Human Resources for Health Observer

Issue nº 5

Measuring health workforce inequalities: methods and application to China and India

World Health Organization

WHO Library Cataloguing-in-Publication Data

Measuring health workforce inequalities: methods and application to China and India / Sudhir Anand.

(Human Resources for Health Observer, 5)

1.Health personnel - organization and administration. 2.Health manpower - trends. 3.Health planning - organization and administration. 4.China. 5.India. I.Anand, Sudhir. II.World Health Organization. III.Series.

ISBN 978 92 4 150022 7 (NLM classification: W 76)

Design: Atelier-Rasmussen / Geneva
Printed in Switzerland / September 2010

Contents

Acknowledgements	
Executive summary	
Preface	1
1. Introduction	2
2. The measurement and decomposition of health workforce inequalities	2
2.1 Notation	2
2.2 Three inequality indices	2
2.3 Decomposition of health workforce inequalities	4
2.4 Interpretation of decomposition for the Theil measures	5
3. Application to China and India	6
3.1 Health workforce inequalities in China, 2005	6
3.1.1 Rural China: decomposition by county income	14
3.2 Health workforce inequalities in India, 2001	16
4. Concluding remarks: a comparison of China and India	27
References	28

Tables

Table C-1	China: decomposition of inter-county inequality by province, 2005	7
Table C-2a	Health professionals, China: density and inter-county inequality by province, 2005	10
Table C-2b	Doctors plus nurses, China: density and inter-county inequality by province, 2005	11
Table C-2c	Doctors, China: density and inter-county inequality by province, 2005	12
Table C-2d	Nurses, China: density and inter-county inequality by province, 2005	13
Table C-3	China: decomposition of county inequality by rural–urban stratum, 2005	14
Table C-4	China: density and county inequality by rural–urban stratum, 2005	15
Table C-5	China: decomposition of county inequality by province and rural–urban stratum, 2005	16
Table C-6	China: decomposition of rural inter-county inequality by rural income decile, 2005	17
Table I-1	India: decomposition of inter-semidistrict inequality by state, 2001	18
Table I-2a	All health workers, India: density and inter-semidistrict inequality by state, 2001	20
Table I-2b	Doctors plus nurses, India: density and inter-semidistrict inequality by state, 2001	21
Table I-2c	Doctors, India: density and inter-semidistrict inequality by state, 2001	22
Table I-2d	Nurses, India: density and inter-semidistrict inequality by state, 2001	23
Table I-3	India: decomposition of semidistrict inequality by rural–urban stratum, 2001	24
Table I-4	India: density and semidistrict inequality by rural–urban stratum, 2001	25
Table I-5	India: decomposition of semidistrict inequality by state and rural–urban stratum, 2001	26

Annexes

Annex 1	Decomposition of the squared coefficient of variation C^2	29
Annex 2	Data and health worker categories for China, 2005	29
Annex 3	Data and health worker categories for India, 2001	30

Acknowledgements

This study was researched and written by Sudhir Anand (University of Oxford). Mario Dal Poz, Neeru Gupta and Angelica Sousa (World Health Organization) reviewed the document and provided constructive comments. Financial support for the publication was provided by the United States Agency for International Development. Thanks are due to Victoria Fan for research assistance and to Helen Wang for formatting assistance.

Executive summary

Despite increasing interest in equity in health access and the pathways by which inequities in health outcomes arise and are perpetuated or exacerbated, the global evidence base to inform policy decision-making on maldistribution in the supply, composition and deployment of the health workforce remains weak.

This study proposes methods for measuring inequalities in the distribution of health workers in a country by adapting techniques from the economics literature on income inequality to the measurement of health workforce distribution across geographical units. Calculations use three indices: the Theil L measure and the Theil T index (both of which are decomposable) and the Gini coefficient G which, though not decomposable, is the most well-known and extensively computed measure of inequality.

Decomposition involves a partition of country units (counties or semidistricts) into mutually exclusive and exhaustive groups (such as rural–urban strata, provinces or states) and a calculation of two separate components of overall inequality: a weighted sum of inter-unit inequality within each group, called the "within-group" component, and a "between-group" component that measures inequality due solely to variations in health-worker density across groups. Formulae to provide a consistent definition for the two components are proposed and explained, and interpretations are discussed and evaluated. It is pointed out that when two interpretations lead to the same answer, there is an unambiguous meaning to the between-group contribution to overall inter-unit inequality. For example, the contribution of between-stratum differences in health-worker density to overall inter-unit inequality can be measured *either* as the inequality that arises when these differences are the only source of variation, *or* as the amount by which overall inequality falls when these differences are eliminated but inter-unit inequality within each group is kept constant.

The measures of inequalities in human resources for health were applied and analysed in two countries with large and diverse health labour markets: China and India. The study uses health workforce data from official sources in the two countries obtained in disaggregated form (by country unit, rural–urban stratum and health-worker category). This allows measurement of inequalities by the three selected indices. The final section reports that the overall density of health workers in China in 2005 was much higher than that in India in 2001. The conclusions highlight some of the main differences and similarities that were found between the health workforce distributions in the two countries. The finding of greater inequality in India than in China is especially significant: the urban–rural disparity ratios in health-worker density are almost twice as high in India as in China. In India, 85–90% of overall inequality is explained by just the two variables of state and stratum. In China, however, only some 40–50% of inequality is explained by these two variables. Thus in China there are large variations within province-and-stratum, whereas there are not in India within state-and-stratum. Reducing state-and-stratum density differences in India will thus achieve much greater reductions in overall health-worker inequality than reducing province-and-stratum differences in China.

Different sources of data typically found in countries can be used to establish the order of magnitude and sources of health workforce inequalities. Countries are encouraged to compile and publish the standard indices of inequality, as a means of strengthening measurement techniques and learning from experience.

Preface

There is increasing awareness at the national, regional and international levels of the critical impact that the supply, distribution and quality of the health workforce has on the achievement of health and health system goals, including the Millennium Development Goals. Countries and partners are investing significantly in the development of human resources for health (HRH), including activities for education and training, deployment, management and retention of health workers. Nevertheless, as was recognized in the *World health report 2006 – Working together for health* (World Health Organization, 2006), in many countries the scarcity and fragmentation of data and information on the health workforce remain a major obstacle to identifying, implementing and monitoring the most effective policy and programme interventions.

In 2009, the World Health Organization, in collaboration with the World Bank and the United States Agency for International Development, published the *Handbook on monitoring and evaluation of human resources for health* (Dal Poz et al., 2009). The handbook offers health managers, researchers and policy-makers a comprehensive, standard reference for monitoring the health workforce, bringing together an analytical framework and country experiences with strategy options for improving the human resources information and evidence base to support decision-making.

Following the production of the handbook, it was recognized that further attention was needed to support countries and stakeholders in analysing and using health workforce data to address sub-national distribution imbalances. Thus, additional research was initiated to systematically review the current state of evidence on measuring and monitoring maldistribution of the health workforce, with special attention to low- and middle-income countries.

This publication is an outcome of that research. It is the second issue in the *Human Resources for Health Observer* series with the goal of promoting statistical discourse on measuring inequalities in national health labour markets and the implications for policy and planning. The present study seeks to identify the most appropriate methods to measure inequalities in the geographical distribution of the health workforce in a country by adapting techniques from the economics literature on income inequality. Three main indices are identified: the Theil L measure, the Theil T index and the Gini coefficient. Formulae to provide consistent definitions are proposed and explained, and the methods developed are applied and illustrated with reference to health workforce data in two countries with large and diverse health labour markets: China and India.

In China, although the production of skilled health workers has greatly expanded in recent years, important differences persist in their distribution across the country. In India, the formal health workforce includes professionals trained and qualified in western medicine as well as practitioners of traditional healing systems, including those trained and qualified in the ayurvedic, yoga, unani, siddha and homeopathic traditions. In both countries, monitoring and addressing the geographical distribution of health-care providers to enable the health system to deliver essential services in an equitable and efficient way is a key social and policy concern (Anand et al., 2008; World Health Organization, 2007).

In most countries there are many different data sources that can potentially be used to measure HRH stock and distribution, including administrative records on health facility staffing and professional licensure, population censuses and statistical surveys. The publication of this study is part of the World Health Organization's broader efforts to enhance country capacities to generate, analyse and use data and information to assess health workforce performance and track progress towards HRH-related goals. The ultimate objective is to improve and promote standards and methods of measuring inequalities of health personnel in countries with different health policy environments and data challenges, so that the results are more comparable within and between countries and over time. Countries are encouraged to apply the methods proposed in this study, as a means of strengthening measurement techniques and learning from experience to build on the knowledge gleaned from the examples from China and India presented here.

1

Introduction

The purpose of this study is to present methods for measuring inequalities in the distribution of the health workforce in a country, and to account for the sources of such inequalities. Techniques from the economics literature on the measurement and decomposition of income inequality are adapted to data on the distribution of health workers across geographical units (e.g. counties, provinces and rural–urban strata). The methods developed here are applied and illustrated with reference to health workforce data from China and India. The selection for this analysis was based on the accessibility of county- and district-level health workforce data from these two very large countries, which account for more than a third of the world's population.

The study begins with a description of different indices of inequality of the distribution of the health workforce and their decomposition between and within groups. The literature on economic inequality lists various desirable properties of inequality measures (Sen, 1997), but still leaves a multitude of indices that satisfy these properties. This study therefore reports three measures selected to describe, analyse and decompose health workforce inequalities: the Theil L measure, the Theil T index and the Gini coefficient. The reason for choosing the two Theil measures (Theil, 1967) is that they are additively decomposable (in a sense to be made precise later), which allows accounting for different sources of inequality. The reason for including the Gini coefficient, although it is not decomposable, is that it is the most well-known and extensively computed measure of inequality; it is thus possible to gauge the extent of inequality in the distribution of health workers with that of other distributions whose Gini coefficients have been computed (income, consumption, etc.). The Theil L measure has the most desirable decomposition property, but it is not defined when there are zero health workers in a geographical unit. The Theil T index has a slightly less desirable decomposition property, but it is well defined when there are zero health workers in a unit. Other members of the class of decomposable indices – the Generalized Entropy (GE) class – have significantly less attractive decomposition properties, so are not used here.

The measurement of health workforce inequalities is a new area of research. It is hoped that the methods applied and the results presented here for China and India will be of much intrinsic and comparative interest. The study concludes by comparing distributions of four categories of health workers in these two countries. In due course, the methods in this study should be applied to other countries and to the same countries over time. Such research would begin to reveal the extent, nature and causes of maldistribution of the health workforce in different parts of the world – and changes over time.

2

The measurement and decomposition of health workforce inequalities

The objective is to measure inequality in the distribution of the health workforce in a country. Geographical units within a country (e.g. counties) that have a high health-worker density are better able to serve the health-care needs of their people than units with a low health-worker density. Units (counties or semidistricts) are ranked in terms of their health-worker density, and inequality is measured in the per capita availability of health workers for people in different units.

2.1 Notation

For each of $i = 1, 2, \ldots, n$ geographical units (e.g. counties or semidistricts), let

h_i = number of health workers in geographical unit (county) i
p_i = number of people (i.e. population) in unit (county) i.

Define the health-worker density in unit (county) i as

$$x_i = h_i/p_i.$$

Label the units (counties) in non-descending order of health-worker density as follows:

$$x_1 \leq x_2 \leq x_3 \leq \ldots \leq x_n.$$

Note that this labelling does not imply a monotonic ordering of either h_i or p_i with respect to i. The distributional assumption is that h_1 health workers are available for p_1 persons, h_2 health workers are available for p_2 persons, and so on. This defines the national distribution of health workers across people in the country.

Let H be the total number of health workers in the country. Then $H = h_1 + h_2 + \ldots + h_n = \sum_i h_i$.

Let P be the total number of people (population) in the country.
Then $P = p_1 + p_2 + \ldots + p_n = \sum_i p_i$.

Let the national health-worker density be X.
Then $X = H/P$.

2.2 Three inequality indices

Three inequality indices are estimated to measure inequality in the distribution of health workers across the population: the two decomposable Theil measures L and T, and the Gini coefficient G. Other indices are not computed in this study because they do not lend themselves to appropriate decomposition.

Decomposition involves a partition of units (counties) into mutually exclusive and exhaustive groups (such as rural–urban strata or provinces in the country) and a calculation of two separate components of overall inequality: a weighted sum of inter-unit inequality within each group, called the "within-group" component, and a "between-group" component that measures inequality due solely to variations in health-worker density across groups. Certain inequality indices, such as Theil L and T, have the property that they can be disaggregated as the sum of these two terms – but in general the Gini coefficient cannot (Anand, 1983:319–326). In the class of additively decomposable measures, the weights on the group inequality indices of the within-group component sum to unity only for the Theil L and T measures. Hence, only for these two measures can the within-group component be interpreted as the *average level of inequality within groups*. Other members of the class do not share this property, so are not used in this study.

The Theil L measure

The inequality index with the most attractive decomposition property (see Section 2.4) is the Theil L measure. Theil L attempts to measure the divergence between the health-worker shares (h_i/H) and the population shares (p_i/P) of units (counties) using a specific distance function (shown in the formula below). In terms of the notation in Section 2.1, L is defined as follows:

$$L = \sum_i (p_i/P) \log [(p_i/P)/(h_i/H)] \qquad (1)$$
$$= \sum_i (p_i/P) [\log (p_i/P) - \log (h_i/H)].$$

Therefore, L can be written as:

$$L = \sum_i (p_i/P) \log (X/x_i), \quad \text{since } X = H/P \text{ and } x_i = h_i/p_i$$
$$= \sum_i (p_i/P) [\log (X) - \log (x_i)]$$

Hence, L is the mean of the deviations of the logarithm of health-worker density, i.e. $\log (x_i)$, from the logarithm of the national health-worker density, $\log (X)$. For this reason, Theil L is also known as the mean logarithmic deviation (or MLD) in the literature.

This formula for Theil L reduces to:

$L = \log (X) - \log (\chi)$ since $\sum_i p_i = P$ or $\sum_i (p_i/P) = 1$, where χ is defined through $\log (\chi) = \sum_i (p_i/P) \log (x_i)$. Therefore, $L = \log (X/\chi)$.

By taking anti-logs of the equation $\log (\chi) = \sum_i (p_i/P) \log (x_i)$, we have $\chi = \prod_i (x_i)^{(p_i/P)}$. Hence, χ is the *geometric* mean of the health-worker distribution.

The national health-worker density X is the arithmetic mean of the health-worker distribution, which is a population weighted average of the x_i.
$X = H/P = (\sum_i h_i)/P = \sum_i (p_i x_i)/P = \sum_i (p_i/P) x_i$.

Thus it has also been shown that L is the logarithm of the ratio of the arithmetic mean X to the geometric mean χ of the health-worker distribution. It immediately follows that if all x_i are equal, then X is equal to χ and $L = 0$. And if any x_i is zero, then $\chi = 0$, $\log (\chi) \rightarrow -\infty$ and $L \rightarrow \infty$. In other words, L is not defined for distributions with zero density (health workers) in any unit.

The Theil T index

Theil's entropy index T simply reverses the roles of health-worker share (h_i/H) and population share (p_i/P) in the definition of the Theil L measure – see equation (1). Thus the Theil T index is defined as follows:

$$T = \sum_i (h_i/H) \log [(h_i/H)/(p_i/P)] \qquad (2)$$

which can be written as:

$$T = \sum_i (p_i/P) (x_i/X) \log (x_i/X).$$

Note that unlike the Theil L measure, the Theil T index is indeed defined with zero h_i (or zero x_i), because all terms with a zero health-worker share (h_i/H) tend to zero in equation (2), since $x \log x \rightarrow 0$ as $x \rightarrow 0$.

The Gini coefficient G

The Gini coefficient is perhaps the best-known measure of inequality, but unfortunately it is not decomposable. In the context of the health-worker distribution, and using the above notation and labelling of units i from 1 to n in non-descending order of health-worker density, it can be defined as below (see Anand, 1983).

Let F_i be the cumulative proportion of the population up to unit i (i.e. the sum of p_k/P from k = 1 to i), and let Φ_i be the cumulative proportion of health workers up to unit i (i.e. the sum of h_k/H from k = 1 to i). Thus,

$$F_i = \sum_{k=1}^{i} (p_k/P)$$
and $$\Phi_i = \sum_{k=1}^{i} (h_k/H).$$
Also, define $F_0 = \Phi_0 = 0$.

Then the graph (F_i, Φ_i) for i = 0, 1, 2, …, n traces the Lorenz curve of the health-worker distribution in the country. The Gini coefficient G of this distribution is defined as:

$$G = 1 - \sum_{i=0}^{n-1} (F_{i+1} - F_i)(\Phi_{i+1} + \Phi_i).$$

This can easily be shown to be equal to the ratio between the Lorenz curve and the diagonal of perfect equality, to the area of the triangle below this diagonal. Equivalently, G can be expressed as (Anand, 1983:311–313):

$$G = \sum_{i=1}^{n-1} (F_i\Phi_{i+1} - F_{i+1}\Phi_i).$$

2.3 Decomposition of health workforce inequalities

Hitherto the empirical investigation of inequality in the health-worker distribution has typically been conducted in terms of simply noting the differences in density between groups – such as strata, provinces or states. Thus, for example, the rural–urban disparity ratio in health-worker densities merely reflects the average density differences *between* the two strata. Such ratios completely ignore inter-unit (inter-county) density differences *within* the strata, which could turn out to be quite large. For instance, it would be interesting to know how much of overall inter-county inequality in China is attributable to between-stratum inequality and how much to within-stratum inequality.

The obvious way to answer this question is to decompose overall inter-unit (inter-county) inequality into between-group (between-stratum) and within-group (within-stratum) inequality. The between-group contribution might then be defined as the ratio of between-group inequality to overall inequality (and similarly the within-group contribution). Many inequality indices, including the Gini coefficient, cannot in general be expressed as the sum of a between-group component and a within-group component, where the two components have a natural meaning.

A consistent definition for these two terms can be provided as follows (Anand, 1983). The between-group component can be defined as the value of the inequality index when all within-group density differences are artificially suppressed. In other words, it is the value of the inequality index for the hypothetical distribution which assigns to each unit (e.g. county) within a group (e.g. province) the mean health-worker density of that group (e.g. province). In this way, within-group inequality in health-worker densities is eliminated, and the resultant distribution shows only the inequality arising from between-group (e.g. between-province) differences in density.

Likewise, the within-group component can be defined as the value of the inequality index when all between-group density differences are artificially suppressed. Thus a hypothetical distribution is constructed in which the group health-worker densities are equalized to the overall mean density through an equiproportionate change in the density of every unit within a group. In this way, between-group inequality is eliminated, but the inequality within each group remains constant (for a mean-independent inequality index).[1] The within-group component is then the value of the inequality index for this hypothetical distribution.

An inequality index may be said to be *strictly* additively decomposable if for any grouping overall inequality can be written as the sum of between-group and within-group inequality as defined above. This property allows the unambiguous measurement of the contribution of a particular grouping (or variable) to total inequality.[2]

The only inequality measure that is decomposable according to this strict definition is the Theil L measure. The decomposition formula for the Theil L measure is shown below. The within-group component for this measure is a weighted average of the Theil L measures for each group, where the weights are the *population shares* of the groups in total population.

Keeping the definition of between-group component the same as before, a weaker definition can be provided for the within-group component which extends the class of decomposable measures. It requires only that the within-group component be constructed from the inequality index, population size and number of health workers in each group – as an additively separable function over groups so that the contribution of each to overall inequality can be identified. The only functional form possible, if the measure is also to satisfy the basic desirable properties of an inequality index, is a weighted sum of inequality indices for each group, where the weights depend only on the population *share* and health-worker *share* of the groups. The measures that satisfy this definition of (weak) decomposability are the Generalized Entropy (or GE) class of inequality indices. These inequality measures were originally identified by Theil (1967) and include:

the Theil L measure (Theil, 1967:125–127) – now known as GE(0).

the Theil entropy index T (Theil, 1967:91–95) – now called GE(1).[3]

The decomposition formulae for the two Theil measures are shown below when the population is divided into *two* groups (e.g. rural and urban).[4] The extension is obvious for a partition of the population into more than two groups. Given that the overall population size is P, the overall number of health workers H, and the overall health-worker density X (= H/P), let P_1 and P_2 denote the population sizes of groups 1 and 2 (e.g. rural and urban), H_1 and H_2 the number of

[1] With a mean-independent inequality index, in fact, it does not matter to what level all the group densities are equalized in constructing the hypothetical distribution, provided the level is non-zero.

[2] Notice the analogy with regression analysis, where the total variance is decomposed into the sum of the "explained" variance and the "unexplained" variance, and R^2 is the fraction of the total variance explained by the independent variable(s).

[3] The GE class is parameterized by a scalar α, with GE(α) for α = 0 being the Theil L measure; for α = 1 the Theil T index; and for α = 2 one-half the coefficient of variation squared $\frac{1}{2}C^2$ (Theil, 1967:125–127). The measure $\frac{1}{2}C^2$ is now known as GE(2).

[4] The decomposition formula for C^2 is shown in Annex 1.

health workers in the two groups, and X_1 and X_2 the group health-worker densities. Then $X_1 = H_1/P_1$ and $X_2 = H_2/P_2$. Finally, let the subscripts 1, 2 on an inequality measure denote the value of the measure for groups 1, 2, respectively. Thus,

$$P = P_1 + P_2$$
$$H = H_1 + H_2$$
or $$PX = P_1X_1 + P_2X_2.$$

From equation (1), the Theil L measure can be written as:

$$L = \textstyle\sum_i (p_i/P) \log (X) - \sum_i (p_i/P) \log (x_i).$$

Having defined the arithmetic mean density in groups 1 and 2 as X_1 and X_2, let the geometric mean density in the two groups be denoted as χ_1 and χ_2, respectively. Then

$$X = (P_1/P)\, X_1 + (P_2/P)\, X_2$$
and $$\log (\chi) = (P_1/P) \log (\chi_1) + (P_2/P) \log (\chi_2).$$
Hence,

$$\begin{aligned} L &= (P_1/P) \log (X_1/\chi_1) + (P_2/P) \log (X_2/\chi_2) + \\ &\qquad \log (X) - (P_1/P) \log (X_1) - (P_2/P) \log (X_2) \\ &= (P_1/P)\, L_1 + (P_2/P)\, L_2 + \\ &\qquad (P_1/P) \log (X/X_1) + (P_2/P) \log (X/X_2) \end{aligned}$$

because $L_1 = \log (X_1/\chi_1)$ and $L_2 = \log (X_2/\chi_2)$.
Hence, $L = L_W + L_B$
where $L_W = (P_1/P)\, L_1 + (P_2/P)\, L_2$
and $L_B = (P_1/P) \log (X/X_1) + (P_2/P) \log (X/X_2)$.

The within-group component L_W is a *weighted average* of L_1 and L_2 where the weights are the population shares (P_1/P) and (P_2/P) of the two groups. The between-group component L_B is the L measure of the distribution which has P_1 persons with health-worker density X_1 and P_2 persons with health-worker density X_2.

As indicated earlier in equation (2), Theil's entropy index T reverses the roles of health-worker share and population share in the Theil L measure. T can thus be written and decomposed as follows:

$$\begin{aligned} T &= \textstyle\sum_i (h_i/H) \log [(h_i/H)/(p_i/P)] \\ &= \textstyle\sum_i (h_i/H) \log (x_i/X) \\ &= \textstyle\sum_i (p_ix_i/PX) \log (x_i/X) \\ &= (H_1/H)\, T_1 + (H_2/H)\, T_2 + (H_1/H) \log (X_1/X) + (H_2/H) \log (X_2/X) \\ &= (H_1/H)\, T_1 + (H_2/H)\, T_2 + (P_1X_1/PX) \log (X_1/X) + \\ &\qquad (P_2X_2/PX) \log (X_2/X) \\ &= T_W + T_B \end{aligned}$$
where $T_W = (H_1/H)\, T_1 + (H_2/H)\, T_2$
and $T_B = (P_1X_1/PX) \log (X_1/X) + (P_2X_2/PX) \log (X_2/X)$.

The within-group component T_W is a *weighted average* of the Theil indices T_1 and T_2 of the two groups, where the weights are the health-worker shares (H_1/H) and (H_2/H) of the two groups – and hence the weights sum to unity. The between-group component T_B is the Theil T index of the distribution which has P_1 persons with health-worker density X_1 and P_2 persons with health-worker density X_2.

2.4 Interpretation of decomposition for the Theil measures

The Theil entropy index T is decomposable in a weaker sense than the Theil L measure. Although the T index can be written as the sum of a between-group component and a within-group component, the within-group component is *not* the value of the T index when all between-group differences in health-worker density are suppressed (unlike the case with the Theil L measure). The reason is that the weights of the within-group component for the T index are the health-worker shares of the groups, *not* their population shares.

The interpretation of decomposition for the two Theil measures is immediate if one examines the decomposition formulae for the weakly decomposable Theil T index and the strictly decomposable Theil L measure. As in Section 2.3, divide the population into two groups[5] with subscripts 1, 2 referring to the value of a variable or index for groups 1, 2, respectively. Then the decomposition formula derived earlier for the Theil T index is:

$$T = T_W + T_B$$
where $T_W = (P_1X_1/PX)\, T_1 + (P_2X_2/PX)\, T_2$
and $T_B = (P_1X_1/PX) \log (X_1/X) + (P_2X_2/PX) \log (X_2/X)$.

Here T_B is indeed the value of the T index when all within-group unit density differences are suppressed. This follows directly from putting $T_1 = T_2 = 0$ in the formula. However, T_W is *not* the value of the T index when all between-group differences are suppressed but inequality within each group remains constant. If the group health-worker densities X_1 and X_2 are equalized to the overall density X by equiproportionate changes in the density of health workers in each unit within a group, the T index reduces to $(P_1/P)\, T_1 + (P_2/P)\, T_2$. This follows directly from the definition of the index, or by putting $X_1 = X_2 = X$ in the above decomposition formula for T. Thus the elimination of between-group inequality results in a value for the T index which is not T_W [$= (P_1X_1/PX)\, T_1 + (P_2X_2/PX)\, T_2$]. Hence T is not decomposable in the strict sense, and therefore T_B does *not* measure the reduction in

[5] There is no loss of generality in partitioning the population into two groups only: the extension for more than two groups is obvious and follows by induction.

inequality if between-group density differences are eliminated.[6]

From Section 2.3 the decomposition formula for the Theil L measure is:

$$L = L_W + L_B$$
where $L_W = (P_1/P)\, L_1 + (P_2/P)\, L_2$
and $L_B = (P_1/P) \log (X/X_1) + (P_2/P) \log (X/X_2)$.

L_B is clearly the between-group component, since it is the value of the measure when all within-group differences are suppressed. This follows directly from the definition of the L measure, or by putting $L_1 = L_2 = 0$ in the above formula. Moreover, L_W is indeed the within-group component according to the strict definition. When all between-group differences are suppressed by the equalization of group densities, but inter-unit inequality *within* each group is kept constant, the L measure reduces to $(P_1/P)\, L_1 + (P_2/P)\, L_2 = L_W$. This follows directly from the definition, or by putting $X_1 = X_2 = X$ in the decomposition formula for L. Since L is the sum of the between-group component, L_B, and the within-group component according to the strict definition, L_W, the measure is additively decomposable in the strict sense. This property allows unambiguous interpretation of the between-group component L_B.

By one interpretation L_B measures the extent of inequality that arises if between-group differences in health-worker density are the only source of variation. In other words, it is the value of the L measure if inter-unit inequality *within* each group is eliminated, but the group mean densities are kept constant. By another interpretation L_B measures the reduction in overall inter-unit inequality if between-group differences in health-worker density are eliminated, but inter-unit inequality within each group is kept constant. These two interpretations of the between-group component give the same answer for an inequality index that is decomposable in the strict sense.

The two interpretations are *not* consistent for measures that are decomposable in the weak sense only. For a measure to be strictly decomposable, the weights of the within-group component must be population shares – and not health-worker shares – as verified for L above. Otherwise, the two interpretations of the between-group component lead to different answers. The elimination of inter-unit inequality within each group while holding group mean densities constant yields the between-group component. But the elimination of between-group differences in mean density while holding inter-unit inequality within each group constant does *not* yield the within-group component. Although the inequality indices for each group remain constant, the *weights* on these indices change if group mean densities are equalized. When the mean health-worker density for each group is the same, the share of health workers of a group collapses to its population share, which changes the weight on its inequality index (unless the health-worker and population shares are the same to begin with, in which case there is no between-group inequality).

When the two interpretations lead to the same answer, there is an unambiguous meaning to the between-group contribution to overall inter-unit inequality. For example, the contribution of between-stratum differences in health-worker density to overall inter-unit inequality can be measured *either* as the inequality that arises when these differences are the only source of variation, *or* as the amount by which overall inequality falls when these differences are eliminated but inter-unit inequality within each group is kept constant.

3

Application to China and India

The three measures of inequality were applied to health workforce data in two countries with the world's largest populations: China and India. Data for the smallest comparable geographical units were used for each country – counties in China and semidistricts in India. Four overlapping and approximately comparable categories of health workers were distinguished: (i) doctors; (ii) nurses (including midwives); (iii) doctors plus nurses; and (iv) health professionals (China) or all health workers (India). The statistical computations were conducted using the Stata software package (Jenkins, 1999; StataCorp, 2009).

3.1 Health workforce inequalities in China, 2005

County-level data on the numbers of health workers for 2005 were obtained from the Center for Health Statistics and Information, Ministry of Health, China. The data were based on an administrative data collection system from public and private health facilities (Anand et al., 2008). For this study, four categories of health worker were distinguished: doctors, nurses, doctors plus nurses, and health professionals (Annex 2). These categories are not disjoint – for example, doctors and nurses are a subset of health professionals. Health professionals include doctors, nurses, pharmacists, laboratory technicians, clinical radiologists and other technical staff with advanced education. "County-level" refers to counties,

[6] The elimination of between-group density differences leads to a reduction in overall inequality of
$T - (P_1/P)T_1 - (P_2/P)T_2 = T_B + [T_W - (P_1/P)T_1 - (P_2/P)T_2] \neq T_B$.

autonomous counties, county-level cities, *qi*, autonomous *qi*, and districts. Henceforth, "county" refers to these county-level units. The 2005 county-level dataset used here comprises 2854 counties (Anand et al., 2008).

This dataset permits examination of inter-county, inter-province and rural–urban inequalities in the distribution of health professionals, of doctors plus nurses, of doctors and of nurses. These inequalities may be associated with corresponding economic disparities – for example, between counties, between provinces and between strata (rural–urban). Apart from these geographical categories, however, individual-level information was not available to examine differences in availability of or access to health workers by income class, education, occupation, or other socioeconomic variables.

Separate calculations were made of inter-county inequality among health professionals, among doctors plus nurses, among doctors and among nurses – including their decompositions into within- and between-province inequality and into within- and between-stratum (rural–urban) inequality. As demonstrated earlier for the two decomposable indices, Theil L and Theil T, overall inter-county inequality is the sum of within-group inequality and between-group inequality, where within-group inequality is a weighted *average* of inequality in each group.

Table C-1 shows overall inter-county, within-province and between-province inequality in the distribution of health workers. Several observations about these estimates are in order. First, overall inter-county inequality in the distribution of all four categories of health worker is very high – Gini of 0.3718 for health professionals, 0.4030 for doctors plus nurses, 0.3622 for doctors and 0.4714 for nurses. Secondly, there is consistently higher inequality in absolute terms in the distribution of nurses than in that of doctors or of health professionals – overall, within-province and between-province, and for all three indices. Thirdly, within-province inequality accounts for 82% or more of overall inter-county inequality in the distribution of all categories of health worker (for both Theil L and Theil T). For nurses, within-province inequality is highest at 84–85% of overall inter-county inequality. Hence, more than four-fifths of the inter-county inequality in the distribution of health workers is explained by *within-province* inequalities.

Table C-1. China: decomposition of inter-county inequality by province, 2005[1]

Health-worker category	Inequality measure	Overall inter-county inequality	Within-province inequality	Between-province inequality	Within-province inequality (% of overall)	Between-province inequality (% of overall)
Health professionals	Theil *L*	0.2224	0.1831	0.0393	82.3%	17.7%
	Theil *T*	0.2532	0.2083	0.0449	82.3%	17.7%
	Gini	0.3718				
Doctors plus nurses	Theil *L*	0.2614	0.2174	0.0439	83.2%	16.8%
	Theil *T*	0.2962	0.2451	0.0511	82.8%	17.2%
	Gini	0.4030				
Doctors	Theil *L*	0.2096	0.1717	0.0379	81.9%	18.1%
	Theil *T*	0.2350	0.1920	0.0431	81.7%	18.3%
	Gini	0.3622				
Nurses	Theil *L*	0.3695	0.3121	0.0574	84.5%	15.5%
	Theil *T*	0.4076	0.3402	0.0673	83.5%	16.5%
	Gini	0.4714				

[1] N=2854 counties.

Note: County-level population data were obtained from household registration data for 2005 (Ministry of Public Security, China, 2006).

Given this situation, it is of much interest to examine inequality in the health-worker distribution within each province. Tables C-2a, C-2b, C-2c and C-2d summarize the health-worker distribution in each province for, respectively, health professionals, doctors plus nurses, doctors, and nurses. In these tables there are 31 "provinces", which comprise provinces, municipalities (i.e. Beijing, Shanghai, Tianjin and Chongqing) and autonomous regions (i.e. Xinjiang, Tibet, Inner Mongolia, Ningxia and Guangxi). Data for the two special administrative regions of China, namely Hong Kong and Macau, are not included here. The tables display for each province the number of counties (N); the number of health workers; the mean, minimum and maximum county density; and the three inequality measures for the distribution of the workforce.

Table C-2a shows that the density of health professionals in China is 3.06 per 1000 population, but this varies considerably between provinces. Beijing, Shanghai and Tianjin have the highest densities (9.36, 7.26 and 5.99, respectively), while Guizhou, Chongqing and Anhui have the lowest densities (1.84, 2.14 and 2.25, respectively). Yet these differences do not amount to great inequality across provinces: as seen in Table C-1, between-province inequality as measured by Theil L and T is quite small – approximately 17–18% of overall inequality.

An indication of within-province differences in the density of health professionals is provided in Table C-2a in the columns listing the minimum and maximum county density. In Beijing, the lowest county density is 4.31 health professionals per 1000 population while the highest is 25.57. The absolute minimum–maximum range is widest in Guangdong: 0.58 to 65.25 health professionals per 1000 population. The ratio between the highest and lowest county density in each province is, in general, substantially larger than the ratio between the highest mean provincial density (Beijing) and the lowest (Guizhou). It is not surprising, therefore, that within-province inequality will be substantially larger than between-province inequality (Table C-1).

The absolute difference between the minimum and maximum density, divided by mean density, is itself a measure of inequality (known as the range). Since it ignores the distribution inside the extremes, the range violates an important property of an inequality index, i.e. the Pigou–Dalton condition (Sen, 1997). Nevertheless, it seems to be reasonably well correlated with the rankings of provinces on the basis of our three standard indices of inequality. The provinces with the highest inter-county inequality are Qinghai, Guangdong, Guizhou and Tibet (Ginis of 0.4343, 0.4285, 0.4148 and 0.4138, respectively) and those with the lowest are Shanghai, Hainan, Chongqing and Beijing (Ginis of 0.2379, 0.2464, 0.2569 and 0.2596, respectively). With minor exceptions, the two Theil measures yield essentially similar rankings of provinces.

Table C-2b shows the same statistics as Table C-2a, but for the smaller category of doctors plus nurses. Given that the latter are a subset of health professionals, the densities per 1000 population are lower for doctors plus nurses than for health professionals. For China as a whole there are 2.26 doctors plus nurses per 1000 population, ranging from 1.45 in Guizhou (1.57 in Chongqing, 1.61 in Anhui) to 7.25 in Beijing (5.82 in Shanghai, 4.43 in Tianjin). Again, between-province inequality for doctors plus nurses is low: it accounts for only about 17% of overall inter-county inequality (Table C-1).

Both overall inequality and within-province inequality are higher for doctors plus nurses than for health professionals (Table C-1). The overall Gini for doctors plus nurses is 0.4030, while that for health professionals is 0.3718. Provincial Ginis are also generally higher for doctors plus nurses – 0.4809 in Tibet, 0.4638 in Guangdong, 0.4498 in Guizhou and 0.4485 in Qinghai; with the exception of Shanghai and Beijing, they are also higher in the low-inequality provinces (0.2593 in Hainan, 0.2842 in Chongqing). The rankings of provinces in Table C-2b are slightly different from those in Table C-2a, but the rank correlation coefficient is high. The larger inequality values for doctors plus nurses than for health professionals suggests that categories of health worker other than doctors and nurses are distributed more equally and seem to be compensating for inequality in the former.

Table C-2c and Table C-2d move on to more disaggregated categories of health worker: Table C-2c displays the doctor distribution by province and Table C-2d the nurse distribution. The overall density of doctors in China is 1.30 per 1000 population and the overall density of nurses is 0.96 per 1000 population. The distribution of doctors *overall* and within each province is far more equal than that of nurses. For China as a whole, the Gini for doctors is 0.3622 while that for nurses is 0.4714. Province by province, Table C-2c and Table C-2d show that nurse inequality is, in general, substantially larger than doctor inequality. Particularly big differences occur in provinces that have very high nurse inequality. Thus, Tibet has a Gini for nurses of 0.6102 and for doctors of 0.4397; Guizhou has a nurses Gini of 0.5392 and a doctors Gini of 0.4067; Qinghai has a nurses Gini of 0.5233 and a doctors Gini of 0.3978. Such large differences in many provinces, coupled with uniformly higher inequality for nurses than for doctors in every province, brings about the 11-point Gini difference between nurses and doctors nationally.

It can be verified from Table C-2b that the distribution of doctors plus nurses – *overall* and in each province – shows intermediate inequality between the distribution of doctors (Table C-2c) and that of nurses (Table C-2d). To the extent that doctors (especially those with shorter training) and nurses perform some of the same health-care functions, it is appropriate to consider them as a combined category – as doctors plus nurses – as in Table C-2b. Of course, the combined category of doctors plus nurses would be a strictly valid one only if doctors and nurses were perfect substitutes in medical care, which is not the case. Hence, it is of value to consider the distribution of doctors separately from that of nurses – as in Tables C-2c and C-2d.

Table C-3 shows overall inter-county, within-stratum and between-stratum inequality in the distribution of health workers, and the percentage contributions of within-stratum and between-stratum inequality to overall inequality. For health professionals, the percentage contribution of between-stratum to overall inequality is 34.0% according to Theil L and 28.9% according to Theil T. Doctors show the lowest between-stratum contribution to overall inequality at 31.4% (Theil L) and 27.2% (Theil T), while nurses show the highest at 38.5% (Theil L) and 32.5% (Theil T). Other things being equal, this is likely to be due to a higher rural–urban disparity in the density of nurses than in that of other categories of health worker.

Table C-4 summarizes the health-worker distribution in each stratum – rural and urban – and displays the number of counties, the number of health workers, the mean, minimum and maximum county density per 1000 population, and the three inequality measures. It is clear that inter-county inequality for each category of health worker is much higher in urban areas than in rural areas, according to all three measures. The Gini coefficient in urban areas is almost as high as it is nationally (e.g. 0.3681 for health professionals in urban areas compared with 0.3718 overall, 0.3903 for doctors plus nurses in urban areas compared with 0.4030 overall). According to both Theil L and Theil T, inequality for each category of health worker is between two and three times higher in urban areas than in rural areas. Rural areas, though more homogenous in this respect than urban areas, have an average density of health workers which is half or a third that in urban areas. As seen in Table C-4, the density of health professionals in rural areas is 1.97 per 1000 population compared with 4.31 per 1000 population in urban areas (and 3.06 nationally). For nurses, the density in rural areas is 0.50 compared with 1.49 in urban areas – almost three times as high. This is a much higher factor than for the other categories of health worker.

Comparison of Table C-1 and Table C-3 shows that the absolute value of between-stratum inequality is approximately twice as large as that of between-province inequality for most categories of health worker, and more than twice as large for nurses. For health professionals, the percentage contribution of between-stratum to overall inequality is 34.0% (Theil L) and 28.9% (Theil T), whereas the contribution of between-province to overall inequality is 17.7% (Theil L) and 17.7% (Theil T). The corresponding numbers for doctors plus nurses are 35.9% and 30.3% (between-stratum contributions according to Theil L and Theil T) compared with 16.8% and 17.2% (between-province contributions according to Theil L and Theil T). For nurses the corresponding numbers are 38.5% and 32.5%, compared with 15.5% and 16.5%. It seems clear that rural–urban differences in health-worker density are much more significant than between-province differences in health-worker densities.

A *two-way* decomposition by province *and* stratum was also computed (Table C-5). Of course, between-province-*and*-stratum inequality will be larger than or equal to both between-stratum inequality and between-province inequality (since the former leads to a sub-partition of the population according to either stratum or province alone). However, between-province-*and*-stratum inequality turns out to be less than the *sum* of between-stratum inequality and between-province inequality, the reason being that the variables – province and stratum – are partitioning the population differently. If the two variables led to exactly the same partition of the population (e.g. if they were perfectly correlated), *both* one-way contributions *and* the two-way contribution would be identical, and the sum of the two one-way contributions would in this case be twice the two-way contribution. Thus an interpretation for the extent of divergence of between-province-*and*-stratum inequality from the sum of between-province inequality and between-stratum inequality might then be the extent of collinearity.

As shown in Table C-5, between-province-*and*-stratum inequality accounts for 46.4% (Theil L) and 41.1% (Theil T) of overall inter-county inequality for health professionals. This compares with between-province inequality accounting for 17.7% (Theil L and Theil T), and between-stratum inequality accounting for 34.0% (Theil L) and 28.9% (Theil T). Thus the sum of between-province inequality and between-stratum inequality is 51.7% (Theil L) and 46.6% (Theil T). The two-way between-province-*and*-stratum inequality is 5.3 percentage points less for Theil L and 5.5 points less for Theil T than the sum of the one-way contributions, which would appear to indicate limited dependence between province and stratum in their contributions to overall inequality.

The study found that health workers in China are highly maldistributed according to various partitions of the population: rural–urban, provincial and regional (regional data are not presented here). Also, within-province inequality accounts for more than 80% of overall inter-county inequality, suggesting that inequality needs to be remedied at the provincial level – and at the level of stratum within provinces. Rural–urban disparities in health-worker density account for much more of inter-county inequality than between-province differences. Nurses are more unequally distributed than other health workers, including within-provinces and within-strata

Table C-2a. Health professionals, China: density and inter-county inequality by province, 2005

			Density per 1000 population			Inequality measure		
Province	N	No. of health professionals	Mean	Min county	Max county	Theil *L*	Theil *T*	Gini
Beijing City	18	110 864	9.36	4.31	25.57	0.1086	0.1198	0.2596
Tianjin City	18	56 520	5.99	2.51	17.96	0.1562	0.1623	0.3119
Hebei	172	189 620	2.76	0.83	14.54	0.2118	0.2534	0.3597
Shanxi	119	131 185	3.98	1.02	13.57	0.1745	0.1904	0.3324
Inner Mongolia	101	89 350	3.80	1.05	13.70	0.1864	0.1921	0.3423
Liaoning	100	184 433	4.40	1.47	15.86	0.1730	0.1829	0.3307
Jilin	60	112 681	4.22	1.60	10.58	0.1111	0.1151	0.2629
Heilongjiang	129	135 587	3.60	0.56	12.66	0.1870	0.1935	0.3414
Shanghai City	19	98 747	7.26	4.17	22.23	0.1006	0.1171	0.2379
Jiangsu	106	227 313	3.13	0.96	22.52	0.1731	0.2057	0.3255
Zhejiang	90	178 718	3.88	1.33	28.40	0.1389	0.1745	0.2837
Anhui	105	146 548	2.25	0.93	13.20	0.2050	0.2574	0.3479
Fujian	84	85 405	2.52	0.55	12.90	0.2335	0.2885	0.3745
Jiangxi	99	102 613	2.34	0.81	18.21	0.1878	0.2525	0.3284
Shandong	140	289 704	3.14	1.23	13.88	0.1334	0.1537	0.2882
He'nan	159	266 040	2.66	1.02	13.44	0.1665	0.2052	0.3171
Hubei	101	168 652	2.83	0.41	11.30	0.1356	0.1515	0.2683
Hu'nan	122	188 477	2.82	0.92	14.16	0.1203	0.1514	0.2606
Guangdong	119	239 079	3.20	0.58	65.25	0.3030	0.3740	0.4285
Guangxi	109	113 022	2.31	0.76	14.17	0.2134	0.2778	0.3531
Hainan	21	26 242	3.20	1.65	7.27	0.0998	0.1087	0.2464
Chongqing City	40	67 724	2.14	1.06	11.22	0.1159	0.1500	0.2569
Sichuan	180	197 256	2.29	0.53	13.27	0.1743	0.2001	0.3287
Guizhou	88	71 113	1.84	0.60	12.06	0.2826	0.3558	0.4148
Yun'nan	129	101 562	2.38	0.49	16.26	0.2653	0.3072	0.4008
Tibet	73	7 956	2.97	0.83	13.75	0.2636	0.2854	0.4138
Shannxi	106	116 489	3.16	0.94	13.38	0.1601	0.1829	0.3157
Gansu	87	66 941	2.57	0.89	10.10	0.2056	0.2190	0.3609
Qinghai	43	17 167	3.41	0.81	12.55	0.3212	0.3358	0.4343
Ningxia	21	20 494	3.48	1.24	11.46	0.2051	0.2219	0.3563
Xinjiang	96	85 320	4.35	1.14	13.37	0.1989	0.1861	0.3425
Total China	**2 854**	**3 892 822**	**3.06**	**0.41**	**65.25**	**0.2224**	**0.2532**	**0.3718**

Table C-2b. Doctors plus nurses, China: density and inter-county inequality by province, 2005

			Density per 1000 population			Inequality measure		
Province	N	No. of doctors plus nurses	Mean	Min county	Max county	Theil *L*	Theil *T*	Gini
Beijing City	18	85 903	7.25	3.38	19.46	0.1066	0.1170	0.2578
Tianjin City	18	41 759	4.43	1.98	13.16	0.1568	0.1625	0.3104
Hebei	172	132 638	1.93	0.54	11.05	0.2643	0.3167	0.4012
Shanxi	119	96 893	2.94	0.69	10.61	0.2053	0.2222	0.3593
Inner Mongolia	101	67 267	2.86	0.70	11.10	0.2264	0.2299	0.3741
Liaoning	100	144 896	3.46	1.08	12.77	0.1963	0.2045	0.3506
Jilin	60	86 244	3.23	1.05	8.63	0.1384	0.1431	0.2928
Heilongjiang	129	98 984	2.63	0.48	10.18	0.2374	0.2448	0.3830
Shanghai City	19	79 184	5.82	3.42	17.62	0.0945	0.1101	0.2297
Jiangsu	106	169 962	2.34	0.64	18.42	0.2019	0.2393	0.3519
Zhejiang	90	133 968	2.91	0.99	22.71	0.1662	0.2062	0.3121
Anhui	105	104 918	1.61	0.44	10.38	0.2754	0.3372	0.4065
Fujian	84	67 651	2.00	0.39	10.73	0.2366	0.2932	0.3767
Jiangxi	99	76 698	1.75	0.50	14.22	0.2127	0.2848	0.3517
Shandong	140	214 950	2.33	0.81	11.00	0.1605	0.1845	0.3171
He'nan	159	174 883	1.75	0.66	10.20	0.2428	0.3059	0.3836
Hubei	101	124 133	2.08	0.29	9.26	0.1497	0.1740	0.2837
Hu'nan	122	134 070	2.01	0.65	11.14	0.1590	0.2015	0.3034
Guangdong	119	174 798	2.34	0.29	50.26	0.3577	0.4331	0.4638
Guangxi	109	86 903	1.78	0.56	10.96	0.2209	0.2853	0.3597
Hainan	21	19 387	2.37	1.21	5.72	0.1105	0.1217	0.2593
Chongqing City	40	49 716	1.57	0.70	8.83	0.1380	0.1773	0.2842
Sichuan	180	145 159	1.69	0.40	10.31	0.1984	0.2286	0.3507
Guizhou	88	55 932	1.45	0.41	11.09	0.3331	0.4238	0.4498
Yun'nan	129	80 470	1.88	0.44	13.13	0.2703	0.3168	0.4059
Tibet	73	5 731	2.14	0.38	11.10	0.3750	0.3796	0.4809
Shannxi	106	84 485	2.29	0.69	10.53	0.1857	0.2118	0.3414
Gansu	87	49 793	1.92	0.55	7.45	0.2245	0.2404	0.3763
Qinghai	43	13 110	2.60	0.54	9.23	0.3501	0.3552	0.4485
Ningxia	21	16 094	2.73	0.94	9.54	0.2139	0.2391	0.3646
Xinjiang	96	63 043	3.21	0.87	9.89	0.2303	0.2128	0.3663
Total China	**2 854**	**2 879 620**	**2.26**	**0.29**	**50.26**	**0.2614**	**0.2962**	**0.4030**

Table C-2c. Doctors, China: density and inter-county inequality by province, 2005

			Density per 1000 population			Inequality measure		
Province	N	No. of doctors	Mean	Min county	Max county	Theil *L*	Theil *T*	Gini
Beijing City	18	44 747	3.78	2.11	9.90	0.0853	0.0951	0.2282
Tianjin City	18	22 517	2.39	1.34	6.36	0.0924	0.1008	0.2389
Hebei	172	83 370	1.21	0.32	6.12	0.2004	0.2346	0.3506
Shanxi	119	58 737	1.78	0.52	5.26	0.1503	0.1585	0.3074
Inner Mongolia	101	43 682	1.86	0.56	5.68	0.1606	0.1640	0.3188
Liaoning	100	79 080	1.89	0.75	6.77	0.1430	0.1527	0.3017
Jilin	60	50 257	1.88	0.70	4.75	0.1120	0.1148	0.2643
Heilongjiang	129	57 365	1.52	0.35	5.19	0.1701	0.1772	0.3268
Shanghai City	19	40 525	2.98	1.95	8.22	0.0756	0.0878	0.2022
Jiangsu	106	94 269	1.30	0.40	8.93	0.1633	0.1862	0.3158
Zhejiang	90	76 994	1.67	0.65	10.51	0.1252	0.1505	0.2714
Anhui	105	60 299	0.93	0.29	5.50	0.2111	0.2564	0.3569
Fujian	84	37 082	1.10	0.28	5.30	0.2066	0.2581	0.3514
Jiangxi	99	42 905	0.98	0.33	6.94	0.1725	0.2276	0.3147
Shandong	140	123 920	1.35	0.49	5.48	0.1244	0.1406	0.2787
He'nan	159	101 491	1.01	0.41	5.19	0.1935	0.2390	0.3431
Hubei	101	70 281	1.18	0.19	4.33	0.1136	0.1224	0.2480
Hu'nan	122	79 647	1.19	0.45	5.52	0.1154	0.1398	0.2600
Guangdong	119	93 021	1.24	0.16	25.34	0.3112	0.3885	0.4337
Guangxi	109	45 853	0.94	0.31	5.08	0.1817	0.2282	0.3267
Hainan	21	9 793	1.20	0.62	2.89	0.1028	0.1109	0.2511
Chongqing City	40	30 898	0.98	0.48	4.67	0.0944	0.1208	0.2299
Sichuan	180	90 483	1.05	0.31	5.57	0.1536	0.1747	0.3081
Guizhou	88	35 525	0.92	0.32	8.00	0.2748	0.3637	0.4067
Yun'nan	129	46 500	1.09	0.29	7.42	0.2249	0.2682	0.3714
Tibet	73	4 046	1.51	0.22	9.49	0.3184	0.3139	0.4397
Shannxi	106	50 163	1.36	0.50	5.58	0.1381	0.1562	0.2949
Gansu	87	29 707	1.14	0.38	4.55	0.1722	0.1864	0.3315
Qinghai	43	7 348	1.46	0.37	4.33	0.2743	0.2749	0.3978
Ningxia	21	9 351	1.59	0.65	4.92	0.1464	0.1666	0.3015
Xinjiang	96	35 514	1.81	0.49	5.49	0.2289	0.2100	0.3642
Total China	**2 854**	**1 655 369**	**1.30**	**0.16**	**25.34**	**0.2096**	**0.2350**	**0.3622**

Table C-2d. Nurses, China: density and inter-county inequality by province, 2005

			Density per 1000 population			Inequality measure		
Province	N	No. of Nurses	Mean	Min county	Max county	Theil *L*	Theil *T*	Gini
Beijing City	18	41 156	3.48	1.27	9.56	0.1399	0.1488	0.2950
Tianjin City	18	19 242	2.04	0.55	6.80	0.2836	0.2660	0.4028
Hebei	172	49 267	0.72	0.12	4.93	0.4360	0.5086	0.5089
Shanxi	119	38 156	1.16	0.17	5.35	0.3376	0.3607	0.4549
Inner Mongolia	101	23 585	1.00	0.14	5.46	0.4308	0.4079	0.4921
Liaoning	100	65 816	1.57	0.33	6.76	0.2954	0.2915	0.4200
Jilin	60	35 987	1.35	0.35	4.13	0.1913	0.1978	0.3427
Heilongjiang	129	41 619	1.10	0.12	5.08	0.3828	0.3751	0.4711
Shanghai City	19	38 659	2.84	1.47	9.40	0.1185	0.1377	0.2608
Jiangsu	106	75 693	1.04	0.21	9.49	0.2653	0.3220	0.4027
Zhejiang	90	56 974	1.24	0.27	12.20	0.2386	0.2991	0.3745
Anhui	105	44 619	0.68	0.15	4.88	0.3905	0.4702	0.4811
Fujian	84	30 568	0.90	0.11	5.43	0.2844	0.3439	0.4133
Jiangxi	99	33 793	0.77	0.16	7.28	0.2833	0.3728	0.4070
Shandong	140	91 030	0.99	0.28	5.73	0.2278	0.2609	0.3775
He'nan	159	73 392	0.73	0.22	5.20	0.3366	0.4215	0.4514
Hubei	101	53 852	0.90	0.10	4.93	0.2148	0.2633	0.3446
Hu'nan	122	54 423	0.82	0.18	5.62	0.2481	0.3195	0.3815
Guangdong	119	81 778	1.09	0.13	24.92	0.4349	0.4972	0.5051
Guangxi	109	41 050	0.84	0.24	6.47	0.2752	0.3614	0.4025
Hainan	21	9 594	1.17	0.59	3.14	0.1237	0.1394	0.2735
Chongqing City	40	18 818	0.59	0.21	4.15	0.2467	0.3047	0.3876
Sichuan	180	54 676	0.64	0.09	4.74	0.3097	0.3530	0.4357
Guizhou	88	20 407	0.53	0.09	3.85	0.4934	0.5729	0.5392
Yun'nan	129	33 970	0.80	0.12	5.71	0.3567	0.3995	0.4605
Tibet	73	1 685	0.63	0.04	3.29	0.6773	0.6428	0.6102
Shannxi	106	34 322	0.93	0.19	4.95	0.2896	0.3206	0.4226
Gansu	87	20 086	0.77	0.16	3.12	0.3438	0.3507	0.4549
Qinghai	43	5 762	1.14	0.06	4.99	0.5058	0.4978	0.5233
Ningxia	21	6 743	1.15	0.26	4.62	0.3617	0.3732	0.4605
Xinjiang	96	27 529	1.40	0.27	4.63	0.2445	0.2262	0.3765
Total China	**2 854**	**1 224 251**	**0.96**	**0.04**	**24.92**	**0.3695**	**0.4076**	**0.4714**

Table C-3. China: decomposition of county inequality by rural–urban stratum, 2005[1]

Health-worker category	Inequality measure	Overall county inequality	Within-stratum inequality	Between-stratum inequality	Within-stratum inequality (% of overall)	Between-stratum inequality (% of overall)
Health professionals	Theil *L*	0.2224	0.1468	0.0757	66.0%	34.0%
	Theil *T*	0.2532	0.1800	0.0733	71.1%	28.9%
	Gini	0.3718				
Doctors plus nurses	Theil *L*	0.2614	0.1676	0.0938	64.1%	35.9%
	Theil *T*	0.2962	0.2063	0.0899	69.7%	30.3%
	Gini	0.4030				
Doctors	Theil *L*	0.2096	0.1439	0.0657	68.6%	31.4%
	Theil *T*	0.2350	0.1711	0.0640	72.8%	27.2%
	Gini	0.3622				
Nurses	Theil *L*	0.3695	0.2274	0.1421	61.5%	38.5%
	Theil *T*	0.4076	0.2749	0.1326	67.5%	32.5%
	Gini	0.4714				

[1] N=2854 counties.

(in both rural and urban areas). These inequalities in the distribution of health workers will obviously contribute to inequalities in health outcomes across counties, strata and provinces. Anand et al. (2008:Table 7) undertook multivariate regression across counties, which confirmed that health professionals were highly significant in explaining infant mortality (see also Anand and Bärnighausen, 2004, 2007).

3.1.1 Rural China: decomposition by county income

County-level income information is available for rural areas in China in 2005 and can be used to investigate the extent to which income differences account for inequalities in the distribution of health workers. It was possible to obtain income data for 1560 of the 1650 rural counties (95%) but only for a much smaller proportion of the 1204 urban counties. Accordingly, the income-related analysis presented here is restricted to rural areas. Although it is intended to be illustrative of the type of analysis possible with such data, the results are likely to be representative of rural China because 95% of rural counties were included.

Rural counties were ranked in ascending order of per capita income and then grouped into population deciles. Thus each decile contains approximately 10% of the rural population, but a variable number of counties depending on the population sizes of counties that belong to the decile. With the rural counties thus grouped by income decile, within- and between-decile inequality can be calculated.

Table C-6 shows the statistics for the decomposition of rural county inequality by rural income decile. The Gini coefficients and Theil indices for the four health-worker categories are all slightly lower for these 1560 rural counties than for the full sample of 1650 rural counties (cf. Table C-4). The Ginis are approximately 1 to 2 Gini points lower, and similarly the Theil indices.

The results in Table C-6 are both surprising and significant. Health-worker density differences between income deciles explain only about 20% of inter-county rural inequality – for all four health-worker categories and both Theil measures. The significance of these findings is that health-worker availability in counties is determined only to a small extent by county income levels, which include both public and private incomes. Provincial and local governments are clearly able to allocate, and are allocating, health workers not just on the basis of county income. Private-sector demand for health workers, which is likely to be correlated with personal incomes, and public-sector incomes still only account for about 20% of inter-county rural inequality. Clearly, the scope

Table C-4. China: density and county inequality by rural–urban stratum, 2005

	Stratum	Rural	Urban	National
	N (counties)	1 650	1 204	2 854
Health-worker category				
Health professionals	Number	1 340 408	2 552 414	3 892 822
	Mean density	1.97	4.31	3.06
	Min county density	0.41	0.53	0.41
	Max county density	13.75	65.25	65.25
	Theil *L*	0.0814	0.2219	0.2224
	Theil *T*	0.0884	0.2281	0.2532
	Gini	0.2194	0.3681	0.3718
Doctors plus nurses	Number	933 602	1 946 018	2 879 620
	Mean density	1.37	3.28	2.26
	Min county density	0.29	0.29	0.29
	Max county density	11.10	50.26	50.26
	Theil *L*	0.0938	0.2524	0.2614
	Theil *T*	0.1049	0.2550	0.2962
	Gini	0.2371	0.3903	0.4030
Doctors	Number	590 312	1 065 058	1 655 369
	Mean density	0.87	1.80	1.30
	Min county density	0.19	0.16	0.16
	Max county density	9.49	25.34	25.34
	Theil *L*	0.0862	0.2101	0.2096
	Theil *T*	0.0944	0.2136	0.2350
	Gini	0.2292	0.3578	0.3622
Nurses	Number	343 291	880 960	1 224 251
	Mean density	0.50	1.49	0.96
	Min county density	0.04	0.09	0.04
	Max county density	5.29	24.9	24.9
	Theil *L*	0.1413	0.3263	0.3695
	Theil *T*	0.1556	0.3215	0.4076
	Gini	0.2878	0.4369	0.4714

Table C-5. China: decomposition of county inequality by province and rural–urban stratum, 2005[1]

Health-worker category	Inequality measure	Overall county inequality	Within-province - *and*-stratum inequality	Between-province - *and*-stratum inequality	Within-province - *and*-stratum inequality (% of overall)	Between-province - *and*-stratum inequality (% of overall)
Health professionals	Theil *L*	0.2224	0.1193	0.1031	53.6%	46.4%
	Theil *T*	0.2532	0.1492	0.1041	58.9%	41.1%
	Gini	0.3718				
Doctors plus nurses	Theil *L*	0.2614	0.1379	0.1235	52.8%	47.2%
	Theil *T*	0.2962	0.1724	0.1239	58.2%	41.8%
	Gini	0.4030				
Doctors	Theil *L*	0.2096	0.1147	0.0949	54.7%	45.3%
	Theil *T*	0.2350	0.1389	0.0962	59.1%	40.9%
	Gini	0.3622				
Nurses	Theil *L*	0.3695	0.1906	0.1789	51.6%	48.4%
	Theil *T*	0.4076	0.2339	0.1737	57.4%	42.6%
	Gini	0.4714				

[1] N=2854 counties.

for allocating health workers on the basis of criteria other than county income, both private and public, exists and seems to be utilized in rural China.

3.2 Health workforce inequalities in India, 2001

District-level workforce data were obtained from the 2001 national population census through the Office of the Registrar General of India. The data are coded by the occupation of workers using the National Classification of Occupations (NCO), a system for classifying and organizing labour data in India according to the type of work performed (Ministry of Labour, 2004). The NCO is based largely on the International Standard Classification of Occupations (International Labour Organization, 2004). For the purposes of this study, four health workforce categories were constructed from the NCO codes: doctors, nurses (including midwives), doctors plus nurses, and all health workers (see Annex 3 for definitions). These categories are intended to match similar categories used in the analysis for China, although the aggregate category of "all health workers" in India is somewhat broader than that of "health professionals" in China (as can be seen from the definitions).

The 2001 Indian census data are available at the level of district within each state of the country. Each district can be subdivided into a rural and an urban part, and all information (population and health workers) is separately available for these two strata. The unit of analysis in this study is the *semidistrict* – i.e. the rural or urban part of a district – and the inequality and other calculations are done across 1167 semidistricts. Of these 1167 semidistricts, there were 584 rural semidistricts and 583 urban semidistricts (see Annex 3 for details). The unit of semidistrict in India is comparable to that of county in China, which can also be either rural or urban (but not both). In China in 2005 there were 2854 counties – 1650 rural and 1204 urban. The next section offers some conclusions on inequality comparisons between China and India, based on inter-county and inter-semidistrict estimates of inequality, respectively.

With this census dataset, semidistrict, state and rural–urban inequality are examined for four health-worker distributions: all health workers, doctors plus nurses, doctors, and nurses. Table I-1 shows overall semidistrict, within-state and between-state inequality for the four categories of health workers. These estimates warrant several comments. First, overall inter-semidistrict inequality is very high for every

Table C-6. China: decomposition of rural inter-county inequality by rural income decile, 2005[1]

Health-worker category	Inequality measure	Overall rural inter-county inequality	Within-decile rural inequality	Between-decile rural inequality	Within-decile inequality (% of overall)	Between-decile inequality (% of overall)
Health professionals	Theil *L*	0.0678	0.0519	0.0159	76.6%	23.4%
	Theil *T*	0.0686	0.0530	0.0156	77.2%	22.8%
	Gini	0.2027				
Doctors plus nurses	Theil *L*	0.0770	0.0600	0.0169	78.0%	22.0%
	Theil *T*	0.0794	0.0624	0.0170	78.6%	21.4%
	Gini	0.2176				
Doctors	Theil *L*	0.0742	0.0606	0.0137	81.6%	18.4%
	Theil *T*	0.0777	0.0637	0.0140	82.0%	18.0%
	Gini	0.2146				
Nurses	Theil *L*	0.1146	0.0900	0.0246	78.6%	21.4%
	Theil *T*	0.1130	0.0891	0.0239	78.9%	21.1%
	Gini	0.2617				

[1] N=1560 rural counties.

health-worker distribution: Gini of 0.4258 for all health workers, 0.4413 for doctors plus nurses, 0.4365 for doctors and 0.5271 for nurses. Secondly, the distribution of nurses is significantly more unequal than that of doctors or of all health workers: this holds nationally, within- and between-state, and for both the Theil *T* and Gini indices. Thirdly, between-state inequality accounts for about 20% of overall semidistrict inequality in the case of all health workers (for both Theil *L* and Theil *T*) and of doctors plus nurses (Theil *T*); for doctors between-state inequality is less than 20% and for nurses it is more than 35% (Theil *T*). Hence, semidistrict inequality in the distributions of health workers is explained largely by within-state, rather than between-state, inequalities.

As most of health-worker inequality arises within states, the health-worker distribution within each state is examined in detail. Tables I-2a, I-2b, I-2c and I-2d summarize the health-worker distribution in each state for the four categories of health workers. There are 35 states in India, which include seven Union Territories. For each state the tables display the number of semidistricts (N); the number of health workers; the mean, minimum and maximum semidistrict density; and the three inequality measures (Theil *L*, Theil *T* and Gini) for the distribution.

Table I-2a shows that the density of all health workers in India is 1.94 per 1000 population, but that it varies considerably between states. The Union Territories and some small states have the highest densities – Chandigarh with 6.72 per 1000 population, Pondicherry 5.23, Delhi 4.73, Mizoram 4.56 and Goa 4.33. In states with at least 1000 health workers, the lowest densities are found in Bihar (1.03 per 1000 population), Uttar Pradesh (1.29), Rajasthan (1.39) and Jharkhand (1.41). As shown in Table I-1, these between-state differences explain about 20% of overall inter-semidistrict inequality – 20.2% by Theil *L* and 21.8% by Theil *T*.

The minimum and maximum semidistrict density for each state is also shown in Table I-2a. There are large variations among the states, which are related to their within-state inequality levels. Thus, among the larger states, Uttar Pradesh has a maximum-to-minimum ratio of 15.2 and a Gini of 0.4307; Karnataka has a max-min ratio of 19.3 and a Gini of 0.4509; Rajasthan has a max-min ratio of 18.2 and a Gini of 0.4280; and Gujarat has a max-min ratio of 16.6 and a Gini of 0.4374. Within-state inequality is high in all the large states (with 100 000 or more health workers) except Kerala, which has both a high mean density of 3.69 per 1000 population and a Gini of only 0.2344. The two Theil indices yield similar rankings of the states.

Table I-1. India: decomposition of inter-semidistrict inequality by state, 2001[1]

Health-worker category	Inequality measure	Overall inter-semidistrict inequality	Within-state inequality	Between-state inequality	Within-state inequality (% of overall)	Between-state inequality (% of overall)
All health workers	Theil *L*	0.3075	0.2453	0.0622	79.8%	20.2%
	Theil *T*	0.2960	0.2315	0.0646	78.2%	21.8%
	Gini	0.4258				
Doctors plus nurses	Theil *L*	…	…	…	…	…
	Theil *T*	0.3187	0.2491	0.0696	78.2%	21.8%
	Gini	0.4413				
Doctors	Theil *L*	…	…	…	…	…
	Theil *T*	0.3087	0.2537	0.0550	82.2%	17.8%
	Gini	0.4365				
Nurses	Theil *L*	…	…	…	…	…
	Theil *T*	0.4670	0.3009	0.1661	64.4%	35.6%
	Gini	0.5271				

[1] N=1167 semidistricts.

Notes:

1. Indian districts have been separated into rural and urban semidistricts, and all inequality measures are calculated across semidistricts.
2. The notation '…' in a cell of the table refers to the situation in which Theil L cannot be calculated because there is a zero observation for the relevant health-worker category.
3. For All health workers: N=1167 semidistricts; for Doctors plus nurses: N=1166 (Sikkim West-urban zero); for Doctors: N=1163 (Sikkim West-urban, Mizoram Mamit-rural, Meghalaya East Garo Hills-rural and Meghalaya South Garo Hills-rural zero); for Nurses: N=1166 (Sikkim West-urban zero).

Table I-2b shows the same statistics as Table I-2a, but for the restricted category of doctors plus nurses. In India as a whole there are 1.36 doctors plus nurses per 1000 population, ranging for the larger states from lows of 0.68 in Bihar, 0.96 in Uttar Pradesh, 1.01 in Rajasthan and 1.06 in Madhya Pradesh – to 1.51 in Tamil Nadu, 2.12 in Maharashtra and 2.67 in Kerala. Some of the Union Territories and smaller states have the highest densities of doctors plus nurses: Chandigarh 5.17, Pondicherry 3.63 and Goa 3.21. Between-state inequality for doctors plus nurses is low (Table I-1), accounting for 21.8% of inter-semidistrict inequality by Theil T.

As seen in Table I-1, both overall inequality and within-state inequality are slightly higher for doctors plus nurses than for all health workers. The overall Gini for doctors plus nurses is 0.4413, while that for all health workers is 0.4258. Statewise Ginis are also generally higher for doctors plus nurses than for all health workers (cf. Table I-2b and Table I-2a); for doctors plus nurses the Ginis are 0.4759 in Karnataka, 0.4610 in Madhya Pradesh, 0.4570 in Rajasthan and 0.4440 in Jharkhand. With the exception of Bihar, Orissa and Andaman & Nicobar Islands, the Ginis for doctors plus nurses are higher than those for all health workers in each state. The larger inequality levels for doctors plus nurses than for all health workers suggest that other categories of health worker may be distributed more equally.

Table I-2c and Table I-2d display the distribution of doctors by state and the distribution of nurses by state, respectively. The overall density of doctors in India is 0.78 per 1000 population and that of nurses is 0.58 per 1000 population. The distribution of doctors *overall* is more equal than that of nurses – the Gini for doctors is 0.4365 while that for nurses is 0.5271. From these two tables it is apparent that nurse inequality is larger than doctor inequality in about half the states. Because *between-state* inequality is much larger for nurses than for doctors (Table I-1), there is much greater overall inequality for nurses than for doctors.

There are some particularly large differences in states between nurse and doctor inequality in both directions. Table I-2c and Table I-2d reveal the following inequality coefficients in states: Bihar – nurse Gini 0.5295, doctor Gini 0.3400; West Bengal – nurse Gini 0.4815, doctor Gini 0.2831; Jharkhand – nurse Gini 0.5530, doctor Gini 0.3729; Maharashtra – nurse Gini 0.4482, doctor Gini 0.3503. In the opposite direction, large differences occur as follows: Orissa – doctor Gini 0.4323, nurse Gini 0.2153; Tamil Nadu – doctor Gini 0.4517, nurse Gini 0.3429; Arunachal Pradesh – doctor Gini 0.4942, nurse Gini 0.3326; Kerala – doctor Gini 0.3399, nurse Gini 0.2486. It can be verified from Table I-2b that the distribution of doctors plus nurses – overall and in most states – shows intermediate inequality between the distribution of doctors (Table I-2c) and that of nurses (Table I-2d). For certain purposes it is useful to consider doctors plus nurses as a combined category, for example for international comparison of health workers who interact with patients.

For different categories of health worker, Table I-3 shows overall semidistrict inequality, within- and between-stratum inequality, and the contributions of within-stratum and between-stratum inequality to overall inequality. For all health workers, the percentage contribution of between-stratum to overall inequality is as high as 68.6% by Theil L and 75.5% by Theil T. Nurses have the lowest between-stratum contribution to overall inequality at 52.3% by Theil T, while for doctors the between-stratum contribution is 73.7% by Theil T. In absolute terms, the between-stratum contribution for all four categories of health worker by Theil T is quite similar (lying in the range 0.21–0.24): it reflects the (mean) density differences for each category between urban and rural areas, which are all approximately 4-to-1 (Table I-4). The reason that nurses have a lower percentage between-stratum contribution to overall inequality compared with the other categories is that they display much larger absolute inequality within strata (Table I-4) and therefore much higher overall inequality.

Table I-4 summarizes the distribution of each category of health worker within each stratum – rural and urban – and shows the number of semidistricts; the number of health workers; the mean, minimum and maximum semidistrict density per 1000 population; and the three selected inequality measures. It is clear that semidistrict inequality for each category of health worker is much *lower* in urban areas than in rural areas, according to all computable measures. The Gini coefficient in urban areas is just over half that in rural areas: 0.1447 for all health workers in urban areas compared with 0.2727 in rural areas; 0.1609 for doctors plus nurses in urban areas compared with 0.2918 in rural areas; 0.1628 for doctors in urban areas compared with 0.2964 in rural areas; and 0.2469 for nurses in urban areas compared with 0.4674 in rural areas. According to the measures Theil L and Theil T, inequality for all health workers in urban areas is one quarter of that in rural areas. (For other categories of health worker, inequality in urban areas is a third to a quarter less than in rural areas as indicated by Theil T.) Urban areas, although more homogenous in this respect than rural areas, have a mean density of health workers which is about four times that in rural areas. As seen in Table I-4, the density of all health workers in urban areas is 4.17 per 1000 population compared with 1.08 per 1000 population in rural areas – a factor of 3.86. For nurses, the density in urban areas is 1.28 compared with 0.31 in rural areas – a factor of 4.13. (For doctors this factor is 3.91, and for doctors plus nurses it is 4.00.)

From a comparison of Table I-3 and Table I-1, it is possible to assess the relative magnitudes of between-stratum and between-state inequality. It is seen that the absolute value of between-stratum inequality is some three to four times larger than that of between-state inequality for all categories of health worker except nurses (for whom it is about 1.5 times as large). For all health workers, the percentage contribution of between-stratum to overall inequality is 68.6% (Theil L) and 75.5% (Theil T), whereas the contribution of between-state to overall inequality is 20.2% (Theil L) and 21.8% (Theil T). The corresponding numbers for doctors plus nurses is 73.6% (between-stratum contribution according to Theil T) compared with 21.8% (between-state contribution by Theil T). For doctors the corresponding numbers are 73.7% compared with 17.8%; for nurses the corresponding numbers are 52.3% compared with 35.6%. In accounting for overall semidistrict inequality in India, it is evident that rural–urban differences in health-worker density are hugely more significant than are between-state differences.

Table I-5 shows a *two-way* decomposition of semidistrict inequality computed by state *and* stratum. We naturally expect between-state-*and*-stratum inequality to be larger than or equal to each of between-state inequality and between-stratum inequality – because the partition of the population by state-*and*-stratum is finer than that by state alone or that by stratum alone. In the case of India, we find between-state-*and*-stratum inequality to be (slightly) less than the sum of between-state inequality and between-stratum inequality. The variables state and stratum have independent explanatory power in accounting for overall semidistrict inequality – the sum of the between-stratum contribution (approximately 75% except for nurses) and the between-state contribution (approximately 20% except for nurses) is only slightly less than the between-stratum-and-state contribution (approximately 85–90% including for nurses).

Examining Table I-5 in detail, between-state-and-stratum inequality accounts for as much as 87.4% by Theil L and 89.9% by Theil T of overall semidistrict inequality for all health workers. This compares with between-state inequality accounting for 20.2% (Theil L) and 21.8% (Theil T), and between-stratum inequality accounting for 68.6% (Theil L)

Table I-2a. All health workers, India: density and inter-semidistrict inequality by state, 2001

			Density per 1000 population			Inequality measure		
State	N	No. of health workers	Mean	Min semi-district	Max semi-district	Theil *L*	Theil *T*	Gini
Jammu & Kashmir	28	21 950	2.16	0.88	12.95	0.2373	0.2580	0.3741
Himachal Pradesh	22	15 517	2.55	1.25	16.13	0.1742	0.2274	0.2883
Punjab	34	63 953	2.63	0.91	6.17	0.1689	0.1687	0.3141
Chandigarh	2	6 052	6.72	2.20	7.23	0.0478	0.0339	0.1234
Uttaranchal	26	17 712	2.09	0.58	7.77	0.1733	0.1716	0.3147
Haryana	38	41 648	1.97	0.12	6.49	0.2698	0.2388	0.3653
Delhi	16	65 488	4.73	1.72	7.04	0.0269	0.0254	0.1247
Rajasthan	64	78 716	1.39	0.34	6.19	0.3116	0.3319	0.4280
Uttar Pradesh	140	213 545	1.29	0.34	5.17	0.2693	0.2899	0.4037
Bihar	74	85 223	1.03	0.36	6.95	0.2647	0.3466	0.3843
Sikkim	8	2 220	4.11	2.17	12.82	0.1262	0.1515	0.2343
Arunachal Pradesh	25	2 877	2.62	1.05	7.72	0.1902	0.2019	0.3298
Nagaland	16	5 327	2.68	0.64	9.16	0.2292	0.2050	0.3148
Manipur	14	5 277	2.44	0.75	6.72	0.2257	0.2296	0.3737
Mizoram	15	4 052	4.56	1.83	9.47	0.1586	0.1407	0.2998
Tripura	8	5 630	1.76	0.74	5.43	0.2314	0.2761	0.2997
Meghalaya	14	3 472	1.50	0.26	6.79	0.5076	0.5146	0.5047
Assam	46	38 327	1.44	0.48	6.76	0.2603	0.3142	0.3848
West Bengal	35	187 138	2.33	0.58	6.57	0.2054	0.2035	0.3452
Jharkhand	36	37 877	1.41	0.40	5.67	0.3148	0.3447	0.4280
Orissa	60	70 797	1.92	1.07	5.90	0.0917	0.1032	0.2353
Chhattisgarh	32	33 324	1.60	0.66	5.95	0.1857	0.2083	0.3222
Madhya Pradesh	90	95 349	1.58	0.55	4.93	0.3115	0.3234	0.4212
Gujarat	49	86 852	1.71	0.29	4.81	0.3439	0.3100	0.4374
Daman & Diu	4	363	2.30	0.53	4.73	0.2128	0.1915	0.2910
Dadra & Nagar Haveli	2	275	1.25	0.79	2.79	0.1694	0.1899	0.1299
Maharashtra	68	274 007	2.83	0.88	7.39	0.2413	0.2223	0.3693
Andhra Pradesh	45	157 617	2.07	0.84	5.38	0.1631	0.1728	0.3121
Karnataka	54	106 325	2.01	0.45	8.67	0.3641	0.3386	0.4509
Goa	4	5 839	4.33	2.86	6.42	0.0552	0.0547	0.1778
Lakshadweep	2	222	3.66	3.36	4.04	0.0043	0.0043	0.0412
Kerala	28	117 561	3.69	1.73	9.06	0.0929	0.0900	0.2344
Tamil Nadu	59	135 266	2.17	0.51	4.79	0.2535	0.2265	0.3726
Pondicherry	6	5 091	5.23	1.93	7.05	0.1356	0.1101	0.2767
Andaman & Nicobar Islands	3	1 687	4.74	3.48	6.77	0.0478	0.0485	0.1264
Total India	**1 167**	**1 992 576**	**1.94**	**0.12**	**16.13**	**0.3075**	**0.2960**	**0.4258**

Table I-2b. Doctors plus nurses, India: density and inter-semidistrict inequality by state, 2001

			Density per 1000 population			Inequality measure		
State	N	No. of doctors plus nurses	Mean	Min semi-district	Max semi-district	Theil *L*	Theil *T*	Gini
Jammu & Kashmir	28	13 325	1.31	0.24	8.73	0.3871	0.3969	0.4595
Himachal Pradesh	22	9 293	1.53	0.61	12.21	0.2483	0.3249	0.3567
Punjab	34	46 919	1.93	0.67	4.79	0.1798	0.1789	0.3242
Chandigarh	2	4 652	5.17	1.24	5.61	0.0716	0.0460	0.1396
Uttaranchal	26	12 831	1.51	0.35	4.39	0.1908	0.1808	0.3265
Haryana	38	29 582	1.40	0.07	4.80	0.2955	0.2566	0.3795
Delhi	16	46 734	3.37	0.86	5.53	0.0457	0.0425	0.1602
Rajasthan	64	56 873	1.01	0.23	4.96	0.3580	0.3760	0.4570
Uttar Pradesh	140	160 226	0.96	0.26	3.66	0.2758	0.2925	0.4090
Bihar	74	56 472	0.68	0.25	3.99	0.2580	0.3382	0.3789
Sikkim	7	1 013	1.87	0.00	9.62	…	0.2067	0.2786
Arunachal Pradesh	25	1 493	1.36	0.43	3.88	0.2314	0.2451	0.3624
Nagaland	16	3 476	1.75	0.30	6.38	0.2926	0.2466	0.3508
Manipur	14	3 328	1.54	0.45	4.56	0.2380	0.2537	0.3820
Mizoram	15	1 440	1.62	0.31	3.05	0.2564	0.2276	0.4090
Tripura	8	3 183	1.00	0.35	3.18	0.2868	0.3399	0.3315
Meghalaya	14	2 447	1.06	0.16	4.79	0.5237	0.5208	0.5132
Assam	46	24 828	0.93	0.26	5.31	0.3054	0.3678	0.4172
West Bengal	35	124 777	1.56	0.43	3.91	0.2012	0.1998	0.3470
Jharkhand	36	25 755	0.96	0.29	4.69	0.3387	0.3845	0.4440
Orissa	60	53 251	1.45	0.78	3.97	0.0859	0.0925	0.2317
Chhattisgarh	32	20 847	1.00	0.32	3.60	0.2633	0.2869	0.3873
Madhya Pradesh	90	63 788	1.06	0.27	3.46	0.3753	0.3823	0.4610
Gujarat	49	56 624	1.12	0.16	3.49	0.3794	0.3369	0.4538
Daman & Diu	4	222	1.40	0.22	3.10	0.3376	0.2881	0.3395
Dadra & Nagar Haveli	2	206	0.93	0.57	2.16	0.1886	0.2112	0.1374
Maharashtra	68	205 330	2.12	0.57	5.65	0.2594	0.2381	0.3823
Andhra Pradesh	45	104 845	1.38	0.54	3.70	0.1718	0.1796	0.3244
Karnataka	54	75 382	1.43	0.27	6.74	0.4128	0.3728	0.4759
Goa	4	4 326	3.21	1.94	4.80	0.0650	0.0629	0.1909
Lakshadweep	2	125	2.06	1.78	2.41	0.0114	0.0114	0.0670
Kerala	28	85 115	2.67	1.12	6.98	0.1141	0.1088	0.2568
Tamil Nadu	59	94 083	1.51	0.34	3.64	0.2687	0.2405	0.3866
Pondicherry	6	3 533	3.63	1.10	5.06	0.1795	0.1398	0.3119
Andaman & Nicobar Islands	3	854	2.40	1.85	3.15	0.0328	0.0328	0.0968
Total India	**1 166**	**1 397 177**	**1.36**	**0.00**	**12.21**	**…**	**0.3187**	**0.4413**

Table I-2c. Doctors, India: density and inter-semidistrict inequality by state, 2001

			Density per 1000 population			Inequality measure		
State	N	No. of doctors	Mean	Min semi-district	Max semi-district	Theil *L*	Theil *T*	Gini
Jammu & Kashmir	28	7 897	0.78	0.12	3.38	0.6139	0.6042	0.5647
Himachal Pradesh	22	5 219	0.86	0.30	4.50	0.2531	0.3075	0.3699
Punjab	34	32 280	1.33	0.54	3.27	0.1850	0.1855	0.3252
Chandigarh	2	2 479	2.75	0.54	3.00	0.0875	0.0530	0.1474
Uttaranchal	26	8 116	0.96	0.19	2.93	0.2440	0.2168	0.3695
Haryana	38	22 281	1.05	0.06	3.27	0.2682	0.2281	0.3635
Delhi	16	26 923	1.94	0.32	2.78	0.0414	0.0368	0.1495
Rajasthan	64	30 596	0.54	0.07	2.07	0.4344	0.4227	0.4940
Uttar Pradesh	140	123 448	0.74	0.17	2.54	0.2685	0.2770	0.4048
Bihar	74	41 725	0.50	0.19	3.01	0.2089	0.2722	0.3400
Sikkim	7	255	0.47	0.00	2.18	…	0.4582	0.4045
Arunachal Pradesh	25	353	0.32	0.08	1.49	0.4462	0.4714	0.4942
Nagaland	16	695	0.35	0.05	1.44	0.4588	0.4102	0.4563
Manipur	14	1 155	0.53	0.08	2.04	0.5380	0.5161	0.5499
Mizoram	14	402	0.45	0.00	0.97	…	0.3034	0.4817
Tripura	8	1 685	0.53	0.10	1.88	0.3754	0.4014	0.3968
Meghalaya	12	636	0.27	0.00	1.28	…	0.7401	0.5934
Assam	46	10 462	0.39	0.08	2.91	0.4220	0.5045	0.4860
West Bengal	35	78 571	0.98	0.31	2.13	0.1353	0.1268	0.2831
Jharkhand	36	12 799	0.48	0.16	1.78	0.2331	0.2532	0.3729
Orissa	60	15 804	0.43	0.07	2.07	0.3363	0.3763	0.4323
Chhattisgarh	32	10 951	0.53	0.10	1.55	0.3142	0.2894	0.4171
Madhya Pradesh	90	38 491	0.64	0.10	2.17	0.3424	0.3442	0.4448
Gujarat	49	31 882	0.63	0.05	2.03	0.4219	0.3602	0.4608
Daman & Diu	4	96	0.61	0.18	1.17	0.2064	0.1918	0.2760
Dadra & Nagar Haveli	2	89	0.40	0.19	1.11	0.3332	0.3648	0.1833
Maharashtra	68	109 067	1.13	0.27	3.22	0.2177	0.2001	0.3503
Andhra Pradesh	45	64 158	0.84	0.31	2.22	0.1724	0.1718	0.3270
Karnataka	54	43 320	0.82	0.12	2.59	0.4068	0.3455	0.4623
Goa	4	1 730	1.28	0.49	2.29	0.2142	0.1890	0.3169
Lakshadweep	2	35	0.58	0.51	0.67	0.0097	0.0097	0.0619
Kerala	28	30 544	0.96	0.39	3.31	0.1926	0.2056	0.3399
Tamil Nadu	59	44 028	0.71	0.10	1.99	0.3917	0.3311	0.4517
Pondicherry	6	1 146	1.18	0.23	1.70	0.2884	0.1982	0.3560
Andaman & Nicobar Islands	3	235	0.66	0.43	1.12	0.1058	0.1095	0.1949
Total India	**1 163**	**799 550**	**0.78**	**0.00**	**4.50**	**...**	**0.3087**	**0.4365**

Table I-2d. Nurses, India: density and inter-semidistrict inequality by state, 2001

			Density per 1000 population			Inequality measure		
State	N	No. of nurses	Mean	Min semi-district	Max semi-district	Theil *L*	Theil *T*	Gini
Jammu & Kashmir	28	5 428	0.54	0.13	6.08	0.2449	0.3034	0.3713
Himachal Pradesh	22	4 074	0.67	0.19	7.93	0.2996	0.4014	0.3923
Punjab	34	14 639	0.60	0.07	1.62	0.2131	0.1933	0.3422
Chandigarh	2	2 173	2.41	0.70	2.61	0.0573	0.0390	0.1308
Uttaranchal	26	4 715	0.56	0.09	2.18	0.3055	0.2643	0.3929
Haryana	38	7 301	0.35	0.01	1.75	0.4934	0.4205	0.4865
Delhi	16	19 811	1.43	0.35	3.53	0.0712	0.0699	0.2048
Rajasthan	64	26 277	0.47	0.11	3.11	0.3505	0.3780	0.4566
Uttar Pradesh	140	36 778	0.22	0.04	1.41	0.3990	0.4322	0.4854
Bihar	74	14 747	0.18	0.01	1.45	0.4979	0.6005	0.5295
Sikkim	7	758	1.40	0.00	8.81	…	0.1477	0.2375
Arunachal Pradesh	25	1 140	1.04	0.29	3.07	0.1941	0.2008	0.3326
Nagaland	16	2 781	1.40	0.26	5.25	0.2709	0.2282	0.3354
Manipur	14	2 173	1.00	0.27	2.51	0.1627	0.1673	0.3113
Mizoram	15	1 038	1.17	0.20	2.12	0.2268	0.1986	0.3826
Tripura	8	1 498	0.47	0.10	1.63	0.3361	0.3473	0.3698
Meghalaya	14	1 811	0.78	0.16	4.17	0.4212	0.4326	0.4702
Assam	46	14 366	0.54	0.17	3.40	0.2599	0.3058	0.3894
West Bengal	35	46 207	0.58	0.10	2.08	0.4272	0.4171	0.4815
Jharkhand	36	12 956	0.48	0.07	3.12	0.5538	0.5920	0.5530
Orissa	60	37 447	1.02	0.52	2.34	0.0744	0.0726	0.2153
Chhattisgarh	32	9 896	0.48	0.14	2.46	0.3797	0.4082	0.4589
Madhya Pradesh	90	25 298	0.42	0.07	1.94	0.4956	0.4833	0.5195
Gujarat	49	24 741	0.49	0.09	1.71	0.3775	0.3412	0.4640
Daman & Diu	4	126	0.80	0.04	1.93	0.5301	0.3823	0.3880
Dadra & Nagar Haveli	2	117	0.53	0.38	1.05	0.1086	0.1214	0.1026
Maharashtra	68	96 264	0.99	0.19	3.23	0.3508	0.3164	0.4482
Andhra Pradesh	45	40 687	0.53	0.20	1.98	0.2393	0.2501	0.3737
Karnataka	54	32 062	0.61	0.08	4.15	0.5790	0.4883	0.5359
Goa	4	2 596	1.93	1.44	2.51	0.0213	0.0211	0.1069
Lakshadweep	2	90	1.48	1.28	1.74	0.0121	0.0121	0.0690
Kerala	28	54 571	1.71	0.68	4.83	0.1072	0.1031	0.2486
Tamil Nadu	59	50 055	0.80	0.18	2.11	0.2044	0.1862	0.3429
Pondicherry	6	2 387	2.45	0.87	3.35	0.1461	0.1184	0.2915
Andaman & Nicobar Islands	3	619	1.74	1.43	2.40	0.0213	0.0217	0.0759
Total India	**1 166**	**597 627**	**0.58**	**0.00**	**8.81**	**…**	**0.4670**	**0.5271**

and 75.5% (Theil T). Thus the sum of between-state inequality and between-stratum inequality is 88.8% (Theil L) and 97.3% (Theil T). Hence the *two-way* between-state-*and*-stratum inequality is just 1.4 percentage points less (for Theil L) and 7.4 points less (for Theil T) than the sum of the one-way contributions.

For all four categories of health worker, the between-state-*and*-stratum contributions to overall semidistrict inequality are between 85% and 90% (Table I-5). For doctors plus nurses the contribution is 88.1% by Theil T; for doctors the corresponding number is 85.7%, and for nurses it is 85.4%. These findings in Table I-5 are very significant indeed. In India, just the two variables of state and stratum seem to account for almost all of semidistrict inequality. A policy focus to redress *average* health-worker density differences between rural and urban areas would itself reduce overall inequality by about 70%; and it *together* with an elimination of between-state differences would reduce overall inequality to approximately 10% of its current levels. Concentrating on these two variables alone would appear to be sufficient to achieve a huge reduction in overall health workforce inequalities in India. Variations arising from other factors account for a relatively minor fraction (about 10%) of overall semidistrict inequalities in the health-worker distributions.

Table I-3. India: decomposition of semidistrict inequality by rural–urban stratum, 2001

Health-worker category	Inequality measure	Overall semidistrict inequality	Within-stratum inequality	Between-stratum inequality	Within-stratum inequality (% of overall)	Between-stratum inequality (% of overall)
All health workers	Theil *L*	0.3075	0.0966	0.2108	31.4%	68.6%
	Theil *T*	0.2960	0.0724	0.2236	24.5%	75.5%
	Gini	0.4258				
Doctors plus nurses	Theil *L*	…	…	…	…	…
	Theil *T*	0.3187	0.0841	0.2346	26.4%	73.6%
	Gini	0.4413				
Doctors	Theil *L*	…	…	…	…	…
	Theil *T*	0.3087	0.0814	0.2274	26.4%	73.7%
	Gini	0.4365				
Nurses	Theil *L*	…	…	…	…	…
	Theil *T*	0.4670	0.2229	0.2442	47.7%	52.3%
	Gini	0.5271				

Notes:

1. The notation '…' in a cell of the table refers to the situation in which Theil L cannot be calculated because there is a zero observation for the relevant health-worker category.
2. For All health workers, N=1167 semidistricts; for Doctors plus nurses, N=1166 (Sikkim West-urban zero); for Doctors, N=1163 (Sikkim West-urban, Mizoram Mamit-rural, Meghalaya East Garo Hills-rural and Meghalaya South Garo Hills-rural zero); for Nurses, N=1166 (Sikkim West-urban zero).

Table I-4. India: density and semidistrict inequality by rural–urban stratum, 2001

	Stratum	Rural	Urban	National
	N (semidistricts)	584	583	1 167
Health-worker category				
All health workers	Number	799 205	1 193 371	1 992 576
	Mean density	1.08	4.17	1.94
	Min semidistrict density	0.12	0.85	0.12
	Max semidistrict density	6.31	16.13	16.13
	Theil *L*	0.1208	0.0340	0.3075
	Theil *T*	0.1302	0.0337	0.2960
	Gini	0.2727	0.1447	0.4258
Doctors plus nurses	Number	549 247	847 930	1 397 177
	Mean density	0.74	2.96	1.36
	Min semidistrict density	0.07	0.00	0.00
	Max semidistrict density	4.75	12.21	12.21
	Theil *L*	0.1383	…	…
	Theil *T*	0.1499	0.0415	0.3187
	Gini	0.2918	0.1609	0.4413
Doctors	Number	318 406	481 144	799 550
	Mean density	0.43	1.68	0.78
	Min semidistrict density	0.00	0.00	0.00
	Max semidistrict density	1.92	4.50	4.50
	Theil *L*	…	…	…
	Theil *T*	0.1406	0.0422	0.3087
	Gini	0.2964	0.1628	0.4365
Nurses	Number	230 841	366 786	597 627
	Mean density	0.31	1.28	0.58
	Min semidistrict density	0.01	0.00	0.00
	Max semidistrict density	3.68	8.81	8.81
	Theil *L*	0.3757	…	…
	Theil *T*	0.4191	0.0994	0.4670
	Gini	0.4674	0.2469	0.5271

Notes:

1. The notation '…' in a cell of the table refers to the situation in which Theil L cannot be calculated because there is a zero observation for the relevant health-worker category.
2. For All health workers, N=1167 semidistricts; for Doctors plus nurses, N=1166 (Sikkim West-urban zero); for Doctors, N=1163 (Sikkim West-urban, Mizoram Mamit-rural, Meghalaya East Garo Hills-rural and Meghalaya South Garo Hills-rural zero); for Nurses, N=1166 (Sikkim West-urban zero).

Table I-5. India: decomposition of semidistrict inequality by state and rural–urban stratum, 2001[1]

Health-worker category	Inequality measure	Overall semidistrict inequality	Within-state-*and*-stratum inequality	Between-state-*and*-stratum inequality	Within-state-*and*-stratum inequality (% of overall)	Between-state-*and*-stratum inequality (% of overall)
All health workers	Theil *L*	0.3075	0.0387	0.2688	12.6%	87.4%
	Theil *T*	0.2960	0.0299	0.2661	10.1%	89.9%
	Gini	0.4258				
Doctors plus nurses	Theil *L*	…	…	…	…	…
	Theil *T*	0.3187	0.0380	0.2807	11.9%	88.1%
	Gini	0.4413				
Doctors	Theil *L*	…	…	…	…	…
	Theil *T*	0.3087	0.0441	0.2646	14.3%	85.7%
	Gini	0.4365				
Nurses	Theil *L*	…	…	…	…	…
	Theil *T*	0.4670	0.0682	0.3989	14.6%	85.4%
	Gini	0.5271				

[1] N=1167 semidistricts.

Notes:

1. The notation '…' in a cell of the table refers to the situation in which Theil L cannot be calculated because there is a zero observation for the relevant health-worker category.
2. The doctor density was 0 in four semidistricts: Mizoram (Mamit-rural), Meghalaya (East Garo Hills-rural, South Garo Hills-rural) and Sikkim (West-urban). The nurse density was 0 in one semidistrict: Sikkim (West-urban).

4

Concluding remarks: a comparison of China and India

This study has presented and illustrated the use of indices for measuring inequalities in the distribution of the health workforce by geographical units. The reader will have noticed that the tables have been prepared and labelled to allow ease of comparison between similar statistics of the health-worker distributions in China 2005 and India 2001. This brief concluding section highlights some of the main differences and similarities between the health workforce distributions in China and India. Although the years for which the Chinese and Indian data were analysed are not the same, the net change in the density of health workers in India between 2001 and 2005, and especially in their distribution, may be expected to be relatively small and not to affect significantly any health workforce inequality comparisons between the two countries.

The overall density of health workers in China in 2005 was much higher than that in India in 2001. The density of health professionals in China is estimated to be 3.06 per 1000 population, whereas the density of all health workers in India is 1.94 per 1000 population. For the combined and comparable category of doctors plus nurses, China has a density of 2.26 per 1000, whereas India has a density of 1.36 per 1000. For the comparable category of doctors, the density is 1.30 in China and 0.78 in India; and for nurses the density is 0.96 in China and 0.58 in India.

The levels of overall inequality in the health-worker distributions are also significantly higher in India than in China. Thus, the Gini for all health workers in India is 0.4258, whereas the Gini for health professionals in China is 0.3718. For the combined category of doctors plus nurses, the Gini in India is 0.4413 and in China 0.4030. For doctors, the Gini is 0.4365 in India and 0.3622 in China; and for nurses the Gini is 0.5271 in India and 0.4714 in China. These inequality differences are particularly significant because the number of units (semidistricts) across which they are calculated in India (N=1167) is *smaller* than the number of units (counties) across which they are calculated in China (N=2854). Other things being equal, one would expect the level of inequality to be higher for a health-worker inequality estimate across a larger number of units – this is analogous to calculating the Gini coefficient for a Lorenz curve defined by a larger number of points.[7] Yet the country (China) with more units (counties) has a *smaller* level of inequality than the country (India) with fewer units (semidistricts). Consequently the finding of greater inequality in India than in China is especially significant.

A point of similarity between the two countries is that nurses are more unequally distributed than the other categories of health worker. It was also found that the within-province contribution to overall inequality in China and the within-state contribution to overall inequality in India are approximately *similar* – at about 80% (except for nurses who show a smaller within-state contribution to inequality in India).

By contrast, the between-stratum contributions to overall inequality are *very* different in the two countries. In India, except for nurses, they are more than *twice* as large as they are in China. Thus, the between-stratum contribution to overall inequality in India is approximately 70–75% (except for nurses for whom it is about 50%), whereas the between-stratum contribution to inequality in China is only about 30–35%. This represents a huge difference between the two countries, which is in part accounted for by the large differences in urban–rural disparity ratios in health-worker densities. The urban–rural mean density ratio for all health workers in India is 3.86, whereas for health professionals in China the ratio is 2.19. For the internationally comparable category of doctors plus nurses, the ratio in India is 4.00, whereas in China it is 2.39. For doctors, the ratio in India is 3.91 and in China 2.07; and for nurses the ratio in India is 4.13 and in China 2.98. The urban–rural disparity ratios in health-worker density are almost twice as high in India as in China.

Perhaps the most important difference between the two countries is the magnitude of the between-state-*and*-stratum contribution to overall inequality in India compared with the between-province-*and*-stratum contribution to inequality in China. In India, 85–90% of overall inequality is explained by just the two variables of state and stratum. In China, however, only some 40–50% of inequality is explained by these two variables. Thus in China there are large variations within province-and-stratum, unlike in India within state-and-stratum. Reducing state-and-stratum density differences in India will thus achieve much greater reductions in overall health-worker inequality than reducing province-and-stratum differences in China.

[7] Given the large absolute number of units in the two countries (2854 in China and 1167 in India), it is unlikely that the Gini for India is underestimated significantly relative to the Gini for China. Other things being equal, the inequality estimate is likely to be a function of the *absolute* number of units used in the calculations – and the absolute number of units is very high in both cases.

References

Anand S. *Inequality and poverty in Malaysia: measurement and decomposition*. New York, NY, Oxford University Press, 1983.

Anand S, Bärnighausen T. Human resources and health outcomes: cross-country econometric study. *The Lancet*, 2004, 364:1603–1609 (http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(04)17313-3/fulltext).

Anand S, Bärnighausen T. Health workers and vaccination coverage in developing countries: an econometric analysis. *The Lancet*, 2007, 369:1277–1285 (http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(07)60599-6/fulltext).

Anand S et al. China's human resources for health: quantity, quality, and distribution. *The Lancet*, 2008, 372:1774–1781 (http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(08)61363-X/fulltext).

Dal Poz M et al., eds. *Handbook on monitoring and evaluation of human resources for health: with special applications for low- and middle-income countries*. Geneva, World Health Organization, World Bank and United States Agency for International Development, 2009 (http://www.who.int/hrh/resources/handbook/en/index.html).

International Labour Organization. *International standard classification of occupations*, ISCO-88. Geneva, International Labour Organization, 2004 (http://www.ilo.org/public/english/bureau/stat/isco/isco88/index.htm).

Jenkins SP. Analysis of income distributions. *Stata Technical Bulletin*, 1999, 48:4–18 (http://www.stata.com/products/stb/journals/stb48.pdf).

Ministry of Health, China. *2006 China human resources for health report*. Beijing, Peking Union Medical Press, 2007 (in Chinese).

Ministry of Labour, India. *National classification of occupations 2004*. New Delhi, Ministry of Labour, Government of India, 2004 (http://dget.nic.in/nco).

Ministry of Public Security, China. Dongli Wu et al., eds. *China county and city population statistical information 2005*. Beijing, Da Zhong Publishing House, 2006.

Sen A. *On economic inequality: expanded edition with a substantial annexe by James E. Foster and Amartya Sen*. Oxford, Clarendon Press, 1997.

StataCorp. *Stata statistical software: Release 11*. College Station, TX, StataCorp LP, 2009.

Theil H. *Economics and information theory*. Amsterdam, North Holland, 1967.

World Health Organization. *The world health report 2006 – Working together for health*. Geneva, World Health Organization, 2006 (http://www.who.int/whr/2006/en/).

World Health Organization. *Not enough here ... too many there: health workforce in India*. New Delhi, World Health Organization Country Office for India, 2007 (http://www.whoindia.org/EN/Section2/Section404_1264.htm).

Annexes

Annex 1: Decomposition of the squared coefficient of variation C²

The squared coefficient of variation C^2 is the variance σ^2 of the health-worker distribution divided by the square of its arithmetic mean, X^2, i.e. $C^2 = \sigma^2/X^2$. The variance σ^2 can be decomposed within- and between- groups 1 and 2 as follows (where subscripts denote values of the measures for the two groups, respectively):

$$\begin{aligned} \sigma^2 &= \textstyle\sum^{n}_{i=1} (p_i/P)\,(x_i - X)^2 \\ &= (P_1/P)\,\sigma_1^2 + (P_2/P)\,\sigma_2^2 + (P_1/P)\,(X_1 - X)^2 + (P_2/P)\,(X_2 - X)^2 \\ &= \sigma_W^2 + \sigma_B^2 \quad \text{as in standard analysis of variance.} \end{aligned}$$

Therefore,

$$\begin{aligned} C^2 &= \sigma^2/X^2 \\ &= (P_1/P)\,(X_1^2/X^2)\,C_1^2 + (P_2/P)\,(X_2^2/X^2)\,C_2^2 + \\ &\qquad (1/X^2)\,[(P_1/P)\,(X_1 - X)^2 + (P_2/P)\,(X_2 - X)^2] \\ &= C_W^2 + C_B^2 \end{aligned}$$

The between-group component C_B^2 is the squared coefficient of variation for the distribution which has P_1 persons with health-worker density X_1, and P_2 persons with density X_2. The within-group component is a *weighted sum* of the squared coefficients of variation for each group, C_1^2 and C_2^2, where the weights depend only on the population share and health-worker share of the groups. However, the within-group component is *not* a weighted *average* of the group inequality indices C_1^2 and C_2^2 because the weights do not add up to unity. In fact, the sum of the weights on C_1^2 and C_2^2 is

$$\begin{aligned} &(1/X^2)\,[(P_1/P)\,X_1^2 + (P_2/P)\,X_2^2] \\ &= 1 + (1/X^2)\,[(P_1/P)\,(X_1 - X)^2 + (P_2/P)\,(X_2 - X)^2] \\ &= 1 + C_B^2. \end{aligned}$$

Thus the weights on the squared within-group coefficients of variation sum to a number greater than unity, by an amount which is equal to the between-group component C_B^2.[1] This implies that the within-group component is not a weighted average of the squared coefficients of variation of the groups. It also suffers from the unfortunate feature of the within-group component depending on the between-group component, making interpretation problematic. The squared coefficient of variation C^2 (or GE(2)) is therefore not used for decomposition analysis in this study.

Reference

Theil H. *Economics and information theory*. Amsterdam, North Holland, 1967.

[1] According to Theil (1967:125) this is "a disadvantage, because one should prefer a measure for which the within-set components, including their weights, are independent of the between-set component".

Annex 2: Data and health worker categories for China, 2005

County-level data on the numbers of health workers for 2005 were obtained through custom files provided by the Center for Health Statistics and Information of the Ministry of Health of China. Details on the data used, including differences between these county-level data and the province-level data published in a Ministry of Health (2007) report on human resources for health, are described in Anand et al. (2008:Webappendix 5).

The categories of health worker used in this study are: health professionals, doctors plus nurses, doctors, and nurses. The definitions of the health worker categories in China are as follows (see also Anand et al., 2008:Webappendix 1).

- *Health professionals:* doctors, nurses, pharmacists, laboratory technicians, clinical radiologists and other technical staff with advanced education.

- *Doctors:* those who pass a licensing examination and are registered at a county or higher-level health authority as either licensed doctors or licensed assistant doctors. Licensed doctors are medical graduates with a bachelor's or higher degree and a year's internship supervised by licensed doctors at a clinic or preventive or health-care institution. Licensed assistant doctors are medical graduates of three-year tertiary medical education programmes with an associate degree, or secondary education programmes with a diploma (two years' medical education after high school), followed by one year of internship supervised by licensed doctors at a clinic or preventive or health-care institution.

- *Nurses:* those who have obtained nursing certification with an associate degree (three years' tertiary nursing education) or higher, or graduates from secondary education programmes with a diploma (two years' nursing education after high school) and recommended by a health authority at provincial level or above. No examination is required.

References

Anand S et al. China's human resources for health: quantity, quality, and distribution. *The Lancet*, 2008, 372:1774–1781 (http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(08)61363-X/fulltext).

Ministry of Health, China. 2006 *China human resources for health report*. Beijing, Peking Union Medical Press, 2007 (in Chinese).

Annex 3: Data and health worker categories for India, 2001

District-level Census 2001 data were obtained through the Office of the Registrar General of India (Primary Census Abstract, B Series, Table 25) on occupation of "main" workers. Main workers were defined as those who worked for six months or more during the previous year. The National Classification of Occupations (NCO) was used to classify the occupation of workers.

The NCO uses a hierarchical structure of occupational titles and codes according to skill level and skill specialization.

There are 19 distinct categories of health worker at the four-digit ("Family") level of the NCO. For the purposes of this study, the census data were aggregated into four categories – doctors, nurses, doctors plus nurses, and all health workers – as seen in Table A.3.1 with corresponding NCO codes and definitions. The category "doctors plus nurses" includes all doctors plus nurses and midwives, while "all health workers" includes doctors, nurses and midwives, and all other health workers.

Table A.3.1. Health worker categories used in the study with corresponding National Classification of Occupations codes and definitions

Health worker categories used in the study	NCO four-digit code	Definition
Doctors	2221	*Physicians and surgeons, allopathic*: diagnose human ailments and treat them allopathically by medicines and surgical operations and specialize in treatment of diseases of particular types or disorders of particular parts of human body.
	2222	*Physicians and surgeons, ayurvedic*: conduct medical examinations, making diagnosis, prescribing and giving other forms of medical treatment based on ayurvedic system of medicine.
	2223	*Physicians and surgeons, homeopathic*: conduct medical examinations, making diagnosis, prescribing and giving other forms of medical treatment based on homeopathic system of medicine
	2224	*Physicians and surgeons, unani*: conduct medical examinations, making diagnosis, prescribing and giving other forms of medical treatment based on unani system of medicine.
Nurses and midwives	2230	*Nursing professionals*: provide professional, general or specialized nursing care for sick, injured and infirm for treatment of physical and mental disorders; give nursing care and advice; assist physicians and perform other nursing tasks and community health service in hospitals, clinics, sanatoria, schools, factories, medical establishments, private homes and elsewhere.
	3231	*Nursing associate professionals*: provide nursing care for the sick, injured, and others in need of such care, and, in the absence of medical doctors or professional nurses, deal with emergencies.
	3232	*Midwifery associate professionals*: deliver or assist doctors or midwifery professionals in the delivery of babies, provide antenatal and postnatal care and instruct parents in baby care.

(Table A.3.1 continues on next page)

Health worker categories used in the study	NCO four-digit code	Definition
Other health workers	2225	*Dental specialists*: conduct research, improve or develop concepts, theories and operational methods, and apply medical knowledge in the field of dentistry.
	2229	*Health professionals (except nursing) not elsewhere classified*: health professionals not classified elsewhere who diagnose and treat human ailments; this category includes health officers, hospital administrators, naturopath physicians, osteopathic physicians and related occupations.
	3221	*Medical assistants*: carry out advisory, diagnostic, preventive and creative and curative medical tasks, more limited in scope and complexity than those carried out by medical doctors; work independently or with the guidance and supervision of medical doctors in institutions or in the field as part of the public health service; and may work mainly with diseases and disorders common in their region or mainly apply specific types of treatment.
	3222	*Sanitarians*: provide technical assistance and advice on measures to restore or improve sanitary conditions, and supervise their implementation.
	3223	*Dieticians and nutritionists*: conduct research and improve or develop concepts and operational methods concerning the preparation and application of diets for general and therapeutic purposes.
	3224	*Optometrists and opticians*: prescribe and fit glasses and contact lenses and advise on their use or the use of other visual aids, as well as on proper lighting for work and reading.
	3225	*Dental assistants*: carry out advisory, diagnostic, preventive and curative dental tasks, more limited in scope and complexity than those carried out by dentists, and assist dentists by preparing and taking care of instruments and other equipment, preparing materials and helping patients prepare for examination and treatment.
	3226	*Physiotherapists and related associate professionals*: treat disorders of bones, muscles and parts of the circulatory or the nervous system by manipulative methods and ultrasound, heating, laser or similar techniques, or apply physiotherapy and related therapies as part of the treatment for the physically disabled, mentally ill or unbalanced.
	3228	*Pharmaceutical assistants*: dispense and prepare medicaments, lotions and mixtures under the guidance of pharmacists, in pharmacies, hospitals and dispensaries.
	3229	*Modern health associate professionals (except nursing) not elsewhere classified*: modern health associate professionals (except nursing) not classified elsewhere who practise, plan and carry out therapeutic activities, e.g. speech therapists and audiologists, orthotists and prosthetists, orientation and mobility instructors, and related occupations.
	3241	*Traditional medicine practitioners*: treat human mental and physical sickness by herbs, medicinal plants and other techniques traditionally used in the community, and believed to cure and heal by assisting or stimulating nature, and advise on methods to preserve or improve health and well being.
	3242	*Faith healers*: endeavour to cure human mental and physical illness by mental influence and suggestion, power of faith and spiritual advice.

Source: Adapted from Ministry of Labour (2004).

Census data on the geographical distribution of health workers are available at the semidistrict level, which corresponds to the rural or urban part of a district. In India in 2001 there were 584 rural semidistricts and 583 urban semidistricts, amounting to a total of 1167 semidistricts. Actually there were 593 districts in India in 2001, so that if each district had a rural and an urban semidistrict there would be a total of 1186 semidistricts. However, there were 19 districts which either did not have an urban population or did not have a rural population, e.g. the districts of Mumbai and New Delhi had a zero rural population. This leaves a total of 1167 semidistricts with non-zero populations. The 19 Indian semidistricts with zero population are shown in Table A.3.2.

Table A.3.2. Indian semidistricts with zero population, 2001 census

State or Union Territory	District	Stratum
Andaman & Nicobar Islands	Nicobars	urban
Himachal Pradesh	Lahul & Spiti	urban
Himachal Pradesh	Kinnaur	urban
Arunachal Pradesh	Upper Siang	urban
Manipur	Tamenglong	urban
Manipur	Ukhrul	urban
Manipur	Senapati	urban
Manipur	Churachandpur	urban
Mizoram	Lawngtlai	urban
Gujarat	The Dangs	urban
Maharashtra	Mumbai (Suburban)	rural
Maharashtra	Mumbai	rural
Delhi	Central	rural
Delhi	New Delhi	rural
Tamil Nadu	Chennai	rural
West Bengal	Kolkata	rural
Andhra Pradesh	Hyderabad	rural
Pondicherry	Yanam	rural
Pondicherry	Mahe	rural

References:

Ministry of Labour, India. *National classification of occupations 2004: code structure*. New Delhi, Directorate General of Employment and Training, Ministry of Labour, Government of India, 2004 (http://dget.nic.in/nco/CodeStructure.pdf).

Office of the Registrar General of India. *Census of India 2001*. New Delhi, Office of the Registrar General and Census Commissioner, Government of India, 2010 (http://www.censusindia.gov.in/).

Human Resources for Health Observer

Previous issues

No. 1 - Who are health managers? Case studies from three African countries

No. 2 - Análise dos recursos humanos da saúde (RHS) nos países africanos de língua oficial portuguesa (PALOP)

No. 3 - Models and tools for health workforce planning and projections

No. 4 - Exploring health workforce inequalities: case studies from three countries

Measuring health workforce inequalities: methods and application to China and India is for users and producers of quantitative data in support of decision-making for health policy and practice, including statistical analysts, researchers, health workforce planners and policy-makers, and development partners. With the aim of strengthening the global evidence base across countries and regions, it is intended to encourage a greater number of countries to monitor and report on health workforce inequalities.

Three indices adapted from the economics literature on measuring inequality – the Theil *L* measure, the Theil *T* index and the Gini coefficient – are elaborated so as to assess the distribution of the health workforce within and between geographical units in a country. A large part of this publication is devoted to presenting the application of these methods to an analysis of health workforce data from China and India, followed by a comparison of the distribution of human resources for health in the two countries.

About Human Resources for Health Observer

The WHO-supported regional health workforce observatories are cooperative mechanisms, through which health workforce stakeholders share experiences, information and evidence to inform and strengthen policy decision-making. The Human Resources for Health Observer series makes available the latest findings and research from different observatories to the widest possible audience.

The series covers a wide range of technical issues, in a variety of formats, including case studies, reports of consultations, surveys, and analysis. In cases where information has been produced in local languages, an English translation - or digest - will be made available.

This publication is available on the Internet at:
http://www.who.int/hrh/resources/observer/en/

Copies may be requested from: World Health Organization
Department of Human Resources for Health
CH-1211 Geneva 27, Switzerland

ISBN 978 92 4 150022 7