Wage and Income Inequalities among Chinese Rural-Urban Migrants from 2002 to 2007

(Revised Version)

RESEARCH PROPOSAL

Presented to

PEP Network

By

Zhong Zhao (Renmin University of China and IZA)

Zhaopeng Qu (Nanjing University)

Juan Liao (Peking University)

Ke Zhang (Renmin University of China)

China

May 15, 2010
1. Abstract

Since 1988, rural-urban migration in China has become an important social and economic phenomenon. According to the National Bureau of Statistics of China, there are around 140 million rural migrants in China. Although the increasing of income inequality in China since the economic reform starting in 1979 is widely observed and extensively examined, to the best of our knowledge, very little study is devoted to study inequality among rural migrants, let alone to investigate the change of inequality across time among the migrants using comparable data sets. In another research project (Qu and Zhao, 2009), we have noticed that from 2002 to 2007, while the wage inequality of urban natives has kept increasing, the wage inequality among rural-urban migrants has decreased from 2002 to 2007. This project is to understand this phenomenon better. Using an existing data set and a unique data set newly acquired, we will carefully and comprehensively document the trends of income and wage inequalities of migrant from 2002 to 2007, analyze factors behind these trends, investigate the reasons why urban natives and rural-urban migrants follow different trends, and formulate relevant policy recommendations.
2. Main Research Questions and Core Research Objectives

Since economic reform of China, especially after 1988, rural-urban migration in China has become an important social and economic phenomenon. According to the National Bureau of Statistics of China, there are around 140 million rural migrants in China, which is probably the largest domestic migration in human history. Along with the rapid economic growth and after China joined the WTO in 2002, more and more rural population join the migration. This phenomenon attracts much attentions from academia, as well as central and local policy makers. Though most people believe rural-urban migration is one of the most significant social phenomena in contemporary China, due to the insufficient statistical methods and data limitation, the population of rural migrants is totally missing in most official statistics in China. To fill up this gap, the team of China Household Income Project (CHIP) has collected data on rural migrants in 2002, and another data designed to study rural migrants, Rural-Urban Migration in China (RUMiC), has been carried out since 2007. This two data sets afford us an unique opportunity to understand the rural migrants group better.

During the same period, the widening income and wage distributions in urban China is one of the most remarkable shifts in the structure of labour compensation in the Chinese labour market since the Chinese economic reform starting in 1979 based on both official publishing and academic researchers. Many economists have been
paying much attention to this phenomenon. They have investigated different aspects of this important phenomenon, such as rural-urban disparity, regional inequality, inequalities among urban residents or rural population. e.g., Zhao et al. (1994) and Zhao, Li, and Riskin (1999), Li, Sicular and Gustafsson (2008). However, due to data limitation as we mention above, there is little study devoted to study inequality among 140 million Chinese rural-urban migrants, let alone to investigate the change of inequality among migrants across time. As a consequence, the dynamics of migrant labour market and its impact of income distribution cannot be monitored and understood well by academia and policy makers.

To fill this gap, this project will devote to explore the wage and income inequality among migrants and its change from 2002 to 2007. In a separate project (Qu and Zhao, 2009) which is to examine the change of migration choice and wage equation from 2002 to 2007 among rural migrants, when we tabulate the descriptive statistics of the data sets, we noticed that from 2002 to 2007, while the wage inequality of urban natives has kept increasing, the wage inequality among rural-urban migrants has decreased during the same period. This is an interesting finding, but there are still so many issues which are not clear from simple descriptive statistics and worth digging more. We do believe a deeper understanding of this phenomenon could be important to both academia and policy makers. The main objective of this research is to understand this phenomenon better, and to study the

\[1\] For example, see the published papers in the symposium on Chinese inequality in the December 2006 issue of the *Journal of Comparative Economics.*
evolution of wage and income inequalities among migrants from 2002 to 2007. More specifically, we will

1) Documents the trends of income and wage inequalities among rural-urban migrants by gender, by skill groups. We will also compare these trends with the trends of urban natives;

2) Study the evolution of wage and income inequalities among migrants from 2002 to 2007. We will decompose the changes into price effects (coefficients effect) and endowment effects (composition effect) using the traditional Oaxaca-Blinder decomposition technique;

3) Extend the above decomposition analysis using the quantile decomposition technique developed by Machado and Mata (2005) and extended by Autor, Katz and Kearney (2006);

4) Investigate why the trends of inequalities are different between urban natives and migrants;

5) Formulate relevant policy recommendations based on our analysis.

3. Related Literature and Scientific Contribution of the Research

Since increasing of inequality and large scale of rural-urban migration are two of the most significant phenomena of Chinese society since the China economic reform, there are numerous studies on these two issues.
For inequality, for example, Gustafsson and Li (2002), Morduch and Sicular (2002) and Benjamin, Brandt and Giles (2005) investigate inequality in rural China, Knight and Song (2003) study the urban inequality from 1988 to 1995, Tsui (1993) and Kanbur and Zhang (2005) examine Chinese regional inequality, Knight and Song (1999) and Sicular, Yue, Gustafsson and Li (2007) investigate the urban-rural income gap, and Ravallion and Chen (2007) provide a comprehensive study on the inequality and poverty in China, among many others. More studies can be found, e.g., in the three volumes (Zhao et al. 1994 and Zhao, Li, and Riskin, 1999, Li, Sicular and Gustafsson, 2008) which are all using the China Household Income Project (CHIP) data sets.

Due to the data limitation, most studies overlook the inequality among rural to urban migrants. The only exception is Kahn and Riskin (2008); they use the supplement survey for migrants in 2002 CHIP to document the inequality among migrants in 2002. ² The evidence in the paper shows that the inequality among migrant households is larger than among rural households or urban households in 2002. However, their focus is nation-wide inequality instead of inequality among migrants, and they do not explore the possible reasons behind their observations. Furthermore, they only have one cross-sectional data in 2002, and it is impossible for them to study the change of inequality of rural-migrants over time.

On the other hand, there are also many studies on Chinese rural-urban migrants. Many researchers focus on the migration decisions, e.g. Hare (1999, 2002),

² See Li, S., T. Sicular, and B. Gustafsson (2008) and the data section of this proposal for details of this survey.
Zhao (1999, 2003) and Zhu (2002). Some others examine the impact of migration on
the source communities, such as Taylor, Rozelle and de Brauw (2003), Du, Park,
and Wang (2005) and de Brauw and Giles (2008). Many studies focuses the labor
market segregation between urban natives and migrants in wage, welfare and other
aspects, such as Meng and Zhang (2001), Knight and Yueh (2008), Demurge et al.
(2009), Deng and Li (2009). The general finding suggests that migrant workers work
longer hours and receive lower payment than urban natives under the same situation.

Nonetheless, as far as to the best of our knowledge, there is no study on the
evolution of the wage and income inequality among Chinese rural-urban migrants
across time. This project is seeking to fill up this gap, and will make following
contributions:

1) First, this project is the first attempt to document the changes of
inequality of migrants from 2002 to 2007. Given our initial findings that
trends of inequalities are different between urban natives and migrants,
carefully documenting these differences and checking if our initial findings
are robust across different gender, education, and age groups could be
important and interesting to both researchers and policy makers. This part
will go well beyond simple descriptive statistics we have at another project.

2) Second, the project will not only apply traditional Oaxaca-Blinder
decomposition technique to study the inequality, but also utilize the recent

---

3 Zhu (2002) is on the impact of urban-rural income gap on migration decision. His focus differs from
our proposal here.
developed quantile decomposition framework. This allows us to look into not only the decomposition at mean but also the decomposition over the whole distribution. This exercise will enable us to see if heterogeneity is important or not in the context of studying Chinese migrants inequality and whether the new sophisticated technique can provide us additional insight. If we find that heterogeneity is important, i.e. we have different findings at different quantiles, we should formulate the policy recommendations accordingly;

3) Third, comparing the evolution of wage and income inequality between urban natives and migrants, this project will provide additional insight for our understanding of the inequality among urban natives, too. Through this project, we also extend the analysis of urban income inequality based on CHIP data from 2002 to 2007, since the 2007 data we will use have similar structure as the CHIP data (more on this point later).

4) Forth, investigating the wage inequalities of migrants and urban natives also help us to have a better understanding of Chinese urban labor market. Arguably, migrants labor market is more competitive, and urban natives labor market is more institutionalized. Comparing wage inequalities of these two groups will enable us to learn the role of labor market institution.

This project is a perfect complement to our current research agenda. With the funding from the Renmin University of China, Zhong Zhao has a project studying the
labor market behavior of Chinese rural-urban migrants, which is focus on occupation choice and self-employment of migrants. Zhaopeng Qu has a project entitled “the Dynamic Evolution of Migrant Labor Market in China" funded by Nanjing University. In this project, working together with Zhong Zhao, Zhaopeng Qu is investigating the change of migration decision from 2002 and 2007. They also are looking into the wage equations (Mincerian equation) of migrants and urban natives from 2002 to 2007.

4. Policy Relevance

Inequality and rural-urban migration are two of the most important social issues in China. This project has clear policy implications:

1) The Chinese government is promoting “harmony society” now. Income inequality and earnings distributions are hotly discussed topics in Chinese media, policy circle and academia. However, without knowing the inequality of the 140 million migrants, the picture of Chinese inequality is incomplete; any policy discussion and formulation must take this group of people into consideration.

2) Chinese government is planning to establish an integrated rural and urban labor market. Our initial findings point out that the Chinese urban labor market not only segmented, but also probably has different wage-setting mechanism for urban natives and for migrants since we have found that the wage inequality of urban natives and rural-urban migrants have followed different trends from 2002 to 2007. This project will help us to
understand the difference between the migrants labor market and urban natives labor market. The results from this project will be very useful to formulate policy targeting at integrating these two labor markets.

5. Methodology

To analyze the evolution of inequality among migrants from 2002 to 2007, we follow the conceptual framework of Oaxaca (1973) and Blinder (1973). We decompose the change of inequality from 2002 to 2007 into price effect (i.e. the change caused by the change of coefficients) and endowment effect (i.e. the change caused by the change of the characteristics of the migrants). However, plus traditional Oaxaca-Blinder decomposition, we apply Machado and Mata Quantile decomposition technique. This technique was first proposed by Machado and Mata (2005) (hereafter “MM”) and then was extended by Autor, Katz and Kearney (2005). Specifically, we use this method to partition the change of distribution of wage/income into “price” components (wage coefficients) and “quantity” components (labor force composition) to interpret the change of overall and residual inequality from 2002 to 2007.

Although there are several alternative approaches based on entire distribution used in wage inequality studies, such as “JMP” proposed by Juhn, Murphy and Pierce (1993) and DFL by Dinardo, Fortin and Lemieux (1996), MM and its extended version has four unique advantages in the construction on the “counterfactual” distribution. First, the “predict” or “counterfactual” distribution based on MM method should be more “accurate” because it allow the estimating coefficients to
change over quantiles through applying a quantile regression while the coefficients in JMP method cannot be like that because it is estimated by a OLS regression. Second, under the convenient partial equilibrium assumption, MM method can be used to study the effect of changing both composition and prices on distribution of wages. Third, MM method can be easily extended to study the residual inequality and readily provides a “uniform and consistent” treatment of both overall inequality and residual inequality. Forth, the two approaches of JMP and DFL are naturally nested in the MM quantile model (Autor, Katz and Kearney, 2005). Therefore, our analysis will be mainly based on this approach.

Now let us turn to the technical part of extended MM method.\(^4\) Let \(Q_\theta(w|x)\) for \(\theta \in (0,1)\) denote the \(\theta\)th quantile of the distribution of the log wage (or income) given the vector of covariates \(x\). So the quantile regression equation is

\[
Q_\theta(w|x) = x'\beta(\theta),
\]

and the unconditional quantile distribution is

\[
\hat{w} = X\hat{\beta}'(\theta).
\]

Following Autor, Katz and Kearney (2005), we define the coefficient vector \(\hat{\beta}(50)\) which denotes the coefficients estimated on 50\(^{th}\) quantile as our measure of between-group inequality and we denote it as \(\hat{\beta}^b(\theta) \equiv \hat{\beta}(50)\). And we also define

\(^4\) More technical details of implementing procedure can be found in the technical appendix.
the measure of within-group inequality as the difference between the estimated coefficient vector $\hat{\beta}(\theta)$ and the median coefficient vector $\hat{\beta}(50)$:

$$\hat{\beta}^w(\theta) \equiv [\hat{\beta}(\theta) - \hat{\beta}^b(\theta)] \text{ for } \theta \in (0,1)$$ (3)

In summary, the distribution of wage/income can be seen as a function of three components: the distribution of covariates, $g(x)$, the vector of between-group prices, and the matrix of within-group (residual) prices, thus the distribution of wage/income, we denote it as

$$f_t(\omega_t) \equiv f(g_t(x), \hat{\beta}^b_t, \hat{\beta}^w_t)$$ (4)

Now we can use this unconditional distribution to construct the counterfactual facts and then do decomposition exercise that can isolate the endowment effect and price effect at quantiles.

Firstly, the overall change of wage/income in distribution at quantiles from 2002 to 2007 is

$$\Delta Q_\theta = Q_\theta(f_{2007}(w)) - Q_\theta(f_{2002}(w)) =$$

$$Q_\theta \left(f(g_{2007}(x), \hat{\beta}^b_{2007}, \hat{\beta}^w_{2007})\right) - Q_\theta \left(f(g_{2002}(x), \hat{\beta}^b_{2002}, \hat{\beta}^w_{2002})\right)$$

Secondly, the quantile wage/income change due to components or endowment from 2002 to 2007 is
\[ \Delta Q_\theta^x = Q_\theta \left( f \left( g_{2007}(x), \hat{\mu}^b_{2002}, \hat{\mu}^w_{2002} \right) \right) - Q_\theta \left( f \left( g_{2002}(x), \hat{\mu}^b_{2002}, \hat{\mu}^w_{2002} \right) \right) \]

Thirdly, the quantile wage/income change due to between group price or coefficients from 2002 to 2007 is

\[ \Delta Q_\theta^b = Q_\theta \left( f \left( g_{2007}(x), \hat{\mu}^b_{2007}, \hat{\mu}^w_{2002} \right) \right) - Q_\theta \left( f \left( g_{2002}(x), \hat{\mu}^b_{2002}, \hat{\mu}^w_{2002} \right) \right) \]

At last, the quantile wage/income change due to within group price or coefficients from 2002 to 2007 is

\[ \Delta Q_\theta^w = Q_\theta \left( f \left( g_{2007}(x), \hat{\mu}^b_{2007}, \hat{\mu}^w_{2007} \right) \right) - Q_\theta \left( f \left( g_{2002}(x), \hat{\mu}^b_{2002}, \hat{\mu}^w_{2002} \right) \right) \]

In this case, the decomposition sums to the total observed change

\[ \Delta Q_\theta = \Delta Q_\theta^x + \Delta Q_\theta^b + \Delta Q_\theta^w \]

6. Data

The data sets used in the project come from two different large household surveys. The first one is China Household Income Project (CHIP) which carried by the Institute of Economics, Chinese Academy of Social Sciences with assistances from the National Bureau of Statistics. CHIP has 1988, 1995 and 2002 three waves of data. The 2002 wave has a supplemental survey for rural migrants, plus standard rural and urban samples. The supplemental survey for migrants in 2002 which
covered 28 cities in 12 provinces. This is the first large scale survey targeted on rural-urban migrants in the cities in China.

The second one is the Rural-Urban Migration in China (RUMiC) research program which is conducted by an international team headed by researchers at Australian National University and with cooperation from Chinese Scholars. RUMiC is aimed to construct the first representative panel data sets target on rural-urban migrants in China. This survey covers 15 cities in 10 provinces in China from 2007 to 2012. Like 2002 CHIP, RUMiC also has rural, urban and rural-migrants there samples. We have already access to the 2007 wave of data and will be able to access the following waves as well.

Both surveys record detailed household information, such as income and expenditure, demographic characteristics, work and employment information. We plan to construct a unique and comparable repeated cross-sectional data sets using these two data sources.

Before starting to introduce our sample in more details, we should briefly discuss comparability between these two surveys. It is important to note that the sampling frames for migrants are different in these two surveys. In 2002, the sampling is based on residence of migrants. Migrants living in a dormitory or workplace such as construction site do not include in the sample. While in 2007, the sampling is based on workplaces. The 2007 sample includes migrants living in a dormitory or workplace, so we have to exclude these observations to make the 2007 sample comparable to the 2002 sample. We do believe that the samples in these two
surveys can match well based on our sample selection criterion since both samples are representative to their universe, i.e. the rural-urban migration population. Nonetheless, we will keep this crucial point in mind throughout the project, carry out sensitivity analysis for our data construction, and use other data sets, such as the mini-census of 2005, to check the comparability of our constructed samples.

We will further restrict the sample among people with age 16-65. On the wage inequality analysis we only keep the employed samples with a positive wage. We will also experiment to limit our observations from the same cities in these two surveys. There are seven common cities. We are sure that the sample size is still large enough after these sample restrictions.

As mentioned before, both data sets also have an urban natives sample. This allows us to compare the urban natives with migrants. For the urban sample, comparability is less a problem, since both CHIP and RUMiC have carried their urban survey with the help of NBS, and drawn their observations from the NBS sampling frame.

7. Consultation and Dissemination Strategy

Consultation and dissemination is an important component of this project.

From the beginning of the project, on the policy side, we will visit relevant government agencies, and seek their input. One the scientific side, we will send out our proposal, especially the methodology part to leading experts in this area for
comments and suggestion. This process will continue until we finish this project. In fact, the constant communication between policy makers and our team serve both consultation and dissemination purpose.

The results of this project will be disseminated through several channels to academics, policy-makers and the public.

First, we will present our research design and results at different universities. We will have seminars at Renmin University of China, Nanjing University, Peking University and Beijing Normal University, and get comments and suggestions from the faculty members at these places. All these universities have first-rate researchers in social sciences, especially economics in China.

Second, we will also submit our work to both domestic and international conferences and workshops, such as PEP meetings and workshops, IZA workshops, the Annual Congress of European Economic Association, Annual International Symposium on Contemporary Labor Economics at Xiamen University, and etc to collect more comments. This will allow our findings to reach wider audiences.

Third, we will circulate our research results as PEP and IZA working paper, and submit it for publication at an international academic journal, such as Journal of Comparative Economics, Journal of Development Economics or Review of Income and Wealth. And we will also disseminate our work in Chinese, and submit the work to a top Chinese academic journal, such as Economic Research Journal and China Economic Quarterly.
Forth, in order to disseminate our research more effective to the policy makers, we also suppose to write non-technical policy briefs based on our research, and send them to relevant government agencies though well-established channels between our universities and related government agencies.

Fifth, we will also plan to present our results to policy gatherings in Beijing such as Quarterly Employment Forum at Renmin University of China. This forum is a joint initiative between the School of Labor and Human Resources, Renmin University and different government agencies which include National Development and Reform Commission, Ministry of Human Resources and Social Security and NBS. With the possible funding from PEP, we are going to organize several small meeting gathering with related policy makers in Beijing.

Last but not least, we also plan to give public lectures in our universities and to arrange some media interview if necessary. Both of them will help us to disseminate our research outcomes to the general public quickly and widely.

8. List of team members

Last name: Zhao, First name: Zhong, Gender: Male, Age: 40

Zhong Zhao is a professor of economics at the School of Labor and Human Resources, Renmin University of China, and a research fellow of the Institute for the Study of Labor (IZA) at Bonn, Germany. He holds a bachelor's degree in economics from the Renmin University of China and a doctorate degree in economics from the
Johns Hopkins University. Before joined Renmin University of China, he had worked for IZA as a Senior Research Associate. He also held faculty position at Peking University, was a visiting assistant professor at the University of Southern California in the spring of 2004 and a visiting scholar at the Tinbergen Institute in October 2005. His main areas of interest are labor economics, applied micro-econometrics, social program evaluation and economy of China. His recent research topics on China include earnings instability and inequality, rural-urban disparity, comparison of wage evolutions in India and China, and rural-urban migration. He has published in both international and Chinese economics journals such as *Review of Economics and Statistics, Review of Income and Wealth, Economics Letters, Journal of Population Economics* and *China Economic Quarterly*.

Zhong Zhao’ past work experience also includes working as a government official at the Bureau of Finance in Yunnan Province for five years. This experience made his research on the China rural-urban migration and poverty has both academic and policy perspectives.

In summary, Zhong Zhao will be team leader. He will contribute to this project with research design, report writing, and mentoring three team members through this project.

**Last name: Qu, First name: Zhaopeng (Frank), Gender: Male, Age: 33.**

Dr. Zhaopeng Qu is currently an assistant professor at Department of Economics, Nanjing University. He obtained his Ph.D. degree in Economics from
Beijing Normal University in 2009. Zhaopeng Qu is competent at using modern econometric methods to analyzing labor economics issues. He has extensive experience with large microeconomic data. In particular, he is specializing in wage decomposition methods and statistical software such as Stata and Matlab.

Zhaopeng Qu will contribute to this project through literature review, model set-up, data analysis and report writing.

**Last name: Liao, First name: Juan, Gender: Female, Age: 30.**

Dr. Juan Liao is a Postdoc researcher at Graduate School of Education, Peking University. She holds a Ph.D. degree in Economics of Education from Beijing Normal University. Her research interests focuses on the effect of human capital such as schooling and training on Chinese labor market under the process of urbanization.

Juan Liao will contribute to this project with model set-up, policy analysis and report writing, especially the content related with education and training of migrants.

**Last name: Zhang, First name: Ke, Gender: Female, Age: 29**

Ms. Ke Zhang obtained her master degree in labor economics in 2004, and has worked for China Banking Regulatory Commission since then. Currently, she is taking a leave of absence from China Banking Regulatory Commission, and is pursuing her doctoral study in the School of Labor and Human Resources at Renmin University of China under the supervising of Zhong Zhao (team leader of this
application). Her research focuses on Chinese labor market. Her working experience in a government agent is a valuable asset for this project.

Ms. Zhang will contribute to this project with literature review, data cleaning, report writing and policy recommendation formulation.

Overall, Zhong Zhao will support the overall development of the research in his role as director of the project. He will work together with Zhaopeng Qu on building decomposition method. Zhaopeng Qu, Juan Liao and Ke Zhang will collaborate with several graduates in Peking University and Nanjing University to clean the data and construct necessary data set, and then carry out statistical and econometrical analysis. At last, four members will write the final report together.

Besides the above mentioned team members, there is also one female master student at Renmin University, Yanjiao Song, who will also work as research assistant for this project.

9. **Expected capacity building**

Expected capacity building from this project has following dimensions:

First, this project will enhance the team members’ econometric skills and knowledge on migration, income inequality and poverty. Due to the lack of data, methodologies and funding support, many important research questions on the rural-urban migration, such as wage inequality of migrants in urban labor market, are still under explored. Our research will enable us to get better knowledge on the
functioning of the migrant labor market, the change of wage structure of migrants, as well as on the urban residents. Obviously, this entails a good deal of knowledge deepening in the theoretical and applied literature on these topics. This project will also help us to develop further econometric and programming skills. Although we gained some experience through our previous research and other projects, having the possibility of developing the research agenda proposed here with the help of the PEP Network, would allow us to deepen our expertise and develop new analytic tools. With the funding from PEP Network, it is possible for us to reach out for outside resource person, such as through a short-term visit. It would also facilitate the study of a topic of great importance to our country.

Second, this project will help us to form a stable and long-lasting research team on China labor market. While our team members come from three different universities in China, we do have very closed academic connection. Zhong Zhao had been the mentor and cooperator to Zhaopeng Qu during Zhaopeng Qu’s one-year research visit at IZA. Juan Liao and Zhaopeng Qu cooperated several times when they pursued their Ph.D. degree in the same university. Ke Zhang is a doctoral student under the supervision of Zhong Zhao. Although we have gained a lot of cooperation experiences through previous projects, with the help of PEP network and through this project, we expect to strengthen our ties and to form a stable and long-lasting research team on China labor market.

Third, this project will help us to attract and to train future Chinese labor economists. Because all of us are working in the top universities in China, identifying
and training potential young researchers are one of our top priorities. For this reason we intend that younger members of our research team get to work together with more senior members. The same applies to research assistants that we will recruit from our best graduates. We do believe that the support of the PEP Network will not only help us to strengthen our research capacity, but also enable us attract and train future Chinese labor economists.

Forth, this project will help us to build the skill to bridge the academics and policy makers. Through this project we intend to bring the rigorous research into policy debate, at the same time, in order to have a policy impact, it is necessary to translate academic research into relevant policy recommendations.

Fifth, this project is taking a novel way to build the capacity of female researchers, i.e. this project will involves female researchers at different stage of their, from master student, Ph.D. student to post doc. The older female researchers in the team will be role model and inspiration for the younger female researchers.

10. Any ethical, social, gender or environmental issues or risks that should be noted.

At this stage of the proposal, and considering the available information, we do not foresee any ethical, social, gender or environmental issues or risks that should be noted or mentioned.
11. List of past, current or pending projects in related areas involving team members

Zhong Zhao


Zhaopeng Qu


Juan Liao

References


Appendix: Technical Appendix for Quantile Decomposition

This appendix is largely based on Machado and Mata (2005) and Autor, Katz, and Kearney (2005).

Let \( Q_\theta(\ln y | x) \) for \( \theta \in (0,1) \) denote the \( \theta \)th quantile of the distribution of the log consumption given the vector \( x \) of covariates. So the quantile regression equation is

\[
Q_\theta(\ln y | x) = x' \beta(\theta),
\]

(A1)

where \( x \) is a \( k \times 1 \) vector of covariates and \( \beta(\theta) \) is a vector of coefficients. We can estimate \( \beta(\theta) \) by the following, as shown in the seminar paper of Koenker and Bassett (1978):

\[
\hat{\beta}_\theta = \arg\min_{\beta} \frac{1}{n} \left[ \sum_{i|\ln y_i \geq x_i \beta} \theta |\ln y_i - x_i \beta| + \sum_{i|\ln y_i < x_i \beta} (1-\theta) |\ln y_i - x_i \beta| \right].
\]

(A2)

The main idea of Machado and Mata (2005) is the probability integral transformation, i.e., if \( U \) is uniformly distributed over the interval \( (0, 1) \), then \( V = F_\beta^{-1}(U) \) has cumulative distribution function \( F_\beta(\cdot) \). So for a given \( X_i \), if \( \theta \) has a uniform distribution, then \( x_i \beta(\theta) \) has the same distribution as the conditional distribution \( y_i | X_i \), and \( x \beta(\theta) \) has the same distribution as the unconditional distribution \( y \) if \( x \) is not given. Based on this idea, we can simulate the counterfactual distributions in the paper as follows:
(1) Draw $\theta_1, \theta_2, \ldots, \theta_i$ from a uniform $(0, 1)$ distribution.

(2) For the sample in the $t$ and $\tau$ time, run quantile regressions at each $\theta_i$, $i = 1, \ldots, l$, and get $l$ estimates of the quantile regression coefficients $\hat{\beta}_i(\theta_i)$, where $t$ and $\tau$ represents the sample in 2002 and 2007 year respectively.

(3) Use equation (4) to calculate the residual price vector $\hat{\beta}^w_t(\theta_i)$ and $\hat{\beta}^w_\tau(\theta_i)$ which denotes the coefficients within-group inequality in 2002 and 2007, and $\hat{\beta}^b_t(\theta_i)$ and $\hat{\beta}^b_\tau(\theta_i)$ which denotes the coefficients of between-group inequality in 2002 and 2007 respectively.

(4) Draw a random sample of size $l$ with replacement from the sample in $t$ and $\tau$ time denoted by $\{x^*_i(t), i = 1, \ldots, l\}$ and $\{x^*_i(\tau), i = 1, \ldots, l\}$.

(5) Combining steps (3) and (4), we can get the unconditional and simulated wage distribution $\{w^*_i(t, \tau)\} = \{x^*_i \cdot \hat{\beta}^w_t(\theta_i) \cdot \hat{\beta}^b_\tau(\theta_i)\}$, which includes all possible wage distributions we want.