.3.2 Axiom 2: The "Haves and Have Nots" test

The upper limit of $H(\boldsymbol{r})$ increases with N, so the maximum value of the inequality measure (3.2) rises with N. Consider two hypothetical countries. *First*, in a two-subnational-region country, there is perfect inequality when one SNG accounts for all revenue, and the other has no revenue. The entropy of the revenue shares is zero, and the value of (3.2) is $\log 2$. *Second*, in a society consisting of 10,000 SNGs, revenue inequality is at maximum when 9,999 SNGs have no revenue. The value of revenue inequality is now $\log 10,000$. It is obvious that revenue distribution in the latter is much more unequal than the first country. In the first country, one-half of the SNGs (one SNG) accounts for all subnational revenue and the other half has no revenue. As a result, revenue inequality of the second country is as unequal as for the first country when one-half of the SNGs account for all subnational revenue and when each of these has the same revenue. The concern is that whether revenue inequality, as expressed in equation (3.2), satisfies this condition. The following material reveals that this is true by showing that as a larger fraction of SNGs join the "revenue" group, revenue inequality falls. This establishes that revenue inequality will be uniquely determined by the size of the revenue group (which we call "the haves") relative to the "no-revenue" group ("the have nots").

Assume there is a set S which consists of M subnational governments where $0 < M \leq N$. It is further assumed that SNGs in set S account for all subnational revenue, so that SNGs outside set S have no revenue. Also, within set S, each SNG accounts for the same amount of revenue (i.e., for $i \in S$, $r_i = 1/M$.). The inequality measure (3.2) then becomes:

$$\sum_{i=1}^{N} r_i \log Nr_i = \sum_{i \in \boldsymbol{S}_i} r_i \log Nr_i = \frac{1}{M} \log N \frac{1}{M} + \frac{1}{M} \log N \frac{1}{M} + ...,$$

or:

(3.3) $$\sum_{i=1}^{N} r_i \log Nr_i = \log \frac{N}{M} = \log \frac{1}{\theta} ,$$

where $\theta = M/N$ is the fraction of SNGs in the country who jointly account for all subnational revenue. The application of the last member of equation (3.3) to the second example above with $N = 10,000$ and $\theta = 5,000/10,000 = 1/2$, reveals that revenue inequality is also $\log 2$.

From these two examples, we can conclude that when revenue is equally distributed among some groups of SNGs in the society, and the remaining SNGs outside these groups have no revenue, revenue inequality of the country is determined solely by the fraction θ - the ratio of the number of SNGs in the group to the total number of SNGs. In both examples above, this ratio is $1/2$, and the revenue inequality is $\log 2$. This result is consistent with intuition: when the number of SNGs receiving revenue, M, increases, revenue distribution becomes more equal. The above discussion shows that as the inequality (3.3) decreases as the share of a number of SNGs which receive revenue rises, this measure satisfies the "Haves and Have Nots" axiom.

5.3.3 Axiom 3: The revenue transfer test

Consider an economy consisting of two SNGs only A (rich) and B (poor) with the revenue shares r_A and r_B, where $r_A > r_B$. Suppose that some revenue is transferred from A to B, such that $dr_A + dr_B = 0$. A reasonable measure of revenue inequality should indicate that such a transfer from the rich SNG to the poor SNG has the effect of decreasing inequality. Does equation (3.2) satisfy this property? The following material shows that it does have this property.

It is assumed that there are G sets of SNGs, to be denoted by $\boldsymbol{S}_1,..,\boldsymbol{S}_G$, and each SNG belongs to one and only one set. Let N_g be a number of SNGs in set $\boldsymbol{S}_g$, with $\sum_{g=1}^{G} N_g = N$. To give some practical significance to the symbols, consider a three-tiered government: tier 1 - national government; tier 2 - state government; and tier 3 – local government. $\boldsymbol{S}_g$ represents the set of state and local governments in the geographical region defined by the jurisdiction of State g. N_g is the total number of state and local governments within the jurisdiction defined by State g. In view of this, the entropy of revenue shares, equation (3.1), can now be expressed as:

(3.4)
$$H(\boldsymbol{r}) = \sum_{g=1}^{G}\left[\sum_{i\in \boldsymbol{S}_g} r_i \log\frac{1}{r_i}\right],$$

where the component inside the square brackets is the entropy of revenue shares within set $\boldsymbol{S}_g$. Let R_g be the sum of revenue shares of all SNGs in set $\boldsymbol{S}_g$, $R_g = \sum_{i\in \boldsymbol{S}_g} r_i$; this R_g is the revenue share of group g with $\sum_{g=1}^{G} R_g = 1$. The entropy of revenue shares within set $\boldsymbol{S}_g$ can be expressed as:

$$\sum_{i\in \boldsymbol{S}_g} r_i \log\frac{1}{r_i} = R_g\left[\sum_{i\in \boldsymbol{S}_g} \frac{r_i}{R_g}\left(\log\frac{1}{r_i/R_g}\times\frac{1}{R_g}\right)\right]$$
$$= R_g \sum_{i\in \boldsymbol{S}_g} \frac{r_i}{R_g}\log\frac{1}{r_i/R_g} + R_g\log\frac{1}{R_g}.$$

Thus, if we define $H_g\left(\boldsymbol{r}_g\right) = \sum_{i\in \boldsymbol{S}_g}\frac{r_i}{R_g}\log\frac{1}{r_i/R_g}$, where $\boldsymbol{r}_g$ is the vector of r_i that fall under $\boldsymbol{S}_g$, as the within-set entropy, we have:

(3.5)
$$\sum_{i\in \boldsymbol{S}_g} r_i \log\frac{1}{r_i} = R_g H_g\left(\boldsymbol{r}_g\right) + R_g\log\frac{1}{R_g}.$$

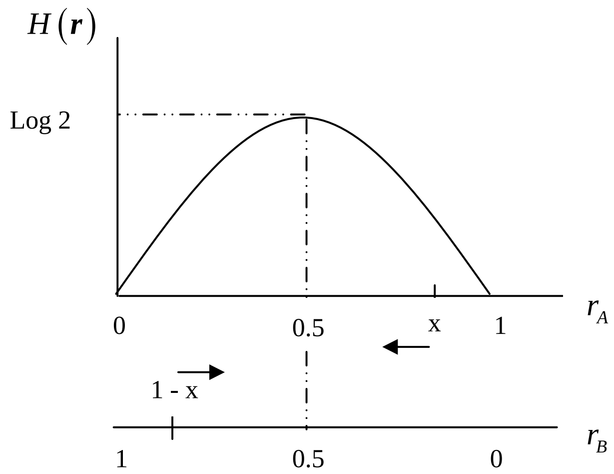
Combining equations (3.4) and (3.5), the total entropy becomes:

(3.6)
$$H(\boldsymbol{r}) = \sum_{g=1}^{G} R_g H_g\left(\boldsymbol{r}_g\right) + \sum_{g=1}^{G} R_g\log\frac{1}{R_g}.$$

On the right-hand side of this equation, the first component is a weighted average of the within-set entropies $H_1(\boldsymbol{r}_1),\ldots,H_G(\boldsymbol{r}_G)$, with the group revenue shares $R_1,\ldots,R_G$ as the weights. The second term on the right of equation (3.6) is <u>the between-set entropy</u>, $\sum_{g=1}^{G} R_g \log\left(1/R_g\right)$.

We consider equation (3.6) in the context of SNGs A (rich) and B (poor) in two situations: (i) when they are the only SNGs of the country, so that $N=2$; and (ii) when the nation is made up of A, B plus all other SNGs, so that $N>2$. When $N=2$, the country comprises two groups, $S_1=A$, and $S_2=B$, which we shall denote by S_A and S_B. Similarly, the revenue shares are $R_1=r_A$ and $R_2=r_B$, with $r_A+r_B=1$. As there is only one SNG in each group, the within-group entropies are zero, $H_A(r_A)=H_B(r_B)=0$, as is their weighted average. Accordingly, in this case, equation (3.6) simplifies to:

$$H(\boldsymbol{r}) = r_A \log\frac{1}{r_A} + r_B \log\frac{1}{r_B}.$$



This entropy is at its maximum when $r_A=r_B=1/2$. In that case, the entropy is $H(\boldsymbol{r})=1/2\log 2+1/2\log 2=\log 2$, as is illustrated below. From the graph, it is clear that any deviations from the equal shares of $r_A=r_B=1/2$ will result in a lower value of the entropy, that is, higher revenue inequality. As A is richer than B, the initial revenue distribution is represented in the graph by the shares $x > 1/2$ and $(1-x) < 1/2$. When revenue is transferred from A to B, both revenue shares move towards $1/2$, the distribution becomes more equal and the entropy increases.

Next, consider the $N > 2$ case where there are three groups of SNGs: (i) Group A with only one SNG A; group B with SNG B; and (iii) group C with $(N-2)$ SNGs comprising every SNG in the economy except A and B. These three groups are denoted by $\boldsymbol{S}_A$, $\boldsymbol{S}_B$, and $\boldsymbol{S}_C$. We assume that the joint revenue share of A and B is a constant, i.e. $r_A + r_B = R_{A+B} =$ constant. This implies that the revenue share of group C, R_C, is also constant at $1 - \mathrm{R}_{\mathrm{A+B}}$. It is further assumed that there are no revenue transfers to or from the other SNGs of the society in $\boldsymbol{S}_C$. We now apply decomposition (3.6) to this economy. The weighted average of the within-group entropies, the first term on the right-hand side of equation (3.6), is:

(3.7) $$\sum_{g=1}^{G} R_g H_g(\boldsymbol{r}_g) = R_A H_A(r_A) + R_B H_B(r_B) + R_C H_C(\boldsymbol{r}_C) = R_C H_C(\boldsymbol{r}_C).$$

where $H_C(\boldsymbol{r}_C) = \sum_{i \in \boldsymbol{S}_C} \frac{r_i}{R_C} \log \frac{1}{r_i / R_C}$, with $\boldsymbol{r}_C$ is the vector of r_i that fall under group $\boldsymbol{S}_C$, is the within-group entropy of group C. The first and second components in the second step of equation (3.7), the within-group entropies for groups A and B, disappear because there is only one SNG in each group. In addition, the between-group entropy, the second term on the right-hand side of equation (3.6), now becomes:

(3.8) $$\sum_{g=1}^{G} R_g \log \frac{1}{R_g} = R_A \log \frac{1}{R_A} + R_B \log \frac{1}{R_B} + R_C \log \frac{1}{R_C}.$$

Substituting equations (3.7) and (3.8) into equation (3.6), the total entropy for this three-group country becomes:

(3.9) $$H(\boldsymbol{r}) = R_A \log \frac{1}{R_A} + R_B \log \frac{1}{R_B} + R_C \log \frac{1}{R_C} + R_C H_C(\boldsymbol{r}_C).$$

When we transfer revenue from A to B, with the distribution within $\boldsymbol{S}_C$ remaining unchanged, equation (3.9) can be expressed as:

(3.10) $$H(\boldsymbol{r}) = R_A \log \frac{1}{R_A} + R_B \log \frac{1}{R_B} + \text{constant}.$$

The constant in (3.10) includes $R_C \log(1/R_C)$ and $R_C H_C(\boldsymbol{r}_C)$. In words, the total entropy of the three-group country is equal to the total entropy of the two-group country plus a constant. Accordingly, the impact on inequality of a transfer from A to B is the same in the $N > 2$ case as it is in the $N = 2$ case.

To summarise this discussion, revenue inequality decreases if there is a transfer of revenue from the rich SNG to the poor SNG. This conclusion holds for a society with two-subnational regions $(N = 2)$, as well as in the higher-dimensional case $(N > 2)$. In short, it is clear that the measure of revenue inequality satisfies the revenue transfer test.

5.4 <u>Decomposing revenue inequality</u>

In the above, we decomposed revenue equality into within-set and between-set terms. We now show that revenue <u>inequality</u> can be similarly decomposed.

Recall from equation (3.6) that the entropy is decomposed into two distinct components: a weighted average of the within-set entropy and the between-set entropy. Furthermore, as in (3.2), inequality is measured by the difference between the maximum value of the entropy, $\log N$ and the entropy $H(\boldsymbol{r})$. Thus, by combining equations (3.2) and (3.6), revenue inequality can be expressed as:

(4.1) $$\log N - H(\boldsymbol{r}) = \log N - \sum_{g=1}^{G} R_g H_g(\boldsymbol{r}_g) - \sum_{g=1}^{G} R_g \log \frac{1}{R_g}.$$

The right-hand side of equation (4.1) remains unchanged if we subtract and add $\sum_{g=1}^{G} R_g \log N_g$, where R_g and N_g are the revenue share of and a number of SNGs in set $\boldsymbol{S}_g$, respectively:

$$\log N - H(\boldsymbol{r}) = \sum_{g=1}^{G} R_g \left(\log N_g - H_g(\boldsymbol{r}_g)\right) + \log N - \sum_{g=1}^{G} R_g \log \frac{N_g}{R_g}$$
$$= \sum_{g=1}^{G} R_g \left(\log N_g - \sum_{i \in \boldsymbol{S}_g} \frac{r_i}{R_g} \log \frac{1}{r_i / R_g} \right) + \sum_{g=1}^{G} R_g \log \frac{R_g}{N_g / N}.$$

As the result, revenue inequality can be expressed as follows:

(4.2) $$\log N - H(\boldsymbol{r}) = \sum_{g=1}^{G} R_g \left[\sum_{i \in \boldsymbol{S}_g} \frac{r_i}{R_g} \log \frac{r_i / R_g}{1/N_g} \right] + \sum_{g=1}^{G} R_g \log \frac{R_g}{N_g / N}.$$

Result (4.2) reveals that revenue inequality consists of two distinct components: (i) a weighted average of within-set inequalities and (ii) a between-set inequality. The right-hand side of equation (4.2) parallels the decompositions given by equation (3.6). The meaning of the two components of equation (4.2) is discussed further in what follows.

5.4.1 The within-set inequalities

The first component of (4.2) is a weighted average of the within-set inequalities:

(4.3) $$\sum_{g=1}^{G} R_g \left[\sum_{i \in \boldsymbol{S}_g} \frac{r_i}{R_g} \log \frac{r_i / R_g}{1/N_g} \right].$$

The term r_i / R_g is the conditional revenue share of SNG i within group $\boldsymbol{S}_g$, that is, SNG i's revenue share within the group. Also, N_g represents a number of SNGs in group $\boldsymbol{S}_g$. Equation (4.3) comprises two weighted averages: (a) $Z_g = \sum_{i \in \boldsymbol{S}_g} \frac{r_i}{R_g} \log \frac{r_i / R_g}{1/N_g}$, the within-set revenue inequality for group $\boldsymbol{S}_g$, and (b) $\sum_{g=1}^{G} R_g Z_g$, the weighted average of the within-set revenue inequalities. We discuss each in turn.

If each SNG in set $\boldsymbol{S}_g$ receives an equal revenue share, then $r_i / R_g = k$ (say). However, as $\sum_{i \in \boldsymbol{S}_g} \left(r_i / R_g \right) = 1$, it follows that $k = 1/N_g$. When each SNG has an equal share of the group's revenue, i.e., $r_i / R_g = 1/N_g$, $i \in \boldsymbol{S}_g$, then there is no dispersion of the revenue distribution within the group, the perfect equality. Accordingly, the extent to which the N_g ratios

$$(4.4) \qquad \frac{r_i / R_g}{1/N_g}, \qquad i = 1, \ldots, N_g$$

deviate from unity is a measure of revenue inequality within set $\boldsymbol{S}_g$. The within-set measure of revenue inequality, the term in square brackets of equation (4.3), is a weighted average of the logarithms of the ratios in equation (4.4), the weights being the conditional revenue shares.

5.4.2 The between-set inequality

The second term on the right-hand side of (4.2) is the between-set inequality:

$$(4.5) \qquad \sum_{g=1}^{G} R_g \log \frac{R_g}{N_g / N}.$$

The basic ingredient of inequality (4.5) is the contrast between two sets of shares, the revenue shares of the G groups, $R_1, \ldots, R_G$ and the corresponding population shares, $N_1/N, \ldots, N_G/N$. If all groups receive their pro-rata shares of revenue based on population, i.e. $R_g = N_g / N$, $g = 1, \ldots, G$, then there is no dispersion of revenue distribution and we have perfect between-set revenue equality.

In summary, total inequality consists of two components: the weighted average of the within-set inequality and the between-set inequality. Interestingly, it is clear that both components are of the form of the expected information content of an indirect message which was previously discussed in Section 5.2.4. For the within-set inequality, the prior and posterior probabilities are $1/N_g$ and r_i / R_g, respectively. Similarly, for a between-set inequality, N_g / N and R_g are prior and posterior probabilities. Furthermore, from equation (4.2), the revenue inequality, can be written as:

$$(4.6) \qquad \log N - H(\boldsymbol{r}) = \sum_{i=1}^{N} r_i \log N r_i = \sum_{i=1}^{N} r_i \log \frac{r_i}{1/N}.$$

The far right-hand side of equation (4.6) reveals that total revenue inequality can also be expressed in the form of the expected information content of an indirect message. In this case, the prior and posterior probabilities are $1/N$ and r_i, respectively. With this perspective, it is clear that the message that transforms the vector $[1/N,...,1/N]'$ into $[r_1,...,r_N]'$ is equivalent to two sub-messages. The first message transforms $[1/N_g,...,1/N_g]'$ into $[r_1/R_g,...,r_g/R_g]'$, $g=1,...,G,$ which could be called "the within-set message", and the second message transforms $[N_1/N,...,N_G/N]'$ into $[R_1,...,R_G]'$, which is "the between-set message".

The entropic analysis of fiscal arrangements can, of course, be extended to the expenditure shares of SNGs in exactly the same manner as applied above to revenue shares[2].

5.4.3 Why not per capita fiscal data?

The above fiscal inequalities are expressed in terms of the number of SNGs, rather than a number of individuals. That is to say, according to our approach, per capita fiscal data are not incorporated in the measurement of fiscal decentralisation. Why?

Consider the two hypothetical countries A and B. Country A has two local councils, each of the same size, whereas country B consists of 100 local councils (again, all of the same size). Government revenue generated by country A is equal to the sum of revenue of the 100 councils in country B. As it has many more local councils, country B is more fiscally decentralised as compared to country A. This conclusion is reasonable and stands independently of the size of the population in the two countries, and how the population is distributed across the 100 local governments in country B.

[2] Fiscal inequality based on SNGs' expenditure shares are presented in Chapter 6.

Next, consider a real-world example from the fastest-growing state in Australia, Western Australia ("WA"). In terms of total expenditure, the largest local government in WA is the City of Stirling, while the smallest is the Shire of Three Springs. Columns 2 and 3 of Table 5.2 show that expenditure in Stirling is almost 100 times greater than that in Three Springs. However, the population in Stirling is almost 250 times larger than that of Three Springs (columns 4 and 5). Spending per capita is therefore significantly higher in Three Springs than in Stirling, as presented in column 6.

TABLE 5.2
EXPENDITURE AND POPULATION IN TWO LOCAL COUNCILS
IN WESTERN AUSTRALIA, 2004

Local council	Expenditure		Population		Per capita expenditure		Fiscal weight
	$' '000	Per cent of WA total	Persons	Per cent of WA total	$	Deflated	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
City of Stirling	100,405	5.97	182,047	9.06	552	0.66	3.94
Shire of Three Springs	1,338	0.08	722	0.04	1,853	2	0.16

Source: Unpublished ABS data.

Let s_i be the expenditure share of local council $i\left(i=1,\ldots,143\right)$, and $S=\sum_{i=1}^{143}S_i$ be total expenditure, $s_i = S_i/S$ be the share, P_i be the population of i and $P=\sum_{i=1}^{143}P_i$ be total population, and $p_i = P_i/P$ be the corresponding share. Then the ratio of the expenditure share to the population share $\frac{s_i}{p_i}=\frac{S_i/S}{P_i/P}=\frac{S_i/P_i}{S/P}$ is "deflated" per capita expenditure of the i^{th} council. If all local councils receive their pro rata expenditure share based on population, then $s_i/p_i = 1$ for each i. Column 7 shows that deflated per capita expenditure of Stirling and Three Springs is 0.66 and 2, respectively, so that Stirling's weighting for fiscal calculations is 3.94 as shown in column 8 (i.e. Stirling's per cent of WA population multiplied by 0.66) and Three Springs' fiscal weighting would be 0.16 (i.e. Three Springs' per cent of WA population multiplied by 2.00). Accordingly, if per capita expenditure is considered, the smallest local government area, Three Springs, would, in effect, play a "relatively" more important role in measuring the degree of fiscal decentralisation in WA. Such an approach, while relevant to many issues in fiscal federalism (e.g. fiscal equalisation), may provide a misleading picture of the degree of fiscal decentralisation. In the context of measuring pure dispersion of revenue and expenditure among SNGs in monetary terms, the number of governments within a given tier appears to be a relevant indicator of decentralisation[3].

5.4.4 A note on notation

In the above discussion, the results are formulated in logarithmic terms. For future reference, it is convenient to take the antilogarithm of the inequality measure.

We start by expressing revenue inequality in terms of information theory as discussed in Section 5.2. Recall the second component on the right-hand side of equation (4.2), the between-set inequality, which is a weighted average of the logarithms of the ratios of the set revenue shares and the corresponding institutional shares, $\sum_{g=1}^{G} R_g \log \frac{R_g}{N_g/N}$. Let m_i and q_i be the revenue share and institutional share of the i^{th} region, that is, $m_i = M_i/M$, where M_i, M are the revenue of the i^{th} region and the total economy, and $q_i = Q_i/Q$, where Q_i, Q are the number of SNGs in the i^{th} region and the total number of SNGs in the economy. As a result, $\frac{m_i}{q_i} = \frac{M_i/M}{Q_i/Q} = \frac{M_i/Q_i}{M/Q}$. The numerator of this ratio is revenue per SNG of the i^{th} region, while the denominator is revenue per SNG. If $\boldsymbol{m} = [m_1,...,m_N]'$ and $\boldsymbol{q} = [q_1,...,q_N]'$, the between-region inequality can be expressed in terms of information theory as:

$$I(\boldsymbol{m}:\boldsymbol{q}) = \sum_{i=1}^{N} m_i \log \frac{m_i}{q_i}.$$

The ratio m_i/q_i is "deflated" per SNG revenue of the i^{th} set. The term "deflated" here means that revenue is expressed as relative to national revenue for SNG. The above

[3] When two tiers of SNGs are considered, the number of governments is still relevant because each set of governments (for between-set analysis) can be specified by hierarchical level (e.g. the "local" level and the "state" level).

$I(\boldsymbol{m}:\boldsymbol{q})$ is the logarithm of a weighted average of deflated revenue per SNG, so that the corresponding geometric mean is:

$$(4.7) \qquad e^{I(\boldsymbol{m}:\boldsymbol{q})} = \prod_{i=1}^{N} \left(\frac{m_i}{q_i} \right)^{m_i}.$$

If all SNGs receive their pro rata share based on a number of SNGs, then $m_i/q_i = 1$ for each i, $\prod_{i=1}^{N} (m_i/q_i)^{m_i} = 1$ and there is no revenue dispersion. Accordingly, the further the mean (4.7) is away from unity, the greater is revenue inequality across sets. Similarly, on the expenditure side, the geometric mean is:

$$(4.8) \qquad e^{I(\boldsymbol{s}:\boldsymbol{q})} = \prod_{i=1}^{N} \left(\frac{s_i}{q_i} \right)^{s_i}.$$

where $\boldsymbol{s} = [s_1, ..., s_N]'$ and $\boldsymbol{q} = [q_1, ..., q_N]'$ with s_i and q_i is the expenditure share and institutional share of the i^{th} region.

5.5 Concluding remarks

The fundamental index of fiscal decentralisation (FDI) and the enhanced index (eFDI) developed in Chapters 3 and 4 take into account the fundamental issues of fiscal decentralisation – fiscal autonomy and fiscal importance of SNGs together with the positive effect of intergovernmental fiscal transfers from the national government, however, those indices are insensitive with the different distributions of revenue and expenditure among SNGs and a number of SNGs. In response to these potential limitations, an entropic approach to the analysis of subnational fiscal inequality has been developed in this chapter.

One of the contributions of this chapter is illustrated with a simple example. Consider two hypothetical nations V and L which exhibit the same degree of fiscal decentralisation, using either the fundamental index (FDI) or the enhanced index (eFDI), (i.e. the structure of revenue, expenditure, total fiscal transfer, total unconditional transfer is the same in two countries). It is now further assumed that these two countries consist of four subnational regions: A, B, C and D, each with different levels of revenue (and expenditure). Table 5.3 provides data for this example.

TABLE 5.3
ILLUSTRATING FISCAL INEQUALITY

	Region/ Measures	Country V				Country L			
		Own-sourced revenue	Share in total			Own-sourced revenue	Share in total		
		($ millions)	Actual	Average	Difference	($ millions)	Actual	Average	Difference
	(1)	(2)	(3)	(4)	(5) = (4) – (3)	(6)	(7)	(8)	(9) = (8) – (7)
1.	A	3,000	0.010	0.250	0.240	3,300	0.011	0.250	0.239
2.	B	125,000	0.427	0.250	-0.177	271,390	0.926	0.250	-0.676
3.	C	97,000	0.331	0.250	-0.081	10,810	0.037	0.250	0.213
4.	D	68,000	0.232	0.250	0.018	7,500	0.026	0.250	0.224
5.	Total	293,000	1.000	1.000	0.000	293,000	1.000	1.000	0.000
6.	Standard deviation		0.178	0.000			0.451	0.000	
7.	Entropy		0.484	0.602			0.146	0.602	
8.	Fiscal inequality		0.118	0.000			0.456	0.000	

Column 2 shows that there is one small region in country V, region A. Revenue from region B is almost double that of D and forty times higher than that of region A. Columns 3 and 4 present the actual and average revenue shares for 4 regions in country V. By contrast, in country L, there is one large and three small regions. Region B accounts for more than 92 per cent of the total revenue of all regions, and the remaining 8 per cent is spread across the three small regions A, C, and D.

- Row 6 presents the standard deviations of the revenue shares of the two, 0.178 and 0.451. This clearly reveals that the distribution of revenue of country L is more dispersed than in V.
- Row 7 gives the values of the fiscal entropy, defined as $-\sum r_i \log r_i$, where r_i is the revenue share of SNG *i*. The entropy value in country V is 0.484 and 0.146 in country L, as shown in columns 3 and 7 of row 7, respectively. If we were to assume alternatively that each region accounts for the same share of 25 per cent,

as shown by columns 4 and 8, there is no inequality, so that fiscal entropy for both countries is $\log 4 = 0.602,$ as in row 8, columns 4 and 8.

- Row 8 presents the fiscal inequality, the difference between the maximum level of the entropy, $\log 4$, or 0.602, and the actual level. Fiscal inequality is 0.118 and 0.456 for countries V and L, respectively. Higher fiscal inequality in L means a greater degree of revenue dispersion among SNGs and, as a result, suggests a lower degree of fiscal decentralisation because revenue is allocated more disproportionately across regions[4].

To summarise this example, countries V and L may exhibit the same degree of fiscal decentralisation as previously discussed in Chapter 1[5]. But as there is much more fiscal inequality in country L, it can be reasonably concluded that the true situation may be different: there is less fiscally decentralised in country L. As such, the fundamental index and the enhanced index may be partial measures of fiscal decentralisation because they ignore the dispersion of revenue (and expenditure) across regions. The issue of subnational fiscal inequality and its relationship to the degree of fiscal decentralisation are investigated further in the context of Australia and Denmark in Chapter 6, by focusing on geographic-region and governmental-hierarchy sets of governments.

[4] To better appreciate the relationship between fiscal inequality and fiscal decentralisation, it is necessary to consider the distinction between "within-set" inequality and "between-set" inequality for geographic regions and political hierarchy, as discussed in the next chapter.

[5] To minimise repetition, this example is not reproduced here. For further details, refer to Table 1.1 of Chapter 1.

Appendix A5.1

THE RANGE OF THE ENTROPY

The Appendix shows that the entropy $H(x)$ falls in the range with a lower limit zero and the upper limit $\log n$, where n represents a number of events or possibilities.

For the lower limit, it is clear that when the event E_i occurs with certainty, $x_i = 1$, and $x_j = 0$ for all $i \neq j$. Thus the probability vector $\left(x_1, x_2, ..., x_i, ..., x_n\right) = (0, 0, ..., 1, ..., 0)$. Then, $x_i \log x_i = 0$ for $i = 1, ..., n$ and $-\sum_{i=1}^{n} x_i \log x_i = 0$. This establishes that the lower bound of $H(x)$ is zero if and only if $x_i = 1$ for some i.

Regarding the upper limit, the task now is to maximise the $\sum_{i=1}^{n} x_i \log x_i$, subject to $\sum_{i=1}^{n} x_i = 1$ where $0 \le x_i \le 1$. To do this, we formulate the Lagrangian function:

$$L\left(x_1, ..., x_n; \lambda\right) = -\sum_{i=1}^{n} x_i \log x_i - \lambda\left(\sum_{i=1}^{n} x_i - 1\right),$$

where λ is the Lagrangian multiplier. The first-order condition is $\partial L / \partial x_i = -\log x_i - 1 - \lambda = 0$. This is equivalent to $\log x_i = -\left(\lambda + 1\right)$. This equation shows that x_i is independent of i. This happens when and only when $x_1 = x_2 = ... = x_i = 1/n$. When $x_i = 1/n$, $H(x)$ takes its upper value of $\log n$.

CHAPTER 6
FISCAL INEQUALITY: AN APPLICATION FOR AUSTRALIA AND DENMARK

6.1 Introduction

The goal of this chapter is to utilise the notion of fiscal inequality, to identify factors that will need to be considered when the fundamental index (FDI) and the enhanced index (eFDI), as developed in Chapters 3 and 4, are extended in future research. This is done by analysing subnational fiscal inequality for Australia and Denmark. Besides Australia, a federal nation, Denmark is selected as this is the only other country whose fiscal data of revenue and expenditure on three levels of government are publicly available. In addition, Denmark is a unitary country which may reveal different implications for SNGs' fiscal inequality when compared to Australia. The entropic factors considered in this chapter account for fiscal dispersions between and within sets, where sets are defined first on a "geographic" basis (i.e. a defined area of land) and then, on a "governmental" basis (i.e. for two hierarchical tiers of subnational governments). The results of fiscal inequality in Australia and Denmark can be used to draw attention to fiscal inequality which may exist in other countries, and, as such, these factors should be reflected in measuring fiscal decentralisation.

Following the introduction, Section 6.2 discusses fiscal federalism in Australia with the emphasis on its revenue and expenditure patterns which lead to an analysis of geographic (regional) and governmental (hierarchical) fiscal inequality. Australia's data are used to illustrate these important notions. In addition, Section 6.3 analyses a similar analytical case study for Denmark which is, in principle, in contrast with Australia in terms of fiscal arrangements and roles of SNGs. Some interesting conclusions drawn from case studies for Australia and Denmark are discussed on Section 6.4, followed by concluding remarks in Section 6.5.

6.2 Australian fiscal federalism

This section applies fiscal inequality measures to Australia represented by equations (4.3), (4.5) and (4.6) from Chapter 5. These inequalities are particularly relevant to Australia because there are great regional fiscal disparities. We start with a brief description of fiscal arrangements in Australia.

The Commonwealth of Australia was established in 1901 as a Federation in which six self-governing British colonies became the six states of Australia. The main purpose of this unification was to form a strong and open country by eliminating tariffs on interstate trade. More than one century after its formation, modern Australia is still seen as a “young” country in comparison with many nations from the “old” world. Australia now consists of six states and two territories (hereafter referred to in aggregate as the “states”) with a total number of local councils of 700. The eight “states” of Australia are New South Wales (NSW), Victoria (Vic.), Queensland (QLD), South Australia (SA), Western Australia (WA), Tasmania (Tas.) and two territories, Northern Territory (NT) and Australian Capital Territory (ACT). The first tier of government is occupied by the Commonwealth Government. The second tier is represented by state governments. The third and lowest tier of government is represented by local councils.

In geographic terms, the three levels of government are not mutually exclusive. The geographic region associated with each state includes a state and many local governments. The geographic area associated with the Commonwealth government also includes state and local governments. The fiscal authority of different levels overlaps – residents in each local government are influenced by fiscal activities of local, state, and federal governments. However, ACT has no local governments because of its special nature of administration. The ACT government performs two roles: one as the “state” government and another role as the “local” government. On this basis, 700 local governments in Australia are allocated to seven “states”, namely NSW (192)[6], Vic. (79), QLD (125), SA (68), WA (143), Tas. (29), and NT (64).

[6] Due to the NSW amalgamation exercise in 2004, local government numbers post-2004 have declined.

6.2.1 Revenue and expenditure patterns

Table 6.1 reveals that the allocation of revenue and expenditure across local councils in Australia is significantly dispersed.

TABLE 6.1
REVENUE AND EXPENDITURE SHARES,
LOCAL COUNCILS, AUSTRALIA, AVERAGE: 2000-2004
(1,000*Per cent of total)

No.	Region	Number of local councils	Revenue shares					Expenditure shares				
			Mean	Median	Standard deviation	Min	Max	Mean	Median	Standard deviation	Min	Max
1.	NSW	192	1.744	0.847	2.082	0.013	12.487	1.704	0.913	1.848	0.015	8.386
2.	Vic.	79	2.559	1.758	2.410	0.237	16.629	2.784	1.946	2.172	0.273	13.578
3.	QLD	125	2.191	0.603	7.495	0.098	77.079	2.039	0.609	6.666	0.144	69.045
4.	SA	68	0.865	0.459	1.155	0.056	7.069	0.899	0.539	1.034	0.079	5.400
5.	WA	143	0.617	0.189	1.016	0.041	6.308	0.637	0.257	0.872	0.069	5.387
6.	Tas.	29	1.004	0.569	1.156	0.163	4.592	1.006	0.602	1.082	0.210	4.587
7.	NT	64	0.203	0.096	0.357	0.003	2.651	0.259	0.139	0.364	0.019	2.665

Source: Unpublished data from ABS. Data are averages over the period 2000–2004.

Brisbane City Council in Queensland is the largest local council in Australia with revenue and expenditure shares of 7.7 per cent and 6.9 per cent of all local councils. The second and third biggest councils are the Gold Coast Council (QLD) and Melbourne City (Vic.). On the other hand, Timber Creek (NT) is the smallest council with revenue and expenditure shares of around 0.0003 per cent and 0.0019 per cent, respectively. As in Figures 6.2-6.3, there is considerable dispersion of revenue and expenditure between and within states.

FIGURE 6.1
DISTRIBUTION OF LOCAL COUNCIL SHARES
AUSTRALIA, AVERAGE: 2000–2004, (1,000*Per cent of total)

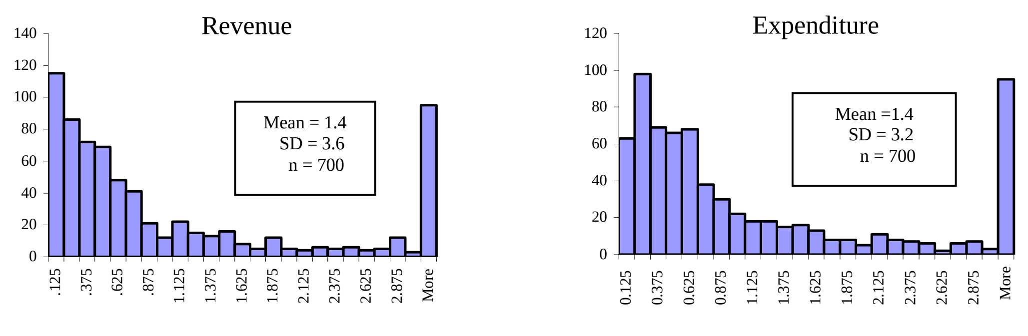


FIGURE 6.2
DISTRIBUTION OF REVENUE ACROSS LOCAL COUNCILS
AUSTRALIAN STATES, AVERAGE: 2000–2004
(1,000*Per cent of total)

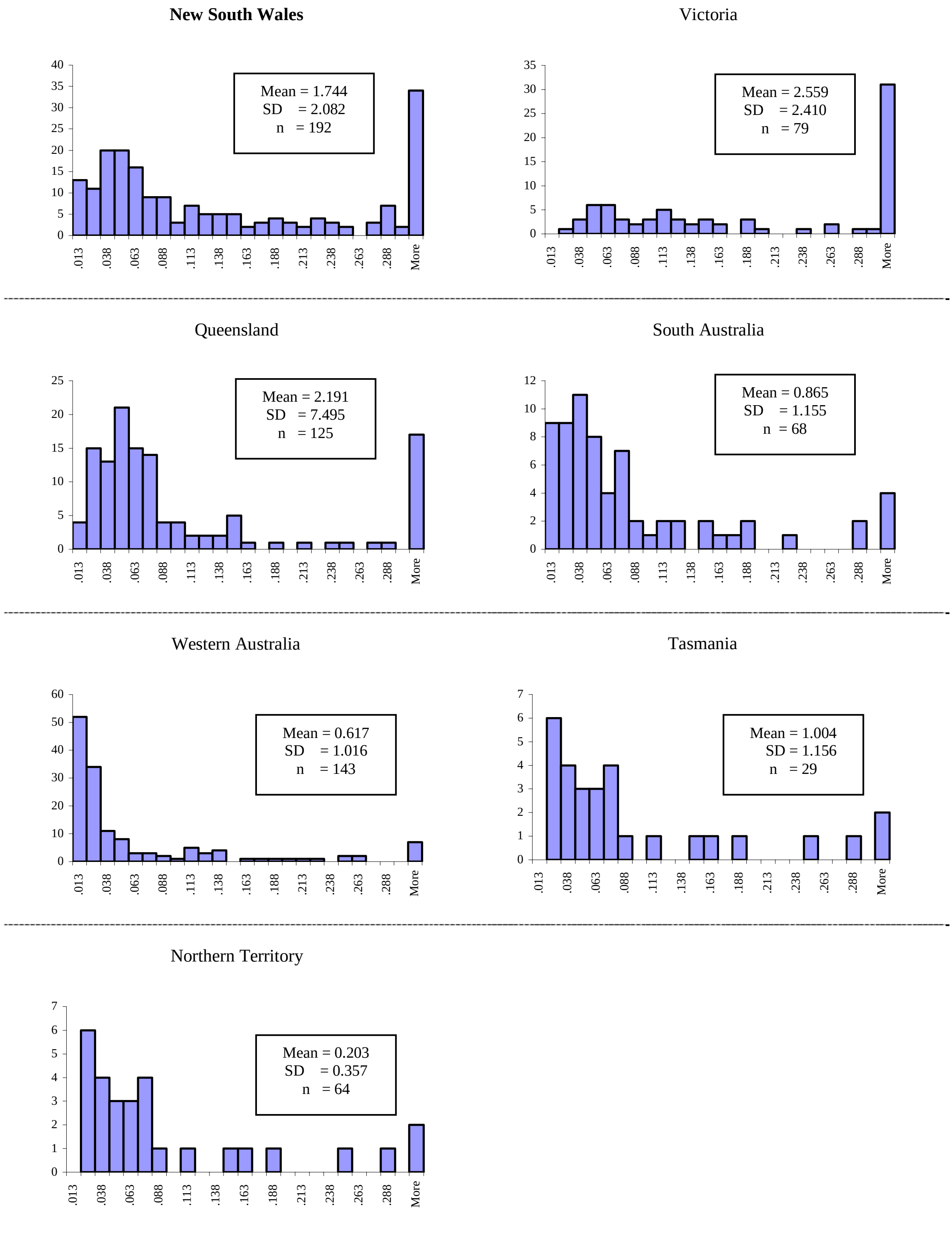
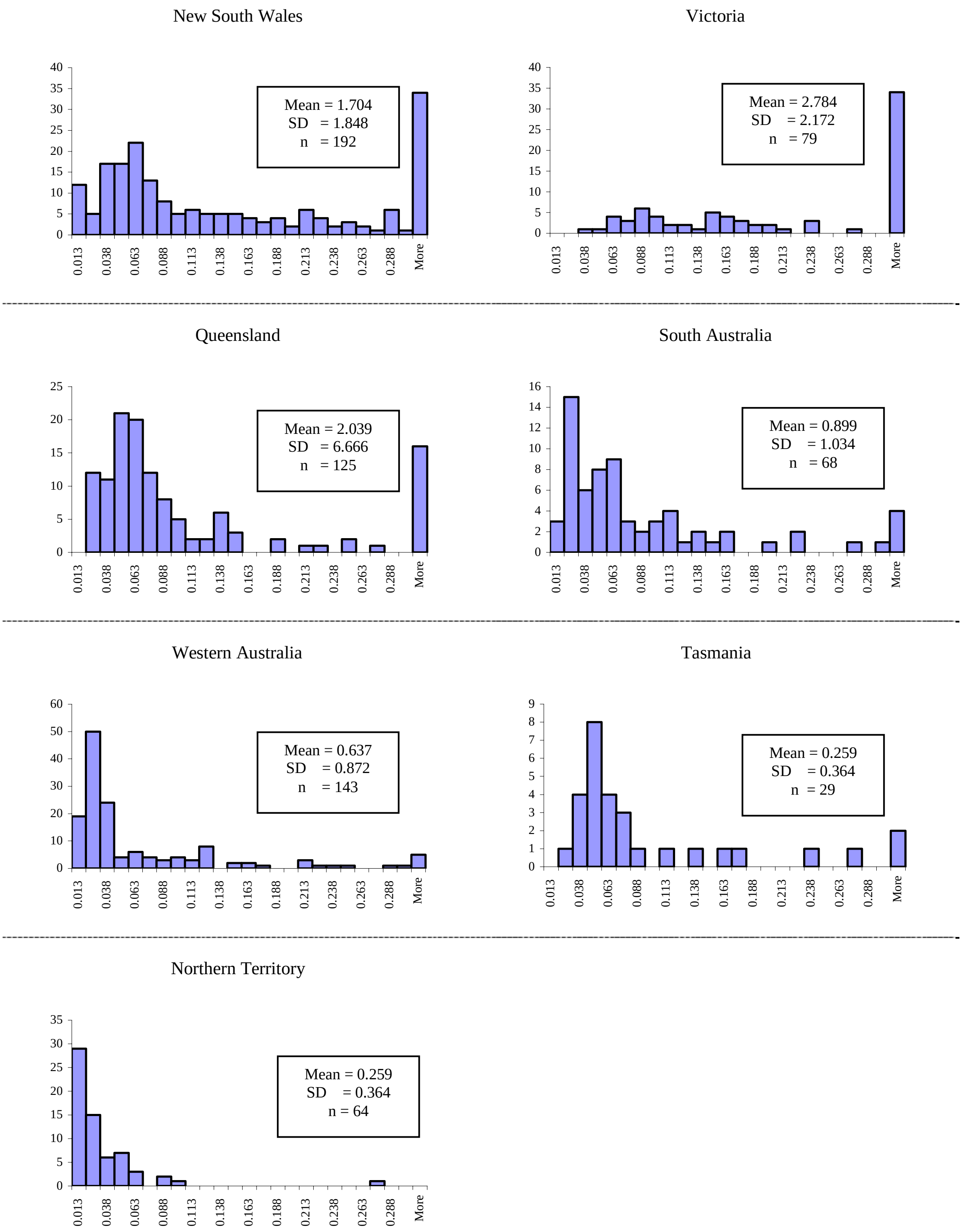


FIGURE 6.3
DISTRIBUTION OF EXPENDITURE ACROSS LOCAL COUNCILS
AUSTRALIAN STATES, AVERAGE: 2000–2004
(1,000*Per cent of total)



6.2.2 Subnational fiscal inequality by geographical region and governmental hierarchy

Table 6.2 provides a framework for the analysis of fiscal inequality when the sets of SNGs are identified by geographically defined region. For this study, each subnational region consists of the geographical region defined by state borders and comprises the state government and a number of local governments within these borders. Each row of the table represents one of the G regions (geographical states) in the country. Consider region g as an example. As indicated in column 2, there are n_g local councils in this region plus one state government, so there are $n_g + 1$ revenue shares, $r_{g1},..., r_{gn_g}, r_{g,n_g+1}$. Total revenue for g is $\sum_{k=1}^{n_g+1} r_{g,k} = R_g$, as presented in column 3. These total regional shares sum over the G regions to unity, that is, $\sum_{g=1}^{G} R_g = \sum_{g=1}^{G} \sum_{k \in \boldsymbol{S}_g} r_{gk} = 1$, as indicated by the last element of column 3. Column 4 of the table presents the number of all SNGs in each region, $N_1,..., N_G$, as well as the total number in the whole economy, N.

TABLE 6.2
THE ANALYTICS OF THE GEOGRAPHIC ALLOCATION OF REVENUE

Region g	Revenue shares of subnational region: Individual shares	Revenue shares of subnational region: Total	Number of subnational regions (state and local councils)
(1)	(2)	(3)	(4)
1	$r_{11},..., r_{1n_1}, r_{1,n_1+1}$	$R_1 = \sum_{k=1}^{n_1+1} r_{1k}$	$N_1 = n_1 + 1.$
⋮	⋮	⋮	⋮
g	$r_{g1},..., r_{gn_g}, r_{g,n_g+1}$	$R_g = \sum_{k=1}^{n_g+1} r_{gk}$	$N_g = n_g + 1.$
⋮	⋮	⋮	⋮
G	$r_{G1},..., r_{Gn_G}, r_{G,n_G+1}$	$R_G = \sum_{k=1}^{n_g+1} r_{Gk}$	$N_G = n_G + 1.$
Total		$\sum_{g=1}^{G} \sum_{k \in \boldsymbol{S}_g} r_{gk} = 1$	$N = \sum_{g=1}^{G} N_g$

To apply the above framework to the Australian case, we have $G = 7$, as there are six states and one territory that contain local councils (the ACT is excluded as it has no local councils). Each of the seven SNGs contains one state government and a number of local councils. Take Western Australia as an example. As there are 143 local councils in this state, there are 143 revenue shares $r_{g,1},..., r_{g,143}$, for $g = WA,$ while the revenue share for the WA state government is $r_{g,144}$. The total of these 144 shares, $\sum_{k=1}^{144} r_{g,k} = R_g$, is the share of national revenue accounted for by WA. For Australia as a whole, there are 700 local councils and 7 states, so $N = 707$.

In accordance with the analysis in Sections 5.3 and 5.4 of chapter 5, total revenue inequality for Australia with $N = 707$ and $G = 7$ is:

$$\log 707 - H(\boldsymbol{r}) = \sum_{g=1}^{7} R_g \left[\sum_{i \in \boldsymbol{S}_g} \frac{r_i}{R_g} \log \frac{r_i / R_g}{1/N_g} \right] + \sum_{g=1}^{7} R_g \log \frac{R_g}{N_g / 707}.$$

The first component on the right-hand side is the within-state inequality for revenue shares of local and state governments:

$$\sum_{g=1}^{7} R_g \left[\sum_{i \in \boldsymbol{S}_g} \frac{r_i}{R_g} \log \frac{r_i / R_g}{1/N_g} \right],$$

The second component on the right-hand side is the between-set inequality:

$$\sum_{g=1}^{7} R_g \log \frac{R_g}{N_g / 707}.$$

Table 6.3 reveals that within-state fiscal inequality accounts for 96.4 per cent and 96.9 per cent total inequality in terms of revenue and expenditure, respectively. Clearly, the within-state fiscal inequality plays a significant role in total inequality of the distribution of revenue and expenditure across subnational regions in Australia. This is partly because each subnational region includes both state and local governments, and the state government is significantly larger than any local government within the same region. For example, the total subnational share of Australian revenue collected from NSW (i.e. R_{NSW}) was 31.7 per cent in 2004, of which 27.1 per cent came from the NSW government (r_{NSW}) and only 4.6 per cent from the 192 local governments within NSW. Another possible reason for the dominance of the within-state component of fiscal inequality is the operation of the system of fiscal equalisation in Australia. Fiscal equalisation has a "tendency" to equalise per capita revenue and expenditure among states, which causes the between-state inequality to be low, or the within-state inequality to be high.

TABLE 6.3
FISCAL INEQUALITIES ACROSS SUBNATIONAL REGIONS
BY GEOGRAPHIC REGION
AUSTRALIA, 2004

Inequality measure	Revenue	Expenditure
Total inequality	1.727	1.763
Between-region inequality	0.063	0.054
Within-region inequality (WRI)	1.664	1.709
Inequality within:		
New South Wales	0.573	0.622
Victoria	0.348	0.343
Queensland	0.365	0.340
South Australia	0.106	0.123
Western Australia	0.228	0.215
Tasmania	0.029	0.034
Northern Territory	0.015	0.032
WRI as the percentage of total inequality	96.4	96.9

Source: Author's calculations.

Total inequality can also be disaggregated in terms of governmental hierarchy with the two sets being considered as: (i) the upper-tier SNGs, the set consisting of the seven states and territories; and (ii) the lower-tier SNGs, the 700 local councils. Table 6.4 presents the results when fiscal inequality is decomposed in this way.

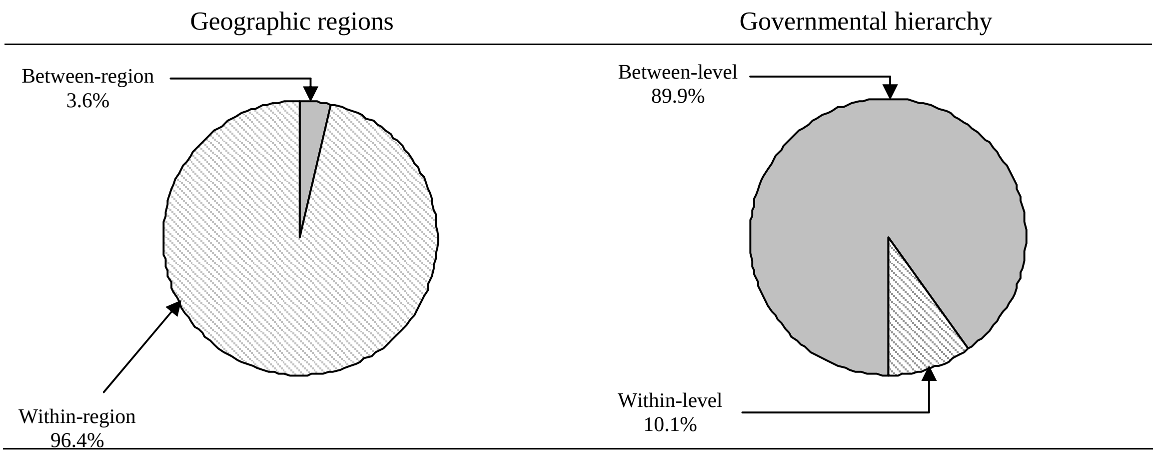
The results show that when local councils and states are completely isolated in this way, the fiscal inequality between the set of state governments and the set of local governments is much larger than the fiscal inequality within the set of states and within the set of local governments. In short, fiscal inequality between sets of the state and local governments accounts for 89.9 per cent of revenue inequality and 91.5 per cent of expenditure inequality. This result also reflects the ideas discussed in the previous paragraph.

TABLE 6.4
FISCAL INEQUALITIES ACROSS LEVELS OF GOVERNMENTS BY GOVERNMENTAL HIERARCHY
AUSTRALIA, 2004

Inequality measure	Revenue	Expenditure
Total inequality	1.727	1.763
Between-level of government inequality (BLI) (i.e. between state and local governments)	1.552	1.613
Within-level of government inequality (i.e. within state and within local governments)	0.175	0.150
Inequality within:		
State governments	0.118	0.108
Local governments	0.057	0.042
BLI as the percentage of total inequality	89.9	91.5

Source: Author's calculations.

FIGURE 6.4
FISCAL INEQUALITY OF REVENUE BY COMPONENT
AUSTRALIA, 2004



The above results have important implications for the measurement of fiscal decentralisation when the indices developed in Chapters 3 and 4 are extended to reflect the dispersion of revenue and expenditure among SNGs. *First*, from the governmental hierarchy perspective, most fiscal "dispersion" is due to fiscal difference between state and local governments. This is consistent with decentralised fiscal constitutions. This is to be expected as state governments are more fiscally important than local governments. This type of dispersion provides a limited but positive impact on measuring fiscal decentralisation. *Second*, the fiscal dispersion due to within-state government and within-local government inequality should have more significant impact on measuring fiscal decentralisation. This reflects the view that the nation is considered fiscally decentralised if fiscal dispersion of revenue and expenditure is identical for subnational units of the same level (i.e. state government versus state government and local government versus local government). When this is not the case, the fundamental and enhanced indices may need to be revised. This will need to be the subject of future research.

6.3 Denmark and fiscal decentralisation

The analytical framework of SNGs' fiscal inequality is now applied to Denmark – a unitary nation with a high per capita income level. The reasons for Denmark's application in the study are as follows. *First*, a test of the framework of SNGs' fiscal inequality is conducted for another country with different fiscal arrangements to Australia (i.e. federal versus unitary nations). *Second*, Denmark is the only country for which data necessary for every level of government are available. Third, SNGs' fiscal inequality in Australia can be "compared" with another country so that some implications of SNGs' fiscal inequality can be drawn.

There are three "autonomous" political and administrative levels in the current structure of the Danish government. *First*, the national government is a parliamentary system of government. Second, the regional level is represented by the counties. Third, the local authorities represent the lowest level of government – the local level. Regional and local governments play an important role and there is very close cooperation between national and subnational governments (Denmark's Government, 2003). At the regional level, Denmark consists of 16 counties, which include Copenhagen (a city) and Frederiksberg (a borough) without local governments. It is generally accepted that the two regional governments of Copenhagen and Frederiksberg also serve as local governments[7]. Table 6.5 presents a summary of 14 counties (excluding the city of Copenhagen and the borough of Frederiksberg) and 273 local municipalities in Denmark.

[7] See Appendix A6.1 for an alternative treatment.

TABLE 6.5
A SUMMARY OF COUNTIES AND MUNICIPALITIES
DENMARK, 2002

No.	County	Abbreviation	A number of local municipalities
1	Copenhagen	COP	18
2	Frederiksborg	FRE	19
3	Roskilde	ROS	11
4	West Zealand	WES	23
5	Storstrøm	STO	24
6	Bornholm	BOR	5
7	Funen	FUN	32
8	South Jutland	SOU	23
9	Ribe	RIB	14
10	Vejle Amtskommune	VEJ	16
11	Ringkøbing	RIN	18
12	Århus	ARH	26
13	Viborg	VIB	17
14	North Jutland county	NOR	27
	Total	14	273

Source: Statistics Denmark (2005).

6.3.1 Distribution of revenue and expenditure across counties

Figure 6.5 reveals the distribution of revenue and expenditure shares of 273 local municipalities in 14 counties in Denmark. Consequently, the distribution of revenue collection and expenditure across the various municipalities are different. As such, fiscal inequality among municipalities is, in terms of revenue and expenditure, expected to be significant. Table 6.5 provides the descriptive statistics regarding revenue shares and expenditure shares for all 14 counties in Denmark.

FIGURE 6.5
DISTRIBUTION OF LOCAL MUNICIPALITY SHARES
DENMARK, AVERAGE: 1990–2002 (1000* Per cent of total)

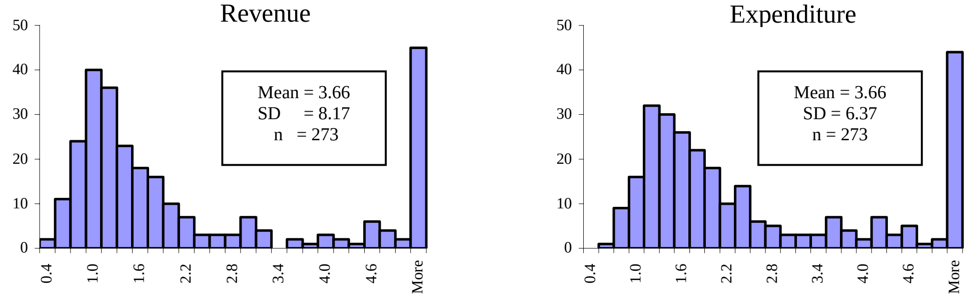


TABLE 6.6
LOCAL MUNICIPALITIES, DENMARK
REVENUE AND EXPENDITURE SHARES, AVERAGE: 1990-2002
(1,000*Per cent of total)

No.	Region	Number of municipalities	Revenue shares					Expenditure shares				
			Mean	Median	Standard Deviation	Min	Max	Mean	Median	Standard Deviation	Min	Max
1	COP	18	6.7526	6.5001	3.8681	1.5027	15.583	7.6390	8.0376	3.9513	1.8763	14.728
2	FRE	19	3.9608	3.0043	3.5472	0.9109	15.234	4.0062	3.7275	3.0670	0.9848	14.564
3	ROS	11	4.2588	1.6566	5.4968	1.1106	18.874	4.2442	2.1691	4.2583	1.3000	13.659
4	WES	23	2.4998	1.6515	2.3295	0.5768	8.754	2.6694	1.7430	2.0739	0.9387	7.948
5	STO	24	2.1879	1.1834	2.3871	0.4919	9.762	2.2607	1.4348	2.0774	0.7508	9.834
6	BOR	5	1.8144	1.3254	1.3290	1.0509	4.177	1.9281	1.5156	1.0311	1.2292	3.739
7	FUN	32	3.1022	1.1189	8.3822	0.5532	48.138	3.1435	1.2733	7.5450	0.6833	43.487
8	SOU	23	1.5908	1.1556	1.3757	0.3269	5.132	2.0083	1.6310	1.5282	0.5560	6.159
9	RIB	14	3.5430	1.3769	7.1086	0.6241	27.915	3.3745	1.7879	5.2296	0.6700	21.147
10	VEJ	16	4.9382	1.6671	6.6169	0.6583	20.142	4.5859	2.1899	5.0546	0.8864	13.976
11	RIN	18	3.4026	1.2277	5.5243	0.4508	23.298	3.0514	1.6300	3.4567	0.6482	14.619
12	ARH	26	6.6623	1.2009	20.2353	0.6666	103.107	5.6366	1.6550	14.7973	0.8513	76.282
13	VIB	17	2.1162	0.9827	2.3414	0.6410	9.209	2.5667	1.3493	2.2544	1.0697	9.034
14	NOR	27	3.7911	1.3856	8.9113	0.6666	47.286	3.7737	1.9123	7.1839	0.6021	38.768

Source: Statistics Denmark.

6.3.2 Fiscal inequality: Geographic region versus governmental hierarchy analysis

The same framework for the analysis of SNGs' fiscal inequality developed in Section 6.2.2 for Australia is again used. However, this time $G = 14$ (not 16) because the city of Copenhagen and the borough of Frederiksberg have been excluded since they have no local councils. Each of the 14 regions contains one county government and a number of local municipalities. Let's take the region of Arhus as an example. This region consists of the county government of Arhus and 26 local municipalities.

When considered from a regional (i.e. geographically defined) perspective, within-region inequality accounts for a substantial level in total fiscal inequality in Denmark, 89.6 per cent and 90.8 per cent for revenue shares and expenditure shares, respectively. The full geographic region allocations of fiscal inequalities across counties in Denmark are presented in Table 6.7 as follows.

TABLE 6.7
FISCAL INEQUALITIES ACROSS SUBNATIONAL REGIONS BY
GEOGRAPHIC REGIONS, DENMARK, 2002

Inequality measure	Revenue	Expenditure
Total inequality	0.3847	0.3274
Between-region inequality	0.0398	0.0300
Within-region inequality (WRI)	0.3449	0.2974
Inequality within:		
Copenhagen county	0.0222	0.0222
Frederiksborg county	0.0129	0.0139
Roskilde county	0.0096	0.0083
West Zealand county	0.0195	0.0151
Storstrøm county	0.0161	0.0143
Bornholm county	0.0004	0.0004
Funen county	0.0491	0.0481
South Jutland county	0.0060	0.0104
Ribe county	0.0160	0.0128
Vejle Amtskommune county	0.0192	0.0178
Ringkøbing county	0.0087	0.0118
Århus county	0.1110	0.0740
Viborg county	0.0096	0.0101
North Jutland county	0.0446	0.0382
WRI as the percentage of total inequality	89.60	90.80

Source: Author’s calculations.

Now, as previously discussed, total inequality can be also disaggregated in a governmental hierarchy. In this approach, all 14 counties (regional government level) are grouped into one set and all 273 local municipalities are grouped into the other set.

TABLE 6.8
HIERARCHICAL ALLOCATION OF
FISCAL INEQUALITIES ACROSS LEVELS OF GOVERNMENT
DENMARK, 2002

Inequality measure	Revenue	Expenditure
Total inequality	0.3847	0.3274
Between-level of government inequality (BLI)	0.1108	0.1310
Within-level of government inequality	0.2739	0.1964
Inequality within:		
County governments	0.0216	0.0156
Municipality governments	0.2523	0.1808
BLI as the percentage of total inequality	28.8	40.0

Source: Author’s calculations.

Table 6.8 presents the results of total fiscal inequality in Denmark when the hierarchical approach is used. In contrast with a geographic region analysis, this approach reveals that fiscal inequality of between-level of government inequality no longer plays a significant role in total inequality: only 28.8 per cent and 40.0 per cent total fiscal inequality for revenue and expenditure shares, respectively.

6.4 Fiscal inequality of Australia and Denmark – a comparison

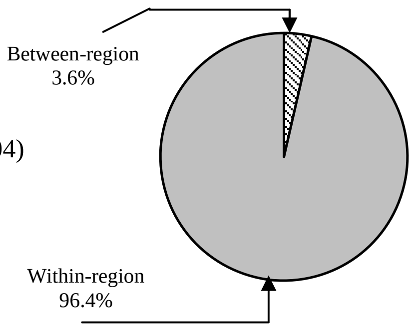
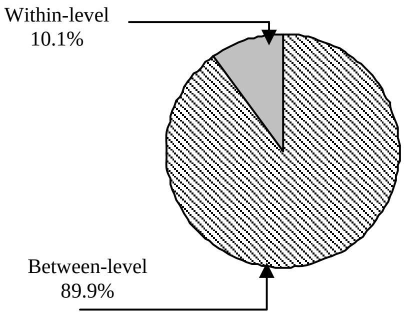
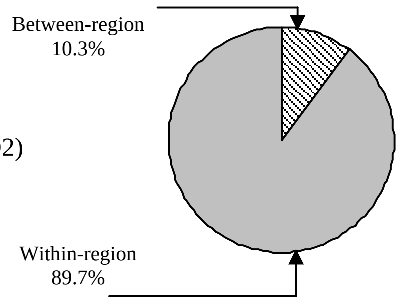
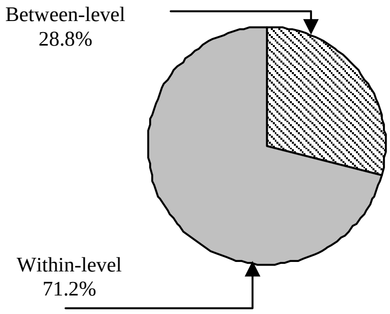
Sections 6.2 and 6.3 analyse fiscal inequality, in terms of revenue and expenditure shares among subnational units, in both countries – Australia and Denmark. In Australia, subnational units include 7 state governments and 700 local governments whereas subnational units in Denmark consist of 14 counties and 273 local municipalities. Table 6.9 presents a summary of the results.

TABLE 6.9
FISCAL INEQUALITY IN AUSTRALIA AND DENMARK

	Australia (2004)		Denmark (2002)	
Inequality measure	Revenue	Expenditure	Revenue	Expenditure
Total inequality	1.727	1.763	0.3847	0.3274
1 *Geographic regions*				
Between-region inequality ("BRI")	0.063	0.054	0.0398	0.0300
Within-region inequality ("WRI")	1.664	1.709	0.3449	0.2974
BRI as the percentage of total inequality (%)	3.6	3.1	10.3	9.2
WRI as the percentage of total inequality (%)	96.4	96.9	89.7	90.8
2 *Governmental hierarchy*				
Between-level inequality ("BLI")	1.552	1.613	0.1108	0.1310
Within-level inequality ("WLI")	0.175	0.150	0.2739	0.1964
BLI as the percentage of total inequality (%)	89.9	91.5	28.8	40.0
WLI as the percentage of total inequality (%)	10.1	8.5	71.2	60.0

In the case of Australia, the within-region inequality accounts for a significant degree of fiscal inequality when subnational units (state and local governments) are considered on a geographical basis (96.4 per cent for revenue and 96.9 per cent for expenditure). However, when sets are defined by state and local governments (i.e. by levels of government), the within-level inequality of government is small (10.1 per cent for revenue and 8.5 per cent for expenditure). That is, fiscal inequality among state governments is relatively modest, and fiscal inequality among local governments is also relatively modest. In this case, fiscal inequality is dominated by the fiscal differences between governments of different levels. In short, state government finances are relatively large in comparison with all local governments.

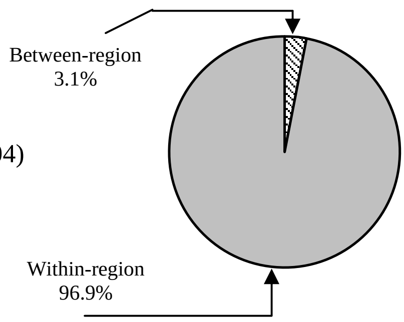
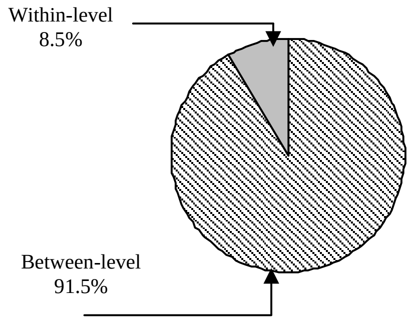
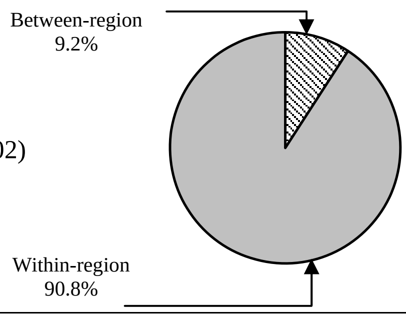
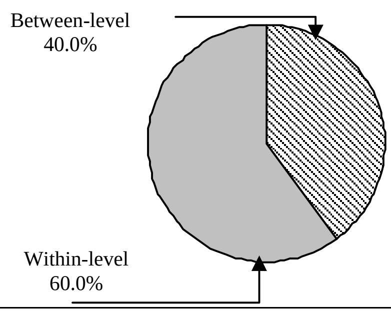
TABLE 6.10
FISCAL INEQUALITY OF SUBNATIONAL REVENUE
AUSTRALIA (2004) VERSUS DENMARK (2002)

Country	Fiscal inequality of subnational revenue	
	Geographic region	Governmental hierarchy
Australia (2004)		
Denmark (2002)		

The pattern of fiscal inequality in Denmark is different to that of Australia. In the geographic-region analysis, within-region fiscal inequality in Denmark accounts for a significant share of fiscal inequality (89.7 per cent for revenue and 90.8 percent for expenditure), but not quite as much as in the case of Australia. However, in contrast to Australia, within-level fiscal inequality accounts for a significant share of fiscal inequality when governmental hierarchy approached is considered (71.2 per cent for revenue and 60 per cent for expenditure). These results suggest regional governments in Denmark (i.e. counties) are not significantly larger compared with local municipalities.

This is in contrast to Australia, where the fiscal size of local governments is much smaller than state governments.

TABLE 6.11
FISCAL INEQUALITY OF SUBNATIONAL EXPENDITURE
AUSTRALIA (2004) VERSUS DENMARK (2002)

<table>
<tr><th rowspan="2">Country</th><th colspan="2">Fiscal inequality of subnational expenditure</th></tr>
<tr><th>Geographic region</th><th>Governmental hierarchy</th></tr>
<tr><td>Australia (2004)</td><td>
</td><td>
</td></tr>
<tr><td>Denmark (2002)</td><td>
</td><td>
</td></tr>
</table>

Tables 6.9, 6.10, and 6.11 reveal the importance of fiscal inequality (or fiscal dispersion). In a generally decentralised federal system, the fiscal difference between state level and lower-level governments will, typically, be large (i.e. first-tier SNGs will be fiscally larger than second-tier SNGs). In entropic terms, this means that fiscal inequality in federal systems will be dominated by: (i) “between-level of government” inequality, when considered from a governmental hierarchy perspective; and (ii) “within-region” inequality when considered from the perspective of SNGs, in particular geographic regions. This is the case for Australia but is not the case for Denmark.

In a non-federal system, the fiscal differences between county and municipality governments will, typically, be relatively small. In entropic terms, this means that fiscal dispersion in total is relatively modest (see Table 6.9). This is the case for Denmark (total revenue inequality of 0.3847) in contrast to the case of Australia (total revenue inequality of 1.727). Moreover, no specific "sets" dominate fiscal inequality. From a governmental hierarchy perspective, between-level of government inequality and within-level of government inequality both contribute to total SNGs' fiscal inequality.

6.5 Concluding remarks

Australia has three distinct levels of government, namely federal, state and local governments. Denmark has only two formal "traditional" tiers[8] as they are in any unitary country, namely national and subnational levels. Using the entropic approach, for both analyses of geographic regions and governmental hierarchy, the results reveal that dispersion of revenue and expenditure shares to each tier of SNGs (state versus local levels in Australia; and county versus municipality levels in Denmark) are substantially different between Australia and Denmark.

The analysis of fiscal inequality in Australia and Denmark in this chapter reveals that fiscal decentralisation is positively correlated with: (i) the extent of fiscal inequality in general; (ii) "within-region" fiscal inequality (when subnational governments units are grouped by geographic region: the New South Wales region compared to the Western Australia region); and (iii) "between-levels of government" fiscal inequality (when subnational government units are grouped in accordance with levels of government: state versus local). Consequently, without an adjustment for the fiscal decentralisation index, there would be a positive bias in the fundamental index (FDI) and the enhanced index (eFDI) in favour of unitary forms of government, such as Denmark, and a negative bias against discrete multi-tiered federal forms of government, such as Australia. This is even evident in the estimates of the fiscal decentralisation index (FDI) developed in Chapter 3 using the "fiscal autonomy and fiscal importance" approach, which provides a fundamental index of fiscal decentralisation for Australia of 0.53, well below the FDI of 0.62 for unitary Denmark.

[8] Local governments vary from large county governments to smaller local authorities.

This Australia and Denmark comparison was primarily designed to demonstrate and test the variability of the fiscal decentralisation index and the notion of fiscal inequality in the entropic approaches as developed and discussed in Chapters 5 and 6. However, as this comparison was only done for the two countries for which adequate data are available to do so, the findings on the relative efficiency of the two sets of fiscal arrangement in Australia and Denmark should be considered provisional in character.

The differences in the level of fiscal inequality among SNGs in Australia and Denmark, using both analyses of geographic regions and governmental hierarchy are clear. As a result, findings from the framework of subnational fiscal inequality developed under the entropic approach will, in future, need to be incorporated into the fundamental and enhanced indices to achieve the most accurate measure of the degree of fiscal decentralisation that is not biased against multi-tiered federal countries.

Appendix A6.1

FISCAL INEQUALITY WITHOUT AND WITH

NO-LOCAL-GOVERNMENT REGIONS

The analyses of fiscal inequality in Australia and Denmark, Sections 6.2–6.4, have consistently treated regions without local governments which are excluded in the analysis of fiscal inequality of the distribution of revenue and expenditure shares across subnational units. In particular, Australian Capital Territory (ACT) is excluded from the analysis of fiscal inequality for Australia since the ACT has no local governments. Also, in the case of Denmark, the cities of Copenhagen and Frederiksberg are excluded since these two cities do not have any local governments (or municipalities in the case of Denmark). This appendix is devoted to the understanding of fiscal inequality in these two countries when their respective no-local-government regions are included in the analysis. It means that, ACT is now incorporated into the analysis of fiscal inequality for Australia and the same principle is used in Denmark with Copenhagen and Frederiksberg. Table A6.1 presents the results of fiscal inequality in Australia and Denmark with these changes.

TABLE A6.1
THE WORLD WITHOUT AND WITH NO-LOCAL-GOVERNMENT REGIONS
AUSTRALIA VERSUS DENMARK

Fiscal inequality	AUSTRALIA, 2004				DENMARK, 2002			
	Revenue		Expenditure		Revenue		Expenditure	
	Without	With	Without	With	Without	With	Without	With
Total	1.7273	1.7124	1.7629	1.7447	0.3847	0.4693	0.3274	0.4094
Between:								
Geographic region	0.0636	0.0703	0.0542	0.0647	0.0398	0.1694	0.0300	0.1486
Governmental hierarchy	1.5525	1.5081	1.6133	1.5683	0.1109	0.1839	0.1310	0.1986
Within:								
Geographic region	1.6638	1.6421	1.7087	1.6800	0.3449	0.2999	0.2974	0.2608
Governmental hierarchy	0.1749	0.2043	0.1495	0.1764	0.2739	0.2854	0.1964	0.2108

Source: Author's calculations.

The examinations of fiscal inequality in Australia and Denmark in this Appendix reveal many interesting and important results. When no-local-government regions are incorporated into the analysis of fiscal inequality, the changes in fiscal inequalities from the two countries are not consistent. Total fiscal inequality does not change substantially for Australia. However, fiscal inequality in Denmark increases substantially, compared with Australia, with this incorporation of these two no-local-government regions. While the Australian Capital Territory accounts for only 0.013 per cent and 0.017 per cent of total subnational revenue and expenditure, respectively, of all subnational units (states and local councils) in Australia, Copenhagen City and Frederiksberg Borough account for a very substantial share of revenue and expenditure in Denmark – 0.107 per cent and 0.104 per cent for Copenhagen City, and 0.029 per cent and 0.030 per cent for Frederiksberg. So, while the incorporation of the Australian Capital Territory does not change the degree of total fiscal inequality in Australia, this is not the same for Denmark when Copenhagen City and Frederiksberg Borough are included. Consequently, results require a cautious treatment in the analysis of fiscal inequality from country to country in the presence of no-local-government regions but they do not alter the general findings outlined in Section 6.4.