This version: 10/24/02

Globalization and Inequality Without Differences in Data Definition, Legal System and Other Institutions

Shang-Jin Wei #
International Monetary Fund

and

Yi Wu
Georgetown University

Abstract

Cross-country studies of the impact of globalization on growth or on inequality have been criticized because differences in legal systems and other institutions across countries are difficult to control for, and the inequality data across countries may not be compatible. An in-depth case study of a particular country's experience can provide a useful complement to cross-country regressions. In this paper, we provide a case study of China using two unique data sets on regions and households. In response to an increase in openness, we find that (a) urban-rural income inequality tends to decline, (b) inequality within a city tends to rise, and (c) inequality within a rural area tends to decline. Putting together the three pieces of information, we find that greater openness contributes a (modest) reduction in the overall inequality. The negative association between openness and inequality holds up when we apply a geography-based instrumental variable approach to correct for possible endogeneity of a region's trade openness.

JEL Codes: F1 and O1
Key Words: Globalization, Inequality, Openness, and Growth

Shang-Jin Wei (corresponding author) is an Advisor at the IMF's Research Department. Mailing address: International Monetary Fund, Room 10-700m, 19th Street, N.W., Washington, DC 20431. Tel: (202) 623-5980. Fax: (202) 623-7271. Email: swei@imf.org.

** We would like to thank Lee Branstetter, David Dollar, Yi Feng, Ed Glaser, Antonio Spilimbergo, Harm Zebregs, and seminar participants at Boston University and the World Bank for useful comments, Azizur Rahman Khan, Tam Ho-Kwong, Zhang Zhanxin, and especially Li Shi at China Academy of Social Sciences, Wang Youjuan, Tang Ping, and Jiang Mingqing at China's State Statistics Bureau for help in interpreting the data, and Heather Milkiewicz, Hayden Smith and Donald Tench IV for editorial assistance. The views are the authors' own, and may not be shared by the IMF or any other organization that the authors are or have been associated with.

"[C]ross-country regressions…are not the best tools for analyzing…the linkage between trade and growth." "[T]he most compelling evidence on this issue can come only from careful case studies…"

T.N. Srinivasan and Jagdish Bhagwati (1999, p9, and the abstract)

1. Introduction

The effect of openness on income inequality sometimes arouses emotion and blood pressure as well as academic curiosity. There exists an active empirical literature on economic growth and income inequality, and a related and equally active literature on openness and economic growth. Most of the papers in these two literatures employ cross-country regressions. Prominent examples include Forbes (2000), Dollar and Kraay (2001a and 2001b), Edwards (1992, 1998), Sachs and Warner (1995), Frankel and Romer (1999), Rodriguez and Rodrik (2000) and other papers cited therein. Useful insights have been gained from this literature.

However, analyses based on cross-country regressions have been criticized on two grounds. The first type of problems has to do with data comparability across countries. This issue is particularly acute for data on income inequality: the definitions and data collection methods can be different across countries. As an illustration, for OECD countries, Atkinson and Brandolini (2001) noted the pitfalls in making cross-country comparisons based on pooled data. For a few countries where multiple measures of income distribution are available (households versus individuals, income versus consumption, etc.), the different measures can give different, sometimes contradictory, patterns even for the same time periods. Since the data that cross-country regressions have to rely on come from potentially different methodologies, they can produce misleading results when pooled together. Atkinson and Brandolini noted further that "in cross-country analysis, use of a dummy variable adjustment for data differences is not appropriate." Atkinson and Brandolini's specific criticism is on pooling data across OECD countries. It is reasonable to assume that the data quality for developing countries is generally inferior to the OECD countries. Therefore, running cross-country regressions involving data from developing countries can only make the quality of inference worse.

Aside from the Atkinson-Brandolini criticism just mentioned, we should note another potential source of data incomparability. The validity of comparing living standards across countries depends on the validity of the so-called purchasing-power-parity adjustment, which in turn depends on the assumption that a common "representative consumption basket" can be meaningfully constructed for all countries. The last assumption cannot always be taken for granted.

The second difficulty with cross-country studies has to do with the fact that differences in cultures, legal systems, or other institutions other than openness may also be relevant for the outcome variable under study (e.g., economic growth or income inequality). These factors are difficult to be quantified and therefore to be controlled for in cross-country regressions. Inclusion of fixed effects in panel regressions helps. However, the myriad of country-specific institutions may also interact with the key regressor under investigation (e.g., openness) to affect the outcome variable (e.g., income inequality). For example, in response to a terms-of-trade shock, some countries would let the poor to fend for themselves, while others would have a social safety net to moderate the negative impact on the poor but the exact size of the income transfer may not be proportional to the size of the shock depending both on the nature of the social safety net and on the size of the shock. In this case, the usual fixed effects are not sufficient to control for the influence of the country-specific institutions.

In an influential paper, T.N. Srinivasan and Jagdish Bhagwati (1999) asserted that cross-country regressions are deficient and cannot be relied upon to understand the impact of globalization on economic growth and presumably nor on income inequality (see the quote at the beginning of the paper). One may not agree completely with the strong assertion by Srinivasan and Bhagwati (1999). Nonetheless, given these criticisms, a careful study of cross-regional experience within a single country can, at a minimum, provide a useful complement to the literature based on cross-country regressions. Within a given country and over a relatively short time period, culture, legal system or other institutions can more plausibly be held constant. So the researchers' ability to isolate the effect of openness is enhanced. Furthermore, the comparability of data definition and collection method is, in principle, also higher within a single country than across multiple countries.

This paper presents a case study of the impact of globalization on income disparity by pooling two unique data sets on Chinese regions. There are five reasons that make China a good case study. Some of them have to do with the fact that China is an important country *per se*. But, perhaps more importantly, other factors (or peculiar features of China) provide a methodological advantage relative to typical cross-country studies.

First, China is the largest developing country that has embraced trade openness in the last quarter of a century. The change in openness over time has been dramatic in magnitude. Before 1978, the country had relatively little trade with the rest of the world. In 1978, Deng Xiaoping-led Chinese government formally adopted "opening-to-the-outside-world" as a national economic policy. Since then, the trade-to-GDP ratio quadrupled (from 8.5% in 1977 to 36.5% in 1999[1]).

Second, due to unequal natural barriers to trade (i.e., distance to major seaports), the effective increase in openness varies widely across different regions in China. This variation across space provides a good opportunity to study the impact of openness on inequality (and potentially other issues as well) while holding constant the legal system, macroeconomic policies, culture and a host of other factors.

Third, China is a large country, which implies a relatively large number of regional observations which is convenient for a statistical analysis. For this reason, China provides better material for an in-depth case study than, say, Argentina, Bangladesh or Costa Rica, whose trade-to-GDP ratios also rose dramatically during the same period, but whose territories are much smaller.

Fourth, the Chinese government imposes a variety of restrictions on the ability of its citizens to move from rural to urban areas, and from one region to another. For example, virtually all local governments maintain a household registration system. If a migrant from a different region has not obtained the requisite local "citizenship," which generally is not easy, he (she) and his (her) family are typically discriminated along multiple dimensions, including access to health facilities, and the ability for the children to attend local schools. As a consequence, the internal migration that is permanent and outside a few "special economic zones" is much lower than it otherwise would be the case. To be sure, these restrictions have been loosened over time, but were more binding during our sample period. While the restrictions on internal migration may have adversely affected the welfare of the migrants, they

[1] Trade data are from IMF's International Financial Statistics, various issues.

paradoxically make regions within China more like separate countries than otherwise. In other words, the cross-regional study reported in this paper is closer to a cross-country study in spirit than a similar study done for a different country that does not have equally stringent restrictions on internal migration.

Fifth, China's geography also turns out to be convenient for the type of statistical analysis that this paper carries out. An important issue in this research is that openness (or increases in openness) may be endogenous. In the literature, Frankel and Romer (1999) pioneered the technique of using geography as an instrument for openness: Geography is an important determinant of trade, and is arguably exogenous with respect to economic growth or income inequality. It would be desirable to adopt a similar instrumental variable strategy for single-country case studies. China is semi-landlocked. It has a coast on the East and Southeast sides, but is surrounded on other sides by tall mountains, deserts, or foreign territories that are minor participants in international trade. To a large degree, the differences across Chinese cities in terms of participation in international trade are due to their varying distance from a major seaport. In fact, two seaports alone - Hong Kong and Shanghai - handled approximately 50% of China's total trade with the rest of the world during our sample period. This makes China a suitable case to apply the Frankel-Romer technique: we can construct an instrumental variable for a city's openness based on its access to major seaports. Such an instrumental variable would be much more difficult to construct for countries like the United States or Indonesia whose access points to international trade are more diffused.

At this point, it is useful to take note from the economic theories on the predicted impact of trade openness on income inequality for a labor-abundant country like China. A natural starting point is the Heckscher-Ohlin model. This model predicts that, subsequent to an increase in trade openness, the return to relatively abundant factors will increase relative to the return to relatively scarce factors. China is a country abundant in unskilled labor. The impact of an increase in trade openness should raise the income of the unskilled labor faster than that of the rest of the population. If one adopts the (seemingly reasonable) assumption that unskilled labor has a lower income to start with, then openness should lower inequality.

However, the prediction from the Heckscher-Ohlin model is not the only one available. In a specific-factor model, the impact of openness on inequality, at least in the short run, is ambiguous. More generally, the theoretical predictions with regard to the effects of openness on growth, and of growth on inequality, are ambiguous, depending on which theoretical model one uses. This point has been amply demonstrated by Rodriguez and Rodrik (2000) and Srinivasan and Bhagwati (1999). Given the ambiguity in the theories, the effect of openness on income inequality is a matter to be settled by the data.

While the Chinese economy has dramatically increased its openness over the last two decades, the overall income inequality has risen as well. For example, Khan and Riskin (1998) estimated that the Gini coefficient in China increased from 38.2 in 1988 to 45.2 in 1995. With a different data set and methodology, the World Bank (1997) also estimated that the Gini coefficient in China had increased (from 28.8 in 1981 to 38.8 in 1995). From these aggregate statistics, it is tempting to conclude that embracing globalization has contributed to the rise in inequality. But this is not correct, as we will show later.

Our main findings can be summarized briefly here. First, openness contributes to a reduction in the urban-rural inequality. Second, openness appears to raise (moderately) the inequality within urban areas. Third, openness is associated with a reduction in the inequality within rural areas. Putting together these three pieces of information, we show that trade openness *per se* has most likely triggered a reduction in the overall inequality.

In addition to the various cross-country studies cited above, there are a large number of research papers that look into either Chinese economic development or foreign trade performance. They are too numerous to cite them all. Qian and Weingast (1996), and Jin, Qian, and Weingast (2001) investigated the role of fiscal decentralization in China's development, using a province-level data set. Perkins and Yusuf (1984), Khan, Griffin, Riskin and Zhao (1992), Khan, Riskin and Zhao (1993), Hussain, Lanjouw and Stern (1994), Kwong (1994), Khan and Riskin (1998 and 2001), Yang (1999), Zhao (1999), and Knight and Song (1999), among others, examined income inequality in China (but not the impact of openness on inequality). Lardy (1992) is a classic reference on the Chinese trading regime during the reform era. Wei (1995) and Wang (1994) produced the first two papers that employed a city-level data set. Of the two, Wei (1995) examined the connection between openness and economic growth across the Chinese cities and argued that openness has promoted growth.

The paper is organized as follows. Section 2 describes the data sets used in the analysis. Section 3 presents the statistical analysis on the urban-rural inequality. Section 4 brings in the discussion of within-urban and within-rural inequalities which uses a different, household survey-based data set. Finally, Section 5 provides some concluding remarks.

2. Data

We employ two distinct sets of data. The first one, which we collect from hard copies of the statistical yearbooks and manually input into machine-readable format, allows us to examine urban-rural income inequality. The second one, which is a combination of two household surveys conducted in 1988 and 1995, allows us to examine inequalities within urban areas and within rural areas, respectively. We explain the two data sets in turn.

Data Used to Examine Urban-Rural Inequality

The first data set comes mainly from two sources: (1) Urban Statistical Yearbook, various issues, published by China's State Statistical Bureau, and (2) Fifty Years of the Cities in New China: 1949-1998, also published by the State Statistical Bureau[2].

The central variable of interest is the gap between urban and rural incomes. In order to explain the data set clearly, it is useful to provide a brief description of the Chinese administrative structure (see Figure 1). The entire country is divided into 27 provinces plus three province-status "super-cities" -- Beijing, Shanghai and Tianjin[3]. In each province (or super-city), the population is further divided into rural counties and cities. In many instances, cities are given the administrative power over the adjacent rural counties. For example, the municipal government of Shanghai administers 10 rural counties in addition to its urban area. The municipal government of Wuhan administers 3 rural counties. In 1994, 783 counties, or approximately 45% of all rural counties are administered by a total of 193 cities. The rest of the

[2] Most data come from the first source, except for GDP and output in second and tertiary sectors in 1992 and 1993, which are missing and need to be supplemented by the second source.

[3] The official term for super-cities are "directly administered cities" – meaning that the city officials report directly to the central government just as the officials in other provinces. Since 1997, Chongqing has been made a fourth "super-city." Note that the data set does not include Hong Kong SAR, Macau SAR and Taiwan Province of China.

counties are not attached to any city (and administered mostly by prefecture governments, which are one level below provincial governments). Our data set consists of information on those cities that administer some adjacent rural counties. In other words, we exclude those counties that are unattached to any cities. We also leave out those (typically smaller) cities that are not authorized to administer any rural counties. In the rest of this paper, the term "city" is used to refer to an administrative unit with an urban area plus the adjacent rural counties under its jurisdiction.

For each variable (say, GDP or population) and each city (say, Wuhan), our data set provides information both at the level of a city (i.e., including the adjacent rural areas) and at the level of urban area of the city (i.e., excluding the rural areas). Let us use GDP(u, k, t) and POP(u, k, t) to denote GDP and population, respectively, for the urban part of city k in year t. Similarly, GDP(c, k, t) and POP(c, k, t) are used to denote GDP and population in year t for the entire city k that encompasses the adjacent rural areas. From this raw data, we can deduce the GDP and population for the associated rural areas alone, denoted by GDP(r, k, t) and POP (r, k, t), respectively. Obviously, GDP(r, k, t) =GDP(c, k, t) – GDP(u, k, t) and POP(r, k, t) = POP(c, k, t) – POP(u, k, t).

We measure the urban-rural income inequality in year t, Q(k, t), by the ratio of the two respective per capita incomes:

$$Q(k, t) = [GDP(u, k, t) / POP(u, k, t)] / [GDP(r, k, t) / POP(r, k, t)].$$

Our data source also reports values of exports by cities, but not information on imports dis-aggregated by cities. We define EXP(k, t) as the value of export in local currency by city k at year t. We then measure openness of city k in year t by

$$OPEN(k,t) = EXP(k, t)/GDP(k, t).$$

Note that the export data are available only at the city level (rather than separately for urban and rural areas). We omit the city subscript "c" in the definition of openness.

In the subsequent statistical analysis, we focus on the change of this inequality measure from 1988 to 1993, Q(k,93) – Q(k, 88). The beginning and ending points are constrained by data availability. 1988 was the first year when it became possible for us to compute Q(k, t) for a sufficient number of cities. 1993 was the last year when EXP(k, t) data were reported. As the reform deepened, an increasing number of corporations (including foreign-invested firms and some Chinese state-owned firms) earned the rights to conduct export and import business directly without having to go through the state-owned trading companies. A fraction of these exports did not get properly recorded at the city level. By 1994, the State Statistics Bureau made the judgment that the export data at the city level lost relevance due to this kind of reporting leakage and stopped reporting them altogether in its subsequent Urban Statistical Yearbooks. The consequence of this reporting leakage for our statistical analysis will be discussed later.

Before we implement the statistical analysis, we undertake a data cleanup. This includes (a) eliminating cities whose jurisdiction (e.g., number of counties under its administration) have changed over the sample period, and (b) correcting observations that appear clearly erroneous to us when checked against related series. The detail of the data cleanup is explained in an appendix. The final sample includes one hundred cities scattered around the country.

Table 1a reports the summary statistics for the key variables in the paper. A number of interesting observations can be made. First, at any given point in time and for a given variable, there is a tremendous amount of heterogeneity across space. For example, in 1988, the urban-rural income ratio, averaged across all the cities in the sample, was 2.87. But the standard deviation of this ratio was 1.17. The ratio spread from 1.09 for the most equal city to 7.33 for the most unequal city. Similarly, the export-to-GDP ratio in 1988 had a mean of 7.76 percent, but a standard deviation of 4.75. This spatial variation in openness and in inequality is necessary for us to conduct a meaningful statistical analysis. Second, over the six-year period from 1988 to 1993, while the average openness increased, the urban-rural income inequality had also gone up. However, as we will show later, it would be misleading to conclude from these statistics that greater openness has contributed to greater inequality.

Data Used to Examine Within-Urban and Within-Rural Inequalities

The second data set consists of two surveys of households conducted in 1988 and 1995, and designed by a group of international economists and the Economics Institute of the Chinese Academy of Social Sciences. The data set is described in more detail in Khan and Riskin (1998 and 2001). Briefly, the 1988 survey covered 10,258 rural households (or 51,352 persons) in 28 provinces and 9,009 urban households (or 31,827 persons) in 10 provinces. The 1995 survey covered 7,998 rural households (or 34,739 persons) in 19 provinces and 6,931 urban households (or 21,694 persons) in 11 provinces.

The original idea of the surveys was to obtain a more accurate measure of income in China by including the market value of various items of incomes in kind (i.e., pigs that farmers raise for own consumption) and the rental value of houses. Many of these items were missed in China's official statistical surveys. With these more accurate income measures at household levels, Khan and Riskin (1998 and 2001) then computed measures of income inequality at the national and provincial levels.

We use the data on household surveys for a purpose that is different from Khan and Riskin. Our objective is not just to measure inequality, but to examine the impact of trade openness on inequalities within rural areas and within urban areas. There are ten provinces in total that cover both urban and rural samples and appear in both 1988 and 1995. Ten observations are not enough to permit sufficiently powerful statistical tests. Instead, we go back to raw survey data and compute measures of inequality (Gini and Theil indexes) at the level of rural counties and urban areas. To be more precise, there are a total of 40 rural counties and 39 urban areas that appear in both 1988 and 1995 surveys. For each of these 79 regional units, we can trace out the evolution of income inequality over time.

Summary statistics for this data set are presented in Table 2. The upper panel reports the information for the 40 rural counties. The average inequality within rural counties as measured by Gini coefficient was 0.24 in 1988 and rose to 0.29 in 1995. Inequality within rural areas measured by the Theil index gives the same impression; it rose from 0.088 in 1988 to 0.178 in 1995. The nominal GDP per capita (in current price) was 809 Chinese Yuan in 1988 and 2,381 Chinese Yuan in 1995[4].

[4] For comparison, the rural per capita income for the whole sample was 784 yuan in 1988 and 2372 yuan in 1995, respectively. So the average income in our sample of 40 counties was somewhat higher than the whole sample in both years.

The lower panel reports the summary statistics across the 39 urban areas that are common in both years' surveys. Measured by the Gini coefficients, the average inequality within urban areas rose from 0.18 to 0.22. Measured by the Theil indexes, the average inequality with urban areas also rose from 0.058 to 0.083. So the inequality within the urban areas started lower than inequality within the rural areas, and rose a bit more slowly as well. The average nominal income, however, was quite a bit higher in the urban areas than in the rural counties in both years. It was 1,760 Chinese yuan in 1988 and increased to 5,564 yuan in 1995 (without adjustment for inflation)[5].

3. Statistical Analysis: Openness and Urban-Rural Inequality

In this section, we present an analysis of urban-rural inequality. Even though urban-rural inequality is only a part of China's overall inequality, it is a dominant part. This is because the poor in China are disproportionately found in the rural areas. The World Bank (1997) estimated that the urban-rural income inequality accounted for more than half of the overall income inequality in 1995, and the change in the urban-rural inequality explained about 75 percent of the change in the overall income inequality during 1984-1995.

Similar conclusions were reached in research papers that investigated a particular region or regions in China. For example, during 1986 to 1994, Yang (1999) estimated that the urban-rural inequality explained 82 percent of the change in the overall income inequality in Jiangsu province, and virtually all of the change in the overall inequality in Sichuan province. Yao and Zhu (1998) found that, in Sichuan and Liaoning provinces from 1988 to 1990, the urban-rural inequality accounted for between 47 to 51 percent of the overall income inequality. The contribution of the urban-rural inequality to the overall income inequality has been more or less stable over time, according to Kanbur and Zhang (1999), or has declined by a moderate amount, according to Khan and Riskin (1998). To summarize, while the exact estimates differ among the

[5] For comparison, the urban per capital income for the whole sample was 1889 yuan in 1988 and 5233 yuan in 1995, respectively.

researchers, all agree that the urban-rural inequality is a major component of the overall income inequality in China.

It is worth noting two features about our data. First, approximately 45% of all rural counties (which numbered 783 nationally in 1994) are administered by some cities. Hence, the urban and adjacent rural areas that enter our data base potentially cover a significant part of China[6]. Second, rural areas administered by the cities typically are more populous than the related urban areas. On average, the population in the "periphery" - adjacent rural counties - is about twice as big as that in the center – the corresponding urban area. The actual ratio of urban-rural populations varies greatly across China, from a low of only 5% to a high of nearly 300% (See Table 1 for details).

Benchmark regression

Let $q(k, t) \equiv \log Q(k,t)$ denote the urban-rural inequality (in log) for city k in year t. We conceptualize that the inequality in a given city depends on a number of factors:

$$(1) \qquad q(k,t) = \alpha + \beta \, open(k, t) + \gamma \, z(k) + h(k, t)\,\Theta + N(t)\,\Lambda + e(k, t)$$

This framework decomposes all the factors into four categories. (1) *open(k, t)* $\equiv$ log[EXP(k, t)/GDP(k, t)] is the degree of trade openness (in log) for city k in year t. (2) z(k) is a summary of city-specific factors that do not change over time. (3) h(k,t) is a vector of factors other than open(k, t) that are specific to city k and do change over time. Average income level of the city is one example in this category. (4) In addition to the city-specific factors, a number of national factors may affect the level of urban-rural inequality in all cities. Inflation rate and the terms of trade for agricultural products are two such examples. These factors are represented by the vector N(t).

In Equation (1), α, β, γ, Θ and Λ are parameters (or vectors of parameters of appropriate dimensions). The effect of embracing globalization on inequality is captured by the parameter β,

6 Due to changes in jurisdictions, missing data or apparent errors in the data, only a subset of the urban and rural areas in our original data base is used in the statistical analyses. A detailed explanation is provided in Appendix B.

whereas the overall level of inequality reflects a confluence of all of the factors in Equation (1). It is possible that other factors (say national inflation rate or national fiscal policy of redistribution) can cause an increase in the overall urban-rural inequality even when the (partial) effect of globalization is to reduce the inequality (i.e., $\beta < 0$).

In our statistical analysis, we implement a first-differenced version of Equation (1).

(2) $Y(k) = \alpha' + \beta X(k) + H(k, t)\Theta + v(k)$

where Y(k) denotes the change in the urban-rural inequality over the sample period,

(3) $Y(k) \equiv q(k, 1993) - q(k, 1988) = \log Q(k, 1993) - \log Q(k, 1988)$

X(k) denotes the change in the log trade openness for city k,

(4) $X(k) \equiv \log[OPEN(k, 1993)] - \log[OPEN(k, 1988)]$

and $\alpha' \equiv [N(1993)\text{-}N(1988)]\Lambda$, is a constant scalar. $v(k) \equiv e(k, 1993) - e(k, 1988)$, is a composite random variable, assumed to be i.i.d. normally distributed. $H(k) \equiv h(k, 1993) - h(k, 1988)$ is a vector of variables other than the change in openness that affect the change in inequality. The most noteworthy feature of the first-difference specification in Equation (2) is that all factors that are common across the cities, and all factors that are specific to a city but invariant over time are eliminated by the differencing process.

Equation (2) is our benchmark specification. Regression results based on this specification are reported in Table 3. In Column 1 of this table, where the change in inequality is regressed on the change in openness alone, the slope coefficient is negative and statistically significant. This means that cities that experienced a bigger increase in openness, tend to witness a decline in urban-rural inequality. Figure 2 presents a scatter plot of the change in inequality against the change in openness across the Chinese cities. A negative association between these two variables is apparent from the graph, and is unlikely to disappear if we remove one or two observations.

Ravallion (2000) suggested the possibility that inequality may regress towards the mean over time. In Column 2 of Table 3, we add the initial inequality in 1988 as another regressor. The coefficient on the initial inequality is negative, suggesting a possible tendency for reversion. However, this effect is not statistically significant. Note that after controlling for initial inequality, the coefficient on openness continues to be negative and statistically significant.

Perhaps economic growth may help to reduce the inequality because richer cities are more able to redistribute (implicitly as well as explicitly) to poor rural areas. In Column 3 of Table 3, the growth rate of city-level per capita GDP is included as another control. The coefficient on this new regressor is indeed negative, consistent with the "re-distributive ability" hypothesis, but it is only marginally significant (at the 15% level). One way to read this low significance level on the growth rate is that, holding openness constant, the per capita incomes in the rural and urban areas tend to grow in similar proportions[7].

One might be concerned that the results are driven by a few special coastal cities. Four cities were designated as "special economic zones." Fourteen coastal cities were designated as "coastal open cities." These cities were allowed to carry out certain types of market reform ahead of the rest of the country. While it is not clear how this would make inequality in these cities different from the other cities, it potentially can. In addition, if the investment in an urban area exceeds that in the adjacent rural areas (on a per capita basis), the gap between the urban and rural residents could widen. If the scale of the investment in the urban areas relative to the adjacent rural areas is correlated with that city's openness, omitting the relative investment measure could generate a spurious correlation between openness and inequality. In Column 4 of Table 3, we add two more control variables. The first is a dummy for cities that have been

---

[7] Using a cross-country regression framework, Dollar and Kraay (2001a) documented that economic growth in terms of per capita GDP is largely uncorrelated with the share of the poor in total income. In their interpretation, therefore, economic growth is good for the poor as their income rises in lock steps with the average income of the country. Of course, as Ravallion (2000) pointed out, no change in the distribution implies that the rich gets a bigger share of the total gain from the growth.

designated as "Coastal Open Cities."[8] The second is the change in log ratio of capital stock (fixed capital per person) in urban relative to rural areas[9]. In the regression, the dummy for "Coastal Open Cities" is not different from zero statistically. On the other hand, the ratio of investment in the urban relative to rural areas does matter as hypothesized: the coefficient is positive and statistically significant. Cities that have invested relatively more in the urban areas would see an increase in the urban-rural inequality. For our purpose, it is important to observe that after these controls, the coefficient on openness does not change and remains negative and statistically significant at the five percent level.

Another channel through which urban-rural inequality could decline is an improvement in the terms of trade for the products of the rural areas. The improvement in the terms of trade might be greater in more open areas. We compute a measure of terms of trade at the city level by the change in log price deflator of urban industrial production minus the change in log price deflator of the rural agriculture production. In Table 3, we add the change in the urban-rural terms of trade as another regressor. The new regressor has a positive sign, consistent with the hypothesis of an improving terms of trade for the rural areas, but is not statistically significant.

<u>Correcting for Possible Endogeneity of the Openness Measure</u>

A city's trade-to-GDP ratio may be endogenous. For example, the ratio may go up as a result of, rather than a cause for, a city's economic growth. Separately, the trade-to-GDP ratio may be mis-measured partly due to under-reporting of trade discussed earlier. To deal with this problem, we adopt a technique that was pioneered by Frankel and Romer (1999), and has subsequently been employed by Irwin and Tervio (2002) and Wei (2000). The basic idea in these papers is this: a country's volume of trade is related to its geography (e.g. proximity to other major trading nations in the world), but its geography is unlikely to be influenced by its income. In our case, we take advantage of the special geographic features of the Chinese territory to construct an instrumental variable for a city's openness, as noted in the introductory

---

[8] The four "special economic zones" are not part of the sample due to definitional change of urban versus rural areas in these cities.

[9] Data on fixed capital investment is not available for 1993. We use the 1994 data instead.

part of the paper. We observe that variation in regional openness reflects largely a different degree of access to major seaports. Furthermore, a small number of seaports carry out a large proportion of freight traffic. In fact, Hong Kong and Shanghai are by far the biggest ports for international trade in China. Together, about half of China's external trade pass through these two ports (and the adjacent smaller ports).

With these observations in mind, we use the distance from a city to either Shanghai or Hong Kong, whichever is smaller, as the instrumental variable (together with other regressors in the main regression as is standard with the two-stage least square approach) for openness for that city. By distance, we mean the great circle distance between a pair of cities, computed by the so-called "oblique spherical triangle method" based on the latitudes and longitudes of the cities[10]. The information on the latitudes and longitudes of the Chinese cities is retrieved from the Defense Mapping Agency (1990).

To be more precise, suppose d(k, Shanghai) [or d(k, HongKong)] is the greater circle distance between city k and Shanghai (or Hong Kong), then, the key instrumental variable for city k is

(5) $$D(k) = \min \{ \log[d(k, \text{Shanghai})] , \log[d(k, \text{HongKong})] \}$$

Note that we use geography as an instrumental variable for the <u>change</u> in openness whereas Frankel and Romer (1999) use it for the <u>level</u> of openness. We justify our approach by noting the peculiar aspect of China's recent economic history. Until the end of the 1970s, China -- which means all its cities -- was very closed to the world trading system. The "opening-up" reform started by Deng Xiaoping in 1979 allowed the various cities to participate in international trade to an extent and in ways that had not been possible before. However, the <u>effective</u> increase in openness varies widely across the cities. We hypothesize that an important part of this

---

[10] Oblique Spherical Triangle Method: Arc Distance D = Cos^{-1}(Sin(latitude1) x Sin(latitude2) + Cos(latitude1) x Cos(latitude2) x Cos(longitude1 – longitude2)). Sign convention: + (-) for north (south) latitude, and +(-) for west (east) longitude. Distance in kilometers = 111.12*D.

difference is explained by the difference in their ability to access the major seaports. We will describe later the extent to which this hypothesis is true (reported in Table 4).

We replicate the regressions in Table 3 with a two-stage-least square approach. Shanghai and Shenzhen (which is next door to Hong Kong) are dropped from the regressions as we want to avoid the problem of having to define the distance for any of these two cities to itself[11]. The results are reported in Table 4 (while the results from the first-stage regressions are reported in Table 5). Openness is negative and significant at the 5% level throughout the table. Figure 3 presents a conditional scatter plot of the change in the inequality against openness based on the IV regression in Column 4 of Table 4. As we can see, removal of any one or two observations on the chart is unlikely to change the negative slope between openness and inequality. If anything, the negative slope would be even steeper if one or two of the apparent outliers in the lower left corner of the graph were removed. We conclude, therefore, that greater openness helps to reduce the urban-rural inequality, and that this pattern is not a consequence of an endogenous trade-to-GDP ratio.

Note that the point estimates on openness are in fact bigger in the IV regressions than in the corresponding OLS regressions. While the reason for this is not immediately clear, we note that this pattern is similar to what Frankel and Romer found in their cross-country sample. If one takes the view that the slope estimates for openness in the OLS regressions in Table 3 are biased, then the IV estimates suggest that the impact of globalization on reducing income inequality is even bigger than what the simple OLS would suggest. According to the point estimate in Column 4 of Table 4, a ten percent increase in the trade-to-GDP ratio (e.g., from 0.2 to 0.22) leads to a three percent reduction in the gap between the urban-rural per capita incomes (e.g., from 5 to 4.85).

In the first stage of the two-stage-least-squared (2SLS) regressions just discussed, we have followed the usual practice in applied statistical work and employed all the regressors in the main regression other than openness as instrumental variables. However, one may be concerned that some of those regressors such as the growth rate of income are themselves susceptible to

[11] Shanghai and Shenzhen are also dropped in the sample in Table 2 so that the OLS and the 2SLS regressions have a comparable sample. Inclusion of these two cities in the OLS regressions makes no qualitative difference.

measurement errors. With this in mind, we also conducted 2SLS with D(k) as the only regressor in the first-stage regression. The results (not reported to save space) are broadly similar to those in Table 4.

So far the key IV variable for a city's openness is defined as the minimum distance from the city in question to either Hong Kong or Shanghai. We note that while Hong Kong and Shanghai are the top two ports for China's international trade, they certainly do not cover all the trade. In particular, these two ports are on the East and South sides of the country. For cities in the north, the distance from a major seaport in the north may be a more relevant determinant for their external trade. The biggest seaport in the northern segment of the coast is Qinhuangdao. As a robustness check, we also define our key IV for city k's openness as the follows:

(6) $$D^*(k) = \min \{ \log[d(k, \text{Qinhuangdao})] , \log[d(k, \text{Shanghai})] , \log[d(k, \text{HongKong})] \}$$

We have replicated the regressions in Tables 4 with this alternative instrumental variable, D*(k). The qualitative results are similar. In particular, the coefficient on openness is still negative and statistically significant, although the point estimates tend to be somewhat smaller (regression results not reported to save space).

Other Robustness Checks and Extensions

As an alternative measure of a city's trade openness, we also use the trade-to-GDP ratio, averaged over the sample period. We replicate the key regressions in Tables 3 and 4, and report the results in Table 6. As one can see, the coefficients on this alternative measure are also negative and statistically significant for all regressions.

We use minimum distance to the two major seaports as an instrumental variable for openness. Note here, geography is used to instrument the level of trade openness, which is similar to the Frankel and Romer's original application. Figure 4 plots the average openness against the minimum distance to Shanghai or Hong Kong. The negative association between the two shows up very strongly. Going back to the main regression (Columns 5-8 in Table 6), we observe that the negative effect of openness on the urban-rural inequality is intact. In fact, the point estimates in the 2SLS regressions tend to be bigger (in absolute values) than the corresponding OLS estimates.

As a further check, we have collected data from additional years between 1987 and 1993 in order to form a panel. The advantage of a panel regression is its ability to make use of all available information on trade and inequality. A disadvantage is that the Frankel-Romer technique for an IV regression cannot be applied here since geographic features of a city (e.g., distance from any given city to Shanghai or Hong Kong) do not change over time. In any case, several different panel regressions are conducted and summarized in Table 7.

In Column 1 of Table 7, where city fixed effects are controlled for, the coefficient on openness is –0.07, which is somewhat smaller than the corresponding regression reported in Table 3. But the estimate is still negative and statistically significant. In the second column of Table 7, where city random effects are controlled for, the point estimate on openness is affected only slightly, remaining negative and significant. In the next two columns, we add year dummies in addition to the city fixed (or random) effects. The openness variable remains negative and statistically significant. The point estimates in the regressions in Columns 3 and 4 become bigger than the corresponding ones without the year dummies. In the last four columns of Table 7, we include an additional regressor, the capital stock per capita in urban relative to rural areas. The coefficients on the openness variable are negative in all regressions. They are statistically significant at the ten and fifteen percent levels, respectively, in the two fixed-effects specifications, but insignificant in one of the two random-effects specifications. A formal Hausman test rejects the null hypothesis that the coefficients in the fixed-effects and random-effects specifications are the same at the one percent level. This suggests that the random effects are correlated with the other regressors, and the random-effects specifications have produced biased estimates. Hence, the estimates in the fixed-effects specifications are more reliable. To sum up, the panel regressions also reveal the same pattern that openness is negatively associated with urban-rural inequality.

Openness and Economic Growth

As noted earlier, some evidence was reported in Wei (1995) that more open cities in China tend to grow faster. However, no attempt was made in that earlier paper to separate urban and rural areas. Here, we examine possibly differential growth rates between the urban and rural areas and their connection with openness. A deficiency of the data on openness is that we only have information at the city level that combines urban areas and surrounding rural counties, which causes a measurement error in the openness variable. We implement a series of two-stage least square (2SLS) estimations where, in the first stage, we relate change in a region's openness to geography (the minimum distance from either Shanghai or Hong Kong). The regressions are analogous to those reported in Tables 4 and 5. We hope that the 2SLS procedure helps to mitigate the bias due to the measurement errors. In addition, the 2SLS estimation should help to correct for the possible endogeneity of openness.

The regression results are reported in the first two columns of Table 8. The rates of economic growth are positively associated with changes in openness for both urban areas and rural counties. The point estimate on the change in openness is almost three times larger in the rural regression than in the urban regression. This is consistent with the view that openness has done relatively more to promote economic growth in previously rural areas than in urban areas. However, the standard errors are relatively large so that neither estimate is statistically significant.

In Columns 3 and 4 of Table 8, we use the log average of the trade-to-GDP ratio over 1988-1993 as an alternative measure of openness. The point estimate on openness in the rural regression is more than twice as big as in the urban regression (and both are positive). Furthermore, the coefficient on openness in the rural regression is statistically significant at the five percent level.

To conclude, there is some (weak) evidence that openness has promoted economic growth in both urban and rural areas. Openness helps to reduce urban-rural income inequality by providing a bigger stimulus to growth in the previously poorer rural areas.

Transition from Agricultural to Industrial and Service Production

Why does openness promote faster growth in rural areas? Our hypothesis is that it offers a good opportunity for residents in rural areas to shift out of agricultural sector and into industrial or service sectors. Because of a peculiar Chinese policy, most of this sectoral shift does not necessarily involve a massive migration into the existing urban areas. Rather, it takes the form of a character change in the previous farm land. Due to the top leadership's concern about possible over-population in the cities, from the very start of the reform two decades ago, the government implemented the policy of "Li Tu Bu Li Xiang," or, the policy of "leaving-the-farm-work-but-not-the-farmland." What it means is that rural townships and villages are permitted to industrialize by converting some of the farmland to factories but farmers are discouraged from migrating to the cities. This policy is not executed perfectly as migration still occurs on a reduced scale, and has been relaxed a bit over time. However, as a result of the policy, there has been a spectacular rise in what are known as the "township-and-village enterprises," or the TVEs, which are industrial entities located in the rural areas. Our presumption is that the rise of income in the rural areas reflects, to a large extent, the growth of the TVEs, though there has also been an improvement in the productivity of agriculture and in the terms of trade for agricultural products[12]. Thus, globalization affects the Chinese rural areas by accelerating the growth of the industrial firms in addition to affecting the agriculture sector directly.

Now we turn to some direct evidence on this. Specifically, we examine the relationship between the growth rate of industrial output in former rural areas relative to their nearby urban areas and the region's openness. More precisely, let G(r, j) and G(u, j) denote the growth rates of industrial output in the former urban and rural areas in region j, respectively, from 1988-1993. We perform the following regressions:

$$(7) \qquad G(u, j) - G(r, j) = a + b \text{ Open } (j) + \text{control variables} + \text{error term}$$

The results are reported in Table 9. In Column 1 of Table 9, only a region's change in openness is used as the regressor. There is a clear negative association between the two variables: The relative growth rate of industrial output in urban versus rural areas tends to lower in regions with a faster increase in openness. In other words, rural industrialization (relative to nearby urban industrial growth) is faster in more open regions.

---

[12] By 1993, the last year in our sample, the TVEs accounted for approximately one-third of China's national exports (Yao and Zhu, 1998) and over one-quarter of China's gross industrial output (State Statistics Bureau, China Statistical Yearbook, 1994).

This basic relationship holds as we successively add other control variables: the initial gap in per capita industrial output between the urban and rural areas, the growth rate of the region's overall industrial output, a dummy for coastal open cities that may have received preferential economic policies from the central government, and the differential investment rate between the urban and rural areas.

In Table 10, we perform the same set of regressions using a 2SLS framework, where the change in openness is instrumented by a region's geographic distance to major seaports (and other control variables in the regression). The negative coefficient on change in openness remains statistically significant. If anything, the point estimates tend to be larger than the corresponding OLS estimates. In other words, a region's increase in openness is likely to have causally led to a faster rural industrialization (relative to its nearby urban industrial production).

As a robustness check, in Table 11, we report 2SLS regression results where a region's openness is measured by the average value of openness during 1988- 1993. The basic conclusion remains the same: rural industrialization is faster in areas with a faster increase in openness.

Industrialization is one way for residents in rural areas to shift out of agricultural production. The other possibility is for them to move into the service (or tertiary) sector. In Table 12, we report the 2SLS results of the growth rate of the tertiary sector in urban relative to rural areas as a function of a region's openness. The specification is parallel to that for the relative industrial growth. The regressions results consistently show a negative relationship. In other words, rural areas' service sector growth (relative to that of adjacent urban areas) is also faster in regions with a faster increase in openness.

Migration

We noted earlier that the Chinese government imposes various restrictions on internal migration (from rural to urban areas, or from one region to another). As a result, migration is smaller than it otherwise would be the case. Nonetheless, migration does take place. As cited in Knight and Song (1999, p268), the Chinese Rural Development Institute estimated that in 1993 approximately ten percent of rural labor forces engaged in short-term or long term migratory work. Migrant workers tend to be poorer than an average citizen in an urban area and at the same time raise the average income in rural areas by sending money homes. We note, however, that this is unlikely to be responsible for the findings reported earlier. This is because rural migrants to an urban area are often not properly recorded as a part of the local urban population.

On the other hand, because the output of the migrant workers is typically recorded as part of the urban output, a particular type of measurement error is introduced. Specifically, the per capita income in urban areas may be upward biased, while the per capita income in rural areas may be downward biased. The bias is probably bigger in more open regions where migration is more extensive. It is also probably more prevalent over time. As an implication, the true gap between urban and rural incomes is smaller than officially recorded. The decline in the income gap is probably faster than recorded, particularly in more open regions. If this is the case, the true effect of openness on the urban-rural inequality is bigger than our point estimates show.

4. Bringing in Inequalities Inside Urban and Rural Areas

We now turn to an analysis of intra-urban and intra-rural inequalities using the household survey data. To examine the effect of openness on inequality, we need to have information on regional openness. Such information unfortunately does not exist in the original surveys. Instead, we construct an openness proxy based on the information discussed in Section 3 that a region's openness is highly correlated with its geographic proximity to major seaports. Specifically, we first find out the geographic coordinates of a region, and then apply the regression coefficient in Column 1 of Table 5 to construct a measure of local openness. We will discuss the statistical implication of this procedure when we get to the empirical results.

Intra-urban inequality

Two common measures of income inequality are the Gini coefficient and Theil index[13]. In Column 1 of Table 13, the inequality is measured by Gini coefficient, and initial inequality is

---

[13] Let y_k be the income level for person k, μ be the average income, and N be the population size. The Gini coefficient is defined as $G = \frac{1}{2N^2\mu}\sum_k \sum_j | y_k - y_j |$. The Theil index is given by $T = \sum_k \frac{y_k}{\mu N} \log \frac{y_k}{\mu}$.

controlled for. The point estimate is positive (0.054) and statistically significant at the five percent level. This suggests that urban areas that experience a faster increase in openness also tend to exhibit a moderate <u>increase</u> in inequality. However, the key regressor -- the city's openness -- is a constructed regressor. Suppose we bootstrap the "true" standard error, reported in a square bracket following the conventional standard error, the coefficient is no longer significant at the ten percent level. The coefficient on initial inequality is negative and significant even with bootstrapped standard error. This suggests a mean reversion tendency: cities that were most unequal to start with tend to be less unequal over time and vice versa. To offer a visual impression, Figure 5 plots the partial correlation between the change in urban Gini coefficient and the change in openness conditional on initial inequality and growth rate. A positive association is apparent from the graph. In Column 2 of Table 13, we add the growth rate of per capital GDP to the specification, the coefficient on the key regressor – change in a city's openness – is still positive, statistically significant at the ten percent level if one uses the conventional standard error, but insignificant if one uses the bootstrapped standard error.

In Columns 3-4 of Table 13, we replace the dependent variable with changes in the Theil index as an alternative measure of inequality. The basic conclusion stays the same qualitatively. Specifically, the coefficients on change in openness are always positive, consistent with the notion that openness may has raised the intra-urban inequality over time. On the other hand, the estimates are not statistically significant once we replace the conventional standard errors with the bootstrapped ones in recognition of the fact that the key regressor is constructed by us.

As a robustness check, we also use a measure of openness, averaged over 1988-1993, as the key regressor, in replace of change in openness. We run regressions with four specifications as in Table 13. The results are reported in Table 14. In this case, the coefficients on average openness variable are positive in all four cases. Moreover, they are typically statistically different from zero as well even when one uses bootstrapped standard errors. This offers a stronger hint that intra-urban inequality tends to worsen as a result of openness.

<u>Intra-rural inequality</u>

The intra-rural inequality can be discussed in a similar fashion. Table 15 reports the main results from regressions of change in inequality on change in openness. Contrary to the case of intra-urban inequality, the coefficients on change in openness in all regressions are negative this time. To see the result visually, we plot the partial correlation between the change in openness and the change in the rural Gini coefficients in Figure 6, based on the regression in Column 2 of Table 15. A negative association between the two variables can be detected. This means that an increase in openness tends to be associated with a reduction in the intra-rural inequality. However, once we use the bootstrapped standard errors, the coefficients are no longer different from zero.

In Table 16, we use average openness as the key regressor (rather than change in openness). In this case, the coefficients are still negative and become significant at the ten percent level (or marginally significant at the fifteen percent level) even when we use bootstrapped standard errors. In other words, with this specification, the evidence is stronger that openness contributes to a reduction, rather than an increase, in rural inequality.

In the previous section, we have hypothesized that openness narrows the urban-rural inequality by offering a better opportunity for rural residents to shift out of agriculture and into industrial and service production. Here, with the household survey data, we can also provide some supplementary evidence on this point. In Table 17, we regress a rural region's increase in per capita income outside agricultural activities on that region's increase in openness or average openness. In all cases, we see a clear positive association between these two variables. When bootstrapped standard errors are used, only the average openness variable is statistically significant at the ten percent level.

Overall Inequality

At this point, we can bring together various pieces from the earlier parts to offer an assessment of the impact of openness on overall inequality. Intuitively, this impact depends on a combination of the effects of openness on urban-rural inequality, on within-urban inequality, and on within-rural inequality. To do this more formally, we adopt a decomposition of the Theil index proposed by Shorrocks (1980) and Mookherjee and Shorrocks (1982)[14]. Let I denote the

[14] Note that the Gini coefficient cannot be linearly decomposed without an inconvenient residual.

overall inequality for the whole country (and assume that there is one urban area and one rural area). Then, Mookherjee and Shorrocks (1982) show that the overall inequality, I, can be decomposed as

(8) $I = V_r\lambda_r I_r + V_u\lambda_u I_u + V_r\lambda_r \log\lambda_r + V_u\lambda_u \log\lambda_u$,

where V_r and V_u are the proportions of population living in rural and urban areas, respectively; λ_r and λ_u are the ratios of rural and urban average incomes to the overall national average income, respectively; and I_r and I_u are within-rural and within-urban Theil indexes, respectively.

Define $X \equiv \frac{y_u}{y_r}$ = urban-rural income inequality. Denote by n_r and n_u the populations in the rural and urban areas, respectively. Then, $\lambda_r \equiv \frac{y_r}{y} = \frac{y_r(n_r + n_u)}{y_r n_r + y_u n_u} = \frac{n_r + n_u}{n_r + n_u X}$. It is easy to verify that $\lambda_u = X\lambda_r$. Therefore, one can rewrite the country's overall inequality (measured by the Theil index, I) to be a function of three variables only (plus other parameters), namely, within-rural inequality, within-urban inequality, and the urban-rural inequality. Or $I = f(I_r, I_u, \lambda_r(X), \lambda_u(X)) = f(I_r, I_u, X)$. In this case, the impact of openness on the overall inequality can be written as a linear combination of its impact on each of the three variables.

(9) $$\frac{dI}{dOpen} = f_1\frac{\partial I_r}{\partial Open} + f_2\frac{\partial I_u}{\partial Open} + f_3\frac{\partial X}{\partial Open} = f_1\frac{\partial I_r}{\partial Open} + f_2\frac{\partial I_u}{\partial Open} + f_3 X\frac{\partial LnX}{\partial Open}$$

where $f_1 = V_r\lambda_r$, $f_2 = V_u\lambda_u$ and $f_3 = (V_r\log\lambda_r + V_r)\frac{-n_u(n_r + n_u)}{(n_r + n_u X)^2} + (V_u\log\lambda_u + V_u)[\lambda_r + X\frac{-n_u(n_r + n_u)}{(n_r + n_u X)^2}]$.

We take the relevant regression coefficients from the previous parts of the paper to quantify the three partial derivatives, $\frac{\partial I_r}{\partial Open}$, $\frac{\partial I_u}{\partial Open}$ and $\frac{\partial LnX}{\partial Open}$, and treat n_r, n_u, (or V_u, V_r), and the initial values of λ_u, λ_r, and X as pre-determined.

According to China's Statistical Bureau (1994, p.59), the urban and rural populations were 29 million and 82 million in 1988, respectively. Hence, we set

$V_r = \frac{n_r}{n} = 0.74$, $V_u = 1 - V_r = 0.26$

The initial average per capita income in the rural and urban areas were 760 and 1842 yuans, respectively, according to the data from the household surveys. Of course, there is uncertainty about which set of regression coefficients to use as the values for the three partial derivatives. Therefore, we experiment with several combinations of the regression estimates. The results of this exercise are summarized in Table 18. In Column 1 of Table 18, we pick regression coefficients from regressions that control for a region's initial level of inequality. Specifically, from Column 2 in Table 4, we pick $\frac{\partial LnX}{\partial Open}$=-0.323; from Table 13, Column 3 in Table 13, we pick $\frac{\partial I_u}{\partial Open}$=0.048; and from Column 3 in Table 15, we pick $\frac{\partial I_r}{\partial Open}$=-0.163. We plug these values into equation (9), the resulting value for the total derivative $\frac{dI}{dOpen}$ is -0.137. In other words, the effect of a rise in openness on the overall inequality is a reduction in inequality as measured by Theil index.

In Column 2 of Table 18, we pick a different set of values for the three derivatives, from the relevant regressions that control for a region's growth rate in addition to its initial level of inequality (hence, Column 3 in Table 4 for $\frac{\partial LnX}{\partial Open}$, Column 4 in Table 13 for $\frac{\partial I_u}{\partial Open}$, and Column 4 in Table 15 for $\frac{\partial I_r}{\partial Open}$). These values are summarized in Column 2, rows 3-5. With these values for the three partial derivatives, the total derivative, $\frac{dI}{dOpen}$=-0.167. Therefore, an increase in openness also leads to an improvement in the equality of the overall income distribution.

We note that in regressions involving household survey data, the coefficient estimates on change in openness in Tables 13 and 15 are not statistically significant once one utilizes the bootstrapped standard errors. As a third experiment, in Column 3 of Table 18, we set

$\frac{\partial I_u}{\partial Open} = \frac{\partial I_r}{\partial Open} = 0$. In this case, $\frac{dI}{dOpen}$ =-0.060. In other words, while the numerical value of the total derivative becomes smaller, the effect of an increase in openness is still a decline in inequality.

In earlier part of the paper, we have used a region's average openness as an alternative regressor to change in openness. Strictly speaking, the coefficient estimates on average openness would not have the interpretation as partial derivatives with respect to openness. But, just as a sensitivity check, we could make the assumption that they approximate the relevant partial derivatives. In Columns 4 and 5 of Table 18, we use the relevant estimates from Tables 6, 14 and 16 to impute the values for $\frac{\partial LnX}{\partial Open}$, $\frac{\partial I_u}{\partial Open}$ and $\frac{\partial I_r}{\partial Open}$. With these changes, the derivative of overall inequality with respect to openness is either -0.056 or -0.072, depending on whether one controls just for a region's initial inequality or also for a region's growth rate. The important message for us is that in both cases, openness is again negatively associated with inequality.

So far, we have used the average urban and rural income levels estimated from the household survey data. As another robustness check, we have also redone the experiment using the initial income level officially published by the Chinese government. According to China's State Statistics Bureau (2001), the average per capita rural and urban incomes in 1988 were 545 yuans and 1181 yuans, respectively. These numbers were considerably lower than the estimates derived from the household surveys, reflecting primarily the fact that certain incomes in kind and the rental value of houses were included in the household surveys but not by the Chinese State Statistics Bureau. We replicate the exercise in Table 18 with the new set of income numbers. The results are reported in Table 19. As it turns out, this change in initial incomes has virtually no impact on the estimated effect of openness on the overall inequality. That is, openness and total inequality are always negatively related. In fact, the numerical estimates are virtually the same in the two tables.

How important is the estimated effect of openness on inequality in economic terms? We note that the median estimate in both Table 18 and Table 19 is -0.07. This means that a ten percentage points increase in trade-to-GDP ratio (e.g., from 15% to 25%) is expected to lead to a reduction in Theil index by 0.007 (e.g., from 0.135 -- the actual level of inequality in 1988 based on the household survey -- to 0.128). So, while openness contributes to a reduction in overall inequality, the effect is modest quantitatively. [Of course, if we were to use the maximum estimate (-0.167 in Column 2 in either Table 18 or 19), the effect would naturally be somewhat bigger. The inequality as measured by the Theil index after a ten percentage points increase in openness would decline from 0.135 to 0.118.]

Finally, we note that we could also treat each province in the household surveys as if it were a separate country. A province naturally has both rural and urban areas. So a change in inequality at the province level automatically reflects a combination of changes of inequality between urban and rural areas, and within urban and rural areas. Such a calculation has several virtues. First, we need to work with only a single, consistent data set. Second, at the province level, we have direct observations on trade to local GDP ratio in 1988 and 1995 (from China's State Statistics Bureau, 1990 and 2001), saving the need to construct a measure of regional openness based on regional geographic features.

Unfortunately, there are only 10 provinces that were common to both 1988 and 1995 household surveys. Ten observations unlikely permit meaningful statistical tests. Nonetheless, we note that the change in inequality from 1988 to 1995 at the provincial level in our data set is negatively associated with a province's change of trade openness. Specifically, the correlation between change in a province's Gini coefficient and change in its log trade-to-GDP ratio is –0.35. The correlation between change in Theil index and change in log trade-to-GDP ratio is –0.55. These negative correlation coefficients are at least consistent with our earlier discussion that openness contributes to a reduction, rather than a rise, in overall inequality.

5. Concluding Remarks

The principal objective of the paper is to overcome the problems typically associated with cross-country regressions in studying the effect of openness on income inequality. One problem has to do with the fact that data on income and inequality across countries may not be comparable. The second problem has to do with the possibility that differences in legal systems and other institutions, culture and macroeconomic policies other than openness may affect cross-country differences in inequality. Fixed effects would not solve the problem if the effects of these factors on inequality interact with that of openness.

The principal innovation of the paper is to utilize cross-regional variation within a large developing country (China) in terms of difference in effective openness resulting from natural barriers to trade such as distance to major seaports. This intra-national study naturally minimizes the two problems associated with the cross-country studies. As an extra bonus, Chinese geography also offers a good opportunity to apply an instrumental variable approach to pin down the causal effect of openness on inequality.

Using two unique data sets on Chinese regions, the paper documents three basic patterns. First, an increase in openness leads to a reduction in the urban-rural income inequality. Second, an increase in openness is associated with a modest increase in within-urban inequality. Third, an increase in openness is associated with a decline in within-rural inequality. Putting together these three pieces of information, the paper shows that the overall effect of openness is to reduce inequality for a labor abundant economy.

This finding is in contrast with the impression one may obtain from the national aggregate figures: during the two decades of economic reform in China, both inequality and trade openness have risen for the country as a whole. For this reason, China is sometimes cited as proof that trade openness leads to a rise in inequality. The results presented in this paper suggest that such a conclusion is unwarranted.

The case study in this paper provides a useful complement to studies based on cross-country regressions. However, the Chinese experience does not necessarily imply that the effect of openness on income inequality should be the same in other countries. In particular, the effect of openness on inequality may be different for a labor-scarce economy. It can be very useful to undertake similar case studies for other countries in the future so that the role of globalization on inequality can be understood better.

References:

Atkinson, Anthony B., and Andrea Brandolini, 2001, "Promise and Pitfalls in the Use of 'Secondary' Data-Sets: Income Inequality in OECD Countries as a Case Study," Journal of Economic Literature, 39(3), p 771-799, September.

Defense Mapping Agency, 1990, Gazetteer of the People's Republic of China, Washington DC: Defense Mapping Agency.

Dollar, David, 1992, "Outward-oriented Developing Economies Really Do Grow More Rapidly: Evidence from 95 LDCs, 1976-85," Economic Development and Cultural Change, 523-544.

Dollar, David, and Aart Kraay, 2001a, "Growth is Good for the Poor," World Bank Policy Research Working Paper 2587.

Dollar, David, and Aart Kraay, 2001b, "Trade, Growth, and Poverty," World Bank Policy Research Working Paper 2615.

Economist, 2001, "Income Distribution In China," June 2, 39-44.

Edwards, Sebastian, 1992, "Trade Orientation, Distortions, and Growth in Developing Countries," Journal of Development Economics 39(1): 31-57.

Edwards, Sebastian, 1998, "Openness, Productivity and Growth: What Do We Really Know?" Economic Journal 108: 383-398, March.

Efron, Bradley, and Robert J. Tibshirani, 1993, "An Introduction to the Bootstrap," Monographs on Statistics and Applied Probability 57, New York: Chapman & Hall.

Forbes, Kristin, 2000, "A Reassessment of the Relationship Between Inequality and Growth," American Economic Review 90(4): 869-887, September.

Frankel, Jeffrey, A., and David Romer, 1999, "Does Trade Cause Growth?" American Economic Review, March, 379-399.

Hussain, Arthur, Peter Lanjouw, and Nicholas Stern, 1994, "Income Inequalities in China: Evidence from Household Survey Data," World Development, 22(12): 1947-57.

Irwin, Douglas A., and Marko Tervio, 2002, "Does Trade Raise Income? Evidence From The Twentieth Century," Journal of International Economics 58: 1-18.

Jin, Hehui, Yingyi Qian, and Barry R. Weingast, 2001, "Regional Decentralization and Fiscal Incentives: Federalism, Chinese Style," Unpublished working paper, Stanford University and University of California, Berkeley.

Kanbur, Ravi, and Xiaobo Zhang, 1999, "Which Regional Inequality? The Evolution of Rural–Urban and Inland-Coastal Inequality in China from 1983 to 1995," Journal of Comparative Economics, 27: 686-701.

Khan, Azizur Rahman, Keith Griffin, Carl Riskin, and Zhao Renwei, 1992, "Household Income and its Distribution in China," China Quarterly, December, 1029-61.

Khan, Azizur Rahman, Keith Griffin, Carl Riskin, and Zhao Renwei, 1993, "Sources of Income Inequality in Post-reform China," China Economic Review, 4(1): 19-35.

Khan, Azizur R. and Carl Riskin, 1998, "Income Inequality in China: Composition, Distribution and Growth of Household Income, 1988-1995," China Quarterly June, 221-253.

Khan, Azizur R. and Carl Riskin, 2001, Inequality and Poverty in China in the Age of Globalization, New York and Oxford, Oxford University Press.

Knight, John, and Lina Song, 1999, The Rural-Urban Divide: Economic Disparities and Interactions in China, New York: Oxford University Press.

Kwong, Tsz Man, 1994, "Markets and Urban-Rural Inequality in China," Social Science Quarterly, 75(4): 821-837.

Lardy, Nicolas R., 1992, Foreign Trade and Economic Reform in China, 1978-1990, Cambridge: Cambridge University Press.

Lin, Justin Yifu, 1992, "Rural Reforms and Agricultural Growth in China," American Economic Review, 82(1): 34-51, March.

Lindert, H. Peter, and Jeffrey G. Williamson, 2001, "Does Globalization Make the World More Unequal?", NBER working paper 8228.

Mookherjee, Dilip, and Anthony F. Shorrocks, 1982, "A Decomposition Analysis of the Trend in U.K. Income Inequality", Economic Journal, 92, 886-902.

Perkins, Dwight and Shahid Yusuf, 1984, "Rural Development in China," Baltimore: Johns Hopkins University Press.

Qian, Yingyi, and Gerard Roland, 1998, "Federalism and the Soft Budget Constraint," American Economic Review, 88(5): 1593-1615, December.

Qian, Yingyi, and Barry R. Weingast, 1996, "China's Transition to Markets: Market-Preserving Federalism, Chinese Style," Journal of Policy Reform, 1: 149-185.

Ravallion, Martin, 2000, "Growth, Inequality, and Poverty: Looking beyond Averages," unpublished World Bank working paper.

Rodriguez, Francisco and Dani Rodrik, 2000, "Trade Policy and Economic Growth: A Skeptic's Guide to Cross-National Evidence," NBER working paper No. 7081, April, in NBER Macroeconomics Annual 2000.

Sachs, Jeffrey D., and Warner, Andrew M., 1995, "Economic Reform and the Process of Global Integration," Brookings Papers on Economic Activity, 1-118.

Shorrocks, Anthony F., 1980, "The Class of Additively Decomposable Inequality Measures," Econometrica 48(3): 613-625, April.

Srinivasan, T.N., and Jagdish Bhagwati, 1999, "Outward-Orientation and Development: Are Revisionists Right?" Yale University, Economic Growth Center Discussion Paper No. 806.

Staiger, Douglas and Stock, James H., 1997, "Instrumental Variables Regression with Weak Instruments," Econometrica, 65(3): 557-86, May.

State Statistics Bureau, 1990, The Forty Years of Urban Development, Beijng: China Statistical Information and Consultancy Service Center.

State Statistics Bureau, 1993, China Statistical Yearbook, Beijng: China Statistics Publishing House.

State Statistics Bureau, 1994, China Statistical Yearbook, Beijng: China Statistics Publishing House.

State Statistics Bureau, 1999, The Fifty Years of the Cities in New China: 1949-1998 (Xin Zhongguo Chengshi Wushi Nian), Beijng: Xinhua Press.

State Statistics Bureau, 2001, China Statistical Yearbook, Beijng: China Statistics Publishing House.

State Statistics Bureau, 1988 to 1995, various volumes, Urban Statistical Yearbook of China, Beijing: China Statistics Publishing House.

World Bank, 1994, China: Foreign Trade Reform, Washington, D.C.: World Bank.

World Bank, 1997, Sharing Rising Incomes – Disparities In China, Washington D.C.

World Bank, 1998, World Development Indicators, Washington D.C.

Welch, Finis, 1999, "In Defense of Inequality," American Economic Review, Papers and Proceedings, 89(2): 1-17, May.

Wang, Tao, 1994, "Market and Reform: The Growth of Chinese Urban Industries," unpublished, a chapter in Ph.D Dissertation, New York University.

Wei, Shang-Jin, 1995, "The Open Door Policy and China's Rapid Growth: Evidence from City-level Data," in Takatoshi Ito and Anne O. Krueger, eds., Growth Theories in Light of East Asian Experience, Chicago: University of Chicago Press.

Wei, Shang-Jin, 2000, "Natural Openness and Good Government," NBER Working Paper No. 7765, June.

Yang, Dennis Tao, 1999, "Urban-Biased Policies and Rising Income Inequality in China," American Economic Review Papers and Proceedings, 89(2): 306-310, May.

Yao, Shujie, and Jirui Liu, 1998, "Economic Reforms and Regional Segmentation in Rural China," Regional Studies, 32(8): 735-746.

Yao, Shujie, and Liwei Zhu, 1998, "Understanding Income Inequality in China: A Multi-Angle Perspective," Economics of Planning 31: 133-50.

Zhao, Yaohui, 1999, "Labor Migration and Earnings Differences: the Case of Rural China," Economic Development and Cultural Change, 47(4): 767-782, July.

Appendix A: List of cities in the sample for urban-rural inequality, 1988-93

Anqing, Baiyin, Baoding, Baoji, Baotou, Beihai, Bengbu, Benxi, Changchun, Changzhi, Changzhou, Chengdu, Chifeng, Chongqing, Dalian, Deyang, Dongying, Foshan, Fushun, Fuxin, Fuzhou, Guangyuan, Guilin, Hangzhou, Hebi, Hefei, Hengyang, Hohhot, Huaibei, Huainan, Huaiyin, Huangshan, Huangshi, Huizhou, Huzhou, Jiaozuo, Jiaxing, Jilin, Jincheng, Jinhua, Jiujiang, Kaifeng, Leshan, Lianyungang, Liaoyang, Liaoyuan, Liupanshui, Liuzhou, Luohe, Luoyang, Maoming, Meizhou, Mianyang, Nanchang, Nanjing, Nantong, Ningbo, Panjin, Putian, Qingdao, Qinhuangdao, Qitaihe, Quzhou, Sanmenxia, Sanming, Shanwei, Shaoguan, Shaoxing, Shaoyang, Suzhou, Tangshan, Tianjin, Tianshui, Tongchuan, Tonghua, Tongling, Wuxi, Wuzhou, Xian, Xiangfan, Xianyang, Xinxiang, Xinyu, Xuchang, Yancheng, Yangquan, Yantai, Yinchuan, Yingkou, Yingtan, Yueyang, Zaozhuang, Zhangzhou, Zhanjiang, Zhaoqing, Zhengzhou, Zhenjiang, Zhoushan, Zhuzhou, Zigong

Appendix B: Data Cleaning

In three cases, observations are dropped to ensure that the quality of the data exceeds a minimum threshold.

a) Cites with at least one change in jurisdiction between 1990 and 1993 as listed in Fifty Years of the Cities in New China: 1949-1998. As an example, the number of rural counties under a city's administration might increase from one to four. Note that changes in jurisdiction prior to 1990 were not listed.

b) Cities whose urban or rural population had a change either by more than 40% or by more than 400,000 people in a single year. We suspect that these cities also experienced a change in their jurisdictions that was not properly documented in the published sources.

c) Cities with obvious errors or major abnormality in one of its key variables. For example, a city may have a virtually constant population in two adjacent years (e.g., 1987 and 1988), but its GDP in later year (1988) is only ¼ of the previous year (1987).

In some rare instances when there is an obvious way to fix a data error, we do that. For example, a city's recorded ratio of the urban-rural GDP per capita in 1993 may be only half of both its 1992 and 1994 levels. There are six such cities in our sample. In this case, we use the average of the 1992 and 1994 levels to replace the recorded 1993 values. However, replacement of the recorded values is done relatively rarely. We generally choose to err on the conservative side: when there is no obvious and non-arbitrary way to fix an error, we choose to drop the observation.

Appendix C: Urban and Rural Areas from the Household Surveys in 1988 and 1995:
(Names of the provinces in parentheses)

39 Urban Areas Beijing, Datong, Yangquan, Changzhi, Taiyuan (Shanxi), Dalian, Jinzhou, Shenyang (Liaoning), Xuzhou, Wuxi, Nanjing, Taixing, Changzhou, Suqian (Jiangsu), Tongcheng, Bengbu, Hefei, Huainan, Wuhu (Anhui), Kaifeng, Xinxiang, Huixian, Zhengzhou, Pingdingshan (Henan), Wuhan, Shashi (Hubei), Shantou, Zhanjiang, Foshan, Guangzhou, Huizhou (Guangdong), Kunming, Gejiu, Dongchuan, Baoshan, Dali (Yunnan), Lanzhou, Pingliang, Wuwei (Gansu)

40 Rural Counties Qing Xian, Wei Xian (Hebei), Qinshui Xian, Hunyuan Xian, Jiexiu Xian, Linyi Xian (Shanxi), Fengchengmanzuzizhi Xian (Liaoning), Yongji Xian (Jilin), Nantong Xian, Shuyang Xian (Jiangsu), Tong Xiang Xian (Zhejiang), Taihe Xian (Anhui), Nanchang Xian, Xingguo Xian, Yichun Xian, Wanzai Xian (Jiangxi), Licheng Xian, Shouguang Xian, Yangguo Xian, Taian (Shandong), Tangyin Xian, Fugou Xian (Henan), Jianli Xian, Zhijiang Xian (Hubei), Xiangtan Xian, Changning Xian (Hunan), Nianxiong Xian, Raoping Xian, Sanshui Xian, Huazhou Xian, Xingning Xian (Guangdong), Dafang Xian (Guizhou), Fumin Xian (Yunnan), Changan Xian, Fufeng Xian, Chengcheng Xian, Xunyang Xian (Shaanxi), Zhangye Xian, Dingxi Xian, Jingning Xian (Gansu)

Table 1: Summary Statistics of the Urban-Rural Inequality Sample

Variables	# Obs	Mean	Std. Dev.	Min	Max
Per Capita GDP in 1988					
Urban GDP p.c. (RMB yuan)	100	2806	1241	663	7626
Rural GDP p.c. (RMB yuan)	100	1058	516	320	3276
Urban/Rural income ratio	100	2.87	1.17	1.09	7.33
Per Capita GDP in 1993					
Urban GDP p.c. (RMB yuan)	100	6358	3252	1355	18547
Rural GDP p.c. (RMB yuan)	100	2483	1719	413	10473
Urban/Rural income ratio	100	2.96	1.34	0.95	8.12
Export-to-GDP Ratio (%)					
1988	100	7.76	4.75	0.90	27.69
1993	100	8.55	10.30	0.43	76.39
Annual Growth Rate of Per Capita GDP (%), 1988-93	100	7.95	3.74	0.65	19.75
Population in 1988					
Urban (millions)	100	0.83	0.76	0.13	5.62
Rural (millions)	100	2.69	1.99	0.28	11.67
Urban/Rural population ratio	100	0.51	0.57	0.05	2.90
Population in 1993					
Urban (millions)	100	0.90	0.80	0.14	5.90
Rural (millions)	100	2.85	2.09	0.30	11.97
Urban/Rural population ratio	100	0.53	0.58	0.05	2.95

Table 2: Summary Statistics of the Household Survey Data

	Obs	Mean	Std. Dev.	Min	Max
Rural counties					
Gini in 1988	40	0.24	0.07	0.13	0.50
Gini in 1995	40	0.29	0.08	0.19	0.48
Theil in 1988	40	0.088	0.043	0.028	0.211
Theil in 1995	40	0.178	0.120	0.058	0.492
GDP p.c. in 1988	40	809	356	380	2214
GDP p.c. in 1995	40	2381	1305	698	7270
Cities					
Gini in 1988	39	0.18	0.03	0.13	0.25
Gini in 1995	39	0.22	0.03	0.17	0.29
Theil in 1988	39	0.058	0.021	0.029	0.129
Theil in 1995	39	0.083	0.025	0.044	0.156
GDP p.c. in 1988	39	1760	449	1157	3052
GDP p.c. in 1995	39	5564	2513	2273	13304

Notes: County towns in the urban sample are dropped (3 of them in both years). GDP p.c. data are in current RMB.

Table 3: Openness and Urban-Rural Income Inequality
(OLS in First Difference with Robust Standard Errors)

Dependent variable: Change in log (urban GDP p.c./ rural GDP p.c.) from 1988 to 1993	(1)	(2)	(3)	(4)	(5)
Change in Log (exports / GDP) over 1988-93	-0.084**	-0.085**	-0.091**	-0.091**	-0.091**
	(0.036)	(0.037)	(0.037)	(0.036)	(0.036)
Initial Inequality in log		-0.030	-0.038	-0.044	-0.056
		(0.053)	(0.054)	(0.054)	(0.062)
Growth rate of GDP p.c.			-0.009#	-0.007	-0.007
			(0.006)	(0.005)	(0.005)
Dummy for costal open cities				0.015	0.016
				(0.071)	(0.071)
Change in log ratio of urban/rural p.c. fixed capital (88-94)				0.073**	0.076**
				(0.029)	(0.030)
Change in urban-rural terms of trade					0.064
					(0.161)
R-squared	0.06	0.06	0.08	0.12	0.12
No. of Obs.	100	100	100	95	95

Notes: Robust standard errors are in parenthesis. **, *, and # denote statistically significant at the 5%, 10%, 15% levels, respectively. An intercept is included in all the regressions but not reported to save space. Special economic zones are not in the sample.

Table 4: Instrumental Variable Regressions

Dependent variable: Change in log (urban GDP p.c./ rural GDP p.c.) from 1988 to 1993

	(1)	(2)	(3)	(4)	(5)
Methodology	IV	IV	IV	IV	IV
Change in Log (exports / GDP) over 1988-93	-0.305**	-0.323**	-0.274**	-0.316**	-0.311**
	(0.128)	(0.134)	(0.114)	(0.125)	(0.125)
Initial inequality in log		-0.051	-0.057	-0.057	-0.068
		(0.065)	(0.060)	(0.061)	(0.068)
Growth rate of GDP p.c.			-0.012*	-0.014*	-0.013*
			(0.007)	(0.008)	(0.008)
Dummy for costal open cities				0.003	0.005
				(0.081)	(0.081)
Change in log ratio of urban/rural p.c. fixed capital (88-94)				0.079*	0.081*
				(0.047)	(0.047)
Change in urban-rural terms of trade					0.061
					(0.165)
No. of Obs.	100	100	100	95	95
First-stage F on the instrument	11.6	10.9	13.5	11.6	11.6
p-value for Hausman test	0.03	0.02	0.04	0.02	0.02

Notes: The null hypothesis for the Hausman test is that the coefficients in the OLS and the IV regressions are not different systematically. First-stage F is the F-statistic for the null hypothesis that the coefficient on the instrument is zero.

Table 5: Openness and Proximity to Major Seaports: The First Stage in 2SLS

Dependent variable: change in log (export/GDP) 1988-93					
Corresponding to the regressions in Table 3	(1)	(2)	(3)	(4)	(5)
Minimum Log distance to Shanghai or Hong Kong	-0.237** (0.070)	-0.236** (0.071)	-0.264** (0.072)	-0.258** (0.076)	-0.260** (0.076)
Initial inequality in log		-0.008 (0.145)	-0.023 (0.144)	0.008 (0.143)	-0.014 (0.160)
Growth rate of GDP p.c.			-0.028* (0.015)	-0.041** (0.017)	-0.041** (0.017)
Dummy for costal open cities				-0.100 (0.189)	-0.097 (0.190)
Change in log ratio of urban/rural p.c. fixed capital (88-94)				-0.042 (0.111)	-0.037 (0.113)
Change in urban-rural terms of trade					0.122 (0.391)
R-squared	0.11	0.11	0.14	0.14	0.15
No. of Obs.	100	100	100	95	95

Table 6: Alternative Measure of Openness

Dependent variable: Change in log (urban GDP p.c./ rural GDP p.c.) from 1988 to 1993								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Methodology	OLS	OLS	OLS	OLS	IV	IV	IV	IV
Log (exports / GDP)	-0.102**	-0.102**	-0.098**	-0.098**	-0.127**	-0.130**	-0.125**	-0.148**
averaged over 1988-93	(0.034)	(0.034)	(0.034)	(0.035)	(0.044)	(0.043)	(0.045)	(0.049)
Initial inequality in log		-0.021	-0.024	-0.040		-0.020	-0.023	-0.035
		(0.051)	(0.052)	(0.057)		(0.051)	(0.051)	(0.056)
Growth rate of GDP p.c.			-0.004	-0.002			-0.003	-0.0002
			(0.006)	(0.006)			(0.006)	(0.006)
Dummy for costal open cities				0.075				0.102
				(0.067)				(0.072)
Change in log ratio of urban/rural				0.089**				0.096**
p.c. fixed capital (88-94)				(0.034)				(0.040)
Change in urban-rural				-0.044				-0.033
terms of trade				(0.139)				(0.137)
R-squared	0.12	0.12	0.13	0.16	.	.	.	.
No. of Obs.	100	100	100	95	100	100	100	95
First-stage F on instrument					68.4	71.6	65.4	53.3
p-value for Hausman test					0.46	0.39	0.43	0.19

See footnotes to Table 4.

Table 7: Panel Regressions: 1988-93

Dependent variable: log (urban GDP p.c. / rural GDP p.c.)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Methodology	Fixed effects	Random effects	Fixed effects	Random effects	Fixed effects	Random effects	Fixed effects	Random effects
Log (exports / GDP)	-0.069**	-0.064**	-0.072**	-0.064**	-0.049**	-0.034#	-0.040#	-0.026
	(0.015)	(0.015)	(0.016)	(0.015)	(0.023)	(0.022)	(0.024)	(0.022)
Log GDP p.c. in constant prices	-0.049*	-0.039#	-0.163**	-0.101**	0.137#	0.127**	0.164	0.134**
	(0.027)	(0.026)	(0.057)	(0.049)	(0.091)	(0.057)	(0.119)	(0.060)
Log urban/rural p.c. fixed capital					0.012	0.169**	0.023	0.174**
					(0.031)	(0.024)	(0.031)	(0.024)
Year dummy	No	No	Yes	Yes	No	No	Yes	Yes
R-squared (overall R-squared for random-effects)	0.04	0.003	0.06	0.002	0.02	0.45	0.05	0.46
No. of Obs.	663	663	663	663	397	397	397	397
p-value for Hausman test		0.46		0.81		0.00		0.00

Notes: The null hypothesis for the Hausman test is that the coefficients in the fixed-effects and random-effects specifications are not different systematically. A rejection of the null implies that the random effects are correlated with the other regressors, and hence the estimates from the random-effects specification are biased.

Table 8: Openness and Economic Growth in the Urban and Rural Areas

	(1) Urban	(2) Rural	(3) Urban	(4) Rural
Change in Log (exports / GDP) over 1988-93	0.685	1.081		
	(1.296)	(1.757)		
Log (exports / GDP) averaged over 1988-93			0.087	0.180**
			(0.069)	(0.085)
Log GDP p.c. in 1988	-0.199	-0.155	-0.114	-0.045
	(0.336)	(0.491)	(0.085)	(0.095)
Dummy for coastal open cities	0.169	0.071	0.013	-0.107#
	(0.307)	(0.282)	(0.071)	(0.070)
Change in log fixed capital Per capita (88-94)	0.143	0.323#	0.161**	0.256**
	(0.174)	(0.217)	(0.068)	(0.046)
Scientific personnel as a share of the population in 88	0.077	0.015	0.041	0.055#
	(0.116)	(0.125)	(0.035)	(0.036)
Remoteness	0.894	1.210	0.268**	0.187**
	(1.293)	(1.800)	(0.096)	(0.093)
Log population in 1988	-0.018	0.110	0.014	0.025
	(0.096)	(0.138)	(0.028)	(0.023)
No. of observations	95	95	95	95
First-stage F on instrument	0.4	0.4	31.5	16.2

Notes: The reported results come from 2SLS estimation where the first-stage regressions relate openness (either changes in log openness or the average of log openness) to the minimum distance from either Shanghai or Hong Kong. "Remoteness" for a given city measures the average distance from the city to all other cities in China, weighted by theses cities' population size in 1988. Special economic zones are not in the sample.

Table 9: Relative Growth of Urban/Rural Per Capita Secondary Sector Output: OLS

Dependent variable: change in log (urban/rural p.c. secondary sector output) from 1988 to 1993

Change in Log (exports / GDP) over 1988-93	-0.089* (0.051)	-0.095* (0.052)	-0.106** (0.053)	-0.108** (0.053)	-0.112** (0.052)
Log (urban/rural p.c. secondary output) in 88		-0.062 (0.056)	-0.089# (0.058)	-0.091# (0.058)	-0.100* (0.058)
Growth rate of p.c. secondary output			-0.225# (0.140)	-0.219# (0.140)	-0.285* (0.159)
Dummy for coastal open cities				-0.072 (0.097)	-0.053 (0.098)
Change in log ratio of urban/rural p.c. fixed capital (88-94)					0.114 (0.079)
R-squared	0.03	0.04	0.07	0.07	0.11
No. of Obs.	99	99	99	99	94

Table 10: Relative Growth of Urban/Rural Secondary Sector Output: IV Estimation

Dependent variable: change in log (urban/rural p.c. secondary sector output) from 1988 to 1993					
Change in Log (exports / GDP)	-0.351*	-0.442**	-0.341**	-0.334**	-0.382**
Over 1988-93	(0.191)	(0.211)	(0.169)	(0.168)	(0.183)
Log (urban/rural p.c. secondary		-0.088#	-0.115*	-0.116*	-0.125**
output) in 88		(0.060)	(0.059)	(0.059)	(0.060)
Growth rate of p.c. secondary			-0.293*	-0.281*	-0.403**
output			(0.156)	(0.155)	(0.180)
Dummy for coastal open cities				-0.093	-0.075
				(0.115)	(0.122)
Change in log ratio of urban/rural					0.124*
p.c. fixed capital (88-94)					(0.072)
No. of Obs.	99	99	99	99	94
First-stage F on instrument	11.4	10.6	14.2	14.1	12.7
p-value for Hausman test	0.11	0.04	0.11	0.11	0.08

Table 11: Relative Growth of Urban/Rural Secondary Sector Output: Average Openness, 2SLS

Dependent variable: change in log (urban/rural p.c. secondary sector output) from 1988 to 1993					
Log (exports / GDP)	-0.145**	-0.182**	-0.157**	-0.159**	-0.182**
averaged over 1988-93	(0.072)	(0.074)	(0.072)	(0.074)	(0.080)
Log (urban/rural p.c. secondary		-0.086*	-0.102*	-0.101*	-0.107**
output) in 88		(0.051)	(0.053)	(0.053)	(0.053)
Growth rate of p.c. secondary			-0.171	-0.174	-0.222#
output			(0.137)	(0.138)	(0.147)
Dummy for coastal open cities				0.033	0.063
				(0.115)	(0.120)
Change in log ratio of urban/rural					0.138**
p.c. fixed capital (88-94)					(0.067)
No. of Obs.	99	99	99	99	94
First-stage F on instrument	68.4	63.8	66.6	69.9	57.9
p-value for Hausman test	0.31	0.15	0.27	0.26	0.15

Table 12: Relative Growth of Urban/Rural Tertiary Sector Output: 2SLS

Dependent variable: change in log (urban/rural p.c. tertiary sector output) from 1988 to 1993						
Methodology	IV	IV	IV	IV	IV	IV
Change in Log (exports / GDP) over 1988-93	-0.517** (0.254)	-0.515** (0.241)	-0.433** (0.210)			
Log (exports / GDP) averaged over 1988-93				-0.214** (0.083)	-0.213** (0.079)	-0.209** (0.086)
Log (urban/rural p.c. tertiary output) in 88		-0.291** (0.084)	-0.301** (0.080)		-0.225** (0.069)	-0.249** (0.070)
Growth rate of p.c. tertiary output			-0.302* (0.176)			-0.105 (0.158)
Dummy for coastal open cities			0.047 (0.145)			0.164 (0.127)
No. of Obs.	99	99	99	99	99	99
p-value for Hausman test	0.02	0.01	0.04	0.58	0.23	0.29

Table 13: Inequality within Urban Areas

Dependent variable: inequality95 - inequality88

	(1)	(2)	(3)	(4)
Measure of inequality	Gini	Gini	Theil	Theil
Change in openness	0.054	0.045	0.048	0.041
	(0.023)**	(0.021)**	(0.024)*	(0.023)*
	[0.042]	[0.037]	[0.040]	[0.038]
Initial inequality (1988)	-0.985	-0.988	-0.987	-0.993
	(0.136)**	(0.140)**	(0.154)**	(0.160)**
	[0.136]**	[0.139]**	[0.154]**	[0.157]**
Growth from 1988-95		0.023		0.017
		(0.017)		(0.016)
		[0.018]		[0.017]
R-squared	0.62	0.64	0.53	0.54
No. of Obs.	39	39	39	39

Notes: Robust OLS standard errors in parenthesis. Bootstrapped standard errors in square brackets (based on 1000 replications). Each bootstrap replication (a) re-samples the 100 city observations used in Table 4 for which we have city-level trade data, (b) estimates the first-stage regression (openness on distance to major ports), (c) constructs change of openness for all regioins in the household suvey sample used in Table 13 using regression coefficients in Step (b), and then (d) re-samples from the 39 urban areas to estimate the second-stage regression.

Table 14: Inequalities within Urban Areas, Average Openness

Dependent variable: inequality95 - inequality88

	(1)	(2)	(3)	(4)
Measure of inequality	Gini	Gini	Theil	Theil
Average openness	0.022	0.019	0.020	0.017
	(0.010)**	(0.009)**	(0.010)*	(0.010)*
	[0.010]**	[0.010]*	[0.011]*	[0.011]#
Initial inequality	-0.985	-0.988	-0.987	-0.993
	(0.136)**	(0.140)**	(0.154)**	(0.160)**
	[0.136]**	[0.139]**	[0.154]**	[0.157]**
Growth from 1988-95		0.023		0.017
		(0.017)		(0.016)
		[0.018]		[0.017]
R-squared	0.62	0.64	0.53	0.54
No. of Obs.	39	39	39	39

See footnotes to Table 13.

Table 15: Inequalities within Rural Areas

Dependent variable: inequality95 - inequality88

	(1)	(2)	(3)	(4)
Measure of inequality	Gini	Gini	Theil	Theil
Change in openness	-0.134	-0.164	-0.163	-0.232
	(0.058)**	(0.081)**	(0.089)*	(0.123)*
	[0.113]	[0.142]	[0.163]	[0.217]
Initial inequality	-0.854	-0.915	-0.784	-0.978
	(0.202)**	(0.221)**	(0.502)#	(0.485)*
	[0.218]**	[0.241]**	[0.515]#	[0.504]*
Growth from 1988-93		0.048		0.108
		(0.061)		(0.081)
		[0.063]		[0.085]
R-squared	0.41	0.42	0.10	0.13
No. of Obs.	40	40	40	40

See footnotes to Table 13.

Table 16: Inequalities within Rural Areas, Average Openness

Dependent variable: inequality95 - inequality88

	(1)	(2)	(3)	(4)
Measure of inequality	Gini	Gini	Gini	Theil
Average openness	-0.056	-0.068	-0.068	-0.096
	(0.024)**	(0.033)**	(0.037)*	(0.051)*
	[0.028]*	[0.037]*	[0.043]#	[0.057]#
Initial inequality	-0.854	-0.915	-0.784	-0.978
	(0.202)**	(0.221)**	(0.502)#	(0.485)*
	[0.218]**	[0.241]**	[0.515]#	[0.504]*
Growth from 1988-93		0.048		0.108
		(0.061)		(0.081)
		[0.063]		[0.085]
R-squared	0.41	0.42	0.10	0.13
No. of Obs.	40	40	40	40

See footnotes to Table 13.

Table 17: Growth of Non-farm Income in Rural Areas

Dependent variable: ln(non-farm income 95) - ln(non-farm income 88)

	(1)	(2)	(3)	(4)
Change in openness	1.029	1.653		
	(0.433)**	(0.453)**		
	[0.905]	[1.232]		
Average openness			0.427	0.685
			(0.180)**	(0.188)**
			[0.206]**	[0.215]**
Log p.c. non-farm income in 88		-0.355		-0.355
		(0.130)**		(0.130)**
		[0.130]**		[0.130]**
R-squared	0.09	0.21	0.09	0.21
No. of Obs.	40	40	40	40

Note: The non-farm income also excludes dividents and interests and other property income, rental value of owned housing and net transfer from the government.

Table 18: Effect of Openness on Overall Inequality

	(1)	(2)	(3)	(4)	(5)
Assumptions					
Rural p.c. income in 1988	760	760	760	760	760
Urban p.c. income in 1988	1842	1842	1842	1842	1842
Intermediate Inputs					
dln(urban-rural gap)/dOpen	-0.323	-0.274	-0.274	-0.130	-0.125
d(urban Theil)/dOpen	0.048	0.041	0	0.020	0.017
d(rural Theil)/dOpen	-0.163	-0.232	0	-0.068	-0.096
Result					
d(overall Theil)/dOpen	**-0.137**	**-0.167**	**-0.060**	**-0.056**	**-0.072**

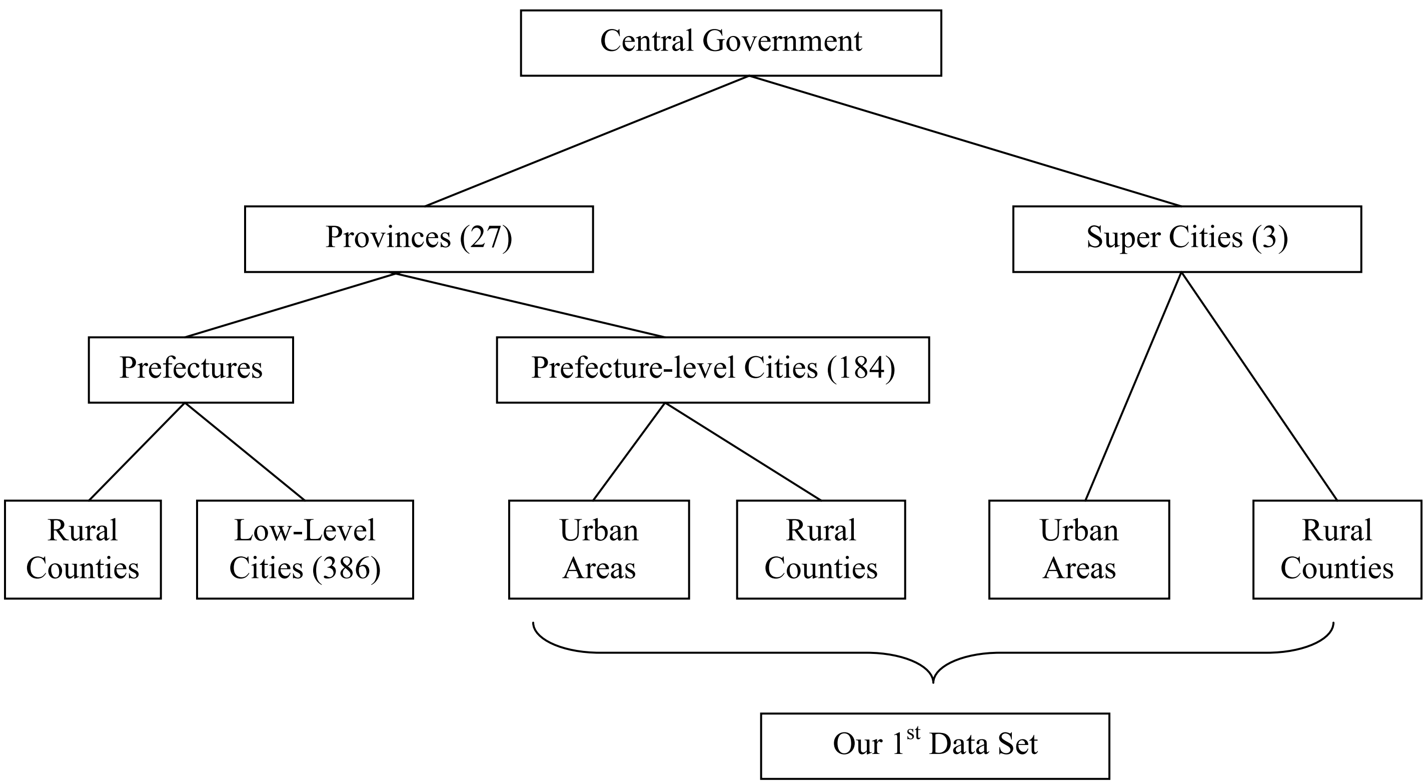
Note: The 1988 incomes are from the surveys.

Table 19: Overall Inequality, Alternative Assumption on Initial Income

	(1)	(2)	(3)	(4)	(5)
Assumptions					
Rural p.c. income (1988)	545	545	545	545	545
Urban p.c. income (1988)	1181	1181	1181	1181	1181
Intermediate Inputs					
dln(urban-rural gap)/dOpen	-0.323	-0.274	-0.274	-0.130	-0.125
d(urban Theil)/dOpen	0.048	0.041	0	0.020	0.017
d(rural Theil)/dOpen	-0.163	-0.232	0	-0.068	-0.096
Result					
d(overall Theil)/dOpen	**-0.134**	**-0.167**	**-0.052**	**-0.055**	**-0.071**

Note: The 1988 incomes are from the China Statistical Yearbook 2001 (p.312).

Figure 1: Administrative Structure in China (1993)



Note: In 1994, there were 783 counties that were administrated by cities, accounting for 45.1% of the total number of counties in China (1993 data not available).

Figure 2: Openness and Urban-Rural Income Disparity: Simple Correlation

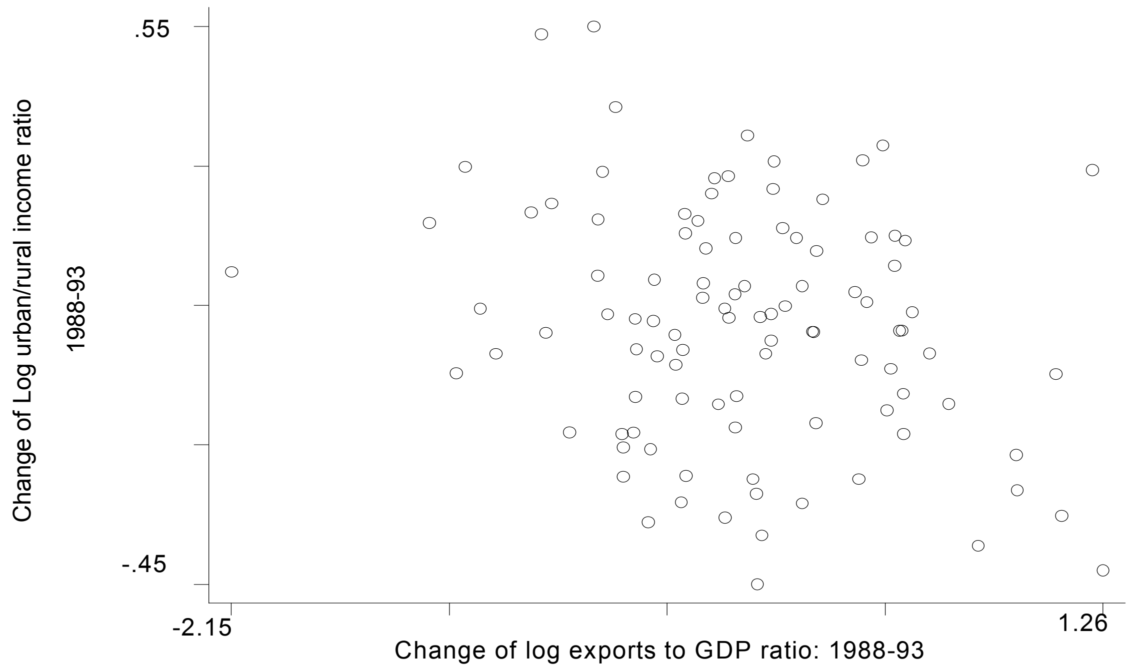


Figure 3: Openness and Urban/Rural Income Disparity:
Conditional Correlation from an IV Regression

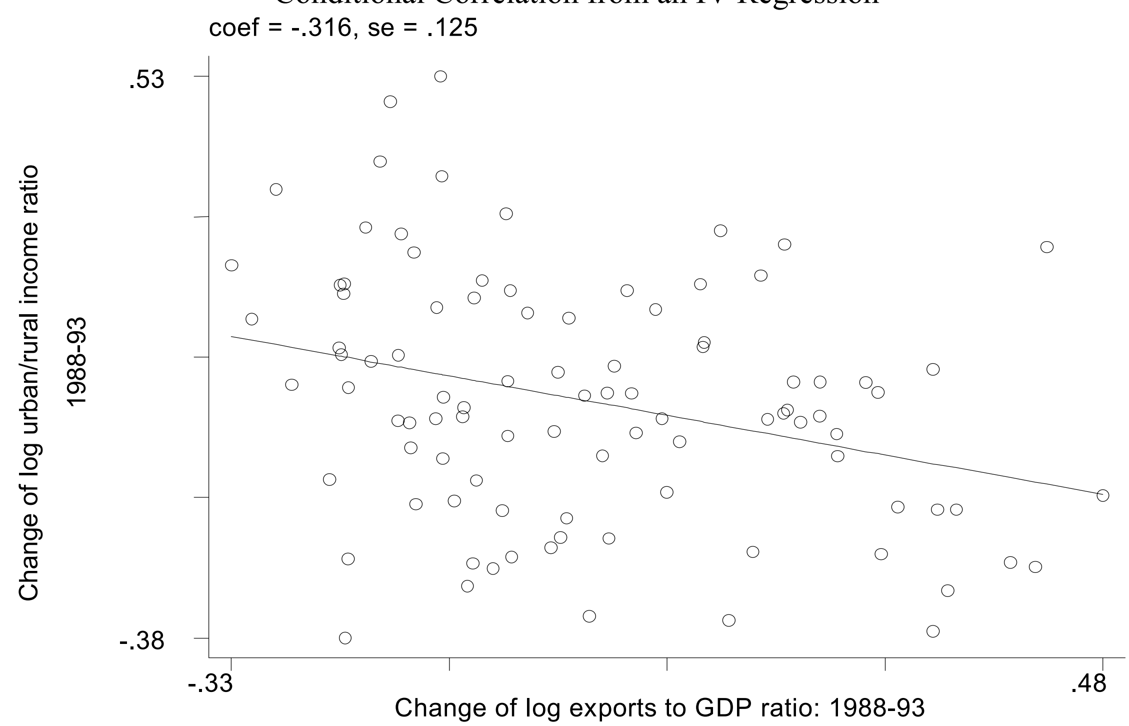


Figure 4: Openness and Minimum Distance from Shanghai or Hong Kong

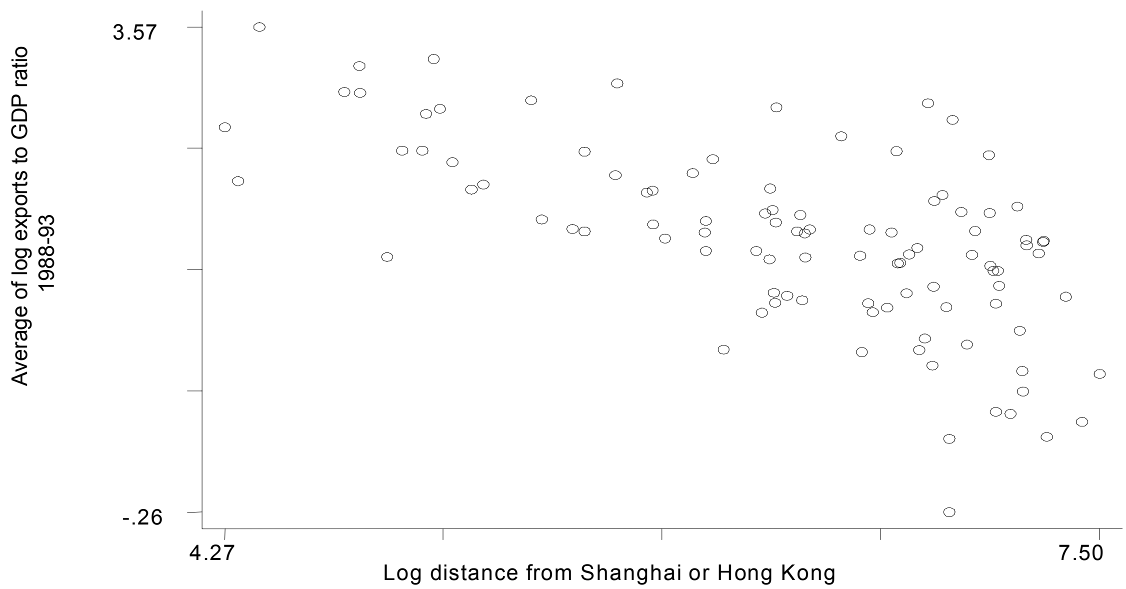


Figure 5: Openness and Within-Urban Inequality: Partial Correlation

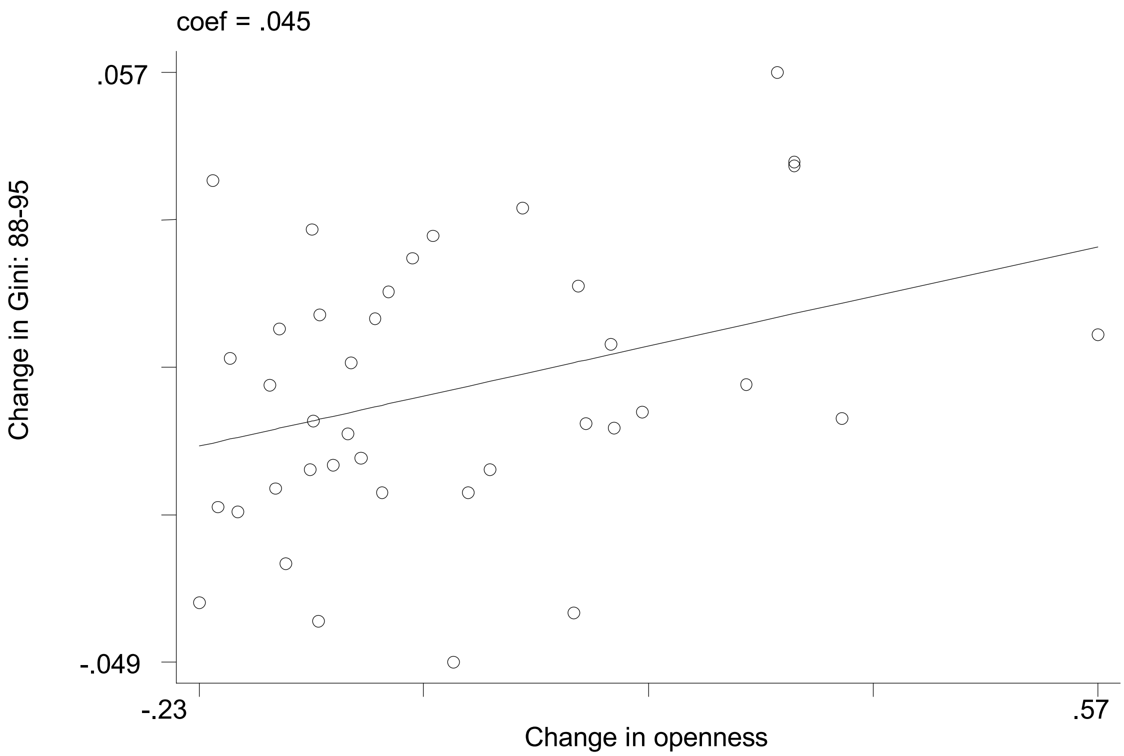


Figure 6: Openness and Within-Rural Inequality: Partial Correlation