Industry specialization, diversification, churning, and unemployment in Chinese cities

Fu, Shihe and Dong, Xiaofang and Chai, Guojun

Southwestern University of Finance and Economics

May 2008

Industry Specialization, Diversification, Churning, and Unemployment in Chinese Cities

Shihe FU*
Xiaofang DONG
Guojun CHAI

Research Institute of Economics and Management
Southwestern University of Finance and Economics
P.R.China

May 16, 2009

Abstract

This paper studies how industry specialization, diversification, and churning affect unemployment rates in Chinese cities. Using a city level panel data set from 1997 to 2006, we find that the specialization of wholesale and retail industry can significantly decrease unemployment rate; however, specializing in finance industry increases unemployment rate. In contrast to the evidence from developed countries, industry diversity is positively and significantly associated with unemployment rates in Chinese cities, possibly due to the higher degree of industry churning during the sample period. We also find that urban economic growth, market maturity measured by the proportion of private sector employment, and human capital can decrease unemployment rate. Industry diversity does not stabilize unemployment; wholesale and retail industry increases unemployment fluctuations; but market maturity and human capital stabilize unemployment.

Keywords: Industry structure; specialization; industry diversity; unemployment; churning

JEL Code: J64; R11; R23

---

*Shihe FU, Corresponding author. Email: fush@swufe.edu.cn; mailing address: Research Institute of Economics and Management, Southwestern University of Finance and Economics, 55 Guanghua Cun Street, Chengdu, 610074, China. Xiaofang DONG and Guojun CHAI's emails are xfangdong@gmail.com and 206020104002@mail.swufe.edu.cn, respectively. The authors thank Li Gan, Yao Meng, the participants of the Eighth China Economics Annual Conference, the Fifth Biennial Conference of the Hong Kong Economists Association, and the seminar at Southwestern University of Finance and Economics for very helpful comments. This work was supported by the Scientific Research Fund of Southwestern University of Finance and Economics.

1. Introduction

Unemployment can basically be classified into three types: structural unemployment, frictional unemployment, and cyclical and seasonal unemployment.[1] Urban industry structure, characterized by the degrees of both industry specialization and industry diversity, and its evolution, are closely related to these three types of unemployment in cities. This paper aims to test how urban industry specialization, industry diversity, and industry fluctuations affect the levels and fluctuations of unemployment rates in Chinese cities.

The degree of specialization of an industry in a city refers to the proportion of total employment or value added of that industry in a city, or the relative proportion standardized by the national average (usually measured by a location quotient index). Industry specialization affects structural and cyclical unemployment. Some industries are considered "unstable" because their employment fluctuates easily, such as mining, durable goods manufacturing; some other industries are considered "stable" in terms of employment, such as education, public utility, with less cyclical fluctuations (Thompson, 1965). In addition, specialized industries have comparative advantages in production and management, leading to stable employment or even stable employment growth in the short run. Malizia and Ke (1993) use the USA quarterly data from 1972-1988 and find that higher proportion of mining and transportation industries are associated with higher degree of unemployment fluctuations in metropolitan areas. Mizuno, Mizutani, and Nakayama (2006) use 118 Japanese urban areas data in 1995 and find that specialization in construction and manufacturing industries significantly reduces unemployment rates, but specialization in transport and telecommunication industry increases unemployment rates.

Industry diversity refers to the degree of unevenness in terms of the distribution of employment or value added across different industries in a city, usually measured by the Herfindahl index. Industry diversity helps reduce frictional unemployment. In a city with a single industry, if random unemployment occurs, the unemployment in that city will increase; however, in a city with many industries, a random unemployment in one industry might be quickly offset by a random hiring in another industry. Since intra-city migration is less costly than inter-city migration, a high degree of industry diversity in a city can reduce frictional unemployment rate. Based on this theory, Simon (1988) uses the USA metropolitan area industry data from 1977-1981 and finds that unemployment rates are lower in metropolitan areas with a higher degree of industry diversity (measured by the Herfindahl index). Malizia and Ke (1993) reach the same conclusion. Izraeli and Murphy

[1] Many macroeconomics textbooks provide short definitions of these types of unemployment. Briefly speaking, structural unemployment is due to the mismatch of labor supply and labor demand, usually exacerbated by economic reform and the change of industry structure. Frictional unemployment results from the search cost and imperfect information in labor markets and it exists even in the full-employment macroeconomic scenario. Frictional unemployment fluctuates randomly. Cyclical or seasonal unemployment is due to the macroeconomic business cycle or seasonal factors.

(2003) use the USA state level data and report that industry diversity is negatively correlated with unemployment rate. Tarzwell (1999) uses the Canadian province level data from 1989-1994 and also find that industry diversity is negatively correlated with unemployment rate.

Urban industry diversity can also be positively correlated with unemployment rate. Barkume (1981) argues that in a city with a high degree of industry diversity, the degree of wage dispersion will be higher too, and workers will be motivated to spend more time searching for jobs. Therefore, the unemployment duration will be longer in more industrially diversified cities. Simon (1987) provides some evidence to support the wage dispersion argument, but he argues that in general, the negative effect of industry diversity on unemployment rate is stronger than the positive effect on unemployment rate. Chiang (2008) uses a city level panel data in Taiwan and finds that industry diversity increases unemployment rate during 1980-2000.

A few other studies find no significant correlation between industry diversity and unemployment rate. For example, Attaran (1986) uses annual data during1972-1981 and finds that the negative relation between industry diversity and unemployment rate is very weak. Mizuno, Mizutani, and Nakayama (2006) show that although in Japan the correlation between urban industry diversity and urban unemployment rate is negative, it is not statistically significant.

Industry diversity is also correlated with unemployment instability (or employment stability). Barth, Kraft, and Wiest (1975) apply the asset portfolio theory to explain how regional industry diversity is negatively associated with unemployment instability: if employment in different industries is correlated in different ways, an optimal industry portfolio can be selected to minimize the fluctuation of the regional total employment, just like minimizing the risk of the aggregate returns by forming an optimal asset portfolio. Malizia and Ke (1993) conclude that industry diversity is negatively correlated with unemployment instability in the USA metropolitan areas. However, Attaran (1986) find that industry diversity is not correlated with unemployment instability.

Urban industry structure changes along time. During any period of time in a city, some industries might be growing, some other industries might be decaying, and new industries might emerge. The fluctuation or turnover of industries in a city is dubbed industry churning. Industry churning results in redistribution of employment across different industries and locations; therefore, industry churning affects unemployment rate and its instability. However, few existing studies have addressed the issue of churning and unemployment.

In summary, although theories provide some deterministic predictions about how industry specialization, diversity, and churning are closely correlated with unemployment rate and unemployment instability, empirical evidence across countries has been mixed. Table 1 further summarizes some representative empirical studies that reach different conclusions.[2]

Table 1
Summary of representative studies on industry structure and urban unemployment

Paper	Industry diversity	Industry specialization	Methodology	Country
Attaran (1986)	insignificant	Not included	Cross section	USA
Simon (1988)	Negative	Not included	Cross section	USA
Malizia and Ke (1993)	Negative	Included	Cross section	USA
Tarzwell (1999)	Negative	Not included	Cross-section	Canada
Izraeli and Murphy (2003)	Negative	Not included	Panel data	USA
Mizuno, Mizutani, and Nakayama (2006)	insignificant	Included	Cross section	Japan
Chiang (2008)	Positive	Not included	Panel data	Taiwan

As indicated in Table 1, most of the existing studies use cross-sectional data. This raises an endogeneity issue since firms and workers may select into more specialized or more diverse locations. Using a panel data approach can partially solve the sorting problem. Some studies use a panel data approach to identify how diversity affects unemployment rate, but fail to control for multiple specializations, creating omitted variable bias. Our paper aims to study how urban industry specialization, diversity, and churning affect unemployment rates in Chinese cities. We employ a panel data approach, controlling for both unobserved city attributes and time-varying macroeconomic shocks that might affect urban unemployment rate.

There is no empirical study on how urban industry structure relates to urban unemployment rate in China. This is surprising given that the massive and rapid urbanization, industry restructuring, and transition to a market economy have resulted in large scale unemployment in cities during the past two decades. In this paper we aim to provide the first empirical evidence. We use the urban industry employment data from 1997－2006 and test how urban industry structure correlates with urban unemployment rate and its instability. We employ a panel data, city fixed effect model to control for unobserved location attributes that might affect unemployment. Controlling for location specific effects is very important in studying urban unemployment in China because there are many different types of cities in China, such as cities directly under the central government, province capital cities, special economic zones, and coastal open cities.

[2] Note that urban industry diversity and specialization are not completely opposite or perfectly linear. For example, cities with the same degree of industry diversity can have very different degrees of industry specialization if the industry categories are different or the distribution of employment varies a lot across industries (Mizuno, Mizutani, and Nakayama, 2006). By the same token, two cities can have the same multiple specializations, but the degrees of industry diversity can be very different if the number and employment distribution of other industries are very different. Also a region can specialize in a particular industry in the short run, but can diversify its industry mix in the long run (Wagner and Deller, 1998).

Besides differences in natural advantages, different types of cities have different status in the urban hierarchy and receive different government policy intervention, which possibly have affected the employment in those cities. Using city fixed effects can control for these differences across cities. In addition, we control for the important urban labor market attributes, such as industry specialization, industry diversity, labor market thickness, human capital level, and the degree of market maturity.

Our empirical results show that in contrast to the evidence from developed countries, after controlling for city fixed effects and multiple industry specializations, industry diversity is significantly positively correlated with unemployment rate; industry diversity is not correlated with unemployment instability. We also find that wholesale and retail industry specialization is negatively correlated with unemployment rate, but finance industry is positively correlated with unemployment rate. In addition, market maturity, measured by the proportion of private employment, and employment growth potential, are significantly negatively correlated with unemployment rate. Although industry diversity can not explain the instability of unemployment rate, industry churning does.

The rest of the paper is organized as follows: section 2 specifies the econometric model and defines the variables; section 3 introduces the data set; section 4 reports estimate results; section 5 further discusses industry churning; and section 6 concludes.

2. Econometric model specification and variable definitions

This section specifies the baseline econometric model that tests how urban industry structure in China is related to unemployment rates. Different from most of the cross-sectional estimates, we adopt a panel data approach with city and year fixed effects to control for unobserved, time-unvarying city attributes that might be correlated with city unemployment and time-varying macroeconomic shocks that are common to all cities. The model is specified as follows:

$$Unemprate_{it} = \beta_0 + \lambda_i + T_t + \beta_1 X_{it} + \varepsilon_{it}, \qquad (1)$$

where the dependent variable is $Unemprate_{it}$, the registered unemployment rate in city i in year t， defined as

$$Unemprate_{it} = \frac{\text{Number of registered unemployed workers}_{it}}{\text{Number of employed workers}_{it} + \text{number of registered unemployed workers}_{it}} \times 100\%$$

, where the number of employed workers is the sum of unit employment and private sector employment.[3] The reason why we use registered unemployment rate is explained in the data section.

β_0: constant term;

[3] "Unit employment" can be simply interpreted as non-private sector employment.

λ_i : city fixed effect, used to control for unobserved, time-unvarying city specific attributes;

T_t : year fixed effect, used to control for unobserved, time-varying macroeconomic shocks that are common to all cities;

X_{it} : the vector of labor market attributes in city *i* in year *t*;

β_1 : coefficient vector of labor market attributes variables to be estimated;

$\varepsilon_{\mathrm{it}}$: disturbance term, possibly not identically and not independently distributed.

Following most of the existing literature reviewed in the introduction section, we select the following variables to represent urban labor market attributes:

$Diversity_{it}$: industry diversity index for city *i* in year *t*, defined as one minus the Herfindahl index in terms of industry employment. Specifically,

$$Diversity_i = 1 - \sum_n e_{in}^2 , \qquad (2)$$

where e_{in} is the ratio of unit employment in industry *n* located in city *i* to the total unit employment in city *i*. The value of this index is between zero and one. When it is close to one, it implies a higher degree of industry diversity. Note that this index takes into account both the number of industries and the employment distribution across different industries. Here we use one-digit industry classification. We should point out that the industry employment is "unit employment" which does not include employment in private sectors. This is because data for private sector employment by industry at the city level are not available in China. Omitting private sector employment by industry creates some measurement error bias, but since private sector employment is much smaller than unit employment (on average 26% of total employment in cities) and private employment also distributes across all industries, we believe that the diversity index measured in terms of unit employment shares is highly correlated with the actual industry diversity in cities.

Log(*pop*): the logarithm of non-agricultural population in a city. Vipond (1974) argues that in large cities wage dispersion is wider, leading to longer job search; also in large cities commuting costs are higher, increasing the cost of accepting new jobs. Therefore, the equilibrium unemployment rate in large cities is higher. Simon (1988) also finds positive correlation between city size and unemployment rate. Duranton and Puga (2000) argue that large cities host more diverse industries, increasing labor demand and reducing unemployment; medium and small cities tend to be specialized, generating smaller labor demand. In addition, labor markets in large cities are thicker or more dense, generating more labor demand and labor supply, and therefore, improving the matching quality of jobs and works and reducing the equilibrium unemployment rate (Gan and Li, 2004; Gan and Zhang, 2006). We add the city size variable to control for both city size effect and labor market thickness effect.

Log(*student*): the logarithm of the number of students at colleges or universities per 10, 000 population in a city, proxy for the average education of labor force in a city. Workers with higher human capital can easily find or retain their jobs even in recession, or can easily move to other cities to get jobs. Therefore, higher human capital stock of the labor force reduces unemployment rate. Simon and Nardinelli (2002) find that human capital stock promotes future employment growth in American cities. A few studies also find that education level reduces unemployment rates (Simon, 1988; Mizuno, Mizutani, and Nakayama, 2006). Human capital is generally measured by schooling years or education level. However, systemic education variables at the city level are not available in China. In Chinese cities where the average education level of labor force is higher, the number and size of colleges and universities are larger too. Also many college graduates tend to stay in the same city to work after graduation. We are clearly aware that *Log*(*student*) is not a perfect measure of the average education of labor force in a city, but we believe it is highly correlated to the human capital stock in a city.

Private： the ratio of total employment in the private sector to the total employment in a city, proxy for the degree of market development. China's transition to a market economy is mostly manifested by the growing private sector (Dougherty and He, 2007). This variable captures the spatial and temporal variations in terms of market maturity. As market economy is more efficient in allocating resources, the sign of the coefficient of *Private* is expected to be negative.

Growth: the annual growth rate of total employment in a city. China has had a rapid urbanization and fast growth since the 1990s. Rapid urbanization increases labor supply in cities and increases employment; however, too fast growth might be at the cost of higher unemployment rate. This variable controls for the impact of urban growth on unemployment. Malizia and Ke (1993) also use this variable and find that the coefficient is negative.

To test the effect of multiple specializations on structural and cyclical unemployment rates, we select eight representative industries. They are agricultural, forestry, and fishery, mining, manufacturing, public utility, construction, finance and insurance, real estate, wholesale and retail industries. The classification of these industries does not change during 1997-2006 period and most of them have been tested in other empirical studies. We construct the location quotient of each industry to measure the degree of specialization of each industry in each city relative to the national average, and label them by *Lqagri*, *Lqmining, Lqmanufactruing*, *Lqutility*, *Lqconstruct, Lqfinance, Lqrealestate*, and *Lqretail*, respectively. Taking construction industry as an example, its location quotient is computed as follows:

$$Lqconstruct_i = \frac{\frac{e_{ic}}{e_i}}{\frac{e_c}{\sum_i e_i}} \quad , \qquad (3)$$

where e_{ic} is the total unit employment in the construction industry in city *i* in a given year, e_i is the total unit employment in city *i* in the same year， e_c is the national unit employment in the construction industry, and $\sum_i e_i$ is the total unit employment in all cities. Location quotient reflects the relative specialization of an industry in a city compared to the national average level. However, if cities are not completely open, using the absolute specialization, the proportion of an industry employment in a city, is more reasonable. We also compute the ratio of unit employment in each industry to the total unit employment in a city, denoted by *Pagri*, *Pmining*, *Pmanufacturing*, *Putility*, *Pconstruct*, *Pfinance*, *Prealestate*, and *Pretail*, respectively.

The disturbance term, ε_{it}, might be heteroskedastic and serially and spatially correlated. We use the Huber/White standard variance estimator clustered by city to get the consistent estimate of standard errors. Since we have only an unbalanced panel, it is not feasible to correct for serial correlation, but we try using lagged dependent variable to partially remedy the serial correlation problem.

3. Data

The data used in this paper are drawn from the *China Urban Statistic Yearbook* from 1998 to 2007. The yearbooks include extensive information for cities above the county level (the so-called "Di Ji Shi") in China, including population, unit employment by one-digit industry, private sector employment, and other economic and demographic variables. The city in our sample is defined as city proper, including inner cities and suburban areas, but excluding suburban counties. According to the official statement in the *Yearbook*, suburban counties do not function as urban areas, and may change from time to time; however, city proper is relatively stable and embodies all the functions of urban areas. We exclude cities at the county level because they are usually small cities (population less than 200,000) and data for employment by industry are not available for these cities. We also drop cities with too many missing values, such as Lasa (the capital city of Tibet). Our final sample is an unbalanced panel dataset spanning 1997 to 2006, including 2,560 observations.

Since 1999 China started publishing the number of registered unemployment in cities. The registered unemployment rate is defined as the number of registered unemployed workers in the local government divided by the sum of employed workers and registered unemployed workers in the end of year. This paper uses the registered unemployment rate as dependent variable. Obviously, registered unemployment is smaller than actual unemployment since many unemployed workers did not register in the local government for different reasons. In addition, registered unemployment excludes unemployed workers who don't have permanent city residence ("Hukou"). The more precise measure is the survey unemployment rate, which is widely used internationally. The China Statistic Bureau started the survey unemployment project since November 2005 but the data on survey unemployment rate are not publicly available so far. We believe that the registered unemployment rate should be highly correlated with the actual unemployment rate: when the actual unemployment rate is higher, there should be more workers willing to register as unemployed. If we assume the measurement error in the dependent variable is not correlated with independent variables, then, the estimate of the constant term will be biased. However, we do not care much about the constant term. The measurement error in the dependent variable can also cause the standard errors of estimated coefficients increase, but we have no way to solve this problem. Note that the significance of estimated coefficients in our models will be smaller than that of using actual unemployment rate as dependent variable. Therefore, using registered unemployment rate provides a more conservative estimate than using actual unemployment rate or survey unemployment rate.

In addition, the 1997 and 1998 Yearbooks do not have registered unemployment data, but contain employment rate data. We use 100 minus the employment rate (in percentage point) to impute the unemployment rate. It is possible that the imputed unemployment rate is not completely consistent with the registered unemployment rate used in later years. We also try estimating models by dropping the 1997 and 1998 observations.

The unit employment data are disaggregated by one-digit industry. There are 15 industries during 1997－2002, and 19 industries since. These industry employment data are used to construct the industry diversity index and location quotient variables. Table A-1 in the appendix provides summary statistics for the variables used in estimation. There are large variations in unemployment rate, industry diversity, and specializations across cities.[4]

4. Results

4.1 Benchmark results

Table 2 reports the baseline model results. Column 1 is a pooled regression without city fixed effects. Many of the coefficients change signs when city fixed effects are added (see column 2). This suggests that omitting unobserved city attributes causes serious bias. For example, in the pooled regression the coefficient of diversity is biased toward zero, as discussed in Izraeli and Murphy (2003). We focus on interpreting column 2 results. The coefficient of *Diversity* variable is 6.986 and significant at the 1% level, meaning that if industry diversity increases by 1 percentage point, unemployment rate will increase by about 0.07 percentage point. This is in contrast to the theory that industry diversity reduces

[4] We find a few outliers in the data. We suspect possibly there are reporting or recording errors in the Urban Statistic Yearbooks. However, we keep those observations in estimation.

frictional unemployment and the empirical evidence from many developed countries (Simon, 1988; Malizia and Ke, 1993). Our conjecture is that since the 1990s China industries have gone through turbulent reform, transition, and restructuring. The reform of stated-owned enterprises has resulted in a large number of lay-offs (both structural and frictional unemployment); industry churning changes the distribution of employment across different industries; and some new industries emerge. All these changes are possibly positively related to industry diversity. Regulation and constraint on labor mobility still exists and prevents unemployed workers from efficient relocation (Knight and Li, 2005). In the next section, we will discuss industry churning and unemployment in detail.

The coefficient of *Log*(*pop*) is negative, suggesting that the labor market thickness effect might be dominant, but it is not significant. The coefficient of *Log*(*student*) is negative but not significant, possibly suggesting that this variable is correlated with average education or human capital stock at the city level, but not highly enough to make it statistically significant. The coefficient of *Private* is -1.627 and significant at the 6% level, suggesting that if the proportion of private sector employment increases by 10 percentage points, the unemployment rate will reduce by about 0.16 percentage point. The negative and statistically significant coefficient of *Growth* is expected and intuitive.

In terms of the effects of multiple specializations on unemployment rate, in column 2 only the coefficient of the location quotient of wholesale and retail industry is significantly negatively correlated with unemployment rate: if the location quotient of wholesale and retail industry increases by one in a city, the unemployment rate will decrease by about 1.1 percentage point. This result is not surprising given that wholesale and retail industry started introducing proprietorship since early 1980s. What is surprising is that the coefficient of the location quotient of finance and insurance industry is positive and significant, probably because in China most of the financial institutions are state-owned with strong monopoly power. Financial industry mostly absorbs high-skilled workers, and over-development of finance industry probably does not help much in generating new employment. Another possible reason is that in China economic growth promotes the development of finance industry, but not vice versa (Liang and Teng, 2006). Most of other industries are weakly or not associated with unemployment rate. The feature of the coefficients of location quotient variables in terms of industry employment is basically similar to what is found in Mizutan, Mizutani, and Nakayama (2006).

Readers may be concerned about the correlation between industry diversity and multiple specializations. Column 3 drops all the location quotient variables that measure multiple specializations, and the pattern of all other coefficients are pretty much the same, except the coefficient of *Private*. This is probably because some industries absorb much of the private sector employment, such as wholesale and retail, real estate, and agriculture. Without multiple specializations, the coefficient of *Diversity* increases from 6.986 to 7.687, increasing by 10%, suggesting that if multiple specializations are omitted, the coefficient estimate of *Diversity* will be biased upward.

Many similar studies did not include the growth rate of employment as one of the control variables. Column 4 drops the *Growth* variable, but the pattern of coefficients remains very similar to column 2. Manufacturing industry is one of the most important industries in most of the Chinese cities. Column 5 includes the location quotient of manufacturing employment and the results are again very similar to column 2 except that the coefficient of *Diversity* becomes much larger and less significant. This is because the correlation between *Diversity* and *Lqmanufacturing* is moderately high (correlation coefficient is -0.55). Therefore, it makes sense to drop the *Lqmanufacturing* variable so as to avoid the multicollinearity problem.

Since the unemployment rates for 1997 and 1998 are imputed by subtracting the employment rate (percentage point) from 100, it is possible that the imputed unemployment rates are not consistent with the registered unemployment rates. Column 6 addresses this concern and drops the observations of 1997 and 1998. Again, column 6 results are very similar to column 2, except that the coefficient of *Private* is twice large, suggesting that during the period of 1999-2006, private sector has played a more important role in reducing unemployment rate.

In summary, Table 1 results show that industry diversity and specialization in finance and insurance industry are positively associated with urban unemployment rate, while market maturity (measured by the proportion of private sector employment), urban economic growth (measured by employment growth), and specialization in wholesale and retail industry are negatively associated with unemployment rate.

Table 2
Benchmark results

Variables	1	2	3	4	5	6
Diversity	0.553	6.986***	7.687***	5.297***	7.058**	6.622***
	(0.33)	(3.21)	(3.74)	(3.12)	(2.29)	(3.08)
Log(pop)	0.006	-0.365	-0.186	-0.403	-0.366	-0.453
	(0.05)	(-1.00)	(-0.54)	(-1.23)	(-1.00)	(-1.23)
Log(student)	-0.114*	-0.055	-0.049	-0.038	-0.055	-0.088
	(-1.86)	(-0.94)	(-0.75)	(-0.72)	(-0.94)	(-1.45)
Private	1.263	-1.627*	0.001	-2.554***	-1.628*	-3.206***
	(1.19)	(-1.90)	(0.00)	(-3.41)	(-1.89)	(-3.95)
Growth	-1.756***	-1.103***	-1.604***		-1.102***	-0.622**
	(-6.78)	(-3.97)	(-5.92)		(-3.94)	(-3.11)
Lqagri	0.014	-0.027		-0.055*	-0.027	-0.048*
	(0.29)	(-0.80)		(-1.69)	(-0.73)	(-1.67)
Lqmining	-0.010	-0.068		-0.067	-0.067	-0.137
	(-0.21)	(-0.58)		(-0.67)	(-0.53)	(-1.16)
Lqconstruct	0.059	-0.517		-0.440	-0.515	-0.628**
	(0.25)	(-1.53)		(-1.45)	(-1.56)	(-2.24)
Lqfinance	-0.018	0.820***		0.978***	0.820***	0.581***
	(-0.06)	(3.36)		(4.16)	(3.35)	(2.18)
Lqutility	0.588***	0.247		0.353*	0.246	0.186
	(2.48)	(1.19)		(1.83)	(1.16)	(0.97)
Lqrealestate	0.610*	0.030		0.236	0.030	0.232
	(1.63)	(0.13)		(0.92)	(0.13)	(1.01)
Lqretail	-0.559*	-1.110***		-0.970***	-1.108***	-1.092***
	(-1.87)	(-3.51)		(-3.36)	(-3.50)	(-3.58)
Lqmanufacturing					0.027	
					(0.04)	
Time period	1997-2006	1997-2006	1997-2006	1997-2006	1997-2006	1999-2006
City fixed effect	No	Yes	Yes	Yes	Yes	Yes
Sample size	2254	2254	2257	2560	2254	2035
Adjusted R^2	0.08	0.50	0.49	0.48	0.50	0.54

Note. Dependent variable: Unemployment rate in percentage point. All models include year fixed effects. Standard errors are estimated using the Huber/White standard variance estimator, clustered by city. t statistics are in the parentheses. Superscript "*", "**", and "***" indicate the significance at the 1%, 5%, and 10% levels, respectively.

4.2 More robust checks

Using location quotient to measure the specialization of an industry in a city is appropriate when cities are open. If cities are not completely open or even closed, using the absolute specialization index, the ratio of employment in the industry in question to the total employment in a city, makes more sense. Chinese cities obviously are not completely open as there still exists much regulation on urban residence (the "Hukou" system that restricts farmers to become permanent city residents and restricts city residents' intra-city migration). We also replace all the location quotient variables by the corresponding proportional employment of each industry in a city and re-estimate all the models in Table 2. The results are very similar and we report only the baseline results in column 1 of Table 3.

Some studies use the Theil entropy index to measure industry diversity (Attaran, 1986). We also replace the *Diversity* variable by the Theil index, computed using the formula $\sum_n e_{in} \ln\left(\frac{1}{e_{in}}\right)$. In our sample the correlation between the Theil index and the *Diversity* variable is about 0.95. As expected, using the Theil index to measure industry diversity does not change the pattern of our estimate results.

Since we have an unbalanced panel data set spanning 10 years, serial correlation among the disturbance terms might be a concern. Usually adding the lagged dependent variable to the right-hand side can partially reduce serial correlation. We also add the lagged one-period unemployment rate to the right-hand side of equation (1). Again, the results, reported in column 2 of Table 3, are very similar to column 2 of Table 2, although the coefficient of *Private* becomes marginally significant now.

We go further and use all the lagged one-period independent variables to test if past industrial structure has impact on current unemployment rates. Results are reported in column 3 of Table 3. Most of the coefficients keep the same sign but become more statistically significant. The coefficient of *Diversity* is still positive and significant with similar magnitude, suggesting that higher industry diversity increases unemployment rate. We will provide some explanation in section 5. In addition, city size, human capital, market development, urban growth, and specialization in wholesale and retail industry all contribute to lower unemployment rates.

Table 3
Robust checks results

	1		2		3
Variable	Coefficient	Variable	Coefficient	Variable	Coefficient
Diversity	6.885***	Diversity	5.780***	$Diversity_{t-1}$	6.217***
	(3.16)		(3.00)		(3.56)
Log(pop)	-0.450	Log(pop)	-0.234	$Log(pop)_{t-1}$	-0.578*
	(-1.19)		(-0.70)		(-1.65)
Log(student)	-0.052	Log(student)	-0.055	$Log(student)_{t-1}$	-0.093*
	(-0.88)		(-1.02)		(-1.86)
Private	-1.771**	Private	-1.246	$Private_{t-1}$	-3.990***
	(-2.02)		(-1.58)		(-5.07)
Growth	-1.102***	Growth	-1.503***	$Growth_{t-1}$	-2.434***
	(-3.87)		(-4.96)		(-8.30)
Pagri	-1.909	Lqagri	-0.023	$Lqagri_{t-1}$	-0.074***
	(-1.37)		(-0.75)		(-2.56)
Pmining	-2.348	Lqmining	-0.057	$Lqmining_{t-1}$	-0.219*
	(-0.97)		(-0.55)		(-1.90)
Pconstruct	-7.018*	Lqconstruct	-0.445	$Lqconstruct_{t-1}$	-0.885***
	(-1.85)		(-1.42)		(-4.65)
Pfinance	20.402***	Lqfinance	0.654***	$Lqfinance_{t-1}$	0.069
	(2.48)		(2.90)		(0.36)
Putility	10.601*	Lqutility	0.218	$Lqutility_{t-1}$	0.283
	(1.26)		(1.16)		(1.09)
Prealestate	-0.498	Lqrealestate	0.098	$Lqrealestate_{t-1}$	-0.286
	(-0.03)		(0.46)		(-1.32)
Pretail	-15.518***	Lqretail	-0.999***	$Lqretail_{t-1}$	-0.708**
	(-3.89)		(-3.48)		(-2.29)
		Unemprate $_{t-1}$	0.189***	Unemp rate $_{t-1}$	0.187***
			(3.86)		(3.83)
Sample size	2254		2254		2254
Adjusted R^2	0.50		0.52		0.53

Note. Dependent variable: Unemployment rate in percentage point. All models include city fixed effects and year fixed effects. Standard errors are estimated using the Huber/White standard variance estimator, clustered by city. t statistics are in the parentheses. Superscript “*”, “**”, and “***” indicate the significance at the 1％, 5％, and 10％ levels, respectively.

4.3 Industry structure and unemployment instability

Some studies show that industry diversity can reduce the fluctuation of unemployment and stabilize employment (Malizia and Ke, 1993). We also test if in Chinese cities industry diversity reduces unemployment instability. We use the standard deviation of unemployment rates during 1997-2006 in each city as the measure of unemployment instability and use it as the dependent variable. Table 4 reports the results of three models. In column 1 the independent variables include all the variables in column 2 of Table 2 but the values are set in 1997 values (except the *Growth* variable uses the average growth rate of employment during 1997-2006 in each city). Column 2 uses the values that are first available for each city. Column 3 uses the average value of each variable during 1997-2006 for each city. The overall results suggest that industry diversity does not reduce unemployment fluctuations. Interestingly, there is some evidence that human capital level, market maturity, specialization in finance industry, and city size might reduce unemployment fluctuations. However, the specialization of wholesale and retail industry increases unemployment instability, possibly because wholesale and retail employment is more sensitive to business cycle and seasonal factors.

Table 4
Industry structure and unemployment instability

	1	2		3
Variable	Coefficient	Coefficient	Variable	Coefficient
Diversity	1.250	0.834	Diversity	6.389***
	(0.55)	(0.41)		(2.97)
Log(pop)	-0.087	-0.267	Lnpop	-0.184
	(-0.55)	(-2.08)		(-1.33)
Log(student)	-0.088*	-0.051	Lnstudent	-0.161**
	(-1.74)	(-1.18)		(-1.22)
Private	-4.912**	-0.580	Private	-0.139
	(-1.99)	(-0.51)		(-0.09)
Average growth	-0.318	-0.108	Average growth	-0.134
	(-1.30)	(-0.41)		(-0.45)
Lqagri	0.651***	0.029	Lqagri	-0.049
	(2.61)	(0.32)		(-0.76)
Lqmining	-0.004	-0.050	Lqmining	-0.016
	(-0.06)	(-0.81)		(-0.31)
Lqconstruct	-0.066	-0.083	Lqconstruct	-0.259
	(-0.22)	(-0.30)		(-0.91)
Lqfinance	-0.427	-0.597*	Lqfinance	-0.889***
	(-1.29)	(-1.81)		(-2.68)
Lqutility	0.459	0.277	Lqutility	0.515**
	(1.33)	(1.02)		(1.90)
Lqrealestate	0.148	-0.004	Lqrealestate	0.517
	(0.66)	(-0.02)		(1.30)
Lqretail	1.578***	0.689**	Lqretail	0.525
	(3.87)	(1.99)		(1.43)
Sample size	222	274		271
Adjusted R^2	0.25	0.20		0.29

Note. Dependent variable: The standard deviation of unemployment rates in percentage point during 1997-2006 in each city. Column 1 variables use the 1997 values; column 2 variables use the first available values during 1997-2006; column 3 variables use the mean over 1997-2006. All models include province fixed effects. Standard errors are estimated using the Huber/White standard variance estimator, clustered by city. t statistics are in the parentheses. Superscript "*", "**", and "***" indicate the significance at the 1%, 5%, and 10% levels, respectively.

5. Industry churning and unemployment

So far we have found that industry diversity increases unemployment rate. This is puzzling since the theory predicts that industry diversity reduces frictional unemployment, as evidenced by many studies. Here we provide a plausible explanation: during the sample period, China urban industries have undergone dramatic structural change and churning. Industry churning changes the proportion of employment in each industry, leading to changes in industry diversity.

In this section, following Duranton (2007) and Findeisen and Suedekum (2008), we compute the urban industry churning index and total employment churning index for Chinese cities and discuss how churning affects unemployment rate. The industry churning index, *Churn*, is computed using the following formula:

$$Churn_i = \frac{1}{T}\left[\sum_{t=1997}^{2002}\sum_{n=1}^{15}\frac{\left|e(n,i,t+1)-e(n,i,t)\right|}{e(i,t)}\right], \tag{4}$$

where $e(n,i,t)$ is the total employment of industry n in city i in year t; T is the number of years that a city has employment data. Since 2003 the number of one-digit industries increases from 15 to 19, and the definitions of some industries changed. Therefore, we compute the churning index using only data from 1997-2002. The *Churn* index describes the average degree of fluctuations of employment in all industries in a city during 1997-2002.

By the same token, we compute the total employment churning index, ΔEMP, for each city, using the similar formula:

$$\Delta EMP_i = \frac{1}{T}\left[\sum_{t=1997}^{2002}\frac{\left|e(i,t+1)-e(i,t)\right|}{e(i,t)}\right], \tag{5}$$

where $e(i,t)$ is the total employment of city i in year t. This index measures the average degree of total employment fluctuations in a city during 1997-2002. [5]

The industry churning index and the total employment churning index will be equal if the employment of each industry is always growing or decreasing during 1997-2002; however, if in a city the employment of some industries grows while employment of some other industries decays, then, the industry churning index will be greater than the total employment churning index as the growth in some industries offsets the decrease in other industries when computing the total employment churning index. Following Duranton (2007) and Findeisen and Suedekum (2008), we define the gap between industry churning index and total employment churning index as the excessive churning of industry employment. In addition, the ratio of industry churning index to the total employment

[5] Here we use only unit employment since employment in private sector is not classified by industry. We also treat private sector employment as an additional industry, and re-compute the industry churning index and the total employment churning index, and the results are similar.

churning index can be used to measure the degree of excess churning too. Table 5 shows that the industry churning index is larger than total employment churning index: the mean of the excessive churning is 3.3 percentage points, and the mean of the ratio of industry churning index to the total employment churning index is about 1.7.

Table 5
Industry churning in Chinese cities (1997-2002)

Index	Mean	Standard deviation	Maximum
Churn	0.206	0.097	0.870
ΔEMP	0.173	0.130	1.187
Excessive churning (*Churn* - ΔEMP)	0.033	0.061	0.248
$\frac{Churn}{\Delta EMP}$	1.727	2.602	34.014

How excessive is industry churning in Chinese cities? Table 6 compares churning indexes of Chinese cities with those of West Germany, the USA, and France, where data for West Germany is cited from Findeisen and Suedekum (2008) and data for the USA and France are cited from Duranton (2007). Obviously, industry churning and total employment churning in Chinese cities are much larger than those in developed countries. One possible explanation is that during the transition to a market-oriented economy, Chinese stated-owned enterprises have undergone a dramatic transformation in institutions and governance. With rapid urbanization and high growth, Chinese urban economies have also undergone dramatic industrial structural change. If urban unemployment is mainly structural, then, we would not observe significant negative correlation between industry diversity and unemployment rate.

Table 6
Industry churning: International comparison

Country	*Churn*	ΔEMP	$\frac{Churn}{\Delta EMP}$	Number of cities or metropolitan areas	Time period
China	20.60%	17.3%	1.19	265	1997-2002
West Germany	4.98%	2.29%	2.17	326	1977-2001
United States	8.26%	4.10%	2.01	272	1977-1997
France	11.4%	5.20%	2.19	217	1985-1993

Note. Data for China are computed by the authors; data for West Germany are from Findeisen and Suedekum (2008); and data for the USA are from Duranton (2007).

High degree of industry churning generates fast turnovers of industry employment. This process may lead to a relatively higher degree of and higher volatility of industry diversity. In fact, the correlation between *Churn* and the mean industry diversity of each city is about 0.13 and the correlation between *Churn* and the standard deviation of industry diversity of each city during 1997-2002 is about 0.41. To present a piece of more intuitive evidence, we estimate a cross section model, using the average unemployment rate in a city during 1997-2002 as the dependent variable. Column 1 of Table 7 uses industry churning index and province fixed effects as independent variables; column 2 adds more city attributes variables valued by the mean over 1997-2002 in each city. We do not include the mean location quotient variables because all of them are moderately correlated with at least one of the mean city attribute variables and most of them are not significant. The results show that industry churning is significantly positively correlated with average unemployment rate, while the coefficient of mean industry diversity becomes insignificant when *Churning* is included. We also do similar estimation for unemployment instability, and the results are presented in columns 3 and 4. Again, industry churning significantly contributes to unemployment fluctuations, while the mean industrial diversity becomes just marginally significant. In summary, we tentatively conclude that the positive, significant correlation between industry diversity and unemployment rates in Chinese cities is most probably driven by the high degree of industry churning in Chinese cities during the sample period.

Table 7
Industry churning and unemployment

	Dependent variable			
	Mean unemployment rate		Standard deviation of unemployment rate	
Coefficient				
	1	2	3	4
Churning	4.310*	5.215*	4.130***	5.450**
	(1.85)	(1.85)	(2.7)	(2.31)
Mean diversity		2.738		3.149
		(0.89)		(1.48)
Mean Log(pop)		-0.072		0.063
		(-0.26)		(0.36)
Mean Log(student)		-0.013		-0.150
		(-0.05)		(-0.78)
Mean Private		10.881		2.133
		(1.60)		(0.86)
Mean Growth		-0.673		-0.685
		(-0.23)		(-0.54)
Sample size	266	248	265	248
Adjusted R^2	0.26	0.50	0.27	0.29

Note. The dependent variables in columns 1 and 2 are the mean unemployment rate in percentage point of each city during 1997-2002. The dependent variables in columns 3 and 4 are the standard deviation of unemployment rates in percentage point of each city during 1997-2002. All the independent variables are valued by the mean over 1997-2002. Province fixed effects are included and robust standard error estimators are used. t statistics are in the parentheses. Superscript "*", "**", and "***" indicate the significance at the 1%, 5%, and 10% levels, respectively.

6 Conclusion

This paper uses the Chinese city level data from 1997-2006 and tests how industry specialization, diversification, and churning affect urban unemployment rate. The panel data estimation shows that in contrast to the evidence from developed countries, industry diversity is positively and significantly associated with unemployment rate. We interpret this as driven by high degree of industry churning in Chinese cities during the sample period. We also find that multiple specializations have different impact on unemployment rate: the specialization of wholesale and retail industry can significantly decrease unemployment rate; however, the specialization of finance industry increases unemployment rate. Urban economic growth, market maturity measured by the proportion of private sector employment, and human capital can decrease unemployment rate. We also find that industry diversity can not stabilize unemployment; wholesale and retail industry increases unemployment fluctuations; but market maturity and human capital can reduce unemployment fluctuations.

We are clearly aware of the two data quality issues: (1) the unemployment rate is measured by registered unemployment rate, which underestimates the actual unemployment rate; (2) the industry employment is “unit” employment, which excludes the private sector employment. However, this data set is the best, systematic, available city level panel data in China for the study of industry structure and unemployment.[6] We have discussed in detail our strategies to deal with these two issues. We have also done a few robust checks. Despite the data quality issues, we believe the empirical regularity we identified is pretty robust. Our findings also remind policy makers to take caution when implementing industry diversification policy to reduce regional unemployment.

[6] Some cities have implemented minimum wages and unemployment insurance. However, systematic city level data on minimum wages and unemployment insurance are not available now.

Appendix

Table A-1
Summary statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
Unemprate	4.12	3.23	0	45.31
Diversity	0.80	0.08	0.41	0.91
Log(pop)	3.84	0.84	1.62	7.05
Log(student)	4.68	2.76	-6.17	9.08
Private	0.26	0.15	0	0.82
Lqagri	1.25	2.58	0	41.05
Lqmining	1.35	2.81	0	18.71
Lqutility	1.31	0.79	0	6.66
Lqconstruct	0.98	0.62	0	6.13
Lqfinance	1.16	0.48	0	4.17
Lqrealestate	0.71	0.57	0	6.22
Lqretail	0.88	0.45	0.12	4.73
Lqmanufacturing	0.93	0.42	0.05	2.33
Pagri	0.05	0.10	0	0.68
Pmining	0.06	0.13	0	0.71
Putility	0.03	0.02	0	0.17
Pconstruct	0.08	0.05	0	0.54
Pfinance	0.03	0.02	0	0.11
Prealestate	0.01	0.01	0	0.11
Pretail	0.07	0.05	0.01	0.48
Pmanufacturing	0.31	0.14	0.02	0.76
Growth	-0.003	0.30	-0.91	3.77

References

Attaran, M., 1986. Industrial diversity and economic performance in U.S. areas. The Annals of Regional Science 20, 44-54.

Barkume, A., 1982. Differentiating employment prospects by industry and returns to search in metropolitan areas. Journal of Urban Economics 12, 68-84.

Barth, J., Kraft, J., Wiest, P., 1975. A portfolio theoretic approach to industrial diversification and regional employment. Journal of Regional Science 15, 9-15.

Chiang, S., 2008. The effects of industrial diversification on regional unemployment in Taiwan: Is the portfolio theory applicable? The Annals of Regional Science, forthcoming.

Dougherty, S., He, P., 2007. Has a private sector emerged in China's industry? Evidence from a quarter of million Chinese firms. China Economic Review 18, 309-334.

Duranton, G., Puga, D., 2000. Diversity and specialization in cities：Why, where and when does it matter? Urban Studies 37, 533-555.

Duranton, G., 2007. Urban evolutions: The fast, the slow, and the still. American Economic Review 97, 197-221.

Findeisen, S., Suedekum, J., 2008. Industry churning and the evolution of cities: Evidence from Germany. Journal of Urban Economics 64, 326-339.

Gan, L., Li, Q., 2004. Efficiency of thin and thick markets. NBER Working Paper 10815.

Gan, L., Zhang, Q., 2006. The thick market effect on local unemployment rate fluctuations. Journal of Econometrics 133, 127-152.

Izraeli, O., Murphy, K., 2003. The effect of industrial diversity on state unemployment rate and per capita income. The Annals of Regional Science 37, 1-14.

Knight, J., Li, S., 2005. Wages, firm profitability and labor market segmentation in Urban China. China Economic Review 16, 205-228.

Liang, Q., Teng, J., 2006. Financial development and economic growth: Evidence from China. China Economic Review 17, 395-411.

Malizia, E., Ke, S., 1993. The influence of economic diversity on unemployment and stability. Journal of Regional Science 33, 221-235.

Mizuno, K., Mizutani, F., Nakayama, N., 2006. Industrial diversity and metropolitan unemployment rate. The Annals of Regional Science 40,157-172.

Simon, C., 1987. Industrial diversity, vacancy dispersion and unemployment. The Annals of Regional Science 21, 60-73.

Simon, C., 1988. Frictional unemployment and the role of industrial diversity. Quarterly Journal of Economics 103, 715-728.

Simon, C., Nardinelli, C., 2002. Human capital and the rise of American cities, 1900-1990. Regional Science and Urban Economics 32, 59-96.

Tarzwell, G., 1999. Canadian city unemployment rates and the impact of economic diversity. Canadian Journal of Regional Science 20, 389-399.

Vipond, J., 1974. City size and unemployment. Urban Studies 11, 39-46.

Thompson, W., 1965. A Preface to Urban Economics. Johns Hopkins Press, Baltimore.

Wagner, J., Deller, S., 1998. Measuring the effects of economic diversity on growth and stability. Land Economics 74, 541-556.