INCOME DISTRIBUTION: MEASUREMENT, TRANSITION AND ANALYSIS OF URBAN CHINA, 1981-1990

Stephen R. Howes

Submitted for the Degree of Ph.D.

London School of Economics and Political Science
University of London

1993

To my parents,

John and Margaret Howes

Abstract

Many aspects of economic analysis require judgements to be made about distributions. When agreement on a single criterion for judgement is not possible, it is necessary to examine whether one distribution is better than another from a number of perspectives. The problem of 'distributional dominance', which Part One addresses, is precisely this problem of ordering two distributions in relation to one or more objective functions, via use of a single 'dominance criterion'. Four themes are pursued.

• It is argued that welfare, poverty and inequality dominance criteria can be fruitfully analyzed within a single framework.

• The need to approach the problem of distributional dominance as a statistical one is stressed. Estimators and a method of inference are proposed and are themselves tested via a simulation study.

• The likely effect of aggregation on the attained ordering of distributions is assessed, also via a simulation study.

• A critical re-appraisal is presented of the most widely-used dominance criterion, second-order stochastic dominance, and alternative criteria are proposed. The usefulness of thinking of dominance criteria in terms of curves within bounds is emphasized.

Part Two of the thesis is a study of the distribution of income in urban China in the eighties, using both aggregated, nationwide data and disaggregated data for two provinces. This study is both an application of the methods developed in Part One and a case-study of the dynamics of income distribution in a transitional economy. Evidence is found that cash-income inequality has grown over the decade, and this is linked to the reform process. However, inequality remains exceptionally low by international standards. Moreover, both the system of price subsidies and that of cash compensation introduced to replace the subsidies are shown to have exerted an equalizing influence on the urban distribution of income.

Table of Contents

Tables

Chapter Five

Figures

Acknowledgements

In the course of writing this thesis, I have built up a large number of debts. The longest-standing is to my supervisor, Professor Nicholas Stern, whose advice, comments and patience have all been of enormous benefit.

Many others have also provided help along the way, including: Ramses Abul Naga, Chen Rong, David Coady, Frank Cowell, Angus Deaton, Peter Drysdale, Toni Hamiotis, Athar Hussain, Jenny Lanjouw, Peter Lanjouw, Stephen Ludlow, Stephen Pudney, Ren Yun, Sebastian Schich, Robin Sherbourne, Wang Limin and Zhu Liwei. Of these, I must single out my former office-mate, Peter Lanjouw. It was through collaboration with Peter in 1990 that I was first introduced to many of the issues pursued in Part One of this thesis. In the same year and through the same process, he taught me the ropes of SAS, the software package of which I have made extensive use throughout the thesis. Since those early days, discussions with both Peter and Jenny Lanjouw have been very beneficial in developing and extending my ideas.

I of course take full responsibility for any errors remaining in the thesis.

My study at the LSE was made possible by the financial support of the Association of Commonwealth Universities. While at LSE, I have been based at STICERD, the Suntory Toyota International Centre for Economics and Related Disciplines. I am extremely grateful to both institutions. In particular, numerous colleagues at STICERD have provided me with a great deal of useful advice and assistance.

Finally, I would like to thank my wife, Clare Holberton, for keeping me going.

Introduction

This thesis is divided into two parts. The first is concerned with measurement issues which arise in the analysis of income distribution, in particular with the problem, as I term it, of distributional dominance. The second is an analysis of the distribution of income in urban China between 1981 and 1990. The two parts complement each other: the ideas and tools developed in the first part assist in the analysis of the second, and the analysis of the second provides a case study of the methods presented in the first.

I Part One

Taking an overview of the burgeoning literature on the measurement of living standards, one can discern four strands of development over the last two decades. The first has been a proliferation of measures, especially for the purpose of equality and poverty analysis. Focus on the former came first, culminating in the derivation of the generalized entropy indices reported more or less contemporaneously by Bourguignon (1979), Cowell (1980) and Shorrocks (1980). Sen's 1976 paper was the spur for much work on poverty measures which took as its starting point the need to improve on the perceived inadequacies of the two traditional measures, the head-count ratio and the poverty gap. Chakravarty (1990) provides good summaries in both fields.[1] More recently, attention has turned to the statistical properties of these functions, with papers by Cowell (1989) and Thistle (1990) building on earlier contributions of authors such as Nygård and Sandström (1981).

The second strand begins with Atkinson (1970) and is the attempt to show under what conditions different measures (of, say, poverty or inequality) will rank two distributions in the same way. The popularity of this approach is linked to the proliferation of measures. The greater the numbers of plausible ways in which one might measure something, the more pressing the need to know whether one's conclusion is sensitive to the particular measure chosen. Atkinson introduced the fundamental idea that an income distribution can be analyzed in the same way as a probability distribution. This insight enabled him to draw on advances in the theory of risk by, among others, Hadar and Russell (1969) and Rothschild and Stiglitz (1970), who had shown how the criteria of stochastic dominance could be used to determine under which general conditions

1. Coulter (1989) provides a survey of inequality measures used in all the social science disciplines, not just economics.

one distribution of uncertain prospects would be preferred to another. Atkinson's work was in the area of equality analysis, where the second-order stochastic dominance criterion could be put in terms of the well-known Lorenz curve. His analysis has since been extended to both welfare and poverty analysis (Shorrocks, 1983, Atkinson, 1987) and, as we will examine in detail, has been applied using statistical tools. As Lambert writes, use of the second-order stochastic dominance criterion has now "become standard among researchers" (1989, p.5).

The third area of research has been a thorough re-working of the philosophical foundations of welfare economics. The hitherto-dominant utilitarian presupposition that welfare can be measured simply by adding a sum of utilities has come under serious criticism, most notably by Rawls (1971), but also by economists such as Sen (1987). Newer criticisms focusing both on utilitarianism's exclusive concern with utility (as against freedom (Nozick, 1974) or primary goods (Rawls, 1971) or capabilities (Sen, 1987)) as well as on its aggregation procedure (summation) have emerged alongside the older objection of the difficulty of making the interpersonal comparisons necessary to get utilitarianism off the ground.[2] For a review of the debate, see Sen and Williams (1982).

Finally, there has been a great deal of research into and thinking about which variable or set of variables should be utilized by economists to measure the standard of living. Techniques have long been available to enable the researcher to take into account differences in prices (see Deaton, 1980, for a survey). More recently, newer methods have been pioneered to take into account differences in family composition, both via the estimation of specific adjustment factors or 'equivalence scales' (Deaton and Meullbauer, 1980)[3] and by the extension of the stochastic dominance framework from a single- to a multi-variable framework (Atkinson and Bourguignon, 1987). Much work has also been done to show how it is possible to take into account such real-world departures from the perfect competition paradigm as rationing, fixed prices and the consumption of public goods (see Cornes, 1992, for a recent, critical survey). In addition, and in part stemming from the theoretical debate outlined above, it has been argued that other variables, not traditionally considered as being in the economist's domain, such as literacy, life-expectancy and infant-mortality (UNDP, 1990), and one's own perception of well-being (Goedhart *et al.*,

2. Interestingly, interpersonal comparability is no longer seen as such an obstacle: "For many years, the majority of economists took the position that the making of interpersonal comparisons, if not impossible, was certainly no part of the economist's trade. In view of Arrow's theorem, such a view leaves very little for welfare economics to do, and much of the so-called new welfare economics of the 1940's and 1950's that embodied this position makes sterile reading by contemporary standards. Modern approaches, by contrast, are firmly based on explicit interpersonal comparisons." (Deaton and Muellbauer, 1980, p.217)

3. Equivalence scales themselves have a long history dating back into the last century - see Deaton and Muellbauer (1980, p.193).

1977) should also be used to define the standard of living.

Part One of this thesis is concerned mainly with the second of these four strands. For convenience, the problem it addresses is given a label, that of 'distributional dominance'. Distributional dominance is defined in relation to two distributions, both defined over a single variable, and a so-called 'dominance criterion', each of which covers some set of living-standard functions. The problem of distributional dominance is the problem of ordering these two distributions in relation to the dominance criterion: if and only if the criterion ranks the two distributions can the one distribution be said to dominate the other, that is, be reckoned to be no worse than the other by all living-standard functions in the relevant set and better by at least one such function.

The term 'standard of living' is used in this thesis to mean 'welfare or equality or (inverse) poverty'. A standard-of-living function is thus a welfare or an equality or an inverse-poverty function (where an inverse-poverty function is the negative of a poverty function). The purpose of introducing this umbrella function is precisely to show how these three types of functions can be viewed as special cases. The aim here is not to suggest that welfare, poverty and inequality are three aspects of the one, necessarily amorphous concept. Rather it is to provide a unified framework and aid to understanding. Many of the problems which crop up within the one field of analysis are also present in the other two, and it is both efficient and illuminating to be able to deal with all three at the one time. Having first established this general framework, Chapter One goes on to introduce and where necessary derive the various dominance criteria which will be analyzed in the rest of Part One and put to work in the second part of the thesis.

The problem of choosing a dominance criterion cannot be viewed in isolation from the other decisions which have to be made when two distributions are to be compared. Above all, one has to choose by which variable or variables one will judge the standard of living. Adjustments may be required to take into account different price levels and household sizes in each of the two distributions. If one has raw data, one can decide on the degree to which one will aggregate (if at all) prior to analysis. If one has only an aggregated data set, one must decide how to utilize it and how to interpret the results. One must also decide whether or not to apply statistical tools or whether one will be content with drawing conclusions about the data to hand.

These last two issues, of statistical inference and aggregation, are the focus of the second and third chapters of Part One. Chapter Two presents and where necessary derives consistent estimators and their asymptotic variance-covariance matrices which, when combined with the general method of inference also given in the chapter, can be used to analyze distributional dominance as a statistical problem. These estimators have the important feature of being utilizable in the presence of randomly weighted data. So they can be used if, as typically occurs, one has data at the household level, but wishes to make living-standard comparisons using the individual as the unit of analysis.

Chapter Three reports the results of a simulation study which puts the estimators and inference method of Chapter Two to work. It provides evidence for a number of claims made in that chapter and also answers various questions arising from it. Chapter Three also takes up the issue of aggregation. Most empirical work on income distributions is based on aggregated data. The simulation study answers the question of what effect the use of aggregated data is likely to have on the nature of the resulting ordering, in particular on the probability of being able to rank any given pair of distributions.

The key restriction imposed by the theoretical analysis of the first three sections is that the standard of living is defined over only one variable. This restriction is in itself controversial. In addition, although the variable could be almost any standard-of-living determinant or indicator, Part Two applies the tools developed to the analysis of income data. And it must be said that the sort of variables to which the methods of Part One are most likely to be applied are income and consumption, simply because these are the variables on which distributional data are most likely to be available. But, as suggested earlier, reliance on standard purchasing-power variables is no longer uncontroversial. Sen has led the attack, in relation to both inequality and poverty measurement:

> An important and frequently encountered problem arises from concentrating on inequality of *incomes* as the primary focus of attention in the analysis of inequality. The extent of real inequality of opportunities that people face cannot be readily deduced from the magnitude of inequality of *incomes*, since what we can or cannot do, can or cannot achieve, do not depend just on our incomes but also on the variety of physical and social characteristics that affect our lives and make us what we are. (1992, p.28)
>
> Similarly, if our concern is with the failure of certain minimal capabilities because of lack of economic means, we cannot identify poverty simply as low income, dissociated from the interpersonally-variable connection between income and capability. It is in terms of capability that the adequacy of particular income levels has to be judged. (1992, p.112)

Providing this sort of general criticism with some specific content, Drèze and Sen write:

> ...many essential commodities are not bought and sold in the market place in the usual way, and conventional estimates of real income may not give us a good idea of the command over a number of inputs which, as we have seen, can play a crucial role in the removal of hunger, such as educational services, health care, clean water, or protection from infectious epidemics ... Income is a rather dubious indicator of the opportunity of being well-nourished and having nutrition-related capabilities. (1989, p.179)

Given that even the most advanced estimates of real income are unlikely to be able to capture the benefits of "clean water, or protection from infectious epidemics", what can one say in defence of reliance on conventional purchasing power data? One can only agree that if other information is available it should be used. But the same problems which arise when just one variable is being used for analysis, and which are analyzed in the next three chapters, will also occur when more than one variable is available. There are outstanding problems in relation to single-variable analysis and their solution is a necessary condition for progress on the more general front of dealing with a vector of living-standard-relevant variables. In addition, there are cases, of which Part Two's analysis of urban China is one, in which one does have a particular interest in the distributions of income and/or consumption not (or not just) because of the partial light they might shed on social well-being but because these are economic phenomena meriting analysis in their own right.

Assume then one does have data on a purchasing power variable, y, either consumption or income. The researcher may have no choice at all about how these data are defined, but, if he or she does, it will be to do with the choice of time-period for analysis or with the choice between consumption and income or with demographic issues, where these three are listed in ascending order of frequency. All the choices which can be made in response to these definitional questions are consistent with the general method of Part One - as is, to re-emphasize, the use of data other than purchasing power - but nevertheless the issues are important ones, and for completeness they are addressed briefly below.

The first choice is the time-period of measurement. Is y to be measured over a month, or a year, or even a longer period? Generally, the greater the possibilities for income-smoothing the longer the desirable period of measurement, but it is also very rare to have a single survey covering more than a year.

The second choice one is sometimes in a position to make is whether to define y as consumption (expenditure) or income. Ravallion writes that the development literature demonstrates "a preference for consumption as the welfare indicator", one reason for which is the "importance attached to specific forms of commodity deprivation, especially food insecurity." By contrast, "Ideas such as 'opportunities' and 'rights' seem to have carried relatively more weight in the developed country literature, particularly in Europe, and they are generally seen to indicate a preference for income" (1992, p.8). See Glewwe and van der Gaag (1988) for an illustration of the different conclusions which can follow from the choice of different measures.[4]

These first two issues are related. To quote Ravallion again, because of the possibility of engaging in consumption smoothing (lending and borrowing) "current consumption will almost certainly be a better indicator than current income of current standard of living, and ... current consumption may then also be a good indicator of long-term well-being, as it will reveal information about incomes at other dates, in the past and future." (1992, pp.13-14). This is, he writes, "probably the main reason" why consumption is preferred to income as an indicator of well-being.

The third and final issue concerns household demographics. It is probable that any available data will have been originally collected at the household level (though of course 'household' might be variably defined). One can then choose whether to work with the household as the unit for analysis or the individual, and, whatever the unit, whether to allocate to it total household y or per capita y or 'equivalent' y, that is, total household y divided by some denominator other than, and typically less than (to account for economies of scale and the existence of public goods), household size. Note that these two choices, relating to, respectively, 'weighting' and 'the treatment of household size' (and, possibly, other household characteristics such as age), to use the terminology of Atkinson and Micklewright (1992, pp.69-71), are distinct: any combination of them is feasible.

II Part Two

It is an irony that despite the massive amount of household data collected in China, probably more (even on a per capita basis) than in any other developing country, our understanding of the country's income distribution, especially of how it has changed over the course of the reform period, remains at a rudimentary level. The outstanding obstacle has been the non-availability rather than the non-collection of data. But much more data have become available in recent years, making progress possible.

---

4. The consumption-income distinction does of course become blurred once one starts considering imputed income.

In addition to the constraints of space and time, there are three reasons why this thesis focuses on *urban* China.[5] First, Chinese household income data are collected separately for rural and urban areas, and the quantity of urban data available for analysis greatly exceeds that of rural data.

Second, although large parts of China's economy are dominated by agriculture (with 70% of the workforce and 30% of total output), living standards in the urban areas of China in many ways resemble those of a semi-industrial country. Hence, for the purpose of making international comparisons, it is useful to deal with urban China separately. In particular, the experiences of urban China are of particular relevance if we wish to compare the effects of reform in China with the experience of the transitional, and largely industrial, economies of Eastern Europe and the former Soviet Union. Although the claim is often made that inequality will increase in an economy in transition from a planned to a market economy, evidence for this proposition has been scanty if not non-existent. China's experience of transition has of course differed in many ways from that of Eastern Europe and the former Soviet Union and there is no reason to expect that all transitional economies will experience the same distributional dynamics. Nevertheless, with the possible exceptions of Hungary and Poland, transition in China has been underway longer than in the other ex-centrally-planned economy, and an examination of urban China may help to inform our judgements on Eastern Europe and the Soviet Union.

Third, the policy implications of distributional analysis are more pressing in urban China. For it is in urban areas that the Chinese state has established such elements of the welfare state as subsidized housing and pensions. Reform in all of these welfare-related fields is currently under consideration in Chinese policy circles; in some cases, implementation of reform is already underway. To assess the distributional implications of reform the sort of analysis conducted in this paper is required.

The two chapters on China both use the same data source, the annual household surveys of China's State Statistical Bureau (SSB). Chapter Four uses nationwide data collected between 1981 and 1990. Chapter Five uses data for the years 1987 to 1990 for the urban areas of two of China's largest provinces (in terms of urban population), Liaoning and Sichuan. Chapter Four provides a broader picture, which covers both the entire decade of the eighties and all of urban China. Although the data set used by Chapter Five is only a subset of that on which Chapter Four is based, it contains detail on implicit subsidy income as well as cash income. As is typical for

5. The focus on the distribution of income rather than consumption is for reasons of data availability.

a socialist economy, cash is only one component of total income in urban China.

Another important difference between the two chapters is that Chapter Four is based on aggregated data, whereas Chapter Five uses disaggregated data. Seen as case-studies using the methods developed in Part One, Chapter Four is based on the findings of Chapter Three concerning aggregation, while Chapter Five uses the statistical tools developed in Chapters Two and Three. Both chapters use the dominance criteria presented in Chapter One.

The reliance of the thesis on one source - data collected by the SSB, an official body - might be seen as a source of concern. A critical assessment of the SSB surveys is presented in Chapter Four. It is argued that, although the surveys do have some defects and although there are gaps in our knowledge concerning survey procedure, they compare favourably to the East European surveys. Importantly, there is no other longitudinal data source for urban China of remotely comparable coverage or depth.

China's period of reform can be dated from the late seventies, though rural preceded urban reform. Right from the start, the official thrust of the entire reform programme has been anti-egalitarian. Of course, inequality has not been proposed as an end in itself, but an increase in inequality has been openly and enthusiastically sought by China's leadership as a necessary correction to the excessive levelling associated with the lurches to the extreme left preceding the reform period, and as essential stimulus to growth. As early as 1978, Deng Xiaopeng uttered his famous entreaty:

> We must permit some regions, some enterprises and some workers and peasants to have a greater income first and have better lives first as a result of their hard work and achievements.

With them as models to spur others on, he continued,

> ... the entire national economy will constantly move forward like a series of waves and the peoples of every nationality in China will then quickly become rich. (from Deng's speech 'Emancipate One's Thinking, Seek Truth from Facts, Unite and Look Forward' in Selected Works of Deng Xiaopeng; extracted in Guo, 1984)

In 1984, when the Party began putting greater emphasis on urban reform, then Premier Zhao Ziyang spelt out the implications for China's non-agricultural work-force

> The central theme of the present reform of the urban economic system is to thoroughly change the situation of no difference between good and bad economic management and no difference between workers who work a lot and those who work very little. We want to ensure that enterprises do not eat from the state's 'big rice bowl' and that workers do not eat from the enterprise's 'big rice bowl'... Letting some enterprises and some workers get rich first is the road we must take to smash these two 'big rice bowls'. (From Zhao's address to the Sixth National People's Congress, 1984; quoted in Guo, 1984.)

A large number of measures was introduced in the course of the urban reform process. These were characterized by decentralization, both to the firm and to lower levels of government, and by attempts to orient commercial activities towards profit-maximization. Contracts were introduced for managers of state-owned firms - under the 'contract responsibility system' - and the widening of the system of bonus payments allowed firms to link pay to profits. Yet the consensus now is that, like so much of China's urban reform and in contrast to the conspicuous success in the countryside, the attempt to trade off equality for efficiency has failed. General Secretary Jiang Zemin is representative:

> On the one hand, egalitarianism in distribution has not yet been completely overcome among wage earners in enterprises, public undertakings and party and government departments and has become even worse in some localities, departments and economic fields. On the other hand, new and unfairly wide gaps in social distribution have also emerged (quoted in Zhao, 1990a, p.34).

Such changes as have occurred are regarded as being for the worse. The comments of Zhao Renwei, one of China's most prominent economists, are typical in this regard.

> People have complained that 'self-employed street peddlers are the rich people and employees are the poor ones' and that 'the remunerations for atom bomb producers are not as good as those for peddlers of eggs boiled with tea' (1990a, p.36)

> Under the situation in which it has been difficult to increase incomes within the system of state plans or in which real incomes have fallen due to inflation, some units and individuals have tried in every way to get supplementary earnings from the activities of seeking rents, thus creating the 'grey incomes' on top of wages, bonuses, and normal business earnings. The formation and distribution of grey incomes are extremely irregular and are very much devoid of transparency, thus creating huge income disparities that have nothing to do with labour contributions... (1990a, pp.36-37)

But these changes due to the rising importance of 'black' and 'grey' income are seen as being at the margin. There has, it is claimed, been little systemic change in the distribution of income among the great mass of factory and white-collar workers. Economists, both Western and Chinese, have argued that overall urban inequality has been stagnant or indeed has fallen over the decade. Gale Johnson (1990) has argued that "the urban and industrial reforms did little to change the structure of compensation" (1990, p.76). Zhao himself has written of a 'new egalitarianism', arguing that

> The comprehensive wage reform in 1985 ... further narrowed wage disparities for urban staff members and workers, thus intensifying the tendency of egalitarianism. (1990a, p.34)

Others, such as this author from the Chinese-language Economic Daily (equivalent to the British Financial Times), are even more sweeping:

> After reform was carried out for 10 years, [the] problems [of an egalitarian distribution] remained unsettled and even worsened... [T]he egalitarian tendency in people's income became more obvious... [T]he income distribution system basically remains unchanged. (Jia, 1990, pp.32-3)

Though such claims seem sufficiently numerous to be persuasive, many of them are based on at most fragmentary evidence. Chapter Four carries out a systematic review of the evidence on the urban distribution of cash income available through the SSB's household survey and emerges with a conclusion which challenges the consensus described above.

As is typical for a socialist economy, cash is only one of a number of important sources of income in urban China. There are two main types of non-cash sources of income for the urban resident. First, there are subsidies and in-kind benefits provided by China's enterprises. Most prominent here is housing. Houses are owned by enterprises and rents are heavily subsidized, so much so that rental outlays come to only 1% of total consumer expenditure. Education and medical care are also firm-based and are provided free or at nominal cost. The second major source of non-cash income is from subsidies provided by the government to urban residents. Most prominent here have been food subsidies, though a number of facilities, such as power, water transport and postal services, are also subsidized.

Attempts have been made to estimate the total income of urban residents with attention to these numerous subsidies and in-kind benefits. As indicated in Chapter Four, such estimates typically show that non-cash income comes to around half or more of cash income. However, the data requirements for such estimates are forbidding and often there is recourse to a number of fairly arbitrary assumptions. The final chapter of the thesis - as well as re-examining the trends in cash income - focuses on just one source of non-cash income, food subsidies. Although food subsidies became more important in the late eighties as, in an inflationary environment, free-market prices rose rapidly, in the mid-eighties and again in the early nineties the Chinese government sought to reduce their importance by allowing state-prices to rise, and issuing cash compensation. By 1990, these compensation payments had become worth some 10% of total cash income. The effects of this important but neglected transitional policy are also investigated and their policy implications considered.

China is a massive country, and income distribution a wide-ranging subject. Any treatment of the two must perforce be selective. This study should be seen as only one contribution to a wider debate. Topics of importance not addressed herein include changes in remuneration by occupation (it is widely believed that government-determined wage rates do not adequately reflect skill differentials between occupations) and the alleged growth of corruption, fringe benefits and rent-seeking, especially among the political elite.

PART ONE DISTRIBUTIONAL DOMINANCE

Chapter One A Unified Framework and Some New Criteria

I Introduction

Many aspects of economics require comparisons of distributions. If one is examining the distributional impact of different policies, judgements are required about what the distributions of interest would be like under the respective policies. Or one might be interested, as a social scientist, in the distribution of income under different economic arrangements. But agreement on a single criterion by which to judge, or order, the distributions of interest may not be possible. One way around this is to examine whether one distribution does better than another by a number of objective functions. The problem of distributional dominance is precisely the problem of ordering two distributions in relation to one or more objective functions, summarized into a single dominance criterion. These distributions could be of many different sorts of variables: for example, distributions of possible returns on risky assets, of market share or of living-standards-relevant variables such as income. Although what follows in this chapter, and indeed in the rest of Part One, is likely to be of relevance whatever the distributions, the focus throughout is on the latter category of variables. Hence the dominance criteria will be used in relation to sets, Σ, of living-standard functions, S. If and only if a criterion ranks two distributions is the one distribution said to dominate the other, that is, reckoned to be no worse by all S in the relevant Σ and better by at least one S.

A 'living-standard function' sounds, it must be said, like an unlikely beast. But the adjective 'living-standard' is simply a label, used to indicate that the function is one which can be used to order distributions on the basis of either welfare or inequality or poverty. One of the aims of this chapter is to show how all three of these can be analyzed within the one framework, not only if the researcher wants to use a specific functional form, but also more generally in relation to this question of distributional dominance. The argument here is not that welfare, equality and poverty measures can somehow be aggregated to give a summary measure of a society's well-being. This would be a totally misdirected endeavour: welfare, poverty and equality measure different things. Rather the argument is that viewing these three families of functions as special cases of a more general function provides a useful framework and aid to understanding. Particularly in relation to the problem of distributional dominance, many of the same problems crop up whichever one of the three is being dealt with. These problems require common solutions and common solutions demand a common framework.

All this raises the obvious question of what is meant by the words 'welfare', 'poverty', and 'inequality'. I take it as read that there are fairly clear, everyday meanings attached to these words. When we measure a society's welfare, we are aggregating over the well-being, however measured, of all the members of that society. When we measure a society's poverty, we examine the proportion and/or well-being of those below some poverty line. And when we look at inequality, we examine the dispersion of income, or some other measure of well-being, abstracting from differences in means. The more precise definitional assumptions given in the next section aim to be in accord with these everyday meanings and to show, as simply as possible, in what respects poverty, welfare and inequality functions are similar and in what respects they differ.

An initial obstacle to dealing with welfare, inequality and poverty in the same breath is that the former is a good, whereas the latter two are bads: we want higher welfare, but less inequality and less poverty. To negotiate this hazard, the formal results of the chapter deal not with inequality but with *equality* functions, and not with poverty but with *opulence* or *inverse-poverty* functions. An equality function is simply the negative of an inequality function, and an opulence or inverse-poverty function the negative of a poverty function. This eases exposition, without loss of generality. (In the more informal discussions, I will still often use the more natural terms 'inequality' and 'poverty'.)

In fulfilling this aim of providing a general framework, the chapter necessarily takes on the role of a literature survey. However, it also contains a number of original contributions. These are highlighted below in a summary of the chapter's structure.

Section II presents the properties the living-standard function is presumed to possess and shows how in special cases it can be thought of as a welfare, equality or one of two types of opulence functions. Various functional forms are presented to provide examples of the similarities and differences between the three families of functions. In addition, a new function is introduced, which generalizes the Clark *et al.* (1981) index and which can be parameterized to be either of the two types of opulence functions.

Second-order stochastic dominance is the best known of the dominance criteria. It is introduced in Section III alongside first-order stochastic dominance. Also in this section, the results of Atkinson (1987) pertaining to poverty second-order stochastic dominance are generalized. The resulting criterion, labelled 'mixed stochastic dominance', covers all well-known poverty functions (including both types introduced in Section II), and thus can be regarded as playing a similar role *vis-a-vis* poverty functions as second-order stochastic dominance does *vis-a-vis* welfare and equality functions. Finally, the pros and cons of using deficit and Lorenz curves for analyzing second-order stochastic dominance are considered.

In Section IV, alternatives to second-order (and mixed) stochastic dominance are considered. These are divided into two types: those requiring 'extreme' forms of dominance and others. Under the latter heading comes the 'restricted dominance' category - a new class of dominance criteria based on a generalization of Atkinson and Bourguignon (1989) - and the criterion of 'e-dominance', based on the isoelastic function.

II The standard of living: welfare, equality and poverty

II.1 provides the notation, assumptions and definitions which will be used throughout the chapter. II.2 defines the three types of functions, and argues that they are based on the everyday meanings we attach to the words 'welfare', 'equality' and 'poverty'. II.3 examines a number of the definitional assumptions made in greater detail. II.4 illustrates the discussion with a variety of functional forms and introduces the new 'generalized Clark' opulence function.

II.1 Notation, assumptions and definitions

S is in general a standard-of-living function. It is an aggregate function: that is, it measures a society's overall standard-of-living. S is defined over the distribution function of a variable y, which itself is defined over the set of real numbers. I will for convenience label y 'income' - or, importantly, some transformation of income - but here income is to be understood loosely to represent almost any variable of relevance to the standard-of-living. It must of course be a variable representable by the real number system. It is also assumed that y can take on a continuum of values. How y should in fact be defined is not discussed in detail, although some comments on the applicability of the framework developed for different y are discussed at the end of II.3. Questions such as whether, if y is a purchasing-power variable, it should be income or consumption, the relevant time-period over which these should be measured (a year or a life-time?), the definition of the recipient unit (an individual or a household?) and how differences in household size and composition should be controlled for have been briefly considered in the

Introduction (Section II thereof), but are left completely open insofar as this chapter is concerned.[1]

Consider a pair of distributions, which are defined over y and denoted by their distribution functions, or cumulatives, F and F^*. All F and F^* in $\mathcal{F}$ are non-decreasing and right-continuous, and bounded by zero and one. Let p=F(y) so that p is the proportion with income less than or equal to y. F and F^* may be continuous, discrete or mixed distributions. They are assumed to have finite means and variances. Moreover, let $\eta=\sup\{y:F(y)=0\}$ and $\theta=\inf\{y:F(y)=1\}$, similarly for η^* and θ^*. Let $\eta^-=\min(\eta,\eta^*)$ and $\theta^+=\max(\theta,\theta^*)$.[2]

As stated earlier, use of a dominance criterion generates an ordering over the set of distributions. There are three possible outcomes. Either

(i) $S(F)\geq S(F^*)\ \forall S\in\Sigma$ and $S(F)>S(F^*)\ \exists S\in\Sigma$

or (ii) $S(F)\leq S(F^*)\ \forall S\in\Sigma$ and $S(F)<S(F^*)\ \exists S\in\Sigma$

or (iii) $S(F)>S(F^*)\ \exists S\in\Sigma$ and $S(F)<S(F^*)\ \exists S\in\Sigma$ or $S(F)=S(F^*)\ \forall S\in\Sigma$.

In the first case F dominates (FDF^*), in the second case F^* dominates and in the third case neither distribution dominates. The result of a pairwise evaluation of all pairs of distributions in a set of size at least two will be a strict **partial ordering** in D (Sen, 1970). Such an ordering has the characteristics of being (i) partial - for each pair, either FDF^* or F^*DF or neither; (ii) transitive - if FDF^* and F^*DF^{**} then FDF^{**}; and (iii) asymmetric - if FDF^* then not the case that F^*DF. If for any pair FDF^* or F^*DF then a **ranking** is said to be achieved. Using this terminology, a strict partial ordering consists of rankings and non-rankings.

The key assumptions made on S also require definition. The assumptions are only defined in this sub-section. They will be discussed at length later in the section. Assumptions 2 and 3 compare distribution functions which are, at least over some range, step functions. But this does not mean they are not useful for comparing continuous distributions. Rather, as we will see, the restrictions in relation to step-functions imply more general restrictions in relation to all distributions.

1 The function $S:\mathcal{F}\rightarrow\mathbf{R}$ indicates the standard-of-living associated with any distribution F where F is in $\mathcal{F}$ and so is defined over y where

1A **y is income**

or 1B **y is mean-normalized income** (income divided by the mean).

1. The argument that S should be defined over a vector of variables, rather than a single variable as here, is also considered in Section I of the Introduction.

2. The assumption that F has a finite maximum and minimum eases the exposition. The modifications to theorems required if this assumption does not hold are indicated by footnotes 17 and 36.

2 S is weakly increasing in y

Let $F(y)-F^*(y)=c>0$ for $y_1 \leq y < y_2$ and $F^*(y)=F(y)$ $y<y_1$ and $y \geq y_2$. Then, for all F, $S(F^*) \geq S(F)$.

3A S satisfies the transfer principle

Let the mean of F (F^*) be μ (μ^*). Let $\mu=\mu^*$, $F(y)-F^*(y)=c>0$ $y_1 \leq y < y_2$, $F(y)-F^*(y)=d<0$ $y_3 \leq y < y_4$, where $y_2 \leq y_3$, and $F^*(y)=F(y)$ $y<y_1$, $y=y_2$, $y_2<y<y_3$ (if $y_3 \neq y_2$) and $y \geq y_4$. Then, for all F, $S(F^*) \geq S(F)$. Note that F is generated from F^* by a single *mean-preserving spread* (Rothschild and Stiglitz, 1970) or rank-preserving regressive transfer (if the distributions have disrete members).[3]

3B S satisfies the transfer principle except possibly at Z^P

Define F^* and F as in 3A. If in addition it holds that $F(Z^P)=F^*(Z^P)$ then $S(F^*) \geq S(F)$.

4A S is insensitive to changes in y at or above Z^P

If $F(y)=F^*(y)$, $y<Z^P$, then, for all F, $S(F)=S(F^*)$.

4B S is insensitive to changes in y above Z^P

If $F(y)=F^*(y)$, $y \leq Z^P$, then, for all F, $S(F)=S(F^*)$.

II.2 General framework

We are now in a position to define welfare, equality and opulence functions as special types of S functions. The different combinations of assumptions used are summarized in Table 1 below.

1. If S satisfies assumptions 1A, 2 and 3A, S is a **welfare** function.
2. If S satisfies assumptions 1B and 3A, S is an **equality** function.
3. If S satisfies assumptions 1A, 2, 3B and 4B, S is an **opulence** function. (Note that each opulence function is defined for a particular poverty line, Z^P.)

3. See Rothschild and Stiglitz (1970, pp.230-231), especially Figure 5 (which should be labelled Figure 6 to be consistent with the text). Rothschild and Stiglitz also have another definition of a mean-preserving spread which if F^* is continuous keeps F continuous (see their Figure 6 - which should be labelled Figure 5). However, this definition is, as they show, not required even for the comparison of continuous distributions, since these can be "approximated arbitrarily closely by step functions" (p.232). See also footnotes 19 and 20.

Table 1 Types of Living-standard Functions

Type of living standard function	Assumptions made concerning...			
	Definition of y	Function being weakly increasing in y	Satisfaction of transfer principle	Function being insensitive to changes in y above some level of y
Welfare	Income (1A)	Yes (2)	Yes (3A)	Not in general
Equality	Mean-normalized income (1B)	Not applicable	Yes (3A)	Not in general
Opulence	Income (1A)	Yes (2)	Yes except possibly for crossings of Z^P (3B)	Yes for $y>Z^P$ (4B).

Note: Numbers in parentheses refer to number given to assumption in text.

Why these assumptions? Note first that, through either 1A or 1B, all three types of functions are defined over F. This has two important implications. First, if instead of agent A having y_1 and agent B y_2, B had y_1 and A y_2, F and thus S would be unchanged. Hence S satisfies the uncontroversial assumption of anonymity. Second, S is based only on proportions with y less than or equal to different levels, not on total numbers. This assumption - one of so-called 'replication invariance' - enables us to abstract from differences in population and to compare continuous distributions (which do not have a finite population size). Whether or not one should abstract from population size when making standard-of-living comparisons is a matter of debate, but not one entered into here.

Turning to the assumptions which differentiate the three families of functions from each other, the first major distinction is that welfare and opulence are defined over income, in which they are increasing, whereas equality functions are defined over mean-normalized income, in which they are not restricted to being increasing. (Note that here, and throughout, when S is said to defined over some y, this is short-hand for saying that S is defined over F which is the distribution function of that y). This is consistent with the common-sense view that richer people are better off. It also accords with the everyday meaning attached to 'equality'. If my income goes up, and I am rich, my mean-normalized income will rise but inequality should also rise, so equality functions cannot be assumed to be increasing. More fundamentally, assumption 2 couldn't be applied to y as mean-normalized income even if we wanted it to since, if F is defined over mean-normalized income, the integral of F is always of constant size (equivalently, if F has discrete members, it is impossible to change the y of only one member of F).

What marks out opulence functions is assumption 4B (for the distinction between 4A and 4B see below). Whereas welfare and equality functions are based on the complete distribution of y, opulence functions are independent of the distribution above the poverty line. Again this accords with our everyday usage. The rich getting richer doesn't reduce poverty.

If 1B marks out equality functions, and 4B opulence functions, assumption 3 (A or B), relating to the transfer principle, plays a key role in characterizing all three types of living-standard functions. The transfer principle was introduced by Dalton (1920) and its centrality is evident from even the most cursory glance at the literature. Assuming that S satisfies 3A rules out 'positively anti-egalitarian' functions (Sen, 1973, p.64) by making it impossible for transfers from the poorer to the richer to improve aggregate living-standards. Equality and welfare functions both satisfy 3A and opulence functions satisfy the weaker 3B. Assuming 3B has the same implications as assuming 3A, with the possible exception of transfers across the poverty line. To distinguish between them, I will call any function satisfying 3A an *egalitarian* function, and any function satisfying 3B an *almost-egalitarian* function, since the latter satisfies all the properties of the former except possibly where crossings of the poverty line are concerned.[4]

Just as the distinction between 3A and 3B is obviously only relevant to opulence functions so too is that between 4B, which embodies a weak definition of the poor, and 4A which embodies a strong. And just as any function satisfying 3A automatically satisfies 3B so any function satisfying 4A automatically satisfies 4B. This enables us to define a sub-class of egalitarian opulence functions which satisfy (in addition to 1A and 2) not just 3B and 4B but also 3A and 4A. II.3.4 examines the relative merits of egalitarian and almost egalitarian opulence functions, and explains the pairing of 3A and 4A.

Finally, note that all egalitarian opulence functions are also welfare functions, since the former satisfy all the assumptions of the latter. To relate the orderings of equality and welfare functions, we need to restrict attention to those S which are unit-invariant.[5] Then any comparison of two distributions with the same mean using an egalitarian welfare function will give the same

4. Egalitarian functions are S-concave functions: see, for example, Sen (1973).

5. This is not to be confused with the stronger assumption of scale invariance, which requires the ordering generated by S over a pair of distributions to be invariant to the multiplication of the first distribution by one positive factor, and the second distribution by another. 1B invokes scale invariance, but IB is only taken to be an attribute of equality functions.

ordering as the same function if transformed into an equality function by defining it over mean-normalized income. In this sense, equality analysis is welfare analysis applied to distributions with the same means.

The assumptions utilized in this section are by no means the only ones available by which welfare, equality and opulence functions can be defined. However, the above discussion does indicate, it is hoped, that their choice is consistent with the everyday meanings attached to the words 'welfare', 'equality' and 'poverty'.

II.3 A closer look at the assumptions

II.3.1 The definition of y

Since it is assumed that y can take on a continuum of values, use of the framework cannot be made for analysis of zero-one variables such as literacy, to which, for example, the transfer principle (which requires that y can take on at least three values) cannot be applied.

Although 3A and 3B talk about y being transferred, from a formal perspective it is a matter of indifference whether y actually is transferrable. For saying that F^* is generated from F is simply shorthand for saying that, if y were the sort of variable that could be transferred, then F^* could be generated from F. Hence the framework can be applied to variables such as life-expectancy, which we do not normally think of as being transferrable.

The objection might be raised that if y cannot in fact be transferred, any judgements of distributions by use of a transfer principle, though possible, are of no relevance. But this is to claim that hypothetical "what if?"s have no persuasiveness. Surely it is plausible that if life-expectancy, for example, could be transferred we would, from an impartial perspective, prefer a society in which it was evenly spread? On the other hand, it might be argued that questions of rights, not addressed within the framework, should figure more strongly for less appropriable y's than for more. To the extent that this position is held, one can only say that, while the formal applicability of the framework developed is broad, its appeal will vary with different definitions of y.

II.3.2 Separability

S need not be restricted to being in the class of additively separable functions. If it is, then the standard-of-living in a distribution can be defined to be

$$S(F) = \int_{\eta}^{\theta} s(y)dF(y) \tag{1}$$

where s is an 'individual' standard-of-living function and S the mean of these individual levels. Assuming that S is separable simplifies analysis. All the assumptions pertaining to it can then be simply put in terms of s, especially if the latter is also assumed differentiable. In this case, 2 (combined with 1A) becomes the requirement of a non-negative first derivative on s; 3A becomes that of a non-positive second derivative (for 3B, a non-positive second derivative except at Z^P); and 4 can be put as the requirement that s(y) is constant above Z^P. However, the assumption of separability is by no means uncontroversial, especially for equality measures. Broome (1989) goes as far as to say that any function which is separable cannot be an acceptable measure of equality. Foster and Shorrocks (1987) argue the reverse in relation to poverty functions.[6] The issues are too deep to go into here. Separability simplifies analysis and allows for various decompositions by population sub-groups to be made. On the other hand, non-separability may capture better social inter-dependence, by allowing my welfare, for example, to depend on your as well as my income. It suffices for our purposes to say that, due to the controversy surrounding it, separability is not invoked here as a primitive assumption, though it is imposed at various points.

II.3.3 The mean-normalization of income: assumption IB

As already indicated, analysis of inequality requires abstraction from differences in mean. Working with mean-normalized income is a natural way of doing this and several authors have introduced it as a primitive axiom (see, for example, Cowell and Kuga, 1981 and Lambert, 1989). An alternative response is to seek to give equality functions a basis in welfare analysis. From this perspective one could argue that we should abstract from differences in means by dividing by the mean not income itself, but equally distributed equivalent income (EDE), introduced by Atkinson (1970) and defined as the amount of income which if had by everyone would give the actual aggregate welfare level. One could then only justify analyzing equality with mean-normalized income if it gave the same ranking as an analysis based on EDE divided by the

6. In fact, Foster and Shorrocks argue for sub-group consistency, but this implies separability.

mean. A ranking based on mean-normalized income is of course independent of the actual mean level. If we take advantage of separability, the only EDE-mean ratio which is independent of the mean is that in which EDE is measured using the isoelastic function (see Table 2 below for a definition). So from this welfare-based perspective and assuming separability, the restriction to the class of relative equality measures is consistent only with S being a monotonic transform of the isoelastic function.[7]

It is also possible to argue that one should abstract from differences in mean not by dividing either actual or EDE income by the mean, but by subtracting the mean from either of these two. This is a method pioneered by Kolm (1976) - see Bishop, Formby and Thistle (1989) for a recent summary and references. Again the possibility can only be raised here. Certainly, working with differences from rather than ratios of the mean introduces the extra problem of choosing a unit of measure. Ratios of dollars are also ratios of cents, but not so for differences. But the real issue lies deeper. If one thinks inequality is unchanged by adding one pound to or subtracting one pound from everyone's income, then one wants to work with differences. If one thinks inequality is unchanged by adding 1% or subtracting 1% from everyone's income, then one wants to work, as here, with ratios.[8]

II.3.4 Egalitarian and almost-egalitarian opulence functions: assumptions 3 and 4

All opulence functions satisfy 3B and 4B. The egalitarian sub-class satisfies in addition 3A and 4A. 3B treats the poverty line as a potential threshold, able to have a discrete impact on well-being. 3A rules this out and forces one to regard poverty entirely as a matter of degree. If the opulence function satisfies 3A, it can never be increased by a regressive transfer. If it satisfies 3B only, it can be if the regressive transfer reduces the number poor. Formally,

Theorem 1 If S satisfies 2, 3B and 4B then the only mean-preserving spreads which can increase S are those which reduce the proportion with income less than or equal to Z^P.

7. "Now suppose that we were to require that the equally distributed measure I [1-EDE/μ] were invariant with respect to ... proportional shifts, so that we could consider the degree of inequality independently of the mean-level of incomes....[T]his requirement implies that [s(y)] has the form: [isoelastic]" (Atkinson, 1970, republished in 1983, p.21).

8. Frameworks can also be developed to support the view that inequality can be changed both by proportionally and by absolutely equal shifts in income. Again see Kolm (1976, Section III).

Let F be generated from F^* by a mean-preserving spread. There are three possible cases. First, $F(Z^P)=F^*(Z^P)$. Then the mean-preserving spread cannot increase S by 3B. Second, $F(Z^P)>F^*(Z^P)$. Then 3B cannot be appealed to. But it must be the case that $y_1 \leq Z^P < y_2$. The change from F^* to F at and above y_3 can be ignored by 4B. The change below y_2 cannot increase S by 2. Third, $F(Z^P)<F^*(Z^P)$. Then $y_3 \leq Z^P < y_4$. Now the change from F^* to F at and above y_3 cannot be ignored (4B cannot be invoked) and S may rise or fall.■

Whether 3A or 3B is the more appropriate assumption is a matter for debate (on which see especially Sen, 1982, pp.32-33 and Atkinson, 1987, p.759). On the one hand, it may be held objectionable to allow for the possibility of regressive transfers increasing S. On the other, the very suggestion that there is a poverty line does seem to carry with it the idea of a threshold, and the existence of a strong asymmetry between being just below or just above the poverty line. As Sen writes, an assumption such as 3A

> ... takes no note whatever of the poverty line, and while that is quite legitimate for a general measure of economic inequality for the *whole* community, it is arguable that this is not so for a measure of poverty as such. (1982, p.33)[9]

At the very least, 3B is an assumption which should be allowed for when building a general framework. The three best known opulence functions - the negatives of the headcount ratio, the Sen index and the poverty gap - all satisfy 3B but only one, the negative of the poverty gap, satisfies 3A.

The relative merits of 4A and 4B are less clear-cut. 4B, incorporating a weak definition of the poor, goes naturally with the weaker assumption 3B, since the (negative of) the head-count ratio, $F(Z^P)$ is a well-known almost-egalitarian opulence function. 4B could also be used with 3A. In fact though, all well-known opulence functions which satisfy 3A also satisfy 4A (see Table 2). This is not a coincidence. The well-known egalitarian opulence functions are all separable, and one can easily show that if a function satisfies 2, 3A and 4B and is separable, then it must satisfy 4A (since 3A and separability imply concavity which implies continuity). This need not necessarily be so for non-separable functions, though it must be the case for functions which satisfy 3A where the weak are replaced by strong inequalities (i.e., where progressive transfers must increase S not just not decrease it) - see Donaldson and Weymark (1986) for a proof. Since, as we will see (in III.1.2), assuming 4A together with 3A simplifies matters and since it seems to cause very little

9. Sen (1982, pp.32-33) considers a strong and a weak transfer axiom. If we replace his 'increase' by 'not decrease', then 3A corresponds to his strong transfer axiom and, by Theorem 1, the combination of 3B with 1A and 2 to his weak transfer axiom.

loss of generality, the pairing is made.

II.4 Some illustrations and a new poverty index

To make the above discussion concrete and to demonstrate the advantages of thinking of welfare, poverty and equality measures within a single framework, Table 2 below gives a number of functions for all three of the classes of standard-of-living functions. Many of the equality and opulence functions are of course better known as inequality and poverty indices. To recover them in their better-known forms, one need only reverse signs (and sometimes take a monotonic transform). The table is illustrative rather than comprehensive: Chakravarty (1990) is probably the most recent and comprehensive survey of the various living-standard functions available.

Table 2 begins with four classes of functions satisfying 3A. They are all separable (so that s(y) rather than S(F) is given) and weakly concave. The first is the isoelastic class, which has already been commented on. It expresses welfare as the product of two terms, one dependent on equi-proportionate shifts in income, the other, measuring the welfare cost of inequality, independent. This assumption, that the welfare cost of inequality is invariant to the level of mean income, is one of constant relative inequality aversion, and indeed the isoelastic class is often labelled accordingly. The degree of relative inequality aversion is given by negative of the elasticity of the first derivative, which is $e=1-\alpha$. Atkinson (1970) applied this function to inequality analysis and based what has come to be known as the Atkinson Index upon it. Clark *et al.* (1981) were the first to apply the function to poverty analysis.

Table 2 Some Welfare, Equality and Opulence Functions

	TYPES OF LIVING-STANDARD FUNCTIONS		
	Welfare	**Equality**	**Opulence**
I. SEPARABLE, WEAKLY CONCAVE FUNCTIONS			
(i) Isoelastic, $\alpha \leq 1$			
$\alpha \neq 0$	$(1/\alpha)[y^{\alpha}]$	$(1/\alpha)[y^{\alpha}]$	$(1/\alpha)[(y/Z^P)^{\alpha}-1]$ (Clark *et al.* index)
$\alpha=0$	$\ln(y)$	$\ln(y)$	$\ln(y/Z^P)$ (Watts' measure)
$\alpha=1$	y (income-utilitarianism)	y (inequality-indifference)	(y/Z^P-1) (poverty gap)
$\alpha \to -\infty$	η	η	$\min(\eta/Z^P-1,0)$
(ii) Generalized entropy (GE)			
$\delta \neq 0$	$1/[\delta(1-\delta)](y^{\delta}-1)$ $(\delta<1)$	$1/[\delta(1-\delta)]*(y^{\delta}-1)$ $(\delta \neq 1)$	$1/[\delta(1-\delta)]*((y/Z^P)^{\delta}-1)$ $(\delta<1)$
$\delta=1$	****	$-y\ln(y)$ (Theil's measure)	****
$\delta=2$	****	$(1/2)*(1-y^2)$ (Coefficient of variation)	****
$\delta \to \infty$	****	$-\theta$	****
(iii) 'Generalized relative deviation'			
$\tau \geq 1$	****	$-\lvert 1-y \rvert^{\tau}$	$-(1-y/Z^P)^{\tau}$ (Foster *et. al* index)
(iv) Exponential			
$r>0$	$-\exp(-ry)$	$-\exp(-ry)$	$-\exp(r(1-y/Z^P))+1$
II. OTHER FUNCTIONS			
Gini-based	****	$-1-(1/N)-\{(2/N^2)*\Sigma_{i=1}^{N}[(y_i)*(N+1-i)]\}$ (Gini)	$-2/[(q+1)N]*\Sigma_{i=1}^{q}[(1-y_i/Z^P)*(q+1-i)]$ (Sen index)
Headcount	****	****	-1
Generalized Clark	****	****	$(1-C)/\alpha[(y/Z^P)^{\alpha}-1]-C$ $\alpha \leq 1$, $\alpha \neq 0$, $1 \geq C \geq 0$

Notes: 1. From assumptions 1A and 1B, y (and hence η and θ) is income in the case of the welfare and opulence functions, and mean-normalized income in the case of the equality functions.

2. For all separable functions, $s(y)$ is presented. For the case of opulence functions, $s(y)$ is presented for $y \leq Z^P$. For $y>Z^P$, $s(y)=0$. For the non-separable Gini, it is assumed that the distribution can be written as a vector of incomes: N refers to the population size, q to the number with $y \leq Z^P$. The GE function with $\delta=0$ is exactly the isoelastic function with $\alpha=0$. For the generalized Clark function with $\alpha=0$, replace $1/\alpha[(y/Z^P)^{\alpha}-1]$ by $\ln(y/Z^P)$.

3. For the isoelastic, GE and generalized Clark functions, it is assumed that $\eta>0$ and $Z^P>0$. It is also assumed for the extreme case in which δ (α) approaches ∞ ($-\infty$) that $\theta \neq \theta^*$ ($\eta \neq \eta^*$).

4. Many of the equality (opulence) functions are better known as inequality (poverty) functions. Their sign should be reversed to convert them to the more familiar form. Other monotonic transforms may also be required. On the coefficient of variation, see Kakwani (1980a, p.81). On the Sen Index and Gini coefficient, see Sen (1982), p.379. For the Watts' measure, see Watts (1968). Other references are given in the text.

The generalized entropy (GE) measure has been axiomatized and advocated for use in inequality analysis by various authors (see Bourguignon, 1979, Cowell, 1980, and Shorrocks, 1980).[10] The GE class is similar to the isoelastic. It displays constant relative inequality aversion equal to 1-δ, except for the single case of δ=1. The GE class applied to equality almost entirely subsumes the isoelastic. The two will give the same ranking if $\delta=\alpha$ as long as $\alpha<1$. Only the ranking obtained by the isoelastic function with $\alpha=1$ cannot in general be attained by the GE function,[11] though none of the GE rankings obtained with $\delta\geq1$ can in general be replicated by the isoelastic function. However, in the case of welfare and poverty analysis, δ must be restricted to being strictly less than one, to ensure the functions are non-decreasing in income, so, in these cases, the isoelastic class (just) subsumes the GE class. Note too that if (as discussed earlier) one demands a welfare basis for one's measurement of equality then one should not use equality functions which have $s'(y)<0$ as this would imply that the underlying welfare function is decreasing in income. This would rule out the use of GE equality functions with $\delta\geq1$. It is also interesting to examine the third derivative of the GE function. If $\delta>2$ then the GE measure has a negative third derivative, which makes it most sensitive to transfers at the upper end of the income scale. For the isoelastic family, the third derivative is always positive, making the measures 'transfer sensitive', that is, increasingly sensitive to transfers as one moves down the income scale. The GE measure is in this regard more flexible. One can choose δ to give different weights to differences in different parts of the distribution, whereas α in the isoelastic function always gives at least as much weight to differences in the lower as in the upper tail.

The 'generalized relative deviation' class of functions generalizes in equality analysis the well-known relative mean deviation, which can be obtained from the more general function by setting $\tau=1$. In poverty analysis, it is known as the Foster index after Foster *et al.* (1984).[12]

Finally, there is the exponential class, with constant absolute-inequality aversion, that is, a constant ratio of the second to first derivative. This class illustrates a different type of inequality aversion. The exponential function can be written to express welfare as the product of two terms, one dependent on any equal absolute additions to income, the other, measuring the welfare cost of inequality, independent. This assumption, that the welfare cost of inequality is

---

10. Its extension to the analysis of poverty has been discussed by Cowell (1988) (though Cowell allows for a much more general form of the poverty function than presented here).

11. Although, to save space, Table 2 restricts $\delta\neq0$, as remarked in the table notes, if $\delta=0$ then the GE function is exactly the isoelastic function with $\alpha=0$.

12. Note that setting $\tau=1$ gives the (negative of the) poverty gap measure. Setting $\tau=0$ gives the (negative of the) headcount ratio.

invariant to equal absolute additions to income, is one of constant absolute inequality aversion, the degree of which is given by negative of the ratio of second to first derivative.

These assumptions on inequality aversion can obviously be applied to welfare or poverty functions. If inequality indices are also to be thought of in the framework of inequality aversion, they should regarded as measures of the welfare cost of inequality for a given type and degree of inequality aversion (Sen, 1982, Chapter 19). There is no clear-cut consensus in the literature as to which inequality aversion assumption is preferable. One need not even be restricted to functions displaying constant inequality aversion, whether relative or absolute. Decreasing and increasing versions of either type of aversion are also of course possible. The issues involved relate to the earlier discussion on mean-normalized income. At the risk of being repetitive, one can summarize as follows: if one wishes to give inequality measurement a welfare basis and thinks that differences in mean should be abstracted from by dividing income by (subtracting income from) the mean, one should normalize appropriately and base one's equality functions on welfare functions displaying constant relative (absolute) inequality aversion.[13] For further discussion, see Atkinson (1970) - on which the above has drawn - for an introduction, and Jewitt (1981) for an advanced treatment. The literature on risk, from which this terminology and classification by 'aversion' stems, can also be consulted: see Deaton and Muellbauer (1981), Chapter 14, for example.

Of the remaining functions in Table 2, the well-known Gini is an example of a non-separable function. It is best known in the domain of equality measurement, but has also been used for poverty measurement, via the Sen index. As mentioned, the negatives of both the Sen index and the headcount ratio are examples of opulence functions which satisfy 3B but not 3A.[14]

The last opulence measure given, which gives the negative of what I call the 'generalized Clark' (since the (negative of the) Clark *et al.* function can be attained from it by setting C=0) represents a simple illustration of the sort of use that can be made, in poverty analysis, of orthodox welfare functions when a discontinuity is allowed at Z^P. In the case of separable functions the extent of the discontinuity at Z^P vis-a-vis the slope of the function, s(y), up to Z^P can provide a measure of the importance of the threshold which the poverty line

13. Using this approach, the exponential equality function given in Table 2 should not be used. It would be more appropriate (though outside the chapter's framework) to write the exponential function for measuring equality as $-\exp[-r(y-\mu)]$, where y is income. This is in fact the function given by Kolm (1976, p.427).

14. The variant of the Sen index proposed by Thon (1979), which still sums over the incomes of the poor, but replaces N by q in the weights used, does satisfy 3A. However, it is not replication-invariant and so does not satisfy 1A.

represents. The head-count ratio is an extreme case: all that matters is whether you are poor or not, which is captured by s being everywhere horizontal, and simply jumping up, from -1 to 0, at the poverty line. At the other extreme are the weakly concave functions with no interior discontinuity, which assume that poverty is entirely a matter of degree. The almost-egalitarian generalized Clark function, with a gap at the poverty line but also a sloping function up to the poverty line, represents an intermediate case in which both crossing the poverty line has a discrete impact on one's standard-of-living, by an amount determined by C, and becoming poorer still makes you even worse off, by an amount determined by α. Of course, the Sen index also captures these two effects, but the advantage of the generalized Clark family is that it parameterizes the relative importance of the two changes more transparently. At one extreme one has the head-count ratio (by setting C=1) and at the other one has an egalitarian function (by setting C=0). More generally, one can see the trade-off between C (which determines the discontinuity at Z^P) and α (which, given C, determines the slope up to Z^P) by re-writing the function, denoted by ψ, as follows

$$\psi = \begin{cases} \frac{1-C}{\alpha} \int_{\eta}^{Z^P} (\frac{y}{Z^P})^{\alpha} dF(y) - (\frac{1-C}{\alpha} + C)F(Z^P), & \alpha \neq 0 \\ 1-C \int_{\eta}^{Z^P} \ln(\frac{y}{Z^P}) dF(y) - CF(Z^P), & \alpha = 0 \end{cases} \tag{2}$$

$$= -\left\{\frac{1-C}{Z^P} \int_{\eta}^{Z^P} (\frac{y}{Z^P})^{\alpha-1} F(y)dy + CF(Z^P)\right\}$$

[15] (2) shows ψ to be a weighted average of the head-count ratio and the Clark *et al.* function where the weight is C, bounded by zero and one. The second term of (2) captures the 'fixed cost' associated with being on or below the poverty line, the first the 'variable cost' increasing in the ratio between one's income and the poverty line. Obviously, for given α, as C increases the fixed cost of poverty increases relative to the variable cost (the absolute value of the ratio of the second to first term increases). Moreover, one can show (see Appendix A) that

$$\frac{\partial \psi}{\partial \alpha} = -\frac{1-C}{Z^P} \int_{\eta}^{Z^P} \ln\left(\frac{y}{Z^P}\right)\left(\frac{y}{Z^P}\right)^{\alpha-1} F(y)dy \geq 0 \tag{3}$$

Thus, for any given C, as α increases, the fixed cost of poverty will become less important relative to the variable cost (the absolute value of the ratio will fall). However, note that, no matter how large α, one will always be able to find progressive transfers which increase poverty, as long as C>0.

15. As mentioned in the notes to Table 2, it is assumed η>0 and Z^P>0.

One can use this last fact to fix C and α. First say one chooses α, using the usual 'leaky bucket' thought experiment of Okun (as described, for example, in Ahmad and Stern, 1991, p.129) applied to transfers below the poverty line. Then fix some point below the poverty line, say half thereof, and ask how by how much would a person with income $\frac{1}{2}Z^P$ have to have their income increased, to, say, $(\frac{1}{2}+x)Z^P$, to neutralize the social loss from bringing someone with income $(1+x)Z^P$ down to the poverty line. The answer, which fixes C, is given by the following equation based on that in Table 2.

$$\frac{C}{1-C} = \frac{1}{\alpha}\left((\frac{1}{2}+x)^\alpha - \frac{1}{2}^\alpha\right). \tag{4}$$

For example, assume α=-1 and x=.1 (so that C=.25). Then if Z^P=\$100, it is just worth taking \$10 from someone with \$110 and giving it to someone with \$50. We can interpret this as saying that the social value of a dollar to someone at $(1+x)Z^P$ is equal to that of someone at $\frac{1}{2}Z^P$ since a (non-infintisimal) transfer of x between them leaves poverty unchanged. One can see the influence of C countering that of α, since, with α=-1, the social value of a dollar to someone at just below Z^P is only a quarter of its value in the hands of someone with income $\frac{1}{2}Z^P$. Note that C is increasing in x since the larger x the greater the gain to the poor beneficiaries while the increase in poverty due to increasing the number poor remains constant.

The language of dividing the 'cost' of poverty into fixed and variable components may be new, but the idea is not. Foster and Shorrocks (1987) show that any separable poverty function, satisfying 2 and 4, can be represented by some function whose arguments are a continuous poverty function and the headcount ratio.[16] The generalized Clark index is simply an example of this, but it is one of interest on account of its particularly simple structure.

To conclude, the presentation of welfare, equality and opulence functions as special cases of an over-arching standard-of-living function has the advantage of enabling one to see just how closely different measures of the three are related. It also makes clear which differences between the three lead to differences in functional form. A broader range of equality functions is available because equality functions, being based on mean-normalized income, are not restricted to the set of increasing functions. And poverty and opulence functions can be sensibly based on the notion of relative distance from, or in the case of poverty crossing of, a 'line' (the poverty line and mean respectively), something which is unavailable for welfare functions.

16. As mentioned earlier, Foster and Shorrocks use the assumption of sub-group consistency, which implies separability.

So far, our 'unified framework' has been applied only to specific functions. In the remainder of the chapter it is shown how not only different functional forms but also different dominance criteria can be usefully viewed from within this single framework.

III First-order, second-order and mixed stochastic dominance

The best known and most-used dominance criteria are those of first- and second-order stochastic dominance. These are presented in III.1 below, as is a new stochastic dominance criterion for poverty indices. III.2 discusses different graphical approaches to the analysis of second-order stochastic dominance.

III.1 The criteria

III.1.1 First-order stochastic dominance

Definition: Iff $F(y)\leq F^{*}(y)$ $\forall y\in[\eta^{-},Z]$ and $F(y)<F^{*}(y)$ $\exists y\in[\eta^{-},Z]$[17] there is first-order stochastic dominance (1OSD) by F of F^{*} up to Z ($FD_{1}F^{*}(Z)$).

Theorem 2: Iff $FD_{1}F^{*}(Z)$ then $S(F)\geq S(F^{*})$ $\forall S\in\Sigma$ and $S(F)>S(F^{*})$ $\exists S\in\Sigma$, where,

(welfare 1OSD) if y is income and $Z\geq\theta^{+}$, Σ is the set of weakly increasing functions (living-standard functions conforming to 1A and 2);

or (poverty 1OSD) if y is income, Σ is the set of weakly increasing functions insensitive above Z^{P} (living-standard functions conforming to 1A, 2 and 4B) with $Z^{P}\leq Z$.

1OSD is a simple criterion, requiring simply that the dominating distribution have, in the relevant range, a no higher and somewhere lower distribution function. Note that 1OSD cannot be meaningfully applied to equality functions: the distribution functions of two mean-normalized distributions either cross or are entirely coincident.[18] Since 1OSD covers functions which do not satisfy either 3A or 3B, it covers sets of functions wider than those defined earlier as welfare or opulence functions. The proof for welfare 1OSD (for which see Thistle (19) and the references therein) rests simply on the fact that if $FD_{1}F^{*}(\theta^{+})$ then F^{*} can be obtained from F by a series of

17. In this and the other stochastic dominance theorems (2 to 4) η^{-} (θ^{+}) can be replaced by $-\infty$ ($+\infty$).

18. Alternatively, 1OSD implies the dominating distribution has a strictly greater mean.

reductions in income as defined in assumption 2.[19] For poverty 1OSD, see Atkinson (1987), the only case not catered for by whom is that of non-separable functions. (Atkinson deals only with separable functions, though he does mention the possibility of extension to the more general case (p.759)). In fact, the extension is very simple. Only sufficiency need be shown, since necessity need only be demonstrated in relation to some subset of the relevant Σ, and this has already been done by Atkinson. The method of proof, which will be used more than once in this chapter, is to show how, under a certain distributional transformation, stochastic dominance over some restricted range implies stochastic dominance over the whole income range and how indifference to such a transformation restricts one to the sub-class of weakly increasing functions insensitive to changes above some level.

Sufficiency proof for poverty 1OSD Assume $FD_1F^*(Z)$, and choose some poverty line $Z^P \leq Z$. Now generate F_1 from F and F_1^* from F^* so that, for $t>0$, $F_1^{(*)}(y)=F^{(*)}(y)$, $y \leq Z^P$, $F_1^{(*)}(y)=F^{(*)}(Z^P)$, $Z^P<y<Z^P+t$ and $F_1^{(*)}(y)=1$, $y \geq Z^P+t$. 1OSD up to Z^P by F then implies 1OSD by F_1 over the whole income range. So for all functions satisfying 1A and 2, $S(F_1) \geq S(F_1^*)$ (this is the welfare 1OSD theorem above). If in addition S satisfies 4B, then $S(F)=S(F_1)$ and $S(F^*)=S(F_1^*)$. So if S satisfies 1A, 2 and 4B, $S(F) \geq S(F^*)$. This will hold for all $Z^P \leq Z$ and the inequality $S(F) \geq S(F^*)$ will hold strictly for at least one S since -F(y) itself satisfies the above assumptions.■

III.1.2 Second-order stochastic dominance

If one is prepared to add the additional assumption 3A, that the functions in Σ are egalitarian, then one can use the criterion of second-order stochastic dominance.

Definition: Let

$$G(y_k) = \int_{\eta^-}^{y_k} F(y)dy \tag{5}$$

Iff $G(y_k) \leq G^*(y_k)$ $\forall y_k \in [\eta^-, Z]$ and $G(y_k) < G^*(y_k)$ $\exists y_k \in [\eta^-, Z]$ there is second-order stochastic dominance (2OSD) by F of F^* up to Z ($FD_2F^*(Z)$).

Theorem 3: Iff $FD_2F^*(Z)$ then $S(F) \geq S(F^*)$ $\forall S \in \Sigma$ and $S(F)>S(F^*)$ $\exists S \in \Sigma$, where

(welfare 2OSD) if y is income and $Z \geq \theta^+$, Σ is the set of welfare functions (living-standard functions conforming to 1A, 2 and 3A);

19. Even though 2 is defined only for step functions, welfare 1OSD can still be applied to comparisons of continuous distributions since if $FD_1F^*(Z)$ where F and F^* are continuous, then $F_nD_1F_n^*(Z)$ where $F_n^{(*)}$ is a step function which approximates $F^{(*)}$ arbitrarily closely (see Rothschild and Stiglitz, 1970, pp.232-234).

or (equality 2OSD) if y is mean-normalized income and $Z \geq \theta^+$, Σ is the set of equality functions (living-standard functions conforming to 1B and 3A);

or (poverty 2OSD) if y is income, Σ is the set of egalitarian opulence functions (living-standard functions conforming to 1A, 2, 3A and 4A) with $Z^P \leq Z$.

2OSD is defined in relation to the integral of the distribution function - known as the deficit curve - and requires that the dominating distribution have, over the relevant range, a no-greater and at least one point smaller integral. Note that 1OSD up to Z implies 2OSD up to Z but not vice versa. The first application of 2OSD in welfare economics was to equality analysis by Atkinson (1970), whose results were subsequently generalized by Dasgupta, Sen and Starrett (1973). Rothschild and Stiglitz (1973, pp.192-193), Kolm (1976, pp.90-91) and later Shorrocks (1983) and Kakwani (1984) extended the application to welfare analysis and Atkinson (1987) made the extension to poverty analysis.[20] Proofs can be found in these references.[21] They rest for equality and welfare 2OSD on the fact that if $FD_2F^*(\theta^+)$ then F^* can be obtained from F by a series of mean-preserving spreads if y is mean-normalized income and mean-preserving spreads and reductions in y if y is income. For poverty 2OSD, as for poverty 1OSD, the proof rests on the fact that opulence functions are special types of welfare functions. Also as with 1OSD, the only case not catered for in the literature is that of non-separable egalitarian opulence functions. Again Atkinson's proof can be simply generalized.

Sufficiency proof for poverty 2OSD Assume $FD_2F^*(Z)$, and choose some poverty line $Z^P \leq Z$. Now generate F_1 from F and F_1^* from F^* so that $F_1^{(*)}(y)=F^{(*)}(y)$, $y<Z^P$ and $F_1^{(*)}(y)=1$, $y \geq Z^P$. 2OSD up to Z^P by F then implies 2OSD by F_1 over the whole income range. So for all functions satisfying 1A, 2 and 3A, $S(F_1) \geq S(F_1^*)$ (this is the welfare 2OSD theorem given above - see any of the references quoted for a proof). If in addition S satisfies 4A, then $S(F)=S(F_1)$ and $S(F^*)=S(F_1^*)$. So if S satisfies 1A, 2, 3A and 4A, $S(F) \geq S(F^*)$. This will hold for all $Z^P \leq Z$ and the inequality $S(F) \geq S(F^*)$ will hold strictly for at least one S since $-G(y_k)$ itself satisfies the above assumptions.■

An illustration of 2OSD is given in Figure 1a, while the proof is illustrated in Figure 1b. Note the role played by the strong definition of the poor in assumption 4A. If 4B was

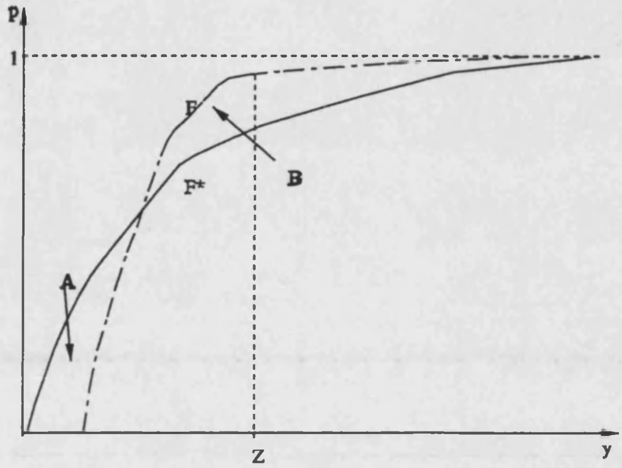
20. Even though 3A is defined only for step functions, welfare 2OSD (like 1OSD - see footnote 19) can still be applied to comparisons of continuous distributions since if $FD_2F^*(Z)$ where F and F^* are continuous, then $F_nD_2F_n^*(Z)$ where $F_n^{(*)}$ is a step function which approximates $F^{(*)}$ arbitrarily closely (see Rothschild and Stiglitz, 1970, pp.232-234).

21. Lambert (1989) gives a review of these stochastic dominance conditions in relation to welfare and equality analysis, while Ravallion (1992) focuses on poverty analysis. For more general reviews of stochastic dominance itself, which has a longer history and which was originally applied to risk analysis, see Whitmore and Findlay (1978) and Fomby and Seo (1989).

assumed, one would require $F_1^{(*)}(y)=F^{(*)}(y)$, $y \leq Z^P$ and $F_1^{(*)}(y)=1$, $y>Z^P$, but then F_1 and F_1^* would not be distribution functions as they would not be right-continuous.

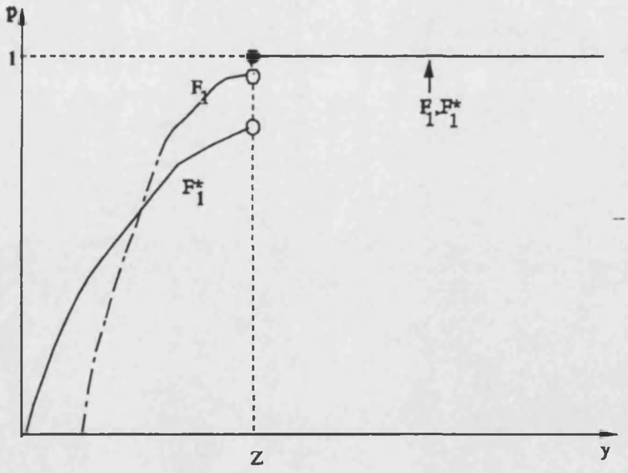
Figure 1 Illustration of Second-order Stochastic Dominance and Sufficiency Proof

1a. $FD_2F^*(Z)$



Notes: On the assumption that area A is no smaller than area B, F has 2OSD over F^* up to Z.

1b. $F_1D_2F_1^*(\theta^+)$



Notes: F_1 and F_1^* are generated respectively from F and F^*. The latter and the former in each pair are identical up to Z^P which has been chosen as equal to Z. At Z, the distribution functions go up to 1. 2OSD by F of F^* up to Z implies 2OSD by F_1 of F_1^* up to the maximum income value of the two distributions.

III.1.3 Mixed dominance

Whereas it is hard to imagine wanting sets of welfare and (relative) equality functions defined more widely than those covered by the second-order stochastic dominance criteria, the situation is not quite the same for opulence functions, since the class of functions satisfying 3B but not 3A is excluded. While, for example, the 2OSD criterion for inequality covers all of the well-known equality functions (including all of those in Table 2),[22] the same cannot be said for the 2OSD criterion for poverty, which, as mentioned, excludes the negative of the head-count and the Sen index. The exclusion of functions satisfying 3B but not 3A not only limits the relevance of the poverty 2OSD criterion. It is also unsatisfactory since, as argued, almost-egalitarian functions have a claim to our attention in poverty analysis which is absent in the cases of equality and welfare analysis. For these reasons, a third criterion is introduced below, which is more demanding than second- but less so than first-order stochastic dominance, and which covers all opulence functions, as defined in Table 1.

Definition: Iff $G(y_k)\leq G^*(y_k)$ $\forall y_k\in[\eta^-,Z^-]$ and $F(y)\leq F^*(y)$ $\forall y\in[Z^-,Z]$ and either $G(y_k)<G^*(y_k)$ $\exists y_k\in[\eta^-,Z^-]$ or $F(y)<F^*(y)$ $\exists y\in[Z^-,Z]$, then there is poverty mixed dominance by F of F^* between Z^- and Z, or $FD_mF^*(Z^-,Z)$.

Theorem 4: Iff $FD_mF^*(Z^-,Z)$ then $S(F)\geq S(F^*)$ $\forall S\in\Sigma$ and $S(F)>S(F^*)$ $\exists S\in\Sigma$, where, if y is income, Σ is the set of opulence functions (living standard functions conforming to 1A, 2, 3B and 4B) with $Z^P\in[Z^-,Z]$.

If we rule out the possibility that $F(y)=F^*(y)$ $\forall y\in[Z^-,Z]$ this criterion is equivalent to 2OSD up to Z and 'restricted' 1OSD between Z^- and Z (as defined before Theorem 2, but replacing η^- by Z^-).[23] Hence the name: mixed dominance requires a mixture of dominance conditions to hold. Figure 2a illustrates. Up to Z^- we need to look at the area under the distribution curve, and, from Z^- to Z, at the distribution function itself. (Unrestricted) 1OSD implies mixed dominance implies 2OSD, all up to Z. The bounds Z^- and Z can be understood as limiting the set of reasonable poverty lines we are prepared to consider. This is contrast to 1OSD and 2OSD where only an upper bound, Z, is given on this set. It might be thought, therefore, that the Σ covered by the mixed dominance criterion is smaller than that covered by the 2OSD criterion. However, since any egalitarian opulence function insensitive to changes above some Z is also an egalitarian opulence function with $Z^P\geq Z$, the class of all opulence functions with $Z\geq Z^P\geq Z^-$ includes,

---

22. The only exception is the standard deviation of logarithms.

23. The 'restricted' 1OSD condition is Atkinson's (1987) Condition I.

by relabelling, all egalitarian opulence functions with $Z^P<Z^-$.[24]

The necessity of mixed dominance for the class of functions given follows simply from noting that egalitarian opulence functions are a subset of this class - so justifying the need for 2OSD up to Z - as is the negative of the head-count ratio - hence the need for 1OSD over the permissible range of the poverty line. Sufficiency can be most easily proved in the case in which S is a separable and, up to Z^P, weakly concave and differentiable function. This enables us to set $s(y)=0$ for $y>Z^P$. Let $\Delta F=F-F^*$ and $\Delta G=G-G^*$. Using repeated integration of parts on (1) gives the result that

$$S(F)-S(F^*) = \Delta F(Z^P)s(Z^P) - \Delta G(Z^P)s'(Z^P) + \int_{\eta^-}^{Z^P} \Delta G(y)s''(y)dy \tag{6}$$

$s(Z^P)$ must be non-positive by the assumption of s being weakly increasing in income. If the function satisfies 3A, $s(Z^P)$ will be zero - since otherwise regressive transfers across the poverty line could raise S - and we can ignore the first term. Otherwise, 1OSD at Z^P will make this term non-negative. If there is 2OSD up to Z^P, the last two terms of (6) will be positive. Allowing Z^P to vary between Z^- and Z proves sufficiency.

A more general proof of sufficiency requires that the assumptions of separability and differentiability be dropped and can be given in three steps, using a method similar to that utilized earlier for the poverty 2OSD sufficiency proof.

Sufficiency proof for Theorem 4

1. Let y be income, assume $FD_mF^*(Z^-,Z)$ and choose $Z^P\in[Z^-,Z]$. Generate F_1 from F so that $F_1(y)=F(y)$, $y\leq Z^P$, $F_1(y)=F(Z^P)$, $Z^P<y<Z^P+t$, and $F(y)=1$, $y\geq Z^P+t$, and similarly for F^*.

2. From the welfare 2OSD theorem, all welfare functions satisfying 1A, 2 and 3A weakly prefer F_1 to F_1^*. Since such functions are weakly-increasing and egalitarian, it is possible, via a sequence of mean-preserving spreads and reductions in income to generate F_1^* from F_1 (see Rothschild and Stiglitz, 1973, pp.192-3). If S satisfies 1A and 2, it cannot be increased by reductions in income. If S also satisfies 3B, the mean-preserving spreads can only increase S if they involve crossings of the poverty line. But since $F_1^*(Z^P)\geq F_1(Z^P)$, no Z^P-crossing mean-preserving

24. By the same argument, I could have re-written the 1OSD (2OSD) poverty theorem as applying only to functions satisfying 4B (4A) with $Z^P=Z$ (since if they are insensitive to changes above some $Z\leq Z^P$, they are certainly insensitive to changes above Z^P). Since this is only a matter of labelling, I have chosen for the different theorems what seems to me to be the most transparent way of expressing the Σ covered.

spreads will be required.[25] So for all S satisfying 1A, 2, and 3B, $S(F_1) \geq S(F_1^*)$.

3. Finally, for any S which satisfies 4B, $S(F)=S(F_1)$ and $S(F^*)=S(F_1^*)$. So for all S satisfying 1A, 2, 3B and 4B, $S(F) \geq S(F^*)$. Varying Z^P within $[Z^-,Z]$ will give the same result, and the inequality will hold strictly for at least one S. The assumptions used to get this result restrict the set Σ to being that of all opulence functions with poverty lines between Z^- and Z.■

The proof is illustrated in Figure 2b. Note that the requirement that $F(Z^P) \leq F^*(Z^P)$ enables assumption 4B to be used without encountering the problems described at the end of III.1.2.

Requiring F to have restricted 1OSD (a no-higher head-count ratio over the set of poverty lines) as well as 2OSD enables one to expand greatly the class of functions over which dominance is guaranteed. One can add not only the negative of the head-count ratio to the egalitarian class of opulence functions,[26] but also an entire additional class of functions, which can be viewed as intermediate to the two extremes of the head-count and the egalitarian class, and which includes the negatives of Sen's index and the generalized Clark function introduced in the previous section.[27] For this reason, the mixed dominance result for poverty should be seen as giving a similar coverage of poverty functions as is given by the second-order stochastic dominance results for welfare and equality functions.[28]

The extent to which mixed dominance will increase the number of distributions which can be ranked compared to 1OSD and decrease the number compared to 2OSD is increasing in

---

25. To see this, note that one can modify F_1 by reducing incomes so that for the 'new' F_1, $F_1(y)=F_1^*(y)$, $Z^- \leq y \leq Z^P+t$. Then the new F_1 will still have welfare 2OSD over F_1^* and the two distributions will be identical for $y \geq Z^-$. Then only mean-preserving spreads (and possibly further reductions in income) up to $Z^- \leq Z^P$ will be further required to generate F_1^* from F_1.

26. Here, of course, reference to the 'egalitarian class of poverty functions' is reference to that class of poverty functions the negatives of which are egalitarian opulence functions.

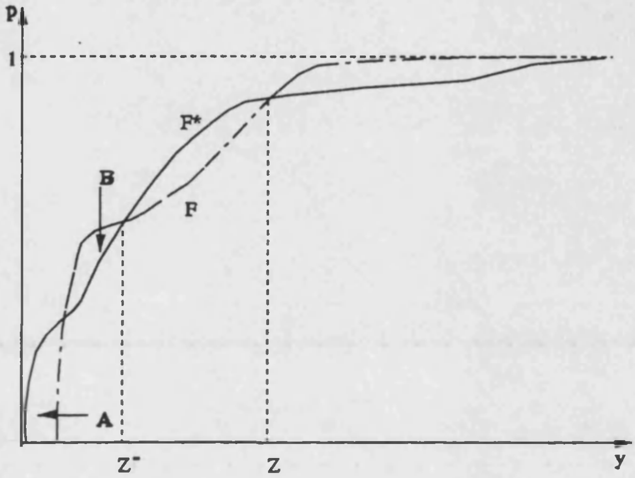
27. The importance of the head-count ratio in this regard can also be seen in the argument of Foster and Shorrocks (1987) that any 'sub-group consistent' poverty function can be written as a function whose arguments are a continuous poverty function and the head-count ratio. (The generalized Clark function given earlier is precisely an example of this.) Although the mixed dominance criterion covers a wider class than the sub-group consistent, one can see 2OSD as ensuring dominance by all continuous functions and restricted 1OSD as ensuring dominance by the head-count ratio.

28. Other lesser-known poverty functions will also be captured by the mixed dominance criterion, including the Kakwani (1980b) and Blackorby and Donaldson (1980) indices. Chakravarty (1989) groups these together with the Sen index as, in our terminology, having negatives which satisfy 3B but not 3A. But not all poverty functions will be covered. Some have negatives which do not satisfy 2, such as the Hamada and Takayama (1977) class of indices. Others, such as the Thon index, do not satisfy 1A (see footnote 14). But since 1A and 2 are uncontroversial, this omission is not serious.

the magnitude of Z^- but is otherwise an empirical matter. It is easily shown that if mixed dominance is to obtain where there is no 1OSD, it must be the case that the distribution functions being compared cross an even number of times below Z^-.[29]

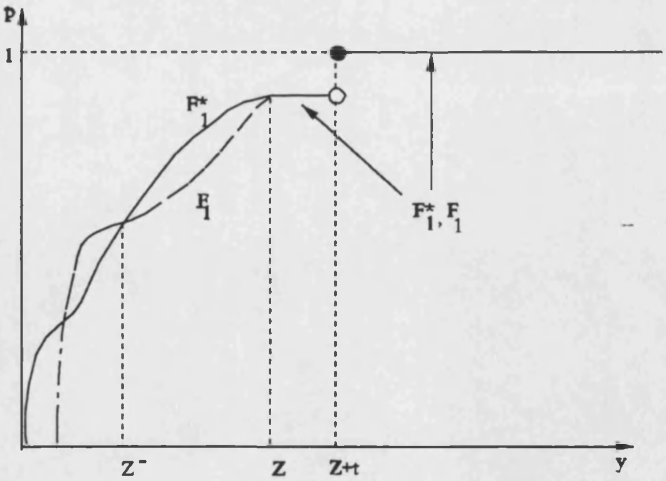
Figure 2 Illustration of Mixed Dominance and Sufficiency Proof

2a. $FD_mF^*(Z^-,Z)$



Notes: Assuming that area A is no smaller than area B, F has 2OSD over F^* up to Z. In addition, F has a lower distribution function between Z^- and Z so F has mixed dominance over F^* between Z^- and Z.

2b. $F_1D_2F_1^*(\theta^+)$



Notes: F_1 and F_1^* are generated respectively from F and F^*. The latter and the former in each pair are identical up to Z^P which has been chosen as equal to Z. At Z+t, the distribution functions go up to 1. 2OSD by F of F^* up to Z implies 2OSD by F_1 of F_1^* up to the maximum income value of the two distributions.

---

29. 2OSD to Z requires that the dominating distribution have a no lower minimum income, which in turn implies a no higher distribution function at η^- (see IV.1).

III.2 Use of Lorenz and deficit curves to evaluate second-order and mixed stochastic dominance

1OSD is easily analyzed by comparing distribution functions. 2OSD could, as already mentioned, be similarly analyzed by examining deficit curves.[30] For welfare or poverty 2OSD, one should define the curve over income, in which case one has a *poverty deficit* curve. The poverty deficit curve can also be used in conjunction with the cumulative density curve to evaluate mixed dominance: in this case one needs to examine the poverty deficit curve up to only the lower bound of the set of poverty lines. For examination of equality 2OSD, the deficit curve should be defined over mean-normalized income resulting in an *equality deficit* curve.

In practice, however, these deficit curves are rarely used. The more common approach exploits the close relationship between 2OSD and Lorenz dominance, and works with the more familiar family of **Lorenz** curves. The usefulness of Lorenz curves for this purpose was first demonstrated by Atkinson (1970) in relation to equality analysis, and the duality between the two in the context of welfare analysis has recently been shown by Atkinson and Bourguignon (1989) and Thistle (1989). All Lorenz curves are constructed according to the same formula, but they can be given three different names associated with the three different types of variable used to construct them: the *ordinary* Lorenz curve constructed using mean-normalized income, the *generalized* Lorenz (GL) curve using income and the *censored* Lorenz curve using censored income, that is equating y with Z^P for that proportion of the population whose income is greater than Z^P. The purpose of this sub-section is to analyze the extent to which the Lorenz and deficit family of curves can act as substitutes for each other.

Following Gastwirth (1971), define $Q(p)=\inf\{y:F(y)\geq p\}$.[31] Then the family of Lorenz curves is defined by

$$\phi(p_k) = \int_0^{p_k} Q(p)dp \tag{7}$$

(7) shows clearly the sense in which the Lorenz curve is a dual for the deficit curve (Atkinson and

30. The deficit curve, $G(Z^P)$, is equal to Z^P times the poverty gap (PG). Using integration by parts:

$$G(Z^P)=\int_{\eta}^{Z^P} F(y)dy=Z^P[F(Z^P)-\frac{1}{Z^P}\int_{\eta}^{Z^P} ydF(y)]=Z^P\int_{\eta}^{Z^P}\frac{Z^P-y}{Z^P}dF(y)=Z^P PG(Z^P)$$

31. If the distribution is continuous then $Q(p)=F^{-1}(p)$.

Bourguignon, 1989). In geometrical terms, whereas the deficit curve measures the area between the distribution function and the income axis, the Lorenz curve measures the area between the distribution function and the p-axis: see Figure 3 at the end of the sub-section.

If

$$\phi(p_k)\geq\phi^*(p_k)\ \forall p_k\in[0,q] \text{ and } \phi(p_k)>\phi^*(p_k)\ \exists p_k\in[0,q] \tag{8}$$

then F **Lorenz dominates** F^* up to q ($FD_LF^*(q)$). Note again the duality between 2OSD (the primal approach) and Lorenz dominance (the dual): the inequalities in (8) are the reverse of those used to define 2OSD - see Theorem 3.

But is the choice between deficit and Lorenz curves of no consequence? We will define a one-to-one relationship between Lorenz dominance and second-order stochastic dominance as existing if one can, for all distributions, find some general rule relating Z and q, such that

$$FD_2F^*(Z) \text{ iff } FD_LF^*(q) \tag{9}$$

By a 'general' rule here is meant one which equates q either to F(Z) or $F^*(Z)$ for all distributions. If and only if (9) holds can any particular members of the deficit and Lorenz curve families be thought of as substitutes for analysis. The results of a comparison along these lines are presented in Table 3.

Table 3 The Deficit and Lorenz Families of Curves as Substitutes for Analysis

For the analysis of ...	**the...Lorenz curve**	**can be substituted for the ... deficit curve**	**References**
Equality 2OSD	ordinary	equality	Atkinson (1970)
Welfare 2OSD	generalized	poverty	Atkinson (1970), Shorrocks (1983), Thistle (1989)
Poverty 2OSD	censored	poverty	Foster and Shorrocks (1988b)
Poverty Mixed Dominance	generalized or censored	poverty	

To derive these results, change the variable of integration in (7) to give

$$\phi(p_k) = \int_{\eta}^{y_k} y dF(y) \tag{10}$$

where y is defined in one of the three ways given earlier, that is as income, mean-normalized or censored income. This gives the Lorenz curve at p_k as the mean of y conditional on $y \leq F(p_k)$ divided by p_k. Applying integration by parts to (10), one has

$$\phi(p_k) = -G(y_k) + y_k F(y_k) \tag{11}$$

Using (11), one can derive the following two relationships between second-order stochastic and Lorenz dominance, where, as before, $p_k = F(y_k)$ and now, in addition, $p_k^* = F^*(y_k)$, $\hat{q} = F(Z)$, and $q^* = F^*(Z)$. First,

$$\text{if } G^*(y_k) \geq G(y_k) \text{ then } \phi(p_k^*) > \phi^*(p_k^*) \tag{12}$$

which implies

$$\text{if } FD_2F^*(Z) \text{ then } FD_LF^*(q^*). \tag{13}$$

Second,

$$\text{if } \phi(p_k) \geq \phi^*(p_k) \text{ then } G^*(y_k) > G(y_k) \tag{14}$$

which implies

$$\text{if } FD_LF^*(\hat{q}) \text{ then } FD_2F^*(Z). \tag{15}$$

The proofs are given in Atkinson (1970), and rely on application of the mean-value theorem.[32]

From (13) and (15), it follows that:

Theorem 5 There exists no general rule to relate q and Z such that, for any pair of distributions F and F^*, $FD_2F^*(Z)$ iff $FD_LF^*(q)$

From (13) if such a q did exist it would have to be equal to $q^* = F^*(Z)$. So assume $FD_LF^*(q^*)$. But from (15) one may not have $FD_2F^*(Z)$ despite $FD_LF^*(q^*)$, as long as $q^* < \hat{q} = F(Z)$ and it is not the

32. Note that Atkinson, who is concerned with the ordinary Lorenz curve, explicitly divides through by the mean which is not necessary here. See also Thistle (1989) who stresses that the these relationships will hold for the distributions over which they are defined, regardless of whether the distributions are continuous, partial or mixed.

case that $FD_L F^*(\hat{q})$.■

This theorem has a useful corollary:

Corollary to Theorem 5 If $F(Z) \leq F^*(Z)$ (i.e., $\hat{q} \leq q^*$) then $FD_2F^*(Z)$ iff $FD_LF^*(q)$ for $q=\hat{q}=F(Z)$.

Figure 3 and its notes illustrate these results, which provide the basis for Table 3. For the cases of welfare and equality, $Z \geq \theta^+$ (see the 1OSD and 2OSD theorems), so the corollary can be applied, as we will have $F(Z)=F^*(Z)$. Also for the case of the censored Lorenz curve, one is restructuring the distributions so that if there is 2OSD up to the upper bound on the set of possible poverty lines, there is 2OSD up to $Z \geq \theta^+$, so again the corollary can be applied. For the case of mixed dominance, one can first look at the distribution functions to see if $F(y) \leq F^*(y)$, for all $y \in [Z^-, Z]$. If not, neither the deficit nor Lorenz curve need be consulted, as mixed dominance cannot hold. But if so, again by the corollary, the two are perfect substitutes. For the case of poverty 2OSD, however, the corollary cannot be applied to justify use of the GL curve since one may have a case in which $F(Z)>F^*(Z)$, and yet be unsure whether or not F dominates - indeed, this is precisely the case which Figure 3 illustrates.[33] This result that the GL curve cannot be used for poverty 2OSD analysis has not yet been stated in the literature. It is one which limits the usefulness of the Lorenz family of curves.[34]

It is no doubt true the Lorenz curve family has presentational advantages over the deficit curve family. In particular, the former has the advantage of always being bounded on the horizontal axis between 0 and 1, and often also covers less vertical distance.[35] Nevertheless the comparisons of Table 3 do reveal some compensating benefits from working with the deficit

33. Labelling as F the distribution with the lowest head-count at Z (so that by definition $F(Z) \leq F^*(Z)$) does not save the Lorenz curve as a tool for analysis. For it may be that due to a lower intersection we know it is not the case that the distribution with the lowest head-count at Z is the dominating distribution.

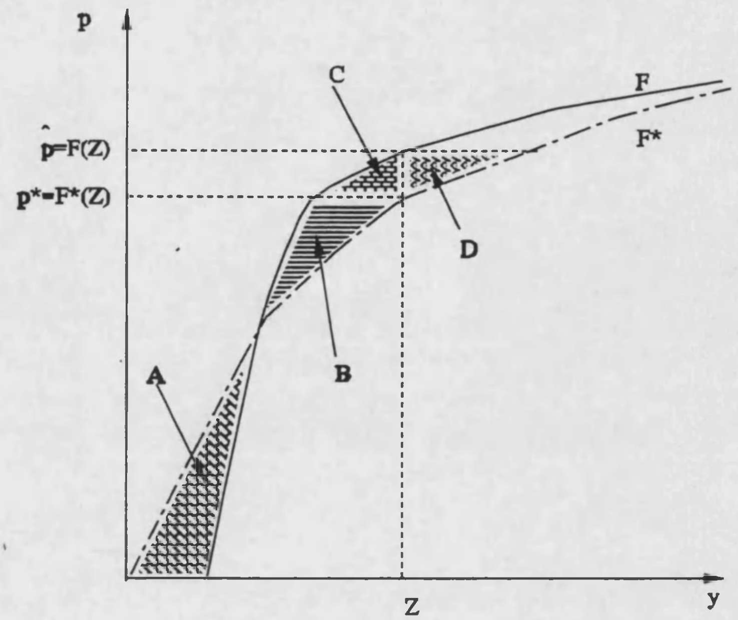
34. Foster and Shorrocks (1988a) analyze poverty 2OSD in terms of the generalized Lorenz curve, but their case is a special one in which no upper bound on the set of poverty lines is allowed; in this case poverty and welfare comparisons are equivalent and the corollary to Theorem 5 can be applied. But this result, while conceptually interesting, is of limited relevance for the analysis of poverty. The set of 'reasonable poverty lines'., which Foster and Shorrocks refer to in their more general discussion (pp. 173, 174), certainly does not include infinite income.

35. From (11), the maximum point of the Lorenz curve is the mean of the distribution, whereas the maximum point of the deficit curve is the maximum y minus the mean. Note that as an alternative to drawing the Lorenz curve, one can draw the curves divided by p (to give a graph of conditional means - see Atkinson and Bourguignon, 1989, and equation (10)), and as an alternative to the deficit curves, one can draw $G(y_k)/y_k$ (to give a graph of the poverty gap - see footnote 29.)

family. In particular, whereas the same deficit curve may be used to evaluate welfare 2OSD, poverty 2OSD and poverty mixed dominance, different Lorenz curves must be used: one for welfare 2OSD and poverty mixed dominance and one for poverty 2OSD. To this one may add another two points. First, a different censored Lorenz curve will need to be drawn for every time the upper bound on the set of poverty lines is increased. In addition, the same difficulties of interpretation as discussed above will arise for crossings of the censored Lorenz curve if one wants to consider using a lower poverty line. By contrast, a single poverty deficit curve allows analysis over all possible poverty lines. Second, as Atkinson (1987) has commented, the deficit curve seems to be a more natural tool for the analysis of poverty, as it is defined directly over income rather than over population proportions.

As Atkinson and Bourguignon (1989) remark, the relative attractiveness of working with deficit and Lorenz curves will depend on the particular problem to hand. It is unlikely that the equality deficit curve has much to recommend it for use *vis-a-vis* the Lorenz curve. But for the analysis of welfare and especially that of poverty, the foregoing would suggest that the family of deficit curves should receive more attention from researchers than has been the case to date. (Howes and Lanjouw (1991) and Lanjouw (1992) provide empirical illustrations. See also Chapter Five.)

Figure 3 Generalized Lorenz and Deficit Curves



Notes: 1. Assume y is income so the area between the horizontal axis and the distribution function gives the poverty deficit curve, that between the vertical axis and the distribution function the GL curve.

2. To see why the GL and deficit curves are not perfect substitutes for poverty 2OSD analysis, take the point Z as an example. Could we tell from looking at the GL curve whether F has 2OSD over F^* up to Z? We would need first to look at the GL curve at p^*. If F doesn't have GL dominance up to p^* it certainly doesn't have 2OSD up to Z. This is what (13) tells us. It is also evident from the graph: if area A is not greater than area B, it is certainly not greater than B and C combined, which is required for 2OSD up to Z. Say F does have GL dominance up to p^*. We could then look at $\hat{p}$. If F does have GL dominance up to $\hat{p}$ it will have 2OSD up to Z, but it may have the latter without the former. This is what (15) tells us. It can also be seen from the graph, as GL dominance up to $\hat{p}$ requires the subtraction from area A of areas D as well as B and C. So there is no single p one can look at to determine if there is 2OSD up to Z. Of course for each distribution one can find a p, but no general rule can be given. Hence Theorem 5. Finally, note that this lack of substitutability between the two curves arises when the possibly dominating distribution has a higher value of p at Z: in this case $\hat{p}>p^*$. If we can rule this out, the problem disappears. This provides the basis for the corollary to Theorem 5.

IV Alternatives to second-order stochastic dominance

As Lambert notes, use of the second-order stochastic dominance criterion has "become standard among researchers" (1989, p.5). But there are many other criteria which could be used. Several are explored in this section, using a basic distinction between those criteria which require 'extreme' forms of dominance, explained below, and those which do not.

IV.1 'Extreme' forms of dominance and alternatives requiring them

The 'extreme' forms of dominance are defined to include mean dominance, minimum dominance, maximum dominance, and any combination of these three. One distribution has mean dominance over another if its mean y is no lower. It has minimum (maximum) dominance if it has a no lower minimum (no higher maximum) y than another, i.e., if $\eta \geq \eta^*$ ($\theta \leq \theta^*$).[36] As usual, y can be defined to be either income or mean-normalized income. The label 'extreme' is intended to convey two meanings. Minimum and maximum dominance are extreme forms of dominance in the sense that they refer to the tails of the distributions being compared. And, as we will see, all three are extreme in reference to the living-standard functions they ensure weak preference over.

Poverty 2OSD and hence mixed dominance require minimum dominance, welfare 2OSD requires mean-minimum dominance (a no lower minimum and no lower mean), and equality 2OSD requires maximum-minimum dominance (defined over mean-normalized income).[37] All these requirements can be explained in terms of the isoelastic and GE functions of Table 2. If F does not have minimum dominance over F^*, S will be lower in F for the isoelastic function with $\alpha \rightarrow -\infty$. If F doesn't have mean dominance, S will be lower for $\alpha=0$. If F doesn't have maximum dominance, S will be lower for $\delta \rightarrow \infty$. The requirements can also be understood in terms of the curves of III.2. If F doesn't have minimum dominance, it will have a higher deficit curve and lower GL curve at some point arbitrarily close to η^- and 0 respectively. If y is mean-normalized income, and F has a higher maximum, it's Lorenz curve will be lower for some p arbitrarily close to one (see Atkinson and Bourguignon, 1989).

Sometimes the requirement of minimum dominance is overlooked in discussions of 2OSD. Bishop, Formby and Thistle (1989), for example, note that "a distribution with a greater mean can never be [second-order stochastically] dominated by a distribution with a lower mean" (p.76). Since replacing 'mean' by 'minimum' preserves the truth of this proposition, their use of it as an example to support their claim that the 2OSD criterion "is heavily weighted towards efficiency preference" is unjustified.

---

36. (a) Strictly speaking, these types of dominance cannot be regarded as dominance criteria since they do not ensure that at least one S in some Σ strictly prefers one distribution to another. For this, one would need to require, e.g. for mean dominance, that the dominating distribution have a higher, not just a no lower, mean. But then these types of dominance would no longer be necessary conditions for 2OSD.

(b) If it was not assumed that η or θ exist, then a more general definition would be required of minimum and maximum dominance. For example, if one restricts attention to distribution functions which cross a finite number of times, one could say that F has minimum dominance if there exists z such that for all $y \leq z$, $F(y) \leq F^*(y)$ (see Lambert, 1989, p.71).

37. Equality 2OSD also requires mean dominance, but this requirement is met automatically since two mean-normalized distributions have the same (derived) mean, equal to one. In addition, welfare 1OSD also requires the dominating distribution to have a no lower maximum y, not a no higher maximum y as in the case of equality 2OSD. However, the focus of the section is on 2OSD.

One set of alternatives to 2OSD which requires at least certain extreme forms of dominance is that of higher orders of stochastic dominance. For example, Whitmore (1970) puts forward the criterion of third-order stochastic dominance, which limits Σ to those functions which display the property of 'transfer sensitivity' (a positive third derivative if separable and differentiable). The usefulness of this approach is probably greatest in inequality analysis, since it removes the requirement of maximum dominance. Obviously one can go on and on in this direction, looking for fourth and higher orders of stochastic dominance. Each order of stochastic dominance requires simply another round of integration, though there may also be end-point conditions to meet. For example, welfare third-order stochastic dominance can be checked by integrating over the deficit curve, though one needs to check separately for mean dominance.[38]

Another alternative is the mean-Lorenz criterion analyzed, for example, by Shorrocks (1983). As the name suggests, this dominance criterion applies to welfare analysis, and requires one distribution to have a no lower mean and Lorenz dominance (or a strictly higher mean if the Lorenz curves are identical). In contrast to the use of higher-than-second orders of dominance, mean-Lorenz dominance, if it obtains, gives dominance over an even larger set of welfare functions than the 2OSD set: not only do all egalitarian functions which are weakly increasing in income have dominance, but also all those which are weakly increasing only in mean income. Bishop, Formby and Thistle (1989) argue in favour of this criterion and against 2OSD by reference to the alleged 'efficiency bias' of 2OSD, shown by the fact that if F is obtained from F^* by increasing the income of the richest individual, then F has welfare 2OSD over F^*, even though inequality has increased (F^* Lorenz dominates F). Using the mean-Lorenz criterion, F and F^* are unrankable.

Both the higher-than-second order and the mean-Lorenz criterion can be criticized on various grounds. Assuming separability and differentiability, the former requires the observer to make judgements as to the third or higher derivatives of s(y), which he or she may not feel confident to do. And the latter, by abandoning 2, sacrifices the Pareto principle. One criticism which can be levelled against both is precisely the requirement they share of reliance on extreme forms of dominance. This is obvious for the mean-Lorenz criterion. The higher orders of stochastic dominance also require mean-minimum dominance if applied to welfare analysis, and require minimum dominance in the cases of poverty and equality. (Indeed, all orders of stochastic dominance require minimum dominance.) There are two reasons why such a reliance is problematic.

38. Lambert (1989) provides an algorithm for the detection of third-order stochastic dominance which can be applied to both welfare and equality analysis.

First, it may be very hard to tell which of a pair of distributions displays minimum or maximum dominance. This will be particularly the case if one is using disaggregated data and trying to infer dominance between two populations from sample data. A moment's reflection reveals that it will not be an easy matter to infer that one distribution has a higher minimum income or a lower maximum-to-mean income ratio than another.

Second, the requirement of each form of extreme dominance represents an extreme normative judgement. Distributional indifference is implied by the requirement of mean dominance in the case of welfare analysis. So-called 'Rawlsianism' is implied by the requirement of minimum dominance.[39] No label comes to mind for the requirement of maximum dominance, but it is far from clear why one should want to judge the inequality in a society on the basis of the maximum-to-mean income ratio.

The first issue is an empirical one, which is analyzed in the Section IV of the next chapter. In this chapter, with its normative focus, it is appropriate to say a little more on the second objection. There would seem to be very little support for the view of distributional indifference and none for that of the requirement of maximum dominance. But the requirement of minimum dominance is often presented as receiving support from the influential work of Rawls (1971). For example, Lambert writes:

> Rawls has argued that an ethically justifiable approach to social choice of income distribution is to seek to improve the position of the least well-off income unit regardless of all else. (1989, p.70)

But this is simply wrong. Rawls specifically addresses this issue. He begins by giving two suggestions for defining the least advantaged *group*, whose standard-of-living, according to his 'difference principle', should be maximized, and suggests unskilled workers or those with less than half the median income. He then continues in a more general vein:

> In any case, we are to aggregate to some degree over the expectations of the worst off, and the figure selected on which to base these computations is to a certain extent ad hoc. Yet we are entitled at some point to plead practical considerations in formulating the difference principle. Sooner or later the capacity of philosophical or other arguments to make finer discriminations is bound to run out. *I assume therefore that ... the difference principle [is to be interpreted] from the first as a limited aggregative principle... It is not as if [we agree] to think of the least advantaged as literally the*

39. While minimum dominance by F is necessary for $S(F) \geq S(F^*)$ by the so-called 'Rawlsian leximin' principle, it is not sufficient if $\eta = \eta^*$. In such cases, recourse to the more general definition given in 36 is required, as it is if η does not exist.

worst off individual and then in order to make this criterion work [adopt] in practice some form of averaging. (1971, p.98, italics added)

This suggests support from Rawls for ordering distributions not by the criterion of minimum dominance but by some distributionally-insensitive poverty measure such as the poverty gap.[40] Without Rawls' backing, it is difficult to see what support can be garnered for the requirement of minimum dominance, and all too easy to think of counter-examples (involving two-person societies with massively different means and almost identical minimum incomes) which go against it.

If one wishes to avoid criteria which are reliant on extreme forms of dominance, one needs either to restrict the stochastic dominance criteria or to leave the stochastic dominance framework altogether.

IV.2 Restricted dominance

In a recent paper, Atkinson and Bourguignon (1989) suggest the criterion of 'restricted dominance', based on 2OSD but with the difference "that we do not ... concern ourselves with what happens beyond a certain income level or percentage of the population." (p.11). In what follows, I generalize their suggestion to apply, appropriately modified, to both ends of the income distribution, show over what sets of functions criteria thus derived are a necessary and sufficient condition for dominance, and consider justifications for reliance on these sets. Note that the justifications come last. The various assumptions presented are of differing degrees of plausibility, but it is more convenient to analyze them once it is known what results they imply.

Define $\phi(p_k)$ as in (7) to represent the Lorenz Curve family of functions. Using this, define $FD_LF^*(p_1,p_2)$ to hold iff

$$\phi(p_k)\geq\phi^*(p_k)\ \forall p_k\in[p_1,p_2] \text{ and } \Delta\phi(p_k)>\phi(p_k)\ \exists p_k\in[p_1,p_2] \tag{16}$$

Then if p_k is defined over income, we will say that iff (16) holds F has **p-restricted** (second-order stochastic) **welfare dominance** over F^* (between p_1 and p_2). If p_k is defined over mean-normalized income (to be the proportion who have mean-normalized income below $Q(p_k)$), we will say that iff (16) holds F has **restricted** (second-order stochastic) **equality dominance** over F^* (between p_1

40. A similar point is made by Atkinson (1987). But note Rawls' remark that "A distribution cannot be judged in isolation from the system of which it is an outcome or from what individuals have done in good faith in the light of established expectations." (p.88), which suggests he would not be very sympathetic to any tradition as minimalist as that pursued in this chapter.

and p_2). Similarly define $\Delta G(y_k)$ as in (5) to represent the deficit curve family. Using this, define $FD_2F^*(Z_1,Z_2)$ to hold iff

$$G(y_k) \leq G^*(y_k) \ \forall y_k \in [Z_1,Z_2] \text{ and } G(y_k) < G^*(y_k) \ \exists y_k \in [Z_1,Z_2] \tag{17}$$

Then, defining y as income, we will say that iff (17) holds, F has **y-restricted** (second-order stochastic) **welfare dominance** over F^* (between Z_1 and Z_2). [41]

These new criteria are very simple. Restricted equality and p-restricted welfare dominance can be checked via an examination of the distributions' ordinary and generalized Lorenz curves respectively, between p_1 and p_2, and y-restricted welfare dominance can be checked via examination of the deficit curve between Z_1 and Z_2.

Fairly obviously, the criteria are generalizations of the corresponding unrestricted second-order stochastic dominance criteria. Equality 2OSD is a special case of restricted equality dominance with p_1=0 and p_2=1. Welfare 2OSD is a special case of p-restricted welfare dominance with p_1=0 and p_2=1, and of y-restricted welfare dominance with $Z_1 \leq \eta^-$ and $Z_2 \geq \theta^+$. Note that there is no mention here of poverty dominance, even though, as will be argued, restricted dominance can be helpfully understood by reference to poverty lines. Restricted poverty dominance (in relation to egalitarian opulence functions) could be defined to be y-restricted welfare dominance with $Z_2 \leq \theta^+$, but it would complicate the exposition unnecessarily to introduce it as a primitive category. Rather think of the criterion of restricted poverty dominance as having been subsumed by that of y-restricted welfare dominance. Note also that while I allow welfare dominance to be restricted along one of two dimensions, p or y, no such choice is given for equality dominance. One could deal with y-restricted equality functions, but since equality is so often thought about and measured in terms of shares accruing to various proportions, it is natural for the restrictions to be along the dimension of p.[42]

Over which sets Σ do these criteria guarantee dominance? To answer this question, three new assumptions on S need to be added to those of Section II. Assumption 4A is also reproduced for convenience.

4A S is insensitive to changes in income at or above Z^P: if $F(y)=F^*(y)$, $y<Z^P$, then $S(F)=S(F^*)$.

41. This corresponds to Atkinson's (1987) Condition II.

42. In terms of the above terminology, Atkinson and Bourguignon (1989) introduce the criteria of y-restricted and p-restricted welfare dominance, with only upper bounds on y and p. Their reasons for this are considered later in the sub-section.

4' S is insensitive to changes in income at or above p^P: if $Q(p)=Q^*(p)$ $p<p^P$ then $S(F)=S(F^*)$.

5 S is insensitive to **mean-preserving** changes in income at or below Z_{min}: if $F(y)=F^*(y)$, $y>Z_{min}$ and $\lambda=\lambda^*$, where λ (λ^*) is the mean of F (F^*) conditional on $y\leq Z_{min}$, then $S(F)=S(F^*)$.

5' S is insensitive to **mean-preserving** changes in income at or below p_{min}: if $Q(p)=Q^*(p)$, $p>p_{min}$ and $\lambda=\lambda$, where λ (λ^*) is the mean of F (F^*) conditional on $Q(p)\leq Q(p)_{min}$, then $S(F)=S(F^*)$.

(Note that 4' is the dual of 4A and 5' of 5.[43]) With these and the earlier assumptions the following three classes of functions can be defined:

1. A living-standards function which satisfies assumptions 1A, 2, 3A, 4' and 5' - for p_{min} and p^P such that $p_{min}\leq p^P$ - is a p-restricted welfare function.

2. A living-standards function which satisfies assumptions 1A, 2, 3A, 4A and 5 - for Z_{min} and Z^P such that $Z_{min}\leq Z^P$ - is a y-restricted welfare function.

3. A living-standards function which satisfies assumptions 1B, 3A, 4' and 5' - for p_{min} and p^P such that $p_{min}\leq p^P$ - is a restricted equality function.

The following results can now be stated:

Theorem 6 Iff y is income and $FD_LF^*(p_1,p_2)$ then $S(F)\geq S(F^*)$ $\forall S\in\Sigma$ and $S(F)>S(F^*)$ $\exists S\in\Sigma$, where Σ is the set of p-restricted welfare functions with $p_1\leq p_{min}\leq p^P\leq p_2$.

Theorem 7 Iff y is mean-normalized income and $FD_LF^*(p_1,p_2)$ then $S(F)\geq S(F^*)$ $\forall S\in\Sigma$ and $S(F)>S(F^*)$ $\exists S\in\Sigma$, where Σ is the set of restricted equality functions with $p_1\leq p_{min}\leq p^P\leq p_2$.

Theorem 8 Iff y is income and $FD_2F^*(Z_1,Z_2)$ then $S(F)\geq S(F^*)$ $\forall S\in\Sigma$ and $S(F)>S(F^*)$ $\exists S\in\Sigma$, where Σ is the set of y-restricted welfare functions with $Z_1\leq Z_{min}\leq Z^P\leq Z_2$.

Necessity is simply shown and follows from $-G(y_k)$, $Z_1\leq y_k\leq Z_2$, itself being a y-restricted

43. In fact, since F is right-continuous, it would make no difference if the strong inequality in 5 ($y>Z_{min}$) were replaced by a weak inequality. Similarly, since Q is left-continuous it would make no difference if the strong inequality in 4' ($p<p^P$) were replaced by a weak inequality.

welfare function and $\phi(p_k)$, $p_1 \leq p_k \leq p_2$, being a p-restricted welfare or equality function. Sufficiency proofs of these theorems follow the same path as that of the poverty theorems in III.1. If restricted dominance holds, then F and F^* can be manipulated to generate new distributions which give unrestricted 2OSD. The ranking of the new and original distributions will only coincide if Σ contains only those S which satisfy the new restricted dominance assumptions. The proofs follow below, but are given in detail only for the case of y-restricted welfare dominance.

First, notation. $G_i^{(*)}(y_k)$ is the value of the deficit curve for $F_i^{(*)}$ at y_k. Let $\rho = F(Z_{min})$, $\rho^* = F^*(Z_{min})$, and $\rho_i^{(*)} = F_i^{(*)}(Z_{min})$. λ is the mean of F conditional on $y \leq Z_{min}$, λ^* that of F^*, and $\lambda_i^{(*)}$ that of $F_i^{(*)}$.

Sufficiency proof for Theorem 8

1. Assume $FD_2F^*(Z_1,Z_2)$, and choose Z_{min} and Z^P so that $Z_1 \leq Z_{min} \leq Z^P \leq Z_2$. Generate F_1 from F and F_1^* from F^* so that $F_1^{(*)}(y) = F^{(*)}(y)$, $y < Z^P$ and $F_1^{(*)}(y) = 1$, $y \geq Z^P$. Then generate F_2 from F_1 and F_2^* from F_1^* by redistributing incomes below or equal to Z_{min}, maintaining $F_2^{(*)}(y) = F^{(*)}(y)$, $y > Z_{min}$ to give, if feasible, $G_2(y_k) \leq G_2^*(y_k)$, $y_k \leq Z_{min}$.

2. The second step shows that it is feasible to generate F_2 and F_2^* as described to give the result desired (i.e., $G_2(y_k) \leq G_2^*(y_k)$, $y_k \leq Z_{min}$).

2a. Due to (11) and assumptions governing the generation of $F_2^{(*)}$ which require $\rho_2^{(*)} = \rho^{(*)}$ and $\lambda_2^{(*)} = \lambda^{(*)}$, we have

$$\begin{aligned} G_2(Z_{min}) &= \rho_2(Z_{min} - \lambda_2) = G(Z_{min}) \\ G_2^*(Z_{min}) &= \rho_2^*(Z_{min} - \lambda_2^*) = G^*(Z_{min}) \end{aligned} \tag{18}$$

Combining (18) with the assumption of restricted dominance that $G(Z_{min}) \leq G^*(Z_{min})$ gives the result that $G_2(Z_{min}) \leq G_2^*(Z_{min})$.

2b. $G_2(Z_{min}) \leq G_2^*(Z_{min})$ ensures the redistribution is feasible. First, assume that $\lambda \geq \lambda^*$. Then the redistribution can be such that $F_2^{(*)}(\lambda^{(*)}) = \rho^{(*)}$ and $F_2^{(*)}(y) = 0$, $y < \lambda^{(*)}$. That is, everyone with initial income less than or equal to Z_{min} can be given the mean income of all those in this category. In this case, $G_2(y) \leq G_2^*(y)$, $y \leq Z_{min}$ follows from F_2 having minimum dominance (note that if $\lambda = \lambda^*$ then $\rho \leq \rho^*$). Second, assume $\lambda < \lambda^*$, which implies from (18) that $\rho < \rho^*$. In this case, give ρ of F^* income λ and $(\rho^* - \rho)$ income $Z = (\lambda^*\rho^* - \lambda\rho)/(\rho^* - \rho)$, where $\lambda^* < Z < Z_{min}$. Integration reveals this redistribution is feasible ($\lambda_2^*=\lambda^*$)[44]; it also results in F_2 having 1OSD over F_2^* up to Z_{min} and so it is certainly true that $G_2(y)\leq G_2^*(y)$, $y\leq Z_{min}$. Figure 4 illustrates this second case.

3. So $G_2(y_k)\leq G_2^*(y_k)$, $y_k\leq Z_{min}$. Also, recalling that F_2 and F_2^* are generated from the 'intermediate' distributions F_1 and F_1^*, $F_2(y)=F_2^*(y)=1$, $y\geq Z^P$. Finally, since F_2 and F_2^* ultimately are derived from F and F^* we have $F_2D_2F_2^*(Z_{min},Z^P)$. Putting all these together gives $F_2D_2F_2^*(\theta^+)$: F_2 has welfare 2OSD over F_2^*. Since y is income, $S(F_2)\geq S(F_2^*)$ for all S which satisfy assumptions 1A, 2 and 3A. Moreover if S satisfies assumption 4A, $S(F)=S(F_1)$ and $S(F^*)=S(F_1^*)$. If S satisfies 5, $S(F_1)=S(F_2)$ and $S(F_1^*)=S(F_2^*)$. Hence for all S satisfying 1A, 2, 3A, 4A and 5, $S(F)\geq S(F^*)$. Varying Z_{min} and Z^P between Z_1 and Z_2 (maintaining $Z_{min}\leq Z^P$) gives the result. The inequality will hold strongly for at least one S.■

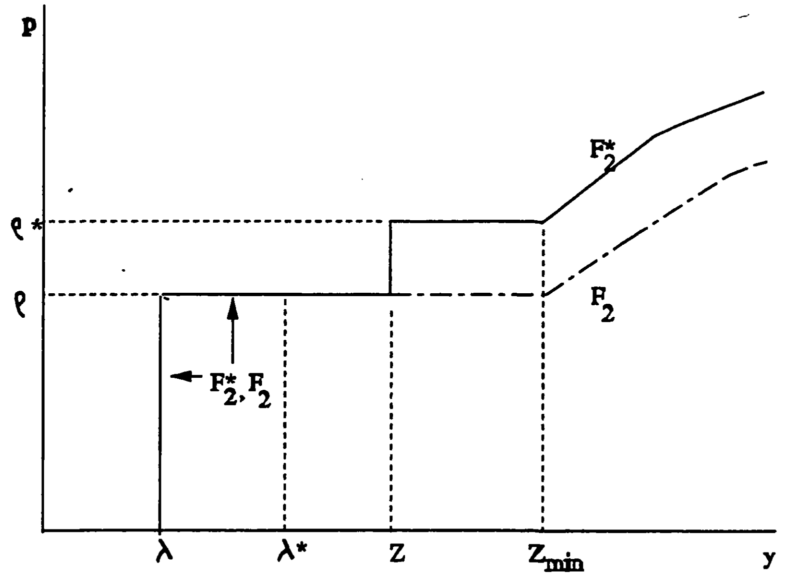
For equality and p-restricted welfare functions, the proofs follow basically by replacing Z by p. That is, assume $FD_LF^*(p_1,p_2)$ and choose p_{min} and p^P so that $p_1\leq p_{min}\leq p^P\leq p_2$. Clearly, however y is defined, assumptions 4' and 5' are sufficient to generate distributions, F_2 and F_2^* characterized by Lorenz dominance. Simply generate F_2 from F and F_2^* from F^* by replacing all y's above $Q^{(*)}(p^P)$ by y's equal to, say, $\max(Q(p),Q^*(p))$ and by redistributing all y's equal to or below $Q(p_{min})$ to give perfect equality of y's over this range (i.e. $Q_2^{(*)}(p)=\lambda^{(*)}$, $p\leq p_{min}$).

These restricted dominance criteria can be put in terms of well-known functions, at least for the cases of equality and y-restricted welfare. For equality, income share (for example, of the bottom 40%) is an oft-used measure. Restricted equality dominance is equivalent to dominance over all income shares between p_1 and p_2. y-restricted welfare dominance is equivalent to dominance by the (negative of the) poverty gap for all poverty lines between Z_1 and Z_2. No similarly simple interpretation is available for p-restricted welfare functions. This is no doubt related to the difficulty of giving a plausible interpretation to the p-restricted welfare assumptions. Say one is comparing two societies, one poor, one rich. It is unclear why the top x% of the poor society should not count in terms of welfare if their income level puts them below that of the top x% in the rich society. On the other hand, the more roughly similar the two societies being compared, the less important this objection, and bounds expressed in terms of p carry much more meaning to those unfamiliar with the distributions being compared than bounds in terms of y.

---

44. $\lambda_2^*=\dfrac{\rho\lambda+(\rho^*-\rho)Z}{p^*}=\dfrac{\rho\lambda}{\rho^*}+\dfrac{\rho^*-\rho}{\rho^*}\dfrac{\lambda^*\rho^*-\lambda\rho}{\rho^*-\rho}=\lambda^*$

Figure 4 Illustration of Step 2b in Sufficiency Proof of y-restricted Welfare Dominance Theorem



Notes: The graph shows F_2 and F_2^* up to Z_{min} on the assumption that $\lambda<\lambda^*$. From the definition of Z (given in the proof) and the assumption that $G(Z_{min})\leq G^*(Z_{min})$, $Z>\lambda^*$, since $\rho(\lambda^*-\lambda)>0$, and $Z\leq Z_{min}$, since, by assumption, $\rho(Z_{min}-\lambda)\leq\rho^*(Z_{min}-\lambda^*)$.

Both the poverty gap and the income share are cases in which, for any given function, the lower and upper bounds coincide. For the poverty gap $Z_{min}=Z^P$, and for the income share function $p_{min}=p^P$. However, these restricted criteria also cover those functions for whom these bounds do not coincide. An example is the following function, for $\eta>0$, $Z_{min}\geq1$, $\alpha\leq1$ and $\alpha\neq0$:

$$s(y) = \begin{cases} \frac{y}{Z_{min}}-1, & y\leq Z_{min} \\ \frac{1}{\alpha}[(\frac{y}{Z_{min}})^{\alpha}-1], & Z_{min}<y\leq Z^P \\ \frac{1}{\alpha}[(\frac{Z^P}{Z_{min}})^{\alpha}-1], & y>Z^P \end{cases} \tag{19}$$

More generally, the criteria cover, *inter alia*, all separable functions which are non-negatively sloped, concave and initially linear, then curved and finally flat.

The pairs of restrictions - 4 and 5, 4' and 5' - treat the two tails in quite different ways and so require different justifications. When Atkinson and Bourguignon introduced the idea of restricted dominance, they allowed only for restrictions at the right-hand tail, which removes the requirement of mean dominance in the case of welfare 2OSD and maximum dominance in that of equality. Imposing an upper bound is clearly an idea with which we are familiar. Ignoring all information about the rich, except for the fact that they are rich, is precisely what characterizes poverty analysis. Hence the labelling of the upper bounds as poverty lines, Z^P and p^P. (In turn, Z_2 and p_2 can be thought of as upper bounds on the range of poverty lines we are prepared to contemplate.) The removal of the requirement of mean dominance provides a normative justification for the imposition of an upper bound. The removal of the requirement of maximum dominance in the case of inequality analysis provides both a normative and a measurement-based justification.[45]

Lower bounds are more unusual than upper bounds. Indeed, Atkinson and Bourguignon would not advocate "similar restrictions at the lower end" since "there is no justification for a similar assumption of indifference at the lower end of the income scale" (p.12). However, as can be seen, restrictions on the lower tail are available which remove the requirement of minimum dominance but which by no means imply indifference. Z_{min} and p_{min} can be thought of poverty lines for the **very** poor (and, in turn, Z_1 and p_1 as lower bounds on the range we are prepared to consider of poverty lines for the very poor). Assumptions 5 and 5' do not remove from the very poor the veto power they are given by 2OSD: if their mean income is less or their poverty gap greater (depending whether the restrictions are along the dimensions of p or y) in one distribution than another, that distribution cannot dominate. What the assumptions do do is remove the veto power each group has (starting with that group, possibly an individual person or household, with the minimum income) over the successively richer group **within** this category of very poor. (A 'group' here is all those with an income below a certain level, whether defined in terms of y or p.) The application of 5 or 5' results in transfers within the group of very poor having no effect on the ranking. This in turn can be justified in two ways. First, from a normative perspective, one can think of there being some level of income or ranking in society which is so low that below which aggregate well-being is not improved by 'robbing poor Peter to pay poor Paul'. Second, from a measurement perspective, we can argue the grouping is justified due to our uncertainty in relation to very low incomes. To avoid attaching too much weight to them, we group.

Although this sub-section has focused on restricting the 2OSD criterion, it is also

45. Note that in the case of inequality analysis, it would make no difference if we replaced assumption 4', that S is insensitive to changes at or above p^P, by the assumption, say 4^*, that S is insensitive to mean-preserving changes at or above p^P. Since the overall mean using mean-normalized income is always equal to unity, and since assumption 4' allows no change in the conditional mean below p^P, there can be no change under 4' in the conditional mean above p^P. If y is income though there is a real difference between 4 and 4^*. In these cases, 4 is the more appropriate assumption since it but not 4^* removes the requirement of mean dominance.

possible to restrict the 1OSD criterion by requiring the distribution function be nowhere higher and somewhere lower between the two bounds Z_1 and Z_2. Indeed this requirement of dominance by the head-count ratio over a range of poverty lines has already been utilized in III.1.3 as a component of the mixed dominance criterion. It is important to bear in mind though that restricted 1OSD between Z_1 and Z_2 does not necessarily imply restricted 2OSD between the same bounds. However, a modified version of the chain of implication for unrestricted dominance does apply: if $FD_1F^*(Z_1,Z_2)$ and $G(Z_1)\leq G^*(Z_1)$ then $FD_2F^*(Z_1,Z_2)$. The additional condition, automatically satisfied in the unrestricted case of $Z_1=\eta^-$, is what provides those with an income less than or equal to Z_{min} with a veto over the ordering of F and F^*. Finally, one could also define a restricted mixed dominance criterion by which one required 2OSD between Z_1 and Z_2 and 1OSD between Z_2 and Z_3. Assuming that y is income, this would then give dominance in relation to all functions satisfying 1A, 2, 3B, 4B and 5 with $Z_1\leq Z_{min}\leq Z^P$ and $Z_2\leq Z^P\leq Z_3$.

IV.3 E-dominance

An alternative means of avoiding the requirements of the extreme forms of dominance is, as mentioned earlier, to leave the stochastic dominance framework altogether and consider explicit restrictions on the functional forms of the members of Σ. To provide a simple illustration of this approach, assume that S is separable and that Σ contains only welfare functions which display constant relative inequality aversion. Then, as stated in II.2, S is restricted to being a monotonic transform of the isoelastic function (given in Table 2):[46]

$$\begin{aligned} \epsilon &= \frac{1}{\alpha}\int_{\eta}^{\theta} y^{\alpha}dF(y) \quad \alpha\neq 0 \\ &= \int_{\eta}^{\theta} \ln(y)dF(y) \quad \alpha=0 \end{aligned} \tag{20}$$

where, as in II.4, $e=1-\alpha\geq 0$ and gives the degree of inequality aversion. If, in addition to assuming constant relative inequality aversion, one takes a welfare-based view of the cost of inequality, one is restricted to (20) (with y defined as mean-normalized income) for the measurement of inequality

46. This means Σ is defined to be a subset of the third-order stochastic dominance set, as s(y) will have a non-negative third derivative.

(see again II.4).[47] For poverty analysis, the generalized Clark function, given by (2), is consistent with constant relative inequality aversion up to the poverty line.

If one combines restriction to (20) with bounds on the range of values, and thus α, can take, the choice of the set, Σ, is reduced, in the case of welfare and inequality analysis, to a choice of minimum and maximum values of e. One can then search for **e-dominance.** Of course, the higher the value of e, the higher the degree of inequality aversion. Typical values of e found in the empirical literature range between 0 and 5 (see, for example, Ahmad and Stern, 1991, who also discuss a number of ways by which values of e might be chosen). As is evident from a glance back to Table 2, finite values of e avoid the requirement of minimum dominance, while strictly positive avoid that of mean dominance. Maximum dominance in the case of equality analysis is avoided by the isoelastic function having a positive third derivative (see again the discussion following Table 2). In the case of opulence analysis, apart from restricting the concavity of s(y) up to the poverty line, by choice of range of e, one can also restrict the range of the poverty line and the extent of the discontinuity at the poverty line. Using the generalized Clark function, all three restrictions can be parameterized, by choice of individual values for or ranges of e, Z^P and C.

This approach certainly has simplicity to recommend it. It is not at all new. Indeed, Atkinson's original 1970 article contained a diagram showing how the rankings of a set of distributions (in his case, countries) changed as e was varied. However, it has not been much applied in empirical work, as authors have preferred to give rankings for one or two values of e (though Ahmad and Stern, 1991, using five values of e is one exception; Howes and Lanjouw, 1991, and Lanjouw, 1992, provide expositions and applications of this methodology).[48]

One of the possible criticisms which can be made against the e-dominance criterion is

---

47. Alternative functional forms could also be used, for example, the generalized entropy function. See the discussion following Table 2 for a discussion of the pros and cons of using the GE and the isoelastic function.

48. A related criterion is presented by Lambert (1989) and Dardanoni and Lambert (1988) who give conditions under which a distribution will have dominance for a Σ which contains all separable welfare functions with a degree of relative inequality aversion greater than some minimum bound. The difficulty with this is that it requires minimum income dominance, the problems associated with which have already been discussed. Jewitt (1981) gives the conditions, in terms of crossings of the distribution functions, under which dominance over constant absolute and relative inequality aversion functions implies dominance over the respectively wider classes of non-increasing absolute and non-decreasing relative inequality aversion functions. But note that his results hold given the widest bounds on the degree of constant inequality aversion, whether absolute or relative, namely from zero to infinity, whereas in this sub-section we want to allow for bounds to be placed on the degree of inequality aversion.

the arbitrariness of bounds. One way to avoid this criticism would be to simply aim to report the 'switch point' values of e, those values of e, if they exist, at which $S(F)=S(F^*)$. Then readers would be free to put their own value-judgements alongside the information provided, and make up their own minds as to which distribution dominates. Locating all the switch points of e may be thought to be very difficult, but in fact one can show that there will often be at most one switch-point. Let ε and ψ be special cases of a general function S_e. Then, as proved in Appendix A,

Theorem 9 If $y>0$ and $F(y)$ and $F^*(y)$ cross only once then $\Delta S_e=S(F,e)-S(F^*,e)$ will change sign at most once as e increases from zero.

Say that F and F^* cross once and that, at low values of y, F lies below F^*, so that, for a high enough value of e, F will be preferred. As e falls, more weight is given to the right-hand portion of the distribution and F^* has a greater chance of being preferred. Note that if y is mean-normalized income and F and F^* only cross once then ΔS_e will not change sign (since one distribution will have 2OSD) which is consistent with, but stronger than Theorem 9.

Since many empirical distributions functions cross more than once, this result may not be thought very useful. However, although the sample distribution functions may cross many times, they may nevertheless come from populations whose distribution functions cross only once (such as the 'two-parameter' family, including the normal, log-normal and gamma). If so, the sample e-dominance curves will often cross only once. (See Appendix B of Chapter 3 for evidence.)

Another criticism of the e-dominance approach is that S_e may only qualify as a standard-of-living function if y is strictly positive. If y is zero, S_e will be defined only if $e<1$. y being negative will result in S_e being undefined if e is not equal either to zero or an integer greater than one. If y is negative and e is an integer greater than one, e is either even in which case S_e has a positive first and second derivative or e is odd, in which case both derivatives are negative. Hence if one has negative income values, and one wishes to use non-zero values of e, one can only use S_e as an equality function, and then only with odd numbered integers greater than one! And if one has both zero and non-negative values one cannot use S_e as an equality function at all (except for the trivial case in which $e=0$). See Anand (1983) for a further discussion of this issue. The only comfort available is that if y is a suitable measure of living-standards it should not be negative. Negative income, for example, cannot be sustained over the long-run. But this is little reassurance if one actually does have a sample containing negative income values.

V Conclusion

This chapter has provided a framework which can be used to understand the differences and similarities between welfare, equality and poverty analysis. Within this framework, various criteria for dominance - some old, some new - have been presented, and the relationship between the criteria analyzed. A distinction has been drawn between those criteria which rank distributions only if various forms of extreme dominance hold and those, such as restricted dominance and e-dominance, which are not thus dependent. The latter are, to use Atkinson and Bourguignon's words - from a different though related context - "less stringent and more realistic criteri[a]" (1989, p.7).

The dominance criteria analyzed in this chapter can also be divided up another way. Recall that these criteria, if they rank two distributions, ensure dominance in relation to some set, Σ, that is, weak preference by all S in Σ and strong preference by at least one S. Most of the criteria analyzed refer to some 'crucial' subset of Σ, dominance in relation to which is a necessary and sufficient condition for dominance in relation to Σ. The negative of the distribution function, the negative of the deficit curve and the Lorenz curve, evaluated at different values of y or p, can all be thought of as different S's. Sets of these functions - evaluated over a range of values of y or p - constitute 'crucial' subsets of Σ. The e-dominance criterion on the other hand refers directly to the relevant Σ. Welfare e-dominance between two bounds means, by definition, dominance in relation to the set of isoelastic functions determined by these bounds but not necessarily dominance in relation to any wider set. Neither of these two types of criteria is intrinsically preferable to the other. But the distinction is a helpful one in understanding the way in which dominance criteria work.

Appendix A Proof of Theorem 9

It is shown that if y>0, and if F(y) and $F^*(y)$ cross only once then $\Delta S_e=S(F,e)-S(F^*,e)$ will change sign at most once as e increases from zero. This is first proved for the case in which S_e is a welfare function. Applying integration by parts to the formula for S_e, given in (20), one can, noting that S_e is differentiable, write

$$\Delta S_e = -\int_{\eta^-}^{\theta^+} s'(y,\alpha)\Delta F(y)dy \text{ where } \Delta F(y)=F(y)-F^*(y) \qquad \text{(A.1)}$$
$$s'(y,\alpha)=y^{\alpha-1}$$

We want to investigate the sign of $\partial\Delta S/\partial e$, which, from (A.1), is given by

$$\frac{\partial \Delta S_e}{\partial e}=\int_{\eta^-}^{\theta^+}\frac{\partial s'(y,\alpha)}{\partial e}\Delta F(y)dy \qquad \text{(A.2)}$$

Since e=1-α,

$$\begin{aligned}\frac{\partial s'(y,\alpha)}{\partial e} &= \frac{\partial(\alpha-1)}{\partial e}\ln(y)*s'(y,\alpha) \\ &= -\ln(y)s'(y,\alpha)\end{aligned} \qquad \text{(A.3)}$$

Substituting (A.3) into (A.2) gives

$$\frac{\partial \Delta S_e}{\partial e} = \int_{\eta^-}^{\theta^+} s'(y,\alpha)\ln(y)\Delta F(y)dy \qquad \text{(A.4)}$$

Since ΔS_e is unit invariant and since y has already been assumed to be positive, we can assume without further loss of generality that the two distributions are such that $\eta^-\geq 1$ so that $\ln(y)\geq 0$ $\forall y\geq\eta^-$. Hence the sign of (A.4) depends on ΔF. Assume F and F^* cross only once, at ξ, and, without loss of generality, choose the labels F and F^* so that, for some ξ, if $\eta^-\leq y\leq\xi$, $\Delta F(y)\leq 0$, and, if $\xi\leq y\leq\theta^+$, $\Delta F(y)\geq 0$. Then (A.4) can be re-written as

$$\frac{\partial \Delta S_e}{\partial e} = \int_{\eta^-}^{\xi} s'(y,\alpha)\ln(y)\Delta F(y)dy + \int_{\xi}^{\theta^+} s'(y,\alpha)\ln(y)\Delta F(y)dy \qquad \text{(A.5)}$$

the first term of which is by assumption positive, the second by assumption negative. If there is some e, $\bar{e}^+$, at which $\Delta S_e=0$ then, from (A.1),

$$-\ln(\xi)\Delta S_{e^+} = \ln(\xi)\left(\int_{\eta^-}^{\xi} s'(y,\alpha^+)\Delta F(y)dy\right) + \ln(\xi)\left(\int_{\xi}^{\theta^+} s'(y,\alpha^+)\Delta F(y)dy\right) = 0 \tag{A.6}$$

By the definition of integration,

$$\int_{\eta^-}^{\xi} s'(y,\alpha)\ln(y)\Delta F(y)dy \le \ln(\xi)\int_{\eta^-}^{\xi} s'(y,\alpha)\Delta F(y)dy \tag{A.7}$$

$$\int_{\xi}^{\theta^+} s'(y,\alpha)\ln(y)\Delta F(y)dy \le \ln(\xi)\int_{\xi}^{\theta^+} s'(y,\alpha)\Delta F(y)dy \tag{A.8}$$

(recalling that both the left- and right-hand sides of (A.8) are negative). From (A.7) and (A.8), both terms in (A.5) are smaller than their respective counterparts in (A.6), so it must be the case that

$$\text{if } \Delta S_e = 0, \text{ then } \frac{\partial \Delta S_e}{\partial e} \le 0. \tag{A.9}$$

Since the derivative is always negative at $\Delta S_e=0$, there can be no more than one point at which $\Delta S_e=0$.

Now consider the case in which S is an opulence function, so S_e is as defined in (2). If the distribution functions cross above Z^P then S_e will not change sign, so consider the case in which $\xi \le Z^P$. The same arguments applied above can be used to show that if the distribution functions cross at most once, the generalized Clark function will change sign at most once as e changes. Note that in this case the derivative is the same as given in (A.4) except that it should be multiplied by (1-C), y should be replaced by y/Z^P and θ^+ by Z^P (see equation (3)). $\ln(y/\underline{Z}^P)$ is of course always negative for $y \le Z^P$, but this in fact leaves the signs in (A.7) and (A.8) unchanged. If $C \neq 0$, then (A.6) must have added to its right hand side $C\Delta F(Z^P)\ln(\xi)$ (see (2)), but since this is positive by assumption, it makes no difference to the derivation of (A.9). This method of proof cannot be applied to equality functions because, for these, ln(y) will change sign as y, mean-normalized income, will be less than 1 for some portion of the distribution and greater for the rest. However, as stated in the text, if y is mean-normalized income and F and F^* cross only once then one of the distributions will have 2OSD and thus e-dominance.

Chapter Two Estimators and methods of inference

I Introduction

As defined in the previous chapter, the problem of distributional dominance is that of ordering distributions in relation to a dominance criterion: if and only if the criterion ranks the two distributions can the one distribution be said to dominate the other, that is, be reckoned to be no worse than the other by all the living-standard functions in the relevant set, Σ, and better by at least one in Σ. The case for engaging in statistical analysis of dominance is twofold. Very often we are in a situation where we have two sets of sample data and wish to use them to draw comparative conclusions about the populations from which they are drawn. To do so requires the use of statistical techniques. In addition, general tests for distinguishing distributions are not sufficient for the statistical inference of dominance. It is of course useful to know if two distributions have significantly different means and variances, and also, more generally, whether the null that they are drawn from the same underlying distribution can be rejected, but this will not tell us whether or not we can infer dominance. If we can parameterize the distribution then analytical techniques can be used to infer dominance, but our conclusions will then be subject to the proviso that the parameterization is the correct one. The alternative path, pursued in this chapter, is to provide estimators and tests which are non-parametric and by which dominance can be directly tested.

Since the dominance criteria presented in Chapter One can be represented graphically in terms of curves, the job of testing for dominance becomes one of investigating differences between curves and asking whether these are significant. Define $\zeta_i=\zeta(F,x_i,\Sigma)$ and $\zeta_i^*=\zeta(F^*,x_i,\Sigma)$ to be the 'dominance curves' of the distribution functions, F and F^*, for given set Σ and evaluated at x_i. Then F will be said to dominate F^* by the set Σ iff $\zeta_i\geq\zeta_i^*$ $\forall x_i$ and $\zeta_i>\zeta_i^*$ $\exists x_i\in[x_{min},x_{max}]$. For future reference, I also define the notion of 'strong dominance'. F will be said to strongly dominate F^* iff $\zeta_i>\zeta_i^*$ $\forall x_i\in[x_{min},x_{max}]$ such that it is not the case that $\zeta_i\equiv\zeta_i^*$.

The conclusion of the previous chapter drew a distinction between two different types of criteria. In the present context, the distinction can be put as follows. The simplest case, the latter of the two types in Chapter One's conclusion, is that in which x is some parameter which defines the S in Σ. For example, if the criterion is e-dominance (see Chapter One, IV.3), x_i will be the value of the inequality aversion parameter, e, and ζ_i the isoelastic function for that value. The other case is that in which x_i is a parameter which defines the S in some 'crucial' subset of Σ such that dominance over the subset is a necessary and sufficient condition for dominance over

Σ. For example, for poverty second-order stochastic dominance, x_{min} would equal the minimum income of the two distributions, x_{max} the upper bound on the set of reasonable poverty lines, and ζ_i the *negative* of the deficit curve at some income level x_i. Note that where, as in this example, dominance requires that one distribution generate a lower value for some curve than another distribution then ζ_i represents the negative of that curve. This simplifies notation.

These dominance curves are defined in terms of a continuous interval $[x_{min},x_{max}]$, but tests can only be carried out over a finite number of ordinates. So make a discrete approximation and evaluate ζ_i at x_i, i=1 to W, where $x_1 \geq x_{min}$ and $x_W \leq x_{max}$. It is assumed that the conclusion which follows from an assessment of ζ_i and ζ_i^* over this finite set of ordinates will be the same as that which would follow from an assessment over the interval $[x_{min},x_{max}]$. So F will be said to dominate F^* by the set Σ (defined in relation to some x_{min} and x_{max}) iff $\zeta_i \geq \zeta_i^*$ $\forall i$ and $\zeta_i > \zeta_i^*$ $\exists i$. F will be said to strongly dominate F^* iff $\zeta_i > \zeta_i^*$ $\forall i$ such that it is not the case that $\zeta_i \equiv \zeta_i^*$. The question of how to make this discrete approximation is discussed in III.4.

To infer whether one curve lies nowhere below and somewhere above another requires a vector of estimators and, depending on the test employed, the variance of each element in the vector or the vector's variance-covariance matrix. The task of Section II is to derive estimators for different definitions of ζ_i - referred to as estimators for different dominance criteria (strictly curves) - and to derive in each case the typical element of the variance-covariance matrix of the vector of estimators. The task of Section III is to provide a test which utilizes these vectors of estimators and their variances to provide a test for dominance.[1]

The dominance criteria for which estimators are provided are those of Chapter One: the stochastic dominance criteria (first-order, second-order and mixed) and the e-dominance criterion (including the use of generalized Clark function for poverty analysis). For both the e-dominance and the second-order stochastic dominance criteria, estimators are given for the case in which income is mean-normalized, so that equality, as well as welfare and poverty, analysis can be conducted. For the second-order stochastic dominance criterion, estimators are given for both the deficit and generalized Lorenz curves so that either can be used. The last chapter introduced the distinction between restricted and unrestricted forms of stochastic dominance. This simply involves the setting of bounds (choice of x_{min} and x_{max}) and both forms are catered for by the testing method of Section III. Finally, all of the estimators given can be used in the presence of randomly weighted data. So they can be used if, as often happens, one has data at the household level but

1. Although only variances are required in the testing method of Section III, for the sake of generality Section II gives the typical element of the various variance-covariance matrices.

wishes to make inferences based on aggregation over individuals, not households.

Statistical testing of dominance relationships between distributions has by no means been ignored in the literature. Some of the estimators required are already available, though only for the special case in which observations are unweighted. These are referred to in the course of Section II. And a number of testing methods have been proposed. These are critically surveyed in Section III alongside the simple alternative put forward.

Although the e-dominance criterion is also considered, much of the chapter's attention is devoted to the family of stochastic dominance criteria. Section IV of Chapter One noted the reliance of unrestricted stochastic dominance criteria on the requirements of 'extreme dominance' (for example, the requirement that the dominating distribution have a no lower minimum income) and criticized this reliance from a normative perspective. Section IV of this chapter examines the inferential problems associated with the requirements of extreme dominance. Finally, Section V concludes.

II Estimators for the statistical analysis of dominance

The derivation of estimators follows a step-by-step procedure. An initial framework which provides the basis for the derivation of (welfare and equality) e-dominance estimators and first-order stochastic dominance (1OSD) estimators is presented in II.1. The special case of e-dominance is then presented in II.2 and that of 1OSD in II.3. II.4 uses II.3's estimators to derive estimators for the poverty deficit curve for the testing of welfare and poverty second-order stochastic dominance (2OSD). II.5 shows how II.4's estimators can in turn be used to derive generalized Lorenz and Lorenz curve estimators utilizable for welfare and-equality 2OSD analysis. II.6 compares the estimators of II.4 and II.5 and asks whether the deficit and Lorenz curve estimators will lead to identical inferences being made. Finally, II.7 uses the derivations of II.3 to present estimators for the generalized Clark poverty function (for use in poverty e-dominance analysis).

One important feature of the estimators provided is that they can be used to analyze weighted data. As Cowell (1989) points out, data will often only be available at the household level, even though it is the individual which is of ultimate interest to the researcher. However, up to now weighted estimators have been largely unavailable for dominance analysis.

The distributions from which one is sampling are, therefore, assumed to be bivariate: each observation can be considered to be a random (though not necessarily independent) combination of a weight, which we will call 'household size', and a living-standards indicator, which will be called 'income'. The random weight will, in the exposition, be assumed to be a discrete variable, though the estimation formulae would be unchanged if it were instead continuous. It is denoted by h_r, $r=1,...,R$. The income variable will be denoted y. As in Chapter One, unless specified to the contrary, the distribution of y may be continuous, discrete or mixed. y may be household income or income per capita (household income divided by household size) or equivalent income, subject to the proviso given above. If $h_r=1$ $\forall r$, we have the special case of no weighting, so the analysis is of the household distribution of income, whether household or per capita.[2]

h_r need not only be interpreted as household size. The method presented will also be consistent with the use of equivalence scales to derive equivalent income and/or weights, though not if the scales are themselves regarded as estimates and so subject to their own sampling error. Since this latter is in fact likely to be the case, the variance-covariance formulae presented should be regarded as simplifications of the true variance-covariance formulae if equivalence scales are used.

As in the last chapter, distributions are denoted by their distribution function. But now F, for example, refers to the population distribution, from which a sample of size N is drawn. The remaining notation generally follows the convention that the vector to be estimated (the special case of the vector $\zeta=(\zeta_1,...,\zeta_W)$) is denoted by a Greek or English lower-case letter (e.g., ϕ), the associated asymptotic covariance matrix by the same letter in upper case (e.g., Φ), and the typical element of that matrix by attaching sub-scripts to the upper-case letter (e.g. Φ_{ij}). An exception to this is the vector of deficit curve ordinates. Since the deficit curve is the integral of the distribution function, F, the vector is denoted by G, and its estimator's covariance matrix by Γ. Also, the notation for the generalized Clark function make use of superscripts as well as subscripts. Consistent estimators are denoted by caps, except where specified to the contrary. Sometimes the same notation is used for a function (e.g., $p=F(y)$) and for a vector ($p=(p_1,...,p_W)$). However, the meaning should be clear from the context.

2. This also covers the case in which data is directly available at the individual level, in which case the terminology can be stretched to equate 'household' and 'individual'. Note too that, unlike in the previous chapter, in this chapter y never corresponds to mean-normalized income.

II.1 Initial framework

The method of derivation in this sub-section follows Cowell (1989). As is shown in the next two sub-sections, estimators for both e-dominance and first-order stochastic dominance can be based on p_i, i=1,...,W, given by

$$p_i = p(\alpha_i,\gamma_i) = \frac{\mu_i}{\mu_h^{\gamma_i}\mu_y^{1-\gamma_i}} \tag{1}$$

where

$$\begin{aligned}
&\mu_i=\mu(\alpha_i) = \sum_{r=1}^{R} h_r\int_{\eta}^{\theta} g(y,\alpha_i)dF_r(y) \quad i=1,...,W,h,y \\
&\sum_{r=1}^{R}\int_{\eta}^{\theta} dF_r(y)=1 \\
&g(y,\alpha_h)=1 \\
&g(y,\alpha_y)=y
\end{aligned} \tag{2}$$

$F_r(y)$ is the proportion of households with income less than or equal to y and household size h_r. $g(y,\alpha_i)$, i=1 to W, is simply a function relating y and α; μ_i is its weighted mean. It has no general meaning, but is used to represent two different functions for the respective cases of e-dominance and 1OSD: this will become clear in the following two sub-sections. μ_h gives the mean population weight, and μ_y the weighted mean of y.

Using vector notation, let $\mu'=(\mu_1,...,\mu_W,\mu_h, \mu_y)$, $p'=(p_1,...,p_W)$ and $p=p(\mu)$. Define m and $\hat{p}$ analogously to be consistent estimators of μ and p respectively. Then a theorem of Rao's (1973, p.387) tells us that, since p is a transform of μ, $\sqrt{N}(\hat{p}-p)$ is asymptotically normally distributed with mean zero and covariance matrix P given by

$$P = p_\mu M p_\mu' \tag{3}$$

M is the asymptotic covariance matrix of $\sqrt{N}(m-\mu)$ and is W+2 by W+2:

$$M=\begin{bmatrix} M_{11} & \cdots & M_{1W} & M_{1h} & M_{1y} \\ . & \cdots & . & . & . \\ M_{W1} & \cdots & M_{WW} & M_{Wh} & M_{Wy} \\ M_{h1} & \cdots & M_{hW} & M_{hh} & M_{hy} \\ M_{y1} & \cdots & M_{yW} & M_{yh} & M_{yy} \end{bmatrix} \tag{4}$$

p_μ is an W-by-W+2 matrix of partial derivatives defined as

$$p_\mu=\frac{\partial p}{\partial \mu}=\begin{bmatrix} p_{11} & \cdots & p_{1W} & p_{1h} & p_{1y} \\ . & \cdots & . & . & . \\ p_{W1} & \cdots & p_{WW} & p_{Wh} & p_{Wy} \end{bmatrix} \tag{5}$$

Note that for i,j=1, ..., W, and $i \neq j$, $p_{ij}=0$ as, from (1), p_i is independent of μ_j. Combining this result with (3), (4) and (5) gives the typical element of matrix P, P_{ij}, as:

$$\begin{aligned} P_{ij} &= p_{ii}(p_{jj}M_{ji}+p_{jh}M_{hi}+p_{jy}M_{yi})+ p_{ih}(p_{jj}M_{jh}+p_{jh}M_{hh}+p_{jy}M_{yh}) \\ &+ p_{iy}(p_{jj}M_{jy}+p_{jh}M_{hy}+p_{jy}M_{yy}) \end{aligned} \tag{6}$$

To show how this expression can be estimated, one first needs the partial derivatives, available from (1)

$$\begin{aligned} p_{ii} &= \mu_h^{-\gamma_i}\mu_y^{\gamma_i-1},\ i=1,\dots,W \\ p_{ih} &= (-\gamma_i)\mu_i\mu_h^{-\gamma_i-1}\mu_y^{\gamma_i-1},\ i=1,\dots,W \\ p_{iy} &= (\gamma_i-1)\mu_i\mu_h^{-\gamma_i}\mu_y^{\gamma_i-2},\ i=1,\dots,W \end{aligned} \tag{7}$$

The definition of a covariance gives

$$\begin{aligned} &M_{ij} = \mu_{ij} - \mu_i\mu_j \\ &\text{where } \mu_{ij} = \sum_{r=1}^{R} h_r^2\int_{\eta}^{\theta} g(y,\alpha_i)g(y,\alpha_j)dF_r(y),\quad i,j=1,\dots,W,h,y \end{aligned} \tag{8}$$

Substituting (7) and (8) into (6) gives

$$\frac{P_{ij}}{\left(\mu_h^{-\gamma_i}\mu_y^{\gamma_i-1}\right)\left(\mu_h^{-\gamma_j}\mu_y^{\gamma_j-1}\right)\mu_i\mu_j} = \frac{\mu_{ij}}{\mu_i\mu_j}-1$$
$$- \gamma_j\left(\frac{\mu_{hi}}{\mu_h\mu_i}-1\right) - \gamma_i\left(\frac{\mu_{hj}}{\mu_h\mu_j}-1\right) + (\gamma_i\gamma_j)\left(\frac{\mu_{hh}}{\mu_h^2}-1\right)$$
$$+ (\gamma_j-1)\left(\frac{\mu_{yi}}{\mu_y\mu_j}-1\right) + (\gamma_i-1)\left(\frac{\mu_{yj}}{\mu_y\mu_i}-1\right) \tag{9}$$
$$- \left((\gamma_i-1)\gamma_j + (\gamma_j-1)\gamma_i\right)\left(\frac{\mu_{yh}}{\mu_h\mu_y}-1\right) + (\gamma_i-1)(\gamma_j-1)\left(\frac{\mu_{yy}}{\mu_y^2}-1\right)$$

This equation is the key one for our purposes. Asymptotically valid tests can be constructed using consistent estimators of the elements of equations (1) and (9). As throughout this section, the estimators required are all functions of sample means (sometimes of transformed variables), and so are consistent. An estimator of (1) is given, for a sample of size N, by

$$\hat{p}_i = \frac{m_i}{m_h^{\gamma_i} m_y^{1-\gamma_i}} \quad \text{where } i=1,\ldots W \tag{10}$$

where

$$m_i = m(\alpha_i) = \frac{1}{N}\sum_{n=1}^{N} h_n g(y_n,\alpha_i), \quad i=1,\ldots,W,h,y \tag{11}$$

In addition, to estimate (9) one needs a consistent estimator of μ_{ij} given by

$$m_{ij} = \frac{1}{N}\sum_{n=1}^{N} h_n^2 g(y_n,\alpha_i) g(y_n,\alpha_j), \quad i,j=1,\ldots,W,h,y \tag{12}$$

Note that in writing these means I have, out of convenience, not divided the sample into different household types, but instead simply attached a different weight to each household. There are thus by this notation not just R weights, but rather N, though, of course, these N weights can only take on R different values.

II.2 E-dominance: welfare and equality analysis

Continuing on within the framework established above, make the additional assumptions that for all i, i=1,...,W:

(a) $g(y,\alpha_i)=y^{\alpha i}$, $\alpha_i\neq0$; $g(y,\alpha_i)=\log(y)$, $\alpha_i=0$; where $\alpha_i=1-e_i$[3]

and (b.1) $\gamma_i=1$ $\forall i$, $\delta=0$: welfare analysis

or (b.2) $\gamma_i=1-\alpha_i$ $\forall i$, $\delta=1$: equality analysis

Then the isoelastic function, for some value e_i, is given by

$$\begin{aligned} \varepsilon_i &= \frac{1}{\alpha_i}p_i, \quad i=1,\ldots,W, \quad \alpha_i\neq0 \\ &= \tilde{p}_0, \quad \alpha_i=0 \text{ where } \tilde{p}_0=\frac{\mu_0}{\mu_h}-\delta\log\left(\frac{\mu_y}{\mu_h}\right) \end{aligned} \tag{13}$$

Defining ζ_i to be ε_i enables one to analyze welfare and equality e-dominance in relation to the bounds e_{min} and e_{max}. (For more details, see Chapter One, IV.3. For the case of poverty analysis, see II.7 of this chapter.) $\sqrt{N}(\hat{\varepsilon}-\varepsilon)$ is asymptotically normally distributed with mean zero and variance covariance matrix, E, with typical element

$$\begin{aligned} E_{ij} &= \frac{1}{\alpha_i\alpha_j}P_{ij}, \quad \alpha_i,\alpha_j\neq0 \\ &= \frac{P_{ij}}{\alpha_j}, \quad \alpha_i=0, \ \alpha_j\neq0 \\ &= P_{ij}, \quad \alpha_i=\alpha_j=0 \end{aligned} \tag{14}$$

where P_{ij} is defined in (9) except in the single case of equality analysis using $\alpha_i=0$. In this case, the correct covariance formula is obtained by replacing γ_i on the right hand side of (9) by $\gamma_i^0=1-\mu_h/\mu_i$ (similarly if $\alpha_j=0$).[4] P_{ij} and p_i can be estimated using (10) to (12), substituting in the relevant combination of assumptions (a) and (b.1) or (b.2) above.

These estimators can also be used for the estimation of the Atkinson inequality index and the generalized entropy index. Denote the Atkinson inequality index for some particular value of α, α_i, as a_i, where i=1, ..., W. Then $\sqrt{N}(\hat{a}-a)$ is asymptotically normally distributed with mean zero and covariance matrix A, where

3. So μ_i, i=1, ..., W, is the α_ith moment of the random variable y. It is assumed throughout that the moments exist.

4. This follows from (13). The partial derivatives of $\tilde{p}_0$ with respect to μ_0, μ_h and μ_y are given by (7) if γ_i is replaced, other than when it appears as a power (superscript), by γ_i^0 as defined in the text.

$$\begin{aligned}
a_i &= 1 - p_i^{\frac{1}{\alpha_i}},\ \alpha_i \neq 0 \\
&= 1-\exp(\tilde{p}_0),\ \alpha_i = 0 \\
A_{ij} &= \alpha_i^{-1}\alpha_j^{-1} p_i^{\frac{1}{\alpha_i}-1} p_j^{\frac{1}{\alpha_j}-1} P_{ij},\ \alpha_i,\alpha_j \neq 0 \\
&= \alpha_j^{-1} p_j^{\frac{1}{\alpha_j}-1} \exp(\tilde{p}_0) P_{ij},\ \alpha_i = 0, \alpha_j \neq 0 \\
&= \exp(2\tilde{p}_0) P_{ii},\ \alpha_i = \alpha_j = 0
\end{aligned} \tag{15}$$

This result extends those of Nygård and Sandström (1981) and Thistle (1990) dealing with unweighted data.

Denote the generalized entropy measure for some value of α, α_i, as τ_i. where $i=1, ..., W$. Then again using the definitions of this sub-section $\sqrt{N}(\hat{\tau}-\tau)$ is asymptotically normally distributed with mean zero and covariance matrix T, where

$$\begin{aligned}
\tau_i &= \frac{1}{\alpha_i^2-\alpha_i}(p_i-1),\ \alpha_i \neq 0,1 \\
&= -\tilde{p}_0,\ \alpha_i = 0 \\
T_{ij} &= \left(\frac{1}{\alpha_i^2-\alpha_i}\right)\left(\frac{1}{\alpha_j^2-\alpha_j}\right) P_{ij},\ \alpha_i,\alpha_j \neq 0,1 \\
&= -\frac{P_{ij}}{\alpha_j^2-\alpha_j},\ \alpha_i = 0, \alpha_j \neq 0,1 \\
&= P_{ii},\ \alpha_i = \alpha_j = 0
\end{aligned} \tag{16}$$

This result synthesizes that of Cowell (1989) who gives the variance in the case of weighted data and that of Thistle who gives the full covariance structure (again for the case of equality) but on the assumption of unweighted data. See Cowell for the case in which $\alpha_i=1$: in this case, the generalized entropy index, unlike the isoelastic function, does not display constant relative inequality aversion (see Chapter One, II.4), and so requires a separate treatment.[5]

Finally, for future reference, note that under the assumptions made at the start of the sub-section, if $\alpha_1=1$, $\mu_1=\mu_y$. Hence if $\gamma_i=1$ $\forall i$, $p_1=\mu_y/\mu_h$ and P_{11} becomes N times the asymptotic variance of m_y/m_h. Denote p_1 under these assumptions by β - the per capita mean of y - and P_{11} by B[6]. Let $b=m_y/m_h$. Then from (9) the asymptotic variance of b is

5. Note that the generalized entropy and Atkinson indices are weakly increasing in inequality, whereas the isoelastic function is weakly decreasing. If $\delta=0$, $\delta=1$ and $\alpha<1$, the negative of the generalized entropy function can be used for welfare analysis (see Chapter One, Table 2).

6. Although I refer to β as the *per capita* mean of y, its actual definition will depend on that of h.

$$\frac{B}{\mu_{hh}\mu_h^{-2}} = \beta_y - 2\beta_h\beta + \beta^2 \quad \text{where } \beta_h = \frac{\mu_{yh}}{\mu_{hh}}, \quad \beta_y = \frac{\mu_{yy}}{\mu_{hh}} \tag{17}$$

Note that β_h can be interpreted as the 'per capita-squared' mean of y, and β_y as the 'per capita-squared' mean of y squared, since both use h^2 rather than h as the random weight.

II.3 First-order stochastic dominance

Again start with equation (1), but this time make the additional assumptions that, for i=1,...,W,

(a) $g(y,\alpha_i)=1$ for $y \le \alpha_i = z_i$

$=0$ otherwise

and (b) $\gamma_i=1, \forall i$

(b) is assumed since the first-order stochastic dominance criterion is not applicable to equality analysis (see Chapter One, III.1.1).

Applying these assumptions to (2) and substituting into (1) gives

$$\begin{aligned} p_i &= \frac{\mu_i}{\mu_h} = \frac{1}{\mu_h}\sum_{r=1}^{R} h_r \int_{\eta}^{z_i} dF_r(y), \quad i=1,...,W \\ &= F(z_i) \end{aligned} \tag{18}$$

In this equation, p_i takes on its 'natural' meaning as the proportion of the population with an income at or below z_i.

Defining ζ_i to be the negative of p_i enables one to analyze first-order stochastic dominance (1OSD). 1OSD with $z_{min} \le \eta^-$ and $z_{max} \ge \theta^+$ guarantees dominance over the set of all weakly increasing functions defined over income, where, as in Chapter One, η^- (θ^+) is the lower (higher) of the minimum (maximum) incomes of the two distributions being compared. 1OSD with just $z_{min} \le \eta^-$ guarantees dominance over all weakly increasing functions with poverty lines up to z_{max}. Restricted 1OSD guarantees dominance by the head count ratio with poverty lines between z_{min} and z_{max} (see Chapter One, III.1.1 and IV.2).

Replacement of (1) by (18) in turn simplifies the key equation (9) to

$$\frac{P(z_i,z_j)}{\mu_h^2} = \frac{P_{ij}}{\mu_h^{-2}} = \mu_{ij}+\mu_{hh}p_ip_j-\mu_{hi}p_j-\mu_{hj}p_i \tag{19}$$

From (8) and the assumptions above, we also have

$$\mu_{hi}=\mu_{ij}=\sum_{r=1}^{R} h_r^2\int_{\eta}^{z_i} dF_r(y) \quad \text{for } i\leq j \tag{20}$$

Analogously to the definition of p_i in (18), define

$$p_{hi} = F_h(z_i) = \frac{\mu_{hi}}{\mu_{hh}}, \quad i=1,\ldots,W \tag{21}$$

which can be interpreted as the proportion with income no greater than z_k, but using household size squared rather than household size as the weight. (20) and (21) further simplify (19) for $i\leq j$ to

$$\frac{P_{ij}}{\mu_{hh}\mu_h^{-2}} = p_{hi}(1-p_j) + p_i(p_j-p_{hj}),\ i\leq j \tag{22}$$

For $i>j$ we use the fact that $P_{ij}=P_{ji}$. Note that in the case of no-weighting

$$\mu_{hh}=\mu_h=1 \text{ and } p_{hj}=p_j \tag{23}$$

and (22) simplifies to

$$P_{ij} = p_i(1-p_j) \tag{24}$$

which is the standard formula for the covariance of two points of a sample distribution function density (see, for example, Mood, Graybill and Boes, 1963, pp.506-508).

From (18) and (21)

$$\hat{p}_i=\frac{m_i}{m_h} \qquad \hat{p}_{hi}=\frac{m_{hi}}{m_{hh}} \tag{25}$$

Consistent estimators have already been given for μ_h and μ_{hh} - see (11) and (12). By utilizing the indicator function, $I_z(y)$, consistent estimators of μ_i and μ_{hi} can be obtained. These are

$$m_i=\frac{1}{N}\sum_{n=1}^{N} h_n\ I_{z_i}(y_n) \text{ where } I_{z_i}(y_n)=1 \text{ if } y_n\leq z_i \tag{26}$$
$$=0 \text{ otherwise}$$

$$m_{hi}=\frac{1}{N}\sum_{n=1}^{N} h_i^2 I_{z_i}(y_n) \tag{27}$$

Estimators for p_i and P_{ij} thus defined can be used to test for 1OSD. They can also be used to derive estimators for the 2OSD criterion.

II.4 Deficit curve estimators for second-order stochastic dominance analysis

Defining ζ_k to be the negative of the poverty deficit curve - the integral of the distribution function defined over income - at z_k enables one to analyze welfare, y-restricted welfare and poverty 2OSD. If the range is unrestricted ($z_{min}\leq\eta^-$, $z_{max}\geq\theta^+$), the dominance extends over all egalitarian, weakly-increasing welfare functions; if it is restricted from above ($z_{min}\leq\eta^-$ only), the dominance extends over all egalitarian opulence functions with poverty lines less than or equal to z_{max}; and if it is restricted from both ends, the dominance is over all y-restricted welfare functions (defined in relation to the bounds z_{min} and z_{max}). (See Chapter One, III.1.2 and IV.2.)

Let the deficit curve at z_k be given by G_k and let $G=(G_1,...,G_k,...,G_q,...,G_W)$. G_k is given by integrating over the distribution function F, defined in (18),

$$G_k = G(z_k) = \int_{\eta}^{z_k} F(y)dy \tag{28}$$

and can be consistently estimated by

$$\hat{G}_k = \hat{G}(z_k) = \int_{\eta}^{z_k}\hat{F}(y)dy$$
$$= \begin{cases} 0 \text{ if } k=1 \\ \sum_{i=2}^{k} \hat{p}_{i-1}*(z_i-z_{i-1}) \text{ if } k\geq 2 \end{cases} \tag{29}$$

where $\hat{p}_i$ is defined in (25). (See Figure B.1 of Appendix B for an illustration). The z_i of the function I_{zi} (the 'grid') can be chosen to make the step-length arbitrarily small, but since the integration is over a step-function it is sufficient to have them given by all the sample values, $y_1=\hat{\eta}$ to $y_N=\hat{\theta}$, and by any other values of z of interest to the researcher not included in the sample, all combined and ordered from smallest z_1 to largest z_z. Note that now we must distinguish between z_i, i=1 to z and z_k, k=1 to W. G_k is estimated by summing over the z_i up to z_k, where the z_k are the points at which one wishes to evaluate the deficit curve. Note too that if the parent distribution function is differentiable, then, as the sample size goes to infinity, the average step-length of the grid (z_i-z_{i-1}) will go to zero, so that summation over the discrete grid will asymptotically approach integration over the smooth function F(y). If, on the other hand, however, the parent distribution function is a step function (for example, integers only), the average sample step-length will remain finite as its size increases.

From (28), the asymptotic covariance matrix, Γ, for $\sqrt{N}(\hat{G}-G)$ has its typical element given by:

$$\Gamma_{kq} = \int_{\eta}^{z_k}\int_{\eta}^{z_q} \mathrm{Cov}(\hat{F}(y),\hat{F}(u))\,dy\,du \tag{30}$$

where 'Cov' stands throughout for the **asymptotic** covariance. As is shown in Appendix A, by substitution of (22), this works out to give

$$\frac{\Gamma_{kq}}{\mu_{hh}\mu_h^{-2}} = G_{hk}(-G_q+z_q-z_k) + 2H_{hk} + G_k(G_q-G_{hq}) \tag{31}$$

where

$$\begin{aligned} G_{hk} &= G_h(z_k) = \int_{\eta}^{z_k} F_h(y)dy \\ H_{hk} &= H_h(z_k) = \int_{\eta}^{z_k} G_h(y)dy \end{aligned} \tag{32}$$

If there is no weighting then from (23)

$$G_{hk}=G_k \text{ and } H_{hk}=H_k=\int_{\eta}^{z_k} G(y)dy \tag{33}$$

which simplifies (31) to:

$$\Gamma_{kq} = G_k(-G_q+z_q-z_k) + 2H_k \tag{34}$$

As one would expect, the value of Γ_{kq} is constant wherever it is assessed at or above the maximum income.[7] Put formally,

$$\text{If } F(z_q)=1 \text{ (i.e. } z_q \geq \theta) \text{ then } \Gamma_{k,q+s}=\Gamma_{kq},\ k\leq q,\ s\geq 0$$
$$\text{and if } F(z_k)=1 \text{ (i.e. } z_k \geq \theta) \text{ then } \Gamma_{k+r,q+s}=\Gamma_{kq},\ k\leq q,\ r,s\geq 0. \tag{35}$$

In addition, since the value of G_k at or above the maximum income depends only on the per capita mean of the distribution, one has

$$\text{If } F(z_k)=1 \text{ then } \Gamma_{kq}=B \tag{36}$$

where B, defined in (17), is the asymptotic variance of b, a consistent estimator of the per capita mean of y. The proofs of (35) and (36) are also in Appendix A.

Estimators for the terms in (31) other than G_k and G_q are given below. G_{hk} is to be estimated in the same way as G_k in equation (29), replacing $\hat{p}_i$ by $\hat{p}_{hi}$ -also defined in (25) - giving

$$\begin{aligned} \hat{G}_{hk} = \hat{G}_h(z_k) &= 0 \text{ if } k=1 \\ &= \sum_{i=2}^{k} \hat{p}_{h,i-1}*(z_i-z_{i-1}) \text{ if } k\geq 2 \end{aligned} \tag{37}$$

And H_{hk} can be consistently estimated by

$$\hat{H}_{hk} = \int_{\mu}^{z_k} \hat{G}_h(y)dy = \begin{cases} 0 \text{ if } k=1 \\ \sum_{i=2}^{k} \hat{G}_{hi}(z_i-z_{i-1}) - 1/2\sum_{i=2}^{k} \hat{p}_{h,i-1}(z_i-z_{i-1})^2 \text{ if } k\geq 2 \end{cases} \tag{38}$$

(This is illustrated by Figure B.2 of Appendix B.) Since $\hat{H}_{hk}$ is the integral of $\hat{G}_{hk}$ use of equations (29), (37) and (38) to estimate Γ ensures that the properties (35) and (36) apply to the estimator of Γ as well as to Γ itself. At high values of y, where sample values are relatively scarce, use of only the first term in (38) may cause substantial and increasing overestimation of H_{hk}.

Finally, consider the estimators required for the analysis of poverty mixed dominance, which is, basically, the combination of 2OSD up to z^{-} and 1OSD between z^{-} and z_{max}, where these give the bounds on the set of possible poverty lines. Mixed dominance gives dominance in relation

7. This is reassuring from a formal perspective since it would be disconcerting if the value of Γ_{kq} changed between two income values both above the maximum and thus both corresponding to p=1. As stressed in Section IV of this chapter, however, any attempt to test for which distribution has the greater maximum is overly ambitious.

to all well-known poverty functions (see Chapter One, III.1.3). Here we need to estimate G_k for $z_k \leq z^*$ and p_j - as defined in (18) - for $z^* \leq z_j \leq z_{max}$. All that needs to be derived here is the asymptotic covariance of $\hat{G}_k$ and $\hat{p}_j$, where $j \geq k$, which, as Appendix A shows, is given by

$$\frac{\text{Cov}\ \sqrt{N}(\hat{G}_k, \hat{p}_j)}{\mu_{hh}\mu_h^{-2}} = G_{kh}(1-p_j) + G_k(p_j - p_{jh}),\ k \leq j \tag{39}$$

II.5 Lorenz curve estimators for welfare and inequality 2OSD analysis

Defining ζ_k to be the generalized Lorenz (GL) curve - the integral over p of the distribution function - at p_k enables one to analyze welfare and p-restricted welfare 2OSD. If the range is unrestricted ($p_{min}=0$, $p_{max}=1$), the dominance extends over all welfare functions, if it is restricted from above or below, the dominance extends over all p-restricted welfare functions defined in relation to the bounds p_{min} and p_{max}. Lorenz dominance has exactly the same implications as GL dominance, except that, since the Lorenz curve is defined over income divided by the mean, Lorenz dominance relates to unrestricted and p-restricted equality, rather than welfare, functions. (See Chapter One, III.2 and IV.2.)

This sub-section first shows how the covariance for two GL estimators can be based on, indeed equated to, that for two deficit curve estimators. This derivation is based on Beach and Davidson (1983), who, although they do not derive deficit curve estimators and do not deal with the weighted case, make implicit use of this equality.[8] I then show how Lorenz curve estimators can be derived from GL estimators. Again the derivation follows and extends Beach and Davidson.

As in the previous chapter (III.2), let $Q_k = Q(p_k) = \inf\{y : F(y) \geq p_k\}$. Then the GL curve at p_k is given by

$$\phi(p_k) = \int_{\eta}^{p_k} Q(p)dp \tag{40}$$

Define ϕ to be a vector of generalized Lorenz ordinates of interest, $\{\phi_1, \ldots, \phi_k, \ldots, \phi_q, \ldots, \phi_W\}$, where $\phi_k = \phi(p_k)$. Let $\hat{\phi}$ be a consistent estimator of ϕ. Then, from (40) and analogously to (30),

8. Compare in particular Beach and Davidson's equation (13) to this chapter's (A.5).

$\sqrt{N}(\hat{\phi}-\phi)$ is asymptotically normally distributed with mean zero and covariance matrix Φ, with typical element Φ_{kq} given by

$$\Phi_{kq} = \int_0^{p_k}\int_0^{p_q} \mathrm{Cov}(\hat{Q}(p),\hat{Q}(q))dpdq \tag{41}$$

Further, assume that F is strictly monotonic over the range $[\eta,\theta]$ and that therefore a unique inverse function, F^{1}, exists, and is given, at p_k, by $Q(p_k)=z_k$. Assume also that F and therefore Q is differentiable. Then

$$\mathrm{Cov}(\hat{Q}(p),\hat{Q}(q)) = Q'(p)\mathrm{Cov}(\hat{p},\hat{q})Q'(q) \tag{42}$$

[9]Substituting (42) into (41) gives

$$\Phi_{kq} = \int_0^{p_k}\int_0^{p_q} Q'(p)\mathrm{Cov}(\hat{p},\hat{q})Q'(q)dpdq \tag{43}$$

Now change the variable of integration from p to y (and q to u). Since

$$Q'(p) = F^{-1'}(p)=\frac{1}{F'(y)} \tag{44}$$

one obtains

$$\begin{aligned}\Phi_{kq} &= \int_{Q(0)}^{Q(p_k)}\int_{Q(0)}^{Q(p_q)} \frac{1}{F'(y)}\mathrm{Cov}(\hat{F}(y),\hat{F}(u))\frac{1}{F'(u)}dF(y)dF(u)\\ &= \int_\eta^{z_k}\int_\eta^{z_q} \mathrm{Cov}(\hat{F}(y),\hat{F}(u))dydu\end{aligned} \tag{45}$$

(since dF(y)=F'(y)dy). Comparing (45) and (30) proves the fundamental result

$$\Gamma_{kq} = \Phi_{kq} \tag{46}$$

That is, assuming differentiability and strict monotonicity of F, consistent deficit curve and GL curve estimators, themselves not equal, have exactly the same asymptotic covariance matrices.

Since Φ is given by Γ, all that needs to be furnished is a consistent estimator of ϕ_k,

9. Note that if one is using unweighted data, then, by substitution of (24) and setting of $Q'(p)=F^{-1'}(p)=1/f(y)$ - see (44) - this formula is equal to $(f_kf_q)^{-1}p_k(1-p_q)$, the standard formula for the covariance between two sample quantiles - see Kendall and Stuart (1963, p.237) and Beach and Davidson (1983).

which can be given by

$$\hat{\phi}_k = \int_0^{p_k} \hat{Q}(p)dp$$

$$= \begin{cases} \hat{z}_i \, p_i \text{ if } k=1 \\ \sum_{i=2}^{k} \hat{z}_i * (p_i - p_{i-1}) \text{ if } k \geq 2 \end{cases} \tag{47}$$

where

$$\hat{z}_i = \min(y_1, \ldots, y_N) \text{ s.t. } \hat{p}_i = p_i \tag{48}$$

The p_i can be chosen to make the step-length $(p_i - p_{i-1})$ arbitrarily small, but since the integration is over a step-function it is sufficient to have them given by all the sample values $\hat{p}$ (calculated using (25)) and by any other values of p of interest to the researcher not included in the sample, all combined and ordered from smallest (p_1) to largest (p_z). It is important to distinguish between p_i, i=1 to z and p_k, k=1 to W. ϕ_k is estimated by summing over the p_i up to p_k, where the p_k are the points at which one wishes to evaluate the GL curve. This formulation brings out the dualism between the deficit and the Lorenz approaches. In the former, the sample values of y (and any other z of interest) provide the grid used to sum over estimated abscissae, $\hat{p}$. In the latter, the sample values of $\hat{p}$ (and any other p of interest) provide the grid used to sum over estimated quantiles, y: compare (29) and (47).

ϕ_k and Φ_{kq} (and thus Γ_{kq}) can also be defined in terms of conditional means, which is likely to be more convenient. As shown in Appendix B, (46) enables us to rewrite Φ_{kq} as

$$\frac{\Phi_{kq}}{\mu_{hh}\mu_h^{-2}} = p_{hk}[\gamma_{hk}^2 - p_q(z_k - \lambda_{hk})(z_q - \lambda_q) - \lambda_{hk}(z_k + z_q) + z_k z_q] + p_k(z_k - \lambda_k)[p_q(z_q - \lambda_q) - p_{hq}(z_q - \lambda_{hq})] \tag{49}$$

p_{hk} is given by (21). Of the terms not yet defined, λ_k is the per capita mean of y conditional on $y \leq z_k$, that is,

$$\lambda_k = \frac{1}{p_k \mu_h} \sum_{r=1}^{R} h_r \int_{\eta}^{z_k} y dF_r(y), \tag{50}$$

λ_{hk} is the 'per capita-squared' mean of y conditional on $y \leq z_k$, that is,

$$\lambda_{hk} = \frac{1}{p_{hk}\mu_{hh}} \sum_{r=1}^{R} h_r^2 \int_{\eta}^{z_k} y\,dF_r(y), \tag{51}$$

and γ^2_{hk} is the 'per capita-squared' mean of y^2 conditional on $y \leq z_k$, that is,

$$\gamma^2_{hk} = \frac{1}{p_{hk}\mu_{hh}} \sum_{r=1}^{R} h_r^2 \int_{\eta}^{z_k} y^2\,dF_r(y). \tag{52}$$

[10]Use of (50) also enables the GL curve to be expressed in terms of the conditional mean - see, for example, the previous chapter, III.2 - as

$$\phi_k = p_k \lambda_k \tag{53}$$

In the case of no weighting,

$$p_q = p_{hq} \text{ and } \lambda_q = \lambda_{hq} \tag{54}$$

so the second line of (49) cancels and one is left with the formula presented by Beach and Davidson (1983).[11]

If p_k=1, let $\phi_k = \phi_z$ and $\Phi_{kq} = \Phi_{zq}$, and similarly for p_q. Evaluating (49) at the top end of the GL curve gives the result, consistent with (36) and (46), that

$$\Phi_{zz} = B \tag{55}$$

This is proved in Appendix B and is due to the position of the generalized Lorenz curve at p=1 depending only the value of the per capita mean of y, that is, $\phi_z = \beta$.

Consistent estimators of z_k and z_q can be obtained from (48): they will simply be the sample income quantiles. Using these, consistent estimators of p_{hk} and p_{hq} can be obtained using equation (25), and consistent estimators of the three new terms by

10. If p_k=0, λ_k can take on any value. Similarly for λ_{hk} and γ^2_{hk} if p_{hk}=0.

11. Also given by Bishop, Chakraborti and Thistle (1989).

$$\hat{\lambda}_k = \frac{1}{p_k m_h} * \frac{1}{N}\sum_{n=1}^{N} h_n y_n I_{\hat{z}_k}(y_n) \tag{56}$$

$$\hat{\lambda}_{hk} = \frac{1}{\hat{p}_{hk} m_{hh}} * \frac{1}{N}\sum_{n=1}^{N} h_n^2 y_n I_{\hat{z}_k}(y_n) \tag{57}$$

$$\hat{\gamma}^2_{hk} = \frac{1}{\hat{p}_{hk} m_{hh}} * \frac{1}{N}\sum_{n=1}^{N} h_n^2 y_n^2 I_{\hat{z}_k}(y_n) \tag{58}$$

where $I_z(y)$ is the indicator function defined in (26) and m_h and m_{hh} are defined in (11) and (12).

The real advantage of working with the Lorenz family is the possibility of using not the GL curve - since, as we have seen in the previous chapter (III.2), the deficit curve can be used for welfare and poverty analysis - but the ordinary Lorenz curve, which makes statistical analysis of equality dominance possible. The equality deficit curve, evaluated at z_k, is given by

$$G_{eq}(z_k) = \int_{\eta}^{\beta z_k} F(y)dy \tag{59}$$

The ordinary Lorenz curve, evaluated at p_k, is given by.

$$\omega_k = \frac{\phi_k}{\beta} = \frac{\phi_k}{\phi_z} \tag{60}$$

Although (59) and (60) are analytical substitutes in the sense defined in III.2 of the previous chapter, from a statistical perspective (60) is much easier to work with, since, following Beach and Davidson (1983), Rao's theorem can be applied to it. This gives the typical element, Ω_{kq}, of the asymptotic covariance matrix, Ω, of the vector of Lorenz ordinates, $\{\omega_1, \ldots \omega_W\}$, as

$$\Omega_{kq} = \frac{\Phi_{kq}}{\phi_z^2} + \frac{\phi_k\phi_q}{\phi_z^4}\Phi_{zz} - \frac{\phi_k}{\phi_z^3}\Phi_{kz} - \frac{\phi_q}{\phi_z^3}\Phi_{qz} \tag{61}$$

Substituting in $\phi_z=\beta$ and $\Phi_{zz}=B$ gives

$$\frac{\Omega_{kq}}{\beta^{-2}} = \Phi_{kq} + \beta^{-2}\phi_k\phi_q B - \beta^{-1}(\phi_k\Phi_{kz} + \phi_q\Phi_{qz}) \tag{62}$$

As with the GL curve estimators, in the case of no weighting (62) simplifies to the formula given by Beach and Davidson (1983).

II.6 Deficit and generalized Lorenz curve estimators compared

Can vectors of deficit curve and GL curve estimators be expected to give the same inferential results? One might think so, since, at least if the underlying distribution is assumed to be differentiable and strictly monotonic and if quantiles and abscissae are appropriately chosen (so that, for the sample to hand, p_k is the proportion with income less than or equal to z_k), $\hat{\Gamma}=\hat{\Phi}$. This follows from (46) of the previous sub-section: the proof, given in Appendix B, for the equality of Φ and Γ, given appropriately chosen ordinates, applies equally well to estimators of the two. Nevertheless, the answer to the above question must be negative. Differences may arise for the reason that one will not in general be able to choose the p_k so that they give the proportions in *both* samples with income less than z_k.

To give an example of this, consider the case of constructing two test statistics to analyze the GL and deficit curves of two distributions, F and F^*, at a particular point, p_k in the former case and z_k in the latter, given samples of size N drawn from both distributions. Since ϕ_k and G_k can both be considered sample means, albeit of transformed variables, the appropriate statistic is one which tests for differences in sample means, viz, for the GL curve

$$Z_k^{GL} = \frac{\hat{\phi}(p_k)-\hat{\phi}^*(p_k)}{N^{-\frac{1}{2}}(\hat{\Phi}_{kk}+\hat{\Phi}^*_{kk})^{\frac{1}{2}}} \tag{63}$$

and for the deficit curve

$$Z_k^{DEF} = \frac{\hat{G}^*(z_k)-\hat{G}(z_k)}{N^{-\frac{1}{2}}(\hat{\Gamma}_{kk}+\hat{\Gamma}^*_{kk})^{\frac{1}{2}}} \tag{64}$$

(63) and (64) will be equal for the special case in which

$$\hat{F}(z_k)= \hat{F}^*(z_k) = p_k \tag{65}$$

The equality of the numerators in this case can be seen by linking ϕ and G, and their estimators, by integration by parts (as in Chapter One, III.2 and Figure B.1 of Appendix B). That of the denominators follows since, if (65) holds, then the p_k at which the estimators of both Φ_{kk} and Φ^*_{kk} are evaluated will correspond to the z_k at which the estimators of both Γ_{kk} and Γ^*_{kk} are evaluated. If (65) does not hold, however, neither the numerators nor the denominators of (63) and (64) need be equal.

One can conclude that, where the distribution functions cross or coincide (including at the very top end of the distribution where both Z-statistics reduce to tests of differences of means), the GL and deficit curve estimators are perfect statistical substitutes, at least as far as this one statistic is concerned. At other points, however, just as the two curves 'ask' different questions - the deficit curve looking for differences along the income axis, the GL curve for differences along the p axis - so they may also give different answers. Whether the answers differ greatly must be assessed empirically and is done so in the next chapter.

II.7 Poverty analysis using the generalized Clark function

As outlined in IV.2 of the previous chapter, just as one can use the isoelastic function to investigate e-dominance in the cases of welfare and equality analysis, so for poverty analysis one can assess the generalized Clark function, ψ_{ijk}, for a range of inequality aversion parameters, α_i, 'discontinuity weights', C_j, and poverty lines, z_k.[12] Since the function can be varied along three dimensions, if there are v_1 possible values of z_k, v_2 of α, and v_3 of C, the resulting vector ψ has $(v_1 * v_2 * v_3)$ elements. ψ_{ijk} can be considered to be a special case of ζ_i if only one of α, C or z is varied.

Define p^i_k to be the per capita mean of $g(y,\alpha_i)I_{zk}(y)$, where I_z is again the indicator function - see (26). That is,

$$p^i_k = \frac{1}{\mu_h} \sum_{r=1}^{R} h_r \int_{\eta}^{z_k} g(y,\alpha_i) dF_r(y) \text{ where } g(y,\alpha_i) = \begin{cases} y^{\alpha_i}, & \alpha_i \neq 0 \\ \ln(y), & \alpha_i = 0 \end{cases} \qquad (66)$$

In addition, define $p_k = F(z_k)$ as in (18). Then the generalized Clark function can be re-written as

12. The z_k represent various values of Z^P in the terminology of Chapter One.

$$\psi_{ijk} = \psi(\alpha_i,C_j,z_k) = -\frac{1-C_j}{\alpha_i z_k^{\alpha_i}}p_k^i + (\frac{1-C_j}{\alpha_i}+C_j)p_k,\ \alpha_i\neq 0$$
$$= -(1-C_j)p_k^i + ((1-C_j)\ln(z_k)+C_j)p_k,\ \alpha_i=0 \tag{67}$$

[13]. A consistent estimator of ψ_{ijk} can be given by using the estimator $\hat{p}_k$ (see (25)), plus a consistent estimator for p_k^i given by

$$\hat{p}_k^i = \frac{1}{Nm_h}\sum_{n=1}^{N} h_n g(y_n,\alpha_i)I_{z_k}(y_n) \tag{68}$$

The typical element of the asymptotic covariance matrix, Ψ, of $\sqrt{N}(\hat{\psi}-\psi)$ is given by

$$\Psi_{(ijk,efg)} = \frac{(1-C_j)(1-C_f)}{\alpha_i\alpha_e z_k^{\alpha_i} z_g^{\alpha_e}}P_{kg}^{ie} + (\frac{1-C_j}{\alpha_i}+C_j)(\frac{1-C_f}{\alpha_e}+C_f)P_{kg}$$
$$-\frac{(1-C_j)(1-C_f(1-\alpha_e))}{\alpha_e\alpha_i z_k^{\alpha_i}}P_{kg}^i - \frac{(1-C_f)(1-C_j(1-\alpha_i))}{\alpha_i\alpha_e z_g^{\alpha_e}}P_{gk}^e$$

$$\text{where } P_{kg}^{ie} = \mathrm{Cov}\,\sqrt{N}(\hat{p}_k^i,\hat{p}_g^e) \tag{69}$$
$$P_{kg} = \mathrm{Cov}\,\sqrt{N}(\hat{p}_k,\hat{p}_g)$$
$$P_{kg}^i = \mathrm{Cov}\,\sqrt{N}(\hat{p}_k^i,\hat{p}_g)$$
$$P_{gk}^e = \mathrm{Cov}\sqrt{N}(\hat{p}_g^e,\hat{p}_k)$$

where $z_k \leq z_g$ and $\alpha_i\neq 0$ and $\alpha_e\neq 0$. To incorporate the case in which $\alpha_i=0$ or $\alpha_e=0$, assume $\alpha_i=0$. Then, from (67), replacing α_i in (69) (except where it appears as a superscript) by $1/\ln(z_k)$ and $z_k^{\alpha i}$ in the same equation by $\ln(z_k)$ will give the correct formula. The four 'P' terms of (69) are of course all based on the general formula (9) with $\gamma_i=1\ \forall i$. P_{kg} is given by (22). The other three terms can be derived by use of (9) and substitution into it of (18) and (66) (i.e., where appropriate substituting μ_k^i for μ_i and μ_g^e for μ_j):

13. Note that in Chapter One the function given is the negative of (67), since what is presented there is an 'opulence' or inverse poverty function rather than a poverty function.

$$\frac{P_{kg}^{ie}}{\mu_{hh}\mu_h^{-2}} = p_{hk}^{i+e} - p_{hk}^{i}p_g^{e} + p_k^{i}(p_g^{e} - p_{hg}^{e}) \text{ where } p_{hk}^{i+e} = p_{hk}^{q},\ q = i+e$$

$$\frac{P_{kg}^{i}}{\mu_{hh}\mu_h^{-2}} = p_{hk}^{i}(1 - p_g) + p_k^{i}(p_g - p_{hg}) \tag{70}$$

$$\frac{P_{gk}^{e}}{\mu_{hh}\mu_h^{-2}} = p_{hk}^{e} - p_{hg}^{e}p_k + p_g^{e}(p_k - p_{hk})$$

p_{hk} is defined in (21) and p_{hk}^i is the 'per capita-squared' mean of $g(y,\alpha_i)I_{zk}(y)$, namely

$$p_{hk}^{i} = \frac{\mu_{hk}^{i}}{\mu_{hh}} \text{ where } \mu_{hk}^{i} = \sum_{r=1}^{R} h_r^2 \int_{\eta}^{z_k} g(y,\alpha_i)dF_r(y) \tag{71}$$

which can be consistently estimated by

$$\hat{p}_{hk}^{i} = \frac{1}{m_{hh}N}\sum_{n=1}^{N} h_n^2 g(y_n,\alpha_i)I_{z_k}(y_n) \tag{72}$$

Various special cases warrant mention. First, if $C_j=1$ then the generalized Clark function becomes the head-count ratio. If $C_j=C_f=1$, (69) collapses to (22). If $C_j=0$, one has the Clark *et al.* (1981) function. Kakwani (1990) gives the variance for this case, on the assumption of no random weighting. Finally, if $C_j=0$ and $\alpha_i=1$, the generalized Clark function becomes the poverty gap (see Table 2 of Chapter One). If $C_j=C_f=0$ and $\alpha_i=\alpha_e=1$, then (dropping the subscripts held constant) $\Psi_{kg}=\Phi_{kg}/(z_k z_g)$, which generalizes Jäntti's (1992) result for the unweighted case. This last result is of course not surprising since $\Phi_{kg}=\Gamma_{kg}$ and, as noted in III.2 of the previous chapter, G_k is z_k times the poverty gap. It can be checked by noting the following equalities which hold in the case of $\alpha_i=1$:

$$\begin{aligned} p_{hk}^{2i} &= p_{hk}\gamma_{hk}^{2} \\ p_k^{i} &= p_k\lambda_k \\ p_{hk}^{i} &= p_{hk}\lambda_{hk} \end{aligned} \tag{73}$$

where the right-hand side terms, λ_k, λ_{hk} and γ_{hk}^2, are the conditional means of II.5.

III Methods of Inference

III.1 A method for the inference of distributional dominance

As stated in Section I, the analysis of dominance requires a comparison of heights of curves. Subject to a discrete approximation and for some Σ, F dominates F^* (FDF^*) iff $\zeta_i \geq \zeta_i^*$ $\forall i$ and $\zeta_i > \zeta_i^*$ $\exists i$, $i=1,...,W$. Three inferential outcomes are of primary interest: it can be inferred that FDF^*; or it can be inferred that F^*DF; or it can be inferred neither that FDF^* nor that F^*DF.[14] This is in contrast to the standard statistical test where there are only two outcomes: the null is rejected or not rejected. It is therefore necessary to distinguish between the rule for inferring dominance (the inference rule) and that for rejecting the null (the statistical test). Accordingly, I will distinguish between the Type I and Type II errors of the inference rule (inference errors) and those of the statistical test (test errors). In fact, it is evident that the inference rule has associated with it more than one Type I or Type II error. One might either infer that FDF^* when it is not the case, or that F^*DF when it is not the case. I will refer to these as, respectively, Type I and I^* inference errors. Similarly, one might either fail to infer that FDF^* when it is the case or fail to infer that F^*DF when it is the case. These will be referred to as Type II and II^* inference errors.

The aim in devising an inference rule will be to bound from above both the probability of making a Type I inference error and that of making a Type I^* inference error. Since the same bound will be given for both the Type I and the Type I^* inference error probabilities, the exposition below considers only the case in which we are interested in inferring that FDF^*, and will thus consider only the Type I and II inference error probabilities.[15] Chapter Three, II.4 considers how these Type I and I^*, II and II^* errors can be respectively combined to give summary size and power statistics.

The Type I inference error probability can be controlled by basing the inference rule on the rejection of some null. One initial complication is that, for tests based on the estimators given in the previous section, the Type I test error probabilities are all only known asymptotically. However it will be assumed for simplicity that these asymptotic results also hold in small samples. This will be called the assumption of 'no small-sample bias'.

14. One might also want to see whether it can be inferred that it is not the case that FDF^*, for example. See III.2 on this.

15. One could, if one wished, control the Type I and I^* inference error probabilities at different levels: see footnote 20.

Define

$$H_0^i:\zeta_i \le \zeta_i^* \qquad H_1^i:\zeta_i > \zeta_i^* \tag{74}$$

and

$$H_0^{*i}:\zeta_i^* \le \zeta_i \qquad H_1^{*i}:\zeta_i^* > \zeta_i \tag{75}$$

For all those special cases whose estimators are given in the previous section, these nulls can be simply tested using a 'Z-statistic', such as given in II.6. Using a cap to indicate estimators, defining Z_{ii} to be the typical diagonal element of the asymptotic variance-covariance matrix, Z,[16] of some consistent estimator of the vector $\zeta=(\zeta_1,\ldots,\zeta_W)$ and letting N (N^*) be the size of the sample drawn from F (F^*), we can define the Z-statistic at ordinate x_i:

$$Z_i = \frac{\hat{\zeta}_i - \hat{\zeta}_i^*}{(\hat{Z}_{ii}/N + \hat{Z}_{ii}^*/N^*)^{\frac{1}{2}}} \tag{76}$$

Under the null of equality, (76) has asymptotically a standard normal distribution, which can be used to choose its critical value. If Z_α is the upper α point of the standard normal distribution, then, asymptotically, the probability that $Z_i>Z_\alpha$ and the null H_0^i is rejected, given that H_0^i is true will be less than or equal to α (equal if $\zeta_i=\zeta_i^*$ and less than otherwise). It is assumed that (76) is only evaluated where (a) at least one sample covariance is not equal to zero and (b) the dominance curves, whether sample or population, are not equal by definition.

Consider the multiple-comparison hypotheses, based on (74):

$$H_0: H_0^i \ (\text{i.e.,} \zeta_i \le \zeta_i^*) \ \exists i, \ i=1,\ldots,W \qquad H_1: H_1^i \ (\text{i.e.,} \zeta_i > \zeta_i^*) \ \forall i, \ i=1,\ldots,W \tag{77}$$

[17]This null and alternative will be referred to as *intersection-union* (IU) hypotheses. This is the convention for the case in which, as here, the rejection (non-rejection) region of the multiple-comparison null hypothesis is the intersection (union) of the rejection (non-rejection) regions of

16. The matrix Z is not to be confused with the Z-statistic.

17. For the case in which interest is in whether F^*DF, use the null $H_0^*:H_0^{*i}\exists i$ and alternative $H_1^*:H_1^{*i}\forall i$.

the individual null hypotheses (i.e., every individual null has to be rejected for the multiple-comparison null to be rejected) - see Bishop, Formby and Thistle (1989). Let a test of H_0 against H_1 (an IU-test) be labelled T_{IU}, and let the rejection region of the test be labelled R_{IU}. Then H_0 will be rejected iff $T_{IU} \in R_{IU}$. The IU inference rule is

$$\Rightarrow FDF^* \text{ iff } T_{IU} \in R_{IU} \tag{78}$$

where '$\Rightarrow$' means 'it is to be inferred that'. For our problem, in which all the individual hypotheses can be tested by the one type of test - given in (76) - the IU test and thus the inference rule takes on a particularly simple form, viz.

$$\Rightarrow FDF^* \text{ iff } \min(Z_1, \ldots, Z_W) > Z_\alpha \tag{79}$$

where Z_i, i=1, ..., W, and Z_α are as defined in (76).

What significance can be attached to a rejection of H_0? The (very simple) answer is given by Berger (1982).[18] Let T_i be a test of H_0^i (i=1, ..., W), with rejection region R_i. Then

$$P(T_{IU} \in R_{IU} | H_0) \leq P(T_i \in R_i | H_0) \tag{80}$$

[19] This must hold as the LHS implies the RHS (in all states of the world including that in which H_0 is true), but not vice versa. Let the Type I test error probability for each of the tests T_1 to T_W be no greater than α. From the definitions of the IU null and the individual tests, T_i, there exists at least one i such that

$$P(T_i \in R_i | H_o) \leq \alpha \tag{81}$$

Combining (80) and (81) gives the result

$$P(T_{IU} \in R_{IU} | H_0) \leq \alpha \tag{82}$$

Assume that $W \geq 2$ and that the test statistics are less than perfectly correlated. Then the probability of a Type I inference error of the rule (78) can be related to α as follows.

18. Although Berger's article is concerned not with welfare economics but with quality control, the form of his problem - which is that "the consumer must decide whether the product is acceptable, that is, all of the parameters meet the standards, or unacceptable, that is, one or more of the parameters do not meet the standards."(p.295) - is similar.

19. When an hypothesis (H) appears in a probability statement, it is to be read as short for 'H is true' (e.g., $P(x|H_0)$ is short for 'the probability that x given that H_0 is true').

$$P(\rightarrow FDF^*|\neg FDF^*)=P(T_{IU}\in R_{IU}|\neg\, FDF^*)=P(T_{IU}\in R_{IU}|H_0,\neg FDF^*) < \alpha \tag{83}$$

The first equality is given by the inference rule (78), the second follows from the fact that if it is not the case that FDF* then H_0 is certainly true. The final inequality holds strictly since if H_0 is true and it is not the case that FDF*, then either $\zeta_i=\zeta_i^*$ $\forall i$, or $\zeta_i<\zeta_i^*$ $\exists i$. In the first case, the inequality in (80) will hold strictly due to the assumption of less than perfect correlation, and in the second that in (81) will hold strictly for at least one i (see the text following (76)). However, these inequalities may hold strictly by only arbitrarily small amounts. Hence no lower bound on the probability of a Type I inference error can be given than α. An analogous result can be given if F*DF is the proposition of interest. Hence both the Type I and the Type I* inference error probabilities are strictly bounded from above by α.[20]

The Type II inference error probability is given by

$$P(\neg\rightarrow FDF^*|FDF^*) = P(T_{IU}\notin R_{IU}|FDF^*) \tag{84}$$

If it is the case that FDF*, then either the alternative H_1 is true or, if $\zeta_i=\zeta_i^*$ $\exists i$, $\zeta_i>\zeta_i^*$ $\exists i$ and $\zeta_i<\zeta_i^*$ $\nexists i$ (i.e., if there is dominance but not strong dominance - see Section I for a definition), H_0 is true. Hence

$$P(T_{IU}\notin R_{IU}|FDF^*) = P(T_{IU}\notin R_{IU}|H_1)P(H_1|FDF^*) + P(T_{IU}\notin R_{IU}|H_0,FDF^*)P(H_0|FDF^*) \tag{85}$$

The first term is dependent on the Type II test error probability of T_{IU} since the latter is exactly $P(T_{IU}\notin R_{IU}|H_1)$. The second term is dependent on the Type I test error since, from (82), $P(T_{IU}\notin R_{IU}|H_0)\geq 1-\alpha$.[21] The presence of this second term indicates that the inference rule will tend not to infer dominance when there is an absence of strong dominance. This problem cannot be avoided within the adopted framework. Assume that the test statistic is continuous in the value of the minimum difference $MD=\min(\hat{\zeta}_1-\hat{\zeta}_1^*,\dots,\hat{\zeta}_W-\hat{\zeta}_W^*)$. Then when $\overline{MD}=\varepsilon^-$, where $\varepsilon^-<0$ but is arbitrarily small, the minimum test statistic generated is required to be 'small' to limit the probability of a Type I inference error. It will not then be possible to have a 'large' test statistic generated for MD=0, even though this would be desirable to reduce the Type II inference error probability. Nor would the Type II inference error be bounded away from 1-α if one could assume that either there was strong dominance or no dominance. There would remain the problem, generic to statistical testing, that the probability of a Type II error approaches one minus the probability

20. If one wished to control the Type I and I* inference error probabilities at different levels, one would use as critical values, say, Z_α and $Z_{\alpha'}$.

21. Analogous results can again be given for the Type II* inference error probabilities.

of a Type I error if the alternative is true but very similar to null. Hence even if one had strong dominance but $MD=\varepsilon^+$ where ε^+ is positive but again arbitrarily small, the probability of a Type II inference error would be close to one. The Type II inference error probability will only be 'small' when there are 'large' differences between the dominance curves at all ordinates. A necessary but not sufficient condition for this is strong dominance.[22]

III.2 An alternative approach

As mentioned, H_0 is an intersection-union hypothesis: a hypothesis that at least one of some set of individual nulls is true. An alternative type of null hypothesis is the *union-intersection* (UI) hypothesis that all the individual nulls are true. What is required here is a test in which the rejection (non-rejection) region of the multiple-comparison null hypothesis is the union (intersection) of the rejection (non-rejection) regions of the individual null hypotheses. Bishop, Formby and Thistle (BFT, 1989) propose using such a test for the inference of stochastic dominance, but their approach, expounded also in Chow, Chakraborti and Thistle (CCT, 1990), can be generalized to address the more general problem of distributional dominance addressed here. Their null and alternative can be written

$$\begin{aligned} &I_0: H_0^i \text{ (i.e.,} \zeta_i \leq \zeta_i^*) \ \forall i, \ i=1,\ldots,W \\ &I_1: H_1^i \text{ (i.e.,} \zeta_i > \zeta_i^*) \ \exists i, \ i=1,\ldots,W \end{aligned} \tag{86}$$

and[23]

$$\begin{aligned} &I_0^*: H_0^{*i} \text{ (i.e.,} \zeta_i^* \leq \zeta_i) \ \forall i, \ i=1,\ldots,W \\ &I_1^*: H_1^{*i} \text{ (i.e.,} \zeta_i^* > \zeta_i) \ \exists i, \ i=1,\ldots W. \end{aligned} \tag{87}$$

Let $T_{UI}^{(*)}$ reject $I_0^{(*)}$ if it is in the rejection region $R_{UI}^{(*)}$, where

22. The same reason could be given as to why it would not help to respecify the null, H_0, with a strict inequality as $\zeta_i<\zeta_i^* \forall i$. As McFadden writes in a similar context this would "induce the same test statistic and critical region, just as in the textbook case of testing the theoretically distinct null hypotheses $H_0:\theta\leq 0$ or $H_0:\theta<0$ for the mean θ of a normal population." (1989, p.115)

23. It should be noted that this is not the way in which BFT or CCT express the null and alternative. For example, CCT give a single null of $\zeta_i=\zeta_i^*$ $\forall$i, and two alternatives $\zeta_i>\zeta_i^*$ $\exists$i and $\zeta_i<\zeta_i^*$ $\exists$i. But this would not seem to be a material difference. The inference rule is unchanged.

$$T_{UI} \in R_{UI} \text{ iff } \max(Z_1,\dots,Z_W) > C_\alpha; \qquad (88)$$
$$T^*_{UI} \in R^*_{UI} \text{ iff } \min(Z_1,\dots,Z_W) < -C_\alpha.$$

Z_i, i=1,...,W, is defined in (76) and C_α is a critical value which will be defined shortly. BFT's UI inference rule is

$$\rightarrow FDF^* \text{ iff } T_{UI} \in R_{UI} \text{ and } T^*_{UI} \notin R^*_{UI} \qquad (89)$$

In other words, if we reject the null I_0 but not the null I^*_0 then we infer that FDF*.

The Type I error probability which BFT control is that of rejecting either I_0 or I^*_0 when both are true. This is given by

$$P(T_{UI} \in R_{UI} \text{ or } T^*_{UI} \in R^*_{UI} | I_0, I^*_0) = P(\max\{|Z_1|,\dots,|Z_W|\} > C_\alpha | I_0, I^*_0) \qquad (90)$$

This last is the probability of rejecting the null $\zeta_i = \zeta^*_i$ $\forall i$ when it is true. It can be bounded from above by α by appropriate choice of C_α. For example, one can use the Bonferroni inequality and set $C_\alpha = Z_{\alpha/2W}$, or the Sidak inequality and set $C_\alpha = Z_{\delta/2}$ where $\delta = [1-(1-\alpha)^{1/W}]$ (see BFT, p.67). Alternatively, and as BFT recommend, one can set C_α equal to $m_\alpha(W, N+N^*)$, the upper α point of the Studentized Maximum Modulus (SMM) distribution, with parameter W and N+N* degrees of freedom, tables of which are given by Stoline and Ury (1979). With any of these choices, one has the result:[24]

$$P(T_{UI} \in R_{UI} \text{ or } T^*_{UI} \in R^*_{UI} | I_0, I^*_0) \le \alpha \qquad (91)$$

Ultimately, however, we are interested in the probability of wrongly inferring dominance. BFT and CCT do not consider, at least not explicitly, what I have called the Type I and II inference error probabilities. To do so, it is helpful to assume C_α is chosen so that

$$P(T_{UI} \in R_{UI} | I_0) \le \alpha/2; \; P(T^*_{UI} \in R^*_{UI} | I^*_0) \le \alpha/2 \qquad (92)$$

This will be consistent with (91) if, for example, one uses the Bonferroni inequality above (setting $C_\alpha = Z_{\alpha/2W}$).

The probability of a Type I inference error using BFT's rule is, from (89):

24. For any α, the critical value based on the SMM distribution is no higher than that based on the Sidak inequality which in turn is no higher than that based on the Bonferroni inequality. Asymptotically, critical values based on the SMM distribution and the Sidak inequality are equal. See Stoline and Ury (1979) and Savin (1984).

$$P(\rightarrow FDF^*|\neg FDF^*) = P(T_{UI}\in R_{UI}, T^*_{UI}\notin R^*_{UI}|\neg FDF^*) \tag{93}$$

If the two distributions are such that it is not the case that FDF* then either the alternative hypothesis I_1^* is true or both I_0 and I_0^* are true. Combining this with (93) gives

$$\begin{aligned} P(\rightarrow FDF^*|\neg FDF^*) = {} & P(T_{UI}\in R_{UI}, T^*_{UI}\notin R^*_{UI}|I_0,I_0^*)P(I_0,I_0^*|\neg FDF^*) \\ & + P(T_{UI}\in R_{UI}, T^*_{UI}\notin R^*_{UI}|I_1^*)P(I_1^*|\neg FDF^*) \end{aligned} \tag{94}$$

Again assume that there is no small sample bias. Then the first term of the RHS of (94) can be simplified using (92):

$$P(T_{UI}\in R_{UI}, T^*_{UI}\notin R^*_{UI}|I_0,I_0^*) \le P(T_{UI}\in R_{UI}|I_0,I_0^*) \le \alpha/2 \tag{95}$$

The second term of the RHS of (94) can also be decomposed to give

$$\begin{aligned} & P(T_{UI}\in R_{UI}, T^*_{UI}\notin R^*_{UI}|I_1^*) = \\ & P(T_{UI}\in R_{UI}, T^*_{UI}\notin R^*_{UI}|I_1^*,I_0)P(I_1^*,I_0|I_1^*) + P(T_{UI}\in R_{UI}, T^*_{UI}\notin R^*_{UI}|I_1^*,I_1)P(I_1|I_1^*) \\ & \le \frac{\alpha}{2}P(I_1^*,I_0|I_1^*) + P(T_{UI}\in R_{UI}, T^*_{UI}\notin R^*_{UI}|I_1^*,I_1)P(I_1|I_1^*) \end{aligned} \tag{96}$$

(Again the inequality follows from (92).) Combining (94), (95) and (96) gives the result

$$\text{If } P(I_1|I_1^*) = 0 \text{ or } P(T_{UI}\in R_{UI}, T^*_{UI}\notin R_{UI}|I_1^*,I_1) = 0 \text{ then } P(\rightarrow FDF^*|\neg FDF^*) \le \alpha/2 \tag{97}$$

If it is assumed that the dominance curves do not cross, the first equality will hold. However, if one is making no assumption about the parametric forms of the distributions F and F*, this assumption is not permissable. Alternatively, if it is assumed that if there are crossings either both nulls will be rejected or both will not be rejected then the second equality will hold. However, this will depend on the Type II error probabilities of the UI tests, since it is conditional on both alternatives being true. No general expression can be given for this latter. Hence it is not possible to bound the Type I inference error probability for the BFT rule. It is not surprising, therefore, that in a simulation study CCT find that the probability of inferring dominance when there is a crossing can be high using their UI inference rule. The simulations reported in the next chapter confirm this result.

To evaluate the probability of a Type II inference error, it follows from the inference rule (89) that

$$P(\neg\rightarrow FDF^*|FDF^*) = P(T_{UI}\notin R_{UI} \text{ or } T^*_{UI}\in R^*_{UI}|FDF^*) \tag{98}$$

For it to be the case that FDF*, I_1 and I_0^* must both be true. Hence

$$P(T_{UI} \notin R_{UI} \text{ or } T^*_{UI} \in R^*_{UI} | FDF^*) = P(T_{UI} \notin R_{UI} \text{ or } T^*_{UI} \in R^*_{UI} | I_1, I^*_0) \tag{99}$$

This expression will depend on the Type I error probability of T^*_{UI} - since this gives the probability of rejecting I^*_0 when it is true - and on the Type II error probability of T_{UI} - since this gives the probability of not rejecting I_0 when I_1 is true. No general expression can be given for the latter.

To conclude, both the IU and UI inference methods are very simple. Neither requires the calculation of covariances. The IU rule infers FDF^* if the minimum Z-statistic is greater than Z_α. The UI rule infers FDF^* if the maximum Z-statistic is greater than C_α and the minimum is greater than $-C_\alpha$. The key difference between them is that, under the assumption of no small sample bias, use of the former enables one to bound the Type I and Type I^* inference error probabilities, and thus to attach a degree of confidence to any inference of dominance made using it. Since one can bound neither the Type I nor the Type II error probabilities of the UI inference rule, its usefulness is limited.

This limited usefulness arises from the fact that the UI inference rule partially bases any inference of dominance on non-rejection of the null: if I^*_0 is not rejected it is inferred that $\zeta_i \geq \zeta^*_i$ $\forall i$, leaving only $\zeta_i > \zeta^*_i$ $\exists i$ to be inferred via the rejection of I_0 before it can be concluded that FDF^*. Although this may lead to spurious inferences of dominance being made, rejection of the null, I^*_0 (I_0), is nevertheless of interest, as it enables us to infer that it is *not* the case that FDF^* (F^*DF). Assuming there is no small-sample bias, the inference rule

$$\rightarrow \neg F^*DF \text{ iff } T_{UI} \in R_{UI} \tag{100}$$

will have a Type I inference error probability bounded by $\alpha/2$, if the critical value is chosen in accordance with (92). Although a conclusion of no dominance doesn't of course carry the same implications as one of dominance, it may nevertheless be useful information. If one cannot infer that it is the case that FDF^* one need not worry about the possibility of a Type II error if one also infers, on the basis of a UI test, that it is not the case that FDF^*. One can conclude that the IU rule (89) should be used for inferring dominance, the UI rule (100) for inferring no dominance.

III.3 Other tests in the literature

A very early test for differences in curves is provided by Mahalanobis's (1960) method of 'fractile graphical analysis', which he applies to, *inter alia*, the Lorenz curve. Mahalanobis proposes one divide one's sample into two, and consider the distance between the curves associated with each sub-sample to be a measure of sampling error at any given point. If one is comparing two samples, one can compare the distance or 'separation' between each sample curve with these two errors (one for each sample) as a means of discerning the significance of the separation. This approach certainly has an intuitive appeal: one can think of the sub-sample curves as providing confidence intervals for the full sample curve.

A more formal test which has been used in relation to various dominance criteria is the chi-squared, in which the test statistic

$$(\hat{\zeta} - \hat{\zeta}^*)' \left(\hat{Z}/N + \hat{Z}^*/N^* \right)^{-1} (\hat{\zeta} - \hat{\zeta}^*) \tag{101}$$

is asymptotically $\chi^2(W)$. As pointed out by various authors, this is a test of differences between dominance curves (see, for example, Beach and Davidson, 1983) rather than of dominance by one curve of another. As Savin (1984) shows, the chi-squared test is another union-intersection test. It tests the null that $a'(\zeta - \zeta^*) = 0$ for all non-null vectors, a, against the alternative that $a'(\zeta - \zeta^*) \neq 0$ for some a. Rejection of this null is clearly necessary but not sufficient for dominance.

Gastwirth and Gail (1985) propose a test in which the summation of the differences $\zeta_i - \zeta_i^*$ over all i is estimated, and divided by the appropriate standard deviation. Although this test performs well in the simulations they run, its rationale is unclear: a positive sum of differences is a necessary but not sufficient condition for dominance.

CCT write that various authors "have pointed out that the one-sided Kolmogorov-Smirnov (K-S) test can be used to test for first-degree stochastic dominance" but correctly cast doubt on this noting that "the one-sided K-S test can be applied to test for first-degree stochastic dominance only if it is known a priori that the distribution functions do not cross" (p.2, footnote 2). In addition, as the authors of a standard text on statistics, Randles and Wolfe (1979, p.382), write (labelling their two distribution functions F and G), the hypothesis which this statistic tests is "H_0:[F(x)=G(x) for all x] against the general alternative H_1:[F(x)≠G(x) for at least one x]." While the test could of course be modified to give a one rather than two-tailed test, the quotation brings out the key point that the K-S test is another UI rather than IU test. Hence it is useful for testing for no dominance, but not for dominance.

Tolley and Pope (1988) present a test based on asking how many permutations of the two samples could have produced at least a large a difference between the deficit curves as the maximum positive and negative differences actually produced. If these numbers are small - less than the total possible number of permutations multiplied by the critical value - then the null hypothesis that the two samples are both from the same population and are simply "one random outcome of all possible two sample permutations one can obtain from the data" (p.694) is rejected. If there is a large positive difference, and no large negative difference, or vice versa, dominance is inferred. Again this test would seem to be a union-intersection test (compare this inference rule with (89)). An additional problem with this test is that it is likely to become unmanageable even with small sample sizes. Pope and Tolley give an example using two samples of 10 observations each. The possible number of permutations in such a case is 184,756.

McFadden (1989) presents a test which is similar to that of Tolley and Pope, but with two substantial differences. First, McFadden recognizes, at least implicitly, that the test is one for no dominance rather than dominance, by setting the null as the hypothesis that F dominates F^*, and the alternative as the hypothesis that it does not. (By contrast, Tolley and Pope give as the null the hypothesis that F does not dominate F^*, and the alternative as the hypothesis that it does.) Secondly, McFadden provides guidance as to approximate critical values for the permutation test, and a program by which one can control the number of random permutations chosen if one wishes to approximate the critical value for a particular pair of samples more accurately.

In conclusion, none of the tests currently available seem to be adequate as tests of dominance, though some of them, like the UI method of the previous sub-section, can be utilized to test for no dominance.

III.4 The choice of ordinates

As stated in the introduction to this chapter, for the purpose of statistical testing dominance can only be evaluated over a finite number, W, of ordinates. It is assumed that the conclusion which follows from the estimation of ζ_i over a finite set of values, x_1 to x_W, where $x_{min} \leq x_1$ and $x_W \leq x_{max}$, will be the same as that which would follow from an assessment, were it possible, over all points in the interval. The question remains as to how this discrete approximation should be made.

Again, the aim of controlling the Type I inference error probability gives the answer. To do this we need to rule out the possibility that FDF^* is inferred over the discrete grid, x_1 to x_W, even though there is some x such that $x_{min} \leq x \leq x_{max}$ and $x \neq x_i$, $i=1,...,W$, at which $\zeta < \zeta^*$. For this, we require $x_1 = x_{min}$, $x_W = x_{max}$ and a fine grid. Thus it is recommended that if one is assessing deficit curves over some range of incomes, the number of ordinates should equal the number of income values in this range contained by the two samples. If one is assessing Lorenz curves over some range of p, the number of ordinates should equal the number of $\hat{p}$ values (see II.5). And if one is assessing e-dominance, the relevant social welfare functions should be assessed at very close intervals of e, at say a distance at 0.1. If computational resources are limited or the data set very large, the approximation may have to be less fine. There would, however, from the computational view, be a much stronger case for using a small W if one wanted to use a chi-squared test, which requires the construction of a W-by-W matrix. The IU tests advocated for use here, by contrast, only require the calculation of vectors of variances, not matrices of covariances. Moreover, use of graphs enables all the test-statistics to be reported without lengthy tables: see Chapter Five for examples.[25]

The approaches taken by the existing testing methods with regards to the choice of W vary. The Kolmogorov-Smirnov, McFadden, and Pope and Tolley tests make use of all the sample income values in searching for the maximum difference between the relevant dominance curves. The Gail and Gastwirth test, by contrast, only uses three ordinates, while the UI test of BFT and CCT has been used by these authors with between 3 and 30 ordinates. The choice of three ordinates by Gail and Gastwirth is for computational convenience. One reason for BFT and CCT's choice of a relatively small number of ordinates is that the critical values they use are increasing in W.[26] The increase in the critical values is to offset the increased probability of a spurious rejection arising from a larger number of W. However, because all the ζ_i estimators are calculated from a single sample, the correlation between the sample ζ_i will increase with W, making the increased probability of spurious rejection increasingly slight as W rises. From this point of view, the critical value may rise 'too fast' with W. CCT have examined this issue in relation to testing for stochastic dominance using a simulation study. Their conclusions are quite specific: "For the first degree stochastic dominance test, 15 to 28 [ordinates] should be used. For the second and third degree stochastic dominance tests, six and three [ordinates], respectively, should be used." (p.20-21) However, it is unclear whether these results apply across all distributions or hold only for those selected in their simulation study. The question is returned to in the simulation study of

25. One can also use this method to draw confidence intervals around individual dominance curves (in which case dominance requires that the confidence intervals do not intercept) or to draw confidence intervals for the difference between dominance curves. However, note that these would not be 'simultaneous' or 'joint' confidence intervals which fix the probability that **all** the intervals contain their respective population parameter, not, as here, that at least one does. The former, based on $\pm C_\alpha$ (the UI method critical value), will be wider than the latter, based on $\pm Z_\alpha$ (the IU method critical value). See Beach and Richmond (1985) for an analysis of joint confidence intervals in relation to Lorenz curves.

26. For example, if the Bonferroni inequality is used, then $C_\alpha = Z_{\alpha/W}$. For given α, as W increases, α/W will fall and C_α rise.

the next chapter (see IV.2.1.1).

IV Inferring minimum and maximum dominance

Chapter 1, IV.1 demonstrated the dependence of the stochastic dominance criteria on various forms of 'extreme' dominance. For example, in relation to 2OSD, for one distribution to have poverty 2OSD it must have minimum dominance, for it to have welfare 2OSD it must have mean-minimum dominance (a no lower minimum and no lower mean) and for it to have equality 2OSD it must have maximum-minimum dominance (a no lower minimum and no higher maximum, both defined over mean-normalized income). Inferring which distribution has a higher mean is not in general problematic, but inferring which distribution has a higher or lower minimum or maximum may well be. This section analyzes the difficulties of inferring second-order stochastic dominance on account of the need to infer minimum and maximum dominance.

I begin by showing the inadequacy of relying on the estimators presented in Section II for the inference of minimum or maximum dominance. To illustrate, an analysis is presented of the behaviour of Z-statistics based on the GL and deficit curve estimators, given as (63) and (64) in II.6, at the lower end of the tail for two samples of unweighted data each of size N. It is assumed that the sample drawn from population F^* has the smallest income of the two samples combined and that from F the second smallest income.

For the deficit curves, (64) is evaluated, at z_2, the second smallest income level present in the two combined samples. (The test-statistic will be undefined at z_1, since both sample variances will be zero.) The denominator is then given by

$$N^{-1/2}(\hat{\Gamma}_{22}+\hat{\Gamma}^*_{22})^{1/2} = N^{-1/2}[(z_2-z_1)^2\hat{p}_1(1-\hat{p}_1)]^{1/2} \tag{102}$$

and the numerator by

$$\hat{p}_1(z_2-z_1) \tag{103}$$

Hence the Z-statistic is

$$Z_2^{DEF} = \sqrt{N}\,\hat{p}_1^{1/2}/(1-\hat{p}_1)^{1/2} \tag{104}$$

What will happen to this value depends on the change in $\hat{p}_1$ as N becomes large. Assume that only one unit ever receives the minimum income. Then $\hat{p}_1=1/N$ and

$$Z_2^{DEF} = \left(\frac{N}{N-1}\right)^{1/2} \rightarrow 1 \text{ as } N\rightarrow\infty \tag{105}$$

In the case of the GL curve, (63) is evaluated at the second smallest cumulative proportion, p_2 (again the test statistic will be undefined at the first ordinate, p_1, since both estimated variances will be zero). In this case, the denominator is given by

$$[N^{-1}p_1(1-p_1)((y_2-y_1)^2+(y_2^*-y_1^*)^2)]^{1/2} \tag{106}$$

the numerator by

$$p_1(y_1-y_1^*)+(p_2-p_1)(y_2-y_2^*) \tag{107}$$

and the test statistic by

$$Z_2^{GL} = \frac{\left(\frac{N}{p_1(1-p_1)}\right)^{1/2}(p_1(y_1-y_1^*)+(p_2-p_1)(y_2-y_2^*))}{((y_2-y_1)^2+(y_2^*-y_1^*)^2)^{1/2}} \tag{108}$$

Again let $p_i=i/N$. Then

$$Z_2^{GL} = \left(\frac{N}{N-1}\right)^{1/2}\frac{y_1-y_1^*+y_2-y_2^*}{((y_2-y_1)^2+(y_2^*-y_1^*)^2)^{1/2}} \tag{109}$$

For a pair of continuous parent distributions, this is likely to be increasing in N, as the average distance between any two sample values will fall, while that between the pth observation of each sample is more likely to remain constant. Hence as N rises, the numerator will remain the same size, and the denominator will fall. The simulation tests carried out in the next chapter indicate that for a range of pairs of distributions, the GL test statistic often takes its maximum absolute value at p_2, and will often indicate a significant difference at p_2 when none exists (see Chapter Three, IV.2.1.1).

So, not surprisingly, neither the deficit curve nor the GL curve estimators provides a reasonable test for minimum dominance. (104) is independent of the actual sample income values.

(109) at least depends on the two smallest income values. In practice, however, it proves to be an unreliable indicator.

It is more difficult to obtain a general expression for Z_i at the maximum income when one is comparing Lorenz curves. The numerator, however, at p=(N-1)/N is simply given, in this unweighted case, by the difference in mean-normalized maximum incomes divided by the common sample size. Again this seems like a fragile basis on which to hang an inference, and again experience suggests that the value of the test statistic at values of p close to 1 is likely to be large, leading to spurious rejections (see Chapter Five, III.2, footnote 7).

Inferring minimum or maximum dominance need pose no particular problems if one is able to parameterize the distributions. An exact parameterization may not even be required. Restriction of the sampled distribution to some class of distributions may be sufficient to determine the relevant limiting extreme-value distribution (Mood, Graybill and Boes, 1963, pp.256-264). But whether or not such restrictions are convincing in general, the chances of them being convincing in respect of the tails of real-world populations of income recipients seem weak.

Even for non-parametric tests of extreme dominance, some assumption that the functions are at least reasonably well-behaved is necessary. The basic difficulty is that any inferential rule for minimum and maximum dominance will have to be based on the two samples' left and right tails respectively. But then the rule will be based on a small number of observations and so, without any regularity assumption, will be fragile to unsampled outliers.

To illustrate, consider a test for inferring minimum dominance derived by Whitmore (1978).[27] His test statistic can be written

$$T_{min}=\begin{cases}+t \text{ if } y_t \le y_1^* \le y_{t+1}\\ -t \text{ if } y_t^* \le y_1 \le y_{t+1}^*\end{cases} \tag{110}$$

where incomes are arranged in ascending order (with ties broken arbitrarily), so that, for example, y_t is the t-th smallest income of the sample F (assuming no ties). The larger T_{min} is in absolute value, that is, the more observations one has from one sample with income less than the minimum income of the second sample, the less likely it is that both samples are drawn from the same

27. Whitmore presented his test for minimum dominance as one part of a test for welfare 2OSD for distributions which cross at most one. The other test is for differences in sample means.

population. Whitmore shows that under the null that the two samples are drawn from the same distribution

$$P(T_{min} \geq k) = \frac{\binom{m+n-k}{n-k}}{\binom{n+m}{n}}, \quad P(T_{min} \leq -k) = \frac{\binom{m+n+k}{m-k}}{\binom{n+m}{n}} \tag{111}$$

and that, for large sample sizes,

$$\lim_{n\to\infty, m\to\infty} P(T_{min} \geq k) = \lim_{n\to\infty, m\to\infty} P(T_{min} \leq -k) = \frac{1}{2}^k \tag{112}$$

This is a very simple test, and one that one could also use for the inference of maximum income by using the same test but arranging the observations in descending order of y.

Certainly if one does get a run of observations all from the one sample it is reasonable to infer that it and the other sample are drawn from different populations. But does a large test statistic justify the inference of minimum dominance? A necessary condition for this to be the case is that it must be more likely for the distribution with minimum dominance to produce the sample with the smallest income value. It is difficult to specify precisely under which circumstances this condition will hold, but it is easy to think of cases in which it will not. Consider a population F which has one household receiving income of $10 and all other households receiving incomes of $20, and a population F^* which has all households receiving income of $15. Samples taken from these two distributions will probably lead to the inference, if Whitmore's test is used, that F^* has the smaller minimum. Only if the solitary $10 household is sampled will this conclusion be prevented.

Long tails and outliers can of course cause problems for all non-parametric tests. If there are only a very few very rich, significant underestimation of the mean may result. But in cases in which the statistic depends on the entire sample or a substantial sub-set of it one can at least hope that, with a large sample size, the difference which would be made by any outliers not sampled would be small. No such hope can be entertained with tests for minimum and maximum income. The fact that pockets of extreme poverty exist even in the most affluent societies, and pockets of extreme affluence in the poorest societies makes it difficult, if not impossible, for minimum or maximum dominance to be inferred. Even if one could be confident that the poorest Australian was better off than the poorest Indian, what could one hope to say about the comparisons of real interest, such as India ten years ago with India today?

A final issue of relevance here is that of measurement error. Even if incomes are not more poorly measured at either end of the tail than in the body of the distribution (which they probably are), any test for maximum or minimum dominance will be highly susceptible to measurement error (and to data-cleaning decisions), more so than tests reliant on a large number of observations. Using Whitmore's test for example, a single mistake may be sufficient to prevent the inference of minimum dominance.

It is surprising that this issue of the inference of extreme dominance has, with the notable exception of Whitmore, not been addressed in the literature on statistical testing for stochastic dominance. One reason for this lies with the tendency, pointed to in III.3, to choose, in relation to certain tests, a relatively small number of ordinates over which to examine whether dominance holds. In this way it is possible to avoid choosing ordinates from the extreme tails of the distribution. Another reason is the preponderance of tests for no dominance, which focus on the maximum value of the test statistic. In this case, if the test values at the relevant tail are small, they can be safely ignored.

Given the difficulties of finding any satisfactory rule for the inference of maximum or minimum dominance, it would seem advisory to avoid usage in a statistical context of dominance criteria, such as (unrestricted) second-order stochastic dominance, which require these 'extreme' forms of dominance to hold.[28] Alternatives provided in Section Four of the previous chapter include the criteria of restricted second-order stochastic dominance and e-dominance. The problem of sensitivity to the lower tail will not vanish due to the introduction of a lower bound in the case of restricted 2OSD or upper bound in that of e-dominance, but it can at least be controlled.

V Conclusion

This chapter has derived consistent estimators and their asymptotic variance-covariance matrices which can be used to test for dominance in relation to the following criteria in the presence of randomly weighted data:

(i) first-order stochastic dominance, in relation to welfare or poverty;
(ii) second-order stochastic dominance, in relation to welfare, equality or poverty;
(iii) poverty mixed stochastic dominance;

---

28. Alternatively, one can, in cases where the relevant tails generate small test statistics (thus excluding, for example, use of the generalized Lorenz curve), restrict such usage to the inference of no dominance.

(iv) e-dominance, in relation to welfare, equality or poverty (using the generalized Clark function in the latter case);
(v) restricted second-order stochastic dominance, in relation to equality and welfare (p-restricted and y-restricted).

For the stochastic dominance criteria - restricted and unrestricted - estimators have been provided based on both the deficit and the generalized Lorenz curve to allow for maximum flexibility.

This chapter has also presented a general method for the inference of dominance criteria based on the use of intersection-union tests. A very simple test procedure has been recommended. First calculate the Z-statistics for the dominance criterion of interest for a 'large' number of ordinates within the range of interest and check that they are all of the same sign. If not, do not infer dominance. If so, then select the Z-statistic with the minimum absolute value. If it is greater than Z_α (the upper α point of the standard normal distribution) then infer FDF^*. If it is less than $-Z_\alpha$ then infer F^*DF. Either inference, if made, will have an asymptotic probability of being wrong of at most α. This inference rule has been contrasted to others in the literature, and advocated for use on the basis that, abstracting from small-sample bias, it has bounded probabilities of inferring dominance when dominance is absent.

Finally, it has been argued that the difficulties of inferring minimum and maximum dominance render dominance criteria dependent on these 'extreme' forms of dominance, such as second-order stochastic dominance, of limited relevance in a statistical context. More restrictive dominance criteria are likely to be of greater use.

Further evidence on some of the outstanding questions arising from this chapter, as well as in support of some of the claims made herein, is provided in the next chapter, where results from a simulation study are reported.

Appendix A Derivations for II.4

Equation (31), the covariance formula for the deficit curve,

$$\frac{\Gamma_{kq}}{\mu_{hh}\mu_h^{-2}} = G_{hk}(-G_q+z_q-z_k) + 2H_{hk} + G_k(G_q-G_{hq}) \tag{A.1}$$

can be derived as follows. Combining the fact that the covariance between two sums is equal to the sum of the covariances (see for example, Mendenhall, Scheaffer and Wackerly, 1986, p.210) with the fact that G_k is the integral of F up to z_k gives

$$\begin{aligned}\Gamma_{kq} &= \int_{\eta}^{z_k}\int_{\eta}^{z_q} \mathrm{Cov}(\hat{F}(y),\hat{F}(u))\,dy\,du \\ &= \int_{\eta}^{z_k}\int_{\eta}^{z_q} P(y,u)\,dy\,du\end{aligned} \tag{A.2}$$

Making use of the fact that

$$P(y,u)=P(u,y) \tag{A.3}$$

(A.2) can be re-written as

$$\Gamma_{kq} = \int_{\eta}^{z_k}\int_{y}^{z_q} P(y,u)\,dy\,du + \int_{\eta}^{z_k}\int_{\eta}^{y} P(u,y)\,dy\,du \tag{A.4}$$

Using (22), the formula for P_{ij}, (A.4) can be re-written

$$\begin{aligned}\frac{\Gamma^c_{kq}}{\mu_{hh}\mu_h^{-2}} &= \int_{\eta}^{z_k} F_h(y)\,dy\int_{y}^{z_q}(1-F(u))\,du + \int_{\eta}^{z_k} F(y)\,dy\int_{y}^{z_q}(F(u)-F_h(u))\,du \\ &+ \int_{\eta}^{z_k}(1-F(y))\,dy\int_{\eta}^{y}F_h(u)\,du + \int_{\eta}^{z_k}(F(y)-F_h(y))\,dy\int_{\eta}^{y}F(u)\,du\end{aligned} \tag{A.5}$$

Integration over u and use of the notation given in (28) and (32) gives

$$\frac{\Gamma^c_{kq}}{\mu_{hh}\mu_h^{-2}} = \int_\eta^{z_k} F_h(y)(z_q - y - G_q + G(y))dy + \int_\eta^{z_k} F(y)(G_q - G(y) - G_{hq} + G_h(y))dy$$
$$+ \int_\eta^{z_k} (1 - F(y))G_h(y)dy + \int_\eta^{z_k} G(y)(F(y) - F_h(y))dy \qquad \text{(A.6)}$$

Several terms cancel, resulting in further simplification to

$$\frac{\Gamma^c_{kq}}{\mu_{hh}\mu_h^{-2}} = z_q G_{hk} - \int_\eta^{z_k} F_h(y)ydy - G_q G_{hk} + G_q G_k - G_k G_{hq} + H_{hk} \qquad \text{(A.7)}$$

Using integration by parts on the second term gives

$$\int_\eta^{z_k} F_h(y)ydy = z_k G_{hk} - H_{hk} \qquad \text{(A.8)}$$

Substituting (A.8) into (A.7) and re-arranging gives the result required.

To prove the two propositions (35) and (36), note that, from (46), Γ_{kq} can be written in the form given in (49). One can check this equation to show that it is independent of y_q if $F(y_q)=1$ which proves (35). If $F(y_q)=1$ and $F(y_k)=1$ then, as shown in Appendix B, (49) simplifies to B, which proves (36).

Finally, in the case of poverty mixed dominance, one has

$$\mathrm{Cov}\sqrt{N}(\hat{G}_k,\hat{p}_j) = \int_\eta^{z_k} P(y,z_j)dy \qquad \text{(A.9)}$$

Substitution of (22) into (A.9) gives the result (39).

Appendix B Derivations for II.5

To derive equation (49), the following results based on integration by parts are required. Note that the first has already been given in the previous chapter (see III.2) and the third in this chapter as (A.8).

$$\begin{aligned} G_k &= p_k(z_k-\lambda_k) \\ G_{hk} &= p_{hk}(z_k-\lambda_{hk}) \\ H_{hk} &= G_{hk}z_k-\int_{\eta}^{z_k} yF_h(y)dy \\ \int_{\eta}^{z_k} yF_h(y)dy &= \frac{p_{hk}}{2}(z_k^2-\gamma_{hk}^2) \end{aligned} \tag{B.1}$$

Substituting these equations into (31) - (A.1) in the previous Appendix - and combining the result with (46) gives (49). Note that as illustrated in Figures B.1 and B.2 the equations in (B.1) apply whether the distribution is continuous or discrete.

Finally, it is shown that Φ_{zz}=B (as claimed in (55)). This is the case in which $p_k=p_q=1$ (see p.86) so (49) simplifies to

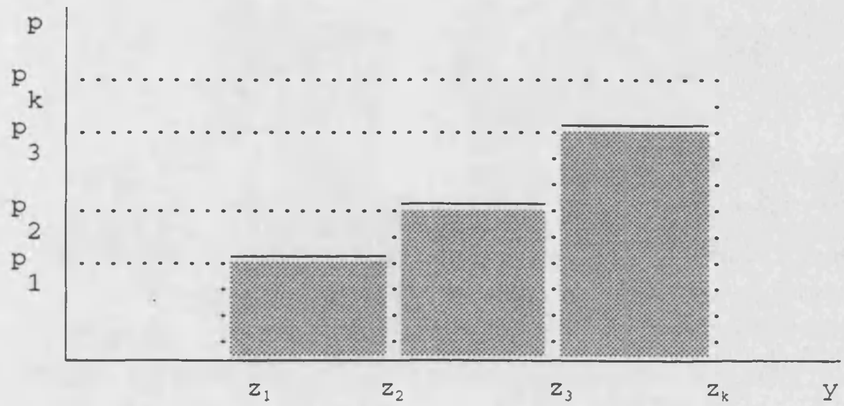
$$\frac{\Phi_{zz}}{\mu_{hh}\mu_h^{-2}} = \gamma_{hz}^2-2\lambda_z\lambda_{hz}+\lambda_z^2 \tag{B.2}$$

Comparison of (17) with (50), (51) and (52) gives

$$\begin{aligned} \lambda_z &= \beta \\ \lambda_{hz} &= \beta_h \\ \gamma_{hz}^2 &= \beta_y \end{aligned} \tag{B.3}$$

Substituting (B.3) into (B.2) gives the result required.

Figure B.1 The Distribution Fumction

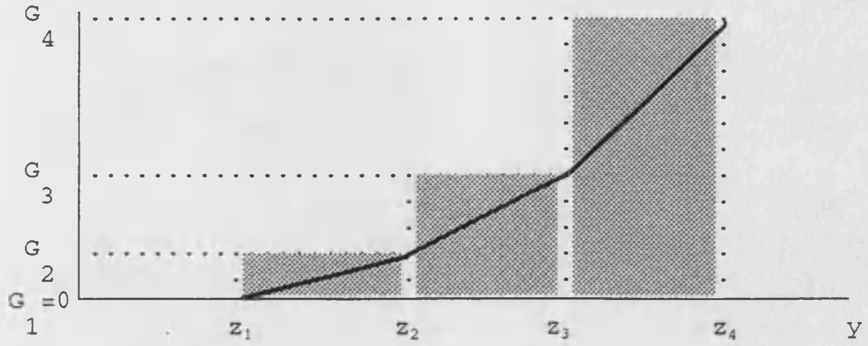


To illustrate how the deficit and Lorenz curves can be linked for a discrete population, consider the graph above. The solid line draws the distribution function (assumed right-continuous) as a step function. The shaded area is G_k. ϕ_k is the blank area in the 'box' bounded by p_k and z_k. So, setting $p_0=0$,

$$p_k z_k = \sum_{i=1}^{k} z_i * (p_i - p_{i-1}) + \sum_{i=2}^{k} p_{i-1}(z_i - z_{i-1}) = \sum_{i=1}^{k} Q(p_i) \int_{p_{i-1}}^{p_i} dp + \sum_{i=1}^{k} F(z_{i-1}) \int_{z_{i-1}}^{z_i} dy$$

$$= \int_0^{p_k} Q(p)dp + \int_0^{z_k} F(y)dy = \phi_k + G_k$$

Figure B.2 The Deficit Curve



The solid line indicates the deficit curve - here the integral of a step function. First, to illustrate (38), note that the area under this line can be decomposed into a difference between rectangles (indicated by the shading) and triangles (the shaded area above the line).

$$H_k = \sum_{i=2}^{k}\left(G_i*(z_i-z_{i-1})-1/2p_{i-1}*(z_i-z_{i-1})^2\right) = \sum_{i=2}^{k}\left(G_i*(z_i-z_{i-1})-1/2(G_i-G_{i-1})(z_i-z_{i-1})\right)$$

$$= \sum_{i=2}^{k} 1/2(G_i+G_{i-1})\int_{z_{i-1}}^{z_i} dy = \int_{0}^{z_k} G(y)dy$$

To illustrate for the discrete case another integration by parts - the third equation of (B.1) - as with Figure B.1 divide the area in the box bounded by $G_k z_k$ into that below G (H_k) and that above it. That is,

$$G_k z_k - H_k = \sum_{i=2}^{k}(G_i-G_{i-1})*z_{i-1} + 1/2(z_i-z_{i-1})^2*p_{i-1} = 1/2\sum_{i=2}^{k}(z_i^2-z_{i-1}^2)*p_{i-1}$$

$$= \sum_{i=2}^{k} F(z_{i-1})\int_{z_{i-1}}^{z_i} ydy = \int_{0}^{z_k} yF(y)dy$$

Finally, one can proceed from this equation to obtain a result similar to that under Figure B.1. The only difference between the two equations is that z has been replaced by its square. This illustrates, for the discrete case, the final equation of (B.1)

$$2\int_{0}^{z_k} yF(y)dy = \sum_{i=2}^{k}(z_i^2-z_{i-1}^2)*p_{i-1} = \sum_{i=2}^{k} p_{i-1}z_i^2 - \sum_{i=1}^{k} z_i^2 p_i + z_k^2 p_k$$

$$= -\sum_{i=1}^{k} z_i^2(p_i-p_{i-1})+z_k^2 p_k = p_k(z_k^2-\gamma_k^2)$$

where p_0 is again equal to zero.

Chapter Three The Influence of Aggregation and Methods of Statistical Inference: Results of a Simulation Study

I Introduction

As stated in the introduction to this thesis, comparison of two distributions in relation to some dominance criterion typically involves much more than choice of criterion. If one has raw data, one can decide on the degree to which one will aggregate (if at all) prior to analysis, or, if the data set is already aggregated, one must decide how to use and interpret it. One can also decide whether or not one will seek to use the data set to hand to investigate a population from which it can be thought to be drawn or whether one will content oneself simply with drawing conclusions about the data set itself. If one is comparing distributions in which two different price levels prevail, one can require the ordering to be invariant to a range of price indices, and, if one is using equivalence scales, a range of equivalence scales, and so on. This chapter is concerned with the first two of these areas of choice: aggregation and statistical method. It uses simulated data to analyze what effect choices concerning them are likely to have on the probability of obtaining a ranking and on the accuracy of one's conclusions.

The problem of statistical inference in relation to dominance criteria was addressed in the previous chapter, in which several estimators were presented as well as a general method for testing for dominance. This chapter provides empirical evidence for a number of the claims made in the previous chapter and also extends the testing framework established in it and addresses various outstanding questions. Areas of interest include the relative performance of the deficit and generalized Lorenz curve estimators, the intersection-union and union-intersection testing methods, and the second-order stochastic dominance (2OSD), restricted dominance, and e-dominance criteria of Chapter One.

Aggregation is achieved by ordering the original distribution from poorest to richest, dividing the ordered units into a number of groups, and using for analysis the means of these groups. Despite its wide use, aggregation has received little attention in the literature on distributional dominance. Cowell (1977) has analyzed the bias it will introduce to summary statistics of dispersion, but its likely impact on the ordering obtained over distributions has not been studied. As Atkinson and Bourguignon write "in practical applications, Lorenz curves are typically drawn for groups of persons" (1989, p.11). The same could be said of virtually all dominance curves. This is, one can surmise, for two reasons. As Coulter, Cowell and Jenkins (1991, p.22) write "most people must rely on published tables rather than microdata". But even if a disaggregated data set is available, researchers commonly choose to aggregate it up prior to analysis, perhaps in an attempt to reduce the influence of individual outliers (possibly subject to measurement error)[1] or simply to ease the computational burden.

Table 1 below provides *prima facie* evidence that the degree of aggregation can be an influential factor on the probability of obtaining a ranking using well-known dominance criteria. This table is based on household income data for 28 provinces - urban areas only - of China, and shows the proportion of the 378 pairs of provinces which can be ranked using different criteria if the data are disaggregated and if the data are aggregated up to deciles (see Howes and Lanjouw (1991) for details of the data set). The effect of aggregation is marked, especially for demanding criteria. Using the e-dominance criterion with bounds of zero and two, three-quarters of all pairs can be ranked using deciles, but only half using disaggregated data. Using the more demanding second-order stochastic dominance (2OSD) criterion, two-thirds of all pairs can be ranked on the basis of deciles, but only one-third if disaggregated data are used.

Table 1 Percentage of Welfare Rankings Achieved Given Different Dominance Criteria and Different Degrees of Disaggregation: an example using Chinese data

Dominance criterion	Degree of aggregation of data	
	Deciles	Disaggregated
E-dominance: $e_{min}=0$, $e_{max}=1$	90.2	85.4
E-dominance $e_{min}=0$, $e_{max}=2$	78.0	53.7
2OSD	67.2	37.3

Note: See Chapter One and II.1 of this chapter for definitions of the dominance criteria

It is interesting in this context to note remarks by Shorrocks in his influential 1983 paper, in which he analyzed a group of 20 countries using data aggregated up to deciles and found he could rank over 80% of all pairs using the welfare 2OSD criterion. This suggested, he wrote, that "the general pessimism concerning the ability to rank distributions is unwarranted" (p.4). To the contrary, the 2OSD criterion was both "sufficiently weak to command a wide degree of support, and sufficiently strong to produce a conclusive ranking in many practical situations"

1. For example, Ravallion (1992, p.43) writes that "It should not be presumed that estimates from unit record data are more accurate than those from grouped data, since the latter can 'average out' errors in the unit-record data."

(p.15).[2] In the light of Table 1, one must ask whether this conclusion is likely to be robust to the degree of aggregation of one's data.

While I am unaware of any other simulation studies which have focused on the issue of aggregation, there are several which have looked at the question of statistical method, particularly in relation to the stochastic dominance criteria. Gail and Gastwirth (1985) and McFadden (1989) study the performance of particular tests (outlined in III.3 of the previous chapter), while Chow, Chakraborti and Thistle (1990) use the criteria of first, second and third-order stochastic dominance to analyze the performance of their general 'union-intersection' inference procedure, also given in the previous chapter (III.2). The latter study is of particular relevance to this chapter and will be returned to at various points.

One important limit on the scope of this chapter is that it is focused primarily on welfare analysis, only secondarily on poverty analysis and not at all on equality analysis. To have thoroughly evaluated the simulated distributions in relation to all three types of criteria would simply have been too onerous. The implications for poverty and equality analysis are, however, mainly straightforward and are briefly considered in the conclusion. All the welfare functions considered satisfy the standard assumptions given in II.1 of Chapter One, that is, they are all defined over a distribution function of a single variable which will be called 'income', are weakly increasing and satisfy the weak transfer principle (are S-concave or 'egalitarian').

The next section sets out in detail the criteria and definitions used in this simulation study and the questions addressed by it. The third section reports the method used and the fourth the results. The fifth and final section concludes.

II Criteria, definitions and questions

II.1 Criteria

The set of dominance criteria examined is restricted to those presented in the first chapter: welfare second-order stochastic dominance (2OSD), unrestricted and restricted, and welfare e-dominance. To briefly recap on these criteria, 2OSD can be defined in relation to either

2. Foster and Shorrocks (1988b) have also since cited this result in relation to poverty analysis, arguing that these "numerous cases" of ranking provide evidence that "there are reasons to be optimistic that [2OSD] will frequently provide the guidance required", that is, that the criterion will produce a "sufficiently strong [ordering] in the sense of producing a conclusive judgement in a large number of cases".

the deficit curve (the integral of the distribution function), using y to define the range over which dominance is required, or the generalized Lorenz (GL) curve (the Lorenz curve times the mean), using p to define the range where p=F(y). One distribution is said to have 2OSD over another if it has a no higher deficit curve or no lower generalized Lorenz curve over the relevant range, and if the inequality holds strongly at at least one point. Unrestricted welfare 2OSD requires dominance from the minimum income of the two distributions to the maximum if the deficit curve is used and from 0 to 1 if the GL curve is used. y-restricted and p-restricted welfare functions require dominance between bounds which lie weakly between the respective unrestricted bounds given above. Unrestricted 2OSD guarantees dominance by the set of weakly increasing, egalitarian welfare functions (that is, all members of the set will regard the distribution with 2OSD as at least as good as the other distribution, and at least one will regard it as better). Restricted 2OSD guarantees dominance by those functions which satisfy the above assumptions but which also are indifferent to the welfare of those with income at or above a certain level (the upper bound, whether given in terms of y or p) and indifferent to transfers of income between those with income at or below a certain level (the lower bound). The rationale for introducing an upper bound is that there may be some segment of the population who are so rich that their well-being, as long as it is above a certain point, is not of concern. Hence one can think of the introduction of an upper bound as changing welfare into poverty analysis. The rationale for introducing a lower bound can also be given a normative basis - there may be some segment of the population so poor that transfers within this group have no effect on welfare - but stems mainly from measurement concerns. Introducing a minimum bound removes the requirement that the dominating distribution have a no lower minimum income. Since it is very difficult to tell which distribution has a no lower minimum, this substantially eases the task of inference.

E-dominance is defined using the isoelastic function $(1/\alpha)y^{\alpha}$, where $\alpha \leq 1$ ($\ln(y)$ if $\alpha=0$), and bounds defined in terms of $e=1-\alpha$, where e is the degree of inequality-aversion (absolute value of the elasticity of the first derivative). If one distribution has e-dominance then all the isoelastic welfare functions within the bounds reckon the dominating distribution to be at least as good as the dominated distribution, and at least one such function strictly prefers it. By controlling the bounds one can control the weight one wishes to give to distributional considerations, with higher values of e indicating that greater weight is being given to equality and less to efficiency.

To use the terminology of the last chapter, the (negative of the) deficit curve, generalized Lorenz curve and e-dominance curves can all be regarded as special cases of dominance curves $\zeta_i=\zeta(F,x_i,\Sigma)$ and $\zeta_i^*=\zeta(F^*,x_i,\Sigma)$, evaluated at x_i and defined for some pair of distribution functions, F and F^*, and set of welfare functions, Σ, itself defined in relation to bounds, x_{min} and x_{max} (x is equal to either y, p or e). Then F dominates F^* by Σ iff $\zeta_i \geq \zeta_i^*$ $\forall x_i$, and $\zeta_i > \zeta_i^*$ $\exists x_i$ such that $x_{min} \leq x_i \leq x_{max}$.

II.2 Definitions

Let the data set actually to hand be referred to as a sample - irrespective of whether it actually is or not - and let the distribution from which the sample is thought of as being drawn be referred to as the population (or the parent), which may or may not differ from the sample in disaggregated form. It is assumed that the researcher has samples, each of size N, drawn from at two populations, each denoted by its distribution function, respectively F and F^*. The three factors of interest to this chapter are (i) the size or range, r, of Σ, (ii) the degree of disaggregation of one's data, DA, and (iii) the use or non-use of statistical techniques, indicated by SI.

Two definitions of 'range' are given, one weaker one stronger. Let $\Sigma_1 \subseteq \Sigma_2$ or $\Sigma_2 \subseteq \Sigma_1$. Let r_1 be the weakly-defined range of Σ_1 and r_2 that of Σ_2. Then $r_1 = r_2$ if $\Sigma_1 = \Sigma_2$, $r_1 < r_2$ if $\Sigma_1 \subset \Sigma_2$ and $r_1 > r_2$ if $\Sigma_2 \subset \Sigma_1$. If neither $\Sigma_1 \subseteq \Sigma_2$ nor $\Sigma_2 \subseteq \Sigma_1$ then the ranges of the two sets cannot be compared. For any set of Σ there could exist many indicators r.

The stronger definition refers only to sets which contain only differentiable and separable welfare functions. Then Σ and thus its range can be defined in terms of s(y), the 'individual' standard-of-living function over which S is obtained by integration (over s(y)dF(y) - see Chapter One, II.3.2). By this definition, any set Σ_1 will be said to have a greater range (strongly-defined) than another set Σ_2 if Σ_1 includes all the functions in Σ_2 plus at least one function which is more concave for all y than any of the functions in Σ_2, and no functions which are less concave. After Pratt (1964), concavity is measured as increasing in the absolute value of the ratio of second to first derivative.[3] Let r_1^s be the strongly-defined range of Σ_1 and r_2^s that of Σ_2. Thus we require that $r_1^s > r_2^s$ iff

$$
\begin{aligned}
&1.\ \Sigma_2 \subset \Sigma_1 \\
&2.\ \min_{\Sigma_1}\left(\frac{-s''(y)}{s'(y)}\right) = \min_{\Sigma_2}\left(\frac{-s''(y)}{s'(y)}\right)\ \forall y \\
&3.\ \max_{\Sigma_1}\left(\frac{-s''(y)}{s'(y)}\right) > \max_{\Sigma_2}\left(\frac{-s''(y)}{s'(y)}\right)\ \forall y.
\end{aligned}
\tag{1}
$$

Note that all r and r^s are only defined up to a positive monotonic transformation, since they are

3. One could equivalently measure concavity using relative rather than absolute risk aversion.

simply measures of rank.[4] Also, since attention is restricted to the set of weakly increasing, egalitarian functions, the set which contains only income utilitarianism, linear in y, has a smaller range (under both definitions) than any other set which contains the linear function and at least one other. And the set of all weakly increasing, egalitarian functions is greater in range than all sets with which it is comparable.

'Aggregation' will be defined by reference to an index of disaggregation, DA, bounded from below by 1 and from above by N, the sample size. Let DA represent the number of quantile groups (vingtiles, deciles, etc.) of a symmetrically aggregated data set. The adverb 'symmetrically' here indicates that the data set is divided into equally sized groups (or at least groups sized as equally as possible subject to any integer constraints). So if DA=1, the data set is fully aggregated - only the mean is given - if DA=10, the data set is aggregated up to deciles - one has a derived data-set of ten observations - and if DA=N the data set is totally disaggregated. Aggregation is carried out by the allocation to each quantile group of that group's mean income.[5]

What are the implications of an ordering of two distributions based on aggregated data? The widespread use of aggregated data notwithstanding, this is not a question much addressed in the literature. Aggregation is generally viewed as a harmless adjunct to analysis. The possibility of aggregated and disaggregated data leading to different conclusions is invariably assumed away without mention. The following, a footnote from Sen, is the only discussion of the issue I could find:

> Non-intersecting Lorenz curves have been often observed in inter-country and inter-temporal comparisons. ... It is, however, worth bearing in mind that the Lorenz curves for actual data are invariably based on size-group averages whereas [the equality 2OSD theorem] would apply to Lorenz curves drawn on a person-by-person basis. There is, therefore, need for caution in facing the usual Lorenz curves armed only with [the equality 2OSD theorem]. (1973, p.58, footnote 8)

Sen's caution has, it must be said, gone unheeded.

This chapter assumes that we are interested in making some sort of conclusion based on the sample distributions about the populations from which they are drawn. Use of aggregated data is therefore interpreted as a technique used, either by choice or lack thereof, to make some sort of inference about the ordering of the two parent distributions. To explain this further, the

---

4. So, the mapping from a particular Σ to a particular value of r can be seen as having three steps. First, a particular indicator r must be chosen which can rank this Σ against others. Then some transform of r must be chosen. And then the value of r corresponding to the Σ of interest can be named.

5. The case in which interpolation is used to approximate the degree of inequality in the original sample is not analyzed, but is discussed briefly in the conclusion.

indicator SI needs to be introduced.

SI is a dichotomous indicator which takes a value of one is statistical techniques are used, and zero otherwise. In more detail, SI takes a value of one if the researcher uses statistical techniques to draw conclusions concerning some two populations from which each of the samples is thought of as respectively being drawn. SI takes a value of zero if the researcher is only aiming to draw conclusions about the two samples or, as I will portray it, if he or she is aiming to make inferences about the populations, but inferences which will be identical to those drawn in relation to the two samples. Sample dominance curves can only be evaluated at a finite number of points, x_1 to x_W. Let $x_1 \geq x_{min}$ and $x_W \leq x_{max}$. Non-statistical or 'analytical' techniques will be said to be used if it is inferred that F dominates F^* (FDF^*) iff

$$\min(\hat{\zeta}_1-\hat{\zeta}_1^*,\ldots,\hat{\zeta}_W-\hat{\zeta}_W^*)\geq 0 \text{ and } \max(\hat{\zeta}_1-\hat{\zeta}_1^*,\ldots,\hat{\zeta}_W-\hat{\zeta}_W^*)>0 \qquad (2)$$

(and analogously for F^*DF) where a cap indicates sample values are used to estimate the dominance curves and the subscript i indicates evaluation at x_i. Use of any other rule will give SI a value of one.

It is assumed that if SI=1 then DA=N, but not vice versa: statistical analysis is assumed to be carried out only with disaggregated data, but disaggregated data may or may not be combined with statistical techniques. If SI=0, then different values of DA can be chosen to evaluate the condition (2).

II.3 The probability of ranking

How are the effects of aggregation and statistical inference to be judged? First, we can examine their impact on the probability of obtaining a ranking, that is, on the likelihood of reaching the conclusion either that FDF^* or that F^*DF. More importantly, we can ask whether each of them leads, on average, to better or worse inferences being made about the ordering of the two populations. With aggregation, our interest is almost entirely restricted to the first issue. In relation to statistical inference, where results relating to the Type I and Type II errors of different statistical methods presented in the previous chapter can be used, we are able to focus more on the second issue.

It is assumed that there are three inferential outcomes of interest: either it is inferred that

FDF* or it is inferred that F*DF or neither is inferred. If and only if either of the first two possibilities occur, we will say that there is a ranking. The probability of achieving a ranking will be labelled ρ. Though ρ can be interpreted as the expected proportion of rankings obtainable from a large set of distributions, based on prior information on the set of distributions, it will be understood in what follows to be the probability of being able to rank any two distributions using some rule of inference.

Which properties can be reasonably attributed to ρ? First, for any combination of values of SI and DA, consider the effect of an increase in r from r_1 to r_2 (where SI and DA are suppressed as constant, as are other arguments held constant such as sample size and prior information on the distributions). Let $\rho(k_1;(\neg)k_2)$ be the probability of obtaining a ranking given k_1 given that there is (is not) a ranking given k_2, where k is either DA or r (or r^s). Then

$$\rho(r_2)=\rho(r_2;r_1)\rho(r_1)+\rho(r_2;\neg r_1)(1-\rho(r_1)) \tag{3}$$

Now assume that $\rho(r_2;\neg r_1)=0$. This is very reasonable, since the only time one can expect to obtain a ranking with r_2 when one is not possible with the smaller r_1 is when living-standards are equal for all S in the smaller set. In this case, use of the bigger set r_2 may break the deadlock. Safely ignoring this remote possibility, it then follows from (3) that ρ will be weakly decreasing in r:

P1. $\rho(r_2) \leq \rho(r_1)$ if $r_1<r_2$

Since this holds for the weaker definition of range it also holds for the stronger.

What about the effect on ρ, for given r or r^s, of an increase in DA? A sufficient condition for ρ to be weakly decreasing in DA, i.e, for

P2. $\rho(DA_2) \leq \rho(DA_1)$ if $DA_1<DA_2$

to hold is that

$$\begin{aligned} &\text{a. } \rho(DA_2;DA_1) \leq \rho(DA_1;DA_2) \\ &\text{b. } \rho(DA_2;\neg DA_1) \leq \rho(DA_1;\neg DA_2) \end{aligned} \tag{4}$$

There are two cases in which (4) is likely to hold.

The first is if the criterion is 2OSD. The shape of the GL curve is determined by the DA conditional means of the distribution for which it is drawn. If $DA_2=nDA_1$, where n is an integer greater than one, then, excluding the case in which disaggregation plays the role of tie-breaking, P2 must hold as a matter of necessity since $\rho(DA_1;DA_2)=1$ and $\rho(DA_2;\neg DA_1)=0$. More generally (i.e. for $DA_2\neq nDA_1$), P2 can still be expected to hold for the 2OSD criterion since the bigger DA, the larger the number of conditional means determining the shape of the GL curve, the less likely all these pairs of conditional means will be ordered in the same way when two GL curves are compared.

The second case is that in which the two distributions from which the samples are drawn are identical, so that the latter are distinguished only by sampling variation. In this case, the plausibility of assuming (4) rests on thinking of disaggregation as adding noise, reducing the correlation between the different estimators of ζ_i and thus making it less likely that the random differences between the samples will all be in the same direction. Then if one **does** achieve a ranking even when the data are noisy, the ranking is likely to survive the reduction of noise occasioned by aggregation. But if one achieves a ranking using aggregated data one can be less confident that it will survive the addition of noise. This gives the first inequality. By the same argument, the probability of achieving a ranking by the reduction of noise, if no ranking is possible in the presence of noise, is greater than the probability of achieving a ranking by the addition of noise when no ranking is possible even when noise is absent. This gives the second inequality.

In the case in which the underlying distributions differ, no definite answer is possible, but again it seems likely that aggregation will increase the probability of attaining a ranking. Consider the case in which S is separable, and an increasing function of the sample's mean, μ_i, and equality Q_i, where i refers to the level of disaggregation. Let the GL curves of F and F^* cross once and let F be the distribution with a higher mean and higher inequality (an everywhere lower Lorenz curve). And let $Q_1=\beta Q_2$ and $Q_1^*=\beta^* Q_2^*$, where β is a measure of the bias upwards in measured equality induced by aggregation and will be no less than one (equal to one in the case of perfect equality). Finally, define the indicator r^s to include the single member Σ of income utilitarianism (which ranks distributions by their means). Then for some Σ with small range, in expectation, FDF^* for both DA_1 and DA_2. For any set with larger range, the higher the expected value taken by the minimum Q_i/Q_i^* for this set, the greater the chance of attaining a ranking. So the probability of attaining a ranking will fall as DA increases if Q_i/Q_i^* falls as DA increases. In turn Q_i/Q_i^* falls as DA increases if $\beta^*<\beta$. This is reasonable since the higher the level of inequality the more one would expect aggregation to exaggerate inequality.[6]

Of course, by this argument one would also expect ρ to be weakly increasing in DA if the high mean distribution is also the low inequality one. But in this case, welfare dominance may be so clear that aggregation has no impact. To summarize, P2 will be likely to hold if Σ is the

---

6. If one restricts oneself to the case in which the distribution functions cross only once, then one can think of aggregation as increasing the value of e required to switch dominance from the high mean to the low mean distribution and thus increasing the probability that a high value of e ranks the distributions in the same way as a low value. (See Chapter One, IV.3 for the notion of 'switch values' of e.)

2OSD set or if the two distributions are identical or if the high mean distribution is also the high inequality one.

If the underlying distributions are equal so that disaggregation is equivalent to adding noise then the more sensitive the criterion used to the presence of noise, the more impact disaggregation should have on the probability of ranking. It is reasonable to think that the more concave a function the more sensitive it is to the addition of noise, as the more it will be dependent on the lower tail. This would give the result

P3 Increasingly decreasing in DA as r^s increases: $\rho(DA_1,r_1^s)-\rho(DA_2,r_1^s)\leq \rho(DA_1,r_2^s)-\rho(DA_2,r_2^s)$ if $DA_1<DA_2$ and $r_1^s<r_2^s$

If ρ is differentiable, this property implies a negative cross derivative between DA and r^s.

P3 would not be expected to hold more generally for samples drawn from non-identical distributions. To return to the case in which the GL curves have a single crossing, here the impact of aggregation may be first increasing and then decreasing. For both small and large r^s, aggregation may have little effect as the higher mean (for small r^s) and the higher inequality (for large r^s) may be the decisive factors at these extremes. It may only be at intermediate values of r^s that aggregation will matter.

The influence of statistical inference on the probability of ranking depends on which method of inference is used. Two methods were compared in the last chapter, the so-called 'intersection-union' (IU) and 'union-intersection' (UI) methods. As defined in that chapter (see equation (76)), let the Z-statistic, Z_i, be a test-statistic for differences in sample dominance curves at ordinate x_i, i=1, ..., W, where $x_1\geq x_{min}$ and $x_W\leq x_{max}$. Then the IU test will infer that FDF^* iff $\min(Z_1, ..., Z_W)>Z_\alpha$, where Z_α is the upper α point of the standard normal distribution. The UI test, advocated by Bishop, Formby and Thistle (BFT, 1989) and Chow, Chakraborti and Thistle (CCT, 1992) will infer FDF^* iff $\max(Z_1, ..., Z_W)>C_\alpha$ and $\min(Z_1, ..., Z_W)\geq -C_\alpha$, where C_α is chosen to control the probability of rejecting the null of equality ($\zeta_i=\zeta_i^*\forall i$) to be no more than α. Put most simply, the IU method requires a large minimum test statistic if dominance FDF^* is to be inferred, whereas the UI method requires only that the maximum test statistic be large and that the minimum test statistic not be large and negative.

As argued in the last chapter (III.4), for the IU method the number of test ordinates should be large. Assume in fact that the Z-statistics are evaluated using the same ordinates used to evaluate (2). For the UI method, a smaller number of ordinates may be appropriate. It follows then from comparison between the IU rule above and (2) that, if the IU method is used, P4 $\rho(SI=1) \leq \rho(SI=0, DA=N)$

However, if the UI method is used the inequality could well be reversed. No general result can be given, but one can note that when BFT used the UI method to order U.S. states in relation to the welfare 2OSD criterion (using the GL curve), they found that it increased the number of rankings obtainable from 60%, based on sample outcomes, i.e., using (2), to 92%.[8] The authors concluded:

> The rankings obtained by applying the statistical tests are much more nearly complete than is common in empirical analyses of income distributions. ... The statistical tests add importantly to the power of these welfare criteria. (pp. 76-77)[9]

II.4 The accuracy of inferences

Ceteris paribus, an increased probability of ranking is to be welcomed: the more discriminating an ordering the better. However, a second desideratum of analysis is a high level of accuracy for one's inferences. Both the use of aggregation and that of statistical tests can have their accuracy as inferential tools measured by their size and power. I define the size of an inference method to be the probability of incorrectly inferring dominance and the power to be the probability of correctly inferring dominance. Hence one wants an inference method which minimizes size and maximizes power. Size and power are in turn dependent on the Type I and II inference error probabilities defined in III.1 of Chapter Two. As there, let the Type I (I^*) inference error be that made if it is inferred that FDF^* (F^*DF) when it is not the case, and a Type II (II^*) inference error as that made if it is not inferred that FDF^* (F^*DF) when it is the case.

There are three possible cases. Either:

(i) FDF^* in which case the size is the probability of a Type I^* inference error and the power one minus the probability of a Type II inference error;

(ii) F^*DF in which case the size is the probability of a Type I inference error and the power one minus the probability of a Type II^* inference error; or,

(iii) neither FDF^* nor F^*DF, in which case the size is the sum of the Type I and I^* inference error probabilities and the power is zero.

Although the simulation results do throw some light on the issue of the accuracy of

---

8. In fact, the 60% proportion of rankings was achieved with aggregation of the data up to vingtiles. The statistical tests also used 20 ordinates.

9. By 'power' here, BFT mean not statistical power but the ability to discriminate or rank.

inferences using aggregated data, more progress can be made in relation to that of the accuracy of statistical inference. Seven areas of investigation are pursued.

1. Most simply of all, it is asked whether there is a need for statistical testing. Are inferences based on sample outcomes reliable?

2. Chapter Two, Section III concluded that the Type I and I* inference error probabilities were controllable using the IU method, but not using the UI method. This study illustrates this conclusion. Note from the definition of size above that if, say, the Type I and I* inference error probabilities are both controlled to be less than α, one can only, on an *a priori* basis, bound the size to be less than 2α (unless one can rule out the possibility of neither distribution dominating). But one might expect the effective bound to be well below 2α. This study asks whether the size as well as the Type I and I* inference error probabilities can be bounded from above by α.

3. No conclusions about the Type II and II* inference error probabilities of the IU and UI methods could be found in Chapter 2 on the basis of theory, but generalizations may be possible on the basis of observation.

4. I investigate whether the IU method is too conservative, that is, whether the desired size can be attained using critical values lower than Z_α, thereby increasing power.[10]

5. One way to increase the power of the IU method is of course to reduce the range over which dominance is required. In the previous chapter, the choice of the bounds, x_{min} and x_{max}, was made prior to testing. Certainly if one is conducting comparisons over a large number of pairs, this approach is the correct one, as one needs to compare each pair using the same criterion. If one is comparing only a small number of distributions, however, one may require a more flexible approach. Say one is carrying out a test for e-dominance and has set one's bounds to 1 and 5. One then finds that one cannot infer dominance over this range, but that one could have if one had modified the lower bound from 1 to 1.1. This is relevant information - F almost dominates F* by the original criterion - but can one then infer dominance between 1.1 and 5? The choice of dominance bounds after inspection of the data is a course of analysis often pursued in non-statistical contexts, at least in a non-formal way. For example, Atkinson and Micklewright, comparing Lorenz curves and using S_x to denote the share of total income of the bottom x%, write

10. This question is not posed in relation to the UI method. Since, as shown, this method can result in very high Type I and I* inference error probabilities, there can be no case for reducing the critical value used.

that "For shares up to S_{70}, Poland does better, but S_{80} and above are higher in the USSR." (1992, p.115). The need here it to turn this approach from simply being one which describes the sample data into one which can be used to make conclusions about the underlying distributions. One natural way in which one might do this is through use of the IU method. One could simply look at a graph of the test statistic, Z, and pronounce dominance in those regions where its absolute value was greater than Z_α. The inadequacy of this approach is that one is not taking into account the possibility of spurious rejection of the null. Two solutions to this problem suggest themselves. The first is to increase the critical value. The second, pursued here for its simplicity, is to set a minimum requirement on the length of the range above the critical value before dominance is pronounced over that range.

If the bounds are endogenous, the null and alternative remain as given in the previous chapter (III.1) though, in this case, x_{min} and x_{max} will be chosen **after** inspection of the sample not before. The following test is proposed:

$$\textbf{Choose } x_{min} \textbf{ and } x_{max}\textbf{. Iff } \min(Z_{min},...,Z_{max})>Z_\alpha \textbf{ and } d(x_{max})-d(x_{min})\geq L \quad \text{then reject } H_0 \textbf{ for chosen } x_{min} \textbf{ and } x_{max} \tag{5}$$

Rejection of the null implies inference of dominance between x_{min} and x_{max}. Z_α has already been defined as the upper α point of the standard normal distribution. The question of interest is the implication of choice of L, the length constraint against which the statistic $d(x_{max})$-$d(x_{min})$ is compared, for the Type I and I^* inference error probabilities. This does not seem to be a question amenable to standard statistical analysis, but guidance may be possible on the basis of examination of simulated data and I return to it in Section IV (IV.2.4). But note for now that L is not necessarily the distance x_{max}-x_{min}, but rather the distance between a function d evaluated at x_{max} and at x_{min}. One might not want to express all distances as linear functions of the ordinates. In particular, dominance over some distance measured in income has a significance which depends on the level of income. $100 might separate the richest from the second richest person in the combined sample, but might also encompass over half of the population in the main body of the distribution.

6. In relation to testing for stochastic dominance, the last chapter presented estimators for both the deficit and the GL curves and showed that test statistics based on these would in general differ (see II.6). It will be of interest to see if any generalizations can be made with regards to the performance of the two types of estimators.

7. Finally, I look at the problem of sensitivity to the lower tail. Note that the lower y_{min} or p_{min} or the higher e_{max}, the more sensitive the respective criteria to the lower tail. Does this cause a

problem for inference? On the basis of empirical work on Côte d'Ivoire, Kakwani (1990, p.17) concluded that the more concave the evaluation function, the less precise the resulting estimate (i.e., the lower the Z-statistic and the less likely the rejection of any null). Cowell (1977, p. 134) has also suggested this might be a problem. Preston (1992, p.7), on the basis of simulation studies, found that choice of evaluation function "does have statistical implications but these depend upon the nature of the ... differences" between the distributions under investigation (p.7). Nevertheless, on the basis of analysis of UK data he did find "slight evidence that ... concavity in welfare functions may be costly in terms of lost efficiency " (p.9), again in the sense of producing a smaller number of significant comparisons.

An extreme form of sensitivity to the left-hand tail is of course the requirement, applicable to the 2OSD criterion, that a distribution can dominate only if it has a no-lower minimum. The difficulty of inferring minimum dominance was discussed in Section IV of Chapter Two. This chapter does focus on the performance of the GL and deficit estimators at minimum values of p and y respectively, and it illustrates the problems thereby caused for the inference of unrestricted 2OSD. However, it does not directly tackle the question of whether minimum dominance can be successfully inferred. In the last chapter, it was argued that the presence of outliers (long tails) and the likelihood of measurement error will make the inference of minimum dominance an impossible task. Since the distributions of this chapter come from two distributions, lognormal and beta, neither of which have long left-hand tails, and since there is no measurement error, the claims of the last chapter in this regard cannot be further pursued.

III Method

To explore the issues raised in II.3 and II.4, inferences concerning the criteria given in II.1 were made on the basis of samples drawn repeatedly from various pairs of parent populations. The simplifying assumption that the observations are unweighted was invoked. Although the derivations of Chapter Two allow for the common case in which observations are randomly weighted, it is simpler to assume that they are not, and unlikely to change substantially any of the conclusions drawn. In all, seven pairs of populations were chosen. Their characteristics are summarized in Table 2. The first distribution of any pair is named F, the second F^*. The first column of Table 2 gives the letter, A to G, by which each pair of distributions is labelled. The second gives the distributional family for each distribution. The third column gives a general characterization for each pair, which will be explained below. The fourth and fifth columns give the mean and coefficient of variation (standard deviation divided by the mean) for each distribution. The sixth to ninth columns give information about the ordering between the two distributions in each pair in relation to various dominance criteria: welfare first-order stochastic dominance (1OSD) and welfare and equality 2OSD and welfare e-dominance. A 'D' indicates dominance, 'X' no dominance, and '=' identity (a special case of no dominance). If there are crossings, the point at which these occur is given for the 2OSD and e-dominance welfare criteria. The final column gives Gini coefficients.

Greater familiarity with the seven pairs of parent distributions can be obtained from Figures 1.A to 1.G, presented at the end of this section, which provide, for each pair, the poverty gaps and conditional means for the two distributions. The poverty gap curve is the deficit curve divided by y, and the conditional mean curve is the generalized Lorenz curve divided by p. In each case, the former curve tends to bring out more clearly the differences between the two distributions and hence is used in preference.

Table 2 Parent Distributions

ID	Type	Character-ization	Means (K=shift)	CVs	1OSD welfare	2OSD welfare	2OSD equality	E-dominance welfare	Ginis
A	LogN v. LogN	Identical, high inequality	100 100	.7 .7	F=F*	F=F*	F=F*	F=F*	.344 .344
B	LogN v. LogN	Identical, low inequality	100 100	.5 .5	F=F*	F=F*	F=F*	F=F*	.261 .261
C	LogN v. LogN	Clear single crossing	110 100	.7 .6	FXF*	FXF* y=86 p=.49	F*DF	FXF* 2.1	.344 .305
D	LogN v. Shifted LogN	Slight single crossing (at left tail)	115 100 (K=15)	.6 .6	FXF*	FXF* y=48 p=.11	F*DF	FXF* 4	.305 .295
E	LogN v. Beta	Slight double crossing	110 100	.76 .58	FXF*	FXF* y=121, 139 p=.68,.76	FXF*	FDF*	.365 .323
F	Shifted LogN v. LogN	Strong dominance	110 (K=10) 100	.64 .7	FDF*	FDF*	FDF*	FDF*	.314 .344
G	LogN v. LogN	Dominance but not strong dominance	100 100	.6 .7	FXF*	FDF*	FDF*	FDF*	.305 .344

Notes: 1. The parameters for the beta distribution are v=2, w=6, M=400. See below for definitions.

2. The orderings by the dominance criteria, and the crossing points, were based on construction, by discrete approximation, of, respectively, deficit curves, generalized Lorenz curves, Lorenz curves and e-dominance curves given the known distribution functions of the populations. The Ginis were similarly calculated.

3. 'LogN' stands for lognormal.

As can be seen from Table 2, the two distributions used to generate these samples are the lognormal and the beta, which can be characterized as follows:

Lognormal

Range: $y-K>0$

Relative frequency function: $f(y;m,\sigma,K) = ((y-K)\ \sigma\ \sqrt{2\pi})^{-1}\exp(-½(\log(y-K)-m)/\sigma^2)$

Mean: $\mu(y) = \exp(m+½\sigma^2)+K$

CV: $CV(y) = \sqrt{}(\exp(\sigma^2)-1)*(1-K/\mu(y))$

Beta

Range: $0<y<M$

Relative frequency function: $f(y;v,w,M) = (y/M)^{v-1}(1-(y/M)^{w-1}/B(v,w)$

where $B(v,w)=\int_0^1 u^{v-1}(1-u)^{w-1}du$

Mean: $\mu(y) = M*v/(v+w)$

CV: $CV(y) = \{w/[v(v+w+1)]\}^{.5}$

A variable has a shifted lognormal distribution with mean and coefficient of variation as given above if the variable minus some constant shift factor, K, has a natural log which is distributed normally with mean m and coefficient of variation σ/m. Non-shifted lognormal distributions have the characteristic that their Lorenz curves never cross. Hence their GL curves and deficit curves cross at most once (as do all members of the so-called two-parameter families of distributions, including the normal and gamma as well as lognormal). The beta distribution can generate Lorenz curves which cross and is defined up to a scalar factor, M, by the two parameters v and w.

Although also involving a large degree of arbitrariness, the choice of distributions was governed by the pursuit of five objectives:

1. While dominance criteria can be applied to various fields, the focus of this study is the analysis of living standards. Hence the choice of distributions and degrees of dispersion they display should reflect what is found in real-world purchasing-power distributions.

2. This study aims to simulate the position of the researcher working with two samples of - aggregated or disaggregated - household survey data for two regions in the same country or for two fairly similar countries or for the same spatial population at different, but relatively close periods of time. Hence there should be 'realistic' differences between mean incomes and degrees of inequality within a pair. The underlying assumption here is that most comparisons of interest are between populations not where it is obvious that one distribution is welfare-superior, but where there is genuine uncertainty and where data analysis is required to examine, say, the effects of a change in government policy on living standards.

3. Each pair of distributions should be characterised by, and thus illustrate, an archetypal ordering of populations. This is discussed further below.

4. As few changes as possible should be made between each pair of distributions. The fewer the changes, the easier it should be to isolate the reason for a change in the performance of some or other inference technique.

5. Computational convenience and simplicity should be maximized.

The first, fourth and fifth aims led to the choice of the lognormal distribution as the basic distribution on which to base this simulation study. The lognormal is generally believed to represent fairly well real-world income distributions, except possibly at the tails.[11] It is also a simple distribution to work with.

The third aim, to have each pair characterised by an archetypal ordering of populations, was crucial. For this to be realized, a dominance criterion first needed to be chosen, so that the ordering could be determined. The natural candidate was the unrestricted 2OSD criterion, both because of the range of the set of welfare functions over which 2OSD gives dominance and because it is a focus of much of this study. For convenience, the GL curve was taken as the 2OSD dominance curve. Pursuit of this aim also required a categorization of the various ways in which two dominance curves might be related. While other categorizations are possible, one tripartite grouping is into: those cases in which the dominance curves are identical; those in which one curve is dominant; and those in which the curves cross. Further subdivisions within these three groupings are also possible. Within the first group, one might have the identical curves generated by identical distributions with 'high' inequality or by identical distributions with 'low' inequality. (Since all the dominance criteria used are unit-invariant, there is no need to distinguish between high-mean and low-mean pairs of distributions). Within the second group, one might have strong dominance (that is, one dominance curve always lying strictly above the other, except where they are equal by construction - see Chapter Two, Section I) or one might have dominance without strong dominance (by having some point at which the dominance curves touch). Within the third category, one might allow for any number of crossings, and one might also distinguish between a 'clear' and a 'slight' crossing. A 'clear' crossing can be defined as one which results in both F and F^* dominating for a substantial distance and a 'slight' crossing as one which results in one distribution dominating for a significantly greater distance. The seven pairs chosen from these possible characterizations were as follows:

A. identical, high inequality
B. identical, low inequality

11. See Aitchinson and Brown (1957, Ch.11).

C. clear single crossing
D. slight single crossing
E. slight double crossing
F. strong dominance
G. dominance, but not strong dominance

With the functional forms chosen, and the pairs characterized, there remained the requirement of parameterization. As already mentioned, the choice of means is only material up to a scale factor. The range of inequality levels was chosen, with the first aim in mind, to approximate levels of inequality found in real-world distributions of income, using the Gini coefficient as a guide.Somewhat arbitrarily, a Gini of around .35 was taken to be relatively high, while one of around .25 relatively low.

The CV's of pair A were set at .7 (giving a Gini of .344) and those of pair B at .5 (giving a Gini of .261). The means of the distributions of these first two pairs were all set at 100. Pair C (characterizing a clear crossing) was obtained by raising the mean of the first distribution by 10 to 110, and decreasing the CV of the second distribution to .6, thus giving a crossing of the GL curves at $p \approx .5$. The second distribution of Pair D also has a mean of 100 and CV of .6, but also has a shift factor of 15, making income more equally distributed. The first distribution of the pair has the same CV, and a higher mean (of 115) but no shift factor, all of which combine to give dominance except at low levels of income. The GL curves cross at $p \approx .1$.[12] In terms of means and CVs, pair E is very similar to C. The means are the same (110 and 100), the first distribution is given slightly greater inequality (.76 to .7), and the second slightly less in terms of the CV (.58 to .6), but due to the functional form (beta rather than lognormal) slightly more in terms of the Gini (.323 to .305). This parameterization ensures a slight double crossing. F's generalized Lorenz curve lies everywhere above that of F^*'s except between approximately .7 and .8. Note that the curves are very close over this range: from Figure 1.E, they appear to be coincident. Since the crossing is so slight the pair is characterized by third-order stochastic dominance and thus be e-dominance for all non-negative values of e (see Chapter One, IV.3). Pair F (characterizing strong dominance) differs from pair A simply by the fact that one of the distributions is 'shifted' by 10 to give it strong dominance. The final pair G returns both means to 100, but gives F a lower CV (.6 to F^*'s .7), thus giving F dominance, but not strong dominance: the generalized Lorenz curves

12. Introducing a shift makes the graphical characterization clearer. If two non-shifted log-normals are used to illustrate the case of a crossing at the lower tail, it becomes very hard to distinguish their GL curves at the lower tail.

touch at p=1.[13]

The remaining issues of method can be dealt with briefly. The choice of sample size was dictated by the second and fifth aims. Some household surveys are very large. For example, Ravallion (1992, p.86) utilizes two Indonesian household survey data sets from 1984 and 1987, each with 50,000 households. But the range is also large. The Family Expenditure Survey of the U.K. has an annual sample size of 7000 households (Atkinson and Micklewright,1992, p.273) and the provincial samples I utilize in Chapter Five of this thesis range in size from 700 to 1500 households. I work at the lower end of this scale of sample size with three sample sizes of 500, 2000 and 4000. Compensating for this somewhat are the facts that the differences between the distributions are, the third objective notwithstanding, relatively large. Mean incomes differ by between 10 and 15%, and Ginis by between .03 and .04 points.

Three rounds of simulations were conducted. Within each round, the sample size was held constant, both within each pair and from one simulation to another. For the first round, a sample size of 500 was used, for the second, the sample size was increased to 2000 and, for the third, it was increased again to 4000. All rounds consisted of 1000 simulations or draws.

Finally, the samples were generated using random number generation routines in SAS. The lognormal samples were obtained by transforming generated normal distributions. To generate the beta samples, I took advantage of the fact that if, as here, the beta parameters, v and w, are integers and the scale factor is set equal to one, then the beta distribution β:v,w is distributed as γ:v/(γ:v+γ:w), where γ:v and γ:w are gamma distributions with parameters v and w respectively (see Hastings and Peacock, 1974).

13. The choice of distributions can be compared to those of Gail and Gastwirth (1985) and CCT. The former use the uniform, unit exponential, Pareto, and gamma distributions. Of these, only the gamma can be said to approximate real-world income distributions, though the Pareto is often said to approximate well the upper-tail of many income distributions. The gamma is, like the log-normal, a two-parameter distribution and should give quite similar results. CCT use three distributional families, the lognormal, normal and uniform. Again, if the first objective is applied, only the lognormal is relevant. CCT's choice of relative means is similar to ours (their three pairs of means are between 1.3 and 1.1), but their CV's are lower, between .46 and .29, corresponding to Ginis of between .16 and .24, and somewhat further apart.

Figures 1.A-1.G Population Distributions

Legend — F ---- F*

Figure 1.A Identical, High Inequality

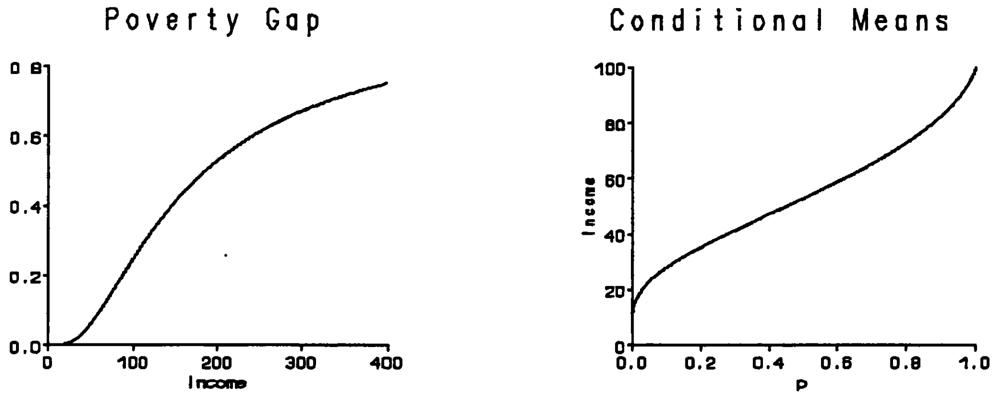


Figure 1.B Identical, Low Inequality

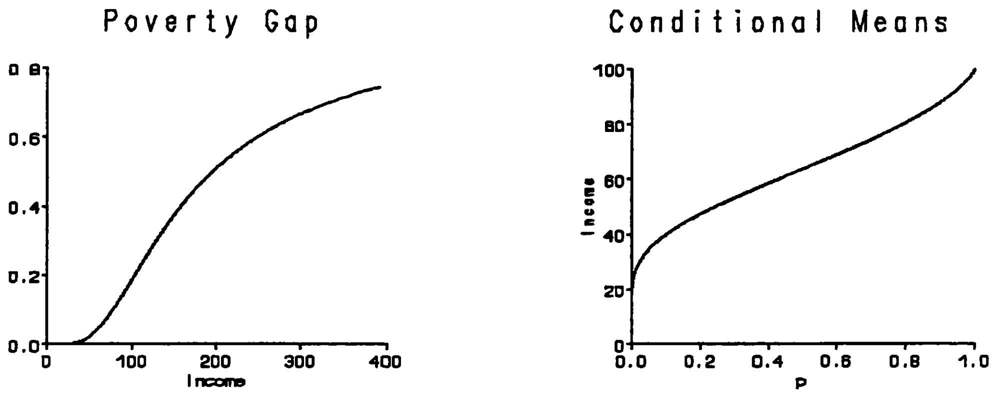
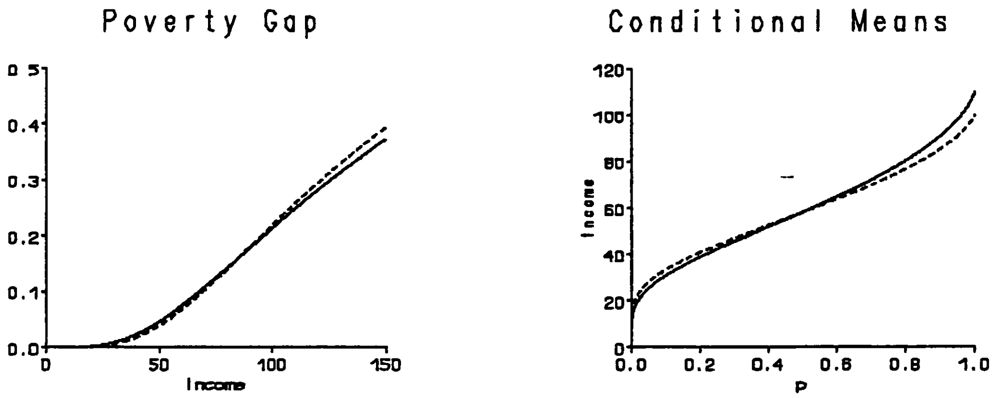


Figure 1.C Clear Single Crossing



<u>Notes for Figures 1.A-1.G</u>: The labels (A to G) give the I.D. of each pair and the titles give each pair's characterization. The poverty-gap curve is the deficit curve divided by the income level, and gives the amount of income required, as a fraction of the relevant income level, to bring all incomes up to that level. The conditional-mean curve is the GL curve divided by the cumulative proportion, p, of the population (ordered in ascending order of income), and gives the mean income conditional on being in the poorest p. Since in pairs A and B the two populations are identical, the two curves are entirely coincident.

Figure 1.D Slight Single Crossing

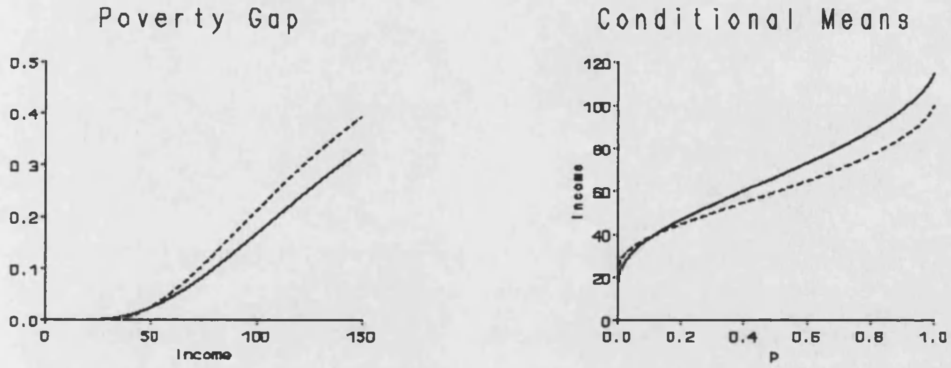


Figure 1.E Slight Double Crossing

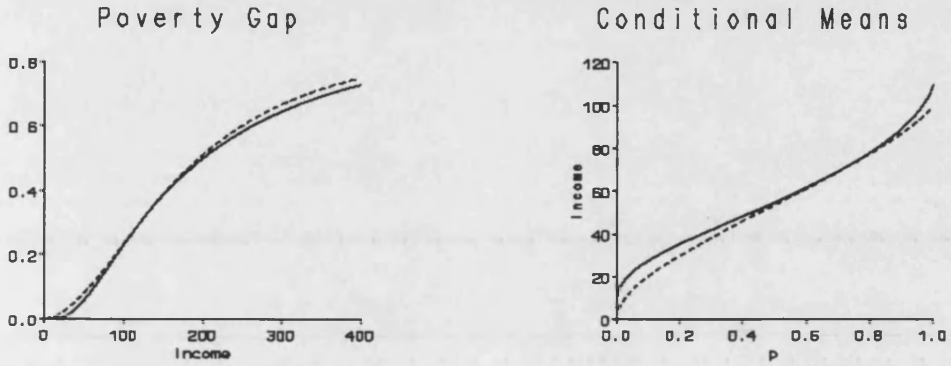


Figure 1.F Strong Dominance

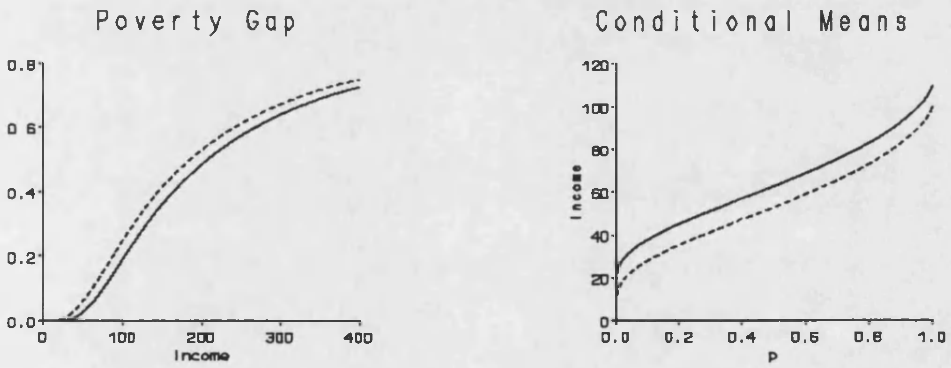
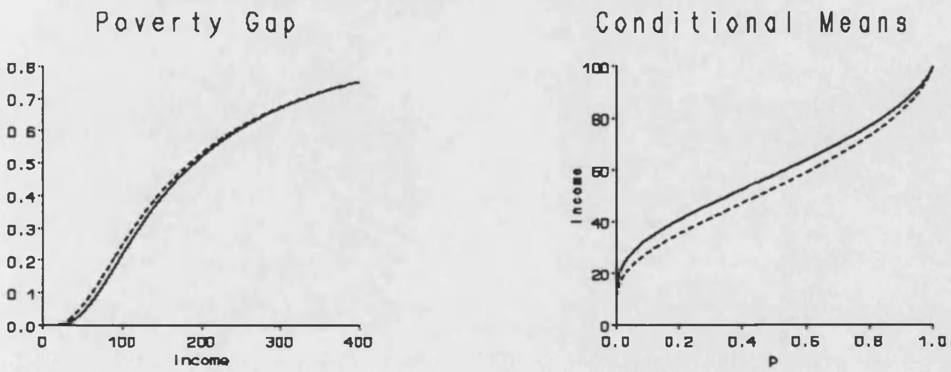


Figure 1.G Dominance, but not Strong Dominance



Legend —— F ---- F*

IV Results

The results are presented in a series of figures and tables analyzed below under two headings, aggregation (IV.1) and statistical method (IV.2). For the reader who is only interested in the main conclusions or in just one or two areas, precis tables, at the start of each of the two sub-sections, summarize: the questions asked of the results; the answers; the implications of the answers; and where further details can be found.

A brief introduction to the general format of the tables of results is required. The first column of each table gives the label of the pair of distributions (one of seven, A to G), as well as that of the simulation round (one of three, with sample sizes respectively of 500, 2000 and 4000). Thus D_2 refers to the second round of simulations (based on a sample size of 2000) drawing (1000 times) from Pair D's distributions. The next column reminds the reader of the ordering within the pair.

For those pairs in which, for the given criterion, neither distribution dominates, the figure given in each column is typically the size of the inference method, the proportion (expressed as a percentage) of sample pairs within each round which (wrongly) lead to the inference of dominance (either that FDF^* or that F^*DF). Included in this category are always the first two pairs, A and B (since these are pairs of identical distributions), and, usually, pairs C, D and E in the case of 2OSD and pairs C and D in the case of e-dominance, though here it depends on how the bounds are drawn. For those pairs characterized by dominance, the figure given is each column is typically the power of the inference method, the proportion (expressed as a percentage) of sample pairs within each round which correctly lead to the inference of dominance. Each table separates the size from the power figures.

As mentioned in II.4, the size and power figures provide very simple means of evaluating the different inference methods. If neither distribution dominates, we do not want dominance inferred, so we want a small size. If one distribution dominates, we want dominance by that distribution inferred, so we want high power. Since the convention is followed that if there is a dominating distribution it is F, the power is typically one minus the Type II inference error probability, expressed as a percentage. And since the size is only given in cases in which there is no dominance, it is equal to the sum of the Type I and Type I^* inference error probabilities, again expressed as a percentage. It would of course be possible to present information directly on the Type I and I^* errors instead of on the size, but this would double the number of tables for little gain. In addition, one could also, for example, give the estimated Type I^* error probabilities in cases in which FDF*. However, these were consistently found to be negligible if not zero.[14] Note that for this reason, the power of the tests is almost if not exactly ρ, the probability of ranking. Where the size is given, it is always equal to ρ.

IV.1 Aggregation

Five different levels of aggregation were chosen: complete disaggregation, and aggregation up to, respectively, percentiles, vingtiles, deciles and quintiles. These correspond to values of DA of, respectively, N, 100, 20, 10 and 5. Tests for 2OSD, using the generalized Lorenz curve, and for e-dominance were carried out using each of these. The key results are summarized in the precis table immediately below. The sub-sections following give more details.

Table 3 Precis of Results Pertaining to Aggregation

Question	Answer	Implications	For details
1. Does P2 hold (is ρ falling in DA using the e-dominance and 2OSD criteria)?	Nearly always yes; often by large amounts, especially for 2OSD; violations are insignificant.	Claims of discriminatory power of dominance criteria unlikely to be robust to degree of aggregation of data.	Table 3 and Figures 2.A-2.G
2. Does P3 hold (is ρ increasingly falling in DA as r_s increases)?	Yes for identical distributions; but not in general.		Figures 2.A-2.G
3. Are inferences based on aggregated data more accurate than those based on disaggregated data?	No general conclusions, but aggregation will exaggerate the differences between similar distributions, and tend to overlook crossings at low levels of p and high levels of e.	Dominance using aggregated data for given lower bound on p and upper bound on e should be taken as a basis for inferring dominance using disaggregated data in relation to higher lower bound on p and lower upper bound on e.	Table 3 and Figures 2.A-2

14. The reason for this is simply that if FDF*, although one might not infer FDF*, it is highly unlikely that one would infer the opposite, i.e., that F*DF.

IV.1.1 Second-order Stochastic dominance

Each of the five columns of figures in Table 4 is headed by the degree of disaggregation on which the analysis was based. 'Disagg' refers to the use of disaggregated data, '100' to that of percentiles, and so on. The main results can be most conveniently presented in bullet-form.

• Table 4 illustrates the truth of P2 of II.2 when the criterion is 2OSD: the probability of ranking is non-decreasing in the degree of aggregation.

• The degree to which this probability is increasing if at all depends on the nature of the underlying distributions. In three out of the seven pairs the effects of aggregation are marked. In pairs A and B, where the distributions within each pair are identical, even aggregating only up to percentiles increases ρ substantially, from around .1-.2 to .25-.3. Moving up to deciles gives a ρ of .4-.5 and up to quintiles one of .5-.6. In other words, one goes from having a one-in-five or lower chance of obtaining a ranking using disaggregated data to having a one-in-two or greater chance using quintiles. Aggregation has an even more marked effect in pair D, characterized by a crossing in the left-hand tail. ρ starts off at around zero, increases to .4 if deciles are used, and to between .8 and 1 with quintiles. This dramatic change is to be explained by the fact that F's GL curve lies above that of F^* for p>.11. In the other four cases, aggregation is not as important a consideration. In pair F, the dominance of the first distribution is sufficiently clear to be manifest even in disaggregated data. Pairs E and G are of special interest. Despite the crossing of pair E being slight it is not at the left-hand tail and so it not overlooked by aggregation. Pair G shows very little change, since ρ is restricted by the two sample means being equal in expectation, a feature which is of course invariant to the degree of aggregation. The conclusion must be that it is when the GL curves are either identical or close at the lower tail that aggregation will have its biggest impact.

• There is no determinate relationship between sample size and the effect of aggregation on the probability of ranking. Sample size does matter, but its effects can go either way. In pair C, for example, the impact of aggregation, though always positive, is falling with sample size, whereas in pairs A, B and D it is increasing.

• In the cases given here, use of aggregated data would, on average, worsen the accuracy of the inferences being made. Where there is dominance, aggregation increases power only slightly, but where there is no dominance aggregation increases size more markedly (except in pairs C and E). However, this is not a robust conclusion. If one had the case in which F dominated F^*, but the two distributions had identical lower tails, then use of disaggregated samples could give a ρ of around .5. Use of aggregated data on the other hand could give a much higher ρ. But two conclusions concerning accuracy are possible:

(a) From pairs A and B, aggregation will exaggerate the differences between similar distributions.

(b) From pair D, dominance in relation to aggregated data will be a more accurate predictor of restricted than unrestricted dominance. Of course, this follows as a matter of necessity since unrestricted implies restricted dominance, but not vice versa - hence one will always be safer inferring the latter. But there is a more specific and compelling reason for this conclusion. Let the proportion of the sample covered by the poorest quantile group , e.g., .1 in the case of deciles, be referred to as the 'first aggregation point'. Then, as pair D illustrates, unless one uses restricted dominance and sets the lower bound, p_{min}, above the first aggregation point, one is in danger of inferring dominance even if there is a crossing below p_{min}.

IV.1.2 E-dominance

To examine the effect of aggregation on rankings using criteria other than 2OSD, the e-dominance criterion was used. First, some lower bound on e, the inequality aversion parameter, was chosen. Then e was increased, and the change in the probability of ranking observed for the various levels of aggregation given earlier. Note that as e is increasing so too is r^s, the range of Σ, strongly defined. Hence we can use this exercise to investigate the relationship between r^s and DA. e=0 was chosen as the lower bound. (A lower bound of one was also experimented with, but with similar outcomes.) The maximum value of e used was five. The results can be seen in Figures 2.A to 2.G, where the letters refer to the seven pairs. For pairs A to D, graphs are drawn for all three sample sizes. For pairs E to G, where there is very little change with sample size, only the results for a sample size of 500 are shown. Each graph displays the probability of ranking starting with e equal to zero, and increasing the range of e along the x-axis, both for disaggregated data and for data aggregated up to percentiles, vingtiles, deciles and quintiles. For comparative purposes, the value taken by ρ when the 2OSD criterion is used is also given. A vertical line marks off the value 'SD' from the rest of the x-axis which gives different values of e.

Table 4 The Effects of Aggregation on 2OSD Orderings

	Pop'n	Disagg	100	20	10	5
			Size			
A_1	F=F*	19.6	25.7	38.9	47.3	56.9
A_2		13.9	26.8	36.2	42.9	53.1
A_3		12.2	27.2	39.6	46.3	55.0
B_1	F=F*	22.5	30.3	40.6	48.8	58.3
B_2		16.9	29.4	40.5	48.4	57.4
B_3		13.9	27.1	37.9	46.5	56.3
C_1	FXF*	2.6	3.6	5.9	8.5	17.5
C_2		0.0	0.2	0.4	0.5	2.5
C_3		0.1	0.1	0.1	0.1	0.6
D_1	FXF*	1.0	3.2	15.6	44.0	85.4
D_2		0.0	0.0	2.7	39.4	97.8
D_3		0.0	0.0	0.3	34.6	99.4
E_1	FXF*	45.9	46.3	46.6	47.2	49.0
E_2		46.6	46.7	47.4	48.2	52.1
E_3		47.9	47.9	48.0	49.3	55.9
			Power			
	Pop'n	Disagg	100	20	10	5
F_1	FDF*	98.5	98.9	99.1	99.1	99.1
F_2		100	100	100	100	100
F_3		100	100	100	100	100
G_1	FDF*	42.5	49.4	51.7	52.1	52.2
G_2		44.0	49.9	50.0	50.0	50.0
G_3		45.1	51.2	51.2	51.2	51.2

Notes: 1. See the introduction to Section IV for an explanation of the first two columns.
2. The figures given under the heading 'Size' ('Power') are the percentage of cases in which dominance was wrongly (correctly) inferred.
3. The sub-headings 'Disagg', '100', '20', '10' and '5' refer to the degrees of aggregation of data used to conduct the analysis. 'Disagg' stands for disaggregated, '100' for percentiles and so on.

There are three main results.

• The evidence from using the isoelastic function is that P2 will hold in general for all Σ and not just in relation to 2OSD criterion. It will not hold as a matter of necessity, but any violations of it seem to be mild. Figures 2.A and 2.B provide strong evidence to support the proposition (put forward in II.3) that the probability of ranking will fall with the degree of disaggregation of the data if the distributions are identical. Figures 2.C and 2.D, especially the latter, bear out the argument of II.3 that aggregation will increase the probability of ranking if the high-mean is also the high-inequality distribution. However, the upward bias generated by use of aggregated data is smaller if criteria more restrictive than 2OSD are used. This can be explained by the much higher levels of ρ obtained with disaggregated data using e-dominance (even with a relatively large e_{max}, such as 5) than using 2OSD. In the first two pairs, for example, ρ is around .5 for e_{min}=0 and e_{max}=5, compared to the .1 to .2 obtained using 2OSD. These higher levels provide an upper bound on the degree of bias introduced by aggregation. Figure 2.E shows that sometimes aggregation will systematically *decrease* the probability of ranking. (Since the Lorenz curves of pair E's distributions cross, the argument of II.3 cannot be applied.) However, note that the bias downwards is very small, just a couple of percentage points.

• P3 - the proposition that the increase in ρ due to aggregation is increasing in r^s - does not hold generally. If P3 holds, the vertical differences between the different curves in each of Figure 2's graphs should be increasing with ρ. For pairs A and B these differences are everywhere increasing, but for pair C they first rise and then fall, while for pair D it depends on the degree of aggregation. This suggests that P2 will hold if the underlying distributions are very similar (A and B) and possibly if there is a crossing at the lower tail (C), but not more generally.

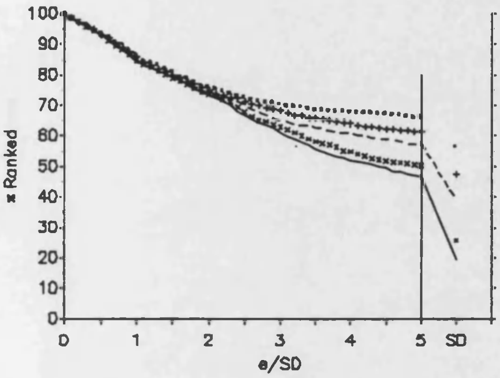
• In relation to accuracy, the conclusions are the same as those drawn for 2OSD. Differences between similar distributions will be exaggerated by aggregation. And e-dominance using a given upper bound based on aggregated data is a better predictor of e-dominance based on disaggregated data with a lower upper bound than with the same upper bound. Again this is not simply because e-dominance using a higher upper bound implies e-dominance using a lower upper bound but not vice versa, but because aggregation will ignore crossings in the GL curves below the first aggregation point. Pair D illustrates. If one uses the ordering based on aggregated data to predict the ordering based on disaggregated data and uses bounds in both cases of 0 and 5, one will be led badly astray, as one will tend to predict dominance where there is none. But if one retains the wide bounds in relation to aggregated data, but uses them to predict dominance based on disaggregated data using narrower bounds, say between 0 and 3, one will be much more accurate.

Figures 2.A-2.G The Probability of Ranking, e-dominance and 2OSD

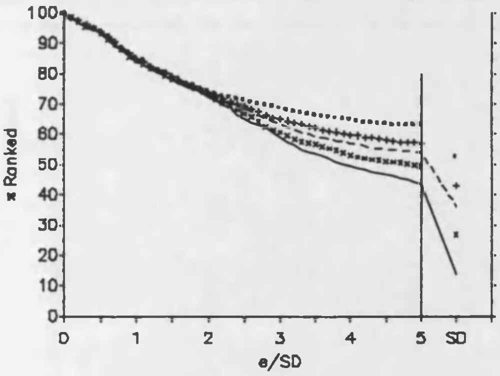
Legend: ___ Disaggregated xxx Percentiles --- Vingtiles +++ Deciles • • • Quintiles

Figure 2.A Identical, High Inequality

Sample Size=500



Sample Size=2000



Sample Size=4000

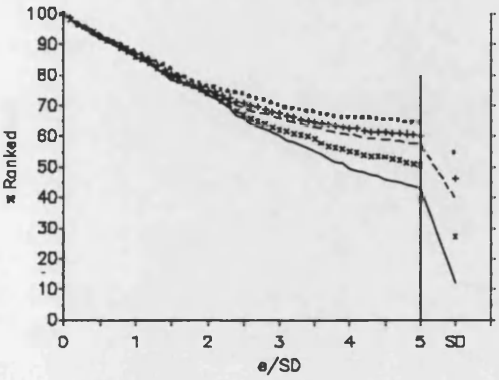
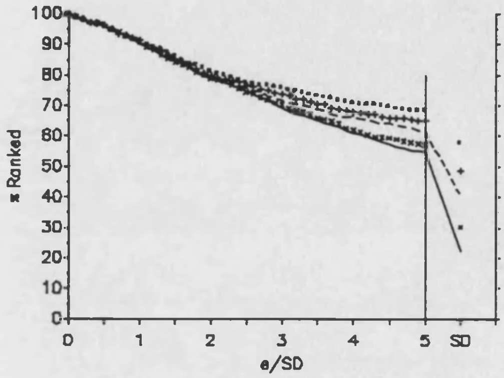
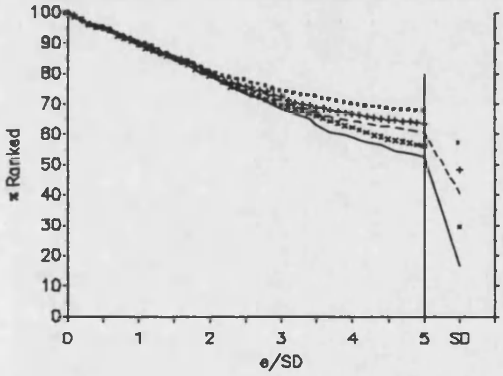


Figure 2.B Identical, Low Inequality

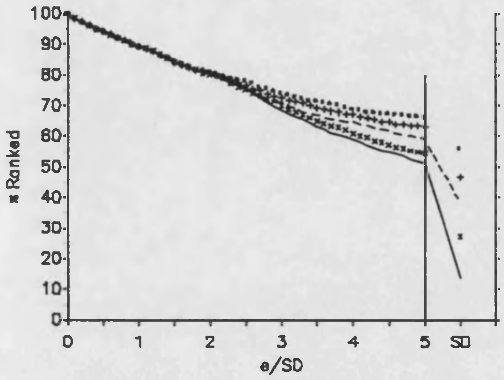
Sample Size=500



Sample Size=2000



Sample Size=4000



Notes: See two pages on.

Legend: ___ Disaggregated xxx Percentiles --- Vingtiles +++ Deciles • • • Quintiles

Figure 2.C Clear Single Crossing

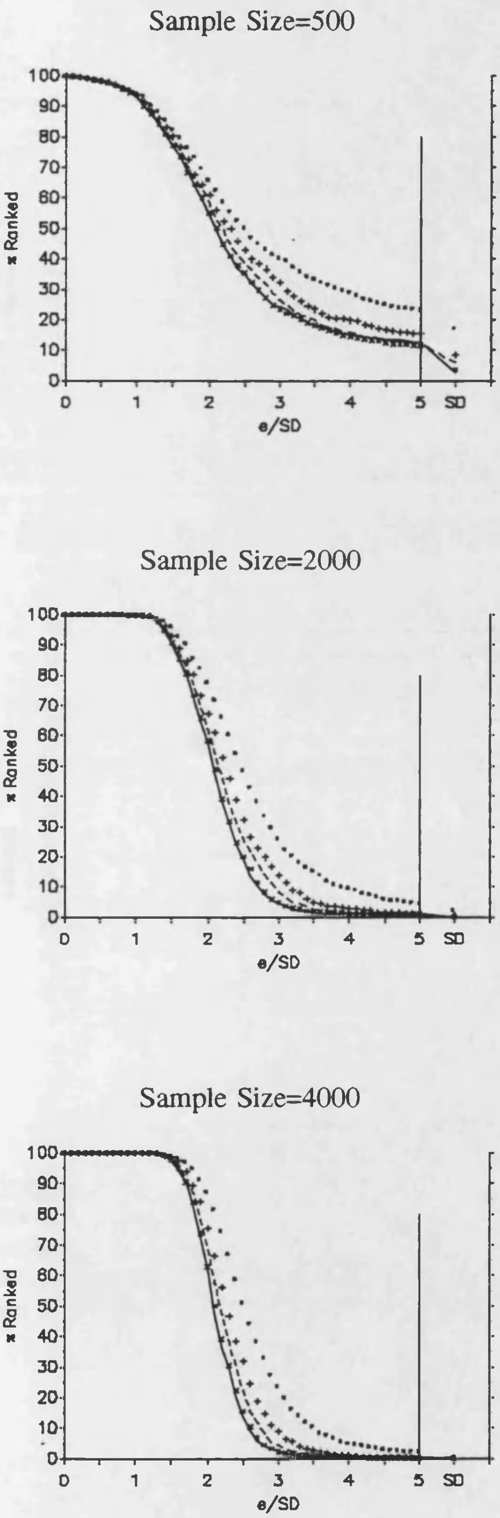
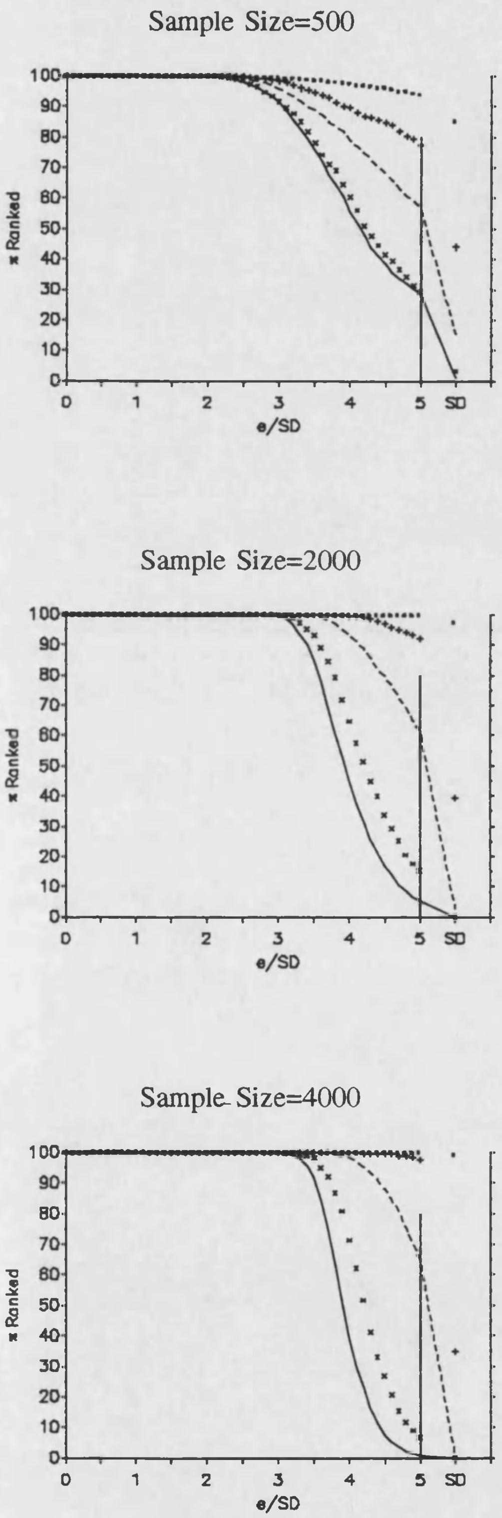


Figure 2.D Slight Single Crossing



Notes: See next page.

Legend: ___ Disaggregated xxx Percentiles --- Vingtiles +++ Deciles ••• Quintiles

Figure 2.E Slight Double Crossing

Sample Size=500

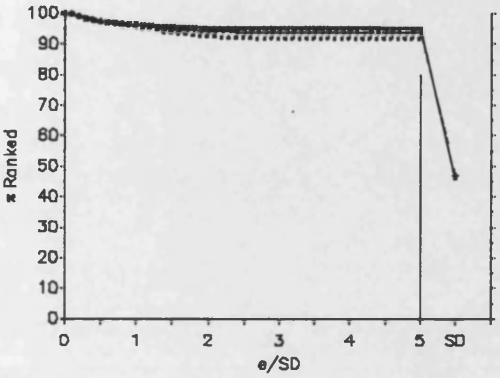


Figure 2.G Dominance but not Strong Dominance

Sample Size=500

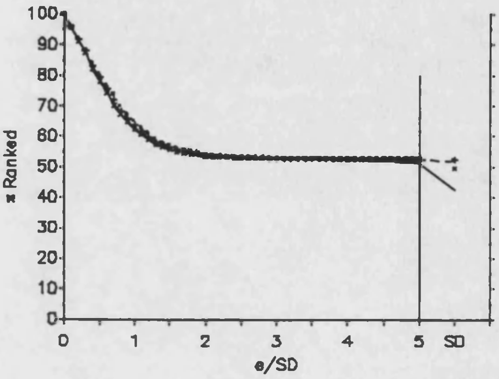
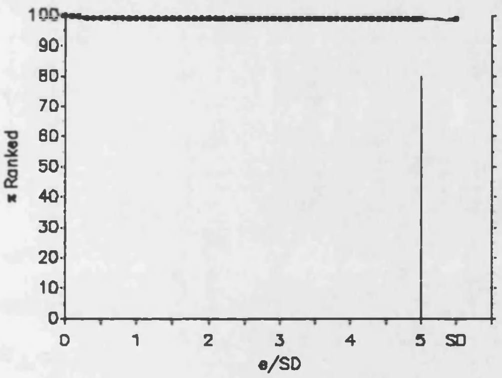


Figure 2.F Strong Dominance

Sample Size=500



Notes for Figures 2.A to 2.G: The letters A to G refer to the ID of each pair. The titles give each pair's characterization in terms of the GL curve. The graphs give the percentage of draws rankable, given different levels of aggregation for increasing values of e_{max}, given e_{min}=0, and, at 'SD', the percentage rankable by the criterion of 2OSD. Only the results for a sample size of 500 are given for the last three pairs, as other two sample sizes give near-identical results.

IV.2 Statistical method

Even with the largest sample size, the probability of ranking samples drawn from the two identical distributions is 12-14% using 2OSD and 43-51% using e-dominance with bounds of zero and five. These unacceptably high sizes even when the dominance criterion is demanding, the data disaggregated and the sample size large answer the first question posed in II.4: there is a need for basing inferences of dominance on a statistical footing.[15] This sub-section looks at both the intersection-union (IU) and union-intersection (UI) methods in relation to both 2OSD (IV.2.1) and e-dominance (IV.2.2). IV.2.3 examines the shape of the test statistics. The extension to the IU method introduced in II.4, concerning dominance over endogenous bounds, is considered in IV.2.4. Since only variances (and not covariances) need to be estimated, and since the observations are assumed to be unweighted, the actual formulae used for the test statistics are simplified versions of those given in Section II of Chapter Two: see (28) and (31) for the deficit curve, (49) and (53) for the GL curve and (13) and (14) for e-dominance. Table 5 below summarizes the conclusions reached in relation to the seven areas of investigation set out in II.4.

IV.2.1 Second-order stochastic dominance

IV.2.1.1 Union-intersection method

The UI method of BFT and CCT was implemented using both the deficit and the GL curves. A range of numbers of ordinates, W, was also used, to see how this choice affects performance. The chosen values of W were 5, 10, 20, 100 and W_{dis}. In the case of the GL curve, W_{dis}=N: the curves are evaluated at every sample point. In relation to the deficit curve, W_{dis} is equal to the number of sample income values, which is bounded from above by 2N. Where $W<W_{dis}$, the ordinates were chosen to give the same number of observations between each ordinate. This was done in the case of the GL curve by setting the ordinate, $x_i=p_i=i/W$. In the case of the deficit curve, the ordinate $x_i=y_i$ was chosen so that i/W of the two samples combined had an income less than or equal to y_i. Note that in both cases this results in $x_W=x_{max}$, but $x_1 \geq x_{min}$. The critical values for the UI method were based on the SMM distribution, as suggested by BFT and CCT (see III.2 of Chapter Two). A list of critical values used is given in Appendix A. All the critical values were chosen following BFT and CCT to bound the size of the inference rule at 5% in the case in which the null of equality ($\zeta_i=\zeta_i^* \forall i$) holds (see (91) of Chapter Two).

15. Note too the much higher size using the 2OSD criterion for pair E: 48%.

Table 5 Precis of Results Pertaining to Statistical Inference

Question	Answer	Implications	For details
1. Is there a need for statistical testing?	Yes. The probability of attaining a sample ranking for the two pairs of identical distributions is 10-20% for 2OSD and 40-50% for e-dominance criterion (bounds of 0 and 5).	Reliance on sample rankings will often lead to the spurious inference of dominance. Dominance should be analyzed as a statistical problem if possible.	Tables 4 and 9.
2. How do the size of the IU and UI methods compare?	The size of the IU method is always below α (given a critical value of Z_α); that of the UI method can range from below α to 1 (given a critical value of C_α).	The advantage of the IU method is that its size can be controlled.	Tables 6a, 6b, 7 and 9; IV.2.1,2
3. How do the power of the IU and UI methods compare?	The IU method's power is nearly always below the UI's. However, the power of both methods can be low.	The advantage of the UI method is its greater power.	Tables 6a, 6b, 7 and 9; IV.2.1,2
4. Can the power of the IU method be increased by lowering the critical value below Z_α while still controlling size at α?	No, pair E illustrates.	The IU method is not in general too conservative.	Table 7; IV.2.1.2
5. Can the power of the IU method be increased by using the method of endogenous bounds while still controlling the size?	Yes, size increases to 2α in return for large increases in power. Power can also be increased by exogenous restrictions on the bounds.	Using the IU method with endogenous bounds gives good power compared to the UI method, while still enabling size to be controlled.	Tables 10a to 12b; IV.2.4
6. Does it matter from a statistical perspective whether one uses deficit curve or generalized Lorenz estimators?	Yes if the UI method is used; no if the IU method is used with a lower bound.	If the IU method is used, and a lower bound imposed, the choice between GL and deficit curves can be made on the basis of analytical convenience.	Tables 6a, 6b, 7 and 9; IV.2.1,4
7. What problems are caused for the different estimators by sensitivity to the lower tail?	On account of behaviour at the lower tail, if the IU method is used with the deficit curve, unrestricted 2OSD will never be inferred; if the UI method is used with the GL curve, 2OSD will often spuriously be inferred. For a pair displaying dominance, there is likely to be a loss of power (lower test statistics) at low levels of y and p and high values of e, even if the two distributions seem to differ most at the lower tail.	Use of low values of p and y and high values of e will tend to result in a loss of statistical power.	Figures 3.A-3.G; IV.2.3

Tables 6a and 6b contain the results for the deficit curve and GL curve respectively. Inference of dominance using the UI method requires the satisfaction of two conditions: one null must be rejected, the other must not (see II.3). The second column for each choice of W (headed 'ND') gives the proportion of simulated pairs in which at least one null was rejected. The first column for each choice of W (marked 'UI') gives the estimated size or power of the UI method for the particular distributions. Since rejection of at least one null is a necessary but not sufficient condition for the inference of dominance, the figure in each second column is never lower than that in each first. Note that the figure given in the second column can be thought of as giving the number of times in which the inference of no dominance is made - see III.2 of the previous chapter.

• If the UI method is implemented using GL curves, a lower bound needs to be imposed on the first ordinate at which the test statistic is assessed. If not there will be spurious inference of dominance when the samples are drawn from identical distributions on account of the tendency, explained in the last chapter (Section IV), of the test statistic to be very large when p is very small. For example, the size is never less than 15% in the first two pairs, though it should never be greater than 5%. This problem can be solved by reducing W: since the ordinates are symmetrically placed, a lower total number implies a higher first ordinate. In fact, the problem of spurious rejection can be avoided by the imposition of even a very low bound. For example, with a sample size of 500, setting W=100 removes the problem, even though there will only be 10 observations (5 from each sample) below the first ordinate. For the remainder of this discussion, references to GL curves will be for $W \leq 100$ only.

• Different conclusions can follow depending on whether the deficit curve or the GL curve is used. The percentages in the 'ND' columns are never higher using the deficit curve than the GL curve, indicating that the GL estimators on average lead to higher maximum test-statistics. The difference in performance falls as W falls, which is consistent with the biggest differences between the two estimators arising at the lower tail.

• The size of the UI method is well below the nominal 5% level in the cases in which the distributions are identical. However, for reasons given in the previous chapter (III.2), this probability can rise sharply in cases in which the distributions cross. Even when the crossing is a clear one, as in pair C, there is a tendency to infer dominance. The size for this pair is at least 13% and at most 76%, depending on the combination of curve, sample size and size of W. The size is even higher in pairs D and E characterized by slight crossings. The minimum size for D is 9% and for E 49%, but the maximum in both cases is 100%. This finding can be compared with the conclusion of CCT that, using the UI method, if the dominance curves cross "the probability of rejecting in favour of dominance can be high, especially in small samples" (p.22). CCT accordingly recommend sample sizes of at least 90 observations for the analysis of 2OSD. However, this study illustrates that the probability of spurious inference of dominance may be high even with a much larger sample size. Indeed, for any finite sample size, it will be possible to find dominance curves which result in a size close to 1-α, since one can always find pairs of distributions which generate sample dominance curves such that one curve significantly dominates at at least one point and is also dominated at some other point, but nowhere significantly. The results also suggest that the problem of large size may become worse as the sample size is increased before it becomes better. A good example is provided by pair C, with, say, W=5 and using the GL curve. With N=500, the probability of inferring dominance is 41%. This increases to 76% when N increases to 2000. But then it falls when N is increased further to 4000, back to 51%. What is happening here is that for low sample sizes neither null is rejected. But then as the sample size increases one moves from rejecting neither null to rejecting only one, thus leading to the inference of dominance. The sample size needs to be increased still further to result in both nulls being rejected and a reduction in size.

• The impact of changing the number of ordinates on the performance of the test is mixed. CCT recommend using six ordinates for the inference of 2OSD to minimize the probability of inferring dominance when curves cross (pp. 20-21). However, there is no evidence to suggest that setting W=5 gives a better performance in this regard than giving it a higher value of 10, 20 or even 100. It does for pair E, but for pairs C and D the highest size is obtained at this lowest level of W.

• The power of the UI method seems high (above 90%) except in the case of pair G if the smallest sample size is used, when it falls to between 28 and 60% depending on the curve used and the size of W.

Table 6a Union-intersection Method: 2OSD using deficit curves

ID	Popn	-----N------		----100-----		-----20-----		-----10-----		-----5------	
					Size						
		UI	ND	UI	ND	UI	ND	UI	ND	UI	ND
A_1	F=F*	0.1	0.1	0.1	0.1	1.3	1.3	2.4	2.4	3.8	3.8
A_2		0.2	0.2	0.2	0.2	1.0	1.0	1.4	1.4	2.8	2.8
A_3		0.4	0.4	0.3	0.3	1.7	1.7	2.2	2.2	3.0	3.0
B_1	F=F*	0.3	0.3	0.3	0.3	1.3	1.3	2.4	2.4	2.9	2.9
B_2		0.1	0.1	0.1	0.1	1.2	1.2	2.4	2.4	2.6	2.6
B_3		0.4	0.4	0.4	0.4	1.7	1.7	2.1	2.1	3.0	3.0
C_1	FXF*	13.1	13.1	13.0	13.0	25.4	25.8	33.4	34.2	40.6	41.3
C_2		76.0	91.4	76.0	91.1	70.3	96.9	68.2	98.7	75.5	98.9
C_3		47.4	100	48.3	100	33.8	100	34.8	100	50.8	100
D_1	FXF*	75.6	75.9	75.5	75.8	86.4	87.1	91.1	91.3	93.4	93.4
D_2		63.0	100	65.9	100	86.4	100	99.7	100	100	100
D_3		8.7	100	11.4	100	61.0	100	98.3	100	100	100
E_1	FXF*	48.9	48.9	48.1	48.1	67.6	67.6	65.6	65.6	51.4	51.4
E_2		100	100	100	100	99.9	100	99.8	100	99.7	99.7
E_3		100	100	100	100	100	100	100	100	99.9	100
					Power						
		UI	ND	UI	ND	UI	ND	UI	ND	UI	ND
F_1	FDF*	94.3	94.3	94.2	94.2	98.4	98.4	98.8	98.8	98.8	98.8
F_2		100	100	100	100	100	100	100	100	100	100
F_3		100	100	100	100	100	100	100	100	100	100
G_1	FDF*	28.2	28.2	28.1	28.1	46.9	46.9	54.3	54.3	59.8	59.8
G_2		97.3	97.3	97.3	97.3	99.1	99.1	99.5	99.6	99.4	99.9
G_3		99.8	99.9	99.8	99.9	99.3	99.9	99.2	99.9	98.8	99.9

Notes: 1. For the first two columns, see the introduction to Section IV. The figures given under the heading 'Size' and sub-heading 'UI' are the percentage of cases in which dominance was inferred. Those given under the heading 'Power' and sub-heading 'UI' are the percentage in which the correct inference of dominance was made. Those under the sub-heading 'ND' give the percentage of cases in which at least one null was rejected (either $\max(Z_i)>C_\alpha$ or $\min(Z_i)<-C_\alpha$). One and only one null must be rejected for dominance to be inferred.The sub-headings 'N', '100', '20', '10' and '5' refer to the number of ordinates used to conduct the tests. See the text for more details.

2. The critical values used are given in Table A.1 of Appendix A. They were chosen, following BFT and CCT, to ensure the asymptotic probability of inference of dominance given identical dominance curves was no more than 5%.

Table 6b Union-Intersection Method: 2OSD using generalized Lorenz curves

ID	Popn	-----N------		----100-----		-----20-----		-----10-----		-----5------	
						Size					
		UI	ND	UI	ND	UI	ND	UI	ND	UI	ND
A_1	F=F*	16.9	16.9	2.9	2.9	2.2	2.2	2.7	2.7	4.0	4.0
A_2		17.8	17.9	0.7	0.7	1.2	1.2	1.5	1.5	2.9	2.9
A_3		21.4	21.4	0.4	0.4	1.7	1.7	2.2	2.2	3.0	3.0
B_1	F=F*	19.6	19.7	3.0	3.0	2.1	2.1	2.4	2.4	3.1	3.1
B_2		18.2	18.2	0.1	0.1	1.4	1.4	2.5	2.5	2.6	2.6
B_3		21.3	21.3	0.4	0.4	1.7	1.7	2.1	2.1	3.0	3.0
C_1	FXF*	35.5	37.7	20.5	20.7	27.0	27.7	34.0	35.3	41.4	42.1
C_2		55.6	94.7	72.1	91.9	68.4	97.0	67.3	98.9	75.2	98.9
C_3		31.0	100	44.3	100	32.5	100	34.4	100	50.3	100
D_1	FXF*	50.7	88.6	67.9	81.3	85.8	87.7	90.9	91.2	93.3	93.3
D_2		8.0	100	40.2	100	85.3	100	99.7	100	100	100
D_3		0.5	100	6.3	100	61.3	100	98.3	100	100	100
E_1	FXF*	93.1	93.2	84.0	84.0	77.4	77.5	69.5	69.5	52.9	52.9
E_2		100	100	100	100	99.9	100	99.8	100	99.7	99.7
E_3		100	100	100	100	100	100	99.9	100	99.9	100
						Power					
		UI	ND	UI	ND	UI	ND	UI	ND	UI	ND
F_1	FDF*	99.5	99.5	98.4	98.4	98.8	98.8	98.9	98.9	98.8	98.8
F_2		100	100	100	100	100	100	100	100	100	100
F_3		100	100	100	100	100	100	100	100	100	100
G_1	FDF*	57.1	57.4	44.8	44.8	52.1	52.1	56.5	56.5	60.9	60.9
G_2		98.2	98.5	97.7	97.7	99.4	99.4	99.5	99.6	99.4	99.9
G_3		99.5	100	99.9	100	99.4	100	99.3	100	98.9	100

Notes: See the notes to Table 6a.

IV.2.1.2 Intersection-union method

Tables 7 and 8 contain the results of applying the IU method using both the deficit and GL curves. Dominance is required over the entire range of the curves (in Table 7), which corresponds to welfare 2OSD, and then in relation to bounds (in Tables 7 and 8) corresponding to y-restricted and p-restricted welfare 2OSD. The critical value was chosen by bounding both the Type I and I^* inference probabilities by α=.05, resulting in Z_α=1.65. The IU method, as defined in II.3, requires a minimum test statistic greater than Z_α for it to be inferred that FDF^*. It was implemented, as recommended in Chapter Two, III.4, with W equal to the number of sample y or p values in the range of interest. Hence $x_1=x_{min}$ and $x_W=x_{max}$. In Table 7, only lower bounds are used. Four different lower bounds were chosen to analyze restricted dominance. In the case of the GL curve, in which $x_{min}=p_{min}$, these were set equal to 1%, 5%, 10% and 20%. Using the deficit curve, so that $x_{min}=y_{min}$, the lower income bounds were chosen so that, respectively, 1% 5%, 10% and 20% of the two samples combined had an income less than or equal the bounds. Strictly speaking this means that for each sample we are analyzing a different type of y-restricted welfare dominance, since the lower income bound, thus defined (equal, let us say, to y(t%)), will change from sample to sample. There are practical advantages, however, in expressing the income bounds in this way for the purpose of this simulation exercise. First, the income bounds thus chosen should divide up the combined sample in a way which corresponds fairly closely to the division using bounds in the space of p, enabling comparison of the deficit and GL curves. Second, as will become evident, this method enables us to compare these results with those obtained using 'endogenous bounds', given in IV.2.3.

Table 8 uses an upper as well as lower bound on y and p. The upper bound is selected in the same way as the lower bound for Table 7. Tables 7 and 8 are arranged in the standard way. Note that the results from pair D are split between the two size and power sections of the tables. Excluding the left tail by setting $x_{min}\geq y(20\%)$ for the deficit curve and $x_{min}\geq 20\%$ for the GL curve results in dominance by F. For the lower bounds used, no ranking is possible.

• If deficit curves are used in conjunction with the IU method, they need to be bounded from below or dominance will never be inferred. This is because, as shown in Chapter Two, Section IV, the value of the first test statistic tends to approach unity as N becomes large. Again the imposition of a very low bound removes this problem. In Pair F, for example, the power increases from 0 for the criterion of unrestricted 2OSD to 70% when $x_{min}=y(1\%)$, putting only 10 observations below the lower bound. From now on, references to deficit curve will be for $x_{min}\geq y(1\%)$.

• With this exception, it makes no difference whether the deficit or GL curve is used. They give very similar results - to within one percentage point. This is not surprising in the cases of no dominance, since here both curves would be expected to display test statistics of value zero. It is more striking that the power values are so similar.

• Although we know from theory that the size of the IU method is asymptotically bounded by 2α (see II.4), in fact it is never above α and often well below.

• However, one cannot conclude that a lower critical value than Z_α can be used while maintaining the size at α. Pair E illustrates. In all the restricted dominance cases, the size is never less than 4.4% and reaches as high as 4.8%. (All these cases of inference are Type I errors.) In many cases, the actual size will be well below α, but one cannot rule out the possibility that it will be arbitrarily close to α.

• The power of the IU method seems to be below that of the UI method. This is true for pair F characterizing strong dominance, though only for the smallest sample size. It is also of course true in the case of pair G where the dominance curves touch but do not cross. In this case, the power of the IU method is bounded from above by α, as shown in Chapter Two, III.3 and as borne out in Table 7.

• Use of restricted dominance can increase the power of the IU method. This is particularly dramatic with pair G and the largest sample size, where reducing the upper bound from 99 to 80% increases the power from 14 to 97%. The size of the method seems little affected however.

Table 7 Intersection-union Method: Unrestricted and Restricted 2OSD, with Lower Bounds Only

ID	Popn	Deficit Curves					Generalized Lorenz Curves				
		Size									
		2OSD	≥1%	≥5%	≥10%	≥20%	2OSD	≥1%	≥5%	≥10%	≥20%
A_1	F=F*	0.0	0.2	0.4	0.5	1.2	0.2	0.2	0.4	0.5	1.2
A_2		0.0	0.1	0.3	0.7	1.1	0.1	0.1	0.3	0.7	1.1
A_3		0.0	0.1	0.4	0.5	0.8	0.0	0.1	0.4	0.5	0.8
B_1	F=F*	0.0	0.1	0.5	1.1	1.7	0.1	0.1	0.5	1.2	1.7
B_2		0.0	0.0	0.2	0.3	0.7	0.0	0.0	0.2	0.3	0.7
B_3		0.0	0.3	0.3	0.7	1.4	0.0	0.3	0.3	0.7	1.4
C_1	FXF*	0.0	0.0	0.1	0.2	0.7	0.0	0.0	0.1	0.2	0.7
C_2		0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.1
C_3		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
D_1	FXF*	0.0	0.0	0.7	4.0	.	0.0	0.0	0.7	4.0	.
D_2		0.0	0.0	0.0	2.8	.	0.0	0.0	0.0	3.1	.
D_3		0.0	0.0	0.0	1.8	.	0.0	0.0	0.0	1.8	.
E_1	FXF*	0.0	4.2	4.7	4.8	4.8	3.8	4.3	4.5	4.5	4.5
E_2		0.0	4.4	4.4	4.4	4.4	4.3	4.4	4.4	4.4	4.4
E_3		0.0	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4
		Power									
		2OSD	≥1%	≥5%	≥10%	≥20%	2OSD	≥1%	≥5%	≥10%	≥20%
D_1	FDF*	.	.	.	.	24.7	.	.	.	.	24.9
D_2		.	.	.	.	62.2	.	.	.	.	63.7
D_3		.	.	.	.	86.3	.	.	.	.	87.1
F_1	FDF*	0.0	68.2	71.0	71.0	71.0	66.9	69.5	71.1	71.1	71.1
F_2		0.0	99.8	99.8	99.8	99.8	99.6	99.8	99.8	99.8	99.8
F_3		0.0	100	100	100	100	99.9	100	100	100	100
G_1	FDF*	0.0	1.7	4.3	4.7	4.9	1.8	2.0	4.3	4.7	4.9
G_2		0.0	4.7	4.9	4.9	4.9	3.3	4.6	4.9	4.9	4.9
G_3		0.0	4.9	4.9	4.9	4.9	3.3	5.0	5.0	5.0	5.0

Notes: 1. For the first two columns, see the introduction to Section IV. The figures given under the heading 'Size' ('Power') are the percentage of cases in which dominance was wrongly (correctly) inferred. The sub-headings '2OSD' '≥1%' etc. refer to the presence and positioning of the lower bound. See the text for more details.

2. The critical value used was 1.65, chosen to ensure that the Type I and I* inference error asymptotic probabilities were no more than 5% each.

Table 8 Intersection-Union Method: Restricted 2OSD, Lower and Upper Bounds

ID	Popn	Deficit Curves				Generalized Lorenz Curves			
		Size							
		1-99%	5-95%	10-90%	20-80%	1-99%	5-95%	10-90%	20-80%
A_1	F=F*	0.2	0.5	1.2	2.7	0.2	0.5	1.2	2.7
A_2		0.2	0.5	1.0	2.3	0.2	0.5	1.0	2.3
A_3		0.1	0.7	1.2	2.4	0.1	0.7	1.2	2.4
B_1	F=F*	0.1	0.6	1.6	2.7	0.1	0.7	1.6	2.6
B_2		0.1	0.5	1.1	2.5	0.1	0.5	1.1	2.5
B_3		0.3	0.3	1.0	2.3	0.3	0.3	1.0	2.3
C_1	FXF*	0.0	0.3	0.4	1.0	0.0	0.3	0.4	0.9
C_2		0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1
C_3		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
D_1	FXF*	0.0	0.7	4.0	.	0.0	0.7	4.0	.
D_2		0.0	0.0	2.8	.	0.0	0.0	3.1	.
D_3		0.0	0.0	1.8	.	0.0	0.0	1.8	.
E_1	FXF*	4.2	4.7	4.8	4.8	4.3	4.5	4.5	4.7
E_2		4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4
E_3		4.4	4.4	4.4	4.5	4.4	4.4	4.4	4.5
		Power							
		1-99%	5-95%	10-90%	20-80%	1-99%	5-95%	10-90%	20-80%
D_1	FDF*	.	.	.	24.7	.	.	.	24.9
D_2		.	.	.	62.2	.	.	.	63.7
D_3		.	.	.	86.3	.	.	.	87.1
F_1	FDF*	74.6	84.1	88.2	94.1	75.9	84.1	88.1	94.1
F_2		99.9	100	100	100	99.9	100	100	100
F_3		100	100	99.9	100	100	100	100	100
G_1	FDF*	3.7	12.3	20.6	37.2	4.1	12.2	20.4	36.7
G_2		9.6	28.7	51.8	82.5	9.4	28.5	51.5	82.3
G_3		14.0	45.3	75.4	96.9	13.9	45.1	75.1	96.9

Notes: 1. For the first two columns, see the introduction to Section IV. The figures given under the heading 'Size' ('Power') are the percentage of cases in which dominance was wrongly (correctly) inferred. The sub-headings '1-99%' etc. refer to the presence and positioning of the lower and upper bounds. See the text for more details.

2. The critical value used was 1.65, chosen to ensure that the Type I and I^* inference error asymptotic probabilities were no more than 5% each.

IV.2.2 E-dominance

E-dominance tests were conducted using 5 different bounds on e: 0-5, .5-5, 1-5, 0-4 and 0-3. These were evaluated at intervals of e of .1. Both methods, IU and UI, were utilized. Critical values were chosen in the same way as for the tests for stochastic dominance. Those for the UI method are given in the Appendix. That for the IU method was again 1.65. Table 9 reports the results of both methods, as well as those which would follow if inferences were based purely on sample outcomes.

As we saw in the aggregation graphs, in the case in which the underlying distributions are identical, if inferences are based simply on sample outcomes there is a much greater tendency to infer e-dominance than there is to infer 2OSD. For a sample of 500, and with e bounded by zero and five, the probability is greater than 50%, whereas for 2OSD it is around 20%. Hence, if e-dominance is used, there is an increased need for statistical testing.

The two test methods display, in general, the same features as described earlier, but note that Table 9 provides an exception to the generalization established in relation to testing for 2OSD that the UI method has greater power than the IU method. Again, we see that this is generally the case. But in pair E, with a sample size of 500, the IU method has greater power, except for bounds of 0 and 5.

Table 9 E-dominance: Sample, Union-intersection and Intersection-union Results

ID	Popn	Sample					Union-Intersection					Intersection-Union				
							Size									
		0-5	.5-5	1-5	0-4	0-3	0-5	.5-5	1-5	0-4	0-3	0-5	.5-5	1-5	0-4	0-3
A_1	F=F*	46.7	51.3	56.9	52.0	61.0	0.0	0.0	0.0	0.0	0.1	0.5	0.7	1.1	1.2	1.9
A_2		43.4	47.1	53.0	49.5	59.0	0.2	0.3	0.1	0.3	0.5	0.4	0.7	1.0	0.6	1.8
A_3		43.1	47.1	51.1	49.5	60.0	0.6	0.5	0.4	0.7	0.8	0.2	0.6	0.6	0.8	1.4
B_1	F=F*	54.7	57.5	61.8	61.0	69.2	0.3	0.3	0.3	0.4	0.4	0.7	1.0	1.4	1.3	2.3
B_2		52.5	56.2	60.2	58.8	68.3	0.1	0.1	0.1	0.1	0.3	1.1	1.6	1.9	1.6	2.8
B_3		51.3	55.1	58.4	59.1	68.4	0.3	0.3	0.4	0.3	0.5	1.1	1.2	1.6	1.5	2.8
C_1	FXF*	12.7	14.5	19.3	15.1	24.0	15.1	8.9	3.9	16.2	17.7	0.2	0.2	0.3	0.3	0.9
C_2		1.0	1.0	1.3	1.1	4.6	84.4	70.0	34.3	84.4	88.8	0.0	0.0	0.0	0.0	0.1
C_3		0.5	0.5	0.5	0.5	2.3	75.4	76.0	72.7	73.7	87.0	0.0	0.0	0.0	0.0	0.1
D_1	FXF*	20.5	20.5	20.5	.	.	72.5	72.2	68.1	.	.	3.1	3.1	3.1	.	.
D_2		5.5	5.5	5.5	.	.	99.9	99.9	99.8	.	.	0.1	0.1	0.1	.	.
D_3		0.8	0.8	0.8	.	.	97.0	96.8	96.6	.	.	0.0	0.0	0.0	.	.
							Power									
		0-5	.5-5	1-5	0-4	0-3	0-5	.5-5	1-5	0-4	0-3	0-5	.5-5	1-5	0-4	0-3
D_1	FDF*	.	.	.	50.6	89.8	.	.	.	74.2	77.1	.	.	.	8.1	36.9
D_2		.	.	.	45.4	99.5	.	.	.	100	100	.	.	.	7.1	80.9
D_3		.	.	.	39.5	100	.	.	.	99.9	100	.	.	.	5.0	95.3
E_1	FDF*	95.0	95.0	95.7	95.3	95.3	21.3	17.2	16.5	23.4	27.3	18.0	18.0	18.7	28.2	40.3
E_2		100	100	100	100	100	97.9	96.9	97.0	98.6	98.9	36.7	36.7	36.7	57.4	82.2
E_3		99.9	99.9	99.9	99.9	99.9	99.9	99.8	99.8	99.9	99.9	40.2	40.2	40.2	61.2	89.0
F_1	FDF*	99.1	99.9	100	99.1	99.1	95.5	96.2	96.6	96.6	96.9	66.6	83.3	90.4	71.3	72.3
F_2		100	100	100	100	100	100	100	100	100	100	98.3	98.4	98.4	99.9	99.9
F_3		100	100	100	100	100	100	100	100	100	100	99.4	99.4	99.4	100	100
G_1	FDF*	50.9	72.3	87.1	52.1	52.3	20.9	21.8	22.9	22.9	25.0	3.3	10.9	23.3	4.6	5.6
G_2		50.0	88.8	99.0	50.0	50.0	95.4	95.7	95.7	95.7	96.0	5.0	31.2	72.3	5.1	5.1
G_3		51.1	95.1	99.6	51.1	51.1	99.8	99.9	99.9	99.8	99.8	4.8	51.0	92.0	4.9	4.9

Notes: 1. For the first two columns, see the introduction to Section IV. The figures given under the heading 'Size' ('Power') are the percentage of cases in which dominance was wrongly (correctly) inferred. The first-level sub-headings refer to the cases in which, respectively, inferences are based on sample outcomes, the union-intersection method, and the intersection-union. The second-level sub-headings refer to the bounds on e over which dominance is sought.

2. Appendix A gives the critical values used for the UI method. Z_α=1.65 was used for the IU method. See the notes to Tables 6 and 7 for the choice of critical values.

IV.2.3 The shape of the test statistics

Many of the results given so far can be surmised from the set of Figures 3.A to 3.G. The three graphs making up each of these figures give the average test statistic for 100 draws in each round for different values of y, in the case of the deficit curve, p, in the case of the GL curve, and e, in the case of the e-dominance curves. Since the sensitivity of the test statistic to the lower tail is of particular interest, the average is given both for the 100 as a whole and for the two sub-groups defined by having a test-statistic which is, respectively, negative or positive at the, respectively, minimum income level, minimum value of p or maximum value of e.[16] To save space, only the graphs for the middle sample size, 2000, are shown.

• It is striking how similar, except at very low values of y and p, the average test statistic curves for the different criteria are in magnitude and slope within each pair (i.e. within each row of three graphs). The exception is pair E, which is characterized by e-dominance but not by 2OSD.[17]

• These graphs show the tendency, already noted, of the GL estimator to take a large value for low p and for the deficit estimator to equal (plus or minus) one for low y. The graphs are also useful for addressing the question raised in II.4 of whether there is a loss of power as more weight is given to the lower tail, that is, as e increases and y and p fall. The shape of the test statistic of course depends on the underlying distribution. Nevertheless, for pairs F and G, where there is dominance, the test statistic first rises and then falls in absolute value as e falls and as y and p increase (I exclude here the initial sharp fall in the value of the GL test statistic). This is surprising since one would think that the distributions in these two pairs differed most at the lower tail, so that the statistical differences between them should be greater the more weight given to the lower tail. (The distributions in pair F differ by a shift factor, while those in pair G have the same mean, but different CV's.) At least for rankable distributions, then, there does seem to be a tendency for the test statistic to begin to fall if y or p become small enough or e large enough. This could make the inference of dominance more difficult, although it does not seem here to have a strong influence on the results reported. These findings are in accordance with those of Preston (1992), quoted earlier, who found that the statistical implications of the choice of evaluation function "depend[s] upon the nature of the ... differences" between the distributions under investigation (p.7). but that there was "slight evidence that ... concavity in welfare functions may be costly in

16. In addition, these two sub-groups are constrained to include at least 5 draws. If not, they are not drawn, as it becomes too difficult to tell apart the lines.

17. In pairs C, E, F and G, the second derivative of the deficit curve average test statistic is the negative of the GL curve test statistic.

terms of lost efficiency" (p.9). (Increased concavity corresponds in this context to increased sensitivity to the lower tail.)

Figures 3.A-3.G Average Test Statistics (for 100 rounds, S.Size=2000)

Figure 3.A Identical, High Inequality

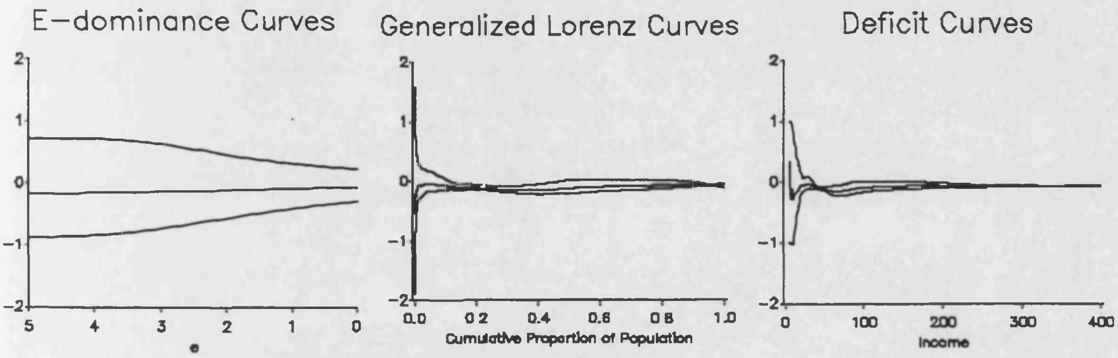


Figure 3.B Identical, Low Inequality

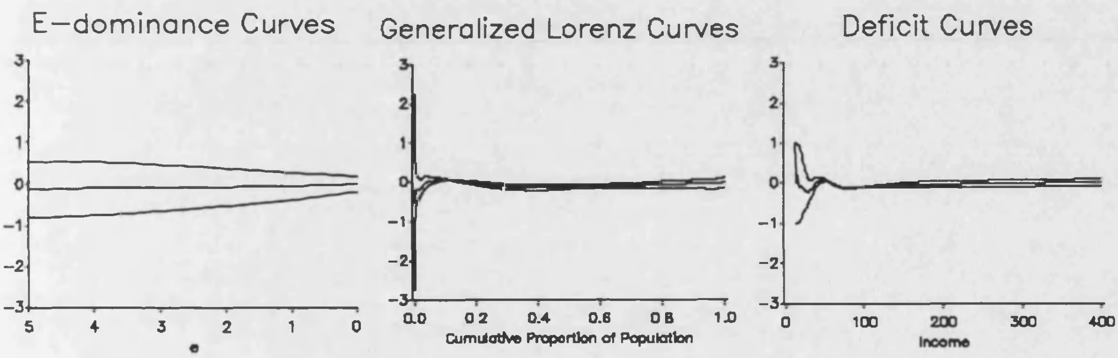
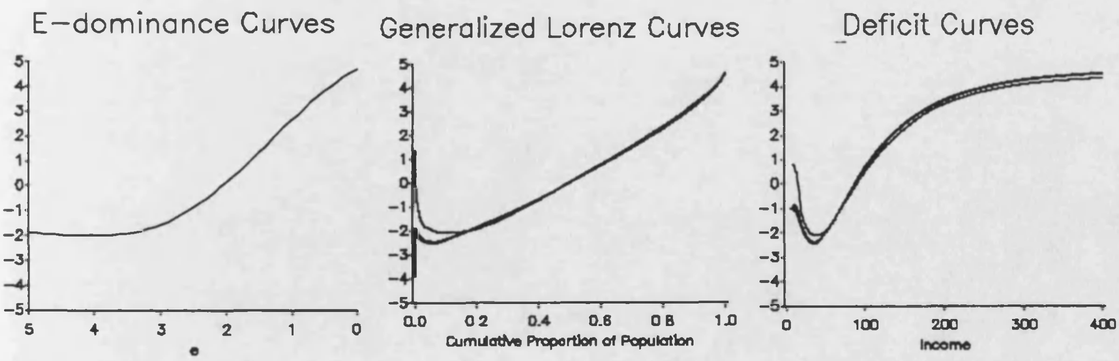


Figure 3.C Clear Single Crossing



Notes for Figures 3.A to 3.G: The test statistics are derived by substituting the various formulae given in Chapter Two, Section II into the formula for Z_i given in Chapter Two, III.1. A positive value for the test statistic indicates dominance by F, negative dominance by F^*. If only one line is shown it is the average test statistic for 100 draws. If there are three lines, the additional two are averages for sub-groups (drawn if they each include at least five draws) based on the sign of the test-statistic at the initial value of x (y, p or e). See the text for further details.

Figure 3.D Slight Single Crossing

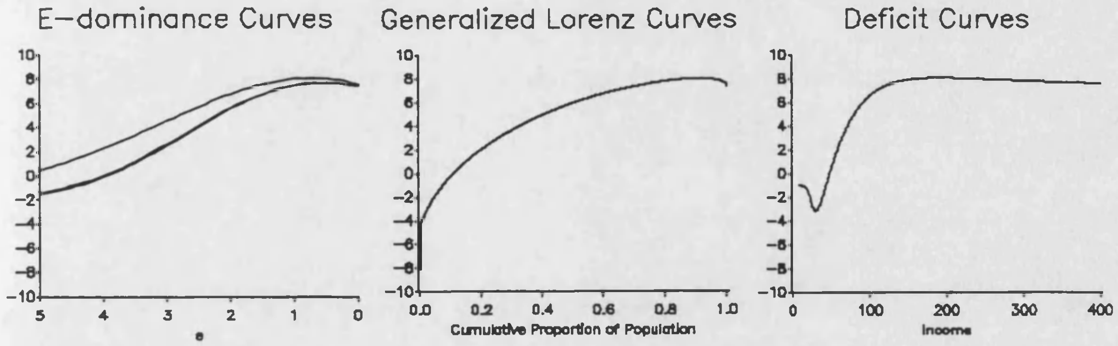


Figure 3.E Slight Double Crossing

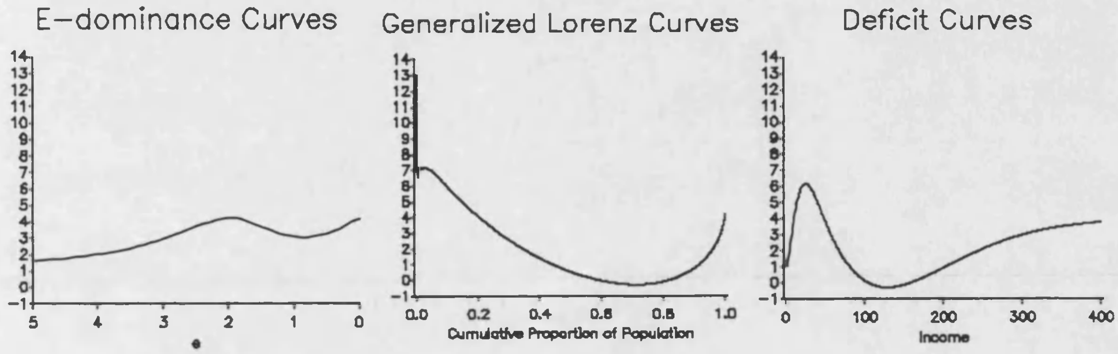


Figure 3.F Strong Dominance

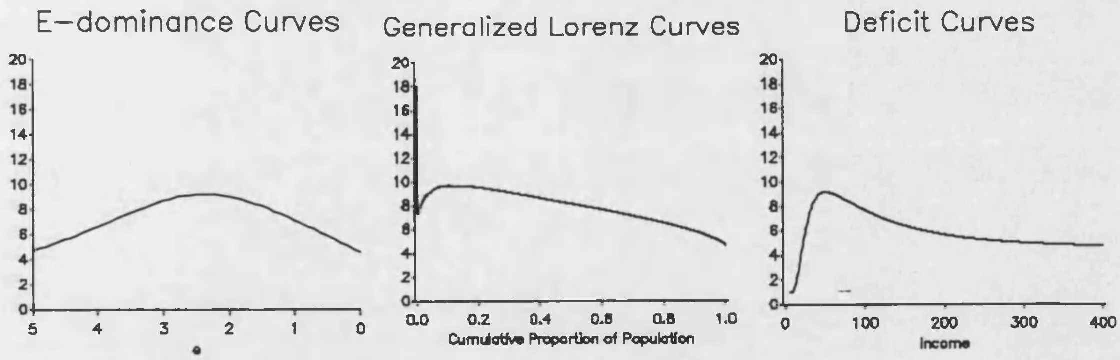
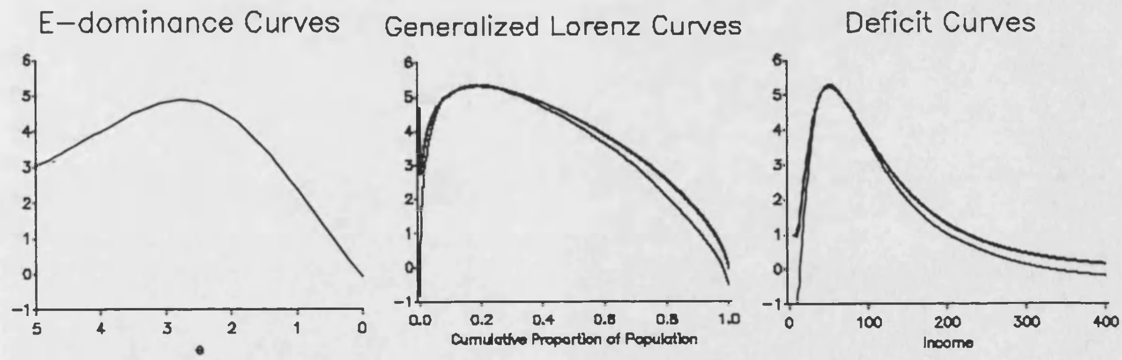


Figure 3.G Dominance but no Strong Dominance



IV.2.4 Dominance with endogenous bounds

The possibility was raised in II.4 of searching for dominance between bounds not pre-specified but chosen after inspection of the data. It was argued then that, to control the probability of spurious inference of dominance, a minimum length requirement, L, could be specified which the bounds had to satisfy - see (5). To make the problem manageable a maximum of one significant range per sample pair was allowed. This, as it turns out, is hardly restrictive, since the minimum length turns out to be of a size which makes it highly unlikely that where could be more than one significant range. Let

$$\begin{aligned} &d_j=d(x_{1j})-d(x_{2j})\ j=1,..,J \text{ where for } x_{1j}\leq x_i\leq x_{2j} \text{ either } Z_i>Z_\alpha\ \forall i \text{ or } Z_i<-Z_\alpha\ \forall i \\ &d^*=\max(d_1,...d_J) \end{aligned} \qquad (6)$$

For e-dominance and GL curves, the function d_j was assumed to be linear in e and p respectively. But for deficit curves, distance was measured, as before, in terms of the coverage of combined observations (arranged in ascending order of income). Then, to prevent dominance being found over more than one range, L was chosen to be the maximum of d^* and L^*, where L^* is some exogenously specified bound. If $d^*\geq L^*$, then dominance was inferred over the range L^*. The problem then reduced to one of choosing L^*. Experimentation suggested lengths of 40 to 50% for 2OSD, and of 2 and 2.5 for e-dominance.

Results using these bounds are shown in three sets of tables, 10a and b, 11a and b, and 12a and b, referring respectively to analysis using the GL, deficit and e-dominance curves. The 'a' tables give the size of the tests for those pairs where the population dominance curves are identical or cross. The 'b' tables give the power results. Making the bounds endogenous increases the number of cases for which power statistics exist. Wherever the population dominance curves cross, for some range FDF* and for some range F*DF. However, I only give the power of the tests in relation to the hypothesis that F*DF for pair C. For the other pairs for which it is true that F*DF for some range (pairs D and E in relation to the deficit and GL curves and pair D in relation to e-dominance), the ranges are too small to be picked up by this method. The power tables show over which range the different distributions dominate. This information is given under the sub-heading 'Population'. The column under this sub-heading headed 'Lgth' gives the distance or length over which population dominance obtains, whether in terms of p (Table 10), coverage of combined populations (Table 11) or e (Table 12). (Note that in Tables 10 and 11, the 'Lgth' columns are expressed as percentages.) The next two columns 'Max' and 'Min' give the maximum and minimum points of dominance. In 10 and 12, the length is simply the maximum minus the minimum bound, but not in 11, since the bounds are given in terms of income and the length in terms of coverage of the combined populations. Note that for pair E, and Tables 10b and 11b, there are two entries in the 'Population' columns due to the double crossing.

The key result-containing columns in each table are those marked, respectively, 'Size' and 'Power'. Additional columns give the average length of the range over which dominance is inferred ('Lgth'), the standard deviation of this length ('StdD') and the average maximum and minium bounds ('Max' and 'Min'). In addition, in the power tables, a further column 'Total' is given. This reports the total number of cases in which the length criterion is met, but includes those in which the bounds are such as to cause a mistaken inference to be made.[18] The size tables give the results where L^* is set equal to d^* - this is the case of 'no length restrictions' - as well as for L^* equal to 40 and 50% in the case of the deficit and GL curves and equal to 2 and 2.5 in that of the e-dominance curves.

There are five main results.

• The need is evident for the exogenous minimum length requirement, L^*. Without such a requirement, we would, in the case where the distributions were equal, have a size of upwards of 50% in the two 2OSD cases, and of around 20% in that of e-dominance.

• Imposing a minimum length requirement improves matters greatly. In all the 2OSD cases, with a length requirement of 40%, the size is below 10% (except in the case of B_3 where it is 10.2%); with a length requirement of 50%, the size is below 8%. In the case of e-dominance, $L^*=2$ gives a maximum size of approximately 11%, while $L^*=2.5$ gives one of 8%.[19]

• The GL and deficit curves give almost identical results.

• There is a marked increase in power. Round F_1, F being the pair displaying clear dominance, illustrates this most clearly in relation to the 2OSD criterion. Using an exogenous minimum bound, the power of the IU method was 70% (see Table 7). With $L^*=50\%$, the power increases to 99%. Yet there is very little trade-off in terms of reduced bounds: the average range over which

18. Since the ranges of dominance of the parent populations are calculated by discrete approximations (see the notes to Table 2) there is the possibility of error in the calculation of the size and power percentages. However, small changes to the bounds were experimented with, and were found not to change the results greatly.

19. It is interesting to compare the size of pairs C, D and E in the 2OSD tables (10 and 11). Whichever constraint and curve are used, pair D has the smallest size of 3-5%. Pair E is slightly higher at 5-7%. This is as expected since pair E is characterized by a double crossing, so there is more room for error. Pair C has the highest size of all, 7-9%. Since the two distributions of pair C cross in the middle, it is possible to infer dominance of either, again increasing the room for error.

dominance obtains is 96.7%. Clearly, there are many cases in which dominance just slightly fails to obtain. This method show that one can do better in such cases than simply recording a verdict of no dominance. Pair G provides another illustration. Recall that this is the pair with identical means, which results in the IU method having very low power (5% or less) unless an upper bound is imposed. Whichever 2OSD curve is used, using an endogenous bound (L=50%) gives a great boost in power, up to 70% with a sample size of 500, and to 100% with a sample size of 2000 or 4000. The average length is quite high though: 78%, 88% and 93% in increasing order of sample size. Power is also increased for pair D, characterized by a crossing in the lower tail. In at least 90% of the simulations, 2OSD is correctly inferred. The method is less successful for pair E, where there is a double crossing. Since the crossing is further away from either tail, it is more difficult to infer dominance subject to a minimum length criterion. However, Table 12b does indicate a large increase in power for pair E in relation the e-dominance criterion.

• Both power and the average length of inferred dominance increase with sample size, indicating increased accuracy.

It must of course be true that the power increases when endogenous bounds are used, at least as long as the range covered by the exogenous bounds is no less than L^*. What the tables show though is that the purchase of using endogenous bounds on power may be substantial, and at little cost in terms of range covered. Endogenous bounds also give greater flexibility, since one does not have to choose bounds prior to observation. The cost of this substantial increase in power and flexibility is two-fold. First, if there are many samples to be compared, it will become difficult to compare results across pairs. Second, there is an increase in size. The results suggest that if use is made of a length criterion of 40% in the case of 2OSD and 2 in the case of e-dominance, the resulting size will be at most approximately 2α. To be on the safe side, which is perhaps warranted given the experimental basis of this approach, one might want to replace 40 by 50% and 2 by 2.5.

Table 10a Size of Intersection-Union Method using Endogenous Bounds: generalized Lorenz curves

ID	Popn	---No length restrictions---					-------Length ≥ 40%--------					-------Length ≥ 50%--------				
		Size	Lgth	StdD	Max	Min	Size	Lgth	StdD	Max	Min	Size	Lgth	StdD	Max	Min
A_1	F=F*	64.6	14.7	23.2	31.6	16.9	9.0	65.2	16.7	84.9	19.7	6.9	71.2	14.4	90.4	19.2
A_2		64.4	13.0	22.5	28.7	15.6	8.0	66.1	16.4	87.7	21.7	6.6	70.6	14.4	87.9	17.4
A_3		66.3	12.3	22.2	26.4	14.1	7.8	65.8	17.5	88.4	22.6	6.3	70.8	15.7	90.9	20.2
B_1	F=F*	61.0	14.9	24.0	31.0	16.1	8.1	69.0	17.0	88.2	19.2	6.7	74.0	14.2	90.7	16.7
B_2		62.6	13.2	22.4	28.1	14.9	7.7	65.5	17.4	87.5	22.0	5.8	72.4	14.3	90.2	17.7
B_3		66.5	13.7	23.2	27.0	13.3	10.1	62.7	15.9	85.8	23.0	7.2	69.5	13.9	89.0	19.5
C_1	FXF*	9.7	56.7	20.4	77.4	20.7	8.7	62.2	12.3	83.9	21.8	7.7	64.1	11.7	88.2	24.1
C_2		8.1	55.8	5.6	78.9	23.1	8.1	55.8	5.6	78.9	23.1	7.3	56.6	5.4	82.1	25.5
C_3		8.6	55.0	4.6	78.5	23.5	8.6	55.0	4.6	78.5	23.5	7.7	55.6	4.4	81.8	26.2
D_1	FXF*	4.9	89.1	15.2	96.6	7.5	4.7	92.2	2.4	100	7.8	4.7	92.2	2.4	100	7.8
D_2		4.8	90.6	1.3	100	9.4	4.8	90.6	1.3	100	9.4	4.8	90.6	1.3	100	9.4
D_3		4.1	90.0	0.8	100	10.0	4.1	90.0	0.8	100	10.0	4.1	90.0	0.8	100	10.0
E_1	FXF*	10.0	68.1	32.2	90.8	22.7	7.1	85.1	20.8	93.9	8.8	6.1	91.6	14.0	94.5	2.9
E_2		6.2	83.6	27.3	95.5	11.9	5.2	93.6	16.0	97.2	3.6	4.8	97.6	7.8	97.7	0.1
E_3		5.8	88.8	22.3	95.0	6.2	5.2	95.2	12.2	96.2	1.0	5.1	96.3	9.5	96.4	0.1

Table 10b Power of Intersection-Union Method using Endogenous Bounds: generalized Lorenz curves

ID	--------Popn----------				----------Length ≥ 40%-----------						----------Length ≥ 50%-----------					
		Lgth	Max	Min	Power	Total	Lgth	StdD	Max	Min	Power	Total	Lgth	StdD	Max	Min
C_1	FDF*	51.0	100	49.0	5.4	10.4	45.0	3.0	100	55.0	0.3	5.3	50.3	0.1	100	49.7
C_2		.	.	.	12.8	17.7	44.1	3.1	100	55.9	1.2	6.0	50.4	0.3	100	49.6
C_3		.	.	.	26.8	30.7	44.7	3.0	100	55.3	1.5	5.3	50.6	0.2	100	49.4
C_1	F*DF	49.0	49.0	0.0	2.7	6.4	43.8	2.5	44.7	0.9	0.0	2.7	.	.	.	.
C_2		.	.	.	8.7	12.2	42.7	2.3	43.5	0.8	0.0	2.7	.	.	.	.
C_3		.	.	.	14.8	19.0	43.8	2.6	44.1	0.2	0.0	3.2	.	.	.	.
D_1	FDF*	89.0	100	11.0	90.9	95.6	70.5	11.1	100	29.5	85.7	90.4	72.0	9.5	100	28.0
D_2		.	.	.	95.3	100	81.2	5.0	100	18.8	95.3	100	81.2	5.0	100	18.8
D_3		.	.	.	95.7	100	83.7	3.0	100	16.3	95.7	100	83.7	3.0	100	16.3
E_1	FDF*	68.0	68.0	0.0	15.4	21.2	48.6	7.2	49.0	0.5	5.3	11.1	57.1	4.7	57.6	0.5
E_2		24.0	100	76.0	45.5	50.2	47.6	6.1	47.7	0.1	12.3	17.0	55.8	4.7	55.9	0.1
E_3		.	.	.	72.3	77.3	48.5	6.2	48.6	0.1	24.5	29.5	55.7	4.3	55.8	0.1
F_1	FDF*	100	100	0.0	99.5	99.5	96.5	8.1	96.9	0.5	99.0	99.0	96.7	7.1	97.2	0.5
F_2		.	.	.	100	100	99.9	0.1	100	0.1	100	100	99.9	0.1	100	0.1
F_3		.	.	.	100	100	99.9	0.0	100	0.1	100	100	99.9	0.0	100	0.1
G_1	FDF*	100	100	0.0	74.8	74.8	74.8	16.7	77.8	3.0	67.2	67.2	78.2	14.2	80.6	2.4
G_2		.	.	.	100	100	87.8	9.7	88.1	0.3	99.8	99.8	87.8	9.6	88.2	0.3
G_3		.	.	.	100	100	93.0	5.6	93.1	0.1	100	100	93.0	5.6	93.1	0.1

Notes: See the notes following Table 12b.

Table 11a Size of Intersection-Union Method using Endogenous Bounds: deficit curves

ID	Popn	---No length restrictions---					-------Length ≥ 40%--------					-------Length ≥ 50%--------				
		Size	Lgth	StdD	Max	Min	Size	Lgth	StdD	Max	Min	Size	Lgth	StdD	Max	Min
A_1	F=F*	40.9	23.0	25.9	210	71.5	9.0	65.4	16.5	348	46.7	6.8	71.6	14.1	378	44.9
A_2		43.1	19.5	25.3	202	70.7	7.9	66.8	16.0	434	48.4	6.7	70.6	14.3	422	43.0
A_3		45.4	17.9	24.9	202	56.3	7.8	65.8	17.5	434	49.1	6.3	70.8	15.7	455	46.3
B_1	F=F*	38.3	23.5	27.1	157	67.1	8.3	68.9	16.5	265	56.0	6.9	73.6	13.9	281	53.9
B_2		42.2	19.5	25.0	150	61.9	7.6	65.9	17.3	285	59.4	5.7	72.9	14.0	299	54.8
B_3		46.1	19.7	25.6	149	56.7	10.2	62.5	16.0	307	60.8	7.1	69.8	13.8	333	56.3
C_1	FXF*	9.3	59.3	16.6	395	51.0	8.7	62.1	12.5	417	51.0	7.5	64.6	11.7	470	54.8
C_2		8.7	55.2	6.0	442	47.2	8.7	55.2	6.0	442	47.2	7.1	56.8	5.4	522	53.1
C_3		9.2	54.5	4.8	470	47.2	9.2	54.5	4.8	470	47.2	7.7	55.7	4.4	544	53.2
D_1	FXF*	4.4	88.8	16.2	548	43.2	4.2	92.3	2.4	572	44.0	4.2	92.3	2.4	572	44.0
D_2		4.7	90.5	1.4	723	46.2	4.7	90.5	1.4	723	46.2	4.7	90.5	1.4	723	46.2
D_3		3.3	90.0	0.8	829	46.8	3.3	90.0	0.8	829	46.8	3.3	90.0	0.8	829	46.8
E_1	FXF*	10.0	67.8	33.0	409	47.2	6.9	86.3	20.8	518	22.9	6.0	92.5	14.2	566	14.1
E_2		6.3	83.4	27.0	680	23.9	5.3	93.1	16.1	775	10.2	4.9	97.0	8.8	822	4.2
E_3		5.8	88.8	22.3	759	14.0	5.2	95.2	12.2	827	5.0	5.1	96.3	9.5	840	3.4

Table 11b Power of Intersection-Union Method using Endogenous Bounds: deficit curves

ID	--------Popn---------				-----------Length ≥ 40%-----------						-----------Length ≥ 50%-----------					
		Lgth	Max	Min	Power	Total	Lgth	StdD	Max	Min	Power	Total	Lgth	StdD	Max	Min
C_1	FDF*	51.0	∞	86.0	5.6	10.5	45.3	3.3	639	95.4	0.4	5.3	51.5	1.2	657	88.9
C_2		.	.	.	13.4	17.5	44.6	3.2	824	95.4	0.9	5.0	50.8	0.8	770	87.0
C_3		.	.	.	25.3	29.8	44.6	3.0	914	95.4	1.6	6.1	50.6	0.4	873	87.1
C_1	F*DF	49.0	86.0	0.0	2.7	6.5	43.5	2.5	80.6	21.7	0.0	2.6	.	.	.	.
C_2		.	.	.	8.1	12.7	43.0	2.3	79.7	18.6	0.0	3.0	.	.	.	.
C_3		.	.	.	14.1	18.8	43.7	2.5	80.4	16.0	0.0	3.2	.	.	.	.
D_1	FDF*	89.0	∞	48.0	91.5	95.7	70.6	11.1	582	68.2	86.5	90.7	72.1	9.6	584	66.5
D_2		.	.	.	95.3	100	80.8	4.9	733	57.3	95.3	100	80.8	4.9	733	57.3
D_3		.	.	.	96.7	100	83.5	3.2	811	54.5	96.7	100	83.5	3.2	811	54.5
E_1	FDF*	67.0	121	0.0	15.4	21.0	48.9	7.5	88.5	8.9	5.3	10.9	57.8	5.2	102	8.3
E_2		24.0	∞	139	48.3	53.2	48.0	6.1	86.9	4.3	14.9	19.8	55.7	4.0	98.5	4.4
E_3		.	.	.	72.8	77.8	48.7	6.1	87.9	3.1	25.8	30.8	55.6	4.4	98.6	3.1
F_1	FDF*	100	∞	0.0	99.4	99.4	96.4	7.9	535	19.1	98.9	98.9	96.7	7.0	537	19.1
F_2		.	.	.	100	100	99.9	0.1	826	14.0	100	100	99.9	0.1	826	14.0
F_3		.	.	.	100	100	99.9	0.0	918	12.2	100	100	99.9	0.0	918	12.2
G_1	FDF*	100	∞	0.0	75.0	75.0	74.8	16.7	178	24.6	67.2	67.2	78.2	14.1	189	23.8
G_2		.	.	.	100	100	87.9	9.7	221	16.0	99.9	99.9	88.0	9.6	221	16.0
G_3		.	.	.	99.9	99.9	92.9	5.9	257	13.4	99.9	99.9	92.9	5.9	257	13.4

Notes: See the notes following Table 12b.

Table 12a Size of Intersection-Union Method using Endogenous Bounds: e-dominance curves

ID	Popn	---No length restrictions---					------------e ≥ 2------------					------------e ≥ 2.5------------				
		Size	Lgth	StdD	Max	Min	Size	Lgth	StdD	Max	Min	Size	Lgth	StdD	Max	Min
A_1	F=F*	22.8	2.0	1.2	3.2	1.2	10.5	3.1	0.9	4.1	1.1	6.9	3.5	0.7	4.5	1.0
A_2		22.1	1.8	1.1	3.3	1.4	8.6	3.0	0.8	4.0	1.0	5.6	3.4	0.7	4.2	0.8
A_3		22.8	1.9	1.1	3.4	1.5	9.1	3.0	0.8	4.2	1.2	5.9	3.5	0.7	4.4	0.9
B_1	F=F*	19.2	2.2	1.3	3.5	1.3	10.6	3.2	0.9	4.1	0.9	8.0	3.6	0.8	4.4	0.8
B_2		21.0	2.0	1.4	3.6	1.6	9.2	3.4	0.9	4.3	0.9	7.6	3.6	0.8	4.4	0.8
B_3		21.6	2.0	1.3	3.6	1.6	10.1	3.2	0.9	4.0	0.8	7.0	3.6	0.8	4.3	0.7
C_1	FXF*	8.4	2.8	0.7	3.7	0.9	7.7	2.9	0.6	3.7	0.8	5.1	3.2	0.5	4.3	1.1
C_2		5.3	2.8	0.4	4.3	1.5	5.3	2.8	0.4	4.3	1.5	4.3	3.0	0.2	4.8	1.8
C_3		5.0	2.7	0.4	3.8	1.1	5.0	2.7	0.4	3.8	1.1	3.2	3.0	0.2	4.6	1.7
D_1	FXF*	7.3	4.4	1.1	4.8	0.4	6.5	4.7	0.3	4.7	0.0	6.5	4.7	0.3	4.7	0.0
D_2		4.2	4.4	0.2	4.4	0.0	4.2	4.4	0.2	4.4	0.0	4.2	4.4	0.2	4.4	0.0
D_3		1.4	4.3	0.2	4.3	0.0	1.4	4.3	0.2	4.3	0.0	1.4	4.3	0.2	4.3	0.0

Table 12b Size of Intersection-Union Method using Endogenous Bounds: e-dominance curves

ID	-----Popn-----				------------e ≥ 2------------						------------e ≥ 2.5------------					
		Lgth	Max	Min	Power	Total	Lgth	StdD	Max	Min	Power	Total	Lgth	StdD	Max	Min
C_1	FDF*	2.1	2.1	0.0	2.0	6.0	2.1	0.1	2.1	0.0	0.0	1.8	.	.	.	.
C_2		.	.	.	3.8	5.2	2.0	0.0	2.0	0.0	0.0	0.4	.	.	.	.
C_3		.	.	.	6.7	8.9	2.0	0.0	2.0	0.0	0.0	0.5	.	.	.	.
C_1	F*DF	2.9	5.0	2.1	11.0	14.7	2.4	0.2	5.0	2.6	5.3	8.6	2.6	0.1	5.0	2.4
C_2		.	.	.	35.9	39.8	2.4	0.3	5.0	2.6	14.9	18.8	2.7	0.1	5.0	2.3
C_3		.	.	.	58.8	61.6	2.4	0.2	5.0	2.6	25.2	27.9	2.6	0.1	5.0	2.4
D_1	FDF*	4.0	4.0	0.0	79.3	85.8	2.8	0.5	2.8	0.0	59.3	65.8	3.0	0.4	3.0	0.0
D_2		.	.	.	95.8	100	3.3	0.4	3.3	0.0	95.1	99.3	3.3	0.4	3.3	0.0
D_3		.	.	.	98.6	100	3.4	0.3	3.4	0.0	98.6	100	3.4	0.3	3.4	0.0
E_1	FDF*	5.0	5.0	0.0	77.4	77.4	3.7	1.0	4.3	0.6	67.5	67.5	3.9	0.9	4.4	0.5
E_2		.	.	.	98.9	98.9	4.1	0.9	4.2	0.0	95.2	95.2	4.2	0.8	4.2	0.0
E_3		.	.	.	99.7	99.7	4.2	0.9	4.2	0.0	96.7	96.7	4.3	0.8	4.3	0.0
F_1	FDF*	5.0	5.0	0.0	100	100	4.8	0.4	5.0	0.1	99.9	99.9	4.8	0.4	5.0	0.1
F_2		.	.	.	100	100	5.0	0.0	5.0	0.0	100	100	5.0	0.0	5.0	0.0
F_3		.	.	.	100	100	5.0	0.0	5.0	0.0	100	100	5.0	0.0	5.0	0.0
G_1	FDF*	5.0	5.0	0.0	77.2	77.2	3.6	0.8	4.8	1.2	70.4	70.4	3.7	0.7	4.8	1.1
G_2		.	.	.	100	100	4.2	0.5	4.9	0.7	99.6	99.6	4.2	0.4	5.0	0.7
G_3		.	.	.	99.9	99.9	4.4	0.3	5.0	0.5	99.9	99.9	4.4	0.3	5.0	0.5

Notes: The last two first-level sub-headings give the minimum length requirements; 'Popn' gives the population characteristics. The second-level sub-headings 'Size' and 'Power' are as for the earlier tables; 'Lgth' gives the average length over which dominance is inferred; 'StdD' the standard deviation of this length; 'Max' ('Min') the average maximum (minimum) bound; 'Total' the percentage of cases which met the length criterion, even if the endogenous bounds did not fall within the dominance range. The text gives more details.

V Conclusion

In relation to aggregation, the main finding is that aggregation does increase the probability of attaining a ranking. A reduction in the probability of ranking is possible, but any such reductions appear in practice to be small, whereas the increases can be substantial. In particular, aggregation will exaggerate the differences between similar distributions and will tend to ignore crossings at the lower tail. The probability of obtaining a ranking using the 2OSD criterion is particularly vulnerable to the degree of aggregation. For the identical distributions studied, the probability of obtaining a ranking increased from around 20% in the case of disaggregated data to above 50% if quintiles are used. Even the use of a large number of aggregated groups, such as percentiles, which one might think quite harmless, led to a substantial increase in the probability of ranking - approximately 100% for the two larger sample sizes of the first two pairs. Where the crossing is in the left-tail, the probability of obtaining a ranking can increase from approximately zero to approximately one. Consequently, although the effects of aggregation are very much dependent on the shape of the underlying distributions, claims such as Shorrocks' that the 2OSD criterion is "sufficiently strong to produce a conclusive ranking in many practical situations" (1983, p.15) are unlikely to be robust to the degree of aggregation of one's data.

No attempt was made in the course of analysis to examine the effect of using aggregated data in conjunction with some method of (non-linear) interpolation designed to approximate the actual degree of inequality in the underlying distribution (see Cowell and Mehta, 1982, for a survey). Instead I simply followed common practice in relation to the use of the 2OSD criterion, ranking the distributions as if they consisted only of the quantile group means. It remains to be investigated whether interpolation would reduce the probability of ranking. Non-linear interpolations should result in crossings below the first aggregation point being less overlooked. However, we have also seen that aggregation reduces the 'noise' associated with disaggregated distributions. How any method of interpolation, however sophisticated, could re-create this noise is unclear.

Clearly if one only one has an aggregated data set one has no choice but to use it, but what if one has disaggregated data? One could argue in favour of aggregation that the choice of criterion depends on the trade-off between the probability of ranking (ρ) and the criterion's range, the usefulness of analysis being increasing in both. Then, since aggregation obtains a higher ρ for given range, one could argue that it should be used. However, this is to ignore the question of interpretation: what does a ranking obtained using aggregated data mean? I have interpreted it to be a basis for inferring dominance between (disaggregated) population distributions. But, as we have seen, it is very difficult to assess the accuracy of aggregation as an inferential method and it seems unreliable.

There are other defences of aggregation. The first is computational ease. The second is measurement-error control. The third is that orderings over aggregated samples are of interest in as means of inferring orderings over aggregated as well as disaggregated populations. On the first point, nothing general can be said, though my own experience would suggest little gain except with very large sample sizes. If one wants to protect against measurement error, one should check that the differences between the sample dominance curves are large. This suggests use of statistical methods rather than aggregation.[20]

The third point is the most interesting. For example, it might be argued that focusing on aggregated distributions enables us to focus on 'essentials' by ignoring slight crossings such as those which characterize pair E. However, from pairs A and B we can see that sample outcomes based on aggregated data can be very poor predictors of the actual ordering between aggregated as well as disaggregated populations. In addition, to return to Sen's point, quoted in II.2, the welfare implications of dominance in relation to two non-existent aggregated distributions are not at all clear. Put another way, one can agree that aggregation is an effective way of summarizing distributional data. But it is difficult to make explicit the loss of information aggregation entails, which militates against its use, where possible, as a basis for analysis.

Even though the probability of ranking is dependent on the degree of aggregation of one's data, this study does not find that the former always goes to zero if the data are completely disaggregated, even if sample sizes are large and the criterion demanding. As mentioned above, for example, this probability hovers around the 10-20% mark for the identical pairs A and B using the 2OSD criterion, and around 40-50% using e-dominance despite the wide bounds of zero and five. These figures emphasize the need to adopt a statistical approach to prevent the spurious inference of dominance.

Of the two competing statistical approaches, it was seen that, as predicted in the previous chapter, the IU, but not the UI, method has bounded Type I and I* inference error probabilities. While the size of the UI method approached 100% on several occasions, that of the UI method was never above α, here 5%. This latter finding is of particular importance. From theory, we can

20. Though this is not to say that use of statistical methods can fully neutralize problems of measurement error.

only bound the size to be no greater than 2α (see II.4). In practice, however, when there is no dominance, it seems either that either both the Type I and I^* inference error probabilities are well below α, resulting in a size below α, or, as in pair E, if one of these probabilities is close to α, the other is close to zero.[21] Hence, in practice, use of a critical value of Z_α will give an upper bound on the size of the IU method as well as on the Type I and I^* inference error probabilities, all of α. In many cases, the actual size will be even lower, but, as we saw with pair E, this cannot be guaranteed. On account of pair E, we can conclude that the IU method is **not** too conservative: a lower critical value than Z_α cannot be used if the size of the test is be bounded by α.

A second finding was that the UI method has in general higher power than the IU method. The relative performance of the IU and UI methods can thus be understood as providing two different trade-offs between size and power. If low size is sought, the IU method should be chosen. But if instead the aim is to maximize power, the UI method should be chosen. The reason for this difference in performance is simple: if two dominance curves are close at one point, and far away at another, the UI method will tend to infer dominance, the IU method not. Unfortunately, this trade-off between size and power cannot be cast in precise terms. This reason provides the argument for using the IU method, since by this method at least the size can be controlled.

If one does stay with the IU method, it is important to investigate how its power might be increased. This can be done by making use of less demanding dominance criteria. One way to do this is via the prior specification of narrower bounds. This is straight-forward: as one would expect from theory, narrower bounds increase power while still controlling size. Reliance on endogenously-determined bounds is less clear-cut as the problem does not seem amenable to the application of standard statistical theory. But the simulation studies do provide some evidence that it is a viable approach. The distributions studied suggest that if one specifies minimum length criteria of 50% in the case of GL and deficit curves, and 2.5 in the case of e-dominance, one can, in conjunction with a critical value of Z_α, control the size to be no greater than 2α. Of course, experimentation with a wider set of distributions (and values of α other than .05) will be necessary to confirm this finding, but the results so far are promising.

Turning from the performance of methods to that of criteria, the study provides illustration of some of the difficulties involved in inferring unrestricted 2OSD due to its reliance

21. If there is dominance (e.g., FDF^*), then the size (here the probability of inferring F^*DF) is approximately zero, which is why only power statistics were presented in such cases - see the introduction to Section IV and footnote 14.

on the requirement of minimum dominance. If the deficit curve is used in conjunction with the IU method 2OSD will never be inferred, while if the GL curve is used with the UI method, 2OSD will often by spuriously inferred. It was also shown that a very low bound, covering as few as ten observations, was sufficient to remove this problem. More generally, though the problem of inferring minimum dominance was not analyzed, the sensitivity of inferences to the left-hand tail was examined. The generalization to emerge from pairs F and G is that for a pair displaying dominance, even if the two distributions seem to differ most at the lower tail, there is likely to be a loss of power at low levels of y and p and high values of e.

Finally, from criteria to curves, and to the choice of deficit *vis-a-vis* GL curves for the inference of 2OSD, restricted or unrestricted. The two curves do give different results if the UI method is used as the GL curve tends to give larger maximum test statistics. If the IU method is used, and a lower bound imposed as recommended in the previous paragraph, the two curves give almost identical results. This is a comforting finding, as it means that the choice between the GL and deficit curves can be made on the basis of analytical convenience.

Although this chapter has been concerned primarily with welfare analysis, many of its findings are applicable also to equality and poverty analysis, though in relation to the former there might be the need for some adjustment. For example, aggregation is likely to ignore crossings in the Lorenz curves at the upper as well as lower tail. Also one cannot be sure that the same minimum length criteria found to be suitable for welfare analysis will also be suitable for inequality analysis. These remain topics for future research.

Appendix A Critical values for the UI method

The critical values for the UI tests were obtained using the Studentized Maximum Modulus (SMM) distribution. The degrees of freedom were set equal to infinity, and the critical values obtained for the different numbers of ordinates from the tables in Stoline and Ury (1979).[22] When the UI test is conducted over all ordinates ($W=W_{dis}$), to bound the critical value it was set equal to the maximum value given in the Stoline and Ury tables (for 190 ordinates).

Table A.1 Critical Values for the UI tests

Criterion	No of ordinates (W)	Critical value used
1. 2OSD	W_{dis} (varying,≥500)	3.64
	100	3.47
	20	3.02
	10	2.80
	5	2.57
2. E-dominance, with bounds		
i. 0-5	51	3.29
ii. 0.5-5	46	3.26
iii. 1-5	41	3.23
iv. 0-4	41	3.23
v. 0-3	31	3.15

22. For values of k (the number of ordinates) not given by Stoline and Ury, an average of a linear interpolation in k and in the reciprocal of k was used to obtain the relevant critical values, as recommended by the authors.

Appendix B Crossing of sample curves

In Chapter One, IV.3, it was shown that if two distribution functions cross at most once then the e-dominance curves of the two distributions will cross at most once. It was also claimed that "Although the sample distribution functions may cross many times, they may nevertheless come from populations whose distribution functions cross only once ... If so, the sample e-dominance curves will often cross only once." In this Appendix, evidence is provided in support of this claim. Table B.1 is based on a sub-sample, of size 100 for each round, of the total number of simulations conducted. It gives the number of crossings of the sample GL and deficit curves (which are equal, and given under the heading '2OSD'), the sample e-dominance curves, and the sample distribution functions (under the heading '1OSD'). The number of crossings of the population curves is also given. The very last column of the table shows that in every pair and for every sample size, at least half of the sample distribution functions cross three or more times. Excluding pairs D and F (which include shifted distributions), this figure rises to 80%. By contrast, the e-dominance sample curves very rarely cross more than once. (The 2OSD sample curves are an intermediate case.)

Table B.1 Crossings of Population and Sample Curves

ID	2OSD					E-dominance					1OSD				
	Pop'n	Sample				Pop'n	Sample				Pop'n	Sample			
		0	1	2	≥3		0	1	2	≥3		0	1	2	≥3y
A_1	0	23	35	23	19	0	51	39	9	1	0	0	0	2	98
A_2	.	15	29	25	31	.	48	39	12	1	.	0	0	0	100
A_3	.	14	35	27	24	.	42	48	9	1	.	0	0	0	100
B_1	0	23	44	15	18	0	57	39	4	0	0	1	0	2	97
B_2	.	14	38	21	27	.	51	43	6	0	.	0	0	2	98
B_3	.	6	32	30	32	.	42	49	9	0	.	0	0	0	100
C_1	1	3	75	15	7	1	10	89	1	0	1	2	9	6	83
C_2	.	0	79	17	4	.	0	99	1	0	.	0	6	4	90
C_3	.	0	79	18	3	.	1	99	0	0	.	0	13	6	81
D_1	1	1	93	3	3	1	15	85	0	0	1	0	19	15	66
D_2	.	0	100	0	0	.	6	94	0	0	.	0	20	10	70
D_3	.	0	100	0	0	.	1	99	0	0	.	0	21	4	75
E_1	2	42	1	57	0	0	90	0	10	0	2	6	0	14	80
E_2	.	35	0	64	1	.	100	0	0	0	.	0	0	10	90
E_3	.	35	0	65	0	.	99	1	0	0	.	0	0	10	90
F_1	0	100	0	0	0	0	100	0	0	0	0	16	13	11	60
F_2	.	100	0	0	0	.	100	0	0	0	.	14	15	14	57
F_3	.	100	0	0	0	.	100	0	0	0	.	15	7	15	63
G_1	0	48	41	8	3	0	55	43	2	0	1	0	7	4	89
G_2	.	40	52	7	1	.	44	56	0	0	.	0	16	4	80
G_3	.	44	51	5	0	.	47	53	0	0	.	0	8	3	89

Conclusion to Part One

The three preceding chapters have examined a number of issues within the rubric of distributional dominance, including: the gains from approaching the measurement of welfare, poverty and inequality within a single framework (Chapter One); the necessity and requirements of a statistical approach (Chapters Two and Three); and the need to be sensitive to the influence of aggregation (Chapter Three). A central focus of all three chapters has been the widely-used stochastic dominance criteria. This focus is retained for the conclusion.

The well-known generality of the second-order stochastic dominance criterion in relation to welfare and inequality analysis has been noted. For poverty analysis a new criterion, that of mixed dominance, has been introduced with the justification that it gives a similar coverage of poverty functions as is given by the second-order stochastic dominance criteria for welfare and equality analysis. But difficulties with the use of the stochastic dominance criteria, to do with their reliance on various 'extreme' forms of dominance, such as minimum or maximum or mean dominance, have also been raised. To recap on these, consider, for concreteness, second-order stochastic dominance, the most popular of all the dominance criteria.

All versions of second-order stochastic dominance (for welfare, equality or poverty analysis) require minimum dominance. Whether defined in terms of income or income divided by the mean, the dominating distribution must have a minimum which is no lower than that of the dominated distribution. Second-order stochastic dominance applied to welfare analysis also requires mean dominance, analogously defined, while that applied to equality analysis requires maximum dominance (in terms of income divided by the mean). Two difficulties arise. The first is normative: do we really want to restrict ourselves to saying that one distribution can dominate another only if it has the requisite form of extreme dominance? The second is inferential: can we really be confident which of two distributions has a higher minimum or lower maximum? It has been argued, in relation to the first question in the first chapter, and in relation to the second in the second, that both of these questions should be answered in the negative.

It is important to stress that the two questions are of different orders of importance. Say that there were no problems of inference, so that we could answer 'Yes.' to the second question. Then one could always start out looking for second-order stochastic dominance and would need to move on to looking for dominance over smaller sets of living-standard functions only if it was absent.[1] But if the answer to the second question is 'No.', then this option is unavailable. If dominance by some criterion cannot be inferred, the utility of that criterion must be called into question.

There are at least four possible responses to the inferential difficulties facing the stochastic dominance approach. The first is to leave the stochastic dominance framework altogether. The second is to use aggregated data. The third is to restrict the class of families assumed to contain the parent distributions, and the fourth is to use the criteria of restricted dominance derived in the first chapter. The first response seems like an over-reaction. There is a place for reliance on orderings based on explicit functional forms, and the e-dominance criterion has been put forward precisely to this end, but it is also true that the stochastic dominance framework is, from a conceptual perspective, an extremely useful, unifying one. The second option of reliance on aggregated data is of course reasonable if only aggregated data are available. But, for the reasons given in the conclusion to Chapter Three, it is not recommended that aggregated data be used in preference to disaggregated. The option of assuming some parameterization for the distributions (and then estimating the parameters required) may be attractive, but it is a task which is least likely to be convincing at the tails, precisely where it is most needed.[2] This leaves the fourth option of reliance on the criteria of restricted dominance, derived in Chapter One as a generalization of the work of Atkinson and Bourguignon (1989), which can be advocated on grounds both of transparency and of simplicity.

If one does choose to use either the e-dominance or the restricted dominance criteria, then bounds need to be chosen. Purely for inferential reasons, one requires: lower bounds on income (y) or the poorest fraction (p) if the criterion is restricted welfare dominance; an upper bound on the inequality aversion parameter (e) if the criterion is e-dominance; and lower and upper bounds on p if the criterion is restricted equality dominance. How are these bounds to be set? Here much work remains to be done. Recall the reason why minimum and maximum dominance cannot be inferred. Any inference of maximum or minimum dominance must, if it is to be convincing, focus only on the relevant tail, and thus be based on very few observations. But then the inference will be very sensitive to unsampled population outliers and to measurement error, making it unreliable. Taking a more general view, outliers (long tails) and measurement error will pose a problem for all estimation attempts and the more sensitive a criterion to changes in either tail, the greater the problem posed. From this perspective, the choice of bounds will be

1. This is indeed the approach taken in Chapters Four and Five, where dominance criteria are used which require mean but not minimum dominance.

2. Non-parametric estimation methods will be subject to the same problem.

determined by a trade-off between a desire for criteria with wide coverage, and thus widely-separated bounds, and the constraints of inferability. The optimal solution to this trade-off will be data-set-specific: the greater the confidence one has in the representativeness and accuracy of one's data the wider the bounds can be. Whether any more specific guidance can be given in this regard remains to be seen. Certainly this would seem to be a fruitful area for future research. In the meantime, the rule-of-thumb that at least 30 observations are required to make use of asymptotic results might be applied here as providing rough constraints upon the placing of bounds for restricted dominance analysis. Using this rule, bounds should be placed so that at least 30 observations of each sample lie below the lower bound and another 30 lie above the upper bound in the case of inequality analysis.

If dominance between such widely-spaced bounds is absent, then the choice of bounds can take on a second role of making explicit the assumptions required to be able to rank distributions. In this case it may be more useful not to specify any narrower bounds exogenously prior to analysis but to report the bounds which emerge "endogenously", that is, which emerge, after inspection of data, as giving dominance over some range.

Whether or not the reader agrees with the above re-appraisal of the role of stochastic dominance in the measurement of living standards, I hope it is clear at least that it is useful to think of criteria not only in terms of curves (Lorenz curves, deficit curves and so on) but in terms of curves drawn between bounds. Varying the bounds as well as the curves used can add both flexibility and realism to the task of analysis.

PART TWO ANALYSIS OF URBAN CHINA

Chapter Four Income Inequality in Urban China in the 1980s: levels, trends and determinants

I Introduction

This chapter and the next turn the focus of the thesis away from questions of measurement towards those of analysis. Both examine the distribution of income in urban China over the eighties.

The introduction to the thesis gave the reasons for focusing on urban rather than rural or all China. To recap these are: the greater availability of data; an interest in making international comparisons, especially with the industrial, transitional economies of Eastern Europe and the former Soviet Union; and the policy implications of changes in the urban income distribution for reform of China's urban-based social security system. Although the focus in both chapters is on inequality, my examination of the income distribution is not restricted to its dispersion, and I also comment on trends in welfare and, to a lesser extent, poverty.

The structure of this chapter is as follows. Section II reviews the literature and assesses, from a theoretical vantage, the sorts of changes in inequality that one might expect to find in urban China over the last decade. In particular, the likely effects of transition are examined, focusing on a characterization of it as involving decentralization and new job opportunities. Section III contains 'preliminaries'. It gives the definition of 'urban' to be used in the study and provides an introduction to the State Statistical Bureau (SSB) survey (both of which are relevant to the next as well as this chapter). It also explains the characteristics of the aggregated, published data from the SSB analyzed in this chapter. Section IV places the income distribution in urban China in an international perspective. Section V investigates the trends in inequality in China over the eighties. Section VI puts the results obtained in a wider context by analyzing the interaction between changes in urban inequality and those in urban welfare and poverty, and by comparing the degree of inequality change in urban areas with that observed in rural China. Section VII analyses the determinants of the trends observed. Section VIII concludes.

II Literature review and theory

Section IV's focus on international comparisons notwithstanding, the major explanatory focus in this chapter is on trends in rather than levels of inequality. Hence this section examines research and theoretical issues concerning the question of what has happened to the distribution of income in urban China in the eighties and why.

II.1 Literature review

Perkins (1988) provides an excellent overview of changes in urban inequality from 1949 to the early eighties, which is worth quoting in full:

> ... urban sector inequality was reduced significantly in the first half of the 1950s by the confiscation (sometimes with modest compensation) of most privately held urban property. The within-urban inequality that remained after this state takeover of urban property was due to wage and salary differentials in state enterprises. These differentials were set by the central government and there was little regional variation in either the average urban wage or the differentials between one grade and another. Furthermore neither the average real wage nor the size of these differentials changed much from the time they were first introduced in the 1950s and early 1960s. In the early 1970s, and perhaps earlier, new entrants to the labour force came in at the lower end of the scale and there were few promotions. When wages were unfrozen in the late 1970s, the initial increases were concentrated in the lower wage grades, which should have reduced inequality further. (p.636)

And, indeed, it does seem that inequality fell in the early eighties. Zhao Renwei, who concurs that "in the early stage of reform emphasis was on raising the income of the lower paid workers" (1990b, p.192), presents urban Ginis of .185 for 1977 and .168 for 1984.[1]

It is less clear what has been happening in more recent years. As argued in the Introduction, despite the government's emphasis throughout the eighties on checking the excessive

1. Zhao Renwei's source is Li Chengrui (1986). These figures have received wide publicity. Gale Johnson (1990, p.76) quotes them with the Beijing Review as his source. Adelman and Sunding (1987) also find that the Gini coefficient in urban China fell between 1978 and 1983. However, the two years' Ginis of the latter authors are calculated by different methods, resulting in underestimation of the 1983 figure relative to that of 1978. (The 1978 figure is based on the Kakwani (1976) interpolation method which attempts to give the Gini which would result from using the disaggregated data. For data reasons, the 1983 figure is based on a linear interpolation. The resulting Gini gives the lower bound on the set of Ginis which could result from the disaggregated data.) Perkins (1988, p.637) also presents inequality measures for the early eighties. He calculates the coefficient of variation to be .31 in 1984 and .345 in 1981. This contradicts my findings (see Section V). But Perkins' results are open to the same criticisms as Ahmad and Wang's, which are based on the same 1981 data - see footnote 2.

egalitarianism of earlier decades, the consensus now is that, with the exception of growth in 'grey' and 'black' income, urban inequality has increased only little, if at all. But firm evidence is hard to come by. Ahmad and Wang (1991) compare 1981 and 1987 data. Nearly all their indices indicate an increase of inequality. Their analysis, however, suffers from a paucity of data, which makes interpretation of results difficult, and the authors themselves advise caution.[2] In addition, whatever the truth is for 1981 and 1987, it is difficult to analyze trends via extrapolation between end-points. Ma (1991) provides a time-series for 1981 to 1989 based on the same data as this study and calculates the ratio of the top to the bottom quintile's income share. Comparing the end-years suggests an upward tendency, but overall Ma finds "no sign of distinct deterioration" (p.21).[3]

A different approach is taken by Howes and Lanjouw (1991), who examine the link between reform and inequality by looking at whether the sea-board provinces, which have been at the forefront of reform, display higher levels of inequality. Using cross-sectional data collected by the Institute of Economics of the Chinese Academy of Social Sciences in 1986, they find no correlation (an answer also arrived at by Hussain, Lanjouw and Stern, 1991). However, this approach suffers from several weaknesses. It is difficult to define precisely which provinces have been leading the reform process. There is also much variation at the sub-provincial level. The data is from the mid- rather than late-eighties. And, most importantly, there is the usual problem of drawing time-series conclusions from cross-sectional data: in this case, we have simply no way of knowing what the inequality profile of the Chinese provinces looked like prior to reform and so cannot control for initial conditions.

A further set of interesting findings concerning inequality, both rural and urban, has emerged from the 1988 National Income Household Survey, conducted and analyzed by the Institute of Economics of the Chinese Academy of Social Sciences in collaboration with various

2. Ahmad and Wang draw on the same source as I will be using. However, the only information they use is the proportion of the population belonging to various income classes as presented in the SSB Yearbooks. No class means are available, which is a serious drawback. In addition, the six income classes used by Ahmad and Wang for 1981 are very variable in terms of proportion of the population included. The two largest classes combined contain three-quarters of the population (their sizes are 42% and 32% respectively). This makes any assumptions concerning group means particularly suspect. Finally, compared to the six classes for 1981, there are 16 for 1987. One would expect this to bias in an upwards direction any ratio of 1987 to 1981 inequality indices. An example of the importance of this sort of bias can be seen in Table 11. Using the same number of classes for 1989 and 1990 results in a fall in inequality; using the number of classes given for the two years results in higher inequality in 1990, as a larger number of classes are given for the latter year.

3. Although this chapter reaches different conclusions to those of Ma, its analysis reveals the pertinence of Ma's measure, as it is the top and bottom quintiles which have been most volatile - see VI.1.

foreign scholars. This survey, covering some 20,000 households in both rural and urban areas, provides the most accurate available source of information to date concerning the distribution of income in China at a single, recent point of time. Khan, Griffen, Riskin and Zhao (KGRZ, 1991) report Ginis based on this data of .23, .34 and .38 for, respectively, urban, rural and all China. (The all-China Gini is higher than the urban or rural due to the large urban-rural income gap.) While there can be no doubt that, for the one year and ten provinces covered, KGRZ provide a more accurate picture of the distribution of urban income than I am able to do using aggregated SSB data, the fact that their survey covers only one year means that it can have little to say about the dynamics of urban income distribution.

II.2 Some theoretical considerations

There are a number of factors which might be taken to have influenced the level of inequality in urban China over the past decade. As Table 1 shows, urban like rural China has experienced rapid growth through the course of the 1980s. Real income per capita growth rates have been high at 5%. This has resulted in a fundamental transformation of urban living standards, as can be seen from changes in the possession of durables. In 1981 only 6% of urban households owned washing machines, less than 1% owned a fridge and less than 1% a colour television. By 1990 these ownership levels had increased, respectively, to 76%, 36% and 51% (SSB, 1987s and 1991y).[4] Real growth rates have also been variable and inflation has been variable around an upward trend. There have been two booms in nominal demand, 1984-85 and 1988-89, both followed by periods of reassertion of macroeconomic control in 1985-86 and 1989-90. Whether or not there is an automatic link between growth and inequality for developing nations remains a matter of controversy. However, it is reasonable to expect that growth of the rapid, uneven type urban China experienced in the eighties would at the very least lead to distributional changes (a 'shake up' of the income distribution).

4. To facilitate referencing, SSB yearbooks will have their year of publication suffixed by a 'y', SSB publications based on the urban surveys by a 's', and those translated into English by 'ye' and 'se' respectively.

Table 1 Urban Household Per Capita Income Growth and Inflation, 1981-1990

Year	Household per capita income mean current prices, Yuan		Price index		Household per income mean 1985 prices, Yuan	
	Level	Growth Rate	Level	Growth Rate	Level	Growth Rate
1981	441.9	.	83.6	.	528.7	.
1982	471.2	6.6	85.3	2.0	552.4	4.4
1983	487.3	3.4	87.0	2.0	560.2	1.4
1984	571.4	17.3	89.4	2.7	639.5	14.1
1985	706.7	23.7	100.0	11.9	706.7	10.5
1986	810.1	14.6	107.0	7.0	757.1	7.1
1987	943.0	16.4	116.4	8.8	810.1	7.0
1988	1155.6	22.5	140.5	20.7	822.6	1.5
1989	1302.0	12.7	163.4	16.3	796.9	-3.1
1990	1425.2	9.5	165.5	1.3	861.1	8.1
Average	.	12.4	.	7.1	.	5.0

Notes: The price indices are cost-of-living indices for 'staff and workers' given in the annual China Statistical Yearbooks. The income figures are taken from Table A.3, except for 1982. 1982 figures are calculated from income class means given in SSB (1987s). More detail on the data used is given in III.3 and Appendix A. 1981-1985 figures are adjusted to take into account expansion of the sample base for published figures in 1986 to include towns: see Appendix A, point 2 for detail. The average figures in the last row are calculated on the basis of end-year values.

China's macroeconomic instability can be related to the process of reform (see Gale Johnson, 1990 and Hussain and Stern, 1991), which itself has implications for the pattern of inequality. China's transition can be characterized as a process of decentralization: from higher to lower levels of government and from government to private agents, both households and firms. Although reform in the Chinese context has not meant privatization, there has been an attempt to introduce greater management autonomy by the introduction of a contract system for managers and the granting to enterprises the right to retain post-tax profits. Provincial and local governments have also been given more autonomy to experiment with reform, to build up their own enterprises and to pursue their own foreign trade and investment opportunities. In addition, reform has involved a rise in new earning opportunities. There has been employment growth outside the traditional state and urban collective sectors, and the number of people with second jobs has increased as have the opportunities for earning non-labour income. Both these features of transition might be expected to have an impact on the level of inequality. Two very simple models (or, perhaps more appropriately, frameworks) are developed below to illustrate their possible impacts. Although both are China-specific, they may also be of use in thinking about distributional change in the transitional economies of Europe.

Both models make the assumption that there is no transfer of labour over the course of reform, whether between firms or provinces. This simplifies the analysis, and is realistic for modelling urban China, where, unlike in rural China, labour transfer has indeed been low. There is some migration, but the increase in the survey base has been small (see III.1). Within the urban sector, far and away the largest employer is the state-owned enterprise (SOE) sector. Its share in total employment has remained fairly constant at around 70%. The second largest employer, the collective sector, has also maintained a roughly constant share of employment at around 20%.[5] Transfers between firms within the same sector may be easier than transfer across sectors, but, by all accounts, they are nevertheless uncommon.

Inequality is measured in what follows below and throughout the chapter in accordance with the assumptions laid down in the first chapter (see Section II thereof). So all the measures used will possess the properties of: (a) replication invariance (so that measures are defined over distribution functions); (b) scale invariance (which restricts us to the class of relative inequality measures); and (c) conformity with the transfer principle. In addition, in the models below attention is restricted to the class of additively decomposable inequality measures so that, if the population is divided into groups, inequality can be written as a sum of between-group and (average) within-group inequality.

II.2.1 Decentralization

This model makes the simplifying assumption, relaxed in the next, that each household obtains its income from a single source. It focuses on reform as decentralization both from central to provincial government and from government to firms. Accordingly, each household i is identified with a firm j (from i to J) and province k (from 1 to K). Obviously K cannot change and J is also assumed fixed.

To analyze the effects of decentralization, it is necessary to distinguish between the change in inequality between the units to whom power is decentralized (the *recipient units*), in this case the provinces or firms, and the changes in inequality within the recipient units.

First assume that decentralization is only to provinces and that it has the effect of shaking up the provinces' relative means, as some provinces do better from decentralization than other, but leaves unchanged the distribution within each province. In this case, we treat the province as recipient unit, and, by the assumption of decomposability, total inequality will move

5. Growth in the urban collective sector has not been as rapid as that in the rural (township and village enterprise) collective sector.

with between-province inequality. Let the mean income in province k be Y_k. Then

$$Y_k^1 = Y_k^0 + \varepsilon_k^b \tag{1}$$

where '1' indicates post-reform and '0' pre-reform.

What happens to inequality depends on the relationship between ε^b, the change in income, and Y^0. We consider two archetypal cases. First, let ε^b be a (discrete) stochastic variable. This is an appropriate assumption if one thinks that decentralization is simply a 'noise-generating' process, so that some do well from being left more to their own devices, others do badly, but not in a predictable manner. In this case, one can write

$$\sum_{i=1}^{N_k} p_{ki}\varepsilon_{ki}^b = 0, \quad \sum_{i=1}^{N_k} p_{ki} = 1 \quad \forall k, k = 1,..., K \tag{2}$$

Note the summation is over the different possible states of the world which, for province k, will come about with probability p_{ki}. *Ex ante* inequality will be higher in the sense that the distribution of Y^1

$$\{(p_{11},Y_{11}^1), ..., (p_{1N_1},Y_{1N_1}^1), (p_{21},Y_{21}^1), ..., (p_{KN_K},Y_{KN_K}^1)\} \tag{3}$$

will be Lorenz-dominated by the distribution of Y^0

$$\{Y_1^0, ..., Y_K^0\}, \tag{4}$$

This result follows directly from Rothschild and Stiglitz's (1970) proof that adding 'noise' to a distribution increases its riskiness, or here, *ex ante* inequality.[6] *Ex post* inequality may be higher or lower: one cannot discount the possibility that it will be the poor who end up lucky, and the rich unlucky.However, since any post-reform income profile can be regarded as a random draw from Y^0, the value taken by some decomposable inequality index for this profile can be regarded as a consistent estimator of *ex ante* inequality and so should rise with the latter, at least if K is large enough.

But the distributions of gains and losses from decentralization may be more systematic than the preceding analysis permits. So consider the alternative assumption that ε^b is deterministic, and dependent on Y^0. Then the effect of reform will depend on whether it can be thought of as implying the existence of regressive or progressive transfers. If, for example, we assume that ε^b

6. For the relationship between Lorenz dominance and lower inequality, see Chapter 1, III.2.

is a linear function of Y^0 and that the sum of ε^b is zero, then a positive derivative will ensure a less equal distribution results, while a negative derivative a more equal distribution.[7]

Analysis of decentralization to the firm level can proceed in exactly the same way (simply by replacing k by j), as can analysis of the impact of decentralization on inequality within recipient units. For the latter case, assume that decentralization has the same effect across units on within-unit incentive structures and that all recipient-unit total incomes respond proportionally to whatever change in the incentive structure is made. In this way, one can abstract from between recipient-unit inequality, and total inequality will follow within-unit inequality. Then for any individual in firm j or province k, one can write:

$$Y_i^1 = Y_i^0 + \varepsilon_i^w \tag{5}$$

Clearly, the effects of decentralization within the recipient unit can be analyzed in exactly the same way as before.

The above provides a framework for thinking about the distributional consequences of decentralization. A great variety of outcomes are possible. Which are those likely to be observed in urban China? Beginning with the distribution of income within recipient units, the distributional outcome will depend on the incentives guiding the units. If we assume a fairly egalitarian distribution of within-unit income prior to the reform and if the impact of the reform is to turn the recipient units into profit-seekers, then a worsening of the distribution can be expected. This follows automatically if the initial distribution is completely egalitarian, and some variation in compensation is introduced. More realistically, let there initially be a weak (gently sloped) but positive relationship between payment and productivity, and let this relationship be strengthened (the slope steepened) by the reform process. This is a case in which ε^w depends on Y^0 and in which inequality worsens, as the initially well-off gain at the (relative) expensive of the initial poor, due to the regressive distribution of productivity. Decentralization does not, however, automatically lead to profit-seeking behaviour. In particular, if the recipient unit comes under the influence of its workers, in the case of a firm, or constituents, in the case of local government, there may be little move away from the original egalitarian distribution. Many analysts have argued that this is the case in China with regards to firms. Although firms now have a great deal of wage-setting power, this has not led to a growth in within-firm wage differentials, it is argued,

7. In the first case, since the rich must gain more than the poor, and since there must be winners or losers, the poor must be losers and the rich gainers. This is reversed in the second case. The effects of various deterministic distributional transformations on the level of inequality are considered in the literature on the distributional impacts of taxation (see Lambert, 1989).

due to worker control over the setting of remuneration levels.[8]

Turning to between-unit inequality and beginning at the provincial level, the general consensus would seem to be that regional disparities have widened. Though there has in fact been little solid evidence presented to back this proposition, it does conform with the widespread impression that the relatively well-off coastal provinces have been the largest beneficiaries of reform and is confirmed by the findings of this chapter. There is even less evidence at the between-firm level, but again the impression is one of increased variation in profitability, suggesting that decentralization has indeed increased the importance of 'noise' in the determination of income. In addition, since firms have greater freedom in their distribution of profits as a result of reform, it may well be that some firms are distributing their profits in kind (e.g., better housing) while others have continued to pay in cash. In this case, even if everyone's total income remains (proportionally) constant, their cash income, analyzed here, will display variation.[9]

II.2.2 The distribution of new income earning opportunities

As we will see in Section VII, income from sources other than the two traditional state-owned and collective sectors and pensions accounts for some 7% of total household income. This 'non-traditional' or 'new' income does not, however, tend to go exclusively to households outside the traditional sector. The number of full-time private sector workers remains at less than 2% (see Table 2). Rather, traditional sector workers and households also earn non-traditional income. This may involve second-jobs, property income (interest receipts or rents) or earnings by the officially retired. These non-traditional sources of income have been growing at least as fast as the traditional sources. I now turn to the question of who has been benefiting from the growth in new income.[10]

Since this growth in new earning opportunities can be seen as the result of a type of decentralization, to the household rather than the province or firm, it is not surprising that a

---

8. Ma (1991, p.16) claims in fact, that the "several wage adjustments and the introduction of bonuses ... ended up quite equalizing".

9. This point is further discussed in V.2.3.

10. One could equally ask: who has been benefiting from the growth in traditional income? Such a question should be regarded as being dealt with in the previous model, on the simplifying assumption that all income is traditional income. It could be argued that the two models should be unified. However, it is not clear that additional insight would be gained from this. In addition, owing to data limitations, the distributional consequences of the two aspects of reform considered in the two models can only be examined separately.

framework similar to that employed in II.2.1 can be used to analyze it. To keep things simple, we abstract from the fact that different households will be working in different provinces or firms. Instead it is simply assumed that every household is employed in one of the two traditional (T) sectors (SOE [sector 1] or collective [sector 2]). Each household may also earn non-traditional (NT) income. Thus any household i has income

$$Y_i = Y_i^{T_k} + Y_i^{NT},\ k=1 \text{ or } 2 \tag{6}$$

I will also assume that income from all three sources is growing at the same rate. This is an appropriate approximation for urban China, the reform of which has not meant the decay of the traditional sector. In practice, a positive growth rate is the mechanism by which the distribution of non-traditional income will change, as some come to participate more than others in that growth. For the purposes of the model though, we can abstract from this fact and set the common growth rate equal to zero. It is also (realistically) assumed that the average total (traditional and non-traditional) income of an employee in the SOE sector is greater than that of one in the collective sector. Thus,

$$\mu_{T_1} + \mu_{NT|T_1} > \mu_{T_2} + \mu_{NT|T_2} \tag{7}$$

where μ indicates mean income and the vertical bar indicates a conditional mean. To focus on the non-traditional sector, we assume that traditional incomes are unchanged by reform. To ensure a zero growth rate for non-traditional income, it is assumed

$$Y_i^{NT,1} = Y_i^{NT,0} + \epsilon_i^{NT},\ \sum_i \epsilon_i^{NT} = 0 \tag{8}$$

where the summation is across households. Now consider the possibility that access to new income earning opportunities depends on which traditional sector you are in. So, assume

$$\epsilon_i^{NT} = \begin{cases} \alpha Y_i^{NT,0} \text{ if } i \in T_1 \\ \beta Y_i^{NT,0} \text{ if } i \in T_2 \end{cases} \tag{9}$$

In this case, inequality in each of the two sectors of population will be constant. Between sector and thus total inequality will increase if $\alpha>0$ and $\beta<0$ and fall if the signs are reversed. Inequality will be unchanged if $\alpha=\beta=0$. (To satisfy (8) α and β must be oppositely signed or both zero, so these are the only three cases possible.)[11]

11. In fact, as Figure 2 of Chapter Five shows, many households receive income from both the SOE and collective sectors. For the model to be more realistic, one would need ε_i to be a function of the proportion of total traditional income earned in the SOE sector. However, the same sort of result would emerge.

Which is the most likely outcome for urban China? This is harder to call than the effects of decentralization. On the one hand those in the SOE sector may have the connections and training to make better use of new earning opportunities. On the other hand, those in the collective sector, which is less regulated, may be more entrepreneurial in approach. With lower incomes, they may also have more incentive to search out new earning opportunities.

To conclude this brief analysis, it must be stressed that the aim of this section has not been to provide a comprehensive model of inequality-generating and -suppressing forces in urban China.[12] This would be an extremely ambitious project, and one which could not be supported by the available data. Rather, the aim has been the more modest one of setting forth several considerations of a theoretical nature. From the foregoing, it would seem that the decentralization associated with reform has probably been inequality-promoting:[13] there is no reason to expect between-province and between-firm inequality to have fallen; and while, as many have claimed, within-unit inequality may not be much altered, it is hard to believe it has declined from its very low pre-reform levels. A confident prediction that reform has increased inequality cannot be made, however, on account of the uncertainty as to the distribution of new earning opportunities. If this 'new' income has gone primarily to those households based in the poorer, collective sector, inequality may have fallen. In addition, whether China's macroeconomic instability has promoted inequality also needs to be investigated.

The links between reform and inequality are returned to in VII.3, once the empirical analysis has been completed.

12. For example no role is given to capital income. Instead it is simply assumed that the return to capital is paid to the government (central or local) as owner, and re-distributed by the government in a distributionally-neutral way.

13. Though, as stressed earlier, this is not inevitable. For example, in China in the late seventies and early eighties, the chief beneficiary of decentralization was the rural sector - it began to catch up with the urban sector, and overall inequality fell (see Adelman and Sunding, 1987).

III Definition of 'urban', and data source and availability

This section evaluates both the data collected by the SSB (III.2) and the data available in aggregated, published form (III.3). The first of these tasks is relevant to both this and the next chapter, the second only to this. The definition of the survey base is a particularly complex question and is dealt with separately immediately below.

III.1 The definition of 'urban'

The SSB annual surveys analyzed in this and the next chapter are surveys of households registered as non-agricultural and living in areas defined by the SSB to be urban. To what extent this makes them surveys of *urban* China is not easy to say. Four different definitions of 'urban' are utilized in Figure 1. The first (labelled 'Official'), which gives a rapidly increasing share of urban to total population, is the total number living in areas (cities or towns) designated officially (by the SSB) to be urban. Under this definition, the urban population has increased rapidly since 1984. As Riskin explains

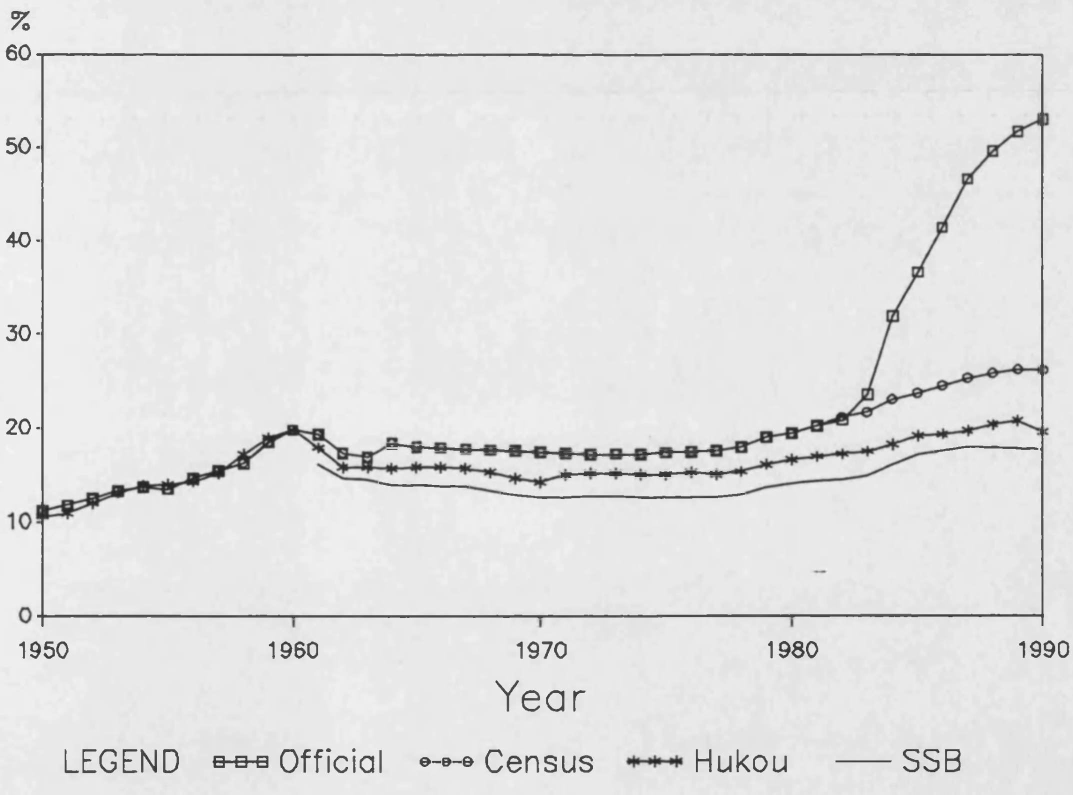
> In 1984, the standard for town designation was considerably relaxed and a change also took place in governing procedures that permitted cities and towns to incorporate large numbers of surrounding farm households within their administrative boundaries. The result was a rapid nominal growth in urbanization resulting in half the population becoming urban by this standard in 1988. (1991, p.21)

As approximately 70% of the Chinese population is engaged in agriculture, this definition now results in some 35% of urban dwellers being peasants. The second definition ('Census') is used in Census-based publications (see, for example, National Population Census Office, 1991a and 1991b) and gives the total number living in areas designated by the Census office to be urban.[14] It gives a proportion urban identical to that under the first definition up to 1982, but remains stable thereafter, slowly increasing to 25% by 1989. The third definition ('Hukou') is based on household registration (*hukou*) status. Each Chinese household is registered as either agricultural or non-agricultural: registration as the latter entitles a household to a number of benefits such as, at least until recently, ration coupons for the purchase of subsidized foodstuffs. The proportion registered non-agricultural has been rising only slowly and is currently around 20%. The fourth and final definition ('SSB') is the survey base from 1985, and, from its definition at the beginning of this paragraph, is the intersection of the first and third definitions. So it excludes those households

14. For a precise definition, see Riskin (1991) or National Population Census Office (1991a).

headed by rural professionals who, on account of the head's university training, have non-agricultural status. Comparing the third and fourth curves this number is around 10% of all non-agricultural-registered households. While, for analysis of urban China, this group should be excluded, the other group omitted on account of the SSB survey base is of greater cause for concern. This is those households living in urban areas but without registration - in the main recently-arrived rural migrants.[15] If we take the census definition of 'urban' as the ideal definition, then this group constitutes one-fifth of the urban population as it should be measured, or some 60 million. There is, however, no alternative to using the SSB survey base and thus excluding this so-called 'floating' population. Even the 1988 CASS survey, mentioned in II.1, uses the SSB survey base as the pool from which to draw its sample.

Figure 1 Urban as a Percentage of Total Population under Four Definitions, 1950-90



Sources: SSB (1990a) for 'Official' and 'Hukou' definitions up to 1989; SSB (1991y) for 'Census' definition up to 1989; Hayase and Kawamata (1990) for 'SSB' definition between 1961 and 1987. National Population Census Office (1991b) for all 1990 data. There are no figures for 1988 and 1989 using the SSB definition.

15. These households might be non-agricultural by occupation but still be without non-agricultural registration.

III.2 Data source

III.2.1 Sample history and method

Each year the SSB carries out two household surveys, one of rural residents and one of urban residents (as defined above). The urban survey was first undertaken in 1955, but was suspended from 1965 to 1978 on account of the Cultural Revolution. Major developments in the survey have taken place since 1978: the survey has grown in both detail and coverage. The urban sample is now a relatively large one, having increased from some 9,000 households in the early eighties to 35,000 households nowadays (the rural sample is even larger). Details of the sampling method are given by Ren and Wang (1992). Summarizing, the annual sample is drawn by selecting annually approximately 10% of the households contained in a larger sample chosen every three years of size 300,000. This larger sample, known as the 'one-time' sample, is drawn in the following way. First, China's provinces are divided into six regions (using the standard administrative division) with the sample size in each region being proportional to that region's population. Next, within each region, all the provincial capitals are chosen for survey and a random selection of other cities and towns is made. Within the chosen cities and towns, a further random selection of neighbourhood committees (the lowest-level administrative organ, covering the population of a single block of flats, for example) and finally households is made.

Basic data on household size, income and employment status are collected from this one-time sample. To select the annual sample, the one-time sample is ordered on the basis of household income and divided into five income classes. Representative samples are then chosen from each of these five classes, subject to the constraint that each city to be included in the annual survey has a minimum of 100 households sampled and each town a minimum of 50 households.[16]

Those chosen for the annual survey are asked to record details of household income and expenditure. Each survey period lasts a year. Interviewers visit the household at least twice every month. At the end of each month, they collect that month's record card and provide a new card for the following month. Information on a large number of variables is collected. The precise number has risen over the eighties and is currently around one thousand. Information on wages is substantiated by the interviewers via examination of work-unit pay records.

16. From 1990, the annual sample has become a rotating one so that a proportion of households are sampled for two or more consecutive years.

The most striking feature of the survey is surely its duration of a year. By contrast, households in Hungary are only asked to record information for two months and in Poland for four (see Atkinson and Micklewright, 1992, 'Sources and Methods'). With such extensive demands being made, one would expect low response rates. Unfortunately, no official information on the response rate is available, but I have been informed by an SSB officer that they are in fact very high. If a household is initially reluctant to participate, it is re-visited by the SSB to encourage it to do so, though households do have the right of refusal. Some financial compensation is also offered to interviewees, but this is relatively small: in 1989, it constituted 1% of the per capita income of the surveyed households (see SSB, 1990s).

Table 2 provides summary data on the decade's urban household surveys. It highlights the most important changes in the structure of the survey between 1981 and 1990. The sample size has increased every year from 8715 in 1981 to 35650 in 1990. The biggest increase came in 1985, where the sample size almost doubled from 12 to 24 thousand. 1985 also represents a watershed in two other respects. First, prior to 1985 the only households represented in the survey were those whose head was employed in the state-owned or collective sectors. This was changed in 1985, when the current survey base was introduced. This meant that those households headed by workers who were self-employed or employed in the private sector were included for the first time. More importantly (since such workers never became more than 2% of the surveyed workforce - see Table 2), the proportion of retired in the sample increased, as households headed by the retired were also included for the first time. The proportion of sample household members retired jumped from 3% in 1983 to 6% in 1985 (data for 1984 is unavailable), around which level it has since remained. Relatedly, the average proportion of employed to total household size fell from 58.4% in 1984 to 55.3% in 1985.[17] Second, prior to 1985 only non-agricultural-registered households resident in cities were sampled. From 1985 onwards between one- and two-fifths of those sampled were drawn from county towns (though households without non-agricultural registration were still excluded). On average, households in towns are poorer and have a lower proportion of workers employed in the state-owned sector, a larger household size and a smaller proportion of retired than households in cities (see Tables 2 and 10).

The survey questionnaire has also been changed through the course of the eighties. It was changed in 1988 and, at least judging by the greater detail of information released from 1985 onwards, in 1985 as well. These changes have not affected the basic definitions of income used, but they have led to the provision of greater detail.

---

17. Apart from the fall in 1985, this ratio has been rising over time due to falling household size, a result of China's one-child policy.

Table 2 The SSB Urban Household Survey, 1981-1990: some basic information

Year	**1981**	**1982**	**1983**	**1984**	**1985**			**1986**		
Coverage	**Cities**	**Cities**	**Cities**	**Cities**	**Urban**	**Cities**	**Towns**[a]	**Urban**	**Cities**	**Towns**
Sample Size	8715	9020	9060	12500	24338	17143	7195	.	.	.[b]
Ave. H/h Size	4.2	4.1	4.1	4.0	3.9	3.8	4.0	3.8	3.8	3.9
% H/hold employed	56.4	57.7	58.6	58.4	55.3	57.4	51.8	55.5	57.9	52.3
Of which: SOE (%)						73.4				
Collective (%)						23.2				
Private and self-employed (%)						.9				
% H/hold retired		2.5	3.0	3.0		6.2	4.3			

Year	**1987**			**1988**			**1989**			**1990**		
Coverage	**Urban**	**Cities**	**Towns**	**Urban**	**Cities**	**Towns**	**Urban**	**Cities**	**Towns**	**Urban**	**Cities**	**Towns**
Sample Size	32855	25265	7590	34945	27655	7290	35235	27795	7440	35650	28885	7580
Ave. H/h Size	3.7	3.7	3.9	3.6	3.6	3.7	3.5	3.5	3.7	3.5		3.6
% H/hold employed	55.9	57.7	52.5	56.0	57.8	53.4	56.3	58.0		56.6		54.3
Of which: SOE (%)	74.2	75.4	71.8	73.7	74.7		74.6	75.4				
Collective (%)	22.0	21.1	23.7	21.9	20.7		21.0	19.9				
Individual (%)	1.4	.9	2.3	1.7	1.1		1.7	1.1				
% H/hold retired	6.7	7.3	5.2	6.6	7.6	5.1	6.8	8.1	5.2			

<u>Notes:</u> a. Given in SSB, 1988se.

b. Conflicting figures are given in different sources. In SSB (1988s and 1989se) the total sample size is 27024 of which cities make up 17046 and towns 9978. In later yearbooks, a larger total is given of 31126 with a smaller total for towns of 7478 (see SSB, 1987y). However, the data given in these different sources otherwise seems identical.

1. Unless indicated to the contrary in the notes above, the figures come from the sources indicated in Table A.1.

III.2.2 Sample representativeness

It is difficult to find flaws in the system used to make the urban survey a random sample of urban China. Note in particular that the sample is household and not workplace based (as in the former Soviet Union - see Atkinson and Micklewright, 1992). On the other hand, very little attention seems to have been paid, inside China or out, to the representativeness of the sample. Unfortunately, sample representativeness is not at all easy to investigate. Consistency checks with other sources can only be carried out if a rural-urban breakdown of the variable of interest is available. For many variables, it is not. Where it is, the definition of 'urban' used is typically based on place of residence not on registration (see III.1). Such progress as has been possible subject to these tight constraints is reported below.

One check which can be made is on sample composition by province, and by city and town residence. The provincial sampling proportions are closely in line with population proportions. However, town-dwellers are under-represented vis-a-vis city-dwellers. The relevant population data are available for 1981-1985 and 1990. As Table 3 shows, the population ratio of city to town dwellers fell dramatically between 1981 and 1985 - from 2.5 to 1.3 - on account of the official redefinition of 'urban'. The population ratio of city-resident-and-non-agricultural-registered to town-resident-and-non-agricultural-registered, on the other hand, is fairly constant at around 2.2. However, the SSB sample gives a value for the latter ratio of between 3.4 and 3.8 between 1987 and 1990.[18]

18. The ratio is lower for 1985 at 2.4, and unclear for 1986: see the notes to Table 2.

Table 3 City and Town Population Percentages and Ratios, Total and Non-agricultural, 1981-1990

Year	Urban		City		Town		Ratio of city to town		
	Total	NA	Total	NA	Total	NA	Total	NA	Sample
1981	20.2	14.3	14.3	9.8	5.8	4.5	2.5	2.2	
1982	20.8	14.5	14.7	10.0	6.1	4.5	2.4	2.2	
1983	23.5	14.9	17.5	10.5	6.1	4.4	2.9	2.4	
1984	31.9	16.1	18.9	11.1	13.0	5.1	1.5	2.2	
1985	36.6	17.2	20.7	11.7	15.9	5.5	1.3	2.1	2.4
1986	41.4	17.5							.
1987	47.3	18.0							3.4
1988									3.8
1989									3.7
1990	53.0	17.7	29.5	12.7	23.5	5.0	1.3	2.5	3.8
1990 census	26.0	16.6	18.7	11.6	7.3	5.0	2.6	2.3	

Notes: The definitions used are the official ones, except for '1990 census' which uses the census definitions (see III.1). All the figures except the ratios are percentages of China's total population in the relevant year. 'NA' stands for 'registered as non-agricultural'. For the official data, the source for 1981-1987 is Hayase and Kawamata (1990); for 1990, see National Population Census Office (1991b). The sample figures are calculated from Table 2.

A rough check by employment is also possible. The SSB sample ratio of SOE to collective employees has been constant at around 3.4 for the second half of the eighties. The ratio of SOE to urban collective employees given in the official statistical yearbook is around 3.3 from 1981 to 1983 and around 2.9 from 1984 onwards (SSB, 1990ye). Recall that, from Figure 1, up to 1983, the official and SSB survey definitions of 'urban' were quite similar. The fact that the official SOE-collective ratio falls in 1984, the year in which the definition of 'urban' was changed, suggests that the fall is due to the change in definition. This makes sense: towns have a higher proportion of their labour-force employed in collectives and, as Table 3 shows, it was the township population which jumped most dramatically due to the definitional change. The upshot is that, although it might seem that the SSB survey is biased towards SOE employees, the bias vanishes, or at least becomes less serious, if one compares the survey figure of 3.4 with the more appropriate official figure, given the sample base, of 3.3. Note too that, whatever the truth about levels, both sample and official data present the same picture with regards to trends: both have a constant SOE-collective ratio for the latter half of the decade.

A similar check is available on household size. The 1% population survey of 1987, using the census definition of 'urban', reported an average household size for cities of 3.8 and towns of 4.1. These compare with survey figures for that year of 3.7 and 3.9, respectively (see Table 2). The discrepancies are again likely to be caused by the presence of urban-dwelling agricultural-registered households, since these are more like rural households and so likely to have a larger household size.

A final check can be made in relation to educational attainment, using the disaggregated data from Liaoning and Sichuan analyzed in the next chapter. The survey estimates for 1990 can be compared with the 10% tabulation of the 1990 census, for both city-dwellers and town-dwellers. It is evident from Table 4 that educational attainment levels are higher in the SSB sample. Of course there is the difficulty of different sample bases. Those without non-agricultural registration are likely to be less educated and this may explain the discrepancy. The fairest comparison possible is in relation to Sichuan town-dwellers in 1990, 88% of whom have non-agricultural registration (see Table 2 of Chapter 5). Even here the gap between the SSB and census results is large. According to the SSB sample, of those with primary education or above, 19% have a university or college education. According to the census figures, this proportion is 13%. It seems unlikely that this difference of a third can be explained by reference to the 12% of Sichuan's town population outside the sample base.[19] A more reasonable explanation is a bias in the SSB survey towards the more educated. How this bias has arisen, if it does exist, can only be a matter for speculation.

Queries have been raised as to whether Party members are over-sampled. This cannot be formally checked, but was strongly denied by a SSB officer I spoke with.

19. The most favourable assumption possible for reaching a conclusion of no bias is that all of the 12% have at least a primary education, but that none have a tertiary or college education. Even with these unrealistic premises, one would expect, on the basis of the SSB figures, a census figure of 16% rather than the 13% obtained.

Table 4 Education Levels, Sample and Official, 1990 (%)

Table 4a Cities

Province		Five	Four	Three	Two	One	All
		-------Attained educational level------					
Liaoning		12.6	8.1	18.8	38.7	21.8	90.3
	Official	6.4	4.4	15.6	44.8	28.8	83.4
Sichuan		11.6	11.7	18.1	34.2	24.4	91.3
	Official	5.8	4.2	12.8	34.1	43.1	79.6

Table 4b Towns

Province		Five	Four	Three	Two	One	All
		-------Attained educational level------					
Liaoning		3.3	12.5	28.1	37.3	18.8	89.1
	Official	3.1	4.0	16.5	43.7	32.7	83.0
Sichuan		7.6	11.7	18.4	38.2	24.1	91.0
	Official	4.7	8.6	15.3	39.0	32.4	84.4

Notes: Unless indicated to be 'Official', the figures are based on the SSB samples for Liaoning and Sichuan analyzed in Chapter Five, which can be consulted for further details. The source for the official data is National Population Census Office (1991b) and the definitions used of cities and towns are those of the Census. The first five figures in each row give the total who have attained the given educational level (one to five) as a percentage of those who have attained at least level one. The final figure is the total of those who have attained at least level one as a percentage of the total population. The different attainment levels are defined in terms of those who have graduated from or who are currently attending: one, primary school; two, middle school; three, high school; four, technical college; five, university.

III.2.3 Sample definitions of 'household' and 'income'

The definition of a household is similar to that used in Eastern Europe, that is, "a group of individuals at the same address who partly or entirely share living expenses" (Atkinson and Micklewright, 1992, p.69).

The key income variable collected in the survey, at least so far as this thesis is concerned, is *shenghuofei shouru*, or 'disposable income'. Disposable income is gross total income (*quanbu shouru*) net of net outgoing gifts and remittances (about 5% of total income - the urban sector is a net remitter), boarding fees paid by friends and relatives (*qinyou huofei shouru*) and subsidies paid to defray recording expenses for surveyees. Both of these latter two are very small. The last is excluded since it is a source of income available only to those surveyed. The former is excluded as it is intended to catch only those transactions paid by non-household members designed to cover their costs. Disposable income is gross of direct taxes, but these are in any case minuscule: less than 0.1% of disposable income in 1989. A more common translation of *shenghuofei shouru* is 'income available for living expenses'. This more literal translation is too clumsy to be satisfactory. Moreover, the term 'disposable income' fits well, as one can think of net gifts as being a prior claim on income, just as direct taxes are in the West. All the published figures available are based on an ordering in terms of disposable income, making this variable the natural focus for this study.

The main sources of gross total income are state-owned sector wages, collective sector wages and pensions. Other sources include income from self-employment and private employment, income from second jobs, and income from sale or rental of property (including interest). Only cash income is included. To get an idea of the importance of the omission of non-cash income, one can compare the SSB figures with those reported by KGRZ, whose survey was based on a sub-sample of the SSB sample and who do try to account for various non-cash income sources. This is done in Table 5 for 1988, the year in which the data set analyzed by KGRZ was collected. Using the SSB definition of income, the two sources give close estimates of mean income, differing by only 3%.[20] However, SSB mean income is less than 2/3 of KGRZ income as defined by the latter. The most important reason for this is that KGRZ include the imputed value of the cheap housing (estimated as a fraction of replacement cost) and ration coupons which non-agricultural-registered households are entitled to. These two account for, respectively, 51 and 15% of the overall difference between the two measures. Also, KGRZ give a more comprehensive coverage of workplace subsidies, including in-kind subsidies, which accounts for 17% of the difference. Finally, KGRZ include the imputed value of owner-occupied housing, which accounts for 11% of the difference.

20. This comparison is itself based on estimates: see the notes to Table 5. The fact that they are so close may more be coincidence than anything else, since the KGRZ survey is based only on 10 provinces, and the SSB estimates give a greater weight to town observations than the KGRZ survey. Nevertheless, the comparison does suggest the sort of figures which would emerge from the SSB survey if non-cash sources of income were assessed.

Table 5 A comparison between KGRZ and SSB data, 1988

	KGRZ (A)	SSB (B)	%Difference (A-B)/B
Total income	1842.0	1192.1	35.3
Total income (SSB definition)	1230.4	1192.1	3.2
(a) Cash income of working members	818.3	807.6	1.3
(b) Cash income of the retired (earnings & pension)	125.8	108.1	14.1
(c) Income of non-working members	8.7	-	-
(d) Income from private/individual enterprises	13.6	16.3	-20.4
(e) Income from property	9.1	7.4	18.8
(f) Subsidies and income in kind	719.9	180.1	75.0
(fi) Ration coupon subsidy	97.0	-	-
(fii) Housing subsidy in kind	334.1	-	-
(fiii) Other subsidy less tax and payments in kind	288.8	180.1	37.6
(g) Rental value of owner occupied housing	71.8	-	-
(h) Other: private transfer and special income from other sources	74.9	72.7	2.9

Notes: 1. KGRZ's definitions are used. The corresponding SSB definitions (explained in Table A1.2) are: for (a), 1+2-1.3-2.3+5.1.1+5.1.3+5.1.4; for (b), 3+5.1.2; for (d), 5.1.1; for (e), 5.2.2; for (fiii) 1.3+2.3; for (h), a residual obtained by subtracting from total income the sum of (a) to (g). Note that (a) to (e) exclude cash subsidies. There are certain, minor discrepancies between the categorizations. In particular, the SSB data do not distinguish between the pensions of earning retired and those of the non-earning retired. Both are put into the equivalent of (b). KGRZ do: the former is included in (b), the latter in (c); also, for KGRZ, (d) excludes labour income from private/individual enterprises which is in (a), but, for my SSB-based classification, (d) includes labour as well as other factor income.

2. Note that the SSB figures used in this table are not those used in the chapter's analysis. Rather they are averages actually published by the SSB (see SSB, 1989s). They differ from the figures used in this chapter (a) by being individual per capita rather than household per capita averages (see Section V) and (b) by using total rather than disposable income (see III.2). This is done to make the figures as comparable as possible with those given by KGRZ.

3. 'Total income (SSB definition)' is obtained by subtracting from the total (fi), (fii), (g) and the difference between the KGRZ estimate of (fiii) and the SSB estimate. (The former includes cash and in-kind workplace subsidies, the latter only cash subsidies; the difference between the two is taken as a proxy for the value of in-kind workplace subsidies, which is not given separately by KGRZ.)

To conclude this assessment, it is instructive to compare the SSB survey with the household surveys of Hungary, Poland and Czechoslovakia from the 1980s analyzed recently by Atkinson and Micklewright (1992). The authors compare these surveys with data available for Britain and conclude:

> The surveys in Eastern Europe typically have had larger samples, have had higher response rates, have been able to substantiate earnings data from employers, have considered the deviations of survey results from macro aggregates, and have set out to collect information on annual income. The Eastern European sources have significant deficiencies, and there are undoubtedly aspects which are not adequately covered, such as private incomes, legal or illegal, which may well have been a growing feature of the 1980s. It is important to bear these in mind when interpreting the data, but it must be

remembered that data from all countries are deficient in some respect whether they come from market or socialist economies. (1992, p.16)

Concerning sample size, substantiation, the collection of annual income data and coverage of the private sector the SSB surveys come out either as well as or better than the East European ones. (For example, neither the Poland nor the Czechoslovakia surveys covered private sector employees, whereas the SSB survey does, at least from 1985 on.) The same concerns arise about the coverage of semi-legal and illegal incomes, and the Chinese survey compares less well with regards to availability of information concerning response rates and "the deviations of survey results from macro aggregates". Insofar as it has been possible to investigate the latter question, there would seem to be an under-representation of town-dwellers and an over-representation of the well-educated.

III.3 Data availability

Unfortunately, the raw data from these annual urban surveys are not in the public domain. Indeed, the size of the surveys is so large relative to the SSB's computing facilities that not even the SSB itself stores all the raw data. Instead, averages are calculated at various locations around the country, and these are collated centrally. Only a sub-sample of each annual survey is received by the central office in disaggregated form. These sub-samples for two of China's provinces are analyzed in the next chapter, but for a China-wide study one is restricted to reliance on aggregated data.

Tabulated aggregations for 1981 to 1985 are brought together in one book (SSB, 1987s). Data for 1986 to 1989 appears in separate books (SSB, 1986s, 1988s, 1989s and 1990s). Some of the information from these books is extracted in China's statistical yearbooks, which provides the study's only source for 1990 (SSB, 1991y). Data for 1985 and 1986 is also available in two English-language publications (SSB, 1988se and SSB, 1989se). Where more than one source was available for any one year, consistency between sources was fortunately only rarely a problem. Table A.1 shows which sources were used.

The variety of publications in which the tabulations from the survey data appear means that different types and levels of detail of data are available for different years. However, quantile group[21] means of per capita 'disposable income' are available for each year. These provide the core data for this chapter. The latter years (from 1985 onwards) have the highest number of quantiles, eight in all: 0-5, 5-10, 10-20, 20-40, 40-60, 60-80, 80-90 and 90-100%. Prior to 1985, the top two quantiles are merged into one; prior to 1983, the bottom two quantiles are also merged.[22]

Although results from the 1982 survey are published, they are not utilized here, since, for some reason, for this one year quantile means for disposable (as against total) income are not given. Though sufficient data are given to gain a rough picture of the situation in 1982, the year does not figure in the more detailed distributional comparisons given in the chapter.

The amount of data available increases over the eighties, with changes coming in 1985 and 1988. For example, pension income is included as a separate category for the first time in 1985. And in 1988, information on income received as compensation for rises in state prices begins to be given. Detailed information on the urban, as against cities-only, sample is available only from 1986 onwards, even though it was first collected in 1985.[23] Hence the SSB publications contain two structural breaks in terms of coverage: from 1985 the sample base widens to include non-traditional sector workers and the retired, and in 1986 it widens still further to include towns as well as cities. The first change has little impact on mean income, but adjustment needs to be made for the second change.[24] To make the pre- and post-1986 data comparable, the mean incomes of the earlier years were adjusted downwards. Appendix A, point 3 explains how this was done.

The fact that the data set is aggregated is not in itself an enormous drawback. It is true that most analyses based on aggregated data use deciles (e.g., Jenkins, 1991) or quintiles (Anand,

---

21. The term 'quantile group', sometimes shortened to 'quantile', is used throughout (as in Phelps Brown, 1988) to refer to the case where data is given for groups of the population fixed by population share. Quintiles and vingtiles are both examples of quantile groups. Groups defined by income bounds are referred to as 'income classes'.

22. See Appendix 1, point 4 for the adjustments made to cope with the missing data.

23. The English translations of the SSB survey publications for 1985 and 1986 use the word 'urban' in two different ways. With reference to the 1985 data (that is, in SSB, 1988se), 'urban' refers only to cities. With reference to the 1986 data (that is, in SSB, 1989se), 'urban' is used as in this paper to refer to cities and county-towns combined. Note also that in fact city and town quantiles are available for 1985 (city in SSB, 1987s and town in SSB, 1988se) and so could be combined - using the SSB aggregation and re-weighting procedures (on which see below) to give urban quantiles. However, this procedure was not followed since the quantile groups were not completely comparable and much more detail was available on the city data.

24. The second change reduced the mean by almost 10% (see point 3 of Appendix A). The first change reduced it by less than 1%, using the provincial data sets as evidence.

1983). Having a maximum of 8 quantiles leaves one with a comparatively sparse amount of information. On the other hand, the uneven size of the quantiles might be considered an advantage (Davies and Shorrocks, 1989). Evidence that the underestimation implied by aggregation is not serious can be found in Table B.1 of Appendix B, using the next chapter's data set. Using the SSB's eight quantiles results in a Gini approximately 97% of the size of the 'disaggregated Gini'.

Of greater concern than the degree is the manner of aggregation. Prior to 1985, quantile means were calculated on the basis of the raw data from sub-samples of the total sample surveyed.[25] From 1985 onwards, however, the aggregated means published are not based on a complete ranking of incomes and then division into groups, but on a two-stage procedure. At the first stage, the SSB asks its officers in each city and town whose residents are being sampled to compute quantile averages for that city or town, and submit them to the central office. At the second, the, say, kth quantile average for all China is calculated by taking the average of the kth quantile of each city and town. As Appendix B shows, both stages of this aggregation procedure cause the true level of inequality to be underestimated. The implications of the second stage are particularly serious. Since this second stage aggregates city and town quantile means into the same quantile group disregarding possible differences in mean value, one would expect a decomposable inequality measure resulting from such a procedure to be similar in size, at least once allowance has been made for aggregation, to the within-city-and-town inequality component of the same decomposable index. Appendix B shows the precise conditions under which this will be true. It also uses the provincial data sets to be analyzed in the next chapter to provide evidence that, whereas the first stage of this aggregation procedure only results in underestimation of the Gini coefficient (compared to that obtained using the conventional method of aggregation) of 1-2%, the second stage results in underestimation of some 25%. With such a large discrepancy, the need to guard against being misled by this aspect of the data is pressing, and the subject is returned to in V.2.

Apparently aware of its survey's bias towards city-dwellers, the SSB has based its published figures from 1986 onwards on a re-weighting. If we decompose the published urban means into a weighted average of the city and town means where the weights are, respectively, α and $1-\alpha$, then in 1985, 1986 and 1988, α is exactly equal to .6, in 1987, α=.656 and in 1989, α=.593. These correspond to city-town population ratios of 1.50, 1.90 and 1.46. Comparing these ratios to those given in Table 2, this re-weighting is appropriate if one's aim is to replicate the official city-town population ratio. However, if the aim is to replicate the city-town ratio of the survey base, then the town-resident observations are given too much weight, as one should be

25. Of size 2350 for 1981 to 1983, and 4420 for 1984.

aiming for a ratio of 2-2.5 rather than 1.5-2. So it must be borne in mind that, although the survey suffers from an under-representation of town-dwellers, the published figures analyzed here suffer from the reverse problem, at least if one takes the survey base as the population of interest.[26]

IV The level of inequality in urban China: some international comparisons

Although, as argued below, urban inequality in China has increased over the eighties, it remains comparatively low. To make this point without access to disaggregated data and also to show in what way the urban Chinese distribution differs from those of other countries, this section provides some international comparisons. 1989 data are used for China, this being the year, it is argued in the next section, in which urban income was most unequally distributed for the decade. The Chinese income distribution analyzed in this section is that of households by household income.[27] This is not because of an intrinsic preference for the analysis of household over per capita income, but for the more pragmatic reason that these figures, though only available for the latter half of the eighties, are, unlike the quantile means of per capita income, properly aggregated (see III.3). For this reason it cannot be argued of them, as it can of the per capita quantile means, that they under-estimate the degree of income inequality, and thus unfairly bias any comparison in China's favour.[28] Interestingly, the household cash income shares used are very close to the per-capita cash-and-non-cash income shares calculated by KGRZ.[29] So the comparisons which follow also do not depend on the exclusion of non-cash income from the

---

26. It is not possible to re-re-weight, as a city-town decomposition is not available for 1990.

27. See Appendix A, point 1 for details.

28. Also, in China there is at least no less and possibly more inequality in the distribution of households by household income than in the distribution of households by per capita income or in the distribution of individuals by per capita income: see the discussion in the text preceding footnote 44. See Anand (1983) for a general analysis of the relationship between inequality in these different distributions.

29.

Decile	Household income (SSB data) 1989	KGRZ income per capita 1988
1	4.4	5
2	6.2	6
3	7.2	7
4	8.0	8
5	8.8	9
6	9.6	9
7	10.5	10
8	11.8	11
9	13.7	14
10	19.9	21

Chinese figures.

International comparisons of inequality are notoriously unreliable. The problem-causing measurement issues are not addressed here at all. The justification for this is in the results: China emerges so clearly as an outlier that one cannot imagine its status as such being vulnerable to changes in the definitions used for measurement.

That inequality is low in China, and in particular urban China, is well-known. China is also unusual amongst developing countries in that urban is lower than rural inequality (see VI.2). In other developing countries, urban appears to be higher or at least no lower than rural inequality (see Zhao, 1990b; India is one example where urban inequality is higher - see Hussain, Lanjouw and Stern, 1991).[30]

Income in urban China is so much more equally distributed than income in urban regions in other developing countries that comparisons with industrialized countries are more illuminating.[31] We begin with a comparison between urban China and the UK. There is no doubt that income is distributed more equally in urban China. More interestingly, we can see from Table 6 in what way the two distributions diverge. This table is based on a division into quintiles. Take the two bottom quintiles to represent the poor, the two middle quintiles the middle-class and the top quintile the rich. The table shows that the middle-class of urban China and the UK get almost exactly the same share of income: 40.7 and 40.9 respectively. Where the two distributions differ is in the shares of income going to the rich and poor groups. Urban China's poor get approximately 9% more of total income than the UK's poor, and urban China's rich get approximately 9% less. The middle-class, as defined above, does not do unusually well in urban China. Rather it is the poor who benefit, at the expense of the rich.

30. Note the Indian data on which this comparison is based relates to consumption, which we would expect to be more equally distributed than income. Yet income in urban China is much more equally distributed than consumption in urban India. Riskin (1987, p.249) has a number of such comparisons with developing countries.

31. Given the small size of the agricultural sector in industrialized countries, one is justified in using national rather than urban data for these countries.

Table 6 A Comparison Between Urban China and the UK

Quintile	Shares of disposable income, UK 1987	Shares of disposable income in UK after 20% tax on top 20% distributed equally among bottom 40%	Shares of disposable income, China, 1989
0-20	6.07	10.30	10.61
20-40	10.74	14.97	15.16
0-40	**16.80**	**25.27**	**25.77**
40-60	16.82	16.82	18.36
60-80	24.05	24.05	22.31
40-80	**40.87**	**40.87**	**40.67**
80-100	**42.31**	**33.86**	**33.56**

Notes: UK data based on disposable income per household (after benefits and direct taxes), and is from Central Statistical Office (1990), Appendix 4, Table 1. For details on the Chinese data, see Appendix A, point 1.

One easy way to represent the difference between urban China and the UK is to ask what sort of additional taxes would be required (on the assumption of a constant pre-additional-tax distribution) in the UK to obtain from its distribution the urban Chinese distribution. As the table shows, a very simple change to the tax structure is required. An additional 20% tax on the richest 20%, distributed equally among the bottom 40% (10% to the bottom 20% and 10% to the 20-40% group) would approximately transform the UK into the urban Chinese distribution.

A wider set of comparisons between China and other countries, in terms of the three-class aggregation developed above, is presented in Figures 2a to 2c. These plot, for 41 developed countries (DCs) and less-developed countries (LDCs) for which information is available, on the vertical axis, the shares of, respectively, the poorest 40%, middle 40% and top 20% (from World Bank, 1991) and, on the horizontal axis, the log (to the base 10) of GDP per capita using Summers and Heston (1988) data based on purchasing power parity conversion factors.[32] Because of the uncertainty as to GDP per capita for China, and *a fortiori* for urban China, the latter's income shares are represented as horizontal lines (their exact values can be read off from Table 6 as, once again, 1989 data has been used). The figures show that whatever its GDP per capita, urban China is an outlier with respect to the shares of the bottom 40% and top 20%. Its share going to the middle 40% is also relatively high by developing country standards. However, if, as seems more

32. All countries which had both income shares and GDP per capita data were used. The GDP per capita figures refer to the year 1980 and are expressed in 1980 US dollars. The income shares refer to various years from the eighties. The figures also show, simply for interest, regression lines: the predicted income share from a quadratic regression of the 41 income shares on the log of GDP per capita.

reasonable, urban China has an income per capita closer to developed country standards,[33] then its share going to the middle 40% seems average.

If urban China is an outlier in relation to both developing and (except for the share of income going to the middle 40%) developed countries, how does it compare to other (ex-) centrally-planned economies? A great deal of diversity is evident among these nations (see Phelps-Brown, 1988, and Atkinson and Micklewright, 1992), but it is striking that Hungary is the one country in Figure 2 which has a higher share for the poorest 40% and lower share for the richest 20% than urban China. Figure 3 compares urban China with Hungary using a 'Pen Parade', which simply shows the mean income of different quantile groups, appropriately normalized to make comparisons possible (here the normalization has been by dividing through by the mean and multiplying by 100).[34] It is striking how closely the two distributions resemble one another.[35] Both have high shares going to the poorest 40% (26.1 in Hungary to 25.8 in China), low shares going to the richest 20% (32.4 to 33.6) and average shares going to the middle 40% (41.5 to 40.7).

One force suppressing inequality in centrally-planned economies is no doubt the confiscation of private property associated with the introduction of revolutionary regimes (see the quotation from Perkins in II.1). The wage-setting powers of the central authorities in centrally-planned economies are also relevant. If, as the evidence suggests (see Phelps Brown, 1988, pp.303-4 and Riskin, 1987, p. 251), these powers were used to raise the wage of unskilled labour and to keep down the top ranks of the salary ladder, a compressed wage structure could emerge, which would raise the share of the poor, lower that of the rich and leave the middle class pretty much unaffected in distributional terms.[36]

33. For example, as stated in II.2, 76 and 36% of all urban households in China owned a washing machine and fridge respectively. In the UK, the figures (for 1984) are 79 and 94% (Central Statistical Office, 1986).

34. See Pen (1971). Note that the Pen Parade is the inverse of the distribution function. Its usefulness as a tool of analysis has recently been demonstrated by Phelps Brown (1988).

35. Atkinson and Micklewright (1992) show that Hungary's income distribution has changed over the years, first falling then increasing, but the changes have not been drastic enough to make the similarity between China and Hungary dependent on the year of comparison. Atkinson and Micklewright also find income to be distributed in Czechoslovakia (which is not included in Figure 2) just as equally as in Hungary.

36. Howe (1973) gives a detailed account of Chinese wage patterns up to 1972. Knight and Song (1990) provide a more recent study, with cross-sectional evidence relating wage levels to education, occupation, gender and region.

Figure 2 Proportion of Total Income Received by Different Quantiles, Urban China and Selected Developing and Developed Countries

Figure 4a Poorest 40%

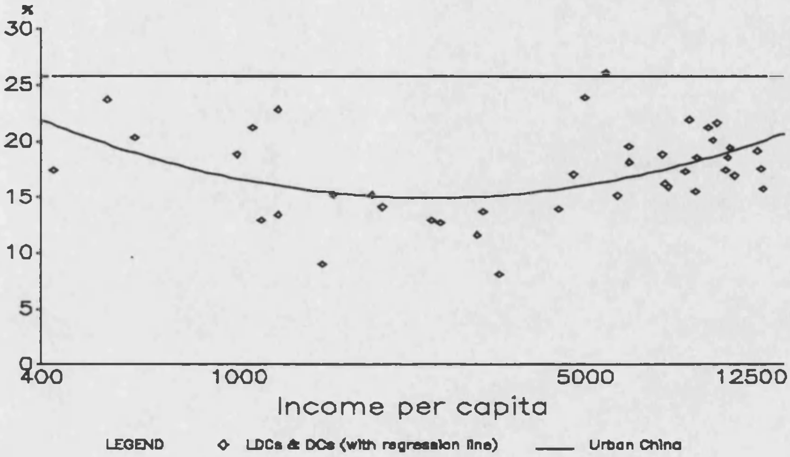


Figure 4b Middle 40%

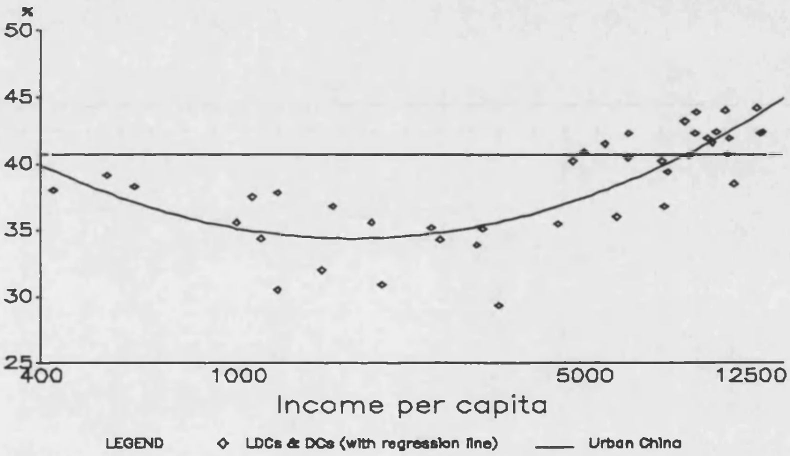
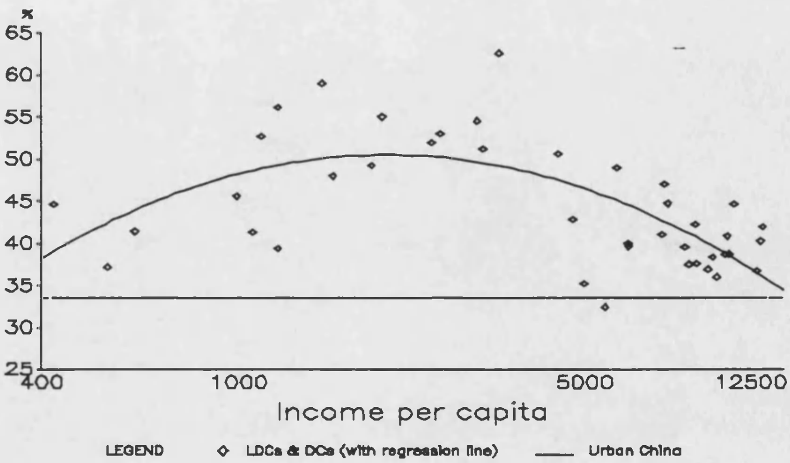
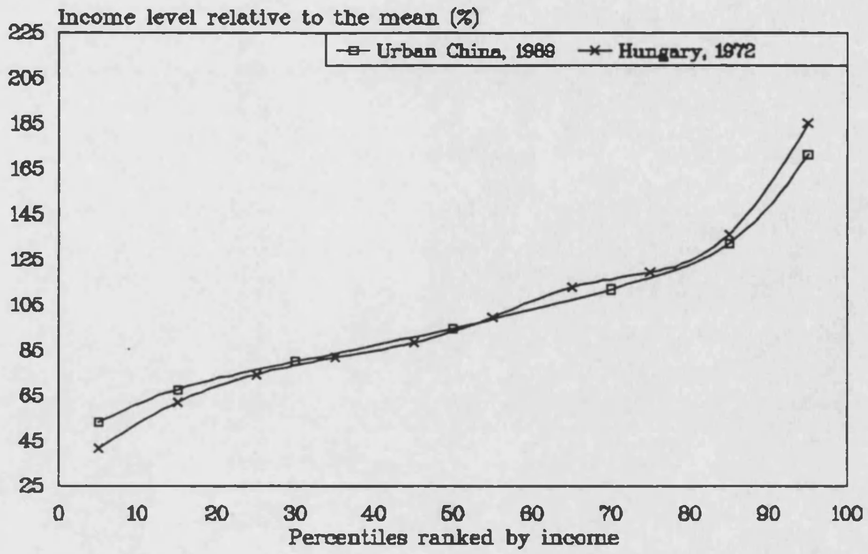


Figure 4c Top 20%



Notes: The horizontal axis is scaled in logs. For sources, see footnote 32 for international data and Appendix A, point 1 for Chinese data.

Figure 3 Income Distribution in Urban China and Hungary



Notes: The Hungarian data are from Szakolczai (1980), as extracted in Phelps Brown (1988). They are similar to the World Bank data used in Figure 2, but less aggregated. For details on the Chinese data, see Appendix A, point 1.

Samuelson and Nordhaus have written of the Soviet Union and Eastern Europe:

> Recent estimates of income distribution indicate that, except for the absence of a super-rich class, [their] income distribution ... shows a striking similarity to that in Western countries (1989, extracted in Atkinson and Micklewright, 1992, p.28)

Whether or not this is actually true for the countries Samuelson and Nordhaus are writing about - and the example of Hungary and the analysis of Atkinson and Micklewright suggests, in fact, that no such generalization can be made - an absence of the super-rich is certainly not the only driving force behind urban China's low inequality. It is also, using this terminology, the absence of a super-poor.

Besides giving us a way of stylizing China's income distribution, these international comparisons are also useful because they show that, although urban China's income inequality figures are low, they are not unbelievably or absurdly low. It is not unreasonable to think of urban China as having developed-country characteristics. And, given this, it is also not unreasonable to expect that a socialist government would be able to 'extract' 20% more from the rich for distribution among the poor than a capitalist government of a developed country would. On the other hand, the differences observed between income distribution in urban China and other countries are too great to be explained away by the argument that the official data underestimate the former's level of inequality.

V Changes in the income distribution over the decade

V.1 analyzes inequality changes evident in the core data (presented in the Appendix in Tables A.3(a) and (b)). V.2 takes account of the possible ways in which the core data may mislead and considers supplementary data to obtain a more accurate picture.

Due to the unusual aggregation procedure described in III.3, it is unclear whether the published per capita means are averaged over individuals or over households. It is indicated to be the former (*pingjun mei ren*) - the average means are obtained from the quantile means by using as weights the number of households times the quantile's average household size (see Appendix A, point 6) - but one cannot be sure that weighting by household size was carried out at both stages of the aggregation procedure. Since the size of the quantile groups is defined in terms of households, not individuals, it is certainly simpler to assume that the income figures are averages over households. This makes year-to-year comparisons of income shares possible without the need for adjustment on account of changing household size. Calculations were also made weighting each quantile group by its average household size but this made a negligible difference (always less than .001 in the case of the Gini for example). Thus, unless otherwise specified, all results (including the means already given in Table 1) are based on the assumption that the quantile means are household per capita means and so should be thought of as summarizing the distribution of households by per capita income.

No attempt is made in the analysis which follows to reconstruct the original data by use of interpolation methods (such as given in Cowell and Mehta, 1982). Although such methods add accuracy, since a number of other adjustments to the data are already required (see Appendix A for a summary), it was decided to use instead the simpler and more transparent method of assuming complete equality within each quantile, thus giving a lower bound on the true level of sample inequality.[37] While this method will, by definition, underestimate the actual level of inequality, it will not be by very much (see III.3) and there is no reason to believe that trends as

37. Interpolation is used at various points in the chapter, but only when the number of income classes differs across years and comparability is sought. In these cases (indicated in the note to the relevant tables (9 and 14) or in Appendix A, point 1 for the case of Figure 2), a cubic spline is used to estimate the desired quantile group or income class means. Any subsequent estimation of inequality measures is then based simply on these new means.

against levels will be underestimated by this use of aggregated data. Trends will be accurately estimated if intra-quantile reflect inter-quantile distributional changes. This is unlikely to be true for any single pair of years, but is more likely to be true on average which makes it a satisfactory assumption if, as here, one has data for a number of years. Having a series of data points also makes it less of a drawback that, without information on sample variance, standard errors for estimates of inequality derived on the basis of aggregated data cannot be calculated. With a decade's data, one should be able to establish, without recourse to statistical methods, whether inequality is systematically changing.

V.1 Changes in inequality: the core data

Due to the vast number of competing inequality indices available for use, reliance on any single measure is unlikely to be convincing. On the other hand, the presentation of a large number of indices, while a step in the right direction, may easily lead to difficulties of interpretation. Use is therefore made of the criteria presented in Part One of the thesis which, if met, guarantee dominance by a range of measures. Foremost among these is the second-order stochastic dominance criterion (or Lorenz dominance criterion in the context of inequality analysis). If one distribution's Lorenz curve is everywhere no lower than another's, and somewhere higher, then the former distribution has no more inequality by all measures and less inequality by at least one measure satisfying the conventional assumptions given in II.2. Also utilized is the e-dominance criterion, which measures equality using the isoelastic function, defined by the parameter e, and looks for whether all indices for a range of values of e give the same ranking.

Table 7 presents the basic information pertaining to equality rankings for the decade of the household income per capita distributions. Years are presented in chronological order. A 'D' ('DB') in cell ij indicates that row i (column j) has second-order stochastic dominance over column j (row i). 'ED' and 'EDB' have the same respective meanings but are used only when stochastic dominance is absent and refer to e-dominance. (Hence a 'D' or 'ED' indicates the row year has less inequality.) An X indicates that no e-dominance, and, *a fortiori*, no stochastic dominance ranking can be made. In these cases the 'switch points', or values of e at which dominance reversals occur, are given. The upper and lower bounds chosen for the e-dominance analysis were taken to be 5 and .1 respectively, which are quite wide relative to those commonly used in inequality analysis.

Table 7 Equality Dominance, Urban China, 1981-1990

	1983	1984	1985	1986	1987	1988	1989	1990
1981	DB	D	X 3.2	D	D	D	D	D
1983		D	D	D	D	D	D	D
1984			DB	ED	X 4.1	D	D	D
1985				D	D	D	D	D
1986					X 2.7	D	D	D
1987						D	D	D
1988							D	D
1989								DB

Notes: 1. 'D' indicates that the row year dominates the column year. 'DB' indicates that the row year is dominated by the column year. A 'D' or 'DB' by itself indicates Lorenz dominance (and therefore e-dominance), a 'D' or 'DB' with an 'E' in front of it indicates e-dominance, but no Lorenz dominance. If a number is also given, this indicates the upper bound up to which e-dominance is present. If no number is given, then e-dominance holds over the entire range.

2. The bounds for the e-dominance analysis are:
upper bound: 5
lower bound: .1

3. Based on data in Tables A.3(a) and (b). See the notes to these tables for further details.

The e-dominance ranking is summarized in Figure 4, a Hesse diagram, in which any one year (i's) distribution dominates the distribution of any other year below it on the page (a) to which it is connected by any combination of downwards and horizontal lines or (b) which are dominated by any other province which i dominates.

Figure 4 Hesse Diagram for Equality E-Dominance, with bounds of .1 and 5

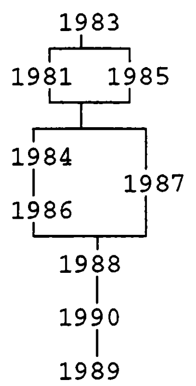


Table 7 and Figure 4 have two striking features. The first is a fall in inequality between 1981 and 1983, and a general though not monotonic trend upwards in inequality thereafter. Inequality at the end of the decade is higher than at the start. Here is our first indication that trends from the early eighties should not be extrapolated over the course of the decade. The second is the completeness of the ordering. Of the 36 pairs, 32 can be ranked by the stochastic dominance criterion, and a further 1 by the e-dominance criterion, leaving only 3 pairs which cannot be ranked at all. As Chapter Three (IV.1) showed, the chances of obtaining a ranking increase with the degree of aggregation of the data. Therefore, we would not expect such a high degree of ranking if disaggregated data were available. Also on the basis of Chapter Three, we should interpret the rankings based on aggregated data as providing a basis for inferring not unrestricted Lorenz dominance in relation to the disaggregated data but rather dominance above the poorest 5% (see IV.1.1). The aggregated data will tend to ignore any crossings below the first point of aggregation, here 5%. Similarly, we should interpret e-dominance based on aggregated data as enabling us to infer e-dominance in relation to the disaggregated data based on a lower upper bound, in this case, say, 3 instead of 5. These caveats notwithstanding, it can still hardly be coincidental that an upward trend emerges post-1983.

By how much did inequality worsen over the eighties? Table 8 presents a number of conventional inequality measures. Together they present an even stronger picture of the trend in inequality than the more general partial ordering of Table 7. The measures all agree that inequality fell between 1981 and 1983, rose in 1984 (to above its 1981 level), fell in 1985, rose to 1989 (to its highest level for the decade) and then fell in 1990 (to about its 1988 level). They show a growth in inequality of between 9 and 19 percent over the decade, depending on which measure is used. The post-1983 growth is of course higher, between 12 and 26%. Table 8 also gives a good idea of over which years inequality grew more quickly. All measures agree there was a relatively big jump between 1983 and 1984, not much change between 1984 and 1987, and then large jumps again between 1987 and 1989.

Table 8 Inequality in the Distribution of Households by Per Capita Income

Year	Gini	Coefficient of Variation	Theil Index	Atkinson Index (e=1)	Atkinson Index (e=2)
			Inequality Measure		
1981	0.1627	0.2952	0.0425	0.0422	0.0831
1983	0.1577	0.2882	0.0403	0.0397	0.0780
1984	0.1660	0.3026	0.0445	0.0439	0.0861
1985	0.1636	0.2986	0.0432	0.0426	0.0834
1986	0.1660	0.3031	0.0446	0.0440	0.0866
1987	0.1671	0.3061	0.0452	0.0444	0.0868
1988	0.1753	0.3210	0.0497	0.0488	0.0953
1989	0.1807	0.3327	0.0531	0.0519	0.1008
1990	0.1769	0.3237	0.0506	0.0498	0.0973
Changes over time					
90/81	1.087	1.097	1.190	1.180	1.171
90/85	1.081	1.084	1.171	1.169	1.166
85/81	1.006	1.012	1.016	1.010	1.005
90/83	1.121	1.123	1.255	1.252	1.247
90/87	1.059	1.058	1.118	1.122	1.121

Notes: Based on data in Table A.3(a). The Atkinson indices are monotonically decreasing transforms of the isoelastic function. The Theil (also called the Theil T) measure is also the generalized entropy measure with parameter $\delta=1$; similarly the square of the coefficient of variation is twice the generalized entropy measure with $\delta=2$. For more detail on the formulae, see Table 2 of Chapter One. To calculate Ginis with different size quantiles, a formula given in Cowell (-) incorporating weights was used, namely

$$\text{GINI} = \sum_{i=1}^{N} W_i \frac{y_i}{\mu} \text{ where } W_i = w_i\left(\sum_{j=1}^{i} 2w_j - w_i - 1\right) \text{ and } \sum_{i=1}^{N} w_i = 1$$

where w_i is the weight on the ith quantile (.05 or .1) and incomes are arranged in ascending order.

V.2 Do the core data mismeasure the rise in inequality?

V.2.1 Structural breaks in 1985 and 1986

As outlined in Section III, there are three changes which take place in the calculation of the published figures in 1985 and 1986. In 1985, the aggregation procedure changes and the sample base widens to include households with heads not employed in either the SOE or collective sectors. Then in 1986, the sample base further widens to include town-dwellers. To make the problem tractable, I assume that these changes all change inequality, as measured by the Gini, by an amount which remains constant over time.

The first and most important change in 1985 is the change in the aggregation procedure. Direct evidence is available in this regard. Prior to 1986, means are given for conventionally aggregated income classes as well as for quantile groups.[38] Gini and Theil indices using both data sets are presented in Table 9. As expected, the two sets of figures are very close (within 5% of each other) for 1981, 1983 and 1984. The 1985 income class figure, however, is 22.8% above the quantile group figure using the Gini and 52.7% using the Theil. Far from inequality falling between 1984 and 1985, as the analysis of the previous sub-section suggests, it increases by the largest single amount for the decade.[39] Whatever the cause of this rise, Table 9 shows that the smoothing procedure underestimates inequality by around 20%, resulting in Ginis of .16 to .18 rather than of .19 to .21. Taking this approach a step further, one crude way to correct for the underestimation introduced by this switch to smoothing would be to add the difference between the income-class-based and the quantile-based measures, which comes to .037. Allowing for the fact that the latter set of measures tend to be very slightly below the former for the years prior to 1985 gives an adjustment factor of .034.[40]

Table 9 Inequality in the Household Distribution by Per Capita Income: measures based on quantile group and income class data, 1981 to 1985.

	Gini Coefficient		Theil Index	
	Quantile group	Income class	Quantile group	Income class
1981	.1627	.1615	0.0425	.0420
1982	-	.1582	-	.0401
1983	.1577	.1581	0.0403	.0400
1984	.1660	.1697	0.0445	.0465
1985	.1636	.2009	0.0432	.0660

Notes: The first and third columns are from Table 8. The second and fourth are based on data in SSB(1987s) - see Table A.1. There are 6 classes for 1981 to 1984 and 10 classes for 1985. To avoid this increased number of classes in 1985 influencing the result, and to give comparability with the figures based on quantile groups, a cubic spline was used for all five years, enabling estimation of income shares based on the eight quantile groups used in this chapter.

How much of the increase between 1984 and 1985, once this adjustment is taken into account, is due to the change in 1985 to the sample base? Here no direct evidence is available, but it would seem the answer is not very much. The change increased the sample base by about 10%. Nearly all the new households must have had a pensioner as head (since the number of self-

38. Post 1985, though income-class proportions are still given (see Table A.1), no class means are given. Due to the difficulty of estimating the class means, this information is not used.

39. Inspection of Lorenz curves constructed using this income class data shows Lorenz dominance of 1984 (and hence 1981 and 1983) over 1985.

40. The adjustment is done by multiplying the 1985 quantile figure by the proportional difference between the 1984 income-class and quantile measures.

employed and private workers in the survey is still very low - see Table 2) and pensioners are not particularly poor in urban China. The Liaoning and Sichuan data sets suggest that this change in the sample base would have raised the Gini by as little as .005.[41]

Direct evidence is available on the effect of the final change to the sample base - the inclusion of county towns. For 1986 to 1989, city and urban inequality indices can be calculated separately. These are presented in Table 10, along with indices for town residents only for 1987 to 1989. Inequality is greater in towns and towns are poorer (see the final column) so urban inequality is higher than city inequality. But again the difference is not a huge one. The average difference over the four years in terms of the Gini is .007.

Table 10 Urban, City and Town Inequality and Relative Means, 1986-1990

	Gini				Theil				City/
Year	Cities	Urban	Town	Diff	Cities	Urban	Town	Diff	town (%)
1986	0.1594	0.1660		0.0066	0.0411	0.0446		0.0035	
1987	0.1645	0.1671	0.1726	0.0026	0.0439	0.0452	0.0483	0.0013	124.1
1988	0.1626	0.1753	0.1836	0.0127	0.0440	0.0497	0.5407	0.0057	124.1
1989	0.1744	0.1807	0.1954	0.0063	0.0493	0.0531	0.0634	0.0038	123.2
Average difference				0.0071				0.0036	

Notes: 'Diff' stands for the difference between the city and urban Ginis. The urban Ginis are from Table 8. The city and town Ginis give the ratio of city to town mean incomes, expressed as a percentage.

So these rough calculations suggest that, assuming these changes have a constant additive effect on inequality, 1985 can be made comparable to the earlier years by increasing the Gini in Table 8 by .029 (.034-.005) and 1986 and later years can be made comparable by increasing their Ginis by .022 (.029-.005). This results in a 1990 comparable Gini of .199, and growth between 1993 and 1990 not of 12%, as concluded using unadjusted data, but of 26%. The adjustments also suggest that inequality actually fell between 1985 and 1986 by .006, similar in magnitude to the fall between 1989 and 1990 (.004). As will be argued later, these falls are as

41. Calculating Ginis using the pre-1985 sample base and the 1985 sample base results in an average difference of .005 for the two provinces combined and the three years of 1988 to 1990. This is the difference observed if the SSB aggregation procedure is used, but no larger difference emerges if alternative, more conventional aggregation procedures are used instead. An alternative estimate of the effect of expanding the share of pension income can be attained by calculating the total Gini on the assumption that pension income is distributed identically to total income, while the distribution of other components of income remains unchanged. This reduces the Gini by a maximum of .006: see VI.2 for a decomposition by source.

significant as the rises in other years in understanding the forces acting on inequality in the eighties (see VII.3).

There is one more change which occurs in 1985 which may have caused the observed rise in inequality. The number of city-residents sampled increases from 12,000 to 17,000 and the number of cities surveyed increases from 47 to 106 (SSB, 1988se). In and of itself, a larger sample, even one covering a larger number of cities, should not lead to an increase in inequality. However, the geographical extension of the survey may have led to the inclusion of a more disparate range of cities. Certainly we find this happening in Liaoning and Sichuan in the next chapter: recorded inequality rises most sharply in 1988, the year in which the number of cities and towns covered by the SSB sub-samples (though not the full samples) increases sharply (see Chapter Five, III.2 for details). This effect cannot be quantified, but should be borne in mind.[42]

Note finally that, since for 1981, 1983 and 1984 the income-class based and quantile based measures are very close (see Table 9), we can take the 1982 income class indices as good estimates of the quantile-based indices for that year (for which the data are missing). If so, we can conclude that inequality falls slightly between 1981 and 1982 and is basically unchanged between 1982 and 1983.

V.2.2 Post-1985 trends

Since the aggregation procedure used from 1985 onwards captures only 3/4 to 4/5 of inequality as measured using conventional means of aggregation, trends based on it could easily be misleading. One check on this can be conducted by comparing the results of Table 8 with those available using quantile total household income data available from the same surveys (for 1987 to 1990 only), but based on standard aggregation procedures (for a sub-sample of the survey sample).[43] Results using this source (already utilized in the previous section for the purpose of making international comparisons) are given in Table 11. All the inequality indices confirm the upward trend in inequality in the latter half of the eighties. In addition, all the indices except the Coefficient of Variation suggest a higher growth rate of inequality than is observed using the

42. In fact, the sample size increases each year after 1983 (see Table 2), but the increase in 1985 is the most dramatic, which may help explain the large rise in the Gini once the appropriate adjustments have been made.

43. The sub-samples are about one-third the size of the full sample: 13,300, 14,000, 13,800 and 13,494 from 1987 to 1990.

smoothed data over the years 1987-1990. However, the 1987-1989 growth rates for the two sets of figures are much closer and in some cases the household per capita figures of Table 8 are slightly higher.

Of course it is possible that total income per household and per capita income per household inequality should differ in both magnitude and trend. However, again the provincial data sets analyzed in the next chapter come in handy: they show that at least for the years 1987-90, the Gini based on the distribution of households by household income is always greater than, by on average .019, that based on the distribution of households by per capita income which in turn exceeds, by on average .012, that based on the distribution of individuals by per capita income. All three Ginis move in tandem.[44]

Table 11 Inequality in the Household Distribution by Household Income in Urban China, 1987-1990

Year	Gini	Coefficient of Variation	Theil Index	Atkinson Index (e=1)	Atkinson Index(e=2)
1987	0.2068	0.3911	0.0714	0.0696	0.1375
1988	0.2157	0.4074	0.0772	0.0749	0.1456
1989	0.2240	0.4212	0.0827	0.0799	0.1547
1990	0.2268	0.4426	0.0875	0.0828	0.1594
1990 comp	0.2233	0.4139	0.0814	0.0800	0.1572
90/87	1.080	1.058	1.139	1.150	1.143

Notes: 'comp' is short for comparable. The first set of 1990 figures are calculated directly from the income classes given in the relevant yearbook. The '1990 comp' figures are based on the same number of income classes as the preceding two years, and thus are comparable. The '90/87' ratios use the '1990 comp' figures. See Appendix A, point 1 for more detail on methods and source. See the notes to Table 8 for details on the measures used.

Although the SSB aggregation procedure leads to underestimation of total inequality, Appendix B provides evidence that it provides an reasonably accurate estimate of average within-city-and-town inequality. Trends in within-city-and-town inequality will in turn provide a good proxy for trends in overall inequality only if inequality between cities and towns moves in the

44. Knight and Song (1990), using the CASS Institute of Economics 1986 urban household survey find very similar Ginis for the distribution of households by household income (.248) and by per capita income (.244). Note that even if these two are exactly the same in urban China, the figures in Tables 8 and 11 are not comparable due to the SSB aggregation procedure and re-weighting both of which affect the Table 8 figures, but not the Table 11.

same direction. Unfortunately, published data for individual cities and towns are unavailable, so we can only go on trends in inter-provincial inequality and in city-town inequality. Section VII.1 shows that there has been a rapid increase in inter-provincial inequality, which, using the Theil index, more than doubles over the decade. City-town inequality, on the other hand, seems to have been basically unchanged. Table 10 shows the city-town mean ratio to have been constant at 124% from 1987 to 1989. The value for 1985 is higher at 128% but only slightly. Inter-provincial inequality is more important than city-town inequality (the Theil for the former in 1987, for example, is .0127 (see Table 15), that for the latter is .0054 (calculated from Table 10)). Hence it may be that within-city-and-town inequality growth may underestimate total inequality growth, since it ignores the increase in inter-provincial disparities. Certainly if it overestimates, it cannot be by very much.

V.2.3 Income-source and sample-base exclusions

It is widely believed that 'black' and 'grey' incomes (both cash and in-kind) have increased in importance over the decade. It is likely that such incomes are under-reported, and that they are more unequally distributed than income from legal activities. Almost certainly then, these omissions lead to an underestimation of not only the level but also the growth in urban inequality. By how much one can only guess.

The importance of in-kind subsidies, excluded by the SSB calculations, has already been seen in the comparisons of SSB with KGRZ mean income (III.2). With regards to their effect on the level of inequality, KGRZ find that the housing subsidy pushes up inequality, while the subsidies attached to low food prices lower it. The overall effect of subsidies is to increase urban inequality, though only slightly. With regards to their effects on changes in inequality, it is helpful to distinguish between those subsidies provided by enterprises (such as housing) and those provided by the state (such as the food subsidies).

Concerning the former, as mentioned earlier (II.2), the observed increase in inequality in received income may simply be the result of some firms distributing productivity gains and rents in the form of wages, others in the form of better housing, health care or education. If this is the case, the recorded rise of inequality over- rather than under-estimates the true rise in inequality. While one cannot refute this as an explanation - certainly investment in housing has grown in both absolute and proportional terms over the eighties (see Hussain, 1990, p.46)[45] - the large growth in interprovincial inequality (to be demonstrated in VI.1) suggests that at most it cannot be the whole story. To maintain this explanation in the face of a doubling of interprovincial inequality one would need to argue that firms handed out extra wages in the form of cash or goods depending on which province they were in. This is unlikely, though direct evidence is unavailable.

The next chapter examines the evolution of food subsidies in the late eighties and shows that their growth has had a dampening effect on the increase in urban cash-plus-subsidy income inequality. In the last couple of years and on various occasions through the eighties, however, government policy has been to replace these implicit food subsidies by cash payments. The evidence of the next chapter provides evidence that, due to the lump-sum way in which these payments are distributed, this policy has basically left unchanged cash-plus-subsidy income inequality, but has reduced cash income inequality.

The possibility must at least be raised that the apparent increase in inequality is due to improved coverage of income sources by the SSB, either because the SSB has become more skilled at unearthing non-obvious sources of income or because less stigma is now attached to 'getting ahead', leading to a more honest revelation of lucrative, previously-hidden income sources. Certainly, the SSB has over the years moved towards adopting a finer categorization of income sources, with changes in both 1985 and 1988. On the other hand, such changes may well be in response to an increased heterogeneity of income sources in China, and hence not so much a source of bias as a reflection of reality.

As explained in III.1, the data set excludes the agricultural-registered but urban-dwelling (the 'floating' population), of which there is certainly a growing number. One would expect this sector to be on average poorer than the SSB survey base. There is also probably greater inequality within the floating than non-floating population (as only some migrants will 'make it'). So incorporating the floating population would tend to increase inequality. If the main change in the

45. And not only housing. Hua, Zhang and Luo are revealing on this point:

> While managerial powers were being extended, corporate consumption ran wild under all sorts of pretexts. There was a widespread story that one enterprise came up with the method of giving free lunches to its workers so that the actual working time was extended and efficiency increased. This was praised by Hu Yaobang [the then General Secretary] when he inspected the enterprise, and soon became widely imitated because of the free lunch rather than the increased efficiency. This kind of practice developed into a raging tide by the end of October 1984 after the grand national celebration of the 35th anniversary of the People's Republic of China, when ... it turned out that everyone who took part in the grand celebration was given a free or heavily-discounted suit. Enterprises, districts and even government ministries then vied with each other in issuing free suits (sometimes a small token charge was made) to their workers and staff. (1992, p.110)

floating population is that their weight in the total urban population has been increasing then we would also expect their inclusion to give a faster rate of inequality growth.

V.2.4 Conclusion

Structural changes notwithstanding, there can be no doubt that using the SSB data and the (post-1985) SSB definitions of 'urban' and 'income', the trend in inequality has been upwards. This is evident even from Figure 5 and our examination of the various structural breaks suggest that their net effect is to bias downwards the rise in inequality. Claims based on data from the early eighties that inequality has not worsened in urban China are misleading. Rather 1983 should be seen as a structural break in the recent history of urban China, as the last year of declining inequality due to the wage compression in place since the late seventies and as providing the base from which inequality grew in subsequent years.

There is only one alternative explanation, namely that various unquantifiable factors have produced a spurious upwards trend. These have already been mentioned and include an increasing sample size and a more widespread disclosure of previously hidden income sources, due to a greater social tolerance of wealth and/or a greater vigilance on the part of the SSB. These are both possibilities - of the type which plague many time-series conclusions. Our conclusion should therefore be not that urban inequality has risen, but that the data are consistent with it having risen: the former can be taken as a short-hand for the latter.

Whether urban inequality has risen once we replace the SSB definitions of 'urban' and 'income' by our own ideal definitions is much more difficult to say. The foregoing has shown that the SSB definitions certainly result in important omissions, some of which probably mean the true level and growth rate of inequality are underestimated, others of which imply the opposite.

VI A wider perspective

It is one thing to observe an upward trend in inequality or even to attain a quantitative estimate of the magnitude of that trend. It is another to get a grasp on its importance. For this, a wider perspective is needed. As a start, the changes in urban inequality can be compared to those which occurred in the UK over the first half of the eighties. This is a period in which unemployment doubled, and in which inequality is generally held to have increased significantly.

The rise in the Gini for disposable per capita income is given in Central Statistical Office (1990) as 11%, which, though from a significantly higher base, is less than even than the growth in inequality given in Table 6 for post-1983 urban China (12%) and considerably less than the growth arrived at once it has been adjusted to take into account the various biases (26%).[46]

As already mentioned, however, international comparisons are often unreliable. The increase in urban inequality can also be placed in perspective by examining the consequences of the growth in inequality for urban welfare and poverty (VI.1), and by comparing urban with rural inequality growth (VI.2).

VI.1 Changes in urban welfare and poverty

To answer the question of whether the post-1983 upward trend in inequality has been large enough to offset, in terms of social welfare, the positive impact of the upward trend in mean income over the same period, use can be made of the same tools of dominance as used for the analysis of inequality in V.1. Only one slight modification of procedure is required. Since we now wish to take into account differences in mean income, we analyze income rather than mean-normalized income and thus deal with welfare rather than equality functions. To do so, we need to make incomes across years comparable. This is done using the official urban price indices given in Table 1, based on a weighted average of state and free-market prices and covering both commodities and services.

Table 12 and Figure 5 are analogous to Table 7 and Figure 4 and utilize the same core data. Table 12 indicates that 35 of the 36 pairs can be ranked with respect to welfare on the basis of second-order stochastic dominance. Only the pair 1987-1988 cannot be ranked. This pair is also unrankable using the criterion of e-dominance with bounds of 0 and 5, since 1988 dominates 1987 for $e \leq 3.3$, but not for higher values of e. The resulting ordering for both criteria is displayed in the Hesse diagram of Figure 5. This mirrors, except for the years 1987 and 1988, the ranking on the basis of mean-income, which suggests that the upward trend in inequality has not been able to offset the welfare gains flowing from an increase in mean income. What has been decisive for welfare has been changes in mean income, not changes in the distribution around that mean.

46. Jenkins (1991) gives a similar growth rate of 10% for Ginis based on per capita income after benefits but before taxes.

Table 12 Welfare Dominance, Urban China, 1981-1990

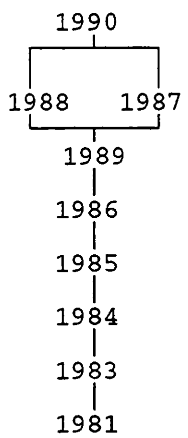
	1989	1988	1987	1986	1985	1984	1983	1981
1990	D	D	D	D	D	D	D	D
1989		DB	DB	D	D	D	D	D
1988			ED 3.3	D	D	D	D	D
1987				D	D	D	D	D
1986					D	D	D	D
1985						D	D	D
1984							D	D
1983								D

Notes: 1. 'D' indicates that the row year dominates the column year. 'DB' indicates that the row year is dominated by the column year. A 'D' or 'DB' by itself indicates second-order stochastic dominance, and therefore e-dominance. A 'D' or 'DB' with an E in front of it indicates e-dominance, but no second-order stochastic dominance. If a number is also given, this indicates the upper bound up to which e-dominance is present. If no number is given, then e-dominance holds over the entire range.

2. The bounds for the e-dominance analysis are:
   upper bound: 5
   lower bound: 0

3. Based on the Tables A.3(c) and (d).

Figure 5 Hesse Diagram for Welfare Second Order Stochastic Dominance



Note: The same ordering is obtained using the e-dominance criterion with bounds of zero and five.

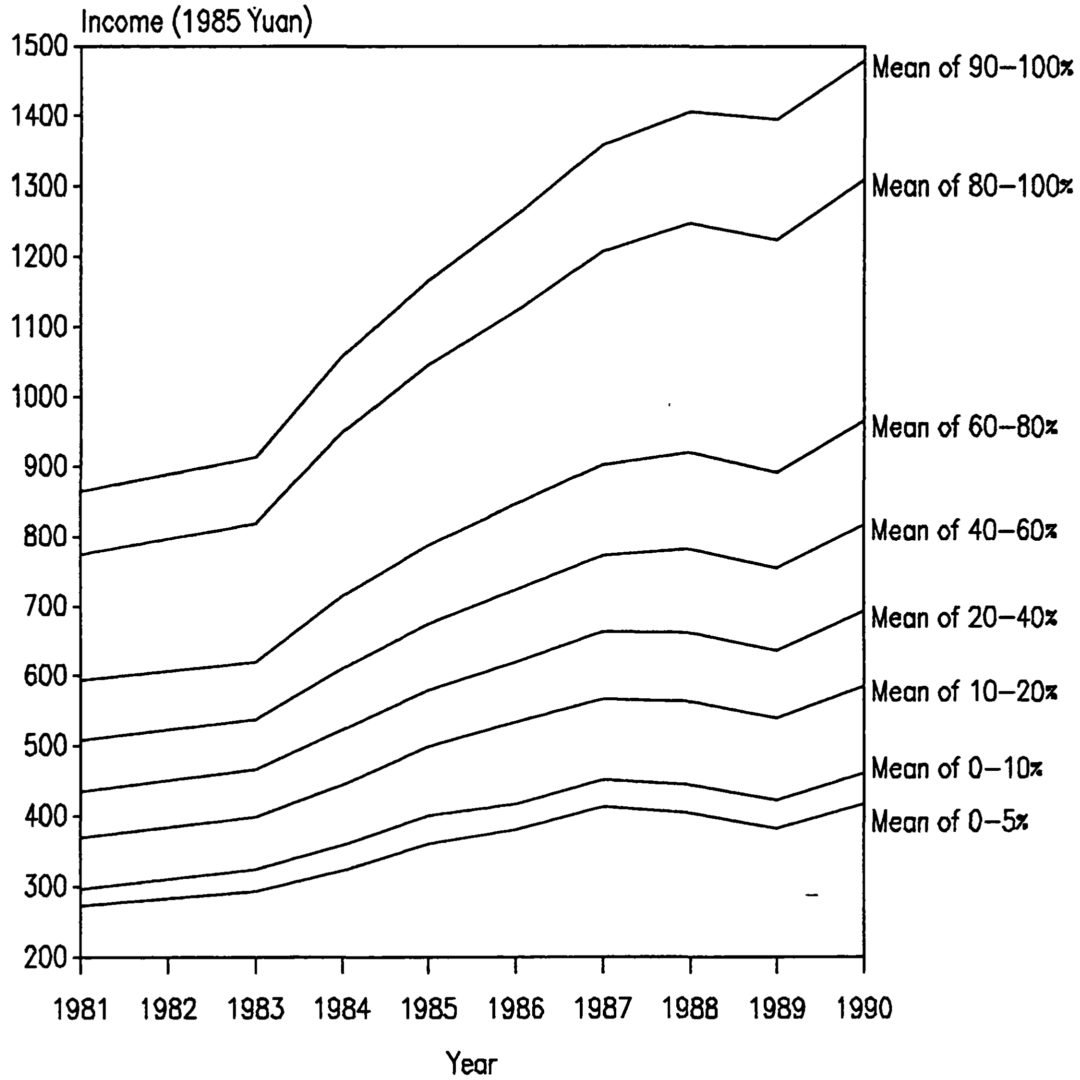
As with the equality analysis, it is necessary to caution against placing too much weight on the completeness of this ordering. Using less aggregated data would almost certainly result in a less complete ordering, and would also, from the analysis of Chapter Three, probably result in 1987 dominating 1988 from a lower value of e than the 3.3 obtained here. In addition, as the SSB aggregation procedure leads to an underestimation of inequality, and as there may be inaccuracies in the estimated growth rate of inequality, one may not be convinced by the above demonstration that the rise in inequality has not prevented a rise in overall welfare and fall in poverty. However, given the size of the increases in mean income which have occurred, it is unlikely that use of properly aggregated data would reverse this result, though it is possible that it would weaken it.

In any case, it is not as if the changes in inequality which can be measured pale into insignificance when set against those in mean income. Figure 6 gives the mean income (in constant 1985 prices) of the different quantile groups between 1981 and 1990. This graph shows that for most years all the quantile means moved in the same direction. The single exception is between 1987 and 1988. This is the one time in the eighties in urban China in which the poor became poorer and the rich richer: those in the bottom 40% saw a fall in their real income, those in the top 60% a rise. Although Figure 6 confirms the relative importance from a welfare perspective of the changes in mean income vis-a-vis those in inequality, it by no means suggests that the latter were insignificant. It is clear that a dispersion in income means is occurring. For example, if one compares the growth rate of mean income of the top 10% with that of the bottom 10% over the second half of the decade (1985-1990) one finds that the latter is more than one-and-a-half times the former, with the richest 10% growing at 5.5% per year on average, and the poorest 10% at 3.5%.

This finding can be related back to the international comparisons of Section IV. Using the tripartite division adopted therein, it is clear from Table 13 below that between 1983 and 1989, the years in which urban inequality worsened in China, the share going to the middle-class (40-80%) was virtually unchanged. The poor (0-40%) saw a decline in their share of 5%, and the rich (80-100%) saw a rise in their share of income also of 5%. Hence the income distribution of urban China has been gradually becoming less of an outlier over the past decade.

Further disaggregating these growth rates by vingtiles and deciles where possible, we see, in Table 13, confirmation that the growth rate of quantile income shares increase monotonically with average income, so that the poor and rich were the important relative losers and gainers, respectively. Data is not available for the richest decile in 1983, but the 1985 to 1989 growth rate suggests that the relative gains among the rich were restricted to the top decile. Over the period 1985 to 1989 their share of income grew by 6%, whereas that of the 80-90 decile grew by only 1%. This contradicts the finding of Ahmad and Wang (1991, p.30) that in urban areas "there is no increased concentration at the upper end of the income scale".[47]

Figure 6 Quantile means, Urban China, 1981-1990



Notes: Based on Table A.3(c). There are no figures for 1982.

47. The gap between these growth rates will underestimate the true gap, due to the introduction of the smoothing procedure.

Table 13 Growth in Income Shares, 1983-1989 and 1985-1989

Quantile	1983-1989 (%)	1985-1989 (%)
0-10	-8.42	-6.28
10-20	-4.86	-4.32
20-40	-3.85	-2.41
0-40	-4.97	-3.62
40-60	-1.32	-0.80
60-80	1.13	0.40
40-80	-0.01	-0.16
80-90		0.96
90-100		6.02
80-100	5.04	3.78

Note: Calculated from Table A.3(a).

If the income of the poor did grow, albeit more slowly than that of the rich, then, at least if one takes the 'absolute' approach of Chapter One, and works with a poverty line fixed in real terms, poverty must have fallen. Welfare implies poverty second-order stochastic dominance (see Chapter One, III.1), so at least a strong an ordering can be made covering all 'egalitarian' poverty functions, wherever the poverty line is placed. The criterion of first-order stochastic dominance can be examined to see if the distributions can also be ranked by a wider set of poverty functions including the head-count ratio. A clear result emerges from analysis on this basis of Table A.3(c): poverty falls to 1987, then rises to 1989 (rising in 1988 for any poverty line which puts 40% or less of the population in poverty) before falling to its lowest level for the decade in 1990.[48]

VI.2 A comparison with trends in rural inequality since 1978

The consensus on the respective courses of rural and urban income distribution in China over the eighties is illustrated by comments of Gale Johnson (1990), who, citing Ginis for the late

48. Head-count ratios or other measures of poverty are not calculated here due to the difficulty of calculating them using aggregated data. These difficulties are illustrated by Ahmad and Wang's (1991) utilization of data on income class sizes (published annually in the State Statistical Yearbooks) from the SSB surveys to calculate urban head-count ratios from 1985 to 1988. The poverty line they use (375 yuan in 1985 prices) falls below the upper bound of first income class given in all years except 1987, making estimation extremely hazardous. In addition, note that Ahmad and Wang's conclusion that there is "an increasing trend [in the headcount ratio] since 1985" (p.15) is not even supported by their own results. The authors estimate headcount ratios of 9.87, 5.42, 5.30 and 7.29 for 1985 to 1988. These figures show a downward trend between 1985 and 1987 and an upward 'trend' only between 1987 and 1988.

seventies and early eighties, argues the increase observed in rural inequality "should have been expected" since "one of the intended results of the reforms was to relate reward more closely to productivity. Such a result would inevitably lead to an increase in inequality within communities in the short-run and perhaps in the longer run as well." The decline in urban inequality he observes also fails to surprise "since the urban and industrial reforms did little to change the structure of compensation" (p.76). This view that inequality has increased rapidly in the countryside, but slowly, if at all, in urban areas is widespread, but misleading. It has already been shown that comparisons between the late seventies and early eighties should not be used if one wants to track the course of urban inequality over the eighties. The task of this sub-section is to compare the findings given above on urban inequality to those available on rural inequality. Its purpose is not to provide a comprehensive treatment of the data available for the analysis of rural inequality but simply to set out, and where possible briefly evaluate the reliability and consistency of, various claims in the literature concerning the path of rural inequality in China.

At least five time-series of Gini coefficients are available for post-reform rural China. These are given with their sources in Table 14. Even though all five estimates draw on the same data source, the picture they present is a bewildering one. Estimates of inequality growth between 1978 and 1986, for example, range from -3 to +25%. The WB(1) figures are based on SSB data given to the reader, though the method of calculation is not provided. The data shown does not include class means, which immediately makes the coefficients obtained less reliable. The WB(2) Ginis are presumably based on the SSB income class data below which they appear in Ahmad and Wang (1991), though this is not stated explicitly. If so, they are again based on (the same) SSB data, again with class means missing. Exactly the same comments as to source apply to the AS figures. The SSB(2) figures are also based on SSB data. They use the same income classes as WB(1) and WB(2). The population sizes diverge slightly, but in such a way (with larger lower income classes and smaller higher income classes) as to be explained by the fact that the SSB(2) figures are based on the division of individuals into classes, while the WB figures are based on the division of households. In addition, and importantly, the SSB(2) data provides class means. No information is given for the method of calculation and exact source of the SSB(1) figures, though presumably the latter is the SSB itself.

Table 14 Inequality in Rural China, Estimates of the Gini Coefficient, 1978-1986

Year	WB(1)	WB(2)	SSB(1)	SSB(2)	AS
1978		.32	.2124	.248	.222
1979	.257	.28			
1980	.237	.26	.2366		
1981	.231	.23	.2388		
1982	.225	.22	.2318		
1983		.25	.2459		.264
1984		.27	.2577	.262	
1985		.30	.2636		
1986		.31	.2800		

Sources: WB(1): World Bank (1985); WB(2): World Bank working group on Poverty in Developing Countries, quoted in Ahmad and Wang (1991); SSB(1): Ma (1991), source given as two agricultural publications by the SSB; Zhao Renwei (1990b) gives exactly the same figures, adding 1984 but lacking 1986, and gives the source as simply SSB; SSB(2): own calculations based on data given by Li (1986) from SSB rural survey; AS: Adelman and Sunding (1987).
Notes: Li also presents his own Gini coefficients from the data from which the SSB(2) Ginis are calculated, giving levels of .237 for 1978 and .264 for 1984. However, there are more income classes in the second year and this biases the comparisons. To remove this problem, decile income shares were calculated on the basis of cubic spline interpolation, and the above figures obtained.

Given that the source for all the Ginis is the SSB, it makes sense to rely on SSB estimates, of which the SSB(1) estimates are the more comprehensive. But the SSB(1) and SSB(2) estimates for 1978 conflict: the former is .21, the latter .25. So we are forced to ignore this initial year.[49] The SSB(1) figures give a 1980-1986 growth rate of 18%.

The proposition that rural inequality has increased faster than urban inequality does not therefore seem to be supported by available inequality estimates. In the first place, such estimates as are currently available are internally contradictory and cannot be adequately assessed. In the second, if we do take what seems to be the most reasonable set of figures, those from SSB(1) from 1980 to 1986, we find an increase in the Gini of .043, comparable in size to our estimate of the increase in the urban Gini of .041 between 1983 and 1990. Since the latter is from a lower base, it corresponds to a higher growth rate: 26 compared to 18%.

49. The WB(2) estimates would have us believe that inequality is higher in 1978 than the rest of the decade. But note that the data given for 1978 by Ahmad and Wang, which I take as providing the basis for the WB(2) figures, disaggregates into only five income classes. 65% of the sample is in the bottom two classes. In such a case, reliance on estimated class means can be particularly misleading.

VII Causes of change

VII.1 Inter-provincial inequality

The Theil index is traditionally used for decompositions of inequality into 'within' and 'between' components. Although a full decomposition is not sensible given the aggregation procedure (which, as shown, results in a neglect of between-city-and-town and thus inter-provincial inequality growth), the Theil index is used in Table 15 to show the growth which has occurred in inter-provincial inequality. The first column of results shows that between-province inequality is on an upward trend throughout the decade more than doubling between 1982 and 1989 (provincial data is unavailable for 1990). Despite the uncertainty as to the share of between-province inequality in total inequality, there can be no doubt that it has risen over the decade, simply because, however much total inequality has risen, it has certainly not doubled.[50]

One would like to know how much of this increase in inter-provincial inequality is due to inflation rates differing across provinces, and how much of it survives once one takes into account differences across provinces in the cost of living. The second set of figures in Table 15 uses cost-of-living indices so that incomes are initially made comparable for the single year 1987 and then comparability is maintained by the use of province-specific inflation rates. In addition, this second set of figures is based on population instead of sample weights. (See Appendix A, point 6 for information on these indices and weights.) Making these adjustments certainly makes a difference. The size of the inter-provincial disparities is reduced significantly, for example, in the final year, 1989, from .016 to .010. However, the general trend remains steep with a 1983-1989 growth rate of 104%.[51]

This exercise shows the importance of deflating to take into account both differences in the cost-of-living across provinces and changes in these differences. It also shows that the eighties saw large rises in inter-provincial nominal and real income disparities.

---

50. Use of the Theil index enables comparisons to be made. Knight and Song (1990), in their analysis of the 1986 Institute of Economics (CASS) household data set, find between-province inequality in 1986 equal to .008, below our .0108, and its share equal to 7%. Hussain, Lanjouw and Stern (1991), analyzing the same data set, find a share of 8%.

51. Experimentation revealed that the use of population weights made very little difference to the results obtained. As reported in III.2.2, sample and population proportions are quite similar. Gini's were also calculated. These showed, without deflation, a generally monotonic increase from .0615 in 1981 to .0989 in 1989.

Table 15 Inter-provincial Inequality

Year	No deflation		Deflated using inflation and cost-of living indices and population weights	
	Level	Growth	Level	Growth
1981	0.0063	.	0.0056	.
1982	0.0061	-4.1	0.0044	-21.8
1983	0.0079	30.7	0.0047	-8.4
1984	0.0078	-2.2	0.0058	24.8
1985	0.0104	34.1	0.0064	8.9
1986	0.0111	6.5	0.0084	32.3
1987	0.0127	14.6	0.0092	9.2
1988	0.0147	15.3	0.0108	17.0
1989	0.0169	15.5	0.0096	-10.8

Notes: The inequality index used is the Theil index (see Table B.1 for a formula), and the inequality is that between provincial per capita mean incomes. For details on the data and methods of calculation, see Appendix A, point 7.

VII.2 Decomposition of inequality by income source

For the years 1985 to 1989, disposable household income per capita, as reported in the SSB survey, can be divided into five types: income from the state-owned sector; income from the collective sector; income from pensions; net gifts and remittances; and a residual, income from other sources. Although a residual, this 'other income' is an important category. It is composed of individual labour income, other labour (including second-job) income, property income (interest payments and rents), income from second-hand sales, and a smaller, unspecified residual. In the analysis which follows, this residual will, simplifying somewhat, be thought of as income from non-traditional or new sources.[52] Each of income from the state-owned sector and income from the collective sector can be further divided into four sub-categories: regular wage income, bonus income, income from subsidies and other income. To simplify an enormously complex system of wage-setting, one can think of these last four categories in the following, stylized way. The regular wage is set by the government, depending on seniority, age and qualifications. Bonuses are set by each firm. They may be established on the basis of productivity, or they may be equally distributed among all workers. Subsidies are of two types. The first is payment by the firm for certain necessities of life, such as train-tickets and haircuts. The second is compensation for price rises, again paid by the firm, but at the order of the government. Finally, 'other payments' consist of a mixture of productivity related and welfare payments. For the years prior to 1985, separate

52. This is in keeping with the terminology of II.2.2, to which the results of this decomposition analysis are related in the next sub-section.

data on pensions are unavailable. In addition, the decomposition of SOE and collective income is not as fine. Table A.2 gives more detail for the different sources, and Table A.4 shows their importance in relation to mean income for the decade.[53]

To analyze the effects of changes in these sources on inequality, a standard technique for decomposition by source is applied to the Gini coefficient of inequality. Fei, Ranis and Kuo (1978) have shown that the Gini coefficient, G, can be expressed as

$$G = \sum_i \mu_i \rho_i \tag{12}$$

where μ_i is the ratio of income from the ith source to total income and ρ is the pseudo-Gini for income from the ith source. The pseudo-Gini for source i is defined analogously to the Gini for source i except that the ranking on which it is based is defined not by source i but by total income. A pseudo-Gini for source i which is less than the Gini for total income indicates, we will say, a relatively *equitable* distribution of source i income, in the sense of a distribution which is equalizing with respect to total income.[54]

Tables 16a to 16f are all based on this decomposition. The first four cover the years 1985 to 1989, the last two the years 1981 to 1984. These two sets of tables are not easily comparable, due to the various structural breaks which occurred after 1984, which are not adjusted for but which affect both the pseudo-Ginis (the smoothing lowers those of latter years) and the relative means (which are unaffected by the smoothing, but which change on account of the changes in coverage). The tables should be read for three different types of information: trends in both relative means and pseudo-Ginis and the relative size of pseudo-Ginis of different income sources in the same year. Increasing importance of one high pseudo-Gini source at the expense of another low pseudo-Gini source will, *ceteris paribus*, raise inequality.

Table 16a decomposes income into the five major sources given at the start of the sub-section. The relative shares of the major income sources are fairly constant, though the share of collective income is slowly declining throughout the period and falls a full percentage point in 1989, which suggests the collective sector was much more weakly placed than the state sector to maintain real income in the face of high inflation and recession.

53. For the most part, the decomposition is the same as that used in the various published sources, but some difficulties are posed by changing definitions: from 1988, a more detailed decomposition of income is available. Note in particular that from 1988, the category of bonuses includes 'floating wage' and 'contract' income, both of which are very small.

54. Ginis for these sources cannot be calculated, as all the data are ordered by disposable income.

Whereas the relative means of income from different sources shows little change, Table 16a reveals that the pseudo-Ginis of the different sources show more definite trends: that for SOE income is slowly rising (from .178 in 1985 to .187 to 1989), while that for other income is rising much more rapidly (from .209 in 1985 to .352 in 1989), but not as rapidly as the pseudo-Gini for collective income is falling (from .032 in 1985 to -.001 in 1989). Other or 'new' income is becoming much less equitably distributed, thus increasing inequality, while collective income is becoming much more so, reducing inequality.

Tables 16b to 16d decompose, respectively, combined SOE and collective ('traditional') income, SOE income and collective income into regular wages, bonuses, subsidies and other income. Similar messages emerge from all three tables. The regular-wage relative mean falls in the traditional sector by 13 percentage points. This is almost entirely compensated for by an increase in 5 percentage points in bonuses and 6 in subsidies. While the increase in bonus income is genuine, that in subsidy income is in part a statistical artifact. The largest increase, in 1988, is due to subsidies for price rises being taken out of wages and counted with other subsidies. This makes comparisons difficult, so the focus in what follows is on bonus rather than subsidy income. Identical trends are apparent in the SOE sector. In the collective sector, however, against a background of falling overall income, subsidies and bonuses have not been able to do much more than maintain their relative importance. Second, turning to the trends in pseudo-Ginis, for the formal sector overall, those for regular wage and subsidies are falling, that for bonuses is rising, and that for other income shows little change. For the SOE sector, the pseudo-Ginis for bonuses and other income are rising, that for the wage is constant, while that for subsidies is falling. For the collective sector, all pseudo-Ginis are falling, though those for bonuses and other income are doing so relatively slowly. In all these three tables, bonuses and other incomes have relatively high pseudo-Ginis. Those for subsidies and regular wages are relatively low.

It is clear that within both collectives and SOE there has been a clear move away from payment by wage and toward payment by bonus. Contrary to much assertion in the literature (see, for example, footnote 11), this movement towards payment by bonuses has increased inequality. The importance of this is implied by the massive contribution of traditional-sector earnings (88%) and especially SOE earnings (73%) to total earnings. Correspondingly, the contribution of bonus payments to total inequality has increased by about 9 percentage points over the course of the latter half of the decade.

Results from the earlier years, though more limited in scope, confirm the above analysis in two important respects. First, Table 16f shows that the increase in bonuses is not a phenomenon restricted to the latter half of the 1980s. The share of bonuses to total income rises in every year, from 10.3 in 1981 to 14.0% in 1984. 60% of this rise occurs between 1983 and 1984. It is only in this latter year that bonuses become less equitably distributed, rising from .211 in 1983 to .230 in 1984. Second, the failure of the urban collective sector to improve its share of urban income is again evident. It is constant at around 18.6% for the four years (see Table 16e). This performance - relative stagnation in the early eighties and then relative decline - stands in stark contrast to the outstanding success of the rural collective sector, whose total output value has risen as a proportion of national output value from 7.7% in 1980 to 16.5% in 1985 and 20.6% in 1987 (Chen, Watson and Findlay, 1990, Table 4). The reasons for the relatively poor performance of the urban collective sector deserve further research.[55]

---

55. The figures of Table 15e may also seem to be hard to reconcile with claims that in the early eighties the urban collective sector did grow rapidly. For example, Riskin (1990, p.354) writes that

> ... virtually all of the 1984 net increase of 4.83 million in non-agricultural employment (including urban but non rural self-employed workers) occurred in collective (4.72 million) and private (1.08 million) establishments. These numbers add up to more than the total increase because state sector employment actually declined.

But recall that the official definition of 'urban' was changed in 1984 (see III.1) resulting in a massive increase in the town population, in which collectives are more important than in cities. As argued in III.2.2, this reported change in employment may reflect nothing more than this definitional change and so would not be apparent in the SSB survey data.

Table 16 Decomposition of Gini by Income Source

Table 16a Income by source: SOE, Collective, pensions, net gifts and other, 1985-1989

		SOE INCOME			COLLECTIVE INCOME			PENSIONS			NET GIFTS			OTHER		
Year	Gini	Cont	μ	ρ	Cont	μ	ρ	Cont	μ	ρ	Cont	μ	ρ	Cont	μ	ρ
1985	0.164	79.5	73.2	0.178	3.3	16.6	0.032	10.1	7.5	0.219	-1.5	-4.2	0.059	8.7	6.8	0.209
1986	0.166	79.2	71.8	0.183	2.2	17.3	0.021	11.2	7.5	0.249	-1.3	-4.7	0.045	8.6	8.1	0.176
1987	0.167	76.6	72.4	0.176	2.3	16.1	0.024	12.4	9.0	0.230	-0.3	-5.2	0.010	9.0	7.7	0.195
1988	0.175	76.2	73.1	0.183	1.0	15.8	0.011	10.1	8.9	0.199	-0.6	-4.6	0.024	13.4	6.9	0.341
1989	0.181	75.7	73.3	0.187	-0.1	14.8	-.001	9.7	8.9	0.197	0.6	-4.2	-.025	14.1	7.3	0.352

Table 16b Income from SOE and Collectives combined, 1985-1989

		REGULAR WAGE			BONUSES			SUBSIDIES			OTHER INCOME		
Year	Gini	Cont	μ	ρ	Cont	μ	ρ	Cont	μ	ρ	Cont	μ	ρ
1985	0.164	43.9	56.0	0.128	17.8	13.6	0.214	11.3	11.4	0.162	9.8	8.8	0.183
1986	0.166	46.0	58.2	0.131	12.6	9.6	0.216	11.4	11.4	0.167	11.4	9.9	0.192
1987	0.167	39.6	53.8	0.123	19.0	14.8	0.213	11.3	11.4	0.166	9.0	8.4	0.179
1988	0.175	31.2	46.2	0.118	22.5	17.6	0.224	13.5	16.5	0.143	9.9	8.5	0.204
1989	0.181	28.1	42.7	0.119	24.0	18.8	0.230	13.6	17.8	0.138	9.8	8.7	0.204

Table 16c SOE Income by Type: regular wage, bonuses, subsidies and other, 1985-1989

		SOE INCOME			REGULAR WAGE			BONUSES			SUBSIDIES			OTHER INCOME		
Year	Gini	Cont	μ	ρ	Cont	μ	ρ	Cont	μ	ρ	Cont	μ	ρ	Cont	μ	ρ
1985	0.164	79.5	73.2	0.178	44.1	45.6	0.158	15.3	10.9	0.231	11.0	9.5	0.190	9.1	7.2	0.206
1986	0.166	79.2	71.8	0.183	46.5	46.5	0.166	10.9	7.7	0.236	11.3	9.6	0.194	10.6	7.9	0.222
1987	0.167	76.6	72.4	0.176	40.1	44.0	0.152	17.0	11.7	0.241	11.0	9.8	0.189	8.4	6.9	0.204
1988	0.175	76.2	73.1	0.183	33.0	38.0	0.153	20.2	13.8	0.256	13.5	14.2	0.167	9.5	7.1	0.235
1989	0.181	75.7	73.3	0.187	30.3	35.5	0.154	22.0	15.0	0.265	14.0	15.5	0.163	9.4	7.3	0.234

<u>Notes:</u> See the next page.

Table 16d Collective Income by Type: regular wage, bonuses, subsidies and other, 1985-1989

		COLLECTIVE INCOME			REGULAR WAGE			BONUSES			SUBSIDIES			OTHER INCOME		
Year	Gini	Cont	μ	ρ	Cont	μ	ρ	Cont	μ	ρ	Cont	μ	ρ	Cont	μ	ρ
1985	0.164	3.3	16.6	0.032	44.1	10.4	-.003	15.3	2.7	0.148	11.0	1.9	0.024	9.1	1.5	0.075
1986	0.166	2.2	17.3	0.021	46.5	11.7	-.007	10.9	2.0	0.140	11.3	1.7	0.014	10.6	2.0	0.071
1987	0.167	2.3	16.1	0.024	40.1	9.8	-.010	17.0	3.1	0.106	11.0	1.6	0.029	8.4	1.5	0.065
1988	0.175	1.0	15.8	0.011	33.0	8.3	-.038	20.2	3.7	0.108	13.5	2.3	0.001	9.5	1.5	0.058
1989	0.181	-0.1	14.8	-.001	30.3	7.2	-.054	22.0	3.9	0.095	14.0	2.3	-.027	9.4	1.4	0.049

Table 16e Income by source: SOE and Collective, 1981, 1983, 1984

		SOE INCOME			COLLECTIVE INCOME		
Year	Gini	Cont	μ	ρ	Cont	μ	ρ
1981	0.1627	75.6	79.5	0.155	11.3	18.5	0.099
1983	0.1577	77.0	79.0	0.154	6.3	18.7	0.053
1984	0.1660	79.9	80.1	0.165	6.8	18.6	0.060

Table 16f Bonus Income by source: SOE and Collective, 1981, 1983, 1984

		TOTAL BONUSES			SOE BONUSES			COLLECTIVE BONUSES		
Year	Gini	Cont	μ	ρ	Cont	μ	ρ	Cont	μ	ρ
1981	0.1627	14.2	10.3	0.223	11.7	8.4	0.227	2.5	1.9	0.206
1983	0.1577	15.8	11.8	0.211	12.5	9.4	0.210	3.2	2.4	0.215
1984	0.1660	19.4	14.0	0.230	16.4	11.3	0.241	3.0	2.7	0.187

Notes: In all tables the Gini figure is that for total inequality. μ gives the mean of the type of income relative to the mean of total income; ρ gives the pseudo-Gini of the type of income (for definition see text). 'Cont' measures the contribution of the distribution of the particular type of income to total inequality and is equal to μ times ρ divided by Gini. For details on the data, see Appendix 1. Pseudo-Ginis are calculated using the standard formula, but with an ordering of income sources by total disposable income.

VII.3 Inequality and reform

From the perspective of possible links between reform and inequality, it is striking that inequality only begins to rise in 1984, for it is often claimed that urban reform also began in 1984. The oft-cited 'urban reforms of 1984' refer to a reform package adopted by the Chinese Communist Party on October 20 1984 (Gale Johnson, 1990, p.51). The package was a significant one, containing as it did all the themes which have guided subsequent reform efforts, namely, greater enterprise autonomy, a reduced role for planning, price reform, and the linking of payment to performance.[56] However, the reform package was not the first presentation of these issues. Prior to 1984, firms had already been allowed to retain their profits, subject to taxation (see Gale Johnson, 1990, pp.47-50 for a summary of the many changes made in the early eighties to enterprise-government relationships), and the importance of bonuses (an attempt to link pay to performance) was, as we have seen, increasing throughout the early eighties.[57] Hua, Zhang and Luo, three analysts who were themselves involved in the process of urban reform, devote an entire chapter of their recent (1992) book on Chinese economic reform to "Early Urban Reform Attempts" covering the pre-1984 period. So any attempt to link the rise in inequality in 1984 to the alleged inception of reform in that year is unlikely to be convincing. However, the possibility that reform has played a longer-term causal role is a different matter, one which is addressed below.

Two such longer-term effects of reform, analyzed in II.2, relate to decentralization and the growth of new job opportunities. The conclusion of the theoretical analysis was that, not universally but at least in the urban Chinese context, decentralization was likely to increase inequality, but that it was not possible to predict the distributional impact of the growth of new job opportunities. We are now in a position to compare these conclusions to the evidence collected in VII.1 and VII.2.

The decomposition by income source is relevant firstly to the analysis in II.2.2 of the impact of growth of new job opportunities. We can conclude from the previous section that the distribution of new earning opportunities (proxied by 'other income') promoted inequality growth in the late eighties in the sense that its pseudo-Gini rose. To try to understand the reason for this, the framework presented in II.2.2 is relevant. Although this framework cannot be directly applied

56. According to Perkins (1988, p.613) the reform package "represented an important step towards major changes in the system".

57. Hussain (1990, p.45) writes that "bonuses linked to enterprise profits ... were reintroduced in 1983". Bonuses as such existed throughout the eighties.

since SOE households cannot be isolated from collective sector households, indirect evidence on its applicability is available. If new earning opportunities have gone mainly to the SOE workers, we would expect two features to emerge from the decomposition: first, 'new' income should become less equitably distributed as it becomes a more important source for the rich; second, collective income should become more equitably distributed as it becomes a more important source for the poor, who are excluded from these new sources of income. Both features are present in the data presented in Table 9.[58] A note of caution must be sounded however. The main increase in the pseudo-Gini of other income comes in a single jump between 1987 and 1988. The judgement that there is an underlying disequalizing trend in the distribution of 'new' income could be made with greater confidence if data was available showing that the trend continued over more recent years.[59] In addition, there may be other factors at work: differential access by households within the two traditional sectors may also be important.[60] Due to the limitations of the data, the analysis cannot be taken any further, and one must remain content simply with noting that the evidence is consistent with the explanation proposed.

On the question of decentralization, clearly the provinces can be viewed as one of the types of 'recipient-units' to whom power has been decentralized as a result of reform (II.2). That inter-provincial inequality has risen is consistent with the widely-held belief that it is the South-Eastern seaboard provinces which have done best from the reform. These were not the richest provinces to begin with (they were not as rich as Beijing or Shanghai, for example) but they certainly had above average incomes prior to the reforms. The rise in inter-provincial inequality - if interpreted as evidence that expected *ex post* inequality has risen - is also what one would expect from the increased 'noise' decentralization may introduce into an economy.

Turning to the disequalizing impact of non-wage payments, it is plain that bonuses have been used by companies to get around centrally-imposed controls on the standard or regular wage.

---

58. Another reason for the falling pseudo-Gini of collective income is the falling share of that source in total income. If collective workers cannot move out of their sector, and are experiencing slower relative growth in their wages, the pseudo-Gini for collective income will rise, as its relative mean falls. However, the fall in the relative mean of collective income is less than 2 percentage points from 16.6 in 1985 to 14.8% in 1989. The fall in the pseudo-Gini over the same period is off a much larger order of magnitude, from .032 to -.001.

59. It is striking that the jump comes in a year when a more detailed break-down of income is given. It is not inconceivable that the disequalizing trend in other income is due to some change in definition. However, note that the share of other income remains roughly constant at between 7-8%.

60. For example, say that some collective workers did very well in earning new income earning opportunities, while most did badly. Then we could still have an increase of collective income among the poor, and an increased concentration of SOE income among the rich (reflecting the failure of most collective workers) and a less equitable distribution of other income.

However, our data is unable two distinguish between two hypotheses: the first that they reflect differences in individual worker productivity (thus increasing within-work-unit inequality), the second that they reflect differences in overall factory profitability (thus increasing between-work-unit inequality). The second explanation is more consistent with the widespread belief that non-wage payments, especially bonuses, are distributed in an egalitarian way within the factory and provides an overlooked reason, even if the belief is true, for bonuses to be disequalizing. Whichever the explanation, the underlying hand of reform is once again evident. We saw in II.1 the argument that in the early eighties inequality fell because "in the early stage of reform emphasis was on raising the income of the lower paid workers." (Zhao Renwei, 1990b, p.192). In the mid and late eighties by contrast, wage policy was a more a matter for the individual firm. The government maintained control over the basic wage, but any further compressing of wage grades, if it occurred, was more than offset by the disequalizing growth of non-wage payments.

The importance of bonuses is also evident from the fact that the two years in which inequality fell after 1983 (1986 and 1990) are the two years in which the share of bonuses in total wage income fell substantially.[61] The fall in 1986 was from 13.6 to 9.6% (see Table 16b). And a report from the SSB suggests a fall in the share of bonus income from 18.8% in 1989 to 15.5% in 1990.[62]

Figure 7 demonstrates the close relationship between the share of bonuses in total income, the level of inequality (measured using the Gini) and, interestingly, the inflation rate in urban China over the eighties. (See the notes to the figure for the adjustments required to the bonus-share and inequality figures to cope with sample changes.) All three trend upwards and all three fall in 1986 and 1990.[63] The two years of declining inequality were *not* years of low real growth: to the contrary both were years of above-average growth. But both were years in which tightened administrative macroeconomic controls were in place precisely to bring inflation down. (In both cases, the controls were introduced in the previous year, but it is reasonable to allow for some lag in their operation.) Greater macroeconomic control in these years meant greater control

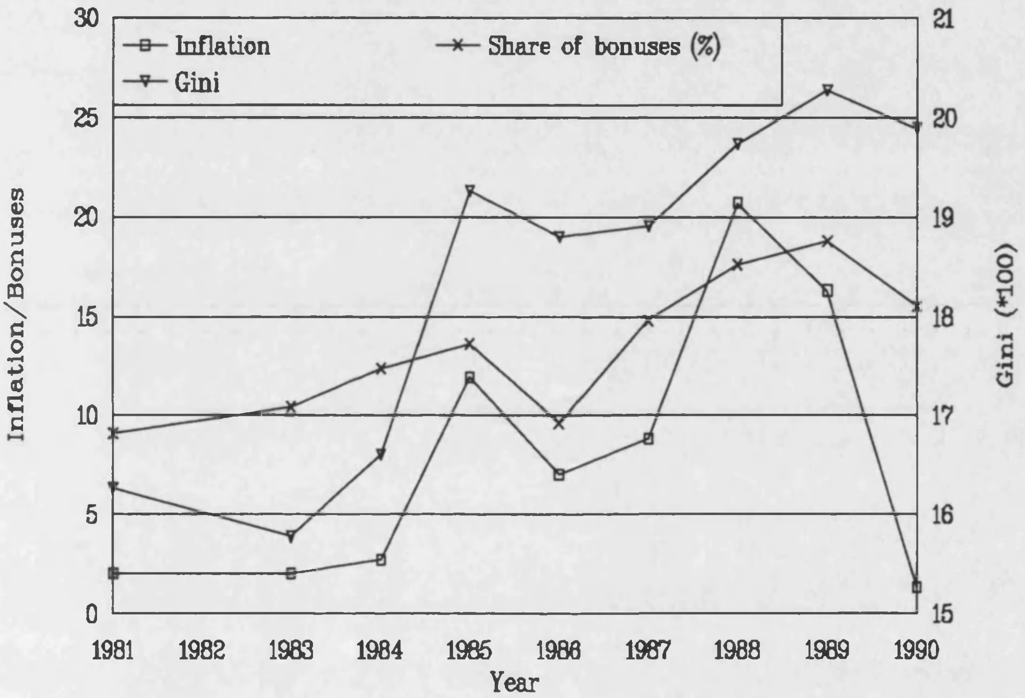
61. The figures of Table 8 show a slight increase in inequality in 1986 and a fall in 1990. But once adjustment has been made for the increase in the sample base in 1986, inequality falls in that year by even more than it does in 1990: see V.2.

62. The SSB reports that "the average standard wage increase in 1990 was 13.5% over 1989, making 1990 one of the better years for wage increases. On the other hand, the average bonus income fell by four percent ... " (from Zhongguo Tongji Xinxi Bao, 4 April 1991, p.1, translated in JPRS-CAR-91-032, 13 June 1991, p.81 [see bibliography for explanation of referencing system]). Taking the same article's reported growth rate for total per capita income of 8.5% gives a fall in the share of bonus income from 18.8% to 15.5%,

63. It is true that inflation leads the bonus-share and the Gini down a year earlier in 1989, but this reflects the deceleration of inflation in the latter half of 1989.

for the government over firms and hence greater control over their payment of workers.[64] Conversely, it seems that the winding-back of central power in the interim, decentralizing years resulted in greater firm-level autonomy which, in turn, resulted in greater payments by bonuses in an attempt by firms to get round remaining central controls on regular wages. The result in these years was both inflation (due to the increase in demand) and inequality (as some firms were better placed to pay-out bonuses than others).

Figure 7 Bonuses, Inflation and Inequality in Urban China, 1981-1990



Notes: Since inequality figures are unavailable for 1982, no data are given for this year. The Gini figures are taken from Table 8, adjusted, to take into account sample changes, as indicated in V.2. The inflation figures come from Table 1. The bonus-share figures from 1985 to 1989 are taken from Table 16b. For pre-1985 figures the figures in Table 16f are used, adjusted (multiplied by a factor of .87) to take into account the changed share of SOE and collective income in total income. The 1990 bonus share is estimated - see footnote 62.

64. The SSB report that the share of bonuses in wage income fell in 1990 (see footnote 62 for the reference) concludes that this " shows that the macroeconomic measures taken by the state to readjust urban incomes have been effective".

Although decentralization does seem to have been very important in promoting inequality, two qualifications are necessary. First, it is not possible to precisely disentangle its effects from those of the growth of new job opportunities. For example, the growth in inter-provincial inequality may be due, in part, to new- income-earning opportunities growing faster in some provinces than others. Alternatively, factors other than the two highlighted may also have been involved. For example, certain provinces may have been given favourable treatment by the central government (allowed to implement certain reform measures before other provinces or given increased investment funds). Second, the official statistics of the share of bonuses in total SOE wage payments show a different time-profile over the eighties than that share reported here, with a fall between 1984 and 1985 and none in 1986 (see Hussain and Stern, Table 3.4). A fall in 1990 is reported, but only a slight one from 17.6 to 17% (1991y). If, therefore, inflation and bonuses cannot be so closely linked, it may be that high inflation has itself contributed to higher inequality. But as long as we accept the standard position that inflation was the result of decentralization,[65] the link between decentralization and inequality growth is maintained.

In summary, although as argued in II.2 there is no necessary link between reform and inequality growth, the Chinese urban reform process does seem in practice to have been disequalizing. It is not satisfactory, it has been argued, to claim as evidence for the link between reform and inequality growth the coincidence of the inception of the two phenomena in 1984. However, two inter-related factors - decentralization and a growth of new job opportunities - have been isolated which do seem to link reform and inequality growth. Macroeconomic instability emerges from the survey data not so much a cause of inequality growth, but rather, like the uneven growth in inequality itself, a symptom of the uneven tempo of decentralization through the eighties.

VIII Conclusion

This chapter has used aggregated household survey data collected annually between 1981 and 1990 to investigate and explain the level of and trends in income inequality in urban China over the decade. The reliability of the data has been evaluated, with particular attention being paid to problems posed by the procedure by which the data are aggregated.

---

65. See, for example, Hussain and Stern (1991, pp.53-4): "The reforms simultaneously loosened control over enterprise investment and wages, thereby engineering both a consumption and an investment boom. The conjunction of the two, together with the extension of the market determination of prices, eventually led to a rapid inflation."

It has been found that income inequality is low in urban China compared to developed countries because the income share of the bottom 40% is relatively large and that of the top 20% relatively small. The share of the middle 40-80% is comparable to that found in most developed countries.

The survey data analyzed challenges the current consensus that urban inequality has not increased in urban China over the last decade. While the share of the middle class has been relatively stable, the income share of the poor has fallen and that of the rich has risen. Although inequality fell from 1981 to 1983, on account of a wage compression in place since the late seventies, it trended upwards thereafter, increasing by, it is estimated, some 25% as measured by the Gini coefficient. The conventional wisdom that growth in rural inequality has been greater than that in urban inequality must be questioned: is it unclear what has happened to rural inequality since 1978, but the best estimates, which cover 1980 to 1986, suggest that rural inequality grew over these years by less than urban inequality did between 1983 and 1990. However, the urban disequalization, while significant, has not been large enough to prevent a rise in urban welfare and reduction in urban poverty over the decade on account of growth in mean income.

On the relationship between transition and inequality, market-oriented reform will not always increase inequality, but in urban China it does seem to have done so. Two features of reform seem to have led to the rise in inequality: decentralization and growth in new income-earning opportunities. The growth of new-earning opportunities may have promoted inequality growth via differential access to such opportunities by state-owned and collective sector households. And decentralization seems to have resulted in increasing regional disparities (interprovincial inequality more than doubles over the decade according to one index) and to an increasing share of income increasingly inequitably distributed as non-wage payments, particularly as bonuses. The estimates derived in the paper also lead to the conclusion that all three of the share of bonuses in total income, inflation and inequality moved together through the eighties. Since the first two are good proxies for decentralization, this provides further evidence for the inequality-promoting role played by the latter.

It has been noted that it is possible to find explanations for the observed rise in inequality other than a rise in underlying inequality. These include the increasing sample size over the decade and the possibility that implicit changes have occurred, for one reason or another, in the survey's coverage of income sources. Although for most of this chapter and in this conclusion we put these alternative explanations to one side, since one cannot exclude them, one can, strictly speaking, only conclude that the evidence is consistent with a rise in underlying inequality.

Nevertheless, the plausibility of the conclusion that inequality has increased is raised by the timing of the growth and that fact that it is not monotonic, but seems rather to be strongly linked to the ebb and flow of the reform process.

The inequality-promoting forces in operation over the eighties are unlikely to go away if reform continues and urban inequality can be expected to rise further. It will, however, at current trends be a long time before urban China 'enjoys' average levels of inequality. To illustrate, taking into account the various biases associated with the data, the income share of urban China's top 20% grew by perhaps 7% between 1983 and 1989. To reach the income share of the top 20% in the UK, at this rate, would take another twenty years. Viewed from this perspective, rising urban inequality would hardly seem to be a pressing concern.

At what rate inequality will increase in urban China, assuming it does, is impossible to say. The rate of increase may well increase though if and when labour mobility is improved, and unemployment becomes a real possibility for formal-sector workers. This would not only have a direct impact on the distribution of income, but also a substantial indirect impact via its influence on patterns of pay bargaining. For this reason the lessons to be drawn from China for transitional economies of Eastern Europe are perhaps limited. One imagines in these countries, where adjustment is occurring without growth, inequality will grow more rapidly, at least in the short-term and in the urban sector, as unemployment rises. In addition the weakening of social security systems in the ex-Communist countries is also likely to promote inequality. However, the experience of China is certainly not irrelevant. Decentralization to lower levels of government and the firm, the growth of new job opportunities, and a weakening of macroeconomic control certainly characterize Eastern Europe and the former Soviet Union.[66] Urban China presents evidence that these are inequality-promoting forces.

While from a positive perspective what is interesting is the apparent rise in inequality, the fact that urban inequality remains low is also important. The gap between urban and rural average incomes in China is often commented on. The existence of a gap *per se* is not unusual: the phenomenon of urban bias is well-documented throughout the developing world. What is unusual about China is, first, the size of the gap, and, second, that the high-mean sector is the low-inequality sector. These two factors combine to ensure that the great bulk of the absolutely poor in China are rural rather than urban dwellers.

66. Examination of income distribution trends in post-1989 East Europe is limited. Atkinson and Micklewright (1992) examine the pre-1989 period and find no clear trends in either Hungary, Czechoslovakia or Poland.

Appendix A Notes on the published data

1. The total household (as against per capita) income data (used in Sections IV and V.2) were obtained from the yearbooks (SSB, 1988y, 1989y, 1990y, 1991y). The number of classes is 12 between 1987 and 1989 and 16 for 1990. Calculations of indices were first made using the given class means. However, to allow for a proper comparison between 1990 and the earlier years, the top 6 classes were aggregated into one for data from that year, so that 1990 had classes with the same bounds as 1989, and had the same number of classes as all the previous years. Calculations for 1990 based on this modification are presented in Table 11 with the heading '1990 comp.'. Although households are ranked by total income, income class means are given in terms of per capita income. To obtain mean household incomes these were multiplied by average household size. This introduces some inaccuracy, as household size will not be constant within an income class, but the distortion is unlikely to be large. To obtain decile income shares from these data (for the purpose of the international comparisons of Section IV), cumulative income shares were calculated and then interpolated using a cubic spline.

2. As mentioned in the text, quantile means of 1982 disposable income are unavailable. The mean income given in Table 1 is based on income class data, as is the Gini in Table 9. See Table A.1 for sources. Provincial overall (not quantile) means are available for this year, and are used in Table 15.

3. As mentioned in the text, income prior to 1986 is the average for cities; income from 1986 onwards is the average for cities and county towns, or all-urban. To provide comparability over time, the pre-1986 data was adjusted downwards. For all the pre-1986 years, a weight of .6 was given to the city mean and .4 to the town mean (see III.3 for the use of these weights). For 1985, the town mean is given, but not for earlier years. For these years, it was assumed that the city-town income mean ratios were the average of those given in Table 10 for 1985 and 1987-89 -the years for which this data is available - which comes out to 1.248.[67]

4. Income shares and means for eight quantiles are used in the chapter's calculations, even though, as stated in III.2, for the early years means are unavailable for some quantiles. The following strategy was adopted to deal with this problem. The second-order stochastic dominance analysis was conducted only over quantile data actually given. Since this showed inequality to be increasing over the decade, the missing shares were calculated on the assumption that, since they

67. Note that a reported town mean is available for 1986, but that quantile-mean town data is not available for this year, so that a constructed mean, as used for the other years, cannot be utilized.

were for early years, they reflected a relatively equal distribution. Specifically, the ratio of the 0-5% to 5-10% quantile was assumed to be equal to its highest for the decade in 1981. Inspection of the data revealed 1987 to display the highest ratio, equal to .844. Similarly, the ratio of the 80-90 to 90-100% quantiles was assumed to be equal to its highest for the decade in 1981 to 1984. Inspection of the data revealed 1985 to display the highest ratio, equal to .793. Use of these ratios enabled the missing 0-5% and 5-10% shares to be calculated for 1981 (given information on the 0-10% share), and the missing 80-90 and 90-100% shares to be calculated for 1981 to 1984 (given information on the 80-100% share). The assumption on the lowest two income groups is not important as it applies only to 1981. The assumption on the highest two income groups is more important. Since the relevant ratio is monotonically falling between 1985 and 1989, the assumption that the years up to 1984 had a 80-90 to 90-100% ratio equal to the maximum for the later years probably results in an underestimation of inequality growth.

5. Quantiles, though treated in the chapter as if bounded by rounded percentiles (eg., bottom 5%, between 20 and 40%), are in fact defined over an exact number of households. This is not a cause for concern, as the exact numbers almost exactly correspond to the rounded percentiles, except for the years of 1987 to 1989 for the category of poorest households. Although for other years (from 1983 onwards), this category is almost exactly 5%, for the three years mentioned it is 5.2%.[68] Income shares given for these years are the share of income for the bottom 5%, calculated on the fiction that the 5-5.2% quantile receive the same mean income as the 0-5% quantile. On this assumption, we can take the 5.2% share and multiply it by 5/5.2 to get the 5% share. Since in fact the 5-5.2% group will get a higher mean income than the 0-5% group this adjustment overestimates the income share of the bottom 5% for the three years. Mean incomes and income shares of the 5-10% group were obtained as a residual (using given 0-10% and calculated 0-5% figures). (Lower bounds on the income share and mean of the poorest 5% were also calculated on the assumption the 5-5.2% group received the same mean income as the 5.2-10% group. These are reported in the notes to Table A.3, as are the original figures).

6. Between-province inequality figures (given in Table 15) were obtained based on reported per capita provincial means. Provincial mean incomes came, for 1981 to 1985, from SSB (1987s). Those for 1987 to 1989 came from SSB, 1988s, 1989s 1990s, respectively. Those for 1986 came from 1989se. When using sample weights (the first two columns in Table 16), the provincial means were weighted by the product of provincial sample size and average provincial household

---

68. Each year's figure was slightly different. Although the modification described in the text was executed using the exact figure for each year, for expositional purposes it is assumed the figure was 5.2 for each of the three years.

size. These came from the same sources. For 1983 and 1984, for which provincial sample size information was unavailable, it was assumed that sample sizes were the mean of those in 1982 and 1985. For the recomputation of interprovincial inequality (the last two columns of Table 15), urban cost-of-living indices were obtained for 1987 from Howes and Lanjouw (1991). Comparability across years was obtained by use of province-specific urban inflation rates using the cost-of-living indices for 'staff and workers' given in the annual China Statistical Yearbooks. These are only available for 1984 onwards. For 1981 to 1984 is was assumed all provinces had the same inflation rates. Provincial totals of non-agriculture-registered population were obtained from SSB (1990a) and used as population weights. It was assumed that all provinces have the same proportion of urban non-agricultural-registered to total non-agricultural registered. Figures for the former (the SSB survey base) are not available for 1988 and 1989. Xizang (Tibet), Qinghai and Anhui are excluded from analysis due to complete or partial unavailability of data.

7. The definitions used in the decomposition by income source are given in Table A.2 below. Most of the definitions given in the SSB publications are self-explanatory; additional details are provided in SSB, 1985. Mean income for the various income sources and quantile groups are presented in Table A.4. Note that the adjustment explained in point 5 was not made for this exercise. Rather, for simplicity, it was just assumed that the figures given for the bottom 5.2% for 1987 to 1989 were in fact given for the bottom 5% for these years. Checks were made to ensure the decomposition by source exhausted total disposable income. This led to the correction of certain typographical errors. For 1985, the 20-40% average collective regular wage was changed from 33.76 to 83.76. For 1986, the disposable income of the 10-20% quantile was changed from 507.36 to 570.36, as it appears in SSB, 1988y. For 1987, the 20-40% average SOE contract income was changed from missing to 2.4. For 1988, income from sources given fell short of total income by 3.3% on average, and for the top six of the eight quantiles. For the 0-5% quantile, the short-fall was 2.8% and for the 5-10% it was 3.2%. Examination of the relevant SSB publication reveals the reason for this. Total disposable income figures given in the summary section are larger by the percentages given above than the total disposable income figures given in the section where incomes are decomposed by source. In other words, the quantile total disposable income figures have been revised upwards by around 3%. This certainly casts some doubt on the mean figure for 1988. However it is not so important for purely distributional analysis. Since the short-fall is almost exactly proportional for all quantiles, it can be safely ignored. All types of income were multiplied by a factor of one plus the proportional shortfall for each quantile. Once these various corrections were made, no residual of more than .2% in absolute value remained for any quantile or any year between total income and income aggregated over all given sources.

Table A.1 Summary of Data Available

Year	Quantile data			Income class data		
	Number of quantiles; whether mean and/or bounds given,	**Coverage: cities, towns, all- urban or all three**	**Source: (SSB ...)[1]**	**Number of income classes; whether mean given**	**Coverage: cities, or all-urban**	**Source (SSB ...)**
1981	6; means only	Cities	1987s	6; means given	Cities	1987s
1982	-	-	-	6; means given	Cities	1987s
1983	7; means only	Cities	1987s	6; means given	Cities	1987s
1984	7; means and bounds	Cities	1987s	6; means given	Cities	1987s
1985	7; means and bounds	Cities and county-towns	1988se	10 classes and means given for sub-sample of 9178 city households; 4 classes and no means for full sample of city households; 9 classes and no means for full sample of all-urban households	Cities and All-urban	1987s for cities sub-sample; 1986ye for cities full sample; 1989y for urban sample
1986	8; means only	Urban and cities	1989se	13; no means	All-urban	1987y
1987	8; means only at national level, means and bounds at provincial level	All three	1988s	13; no means	All-urban	1988y
1988	8; means only	All three	1989s	12; no means	All-urban	1989y
1989	8; means only	All three	1990s	12; no means	All-urban	1990y
1990	8; means only	All-urban	1991y	11; no means	All-urban	1991y

Note: 1. The sources given in this column are also those used in the chapter.

Table A.2 Definitions Used in the Decomposition by Income Source

1.Income from the state-owned sector.
1.1 Regular (or base) wage income (*biaozhun gongzi*).
In 1984, the regular wage is taken as the sum of *shijian gongzi* and *jijian gongzi* (time and piece-rate wages) excluding from the latter *jijian chao'e gongzi* (over-quota piece-rate payments), which are included in 1.2.3.
1.2 Bonus income.
1.2.1 Floating wage income (*fudong gongzi*).
1.2.2 Contract income (*chengbao shouru*).
1.2.3 Unspecified bonuses (*jiangjin, chao'e gongzi*).
Note: 1.2.2 and 1.2.3 are only reported from 1988 onwards. Although they are not bonus income, they have been included in category 1.2 on the assumption that they are productivity-based payments. In any case, the size of each is only about 1%.
1.3 Income from subsidies.
1.3.1 Allocated. Subsidies (*gezhong lutie*) explicitly labelled as assigned to state-owned sector employees.
1.3.2 Unallocated. For 1988 and 1989, a separate unallocated category is introduced of unallocated price subsidies (*jiage butie*). These are compensation for inflation and are not allocated in the SSB publications to sector. They have been allocated by me (to this item of 1.3.2) in proportion to the share of allocated subsidies to the state-owned sector in combined (SOE and collective) allocated subsidy income (i.e., 1.3.1 over the sum of 1.3.1 plus 2.3.1). 1.3.1 is only available from 1983 onwards.[69]
1.4 Other income.
1.4.1 Allocated. The residual item (*qita*) for income from state-owned sector.
1.4.2 Unallocated. Data is given for other work unit income (*cong danwei dedao qita shouru*). This is allocated by me to this item in proportion to the share of given income for the state-owned sector (*quanming souyouzhi zhigong gongzi*) (i.e., 1 minus 1.3.2 minus 1.4.2) in combined given income (i.e., the sum of given SOE income, as defined above, and given collective sector income (*jiti souyouzhi zhigong gongzi*) (i.e., 2 minus 2.3.2 minus 2.4.2).

2.Income from the collective sector.
2.1 Regular wage income.
2.2 Bonus income.
2.2.1 Floating wage income.
2.2.2 Contract income.
2.2.3 Unspecified bonuses.
2.3 Income from subsidies.
2.3.1 Allocated.
2.3.2 Unallocated.
2.4 Other income.
2.4.1 Allocated.
2.4.2 Unallocated.
Note: 2 to 2.4.2 are defined analogously to 1 to 1.4.2. Pre-1985 only 2 and 2.2.3 are given.

3. Retired pension income. Pre-1987, *tuixiu renyuan shouru*; in 1988 and 1989, *litui xiujin*. Not available pre-1985.

4.Net gifts. Net income from gifts (*zengsong shouru* minus *zengsong zhichu*) and net income from dependents (*zengyang shouru* minus *zengyang zhichu*).

---

69. It is something of a simplification to assume that all price subsidies go to workers and none to pensioners and children. The next chapter (IV.3) gives a more accurate account.

5.Other income.

5.1 Labour income.
5.1.1 Income from private and self-employment (*geti laodongzhi shouru*).
5.1.2 Employment income of retired persons (*bei yong hou liuyong de lixiuxi ren yuan shouru*).
5.1.3 Income of other employees (*qita jiuye zhe shouru*).
5.1.4 Other labour income (*qita laodong shouru*).
5.2 Non-labour income.
5.2.1 Income from sales (*chushou caiwu shouru*).
5.2.4 Property income (*caichenxing shouru*).
5.2.5 Other transfer income (*qita zhuanyixing shouru*).
5.2.6 Other special income (*qita tebie shouru*).
5.3 Unspecified (*qita shouru*).

Notes: For 1981 and 1982, only 5.1.1, 5.1.4, 5.3 are available. In 1983, 5.2.1 is added. The unspecified category, 5.3 is abolished after 1987 and decomposed into 5.2.4 to 5.2.6 and 5.1.2 to 5.1.3. Note too that 'other employees' is the residual from subtracting SOE, collective and Private and self-employed from all employees. It includes those in partial employment, and in home-based enterprises (such as small child-minding nurseries). 5.1.4 consists of part-time income of those engaged in full-time jobs elsewhere. 'Transfer' income consists of gross income from dependents, price subsidies, and a residual. 'Special' income consists of gross gift (or remittance) income, income from boarders, income derived from participation in the survey, sales income and a residual. Note finally that income to cover boarding costs of relatives and friends (*qinyou dahuo fei shouru*) and subsidy to cover survey-participation costs (*jizhang butie*), while included in total income, are excluded from disposable income.

Table A.3 Basic Quantile Data

Table A.3(a) Quantile Income Shares

Quantile	1981	1983	1984	1985	1986	1987	1988	1989	1990
0-10%	5.62	5.78	5.60	5.65	5.51	5.56	5.39	5.30	5.34
0-5%	2.57	2.60	2.52	2.55	2.51	2.55	2.46	2.40	2.42
5-10%	3.05	3.18	3.08	3.10	3.00	3.02	2.94	2.90	2.93
10-20%	6.99	7.10	6.95	7.07	7.04	6.99	6.83	6.76	6.80
20-40%	16.42	16.61	16.36	16.36	16.38	16.34	16.09	15.97	16.07
40-60%	19.20	19.16	19.09	19.06	19.08	19.04	19.01	18.91	18.98
60-80%	22.49	22.12	22.30	22.28	22.35	22.26	22.36	22.37	22.44
80-100%	29.28	29.22	29.70	29.57	29.64	29.80	30.31	30.69	30.37
80-90%	12.95	12.93	13.14	13.08	13.02	13.03	13.23	13.21	13.23
90-100%	16.33	16.30	16.56	16.49	16.63	16.77	17.09	17.48	17.14

Table A.3(b) Cumulative Income Shares

Quantile	1981	1983	1984	1985	1986	1987	1988	1989	1990
0-5%	2.57	2.60	2.52	2.55	2.51	2.55	2.46	2.40	2.42
0-10%	5.62	5.78	5.60	5.65	5.51	5.56	5.39	5.30	5.34
0-20%	12.61	12.89	12.55	12.72	12.55	12.55	12.22	12.06	12.14
0-40%	29.02	29.49	28.91	29.08	28.93	28.90	28.31	28.03	28.21
0-60%	48.23	48.65	48.00	48.14	48.01	47.94	47.33	46.93	47.19
0-80%	70.72	70.78	70.30	70.43	70.36	70.20	69.69	69.31	69.63
0-90%	83.67	83.70	83.44	83.51	83.37	83.23	82.91	82.52	82.86
0-100%	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Notes: See next page.

Table A.3(c) Quantile Means

Quantile	1981	1983	1984	1985	1986	1987	1988	1989	1990
0-10%	297.2	324.1	358.0	399.5	417.2	450.7	443.4	422.1	459.9
0-5%	272.0	292.7	322.2	360.6	380.5	412.4	404.0	382.9	416.3
5-10%	322.4	355.2	393.9	438.4	454.0	489.0	482.9	461.3	503.5
10-20%	369.3	397.8	444.6	499.4	533.1	566.2	562.0	538.6	585.3
20-40%	434.0	465.2	523.0	578.1	619.9	661.9	661.7	636.2	691.5
40-60%	507.6	536.8	610.5	673.5	722.4	771.3	781.9	753.3	816.7
60-80%	594.5	619.7	713.2	787.3	845.9	901.7	919.6	891.4	965.7
80-100%	774.1	818.5	949.5	1045.2	1122.3	1207.0	1247.4	1223.7	1310.4
80-90%	684.8	724.2	840.0	924.7	985.6	1055.4	1087.8	1052.4	1141.7
90-100%	863.4	912.9	1059.0	1165.7	1259.0	1358.7	1407.2	1395.2	1479.0

Table A.3(d) Conditional Means

Quantile	1981	1983	1984	1985	1986	1987	1988	1989	1990
0-5%	272.0	291.4	322.2	360.6	380.4	412.4	404.0	382.9	416.5
0-10%	297.2	324.1	358.0	399.5	417.2	450.7	443.5	422.2	460.2
0-20%	333.3	360.9	401.3	449.4	475.2	508.5	502.8	480.4	522.9
0-40%	383.6	413.1	462.2	513.8	547.5	585.2	582.3	558.4	607.4
0-60%	424.9	454.3	511.6	567.0	605.8	647.2	648.9	623.4	677.3
0-80%	467.3	495.7	562.0	622.1	665.9	710.8	716.6	690.4	749.5
0-90%	491.5	521.0	592.9	655.7	701.4	749.1	757.8	730.6	792.8
0-100%	528.7	560.2	639.5	706.7	757.1	810.1	822.6	796.9	861.1

Notes: 1. See Table A.1 for details of sources.

2. The quantile groups are defined over households, not over individuals. For the bottom 5% of 1987, 1988 and 1989, see Appendix A, point 5. The original income shares, for the bottom 5.2% were, respectively, 2.66, 2.56, and 2.50 for 0-5%, and 2.91, 2.84 and 2.79 for 5-10%. The figures given here are upper bounds of the share of the bottom 5%. Lower bounds are 2.52, 2.44 and 2.38, resulting in quantile means of 408.4, 401.2 and 379.5 for the bottom 5%.

3. Means for 0-5 and 5-10% for 1981 and for 80-90 and 90-100% for 1981 to 1984 are unavailable in the original data and have been estimated, using the method given in Appendix A, point 4. (For the absence of 1982 data, see Appendix A, point 2.)

4. Quantile means and conditional means (generalized Lorenz coordinates divided by p=F(y)) are given in terms of 1985 prices: see Table 1 for source of price index.

Table A.4 Income Decomposition by Source for Mean Income

INDEX	NAME	1981	1983	1984	1985	1986	1987	1988	1989
0	TOTAL	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1	SOE	79.5	79.0	80.1	73.2	71.8	72.4	73.1	73.3
1.1	Regular	57.6	55.7	53.6	45.6	46.5	44.0	38.0	35.5
1.2	Bonuses	8.4	9.4	11.3	10.9	7.7	11.7	13.8	15.0
1.2.1	Other	.	.	.	.	.	10.7	12.2	12.8
1.2.2	Floating	.	.	.	.	.	0.7	0.8	1.1
1.2.3	Contract	.	.	.	.	.	0.4	0.8	1.0
1.3	Subsidies	.	8.3	8.1	9.5	9.6	9.8	14.2	15.5
1.3.1	Allocated	.	.	.	.	.	.	8.8	9.2
1.3.2	Unallocated	.	.	.	.	.	.	5.4	6.3
1.4	Other	13.4	5.5	7.0	7.2	7.9	6.9	7.1	7.3
1.4.1	Allocated	9.9	1.4	2.2	1.8	2.8	1.5	1.6	1.5
1.4.2	Unallocated	3.6	4.1	4.8	5.4	5.2	5.4	5.5	5.8
2	COLLECTIVE	18.5	18.7	18.6	16.6	17.3	16.1	15.8	14.8
2.1	Regular	.	.	.	10.4	11.7	9.8	8.3	7.2
2.2	Bonuses	1.9	2.4	2.7	2.7	2.0	3.1	3.7	3.9
2.2.1	Other	.	.	.	.	.	2.4	2.6	2.7
2.2.2	Floating	.	.	.	.	.	0.3	0.3	0.4
2.2.3	Contract	.	.	.	.	.	0.4	0.8	0.8
2.3	Subsidies	.	.	.	1.9	1.7	1.6	2.3	2.3
2.3.1	Allocated	.	.	.	.	.	.	1.4	1.3
2.3.2	Unallocated	.	.	.	.	.	.	0.9	1.0
2.4	Other	16.6	16.4	15.9	1.5	2.0	1.5	1.5	1.4
2.4.1	Allocated	15.7	15.4	14.8	0.3	0.8	0.3	0.3	0.3
2.4.2	Unallocated	0.8	1.0	1.1	1.2	1.2	1.2	1.2	1.2
3	PENSION	.	.	.	7.5	7.5	9.0	8.9	8.9
4	NET GIFTS	-6.0	-5.2	-5.3	-4.2	-4.7	-5.2	-4.6	-4.2
5	OTHER	8.0	7.5	6.5	6.8	8.1	7.7	6.9	7.3
5.1	Labour	1.6	1.9	1.9	2.6	3.2	2.8	4.6	4.9
5.1.1	Private & self-emp'd	0.0	0.1	0.1	1.0	1.5	1.2	1.5	1.6
5.1.2	Other employee	.	.	.	.	.	.	0.3	0.3
5.1.3	Retired	.	.	.	.	.	.	1.2	1.2
5.1.4	Other	1.5	1.9	1.8	1.5	1.7	1.6	1.6	1.8
5.2	Non-labour	.	0.6	0.4	0.7	0.8	0.8	2.3	2.3
5.2.1	Sales	.	0.6	0.4	0.7	0.8	0.8	0.8	0.7
5.2.2	Property	.	.	.	.	.	.	0.7	1.0
5.2.3	Transfer	.	.	.	.	.	.	0.3	0.2
5.2.4	Special	.	.	.	.	.	.	0.6	0.4
5.3	Unspecified	6.4	5.0	4.3	3.6	4.1	4.2	.	.

Notes: See Table A.2 for details and Appendix A, point 7 for further details.

Appendix B The SSB aggregation procedure

This Appendix addresses the biases introduced by the two-stage aggregation procedure outlined in III.3. It will be useful to distinguish between three families of inequality measures, each distinguished by the method by which the quantile means used for their calculation are obtained. The members of each of the families are the conventional inequality indices, satisfying the properties set out in II.2.

- V_1 is the family of *correctly aggregated* inequality measures, that is, the family of measures computed on the basis on quantile groups formed on the basis of a ranking of households by income.
- V_2 is the family of *one-stage aggregated* inequality measures, that is, the family of measures computed on the basis of aggregating households into **urban unit** (i.e., city or town) quantile groups, and then ranking these on the basis of mean income to form overall or all-China quantile groups.
- V_3 is the family of *two-stage aggregated* inequality measures, that is, the family of measures based on the aggregation procedure used by the SSB since 1985. (Recall that under the two-stage aggregation procedure, the kth all-China quantile mean is the mean of all the urban units' kth quantile group means.)

With these definitions, one can show that, for any inequality measure, R, in the three families,

$$R_{V_1} \geq R_{V_2} \geq R_{V_3} \tag{B.1}$$

The two families, V_1 and V_2, are based on aggregation procedures which give different rankings of households. Similarly, the two families, V_2 and V_3, are based on aggregation procedures which give different rankings of urban-unit quantile group means. To cover both of these two cases, let 'income unit' refer to a household in the context of comparing V_1 and V_2, and to an urban-unit quantile group mean in the context of comparing V_2 and V_3. Then the only income units which will change quantiles in going from correct aggregation to one-stage aggregation or from one-stage aggregation to two-stage aggregation will be richer income units moving into lower quantile groups or poorer income units moving into higher quantile groups. Such moves have the same effect as rank-preserving progressive transfers between the quantile groups, and thus reduce inequality.[70]

---

70. The result is equivalent to saying that the distribution based on the two-stage aggregation will have Lorenz dominance over that based on the one-stage aggregation which in turn will have Lorenz dominance over that based on the correct aggregation.

Thus the SSB procedure underestimates twice over. Which source of underestimation is the more important and under which conditions will the problem be more serious? To ease the exposition I simplify to the case in which each urban unit has the same number of households surveyed and all households are of the same size, so that there is no need to worry about weighting. There are assumed to be N urban units, and K equally-sized quantiles (both for each urban unit and for China as a whole). Define Q_p^j, with mean μ_p^j, to be the pth all-China quantile group, where the superscript, j=1,2 or 3, indicates which aggregation procedure has been used. All quantiles are arranged in ascending order of mean income. Define $\max(Q_p^j)$ to be the maximum income unit in the quantile, and $\min(Q_p^j)$ the minimum. Then from the explanation given for (B.1) it follows that, for any R, a necessary condition that $R_{V1}>R_{V2}$ is that, for some p and q,

$$\min(Q_q^2) < \max(Q_p^2) \text{ where } p<q \tag{B.2}$$

Now let Q_{km} be the kth quantile group of the mth urban unit with mean μ_{km}, maximum income $\max(Q_{km})$ and minimum income $\min(Q_{km})$. Let

$$Q_{km} \subset Q_p^2 \text{ and } Q_{no} \subset Q_q^2, \text{ where } p<q \tag{B.3}$$

which implies

$$\mu_{km} \leq \mu_{no}. \tag{B.4}$$

If it is nevertheless the case that

$$\min(Q_{no}) < \max(Q_{km}) \tag{B.5}$$

then (B.2) will hold. But this suggests that the underestimation implied by the one-stage aggregation will not be serious. Underestimation arises from having high income households in low urban-unit quantile groups. But such households cannot be too high, or too many, or it would not be possible for their quantile group to be counted as a low one.

The second stage of the aggregation procedure is more serious, however, since we no longer have the restriction captured by (B.3) and (B.4) that, if one urban unit quantile group mean is higher than another, the former must be ranked no lower than the latter in the ordering from which the all-China quantile group means are obtained. Since the two-stage procedure aggregates urban-unit quantile means into the same quantile group regardless of differences in mean value, one would expect the inequality measure resulting from such a procedure to be similar in size to the average within-urban-unit inequality component of a decomposable index. However, the aggregation introduces complications: how can one decompose without the urban-unit data? To negotiate this problem, a new family of indices, V_4, is introduced. Let

$$\alpha_{ki} = \frac{\mu_{ki}}{\mu_i} \tag{B.6}$$

where μ_i is mean of the ith urban unit. Then V_4 is the family of inequality measures based on the ranking in ascending order of and aggregation over α_{ki}, the mean-normalized urban-unit quantile group means. As will be illustrated later, any decomposable index, R_{V4}, will approximate the average within-urban-unit inequality component of R obtained from decomposing the urban-unit level quantile data.

If each urban unit has the same Lorenz curve, then for each k and all i $\alpha_{ki}=\alpha_k$, so, for any R, $R_{V3}=R_{V4}$. If Lorenz curves differ, no clear cut result emerges, but it would seem likely that $R_{V3}\leq R_{V4}$. To see this, note that the kth quantile all-China mean using the fourth method will be

$$\mu_k^4 = \frac{1}{N}\sum_{ki\in Q_k^4} \alpha_{ki} \tag{B.7}$$

where the k in the composite index ki may take on different values for different i. This is also the kth quantile all-China relative mean, since the all-China mean (averaging over all the α_{ki}) will be unity. Assuming that there is not too strong a correlation between mean income and inequality, one can approximate and write

$$\mu_k^4 \approx \frac{1}{N\mu}\sum_{ki\in Q_k^4} \mu_{ki} \tag{B.8}$$

where μ is the all-China mean (the average of the μ_i). Using the SSB aggregation procedure, the kth quantile all-China relative mean is

$$\frac{\mu_k^3}{\mu} = \frac{1}{N\mu}\sum_{i=1}^{N} \mu_{ki} \tag{B.9}$$

Since, by the same argument used to establish (B.1), (B.9) will differ from (B.8) by a series of progressive transfers, one obtains, on the assumption that the approximation in (B.8) is a good one, the result $R_{V3}\leq R_{V4}$.

Hence the SSB aggregation procedure will completely ignore between-urban-unit inequality and, unless inequality is constant across urban units, will even tend to underestimate average within-urban-unit inequality.

The provincial data sets used in the next chapter illustrate the downward biases involved.

Table B.1 uses data from the combined sample for 1988 to 1990 (Chapter Five, I.1 gives more detail). SSB procedure is closely followed: the eight unequally sized quantiles used by the SSB are employed, and city and town observations are reweighted in line with SSB conventions (see III.3). (At all stages of aggregation, household size is used to weight the per capita income observations.) Ginis and Theil indices are given based on six different methods of calculation. In the first, no aggregation at all is undertaken. In the second, each urban unit's observations are aggregated into quantiles, and these are used without further aggregation (giving what will be called the 'urban unit' family of measures). In the third (defining family one), overall quantiles are created in the conventional way. In the fourth (defining family two), overall quantiles are created using the urban unit quantiles. In the fifth (defining family three), overall quantiles are created using the SSB procedure. In the sixth and final (defining family four), overall quantiles are created from mean-normalized urban-unit quantiles. The following conclusions emerge:

• Consistent with (B.1), the family-three measures are smaller than the family-two measures which are smaller than the family-one measures. The latter are also less than the urban-unit measures (though this is not a matter of necessity), which are in turn (as a matter of necessity) less than the disaggregated measures.

• As predicted above, little downward bias is inflicted by using family-two rather than family-one measures: in all three years, the family-two Ginis are more than 99% of the family-one Ginis. Much more bias is inflicted by using family-three measures: the family-three Ginis are 74, 78, and 71% of the family-one Ginis in the three years.

• The family-four Theil closely approximates the within-inequality component of the urban-unit Theil: the difference is less than 1% in 1988 and 1990 and less than 5% in 1989. This illustrates the claim that the family four indices provide an approximate measure of aggregated average within-unit inequality.

• The family-three and family-four measures are close: using the Gini, the difference is 5% or less each year. But in each year, the family three measure is slightly below the family-four measure, as predicted above.

So, according to these provincial samples, the SSB aggregation procedure results in underestimation even of aggregated average within-urban-unit inequality, though only slightly - of the order of some 5% - and, due to its neglecting between-urban-unit inequality, more serious underestimation of total inequality - of the order of some 25%.

Table B.1 Aggregation-induced Bias: an example using Chapter Five data

Family	Gini	Theil	Within	Between
1988				
Disaggregated	0.1939	0.0612	0.0406	0.0205
Urban unit	0.1917	0.0588	0.0383	0.0206
One	0.1880	0.0575	.	.
Two	0.1860	0.0561	.	.
Three	0.1424	0.0327	.	.
Four	0.1462	0.0374	.	.
1989				
Disaggregated	0.1971	0.0675	0.0507	0.0168
Urban unit	0.1946	0.0619	0.0452	0.0167
One	0.1909	0.0593	.	.
Two	0.1888	0.0580	.	.
Three	0.1521	0.0380	.	.
Four	0.1583	0.0430	.	.
1990				
Disaggregated	0.1889	0.0572	0.0370	0.0202
Urban unit	0.1866	0.0552	0.0349	0.0203
One	0.1833	0.0541	.	.
Two	0.1809	0.0526	.	.
Three	0.1346	0.0291	.	.
Four	0.1414	0.0359	.	.

Notes: For the definitions of the various families see the text of Appendix B. For more details on the data see Chapter Five, I.1. 'Within' and 'Between' refer to the average within-urban-unit and the between urban-unit components of the Theil index. The formula used for the Theil decomposition is

$$
\begin{aligned}
T &= \frac{1}{N}\sum_i \sum_j n_{ij}\frac{\mu_{ij}}{\mu}\log\left(\frac{\mu_{ij}}{\mu}\right) \\
&= \frac{1}{N}\sum_i n_i\frac{\mu_i}{\mu}\left\{\frac{1}{n_i}\sum_j n_{ij}\frac{\mu_{ij}}{\mu_i}\log\left(\frac{\mu_{ij}}{\mu_i}\right)\right\} + \frac{1}{N}\sum_i n_i\frac{\mu_i}{\mu}\log\left(\frac{\mu_i}{\mu}\right) \\
&= T_w + T_b
\end{aligned}
$$

where μ is the overall mean of the total population of size N, μ_i the mean of group i with population n_i and μ_{ij} the mean income of the number (n_{ij}) in group i's jth quantile group. T is the Theil index, T_w the 'within' component and T_b the 'between' component.

Chapter Five The Distribution of Cash and Subsidy Income in Two Chinese Provinces, 1987-90

I Introduction

In the previous chapter, urban China's income distribution was examined using aggregated, nationwide data from 1981 to 1990. This chapter utilizes a disaggregated data set from the urban areas of two of China's most populous provinces, Sichuan and Liaoning, for the four years of 1987 to 1990. So in terms of coverage of both provinces and years, this chapter's data set is more limited in scope. On the other hand, the fact that it is disaggregated makes it more susceptible to statistical analysis than the aggregated data used in the previous chapter. This chapter's data set also contains information which can be used to calculate the income flows implicit in the government scheme of food rationing, which, until its recent dismantling, has been an important source of income for urban residents. The first of these advantages enables this data set to be used as an application of the statistical tools developed in the first part of this thesis. The second enables us to get at least some insight into the important role played in China, as in other socialist economies, by non-cash income.

Of course, the two provinces of Liaoning and Sichuan, whose locations can be seen from the map following, cannot be regarded as representative of urban China as a whole. Nor has either been at the forefront of urban reform. Nevertheless, this data set is an important one, not only because of the size of the two provinces, but also because it would appear to be the only disaggregated, longitudinal (but not panel) data set for urban China currently available for academic study.

The structure of the chapter is as follows. The next section gives an introduction to the data and to the two provinces. The third section focuses on cash income, and re-examines the question taken up in the previous chapter of trends in inequality, and their implications for welfare and poverty. Data from both provinces are pooled and used to examine changes over time. The fourth section extends the analysis to examine both the income implicit in urban residents' entitlements to cheap food and the payments issued by the government to compensate residents for the recent dismantling of these subsidy entitlements. The fifth and final section concludes. An Appendix summarizes the distributions of the provincial samples with tables based on decile income shares and means.

Figure 1 Liaoning and Sichuan

II Background

II.1 The data

Even for the two provinces and four years which it covers, the data set this chapter analyzes is only a sub-sample, approximately one-third the size, of the annual household survey conducted by China's State Statistical Bureau (SSB).[1] The reason is that, due to financial and computing constraints, the SSB collects centrally in disaggregated form only a fraction of the data its provincial branches collect. From now on, I refer to the provincial samples analyzed in this chapter as the 'sub-samples' and the SSB samples from which they are drawn as the 'full samples' or simply the 'samples'.

Table 1 presents information concerning sample size, for both the sub-samples and the full samples. Its figures can be compared with those of Table 2, which presents population figures for the urban areas of the two provinces using the various definitions of 'urban' presented in the last chapter (see III.1). The full SSB Liaoning sample covers some 2,500 households and constitutes around 8% of the nationwide sample, making it the largest sample from any one province. The full Sichuan sample (the third largest after Liaoning and Jiangsu) covers some 2,000 households, around 6.5% of the nationwide sample. The annual sub-samples for Liaoning include on average 700 households, three-tenths of the total surveyed, while those for Sichuan cover 900, which is closer to half the full sample size. Although the sub-sample sizes are more or less constant over the four years, note from Table 1 that the number of city and towns surveyed jumps from 7 in Liaoning and 9 in Sichuan in 1987 to 15 and 17 respectively in 1988. This has quite a large impact on sample results as we will see.[2]

Table 2 shows Sichuan to be China's most populous province with a population of over 100 million. However, the province has a relatively small proportion of urban to rural dwellers. Liaoning on the other hand is a medium-sized province with a total population of around 40 million, but it is heavily urbanized, making it one of the largest provinces in terms of its urban population. How Liaoning and Sichuan compare with each other and with the other Chinese provinces in terms of size of urban population depends on the definition of 'urban' used, a subject

1. In fact, the data set available also incorporates 1986. But see footnote 3 for why this year is not analyzed.

2. These are realized sizes for the sub-samples, after my own cleaning. Although the data, having already been processed by the SSB, did not suffer from internal inconsistencies, there were a handful of cases of replication of observations. Also all households from one city in the Liaoning 1989 sub-sample had unbelievably low incomes and were therefore excluded.

discussed in detail in Chapter Four, III.1. Using the SSB survey definition - *hukou* (non-agricultural-registered) households in SSB-designated urban areas - Liaoning has China's largest urban population with almost 16 million residents. Sichuan is fourth with 12.5 million. Using alternative definitions of 'urban', Sichuan and especially Liaoning slip somewhat in their ranking. For example, if the official definition of urban is used, Sichuan becomes the fifth largest province, and Sichuan the seventh, indicating that both provinces have relatively high ratios of *hukou* to non-*hukou* urban households.

Another feature evident from Table 2 is that Sichuan has a much higher proportion of town- to city-dwellers than Liaoning. This is true whichever definition of 'urban' is used. Using the survey definition, 18.4% of urban residents are town-dwellers in Liaoning, but 39.1% in Sichuan. As we saw in the previous chapter (III.2.2), the full SSB sample suffers from an over-representation of city-dwellers vis-a-vis town dwellers. In 1987, for example, 18.6% of the full SSB Sichuan sample was town-based and 12.2% of the Liaoning sample. The ratios for the sub-samples are similar for Liaoning - around 15% - but higher for Sichuan - between 33 and 41%. So in this respect the sub-samples are more representative than the full samples.[3] (See Chapter Four, III.2.1 on the importance of the town-city distinction: city households are on average richer and smaller than their town counterparts.)

Table 1 Sub-sample and Sample Sizes for Sichuan and Liaoning, 1987-1990

		Liaoning			Sichuan		
		Total	City	Town	Total	City	Town
1987	Sub-sample	700	600 (6)	100 (1)	900	600 (6)	300 (3)
	Sample	2450	2150	300	2150	1750	400
1988	Sub-sample	700	600 (13)	100 (2)	848	548 (11)	300 (6)
	Sample	2450	2200	250	2100	1700	400
1989	Sub-sample	648	548 (11)	100 (2)	899	550 (11)	349 (7)
	Sample	2450	2200	250	2100	1700	400
1990	Sub-sample	699	599 (12)	100 (2)	800	550 (11)	250 (5)

Notes: For sources for SSB figures up to 1989, see Chapter Four, Table A1. The 1990 SSB figures come from the Liaoning and Sichuan 1991 yearbooks (LTN and STN). Figures in brackets in sub-sample columns indicate, respectively, number of towns and cities sampled.

3. As mentioned in footnote 1, the data set includes 1986. However, for this year for Liaoning the town-urban ratio is 25%, ten percentage points higher than for the other three years. To avoid the need for re-weighting the data to take this difference into account, 1986 is not analyzed in this paper. Dropping this year from the analysis is not a particularly heavy price to pay to improve comparability over time since Chapter Four's analysis suggested little or no change in urban inequality between 1986 and 1987: the big changes come both before 1986 (in 1984 and 1985) and after 1986 (in 1988 and 1989).

Table 2 The Urban Populations of Liaoning and Sichuan under Different Definitions

	Liaoning			Sichuan		
	Total (m)	ranking	%	Total (m)	ranking	%
PROVINCE	39.67	12	3.5	105.91	1	9.4
A. OFFICIAL URBAN						
Urban	29.25	7	4.9	35.09	5	5.9
Of which *hukou*	15.86	1	8.0	12.53	4	6.3
City	20.37	3	6.1	19.60	4	5.9
Of which *hukou*	13.39	1	9.4	7.63	2	5.3
Town	8.88	12	3.3	15.49	5	5.9
Of which *hukou*	2.47	7	4.4	4.90	2	8.8
B. CENSUS URBAN						
Urban	20.41	4	6.9	20.89	3	7.1
Of which *hukou*	15.11	1	8.1	12.27	3	6.6
City	17.02	1	8.0	15.00	4	7.1
Of which *hukou*	12.86	1	9.8	7.11	3	5.4
Town	3.38	10	4.1	5.87	4	7.1
Of which *hukou*	2.25	10	4.0	5.16	1	9.2
C. *HUKOU*	16.59	1	7.5	14.76	2	6.7

Notes: 'Total' gives the population in millions, 'Ranking' the ranking among all of China's provinces, '%' the percentage of nationwide population. All three figures are given in terms of the definition provided in each row: see the text for more detail on these. The source is National Population Census Office (1991b).

An analysis of the full SSB household survey has already been presented in Chapter Four (see III.2). In the remainder of this sub-section, the focus is on the question of how representative the provincial sub-samples are of the full samples from which they are drawn. This is a necessary task since the method by which the sub-sample is culled is, unfortunately, not publicly available. (All that is known is that selection takes place at the city and town level.) A range of comparisons between sub-sample and full sample for both provinces is given in Table 3. Comparisons are made for 1987 since, for this year, a full-sample decomposition of mean income by city and town is available. And comparisons are given for 1989 and 1990 since for these two years information is available on the full SSB survey from Sichuan and Liaoning's 1991 yearbooks (LTN and STN).

Table 3 Comparison Between Sub-samples and Full Samples

	1987		1989		1990	
	Sub-sample	**Sample**	**Sub-sample**	**Sample**	**Sub-sample**	**Sample**
Liaoning						
Mean p.c. income	960.0	891.3	1394.0	1289.2	1465.9	1398.9
City	987.7	942.6	1439.5	1355.5	1530.1	1458.7
Town	797.0	757.6				
Gini of p.c. income			0.161	0.172	0.155	0.163
Mean h/h size	3.5	3.7	3.4	3.4	3.3	3.4
% workers	61.2	58.6	61.0	60.0	60.9	60.4
SOE/collective ratio	1.65	1.90	2.24	2.33		
Sichuan						
Mean p.c. income	876.2	899.2	1264.7	1226.3	1424.1	1354.4
City	926.7	952.2	1314.9	1319.9		
Town	779.7	796.0				
Mean h/h size	3.5	3.6	3.4	3.4	3.3	3.3
% workers	56.6	55.9	54.9	55.9	56.4	56.6
SOE/collective ratio	3.42	3.30	3.51	3.46		

Notes: Sources for full-sample information are: SSB(1988s, 1990s), LTN and STN. For calculation of Ginis, see footnote 4. Income in measured in current yuan per annum. '% workers' gives the percentage of household members employed. The 'SOE/collective ratio' gives the ratio of the average number of SOE to the average number of collective employees.

Most of the variables in Table 3 do indicate that the sub-samples are representative. The figures on mean household size, the proportion of the household in employment and the state-owned enterprise (SOE) to collective-enterprise employees ratio are all quite similar. However, the income variables give some cause for concern. For Liaoning, the sub-sample per capita income means are consistently higher than those from the full sample. This is not true for Sichuan, but the 1987 figures, which show higher city and town means for the sub-sample, suggest that this is only because the sub-sample has a higher proportion of (on average poorer) town dwellers than the full sample.

For Liaoning 1988 and 1989, a comparison can also be made of the dispersion within the sub-sample and that within the full sample. For both years, the Gini coefficient is higher for the full sample, by approximately .01.[4] Combining these two findings suggests that the sub-samples are not selected entirely at random, but rather that outliers, especially those at the lower end of the distribution, may well be omitted in the process by which the sub-sample is selected. So, although this brief inspection reveals no gross defects in the two provincial sub-samples, their small size and the evidence that they are not fully representative even of the full SSB provincial samples should lead us to be wary of placing excessive weight on results obtained from their analysis, a conclusion which is reinforced when combined with the assessment of the full SSB sample presented in the previous chapter.

II.2 Liaoning and Sichuan

That Liaoning has a higher degree of urbanization than Sichuan has already been mentioned. Consistent with this, Liaoning is a far more industrialized province than Sichuan. Only the three municipalities have a equal or higher industrial share and lower agricultural share in total GDP than Liaoning, whose respective shares were, in 1989, 62 and 18%. Sichuan is far more typical in this regard with both shares around 40% (see SSB, 1991y). Differences between the two provinces' urban economies are also evident from Table 4 below, which, based on sub-sample data, shows the different occupational structures prevailing in 1987. Whereas Liaoning has half its employees in manufacturing, Sichuan has four-tenths, and a higher proportion employed as government cadres (12 compared to 5%) and in commerce (16 compared to 11%).

Table 4 Occupational Structure, 1987 (%)

Sector	Liaoning	Sichuan
Cadres (gov't officials)	5.1	12.3
Commerce	10.7	16.1
Construction	6.9	3.6
Ed'n & research	8.8	7.8
Manufacturing	50.2	41.4
Power	0.8	1.0
Primary	0.7	3.1
Services	10.6	8.9
Transport	5.7	5.1
Other	0.6	0.7

Notes: The figures give the percentage of employees in the various occupations. Agriculture is included in the category of 'other'. The source is the SSB sub-samples for Liaoning and Sichuan, 1987.

4. The Gini for the full sample is calculated using income-class means calculated from the full sample and presented in LTN. The Gini for the sub-sample is calculated by first dividing the sub-sample into the same income classes as those given in LTN, then obtaining income-class means and then using these to calculate the Ginis. Hence the Ginis for the sub-sample given in this table will differ from those given later in the text and based on disaggregated data.

Table 5 below gives the employment structure by ownership category. In both provinces the SOE sector predominates, but more so in Sichuan. Note also from this table the large gap in both provinces between the average SOE and collective wage. The former is 4/3 times the latter in Liaoning and 3/2 times in Sichuan.

Table 5 Labour Force, by Ownership, Proportion and Wage, 1990

Sector	Liaoning		Sichuan	
	% of emp	Ave wage	% of emp	Ave wage
SOE	74.4	2058.0	82.7	1905.6
Collective	25.6	1515.2	16.4	1271.9
Private & self	0.1	1859.8	0.9	1187.5

Notes: The column headed '% of emp' gives the proportion employed in any particular sector. The average wage is per employee per annum and is given in current yuan. 'Private & self' includes those employed in the private sector and the self-employed and the residual 'other': see the notes to Table A.2, Chapter Four for definitional details. The source is the SSB sub-samples for Liaoning and Sichuan, 1990.

Figures 2a and 2b overpage are occupational 'trees'. They provide more detail on the importance of various sources of employment and income in the two provinces. In both provinces, the largest single block of households are 'SOE households': those households whose working members are employed solely in the state-owned sector. This accounts for just over half of all households in Sichuan and just under half in Liaoning. The next largest category is not 'collective' but 'mixed' households: those with at least one worker in the SOE sector, and at least one in the collective sector. There is a sexual division of labour: it is fairly common to have the male working in the SOE sector and the female in the collective sector. 18% of all Liaoning households and 12% of all Sichuan households are thus characterized. Another important category of households are those whose main income source is the pension: this includes 10% of all households in Sichuan and 5% in Liaoning. Pensioners are not particularly poor: the average annual pension in Liaoning in 1990 was 1298 yuan and in Sichuan an almost identical 1296. This is roughly equal to the average wage of a collective-sector worker in Sichuan.

Consistent with having a lower proportion of pension-reliant households, Liaoning has a slightly younger age-profile than Sichuan. However, the difference is small: mean-age is around 31 in Liaoning and 32 in Sichuan. The effects of the one-child policy can be seen in both provinces. In Liaoning, 69% of all households have children (defined as those with an age under 16). Of these, nearly all, 93%, are one-child households. In Sichuan, only 56% of households have children. 91% of these are one-child households.[5] As is well-known, rural areas have been treated with greater leniency than urban with regards to the one-child policy.

5. All these figures refer to 1990.

Figure 2 Occupation Trees for Liaoning and Sichuan, 1990

Figure 2a Liaoning

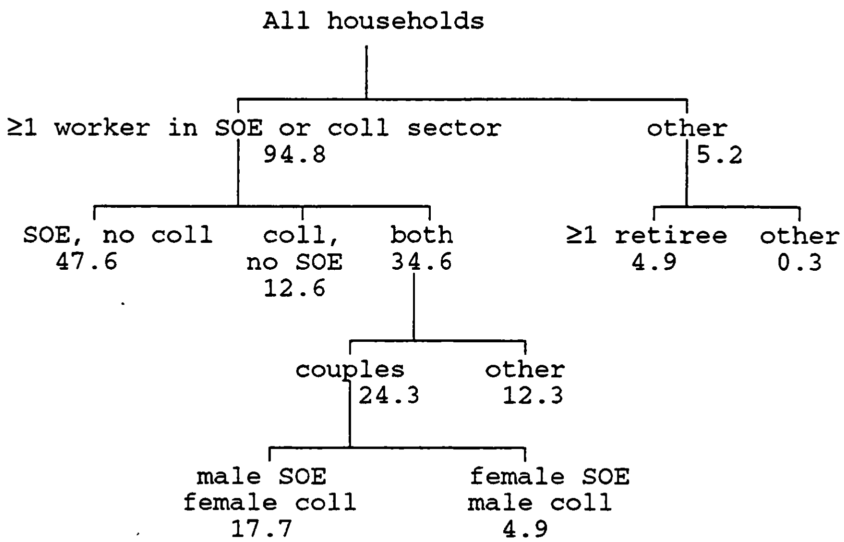
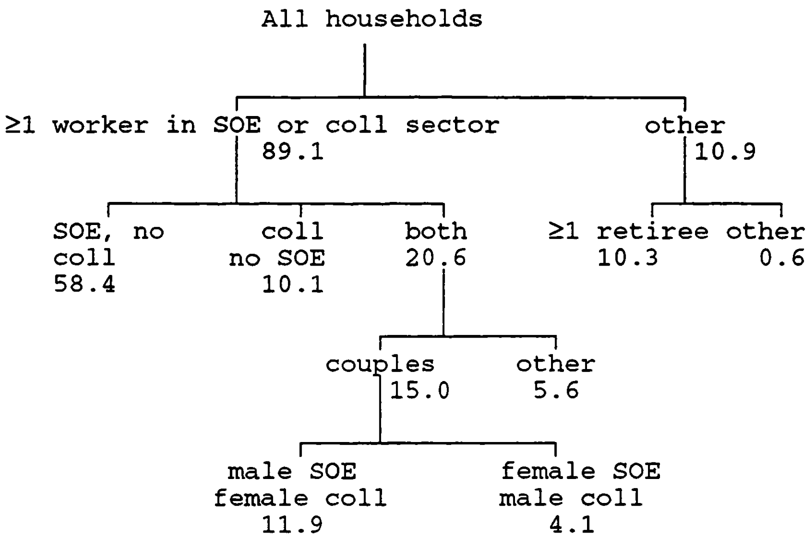


Figure 2b Sichuan



Notes: All figures are percentages. 'coll' stands for collective. Couples were defined to be households whose head was adult and male (female), and whose second member was female (male). The source is the SSB sub-samples for Liaoning and Sichuan, 1990.

III The distribution of cash income

III.1 Mean income

As in the previous chapter, cash income is defined as annual disposable income (*shenghuofei shouru*), which is basically gross income minus net remittances or gifts. For a fuller definition, see Section III.2 of Chapter Four. Income is defined on a per capita basis, and all means and measures of dispersion are calculated by weighting each household's per capita income by its household size.

Given Liaoning's greater degree of industrialization, one might expect it to have a much higher income level than Sichuan. Certainly it has a much higher provincial GDP per capita: 1,977 yuan in 1989, again the highest outside the municipalities and twice as high as Sichuan's 813 yuan, the fifth lowest of all the provinces (see SSB, 1991y). But this refers to rural and urban combined. The two provinces' urban per capita disposable incomes differ by much less. As Table 6 shows, mean per capita income in urban Liaoning is just some 10% above that in urban Sichuan.

To properly compare purchasing power in the two provinces, price differences need to be taken into account. A Fisher cost-of-living index was calculated for 1987 using sub-sample data on consumption quantities and expenditures.[6] Taking prices in Liaoning for this year as the base gave an index value for Sichuan in the same year of 0.872, indicating that prices are substantially lower in Sichuan. This figure can be compared with the very similar ratio of the Liaoning to the Sichuan cost-of-living index, taking all-China urban prices as the base, given in Howes and Lanjouw (1991) for the same year of 1987, using aggregated full-sample SSB data, of 0.877. Also from this source, one can note that in 1987 only Guangdong had a higher price index than Liaoning and only Jiangxi and Henan a lower price index than Sichuan. Although urban China displays relatively little price disparity for its population and size, the two provinces of Liaoning and Sichuan happen to be on either side of such spread as there is and it is important therefore that differences in prices between the two provinces be taken into account.

For years other than 1987, incomes were made comparable using the annual provincial inflation indices given in the annual editions of the China Statistical Yearbooks. These indices (which were also used in Chapter Four) are based on a weighted average of free-market and state prices, and, unlike the calculated cost-of-living indices, include the prices of services as well as goods. Adjusted incomes given in yuan at Liaoning 1987 prices are also presented in Table 6, which shows that the two provinces have almost identical mean incomes once prices have been adjusted for. Table 6 also enables us to examine changes over time. Both Liaoning and Sichuan display high nominal growth rates for 1988 and 1989. However, inflation was also high in these two years, leading to little or even negative real income growth. In 1990, inflation was under control, so that, although nominal growth was also lower, real growth rates were higher, especially in Sichuan where they were above 10%. This pattern for the late eighties is similar to that

6. See Howes and Lanjouw (1991) for the required formulae and for detail on the commodity composition of the index.

observed for urban China as a whole in the previous chapter. However, note that while there was negative growth China-wide in 1989, this was avoided in our two provinces. Sichuan, did, however, suffer from negative growth in 1988. Table 6 also gives the ratio of city to town means for the period. The differential of around 20% is again typical for China (see Chapter Four, Table 10). The ratio shows a good deal of year-on-year variation, but displays no trend.

Table 6 Per Capita Cash Income, Nominal and Deflated, and Inflation Indices

	Nominal Income Level	Nominal Income Growth Rate	Inflation Level	Inflation Growth Rate	Deflated Income Level	Deflated Income Growth Rate	City-town mean ratio
Combined							
1987	912.8	.	92.4	.	985.0	.	117.8
1988	1117.1	22.4	112.5	21.8	989.8	0.5	123.7
1989	1319.8	18.1	131.9	17.2	998.4	0.9	113.1
1990	1443.9	9.4	134.2	1.8	1074.8	7.6	122.7
Liaoning							
1987	960.0	.	100.0	.	960.0	.	123.9
1988	1187.4	23.7	119.6	19.6	992.8	3.4	129.6
1989	1394.0	17.4	140.2	17.2	994.5	0.2	124.9
1990	1465.9	5.2	142.0	1.3	1032.4	3.8	140.5
Sichuan							
1987	876.2	.	87.2	.	1004.4	.	118.9
1988	1058.5	20.8	107.2	22.9	987.3	-1.7	123.1
1989	1264.7	19.5	126.3	17.8	1001.3	1.4	110.6
1990	1424.1	12.6	127.9	1.3	1113.0	11.2	120.1

Notes: Deflated incomes are expressed in Liaoning 1987 prices. See text for details on price indices used. 'City-town mean ratio' gives the ratio of the average city to the average town per capita income. The source is the SSB sub-samples for Liaoning and Sichuan.

III.2 Income inequality

To analyze trends in inequality, use is made of two of the dominance criteria given in Chapter One, restricted equality second-order stochastic (or Lorenz) dominance (2OSD) and e-dominance. Both these criteria are defined in relation to bounds, in terms of p, the cumulative proportion, for the restricted Lorenz dominance criterion, and in terms of e, the relative inequality aversion parameter, for the e-dominance criterion. The bounds for p are set at .05 and .95, those for e at .1 and 3. These bounds are narrower than those used to analyze the aggregated data of Chapter Four, where the bounds for p were set (implicitly since the criterion was unrestricted Lorenz dominance) at 0 and 1 and those for e at .1 and 5. These narrower bounds are required as a result of the data being disaggregated. They prevent results being too sensitive to the tails of the distributions (see Chapter One, IV.2 and IV.3).[7] The narrower bounds notwithstanding, both criteria are demanding and a ranking by either of them, but particularly the Lorenz dominance criterion, should command widespread agreement.

Only results from pooling the Liaoning and Sichuan observations are given.[8] The rationale for this is that, first, pooling the two samples obviously increases the sample size and, second, as we will see below, the sample trends are the same for both provinces.

The results are summarized in Tables 7 and 8 below, and are also presented graphically in Figures 3 and 4. Tables 7a and 8a give the sample outcomes. A 'D' indicates dominance by the row of the column (i.e., less inequality in the row year), and a 'DB' by the column of the row. Using the restricted Lorenz dominance criterion, 1987 comes out as more equal than the other three years. 1990 is the second most equal year, while 1988 and 1989 cannot be ranked against each-other. Using the e-dominance criterion to supplement this ordering gives a complete ranking, as 1988 emerges as more equal than 1989. This picture is very similar to that obtained in Chapter Four: equality deteriorates from 1987 to 1989 but then improves in 1990.

Tables 7a and 8a also indicate which of these rankings are statistically inferable at the 5% level. Here we combine the estimators derived in Section II of Chapter Two - using household size as the random weight - with the method for inferring dominance presented in Section III.1 of the same chapter. The test statistic to be calculated is, as in Chapter Two,

$$Z_i = \frac{\hat{\zeta}_i - \hat{\zeta}_i^*}{(\hat{Z}_{ii}/N + \hat{Z}_{ii}^*/N^*)^{\frac{1}{2}}} \tag{1}$$

where ζ_i is a particular dominance curve, estimated at x_i. The caps indicate estimators and Z_{ii} is the variance of the estimator of ζ_i. $N(N^*)$ is the sample size of $F(F^*)$. For different dominance criteria one simply estimates different ζ and Z, using the formulae of Chapter Two, Section II. The

7. As seen in Chapter Three, IV.2.1.1, spuriously large test statistics tend to emerge for the generalized Lorenz curve at very low levels of p. Experimentation revealed similarly large test statistics for the Lorenz curve at both very low and very high values of p. Use of bounds (lying between zero and one) also avoids the need to represent these large test values graphically: test statistics at these extremes can be so large that they render the rest of the graph uninformative.

8. Although for the purpose of inequality analysis incomes do need to be made comparable over time, comparability across the two provinces is required, and obtained using the cost-of-living and inflation indices given in III.1

inference method can be simply summarized: dominance is inferred only if all the test statistics within the exogenously imposed bounds are of the same sign and if the test statistic with the minimum absolute value is greater in absolute value than Z_α, the critical value obtained from the normal distribution for a one-sided test of size α. An asterisk in Table 5 indicates that dominance can be inferred. Using the Lorenz dominance criterion, 1987 emerges as less equal than 1988, but no other rankings are possible. Using the e-dominance criterion, the additional ranking of 1987 over 1989 is also possible.

Tables 7b and 8b show the largest single range - of p and e respectively - over which the test statistic is either greater than Z_α or less than $-Z_\alpha$. Wider bounds are used for this analysis: of 0 and 1 for p - the widest possible - and of .1 and 5 for e. The information of Tables 7b and 8b is pertinent to the extension to the testing methodology using endogenous bounds presented in Chapter Three, IV.2.4. Since that extension was only given in terms of welfare one should be cautious in applying it here. Nevertheless, the fact that 1987 has statistically significant dominance over 1990 from e=.01 to e=2.5 probably does imply dominance by the former population over at least this range.

The overall picture is clear. On the basis of the statistical analysis, we can infer that 1987 is significantly more equal than the other three years. Though there is evidence of further deterioration between 1988 and 1989 and of subsequent improvement in 1990, it is not strong enough for these changes to be regarded as statistically significant.

Table 7 Equality Second-order Stochastic (Lorenz) Dominance, Combined Sample

Table 7a Sample and Statistical Dominance with Exogenous Bounds of .05 and .95

Year	1988	1989	1990
1987	D*	D	D
1988	.	X	DB
1989	.	.	DB

Table 7b Statistical Dominance, Endogenous Bounds

Year	1988	Lgth	Min	Max	1989	Lgth	Min	Max	1990	Lgth	Min	Max
1987	D	0.97	0.03	1	D	0.92	0.08	1	D	0.62	0.12	0.74
1988	.	.	.	.	X	.	.	.	X	.	.	.
1989	.	.	.	.	X	.	.	.	X	.	.	.

Table 8 Equality E-dominance, Combined Sample

Table 8a Sample and Statistical Dominance with Exogenous Bounds of .1 and 3

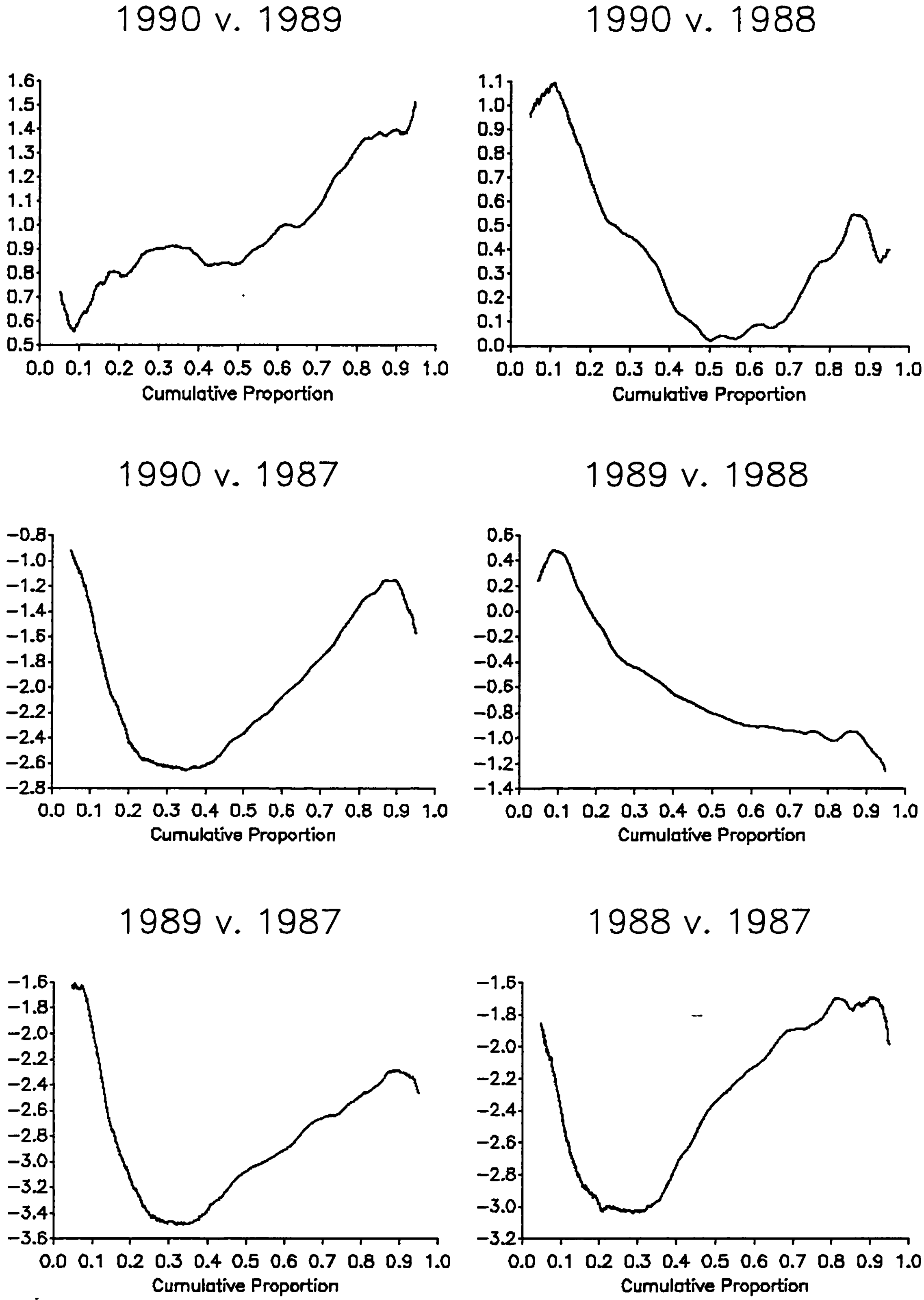
Year	1988	1989	1990
1987	D*	D*	D
1988	.	D	DB
1989	.	.	DB

Table 8b Statistical Dominance, Endogenous Bounds

Year	1988	Lgth	Min	Max	1989	Lgth	Min	Max	1990	Lgth	Min	Max
1987	D	4.2	0.1	4.3	D	4.4	0.1	4.5	D	2.4	0.1	2.5
1988	.	.	.	.	X	.	.	.	X	.	.	.
1989	.	.	.	.	.	.	.	.	X	.	.	.

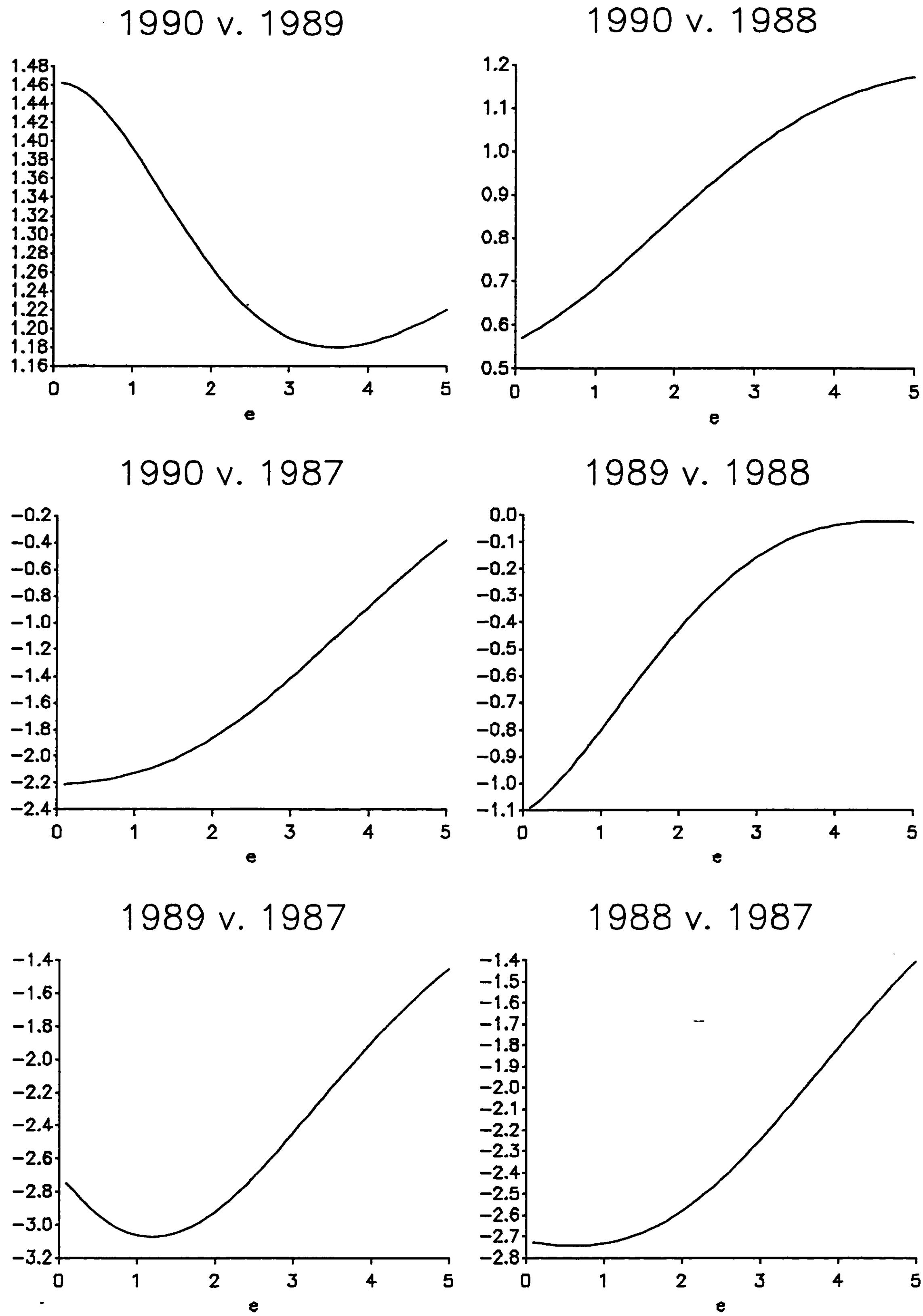
Notes: A 'D' indicates dominance by the row over the column, 'DB' dominance by the column of the row. An 'X' indicates no ranking is possible. An asterisk in Tables 7a or 8a indicates that the dominance is statistically significant at the 5% level. In Tables 7b and 8b the maximum range or length over which one distribution has statistically significant dominance over the other is given, in terms of p for Table 7b and e for Table 8b (this is the variable 'Lgth'). This range may be found anywhere between p=0 and p=1 for 7b, and anywhere between e=0 and e=5 for 8b. 'Min' and 'Max' give the, respectively, minimum and maximum bounds for this range. 'D' and 'DB' have the same meaning in these tables, indicating whether the row or column dominates over the range specified. An 'X' in Table 8b means that there is no point at which the dominance is significant; an 'X' in Table 7b means that there is no range of length greater than .01 at which the dominance is significant (this restriction excludes spuriously large test statistics at either tail). The source is the SSB sub-samples for Liaoning and Sichuan.

Figure 3 Lorenz Curve Test Statistics, Combined Sample



Notes: The vertical axis gives the value of the test statistics on which Table 7 is based. These are calculated using (1) and the appropriate formulae from Chapter Two, II.5: see equations (60) and (62). A positive (negative) value indicates a higher Lorenz curve for the first named (second-named) year for the given value of p. The 'cumulative proportion' (p) is derived from the ordering of each sample in ascending order of per capita income. The source is the SSB sub-samples for Liaoning and Sichuan.

Figure 4 Equality E-dominance Test Statistics, Combined Sample



Notes: The vertical axis gives the value of the test statistics on which Table 8 is based. These are calculated using (1) and the appropriate formulae from Chapter Two, II.2: see equations (13) and (14). A positive (negative) value indicates a higher isoelastic function (calculated using mean-normalized income) for the for the first named (second-named) year at the given value of e, the relative inequality aversion parameter. The source is the SSB sub-samples for Liaoning and Sichuan.

Table 9 Per Capita Cash Income Inequality, 1987-1990, by Province and Place-of-residence

	Year	Gini	Coefficient of Variation	Theil Index	Atkinson Index (e=1)	Atkinson Index(e=2)
Combined						
Urban	1987	0.1714	0.3136	0.0476	0.0475	0.0959
	1988	0.1843	0.3400	0.0556	0.0558	0.1135
	1989	0.1894	0.3773	0.0618	0.0590	0.1168
	1990	0.1824	0.3346	0.0538	0.0536	0.1075
Cities	1987	0.1597	0.2932	0.0411	0.0403	0.0793
	1988	0.1673	0.3098	0.0460	0.0457	0.0921
	1989	0.1695	0.3122	0.0469	0.0466	0.0936
	1990	0.1684	0.3118	0.0462	0.0452	0.0892
Towns	1987	0.1904	0.3493	0.0588	0.0588	0.1191
	1988	0.2072	0.3922	0.0709	0.0687	0.1352
	1989	0.2250	0.5145	0.0961	0.0831	0.1540
	1990	0.2048	0.3758	0.0673	0.0662	0.1306
Liaoning						
Urban	1987	0.1575	0.2852	0.0398	0.0399	0.0805
	1988	0.1630	0.2992	0.0434	0.0434	0.0877
	1989	0.1684	0.3203	0.0477	0.0466	0.0931
	1990	0.1633	0.2984	0.0433	0.0435	0.0880
Cities	1987	0.1479	0.2684	0.0351	0.0350	0.0700
	1988	0.1500	0.2758	0.0370	0.0372	0.0761
	1989	0.1519	0.2791	0.0376	0.0372	0.0741
	1990	0.1489	0.2744	0.0361	0.0356	0.0709
Towns	1987	0.1778	0.3352	0.0526	0.0509	0.0990
	1988	0.1802	0.3733	0.0594	0.0532	0.0971
	1989	0.2120	0.5004	0.0923	0.0771	0.1384
	1990	0.1684	0.3084	0.0460	0.0456	0.0908
Sichuan						
Urban	1987	0.1807	0.3307	0.0528	0.0528	0.1071
	1988	0.2012	0.3708	0.0659	0.0660	0.1337
	1989	0.2041	0.4140	0.0721	0.0681	0.1337
	1990	0.1956	0.3564	0.0614	0.0614	0.1233
Cities	1987	0.1683	0.3089	0.0454	0.0443	0.0866
	1988	0.1846	0.3410	0.0555	0.0548	0.1094
	1989	0.1866	0.3423	0.0564	0.0562	0.1132
	1990	0.1843	0.3376	0.0547	0.0540	0.1070
Towns	1987	0.1907	0.3482	0.0591	0.0599	0.1234
	1988	0.2125	0.3952	0.0737	0.0732	0.1470
	1989	0.2244	0.5131	0.0951	0.0829	0.1556
	1990	0.2043	0.3708	0.0667	0.0669	0.1348

Notes: For notes on the indices, see Table 8 of Chapter Four. The source is the SSB sub-samples for Liaoning and Sichuan.

Table 9 sub-divides the pooled sample both by city and town and by province. It presents, for each sub-division, a number of well-known inequality indices. The pattern observed in the pooled sample - an increase in inequality in 1988 and 1989 but then a fall in 1990 back to its 1988 level - is also observed in each sub-division of the data, almost regardless of measure used. Finally, note the large jump in the Gini of .02 for Sichuan between 1987 and 1988. It is this increase which would seem to be driving the result that 1987 is significantly more equal than the other three years. In addition, in each year, whichever measure is used, and in all three of city, town and all-urban, inequality is lower in Liaoning than Sichuan. Also, within each province, inequality is lower in the cities than in the towns. In all these cases, using the Gini as our measure, the difference is about .03 points.

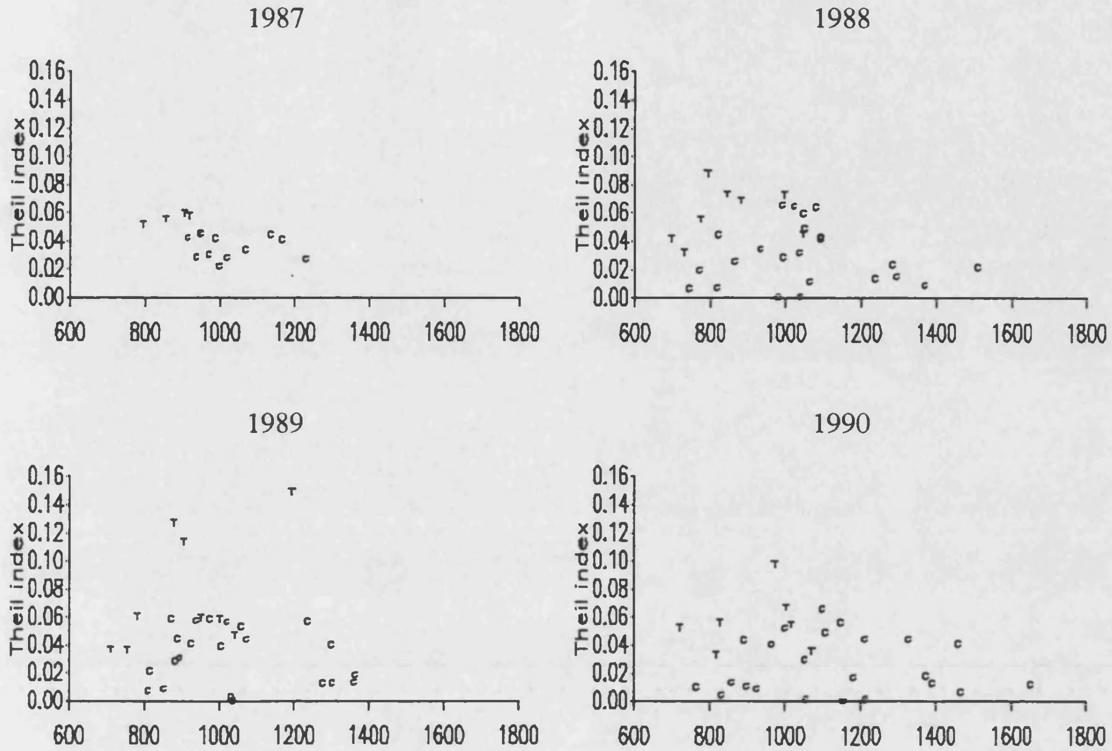
The inference that income has become less equally distributed since 1987 is only valid on the assumption that the different years' samples are randomly drawn. Figure 5 and Table 10 cast some doubt on the validity of this assumption. Figure 5 plots, for each year and for each town and city in the combined sample, the level of inequality, as measured by the Theil index, against mean income, in 1987 Liaoning prices. As mentioned in II.1, the number of towns and cities surveyed approximately doubles in 1988. Figure 5 shows that this brought into the survey a much wider range of mean incomes and inequality levels. Table 10 confirms that the increased dispersion of mean incomes was responsible for much of the increase in inequality observed in 1988. It presents the Theil index decomposed into average-within-cities-and-towns inequality and between-cities-and-towns inequality components (see Appendix B, Chapter Four for the formula used). Between-cities-and-towns inequality increases dramatically between 1987 and 1988: it triples from .006 to .181. Although increasing the number of cities and towns covered should not, if the original sample is random, necessarily lead to increased between-cities-and-towns inequality, it is certainly plausible that widening the geographical base of the sub-samples did in fact increase their representativeness, bringing out more accurately the underlying dispersion of income in the two provinces.[9]

In contrast to the increase in between-cities-and-towns inequality, average within-cities-and-towns inequality actually falls over these two years, from .0415 to .0375. If, as argued in II.1, the selection of the sub-samples lead to under-estimation of the inequality in the full sample, then, since selection takes place at the town and city level, the reduction in size in each city and town's sample in 1988 (typically from 100 to 50 households) to accommodate the increased geographical coverage may well in and of itself have led to a greater under-estimation of inequality, which would explain this fall in average within-cities-and-towns inequality. At the very least, Figure 15 and Table 10 suggest an attitude of agnosticism towards the comparability of 1987 and later years is justified. Bearing this in mind, one is forced to conclude that it is not possible to discern any statistically significant trends in inequality from this provincial data for the late eighties.

9. Note that this increase in geographical coverage is an increase in the sub-sample's coverage of the full sample, not an increase in the full sample's coverage of urban China. So this problem of comparability does not afflict comparisons with 1987 in the previous chapter. A similar problem may arise though with the full sample in 1985 - see Chapter Four, V.2.1.

Figure 5 City and Town Inequality and Means, Combined Sample, 1987-90

Legend: C City T Town



Horizontal axes: Mean income (Liaoning 1987 prices)

Table 10 Inequality Within and Between Cities and Towns

	Year	Theil	Within	Between
Combined				
	1987	0.0476	0.0415	0.0061
	1988	0.0556	0.0375	0.0182
	1989	0.0619	0.0464	0.0155
	1990	0.0537	0.0329	0.0208
Liaoning				
	1987	0.0398	0.0361	0.0038
	1988	0.0434	0.0207	0.0227
	1989	0.0480	0.0307	0.0174
	1990	0.0430	0.0178	0.0252
Sichuan				
	1987	0.0528	0.0455	0.0073
	1988	0.0658	0.0515	0.0144
	1989	0.0721	0.0579	0.0141
	1990	0.0614	0.0455	0.0158

Notes: The index used is the Theil index: see Chapter 4, Appendix B for the decomposition formula. Very small discrepancies between the Theils of this table and those of Table 9 exist on account of rounding errors. The source for both the table and the figures is the SSB sub-samples for Liaoning and Sichuan.

III.3 Welfare and poverty

I turn now to the implications of the combination of changes in mean income presented in Table 6 and the changes in inequality described above for changes in welfare and poverty. The method of analysis is exactly the same as for inequality: first sample and then statistical outcomes are analyzed. The only difference is that different dominance criteria are used. I begin with the first-order stochastic dominance (1OSD) criterion. As with the equality analysis, bounds are used are at either end to prevent minimum and maximum income dominance becoming necessary conditions for dominance (see footnote 37 of Chapter One). Since the 1OSD criterion compares distribution functions, bounds are expressed in terms of income (in constant 1987 prices). A lower bound of 500 yuan per capita is used and an upper bound of 1500. On average, approximately 95% of the annual samples lie above the lower bound and 95% below the upper bound. Restricted 1OSD can be simply interpreted as equivalent to dominance using the head-count ratio for all poverty lines within the two bounds. Another criterion used in this sub-section is the y-restricted welfare 2OSD criterion. This is analyzed using the deficit curve. Only a lower bound is utilized, to prevent the need for minimum dominance, again equal to 500 yuan. Finally, the e-dominance criterion is used with bounds of 0 and 3. Chapter One gives more details on all these criteria and Chapter Two gives details of the estimators required. Results using them are given in Tables 11 to 13 and shown graphically in Figures 6 to 8.

Looking first at sample outcomes, the main result to emerge is that 1990 dominates the other three years, by all three criteria. If the 2OSD or e-dominance, but not the 1OSD criterion is used, the surprising result also emerges that 1989 dominates 1988. If we take a statistical approach and use the same exogenous bounds, 1990 is seen to dominate only 1988 by the 1OSD criterion, to dominate both 1988 and 1989 using the 2OSD criterion, and to dominate all three of 1987, 1988 and 1989 using the e-dominance criterion, which provides a nice example of how using increasingly restrictive criteria can lead to increasingly discriminating rankings. If Chapter Three's methodology concerning endogenous bounds is used, all three criteria give dominance with wide bounds by 1990 over all three of the earlier years (see Tables 11b, 12b and 13b). The sample dominance by 1989 over 1988 turns out to be completely insignificant.

Apart from the clear dominance of 1990 over the other three years, it is also of interest to note from Figures 6 and 7 that for low values of y quite large negative test statistics emerge from the comparison of 1987 with both 1988 and 1989. This suggests a distinct deterioration in the position of the sample poor in these two years. Table 14 allows for further investigation of this question. It presents head-count ratios for two different 'poverty lines', 600 and 750, both in terms of

Liaoning 1987 prices. Although these figures have been somewhat arbitrarily chosen, they do show a clear picture.[10] Poverty, as measured by the head-count, stagnates in Liaoning between 1987 and 1989, and is substantially higher in Sichuan in 1988 and 1989 than in 1987. For example, using the lower of the two poverty lines, the head-count rises from 8% in 1987 to 13.2% in 1988. Again, however, the question arises of how much of this is due to the increase in geographical coverage in 1988.

To summarize, the cost of the statistical approach is that it allows for a less complete ordering than that based purely on sample outcomes. On the other hand, its benefits are that it allows a degree of confidence to be attached to those inferences which are made and enables one to single out features of importance and to discard sample findings of insignificance (such as 1989's welfare dominance over 1988). Once questions of comparability are taken into account, the single feature to emerge from this study as significant is the improvement in welfare and reduction in poverty in 1990 compared to all earlier years.[11] Larger sample sizes and greater comparability would be required for more discriminating orderings to be possible. In this case less demanding criteria probably would not help. Where statistical rankings are not possible, the test statistics tend to be small over the entire range considered.

10. Ahmad and Wang (1991) present a range of poverty lines for urban China from 350-460 yuan for 1987. Somewhat higher lines are chosen here to give a head-count ratio of above 3-5%.

11. Use is not made of the mixed dominance criterion introduced in Chapter One. Mixed dominance has a role only when there is 2OSD but no 1OSD. The only pair in which that obtains is 1988-1989, in which 1988 dominates. However, for this pair the 2OSD is not statistically significant, and, in addition, there is no clear range over which 1989 has a lower distribution function: see Figure 6.

Table 11 First-order Stochastic Dominance, Combined Sample

Table 11a Sample and Statistical Dominance with Exogenous Bounds, 500 and 1500 Yuan

Year	1988	1989	1990
1987	X	X	DB
1988	.	X	DB*
1989	.	.	DB

Table 11b Statistical Dominance, Endogenous Bounds

Year	1988	Lgth	Min	Max	1989	Lgth	Min	Max	1990	Lgth	Min	Max
1987	D	0.11	533	689	D	0.15	616	756	DB	0.88	675	2785
1988	.	.	.	.	X	.	.	.	DB	0.95	478	1912
1989	.	.	.	.	.	.	.	.	DB	0.93	541	1863

Table 12 Welfare Second-order Stochastic Dominance, Combined Sample

Table 12a Sample and Statistical Dominance with Exogenous Lower Bound of 500 Yuan

Year	1988	1989	1990
1987	X	X	DB
1988	.	DB	DB*
1989	.	.	DB*

Table 12b Statistical Dominance, Endogenous Bounds

Year	1988	Lgth	Min	Max	1989	Lgth	Min	Max	1990	Lgth	Min	Max
1987	D	0.34	545	873	D	0.27	657	871	DB	0.84	719	3078
1988	.	.	.	.	X	.	.	.	DB	1.00	306	3470
1989	.	.	.	.	.	.	.	.	DB	0.98	463	7127

Table 13 Welfare E-dominance, Combined Sample

Table 13a Sample and Statistical Dominance with Exogenous Bounds of 0 and 3

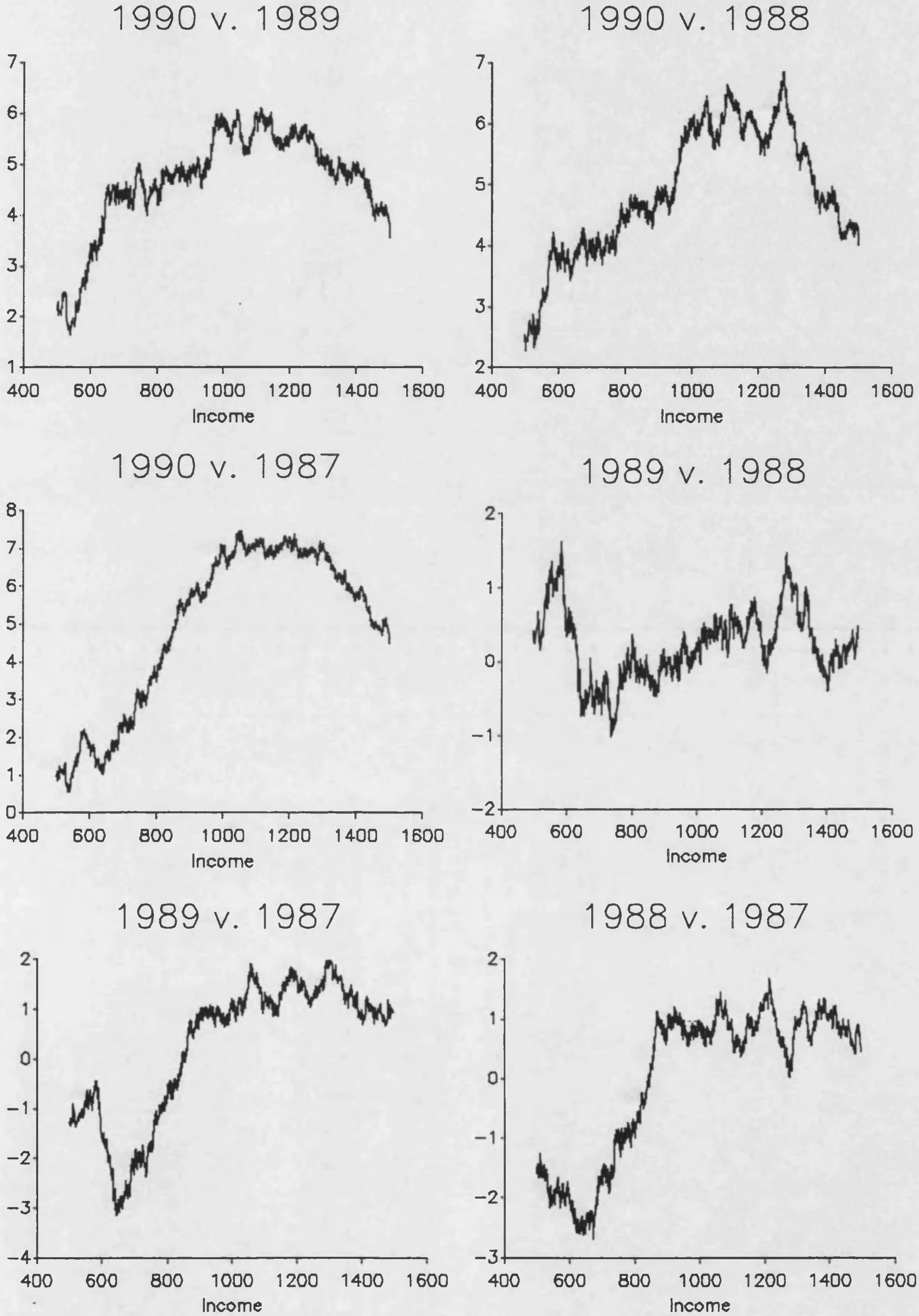
Year	1988	1989	1990
1987	X	X	DB*
1988	.	DB	DB*
1989	.	.	DB*

Table 13b Statistical Dominance, Endogenous Bounds

Year	1988	Lgth	Min	Max	1989	Lgth	Min	Max	1990	Lgth	Min	Max
1987	X	.	.	.	X	.	.	.	DB	5.00	0.0	5.0
1988	.	.	.	.	X	.	.	.	DB	5.00	0.0	5.0
1989	.	.	.	.	.	.	.	.	DB	5.00	0.0	5.0

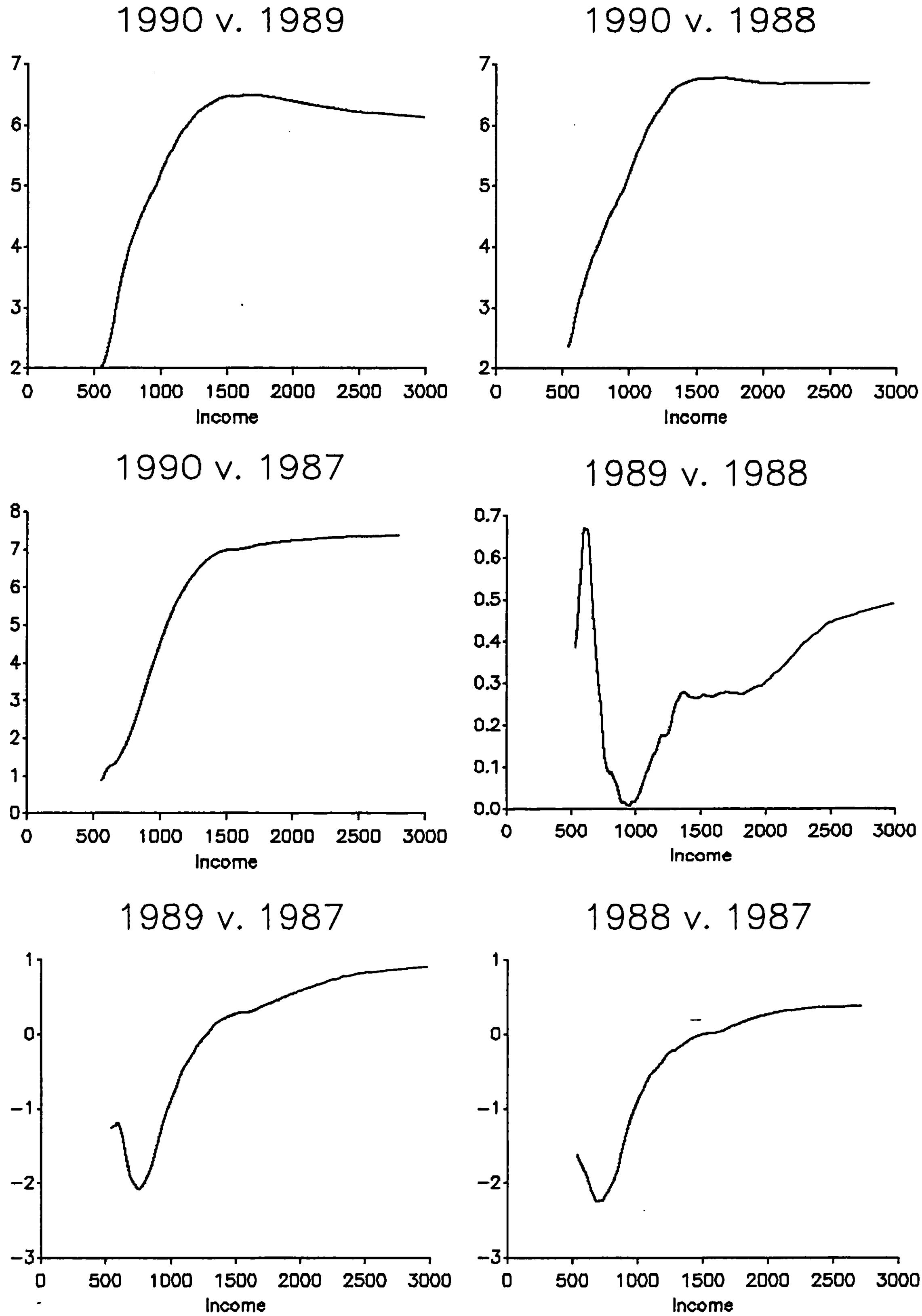
Notes: A 'D' indicates dominance by the row of the column, 'DB' dominance by the column of the row. An 'X' indicates no ranking can be made. An asterisk in Tables 10a, 11a and 12a indicates that the dominance is statistically significant at the 5% level. In Tables 10b, 11b and 12b, the maximum range or length over which one distribution has statistically significant dominance over the other is given, in terms of combined sample coverage for Tables 10b and 11b and e for Table 12b (this is the variable 'Lgth'). This range may be found anywhere between the minimum and maximum incomes for Tables 10b and 11b, and anywhere between e=0 and e=5 for Table 12b. 'Min' and 'Max' give the, respectively, minimum and maximum bounds for this range in terms of income for Tables 11b and 12b and in terms of e for Table 13. In these 'b' tables, 'D' and 'DB' indicate dominance by the row or column year respectively over the range specified. An 'X' indicates that there is no point of statistically significant dominance or, in Table 11b, that there is no range greater than .01 at which dominance is significant. The source is the SSB sub-samples for Liaoning and Sichuan.

Figure 6 Distribution Function Test Statistics, Combined Sample



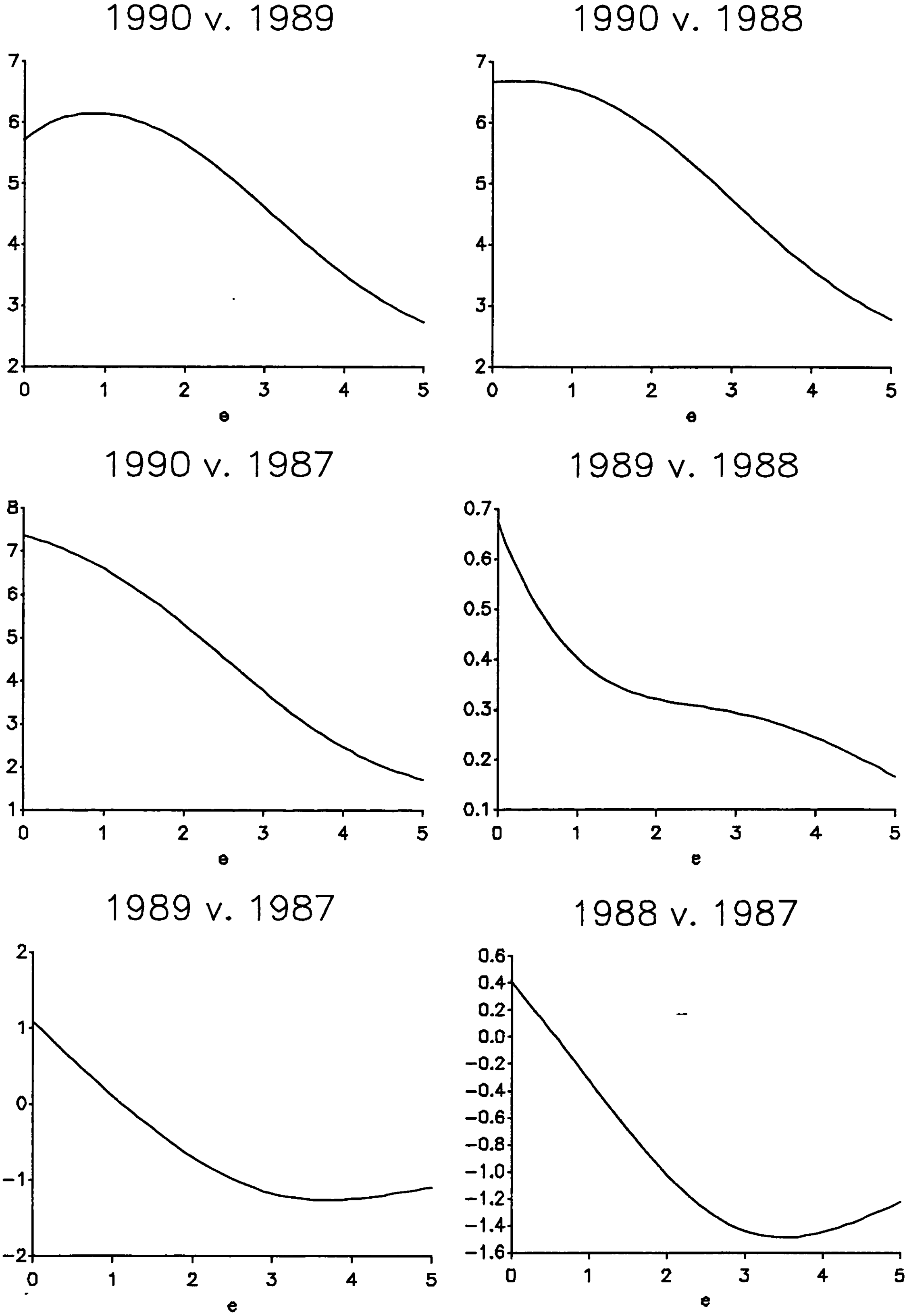
Notes: The vertical axis gives the value of the test statistics on which Table 11 is based. These are calculated using (1) and the appropriate formulae from Chapter Two, II.3: see equations (18) and (19). A positive (negative) value indicates a lower distribution function for the first named (second-named) year at the given income level. The source is the SSB sub-samples for Liaoning and Sichuan.

Figure 7 Deficit Curve Test Statistics, Combined Sample



Notes: The vertical axis gives the value of the test statistics on which Table 12 is based. These are calculated using (1) and the appropriate formulae from Chapter Two, II.4: see equations (28) and (31). A positive (negative) value indicates a lower deficit curve for the first named (second-named) year at the given income level. The source is the SSB sub-samples for Liaoning and Sichuan.

Figure 8 Welfare E-dominance Test Statistics, Combined Sample



Notes: The vertical axis gives the value of the test statistics on which Table 13 is based. These are calculated using (1) and the appropriate formulae from Chapter Two, II.2: see equations (13) and (14). A positive (negative) value indicates a higher isoelastic function (calculated using income) for the for the first named (second-named) year at the given value of e, the relative inequality aversion parameter. The source is the SSB sub-samples for Liaoning and Sichuan.

Table 14 Head-count ratios (%)

Poverty line	Year	Liaoning	Sichuan	Combined
600	1987	6.9	8.0	7.5
	1988	6.4	13.2	10.1
	1989	8.0	12.1	10.4
	1990	5.3	6.9	6.2
750	1987	22.0	22.1	22.0
	1988	21.5	27.4	24.7
	1989	23.7	29.3	26.9
	1990	18.6	18.5	18.6

Notes: The poverty line is measured in Liaoning 1987 prices. The source is the SSB sub-samples for Liaoning and Sichuan.

IV The distribution of food subsidies and price-rise compensation

So far, only the distribution of cash income has been analyzed. As indicated in III.2 of Chapter Four, however, the value of non-cash income is around half of cash income in urban China. In this section, the focus is on just one form of non-cash income, that of consumer subsidies attached to foodstuffs - and, in fact, only on the subsidies attached to two important foodstuffs, grain and oil. In addition, the system of cash compensation which the government has introduced as a substitute for these subsidies is examined. After presenting a historical overview (IV.1), I chart the impact on the distribution of income of the food subsidies (IV.2) and cash compensation (IV.3), and I also simulate the distributional impact of the price reforms of 1991 and 1992 (IV.4).

It must be emphasized at the outset that food subsidies have been by no means the most important source of non-cash income for China's urban residents. Table 5 of Chapter 4 values these subsidies at approximately 10% of total subsidies and income in kind, half of which is due to housing subsidies. However, attempts to estimate the distribution of total non-cash income are unfortunately plagued by measurement difficulties. Food subsidies are focused on here since, one, though measurement difficulties are by no means absent in this area at least free-market and state prices are available and, two, the tracking of the dismantling of food subsidies is of special interest.

IV.1 A brief history

Consumer subsidies on foodstuffs have a long history in urban China. Zhang summarizes their history up to 1978.

> Price subsidies in China date back to 1953. Back then, only ginned cotton was subsidized, to the tune of 50 million yuan. In the eight years between 1953 and 1960, ... the varieties of price subsidies increased from one to four, with grain, edible vegetable oil and tea, for border areas, being added to the list. In the ten years between 1961 and 1970, the varieties of price subsidies further increased from four to eight with vegetables, pigskin and mulberry silk cocoons being added to the list. The amounts involved also increased. In the 1970s, the scope and amount of price subsidies continued to grow, though not by a very large extent. This state of affairs continued right up to 1978 when the reform and open-door policy was first introduced. (1990, p.26)

One might have expected that, with the commencement of reform in 1978, the subsidy system would have been gradually phased out, but the picture is more complex than this and in many ways the opposite. Zhang continues:

> After 1978, price subsidies for [urban] residents showed a sharp upturn, increasing from 5.56 billion in 1978 to the budgeted 35.1 billion in 1989. ... During this period, not only did the amounts of subsidies swell rapidly but the scope of subsidies also expanded significantly. According to incomplete statistics, there are more than 120 different types of subsidized commodities in China. Included under food are fruits, sugar, rice wine, tea leaves, peanuts and beer. Included under clothing are dacron, silk, white cloth and boots. Included under articles for daily use are matches, tissue paper, metal fittings, and thermos flasks. Included under fuel are liquefied gas and firewood. The list covers practically everything. (*ibid.*)

Up to and through the eighties, subsidies were distributed in two ways. For a small group of basic goods - including grain, edible oil, and, up to the mid-eighties, meat - cheap prices were only available on purchases backed by ration coupons, coupons available only to those with official urban (strictly, non-agricultural) registration (*hukou*). Purchases above the coupon limits had to be made on the free-market where prices are higher. For a much wider range of goods, including those mentioned above by Zhang, coupons were not required and unlimited purchases could be made at the prevailing low prices set by the government.

As Zhang indicates, the cost of the subsidy program escalated in the eighties. Bai of the Price Commission Office writes that "Price subsidies as a proportion of financial revenue [i.e. tax revenues plus borrowing] rose from 8.4% to 14%." (1990, p.33) There were attempts in the late seventies and mid-eighties to scale back the program. The first of these two occasions was in November 1979, when "the state raised the prices of eight categories of foodstuffs, including meat, eggs, vegetables, and milk" (Zhao Hongyue, 1991). Then in 1985 "... pork rationing was abolished. ... The prices of fish, chicken, duck, and beef ... and the prices of quality vegetables were also freed." (Hua, Zhang and Luo, 1992, p.119)

However, the effect of these measures was more than offset by the growing costs faced by the government. Procurement prices were increased as part of the rural reform program, procurement quantities rose on the back of strong agricultural growth, but urban prices were held down to protect urban consumers. Han and Lu write:

> A look at changes in the country's agricultural product procurement price indices since reform shows a 181.2% increase in ... 1989 over 1978. This includes a 248.2% increase in the procurement price of grain, a 102.4% increase in the procurement price of cotton and a 145.2% increase in the price of edible vegetable oil. [However] retail prices of some agricultural products were not raised at all. Examples include grain and vegetable oils... Although retail prices of some family products were raised, the amount of increase was far smaller than the increase in procurement prices. Such was the case for fresh eggs, pork and fresh vegetables. (1991, pp.41-2)

Indeed, the price of retail grain remained unchanged for twenty-five years. The problem of runaway costs became chronic in the late eighties, when inflation reached into the high teens and twenties. The government was caught between its desire to stabilize consumer prices and to raise purchasing prices to protect peasant incomes. Their solution was to increase spending on subsidies. In addition, goods whose prices had been freed, such as pork, were once again brought under state control (Hua, Zhang and Luo, 1992, p.119). Once inflation had been brought down, however, the government resumed its policy of price reform. A start was made in 1990 with increases in a number of prices of non-essentials.[12] Then, against a favourable background of falling free-market prices, more substantial price rises were made in 1991 and 1992. In particular, the selling-prices of grain and oil were increased to match the government's procurement prices (more details are given later in IV.4). In addition, a large-scale program of freeing prices from government control was undertaken. According to Yao (1992), of all agricultural goods, only grain and wood now have their consumer prices under government control.

A politically prominent, but little studied feature of the reform of the program of government subsidies has been the cash compensation which the government has used to maintain real income levels in the face of price rises resulting from reform. Compensation has been distributed on each of the three occasions of reform outlined above. First in 1979, workers were given five yuan per

12. In the course of 1990, the government announced price rises covering postage, soap, detergent powder, milk, running water, bus fares, rent and hospital registration fees (Xinhua, 1990).

month in compensation for the various price rises of that year. (Zhao Hongyue, 1991, p.37) Further compensation was introduced in 1985 in response to the meat price rises worth, according to central budgetary figures, about 15 yuan per urban resident per year (SSB, 1991y). Finally, compensation was introduced for the grain and oil price rises of 1991 and 1992: six yuan per month for workers and pensioners in 1991 and five yuan in 1992 (Xinhua, 1991 and 1992). These have not been one-off payments, but commitments by the government to annual payments into the indefinite future.

IV.2 Subsidy income

The focus of this sub-section is solely on the subsidies attached to grain and oil purchases. This is for a number of reasons. First, the plethora of goods receiving subsidies notwithstanding, there can be no doubting the central importance in the food subsidy program of these two commodities. Zhang (1990) indicates that in 1987 over 60% of all government price subsidy expenditure went to subsidizing grain and oil.[13] Second, it is widely recognized that there are quality as well as price differences between subsidized and non-subsidized commodities. If not taken into account, these will lead to an over-estimation of the price gap due to the subsidy. This is a problem which will affect our analysis of grain and oil, but which would plague even more seriously the analysis of less homogenous goods such as meat, vegetables and poultry. Third, whereas the status of certain goods such as pork was changing through the eighties, the situation concerning grain and oil is much simpler.

There are two ways of calculating income received in the form of consumer subsidies. In both cases, one multiplies the total quantity of the good bought at subsidized prices by a per-unit subsidy to obtain the subsidy accruing to an individual or household. Using the first method, the per-unit subsidy is simply calculated as the total government expenditure on subsidizing that good divided by the total consumption of the good at subsidized prices (see World Bank, 1989 and Kupa and Fajth, 1990, both summarized in Atkinson and Micklewright, 1992, for applications of this method to Eastern Europe). Using the second method, the per-unit subsidy is calculated as the difference between free-market and state prices.[14] Whatever the respective merits of the two methods, the cost information required to implement the first method is unavailable for China, and so I rely exclusively on the second. This will give an accurate measure of the cash equivalent of

13. Bai (1990, p.33) puts the proportion at 56% for 1988, an increase from 38.7% in 1978.

14. A more complicated variant of this would be to calculate the equilibrium price of the subsidized good in the absence of the subsidy.

the subsidy program if the relevant variables are measured accurately and if the rations are intra-marginal. Neither assumption can be adhered to with a high degree of confidence. In particular, as will be seen, the free-market in both goods is of little importance. So the estimates presented should be regarded as illustrative rather than definitive: due to the approximate nature of many of the assumptions made, the results are not subjected to statistical analysis.

The main results are reported in Table 15 at the end of this sub-section. Subsidized quantities consumed and prices ('state' prices) were obtained from the survey data, with prices being calculated as average unit values of purchases at state stores. The state price of grain was roughly constant in Liaoning, between 1987 and 1990 at around .5 yuan. That in Sichuan rose slightly from .4 in 1987 to .5 in 1990. The state price of oil was rising in both provinces from 2 yuan in 1987 to 2.5 in 1990 in Liaoning and 2.9 in Sichuan. How is one to reconcile these increasing prices with the absence of official price rises? The most likely explanation would seem to be a change in the quality-composition of goods purchased: although we are analyzing grain and oil as if they were each single commodities, in fact there are different types of grains and oils which can be purchased using ration cards.[15]

Information on free-market prices could be calculated from the survey data but is not on account of the thinness of the free-market in these two commodities, in turn the result of the generosity of the coupon quotas. Table 16 demonstrates. With regards to oil, only one in three or four of all households make at least one purchase on the free-market per year in Liaoning and only one in ten in Sichuan. With regards to grain, the number is much higher, but the importance of the free-market is small, with expenditure in state shops typically constituting 90% of total grain expenditure. In the case of oil, measurement error may cause severe difficulties and those few buying may not be representative. In both cases, there is the danger that the free-market is being resorted to by ration-card holders only for high-quality purchases, which would result in an exaggeration in the gap between the free-market and state prices. Instead of using sub-sample data, published prices, based on SSB price surveys, are used. SSB (1990y) gives the ratio of free-market to state prices for various commodities for 1989, and SSB (1990b and 1991y) gives commodity-specific free-market inflation rates. Both sources use the same definitions of grain and oil as the survey data. Combining these enables us to estimate the free-market price for each year of interest. Note that these inflation rates and ratios are nationwide averages. Complete provincial figures are not available.

15. Three kinds of grain - wheat flour, rice and maize - and six kinds of edible oil - peanut oil, sesame oil, rape oil, refined cottonseed oil, tea-seed oil and soya-bean oil - were rationed in the eighties (Xinhua, 1991).

Since free-market prices rose faster then state prices, the gap between the two widened in the late eighties as Table 15 shows. Whereas in 1987 the state price of grain was about two-thirds the market price, by 1989 it was only a half. The ratio for oil fell over the same period from over .8 to around 2/3. The result of this was that, even though per capita quantities consumed were constant over the years, both nominal subsidy income and the ratio of subsidy to disposable income rose rapidly between 1987 and 1989, the former tripling and the latter doubling (from 2-3% to 6-7%). In the deflationary environment of 1990, when free-market prices actually fell, this trend was reversed and the subsidy-to-income ratio fell back to 4%. Note that the size of the subsidy, though changing over time, is similar for any given year in the two provinces. Grain purchases were a far more important source of subsidy income than oil purchases in both provinces, but especially in Liaoning where they made up 90% of total subsidy income from the two sources.

As mentioned above, the value of income subsidies reached their peak in 1989 at 6-7% of disposable income. It is interesting to note that food subsidies were found to be worth 7% of disposable income in Poland in 1987 (World Bank, 1989, quoted in Atkinson and Micklewright, 1992).[16] By contrast, the subsidies implicit in India's Public Distribution System were estimated by Howes and Jha (1992) to be worth 3.2 rupees per person per month in 1986-87. The same data set reveals per capita mean monthly expenditure to be on average 226 Rupees per month, so that income from food subsidies in urban India constitutes only 1.4% of per capita expenditure and thus even less of income. As in Chapter Four, Section IV, urban China emerges with characteristics more typical of Communist Eastern Europe than of a developing economy.

If one aggregates up and assumes that the subsidies received in Sichuan and Liaoning are typical for those throughout urban China then one arrives at the total benefit accruing from these subsidies to be worth 6 billion yuan in 1987, 10 billion in 1988, 19 billion in 1989, and 13 billion in 1990. This compares with a reported cost for the oil and grain subsidies in 1987 of 16 billion (Zhang, 1990) and of 18 billion in 1988 (Bai, 1990). There could be many reasons for the discrepancy between calculated costs and benefits. In particular, the budgetary cost of the subsidy program will depend not on the gap between the free-market retail and state retail price, but on that between the state procurement and storage costs on the one hand and the state retail price on the other. However, it is unlikely that the latter difference (between state procurement and state retail prices) will be greater than the former (between free-market and state retail prices) since free-market retail prices depend on free-market procurement prices which have tended to lie above

16. Though note the method of calculation is different: the first of the two methods given at the start of the sub-section is used.

the state procurement price. If the reported cost figures are reliable, therefore, the comparison would suggest either that the results of Table 15 under-estimate the difference between the free-market and state prices and thus the value to consumers of the subsidies attached to grain and oil or that Liaoning and Sichuan urban residents receive below-average amounts by way of subsidies.[17]

Table 17 illustrates the equalizing impact of the subsidy-income. It presents two well-known indices, the Gini and Theil, calculated first on the basis of disposable or cash income (that is, as reported in Table 9) and then on the basis of cash-plus-subsidy income. Both indices show that income is more equally distributed once subsidy income is taken into account. The Gini falls by as much as .01 in 1989 once subsidies are taken into account. The indices also suggest that the growth in subsidy income has had a dampening influence on the rise in inequality in 1988 and 1989. Rises are still apparent in 1988, but in neither year are they as great for cash-plus-subsidy as for cash income and in 1989 the Sichuan Gini for the former actually falls.

17. Khan, Griffen, Risken and Zhao (1992) estimated (on the basis of another household survey - see Chapter 4, II.1 for details) the value of ration subsidies to be 97 yuan per capita in 1988, twice our value but including all food-stuffs. Their estimates were based on the reported market value of the ration cards themselves. Zhang (1990) also estimates a higher value of approximately 60 yuan for the grain and oil rations alone in 1987. However, his estimates are based on aggregated SSB data, used to proxy free-market prices by unit values. As argued in the text, this could well lead to error.

Table 15 Implicit Food Subsidies, Grain and Oil

Year	TOTAL Subsidy	As % of disp. y	GRAIN Subsidy	Quantity	Price Market	Price State	OIL Subsidy	Quantity	Price Market	Price State
Liaoning										
1987	26.6	2.8	25.3	114.0	0.72	0.49	1.3	4.5	2.34	2.06
1988	45.2	3.8	41.6	115.7	0.89	0.53	3.6	5.7	2.77	2.14
1989	81.7	5.9	73.9	109.6	1.21	0.54	7.8	5.4	3.75	2.31
1990	64.1	4.4	58.8	124.2	0.99	0.52	5.3	4.7	3.68	2.56
Sichuan										
1987	27.7	3.2	21.5	117.9	0.58	0.39	6.2	6.8	2.95	2.05
1988	45.3	4.3	37.9	126.0	0.72	0.41	7.4	6.8	3.50	2.41
1989	86.0	6.8	75.0	138.2	0.98	0.43	11.0	6.1	4.74	2.92
1990	50.5	3.5	39.1	126.0	0.80	0.49	11.4	6.8	4.64	2.98

Notes: 'State' prices are the subsidized prices per kilogram. 'Quantity' refers to the quantity (in kilograms) of the good purchased at subsidized prices. All subsidies and quantities are per capita per annum. The source for all the tables on this page is the SSB sub-samples for Liaoning and Sichuan.

Table 16 Proportions Consumed

Year	Grain: % of exp	Grain: FM as % of total	Grain: h/hs using FM	Oil: % of exp	Oil: FM as % of total	Oil: h/hs using FM
Liaoning						
1987	7.4	9.8	75.3	1.2	12.0	34.3
1988	6.5	11.0	68.6	1.2	7.6	23.7
1989	6.2	15.1	68.4	1.2	11.0	27.5
1990	6.0	11.1	60.5	1.1	11.7	25.5
Sichuan						
1987	6.6	11.1	89.8	1.7	1.5	13.9
1988	6.2	13.5	92.9	1.6	1.1	8.5
1989	6.1	10.4	93.1	1.7	5.2	10.3
1990	5.7	10.4	95.4	1.6	2.7	8.9

Notes: '% of exp' gives total expenditure on grain or oil as a percentage of total expenditure on all commodities. 'FM as % of total' gives free-market expenditure on grain or oil as a percentage of total expenditure on the good. 'h/hs using FM' gives the percentage of households who purchase grain or oil at least once on the free-market.

Table 17 Per Capita Total Income Inequality

Year	Gini Cash	Gini Total	Theil Index Cash	Theil Index Total
Liaoning				
1987	0.1575	0.1540	0.0398	0.0381
1988	0.1630	0.1577	0.0434	0.0405
1989	0.1684	0.1602	0.0477	0.0433
1990	0.1633	0.1560	0.0433	0.0395
Sichuan				
1987	0.1807	0.1747	0.0494	0.0528
1988	0.2012	0.1936	0.0610	0.0659
1989	0.2041	0.1919	0.0638	0.0721
1990	0.1956	0.1897	0.0577	0.0614

Notes: 'Cash' is disposable income. 'Total' is cash plus subsidy income. The 'cash' figures are taken from Table 9.

IV.3 Price-rise compensation

Table 18 gives the average per capita compensation for rises in state prices received between 1988 and 1990, the years for which this item is recorded separately in the SSB survey. The amounts involved are clearly substantial: at around 10% of disposable income, the cash compensation received is greater than the income accruing from the grain and oil subsidies. Sichuan and Liaoning residents seem to do especially well from the compensation payments: they receive some 50% more than the published nationwide averages based on the full SSB sample: 70 yuan in 1988 and 94 yuan in 1989 (SSB 1989s and 1990s).

The rules by which these subsidies are distributed are not well known, but some insight can be obtained using regression analysis, with household compensation income as the dependent variable. Table 19 gives the results from both provinces using 1990 data.[18] The explanatory variables used are: the number in the household employed in the SOE and in the collective sectors, the number of private employees and self-employed, the place of residence (city or town), the number of retirees and the number of children. Higher R^2's (by about 10 percentage points) were obtained using the number of different types of household members (SOE employees, pensioners etc.) than using different types of household income (SOE income, pension income etc.), which is suggestive of the lump-sum way in which this compensation income is distributed. Most of the results are as expected. SOE employees receive more than collective employees and city-dwellers more than town-dwellers. There is also a great deal of variation between the two provinces. The coefficient on number of pensioners is much larger in Sichuan than Liaoning. And whereas there is a positive coefficient on the number of children for Liaoning this is negative for Sichuan. Indeed, in Liaoning, but not in Sichuan, the price-subsidy is recorded in the sample as being paid directly to children.

Table 20 demonstrates the equalizing impact of this compensation payment via use of a Gini decomposition. As explained in the previous chapter (VI.2), any Gini for total income can be decomposed into a weighted sum of pseudo-Ginis for the various sources of income (calculated in the standard way for a Gini but on the basis of an ordering of units by total income, not by the particular income source).[19] Table 20 shows the contribution to cash-plus-subsidy inequality of three sources: subsidy income, compensation income and other cash income. Although both compensation and subsidy income are distributed relatively unequally (they have relatively high

18. OLS was used. Since less than 1% of the sample in Liaoning and less than 3% of the sample in Sichuan receive no cash compensation, Tobit analysis was not conducted.

19. The actual decomposition formula is given in the notes to Table 20.

Ginis), they are distributed relatively equitably, as indicated by their low pseudo-Ginis.[20] Although compensation income is not distributed as equitably as subsidy income, it still exerts an equalizing influence on total income: its pseudo-Gini is well below the total Gini. Note too the much more equitable distribution in Liaoning, whose cash compensation pseudo-Gini is less than half of Sichuan's. This is consistent with the regression results which showed greater account being taken of children in Liaoning when distributing the cash compensation.

Table 18 Average Per Capita Compensation Payments, 1988-1990

	Liaoning		Sichuan	
Year	**Payment**	**% of income**	**Payment**	**% of income**
1988	110.2	9.3	97.6	9.2
1989	141.5	10.2	138.0	10.9
1990	147.4	10.1	152.1	10.7

Notes: 'Payment' gives the per capita payment received as compensation for price rises in current yuan. '% of income' gives this figure as a percentage of cash income.

Table 19 Regression Results for the Distribution of Compensation Income

Sichuan, 1990

R^2=0.4727

Variable	Estimate	S.E.	T-ratio
Intercept	14.9	22.9	0.7
SOE	188.3	8.8	21.4
Collective	51.2	11.6	13.0
Retired	256.5	12.0	21.3
Self-employed	-27.5	40.9	-0.7
Child	-10.4	11.0	-0.9
City dummy	85.4	13.3	6.4

Liaoning, 1990

R^2=0.4470

Variable	Estimate	S.E.	T-ratio
Intercept	-21.0	24.1	-0.9
SOE	142.9	8.2	17.5
Collective	139.3	9.1	15.2
Retired	62.7	11.4	5.5
Self-employed	-56.6	71.2	-0.8
Child	67.1	10.6	6.3
City dummy	188.8	14.7	12.9

Variable definitions

SOE	Number in household employed in state-owned sector
Collective	" " " " " collective "
Retired	" " " retired
Self-employed	" " " self-employed
Child	" " " who are children (age less than 16)
City dummy	Dummy equal to one if city, zero otherwise

Notes: The source is the SSB sub-samples for Liaoning and Sichuan, 1990.

20. Indicating that total income and, respectively, subsidy and compensation income have a low rank correlation since this can be shown to be measured by the ratio of the pseudo-Gini to the true Gini.

Table 20 Income by Source: price-rise compensation, other cash income and subsidy income

	Total	Cash				Compensation				Other cash income				Subsidy income			
Year	Gini	Cont	μ	ρ	Gini	Cont	μ	ρ	Gini	Cont	μ	ρ	Gini	Cont	μ	ρ	Gini
1988	0.158	99.5	96.3	0.163	0.158	3.8	8.9	0.067	0.189	95.7	87.4	0.173	0.174	0.5	3.7	0.022	0.208
1989	0.160	99.0	94.5	0.168	0.160	3.6	9.6	0.060	0.167	95.4	84.9	0.180	0.181	1.0	5.5	0.030	0.224
1990	0.156	100.1	95.8	0.163	0.156	2.8	9.6	0.045	0.161	97.4	86.2	0.176	0.177	-0.1	4.2	-.004	0.243
1988	0.194	99.6	95.9	0.201	0.194	8.2	8.8	0.180	0.296	91.3	87.1	0.203	0.205	0.4	4.1	0.021	0.173
1989	0.192	99.4	93.6	0.204	0.192	7.7	10.2	0.145	0.294	91.7	83.4	0.211	0.215	0.6	6.4	0.019	0.169
1990	0.190	99.5	96.6	0.195	0.190	7.8	10.3	0.143	0.256	91.8	86.3	0.202	0.204	0.5	3.4	0.027	0.175

Notes: Total income is equal to cash income plus subsidy income. μ gives the mean of the type of income relative to the mean of total income (times 100); ρ gives the pseudo-Gini of the type of income. 'Cont' measures the contribution of the distribution of the particular type of income to total inequality and is equal to μ times ρ divided by Gini. Since Gini=($\Sigma\mu_i\rho_i$)/100, the sum of 'Cont' figures is 100. The source is the SSB sub-samples for Liaoning and Sichuan.

IV.4 Reform

As mentioned in IV.1, in 1991 and 1992 the state retail prices of grain and oil were raised. In 1991, the price of grain was raised by .2 yuan per kilogram, and the price of oil by 2.7 yuan per kilogram (Xinhua, 1991). In 1992, the price of grain was further raised by an additional .22 per kilogram (Xinhua, 1992). Compensation was distributed for these price rises: six yuan in 1991 and five yuan in 1992 per month per worker and retiree.

Comparing these price rises with the free-market and state prices shown in Table 15, the total rise in the state price of grain of .42 leads to an approximate doubling in price, and is more or less sufficient to close the gap between the two retail prices. The price rise for oil of 2.7 yuan again approximately doubles its price. It also takes the state price of oil more than 1 yuan above the calculated 1990 free-market price. One possible explanation for this is that, whereas the state procurement price of grain was raised in 1989, that for oil was raised in 1990 (Bi, 1992, p.23). We would expect this to push up the free-market retail price for oil, but it is possible that this increase did not feed through until 1991.

In this sub-section, results of a simple simulation are presented showing the effects of these reforms. It is also assumed that the reforms leave unchanged the amounts of grain and oil consumed, so that they make consumers worse off by the amount consumed times the price rise, and better off by the amount of the compensation.[21] In accordance with the official announcements, it is assumed that all SOE and collective workers and all pension-recipients receive compensation. In view of the regression results, I assume that private-sector employees and the self-employed do not receive any compensation. Table 21a gives the main results: note that the pre-reform figures refer to cash-plus-subsidy income. The first prominent feature is that, even without assuming any substitution away from grain and oil as a result of the price-rise, there is an over-compensation. The average per capita cost of the price rise is 65 yuan in Liaoning and 70 yuan in Sichuan. The average per capita gain from the compensation is 90 yuan in both provinces. The difference is not large as a proportion of pre-reform income, less than 2%. Nevertheless it represents a significant outlay on the part of the government, and it suggests that, in the short-term at least, the government's budgetary position will deteriorate as a result of the reform. If Liaoning and Sichuan are typical, the reform will, on an annual basis, save the

21. For a behavioural analysis of the effects of food price rises in urban China, see Pudney and Wang (1991).

government 15 billion yuan, but cost it 20 billion.[22]

The second important result is that there is almost no increase in inequality at all as a result of the reform. The rises in the Gini are minimal: from .156 to .157 in Liaoning and from .190 to .191 in Sichuan. Since the pre-reform figures are based on cash-plus-subsidy income, in fact if only cash inequality is considered the reforms will lead to a fall in recorded inequality, since we know from Table 9 that the cash-based Ginis for 1990 are .163 and .196 for Liaoning and Sichuan respectively.

Table 21b sheds more light on the question of who will gain and who lose from the reform by giving a breakdown by decile of gains and losses as a proportion of pre-reform income. The losses are a falling proportion of income. In fact, they are almost constant in absolute size, as we would expect from the almost-zero pseudo-Gini of subsidy receipts indicated in Table 20. On its own, this would increase inequality, but the gains from compensation are also a falling proportion, which almost exactly offsets this effect, resulting in a remarkably even distribution of net gains. There are losers though. The poorest ten percent in Sichuan do end up worse off. This is due to a high dependency ratio: if the reform is amended to give half as much compensation to children as to workers, while keeping the total amount of compensation constant, then the poorest decile in Sichuan become marginal gainers from the reform. In any case, their loss under the actual reform doesn't translate into a very much higher Gini since it is both small and offset by the slight relative gain of the sixth to eighth deciles vis-a-vis the ninth and tenth.

22. This estimate of savings does not rely on accurate estimation of pre-reform free-market prices. It follows simply from taking the average per capita quantities consumed in Sichuan and Liaoning and multiplying them by the increase in price and by the urban population. The cost may be biased upwards if, as suggested by the comparison of IV.3, Liaoning and Sichuan residents tend to benefit disproportionately from compensation payments.

Table 21 Simulated Impact of Reform

Table 21a Summary statistics

	Liaoning	Sichuan
Pre-reform mean	1529.9	1474.6
Post-reform mean	1554.6	1492.5
Pre-reform Gini	0.156	0.190
Post-reform Gini	0.157	0.191
Number gainers	568.0	614.0
Number losers	131.0	186.0
Ave comp"n gain	89.6	89.4
Ave price rise loss	65.0	71.4
Ave gain (among gainers)	36.1	30.2
Ave loss (among losers)	23.0	20.5
Max gain	99.2	115.4
Max loss	207.7	108.8

Table 21b Percentage Gains and Losses by Decile

	Liaoning			Sichuan		
Decile	Gain	Loss	Net gain	Gain	Loss	Net gain
1	8.9	7.3	1.6	7.4	7.7	-0.4
2	7.3	5.9	1.4	7.0	6.3	0.8
3	7.2	5.4	1.8	6.7	5.6	1.3
4	6.3	4.8	1.6	6.2	5.1	1.2
5	6.2	4.8	1.4	5.9	4.9	1.1
6	5.7	4.1	1.5	5.9	4.5	1.5
7	5.6	3.7	1.9	5.8	4.5	1.4
8	5.2	3.7	1.5	5.6	4.0	1.6
9	4.9	3.3	1.6	5.1	3.8	1.4
10	4.5	2.8	1.7	4.8	3.5	1.3

Notes: All monetary values in Table 21a are in 1990 yuan per capita. The figures in Table 21b are percentages of pre-reform income. The simulation assumes fixed quantities consumed, a price rise on both grain and oil, and cash distributed in compensation. For more detail, see text. The source is the SSB sub-samples for Liaoning and Sichuan, 1990.

V Conclusion

The main findings of this chapter can be simply summarized. Concerning trends in inequality, the sample results were in line with the pattern of increased inequality found in Chapter Four: inequality rose in 1988 and 1989, and then fell in 1990. As explained in Chapter Four, there are good reasons for this, relating to the rapid growth and decentralization of 1988-89 and the restrictive policies of 1989-90. However, only the post-1987 increase turned out to be statistically significant and here it was argued that the expansion of the sample coverage prevents full comparability between 1987 and later years. So our main conclusion in the regard must be a negative one: the lack of comparability and relatively small sample sizes, combined with the limited coverage in terms of years, renders the provincial sub-samples unsuitable for statistical analysis of inequality trends. Interestingly though, it was possible to infer the welfare-superiority of the 1990 distribution. This drives home the point that as long as rapid growth rates are maintained - real disposable per capita income growth was 8% in 1990 - increases in inequality, such as recorded between 1987 and 1990, are unlikely to be sufficient to cause a rise in poverty and fall in welfare.

Section IV of this chapter examined the distribution of in-kind subsidy income and of cash compensation. Although statistical analysis was not applied, the main results emerged clearly. In the inflationary environment of the late eighties, the value for the consumer of the government's controls on grain and oil prices increased. Since food purchases take a larger proportion of the total expenditure of the poor than of the rich, this had an equalizing influence on the distribution of cash-plus-subsidy income. One might think therefore that the government's dismantling of the grain and oil subsidies in 1991 and 1992 would have a regressive impact, but simulations do not support this. It is true that the price-rise compensation payments, while still exerting an equalizing influence on the overall distribution, are less equitably distributed than the subsidy income. However, if the government's official guidelines for the distribution of the 1991 and 1992 compensation payments are followed, there will not be a large rise in inequality: just as food purchases are more important to the poor so are the lump-sum compensation payments, and the costs in relation to the former and benefits in relation to the latter largely cancel each other out.

This raises a number of questions, not least about the attitude of the Chinese government to increases in inequality. The quotations in the Introduction (Section II) give the strong impression that the government wants to see a rise in inequality as the price that must be paid for increased efficiency. Yet its own policies in the area of food subsidies do not evince an inclination to move in this direction. The equalizing effect of the expanding value of food subsidies was in part an unintended effect of high inflation. But the unwillingness of the government to increase consumer prices as well as the equitable and generous nature of the compensation payments when it eventually did increase prices also point to the extreme reluctance of the government to impose even short-term income losses on urban residents. The government, it would seem, is only prepared to see inequality rise through differences in the increase rates in the incomes of different urban groups. The policy is indeed one of letting "some workers get rich first" (see the Introduction), rather than one of letting some get rich, and others get poor.

This policy of maintaining real incomes is an expensive one, at least if interpreted as requiring full compensation by the government for all major price reforms. On the basis of the simulations, and allowing for 10% real growth in disposable incomes since 1990, price-rise compensation payments will now account for some 15% of disposable income. In other words, China, or rather non-agricultural-registered China, now has what is essentially, on account of its lump-sum nature, a substantial basic income scheme in place. This is an extremely important development for a number of reasons. It is an expensive fiscal item - costing, by my calculations, up to one-third of the much discussed subsidies to loss-making enterprises. The compensation also maintains the importance of household registration. All urban dwellers now face more-or-less the same prices for their food. But some (with non-agricultural-registration) receive over 200 yuan a year to help them buy at these prices; others receive nothing. It is little wonder that a thriving black market exists in registration permits.[23] And, since companies are responsible for the distribution of these payments to workers, the growth of these compensation payments intensifies the involvement of the commercial sector in the delivery of welfare services, despite the professed aim of the government to move in the opposite direction.

Whether the government will let this *de facto* basic income scheme wither away by failing to maintain its real value in the face of inflation remains to be seen. Perhaps more likely is that the government will continue to use the method of cash compensation to path the way for further reform. With housing rents still only at nominal levels in urban China, and the government seeking reform in this area, compensation payments could increase substantially in the future. Whether this would be a welcome development is unclear. Given the privileged position of China's non-agricultural-registered population (see the conclusion to Chapter 4), the government might do better, from the point of view of equity, by providing compensation only for those already in or pushed into poverty as a result of price changes and otherwise letting this sector take its compensation in the form of growing disposable incomes generated by a growing urban economy. This would further reduce urban inequality, which might run counter to the government's aims and might have adverse incentive effects. On the other hand, it would result in large expenditure savings and would lead to a narrowing of the urban-rural gap.

23. According to one report the sales of permits is "a new 'craze' sweeping the country". China's Economic Information Daily reported that officials in the province of Anhui raised $US 350 million in the first eight months of year by the sale of the permits to more than 500,000 peasants, at prices ranging from $525 to $2,500. (Reported by UPI, 12 June 1992, extracted in CND-Global (electronic information sheet), December 7, 1992.)

Appendix A Basic data tables

Table A.1 Quantile and Cumulative Income Shares, 1987-1990

Combined

	---Decile income shares---				--Cumulative income shares--			
Decile	**1987**	**1988**	**1989**	**1990**	**1987**	**1988**	**1989**	**1990**
1	5.4	4.9	5.1	5.2	5.4	4.9	5.1	5.2
2	7.1	6.8	6.6	6.7	12.4	11.7	11.7	11.9
3	7.9	7.7	7.6	7.7	20.3	19.4	19.3	19.5
4	8.5	8.5	8.4	8.4	28.8	27.9	27.7	28.0
5	9.2	9.3	9.2	9.2	38.0	37.2	36.9	37.2
6	10.0	9.9	9.9	10.0	48.0	47.1	46.8	47.2
7	10.7	10.9	10.8	10.8	58.7	58.0	57.6	58.0
8	11.7	11.8	11.7	12.0	70.4	69.8	69.3	70.0
9	13.1	13.2	13.2	13.2	83.5	83.0	82.6	83.2
10	16.5	17.0	17.4	16.8	100.0	100.0	100.0	100.0

Liaoning

	---Decile income shares---				--Cumulative income shares--			
Decile	**1987**	**1988**	**1989**	**1990**	**1987**	**1988**	**1989**	**1990**
1	5.7	5.5	5.4	5.5	5.7	5.5	5.4	5.5
2	7.3	7.1	6.9	7.0	12.9	12.6	12.3	12.5
3	8.0	7.8	7.8	7.9	20.9	20.4	20.1	20.5
4	8.6	8.7	8.7	8.7	29.6	29.1	28.8	29.1
5	9.4	9.5	9.4	9.3	39.0	38.6	38.2	38.4
6	10.0	10.1	10.2	10.1	48.9	48.7	48.3	48.5
7	10.7	10.7	10.9	10.9	59.7	59.4	59.2	59.5
8	11.6	11.7	11.7	11.7	71.3	71.0	70.9	71.1
9	12.8	12.8	12.9	13.0	84.1	83.8	83.8	84.1
10	15.9	16.2	16.2	15.9	100.0	100.0	100.0	100.0

Sichuan

	---Decile income shares---				--Cumulative income shares--			
Decile	**1987**	**1988**	**1989**	**1990**	**1987**	**1988**	**1989**	**1990**
1	5.2	4.6	4.8	4.8	5.2	4.6	4.8	4.8
2	6.9	6.4	6.5	6.5	12.1	11.0	11.3	11.3
3	7.6	7.4	7.4	7.4	19.7	18.4	18.6	18.7
4	8.5	8.3	8.2	8.4	28.2	26.8	26.8	27.1
5	9.2	9.2	9.1	9.1	37.3	36.0	35.9	36.2
6	9.9	9.9	9.7	10.0	47.2	45.9	45.7	46.1
7	10.8	10.8	10.6	11.0	58.0	56.7	56.3	57.1
8	11.9	12.1	12.0	12.1	69.9	68.8	68.2	69.2
9	13.1	13.5	13.6	13.4	83.0	82.3	81.8	82.7
10	17.0	17.7	18.2	17.3	100.0	100.0	100.0	100.0

Note: Calculated from the provincial sub-samples using, to obtain the 'Combined' results, the price indices described in III.1.

Table A.2 Real Decile and Conditional Means, 1987-1990

Combined

	Decile means				Conditional means			
Decile	**1987**	**1988**	**1989**	**1990**	**1987**	**1988**	**1989**	**1990**
1	530.9	493.1	506.1	555.4	530.9	493.1	506.1	555.4
2	695.0	666.8	663.4	723.5	612.9	579.9	584.8	639.5
3	772.5	760.0	755.1	822.0	666.1	640.0	641.6	700.3
4	841.8	842.7	840.5	908.2	710.0	690.6	691.3	752.3
5	907.6	919.4	918.8	991.3	749.5	736.4	736.8	800.1
6	979.0	988.7	990.9	1075.4	787.8	778.4	779.1	846.0
7	1056.9	1070.3	1076.9	1164.6	826.2	820.1	821.7	891.5
8	1153.3	1169.0	1175.8	1277.8	867.1	863.8	865.9	939.8
9	1288.8	1307.3	1318.7	1421.6	914.0	913.0	916.2	993.3
10	1620.0	1675.4	1733.9	1804.1	984.6	989.3	998.0	1074.4

Liaoning

	Decile means				Conditional means			
Decile	**1987**	**1988**	**1989**	**1990**	**1987**	**1988**	**1989**	**1990**
1	545.9	550.3	544.0	569.8	545.9	550.3	544.0	569.8
2	694.5	702.6	685.5	727.4	620.2	626.4	614.7	648.6
3	769.0	784.4	779.5	819.3	669.8	679.1	669.7	705.5
4	835.0	863.6	854.9	893.9	711.1	725.2	716.0	752.6
5	898.0	935.5	931.5	962.4	748.5	767.3	759.1	794.6
6	962.2	993.0	1007.5	1042.0	784.1	804.9	800.5	835.8
7	1028.6	1069.8	1082.4	1124.0	819.0	842.7	840.8	877.0
8	1113.8	1149.7	1162.0	1212.4	855.9	881.1	880.9	918.9
9	1235.2	1273.3	1278.8	1329.6	898.0	924.7	925.1	964.5
10	1512.4	1596.7	1612.7	1638.6	959.5	991.9	993.9	1031.9

Sichuan

	Decile means				Conditional means			
Decile	**1987**	**1988**	**1989**	**1990**	**1987**	**1988**	**1989**	**1990**
1	519.3	459.3	480.4	542.0	519.3	459.3	480.4	542.0
2	695.0	630.5	646.8	719.2	607.1	544.9	563.6	630.6
3	774.8	735.7	736.6	825.0	663.0	608.5	621.3	695.4
4	847.9	821.0	826.5	927.6	709.2	661.6	672.6	753.4
5	915.0	902.9	909.1	1021.0	750.4	709.9	719.9	807.0
6	996.3	983.8	978.8	1107.7	791.4	755.5	763.0	857.1
7	1082.3	1070.8	1070.0	1219.1	832.9	800.6	806.9	908.8
8	1187.2	1188.0	1188.9	1346.6	877.2	849.0	854.6	963.5
9	1326.3	1333.3	1357.5	1494.0	927.1	902.8	910.5	1022.5
10	1692.9	1736.0	1811.9	1920.5	1003.7	986.1	1000.7	1112.3

<u>Note</u>: Calculated from the provincial sub-samples using the price indices described in III.1.

Conclusion to Part Two

One of the main conclusions to emerge from the analysis of Chapters Four and Five is that, in relation to its income distribution, urban China displays many of the characteristics typical of pre-transition Eastern Europe. We saw in Chapter Four that in our sample of more than thirty developed and developing countries only Hungary could match the compression of the urban Chinese income distribution. And then Chapter Five showed that the urban system of food-rations was equivalent in the late eighties in its value to the urban consumer to a similar scheme operating in Poland and much greater in value than one operating in urban India.

A second key finding is that, notwithstanding the current consensus to the contrary, the SSB survey evidence is consistent with a substantial increase in income inequality in urban China since 1983. This was the major conclusion of Chapter Four. Chapter Five also investigated the question of trends. Although the sample results were consistent with this conclusion of an increase, on closer inspection either incomparability emerged (between 1987 and later years) or the differences were not statistically significant (between 1988 and 1989). However, it must be stressed that Chapter Four did not find evidence the other way, that is, of a downward trend. The conclusion argued for here is that with larger sample sizes and greater comparability, the conclusions of Chapter Four would be borne out with disaggregated as well as aggregated data. Needless to say, a greater availability of data would make a more precise and confident conclusion possible.

Although one cannot be sure, it is tempting to combine these first two conclusions and hypothesise, on the basis of the experience of urban China, that inequality in the countries of Eastern Europe and the former Soviet Union is rising and will continue to do so. The factors behind the increase in urban China - decentralization and weakening of macroeconomic control, and growth of new earning opportunities - would seem to apply more widely to other countries undergoing transition. Indeed, as argued in the conclusion to Chapter Four, one imagines that in those countries in which adjustment is occurring without growth, inequality will grow more rapidly as unemployment rises and social security systems are weakened.

Should the apparent increase in inequality in urban China be a source of concern? A full answer would require an analysis of the trade-offs between equality and efficiency, something which has not been gone into. However, from a simple welfare perspective, the evidence of the eighties suggests that as long as rapid growth continues, welfare will rise and poverty fall despite any rise in inequality.

Not all the forces acting on inequality in urban China in the eighties have been pushing upwards. Chapter Four argued that the re-assertion of macroeconomic control in 1985-86 and 1989-90 led to a halting in the growth in inequality or even a reversal, especially since it led to reversal in the disequalizing growth of bonus payments. And Chapter Five showed that the increasing value of food subsidies in the inflationary environment of the late eighties depressed inequality growth, and that the subsequent replacement of these subsidies by compensation payments has done little, if anything, to widen disparities. Leaving aside the question of whether a rise in inequality is desirable or not, these findings do shed light on the professed aim of the Chinese government to combat what it sees as excessive egalitarianism (see the Introduction). Although this is no doubt a genuine goal of the government's, its pursuit has been tempered both by the need to maintain macroeconomic stability (resulting in limitations being placed on how firms pay their workforces) and by the government's desire to protect real incomes (resulting in a initial reluctance to raise the prices of basic food-stuffs such as grain and oil, and then the provision of extremely egalitarian compensation when they were finally raised). This brings home the point that, even if in general we would expect transition to increase inequality, the way in which transition occurs has a great impact on the transition-inequality relationship. A variety of government policies - macroeconomic as well as distributional - will determine the path inequality takes.

Part Two could not have been written without the data collected and published by China's State Statistical Bureau. This very fact is confirmation of the central role the SSB surveys must have if our understanding of China's income distribution is to be enhanced. Nevertheless, certain weaknesses in this data base have emerged in the course of analysis. Two of these are highlighted here, alongside proposals for change. In relation to the collection of data, we have seen how the SSB's choice of survey base - non-agricultural registered households living in officially designated urban areas - has two negative effects. It means any analysis of urban living standards must leave unaccounted for those living in urban areas without registration. And it makes checks of sample representativeness using known population parameters very difficult, as such parameters are not given using the SSB survey definition. A better survey base would be simply those living in urban areas. If one stays with the SSB definition of urban areas, this more than doubles the sample base, so there is a case for using the more restrictive definition of 'urban' used for Census purposes (see Chapter Four, III.1). This would result in a 25% increase in the survey base. For comparability and policy purposes, it would still be necessary to record whether or not a surveyed household had non-agricultural registration and to publish the main statistics separately for those who do.

If the above suggestion is fairly radical, the next, though just as important, is much more easily accomplished. It is hard to see the justification for the aggregation procedure used by the SSB to compute its published figures and described in III.3. To remind the reader, the data collected centrally by the SSB have, apart from a sub-set, already been aggregated at the city or town level. These city and town quantile means are in turn aggregated into overall quantile means for publication not by placing all the quantile means in ascending order of income, but by aggregating, for example, the poorest decile's mean of each city and town into the poorest overall decile mean. An ideal solution would obviously be to collect all data centrally in disaggregated form, and to aggregate it in the conventional way. But let us assume this is not possible. On the evidence of Appendix B of Chapter Four, the SSB aggregation procedure produces Ginis some 25% smaller in value than those which would result using the 'ideal solution' above. The alternative and no more costly method of aggregating the city and town quantiles properly into overall quantiles results by contrast in underestimation of less than 1%.[1] The case for changing the method of aggregation is compelling, even if figures based on the old method continue to be presented for the sake of maintaining comparability over time.

The SSB has gone some way to mitigating the problems imposed by its aggregation procedure by also presenting data in the yearbooks based on a sub-sample of disaggregated data. But here one must ask what purpose is served by presenting this data on the basis of a ranking of household rather than per capita income. It would also be useful if this data were given on the basis of quantiles, with a fixed number of households or individuals per quantile, rather than on the basis of income classes with fixed bounds, which, because of nominal income growth, can result in rapidly changing class sizes, making comparisons over time difficult.

The link between Parts One and Two is, of course, that the latter provides an application of the methodological points developed in the former. In this respect, I would simply draw attention to three points. The first is the need for large sample sizes if one is to make comparisons of a statistical nature between distributions which are closely related (for example, only a year or two apart). The second is the simplicity of the test statistics and testing method required: in particular, only vectors of variances and not matrices of covariances need to be calculated. The third is the usefulness in the statistical context of a graphical approach. For example, the graphs of the Lorenz curve test statistics in Chapter Five not only convey the same ordinal information as do graphs of pairs of Lorenz curves (namely, whether one distribution has Lorenz dominance or, if not, where the curves cross), but also show at which points, if any, the two curves significantly differ.

1. The SSB now surveys between 30 and 40 thousand households altogether. Each city or town chosen has surveyed at least 50 households. This gives a maximum of 800 cities and towns. A data-set with less than 1000 observations can be easily analysed on a simple personal computer.

Bibliography

Note: A number of translations of Chinese articles and reproductions of English-language articles first published in China are referred to. For these, both the reference for the original article and that for the translation are given. The latter reference is typically either to the Foreign Bureau Information Service, China Daily Report, abbreviated 'FBIS-CHI' or to the Joint Publication Research Service, China Annual Report, abbreviated 'JPRS-CAR'. Both are US Government publications. A numbering system is used for these references. For example, 'FBIS-CHI-90-216' refers to number 216 of the FBIS-CHI, published in 1990. In these cases, the year given in brackets after the authors' names is the year in which the translation was published; unless otherwise indicated, this is also the year in which the original article was published.

Adelman, I. and Sunding, D. (1987) 'Economic Policy and Income Distribution in China', Journal of Comparative Economics, 11, 444-461.

Ahmad, E. and Stern, N. (1991) The Theory and Practice of Tax Reform in Developing Countries, Cambridge University Press.

Ahmad, E. and Wang, Yang (1991) 'Inequality and Poverty in China: institutional change and public policy 1978-1988', LSE, STICERD, China Programme Discussion Paper 14.

Aitchinson, J. and Brown, J.A.C. (1957) The Lognormal Distribution with Special References to its Uses in Economics, Cambridge University Press.

Anand, S. (1983) Inequality and Poverty in Malaysia, Oxford University Press.

Atkinson, A.B. (1970) 'On the Measurement of Inequality', Journal of Economic Theory, 2, 244-263; republished as Ch.1 in (1983) Social Justice and Public Policy, Wheatsheaf, Brighton.

Atkinson, A.B. (1987) 'On the Measurement of Poverty', Econometrica, 55, 749-763.

Atkinson, A.B. and Bourguignon, F. (1987) 'Income Distribution and Differences in Needs', in G.R. Feiwel (ed.) Arrow and the Foundations of the Theory of Economic Policy, Macmillan, London.

Atkinson, A.B. and Bourguignon, F. (1989) 'The Design of Direct Taxation and Family Benefits', Journal of Public Economics, 41, 32-9.

Atkinson, A.B. and Micklewright, J. (1992) Economic Transformation in Eastern Europe and the Distribution of Income, Cambridge University Press.

Bai Fan (1990) 'Strategic Options for Price Reform During the Eighth Five-Year Plan Period', Jingji Yanjiu (Economic Research), 9, 28-34, 20 September, translated in FBIS-CHI-90-216, 32-40, 7 November.

Beach, C. and Davidson, R. (1983) 'Distribution-Free Statistical Inference with Lorenz Curves and Income Shares', Review of Economic Studies, L, 723-735.

Beach, C. and Richmond, J. (1985) 'Joint Confidence Intervals for Income Shares and Lorenz Curves', International Economic Review, 26, 439-450.

Berger, R. (1982) 'Multiparameter Hypothesis Testing and Acceptance Sampling', Technometrics, 24, 295-300.

Bi Jingquan (1992) 'Price Reform in the Course of China's Economic Improvement and Rectification', Liaowang Overseas Edition (Chinese), 1, 10-11, 6 January, translated in FBIS-CHI-92-028, 23-25, 11 February.

Bishop, J., Formby, J., and Thistle, P. (1989) 'Statistical Inference, Income Distributions and Social Welfare', in D. J. Slotje (ed.) Research on Economic Inequality, JAI Press Inc, Greenwich, Conn.

Bishop, J., Chakraborti, S. and Thistle, P. (1989) 'Asymptotically Distribution Free Statistical Inference for Generalized Lorenz Curves', Review of Economics and Statistics, 71, 725-727.

Blackorby, C. and Donaldson, D. (1980) 'Ethical Indices for the Measurement of Poverty', Econometrica, 48, 1053-1060.

Bourguignon, F. (1979) 'Decomposable inequality measures', Econometrica, 47, 901-920.

Broome, J. (1989) 'What's the good of equality?' Ch.9 in Hey, J. (ed.) Current Issues in Microeconomics, Macmillan, London.

Central Statistical Office [UK] (1990) Economic Trends 439, Her Majesty's Stationery Office.

Central Statistical Office [UK] (1986) Social Trends 16, Her Majesty's Stationery Office.

Chakravarty, S.R. (1990) Ethical Social Index Numbers, Springer-Verlag, New York.

Chen, Chunlai, Watson, A. and Findlay, C. (1990) 'One state - two economies: current issues in China's rural industrialization', mimeo, China Economy Research Unit, University of Adelaide.

Chow, K.V., Chakraborti, S. and Thistle, P.D. (1990) 'Testing for Stochastic Dominance: a general approach', mimeo.

Clark, S., Hemming, R. and Ulph, D. (1981) 'On Indices for the Measurement of Poverty', The Economic Journal, 91, 515-526.

Cornes, R. (1992) 'Measuring the Distributional Impact of Public Goods', paper presented to World Bank Conference on 'Public Expenditures and the Poor: incidence and targeting', June 1992, Washington D.C.

Coulter, P.B. (1989) Measuring Inequality: a methodological handbook, Westview Press, Boulder.

Coulter, F.A.E, Cowell, F.A. and Jenkins, S.P. (1991) 'Family Fortunes in the 1970s and 1980s', University College of Swansea, Department of Economics, Discussion Paper Series, 91-12.

Cowell, F.A. (-) Ineq (manual), STICERD, London School of Economics

Cowell, F.A. (1977) Measuring Inequality, Phillip Allen, Oxford.

Cowell, F.A (1980) 'On the Structure of Additive Inequality Measures', Review of Economic Studies, 47, 521-531.

Cowell, F.A. (1988) 'Poverty Measures, Inequality and Decomposability' pp.149-156 in Bös, D., Rose, M. and Seidl, Ch. (eds.) Welfare and Efficiency in Public Economics, Springer-Verlag, Berlin.

Cowell, F.A. (1989) 'Sampling Variance and Decomposable Inequality Measures', Journal of Econometrics, 42, 27-41.

Cowell, F.A. and Kuga, K. (1981) 'Additivity and the Entropy Concept: an axiomatic approach to inequality measurement', Journal of Economic Theory, 25, 131-143.

Cowell, F.A. and Mehta, F. (1982) 'The Estimation and Interpolation of Inequality Measures' Review of Economic Studies, 49, 273-290.

Dalton, H. (1920) 'The measurement of the inequality of incomes', Economic Journal, 30, 3480-361.

Dardanoni, V. and Lambert, P. (1988) 'Welfare Rankings of Income Distributions: a role for the variance and some insights for tax reform', Social Choice and Welfare,5, 1-17.

Dasgupta, P., Sen, A.K., and Starrett, D. (1973) 'Notes on the Measurement of Inequality', Journal of Economic Theory, 6, 180-187.

Davies, J.B. and Shorrocks, A.F. (1989) 'Optimal grouping of income and wealth data', Journal of Econometrics, 42, 97-108.

Deaton, A. (1980) 'Measurement of Welfare: theory and practical guidelines', World Bank, LSMS Working Paper No.7.

Deaton, A. and Muellbauer, J. (1980) Economics and Consumer Behaviour, Cambridge University Press.

Donaldson, D. and Weymark, J.A. (1986) 'Properties of fixed-Population Poverty Indices', International Economic Review, 27, 667-688.

Drèze, J. and Sen, A. (1989) Hunger and Public Action, Clarendon, Oxford.

Fei, J.C., Ranis, H.G. and Kuo, S.W.Y. (1978) 'Growth and the Family Distribution of Income by Factor Components', Quarterly Journal of Economics, 92, 17-53.

Fomby, T.B. and Seo, T.K., eds. (1989) Studies in the Economics of Uncertainty in honour of Josef Hadar, Springer-Verlag, New York.

Foster, J., Greer, J. and Thorbecke, E. (1984) 'A Class of Decomposable Poverty Measures', Econometrica, 52, 761-766.

Foster, J. and Shorrocks, A. (1987) 'Subgroup Consistent Poverty Indices', University of Oxford, Nuffield College, Measurement and Modelling in Economics Discussion Paper 3.

Foster, J. and Shorrocks, A. (1988a) 'Poverty Orderings', Econometrica, 56(1), 173-177.

Foster, J. and Shorrocks, A. (1988b) 'Poverty Orderings and Welfare Dominance', Social Choice and Welfare, 5, 179-198.

Gale Johnson, D. The People's Republic of China, 1978-1990, International Centre for Economic Growth, San Francisco.

Gastwirth, J.L. (1971) 'A General Definition of the Lorenz Curve', Econometrica, 39, 1037-1038.

Gastwirth, J.L and Gail, M.H. (1985) 'Simple Asymptotically Distribution-Free Methods for Comparing Lorenz curves and Gini Indices Obtained from Complete Data', Ch.7 in R.L. Basmann and G.F. Rhodes, Jr. (eds.) Advances in Econometrics, Vol. 4, JAI Press, Greenwich.

Glewwe, P. and van der Gaag, J. (1988) 'Confronting Poverty in Developing Countries: definitions, information and policy', The World Bank, LSMS Working Paper 48.

Goedhart, T., Halberstadt, V., Kapetyn, A., and van Praag, B.M.S. (1977) 'The Poverty Line: concept and measurement', Journal of Human Resources, 12, 503-520.

Guo Yuanxi (1984) 'Let Some Enterprises and Workers Get Rich First', Jingji Yuanjiu (Economic Research), 9, 33-38, 20 September, translated in JPRS-CEA (Chinese Economic Affairs)-84-100, 11-21, 7 December.

Hadar, J. and Russel, W. (1969) 'Rules for Ordering Uncertain Prospects', American Economic Review, 59, March, 25-34.

Hamada, K. and Takayama, N. (1977) 'Censored Income Distributions and the Measurement of Poverty', Bulletin of the International Statistical Institute, Book 1, 47, 617-30.

Han Hongshu and Lu Shengliang (1991) 'Status of and Remedies for China's Agricultural Price Subsidies', Gaige (Reform), 5, 157-164, 20 September, translated in JPRS-CAR-91-069, 41-44, 3 December.

Hastings, N.A.J. and Peacock, J.B. (1974) Statistical Distributions: a handbook for students and practitioners, New York, John Wiley and Sons.

Hayase, Y. and Kawamata, S. (1990) Population Statistics of China, IDE Statistical Data Series 55, Institute of Developing Economies, Tokyo.

Howe, C. (1973) Wage Patterns and Wage Policy in Modern China, 1919-1972, Cambridge University Press.

Howes, S. and Jha, S. (1992) 'Urban Bias in Indian Public Distribution System', Economic and Political Weekly, 27(19), 9 May, 1022-1030.

Howes, S. and Lanjouw, P. (1991) 'Regional Variations in Living Standards in Urban China', LSE, STICERD, China Program Discussion Paper 17.

Hua Sheng, Zhang Xuejun and Luo Xiaopeng (1992) China - From Revolution to Reform, Macmillan, London.

Hussain, A. (1990) 'The Chinese Enterprise Reforms', LSE, STICERD, China Program Discussion Paper 5.

Hussain, A. and Stern, N. (1991) 'Effective Demand, Enterprise Reforms and Public Finance', LSE, STICERD, China Program Discussion Paper 10.

Hussain, A., Lanjouw, P. and Stern, N. (1991) 'The Distribution of Income in China: Evidence from Household Survey Data', LSE, STICERD, China Program Discussion Paper 16.

Jäntti, M. (1992) 'Statistical Inference and Poverty Dominance', mimeo.

Jenkins, S.P. (1991) 'Income Inequality and Living Standards: changes in the 1970s and 1980s', Fiscal Studies, 12, 1-28.

Jewitt, I.D. (1981) Applications of the Theory of Consumer Behaviour under Certainty and Uncertainty, D.Phil. thesis, University of York.

Jia Dechang (1990) 'Squeezing onto a Bus, Putting a Wad of Paper under the Leg of Table to Make it Level, and Eating from the Same "Big Rice Pot"', Jingji Ribao [Economic Daily], 3, 22 June, translated in FBIS-CHI-90-137, 33-33, 17 July.

Kakwani, N. (1976) 'On the Estimation of Inequality Measures from Grouped Observations', Review of Economic Studies, 43, 483-492.

Kakwani, N. (1980a) Income Inequality and Poverty: methods of estimation and policy applications, Oxford University Press.
Kakwani, N. (1980b) 'On a Class of Poverty Measures', Econometrica, 48, 437-446.
Kakwani, N. (1984) 'Welfare Ranking of Income Distributions' in Basmann, R and Rhodes, G. (eds.) Advances in Econometrics vol.3, JAI Press, Greenwich.
Kakwani, N. (1990) 'Testing for Significance of Poverty Differences With Application to Cote d'Ivoire', The University of New South Wales, School of Economics Discussion Paper, 90/3.
Kendall, M.G. and Stuart, A. (1963) The Advanced Theory of Statistics, Vol. 1, Distribution Theory, Second Edition, Charles Griffin and Company Limited, London.
Khan, A.R, Griffen, K., Riskin, C. and Zhao Renwei (1991) 'Household Income and its Distribution in China', University of California Department of Economics, Discussion Paper 92-3.
Knight, J. and Song, Lina (1990) 'The Determinants of Urban Income Inequality in China', University of Oxford, Applied Economics Discussion Paper 91.
Kolm, S.C. (1976) 'Unequal Inequalities' I and II, Journal of Economic Theory, 12, 416-442 and 13, 82-111.
Lambert, P. (1989) The Distribution and Redistribution of Income: a mathematical analysis, Basil Blackwell, Cambridge, Mass.
Lanjouw, P. (1992) 'Inequality, Poverty and Mobility: the experience of a North Indian Village', Ph.D. thesis, London School of Economics.
Li Chengrui (1986) 'Guanyu Zhongguo Jinniande Jingji Zhengce Dui Jumin Shouru he Xiaofei Zhuangkuang Yingxiang de Tongji Baogao' (A Statistical Report Concerning the Influence of Current Policy On Household Income and Consumption) Jingji Yanjiu, 1, January.
[LTN] Liaoning Statistical Yearbook Compiling Committee (1991) Liaoning Tongji Nianjian, 1991 (Liaoning Statistical Yearbook, 1991), Chinese Statistical Publishing House, Beijing.
McFadden, D. (1989) 'Testing for Stochastic Dominance' Ch. 7 in Fomby and Seo (1989)
Ma Guonan (1991) 'China's Income Distribution in the 1980s', mimeo, Department of Economics, RSPacS, Australian National University.
Mahalanobis, P.C. (1960) 'A Method of Fractile Graphical Analysis', Econometrica, 28, 325-351.
Mendenhall, W., Schaeffer, R. and Wackerly, D. (1986) Mathematical Statistics with Applications (Third Edition), Duxbury Press, Boston.
Mood, A.M., Graybill, F.A., and Boes, D.C. (1963) Introduction to the Theory of Statistics, Third Edition, McGraw-Hill, Tokyo.
National Population Census Office Under the State Council (1991a) Major Figures Of The Fourth Population Census of China, China Statistical Publishing House.
National Population Census Office Under the State Council (1991b) Zhongguo 1990 Nian Renkou Diaocha 10% Chao yang Ziliao (10% Sampling Tabulation on the 1990 Population Census of the People's Republic of China), China Statistical Publishing House.
Nozick, R. (1974) Anarchy, State and Utopia, Basic Books, New York.
Nygård, F. and Sandström, A. (1981) Measuring Income Inequality, Almqvist & Wiksell International, Stockholm.
Pen, J. (1971) Income Distribution, trans. T.S. Preston, Allen Lane, London.
Perkins, D.H. (1988) 'Reforming China's Economic System', Journal of Economic Literature, 36, 601-645.
Phelps Brown, H. (1988) Egalitarianism and the Generation of Inequality, Oxford University Press.
Pratt, J. (1964) 'Risk Aversion in the Small and in the Large', Econometrica, 32, 122-136.
Preston, I. (1992) 'Large Sample Estimation and Inference for Poverty Measures', University College London, Department of Economics, Discussion Paper 92-02.
Pudney, S. and Wang Limin (1991) 'Rationing and Consumer Demand in China: simulating the effects of a reform of the urban food pricing system', LSE, STICERD, China Program Discussion Paper 15.
Randles, R.H. and Wolfe, D. (1979) Introduction to the Theory of Nonparametric Statistics, John Wiley and Sons, New York.
Rao, C.R. (1973) Linear Statistical Inference and its Applications, John Wiley, New York.

Ravallion, M. (1992) 'Poverty Comparisons: a guide to concepts and methods', The World Bank, LSMS Working Paper Number 88.

Rawls, J. (1971) A Theory of Justice, Harvard University Press, Cambridge, Mass.

Ren Yun and Wang Limin (1992) 'Chinese Urban Household Survey', mimeo, STICERD, London School of Economics.

Riskin, C. (1987) China's Political Economy: the quest for development since 1949, Oxford University Press.

Riskin, C. (1991) 'Rural Poverty in Post-reform China', paper presented to Conference on China's Reform and Growth, A.N.U., Canberra, mimeo.

Rothschild, M. and Stiglitz, J. (1970) 'Increasing Risk: 1. A definition', Journal of Economic Theory, 2, 225-243.

Rothschild, M. and Stiglitz, J. (1973) 'Notes on the Measurement of Inequality', Journal of Economic Theory, 6, 180-203.

Savin, N.E. (1984) 'Multiple Hypothesis Testing', Ch. 14 in Griliches, Z. and Intriligator, M.D. (eds.) Handbook of Econometrics, Vol.2, North Holland.

Sen, A. (1970) Collective Choice and Social Welfare, Oliver and Boyd, London.

Sen, A. (1973) On Economic Inequality, Clarendon Press, Oxford.

Sen, A. (1976) 'Poverty: an ordinal approach to measurement', Econometrica, 44, 219-231; republished as Ch.17 in Sen (1982).

Sen, A. (1982) Choice, Welfare and Measurement, Basil Blackwell, Oxford.

Sen, A. (1987) Economics and Ethics, Basil Blackwell, Oxford.

Sen, A. (1992) Inequality Reexamined, Clarendon Press, Oxford.

Sen, A. and Williams, B., eds. (1982) Utilitarianism and Beyond, Cambridge University Press, Cambridge.

Shorrocks, A. (1980) 'The Class of Additively Decomposable Inequality Measures', Econometrica, 48, 613-625.

Shorrocks, A. (1982) 'Inequality Decomposition by Factor Components' Econometrica, 50, 193-211.

Shorrocks, A. (1983) 'Ranking Income Distributions', Economica, 50, 3-17.

[STN] Sichuan Statistical Bureau (1991) Sichuan Tongji Nianjian, 1991, (Sichuan Statistical Yearbook, 1991), Chinese Statistical Publishing House, Beijing.

[SSB] State Statistical Bureau (1985) Chengshi Zhuhu Diaocha Tianbao Xize (City Household Survey Questionnaire Guidelines), Guojia Tongjiju Zhiding (Formulated by the State Statisical Bureau).

State Statistical Bureau (1986ye) Statistical Yearbook of China, 1986, Economic Information and Agency, Hong Kong and China Statistical Information and Consultancy, Beijing.

State Statistical Bureau (1987s) 'Liu Wu' Qijian Woguo Chengzhen Jumin Jiating Shouzhi Diaocha Ziliao ('Sixth Five Year Plan' National Urban Residential Households Income and Expenditure Sample Survey Data).

State Statistical Bureau (1988y) Zhongguo Tongji Nianjian, 1988 (Statistical Yearbook of China, 1988), Chinese Statistical Publishing House, Beijing.

State Statistical Bureau (1988s) 1987 Quanguo Chengzhen Jumin Jiating Shouzhi Diaocha Ziliao (National Urban Households Income and Expenditure Survey, 1987), Chinese Statistical Publishing House, Beijing.

State Statistical Bureau (1987y) Zhongguo Tongji Nianjian, 1987 (Statistical Yearbook of China, 1987), Chinese Statistical Publishing House, Beijing.

State Statistical Bureau (1989s) 1988 Quanguo Chengzhen Jumin Jiating Shouzhi Diaocha Ziliao (National Urban Households Income and Expenditure Survey, 1988), Chinese Statistical Publishing House, Beijing.

State Statistical Bureau (1988se) A Survey of Income and Expenditure of Urban Households in China, 1985, China Statistical Information and Consultancy Service Centre, Beijing and the East-West Population Insitute, East-West Center, Honolulu.

State Statistical Bureau (1989se) A Survey of Income and Expenditure of Urban Households in China, 1986, China Statistical Information and Consultancy Service Centre, Beijing and the

East-West Population Insitute, East-West Center, Honolulu.
State Statistical Bureau (1989y) Zhongguo Tongji Nianjian, 1989 (Statistical Yearbook of China, 1989), Chinese Statistical Publishing House, Beijing.
State Statistical Bureau (1990a) Quanguo Gesheng, Zizhiqu, Zhixiashi Lishi Tongji Ziliao Huibian, 1949-1989 (Compilation of Statistics for Provinces, Autonomous Regions, Municipalities and all-China, 1949-1989), China Statistical Publishing House.
State Statistical Bureau (1990b) Zhongguo Wujia Tongji Nianjian, 1990 (China Prices Statistical Yearbook), Chinese Statistical Publishing House, Beijing.
State Statistical Bureau (1990s) 1989 Quanguo Chengzhen Jumin Jiating Shouzhi Diaocha Ziliao (National Urban Households Income and Expenditure Survey, 1989), Chinese Statistical Publishing House, Beijing.
State Statistical Bureau (1990y) Zhongguo Tongji Nianjian, 1990 (Statistical Yearbook of China, 1990), Chinese Statistical Publishing House, Beijing.
State Statistical Bureau (1990ye) China Statistical Yearbook, 1989, International Centre for the Advancement of Science and Technology Ltd. and China Statistical Information and Consultancy Service Center.
State Statistical Bureau (1991a) 'Qiwu' Shiqi Guoming Jingji he Shehui Fazhan Gaikuang (Summary of National Economic and Social Development in the Seventh Five Year Plan Period), Chinese Statistical Publishing House, Beijing.
State Statistical Bureau (1991y) Zhongguo Tongji Nianjian, 1991 (Statistical Yearbook of China, 1991), Chinese Statistical Publishing House, Beijing.
Stoline, M.R. and Ury, H.K (1979) 'Tables of the Studentized Maximum Modulus Distribution and an Application to Multiple Comparisons Among Means', Technometrics, 21, 87-93.
Summers, R. and Heston, A. (1988) 'A New Set of International Comparisons of Real Product and Prices: Estimates for 130 Countries, 1950-1985', The Review of Income and Wealth, 34, 1-27.
Thistle, P.D. (1989) 'Ranking Distributions with Generalized Lorenz Curves', Southern Economic Journal, 56, 1-12.
Thistle, P.D. (1990) 'Large Sample Properties of Two Inequality Indices', Econometrica, 58, 725-728.
Thon, D. (1979) 'On Measuring Poverty', Review of Income and Wealth, 25, 429-439.
Tolley, H.D. and Pope, R.D. (1988) 'Testing for Stochastic Dominance', American Journal of Agricultural Economics, 70, 692-700.
UNDP (United Nations Development Program) (1990) Human Development Report, 1990, Oxford University Press.
Watts, H.W. (1968) 'An Economic Definition of Poverty' Ch.11 in Moynihan, D.P. (ed.) On Understanding Poverty, Basic Books, New York.
Whitmore, G.A. (1970) 'Third Degree Stochastic Dominance', American Economic Review, 24, 617-628.
Whitmore, G.A. (1978) 'Statistical Tests for Stochastic Dominance' Section 8.2, Ch. 8 in Whitmore and Findlay (1978).
Whitmore, G.A. and Findlay, M.C., eds. (1978) Stochastic Dominance: an approach to decision-making under risk, Lexington Books, Toronto.
World Bank (1985) China: economic structure in international perspective, Annex 5 to China, Long-term Development Issues and Options, A World Bank Country Report, The World Bank, Washington.
World Bank (1991) World Development Report 1991, Oxford University Press.
Xinhua (Domestic service) (1990) 'From "Turning Pale at the Mention of Prices" to "Staying Calm at Hearing about Price Hikes", Our Citizens' Ability to Withstand Price Reform is Strengthened', 27 July, translated in FBIS-CHI-90-191, 36, 2 October.
Xinhua (English service) (1991) 'Prices of Rationed Grain, Edible Oil to Rise', 27 April, reproduced in FBIS-CHI-91-082, 29 April, pp.39-40.
Xinhua (English service) (1992) 'State Grain Sales Price to Match Buying Price', 18 March, reproduced in FBIS-CHI-92-054, 19 March, p.53.
Yao Jinguan (1992) 'On the Fluctuation and Regulation of the Agricultural Product Prices under

Chinese Market Economic System', mimeo, Hebei Institute of Finance and Economics, Shijiazhuang, Hebei.

Zhang Ping (1990) 'An Analysis of the Effects of Price Subsidies on the Distribution of Residents' Income in China', Jingji Yanjiu, 4, 20 April, translated in JPRS-CAR-90-060, 26-33, 6 August.

Zhao Hongyue (1991) 'Price Subsidy in China', Liaowang, 25, 16-17, 24 June, translated in FBIS-CHI-91-149, 36-39, 2 August.

Zhao Renwei (1990b) 'Income Distribution', Ch. 11 in Nolan, P. and Dong Fureng (eds.) (1990) The Chinese Economy and its Future: achievements and problems of post-Mao reform, Polity Press, London.

Zhao Renwei (1990a) 'Some Views on the Question of Unfair Income Distribution at Present', Jingji Yanjiu (Economic Research), 12, 6-10, 20 December, translated in FBIS-CHI-90-082, 33-38, 27 April.