ELSEVIER

Returns to dialect
Identity exposure through language in the Chinese labor market

Zhao CHEN [a], Ming LU [a,b,*], Le XU [a]

[a] *Fudan University, China*
[b] *Shanghai Jiao Tong University, China*

ABSTRACT

Though Mandarin is China's common language, each region/city has its own dialect. Using a unique self-collected dataset, this paper estimates returns to dialect familiarity in China's largest and most developed city, Shanghai. We evaluate migrant workers' comprehension and fluency of the Shanghai dialect, and instrument their dialect fluency by determining whether the workers' hometowns were located in the *Wu* dialect region and the distance between those hometowns and Shanghai. We determined that in OLS regressions, the returns to dialect are a consequence of endogeneity bias. After using IV (instrumental variable), dialect fluency was shown to significantly impact one's income in the service industry, in particular affecting sales jobs. In manufacturing and construction jobs, migrants with higher dialect fluency tended to be self-employed in order to earn more income. By distinguishing between listening and speaking abilities, we found that auditory comprehension does not significantly increase one's earning, while oral fluency does. Since local residents in Shanghai can understand Mandarin, migrants who can understand Shanghainese won't have difficulty in the information exchange. Therefore, our results confirm that dialect is a channel through which people expose their identity. Speaking the local dialect is a way for migrant workers to integrate into the local society and also to reduce transaction costs in the labor market.

© 2014 Elsevier Inc. All rights reserved.

"*A language is a dialect with an army and navy.*"[1]

A dialect is a language without an army and navy.[2]

1. Introduction

China is experiencing rapid urbanization. The urban population accounted for only 20.91% of China's population in 1982 but increased to 52.6% of the population in 2012.[3] According to NBS (National Bureau of Statistics of China, 2013), 236 million Chinese

people live in a place where they do not have a local *hukou* (a household registration as a local resident), among whom most are rural-to-urban migrants.[4] In Shanghai, for instance, the 2010 population census shows that the population reached 23 million in 2010, 39% (8.98 million) of which were migrants without local *hukou*. Shanghai's population has increased at an annual rate of 3.24% over the past decade, whereas the population with local *hukou* has experienced a continuous decrease for almost two decades.[5] Due to *hukou* discrimination in the labor market, social security, and public service (Lu, 2013), many migrants choose to return home as they age. Once an individual from a rural area is older than 33, the probability of that individual migrating to a city decreases (Chen, Jiang, Lu, & Sato, 2013). The *hukou* system and the urban–rural divide have led to a lag in urbanization compared to industrialization (Chen, 2013),[6] a shortage of labor in the cities (especially in the coastal areas (Cai & Du, 2011)), a decrease in migrant consumption (Chen, Lu, & Zhong, 2012), and an overall decrease in migrant happiness and trust in cities (Jiang, Lu, & Sato, 2012; Wang, Chen, & Lu, 2009).

Although reform of the *hukou* system has stagnated, especially in big cities, scholars are of the opinion that the *hukou* system should be reformed either by giving migrants local *hukou* identities or by removing discrimination against the non-*hukou* population. However, there are few studies on non-*hukou* factors against migrants' integration into local societies. In reality, a majority of migrants would stay in cities even if they could not obtain local *hukou* (Development Research Center of State Council, 2011). Non-*hukou* factors against social integration would become increasingly important both for migrants living in cities and for urban development in general; needless to say, *hukou* discrimination will be removed in the future.

A great deal of research has been conducted on the assimilation of international migrants in host countries, language return being one of the focuses (Chiswick & Miller, 1995; Chiswick & Repetto, 2000; Mcmanus, 1985, etc.). Language ability is always combined with other factors, such as ethnicity and invisible ability, which causes difficulty in estimating returns to languages. Similarly, in China, many rural-to-urban migrants are moving between regions and speaking different dialects wherever they live. Whether dialect difference is an obstacle to social integration between migrants and local residents remains an open question. Chen et al. (2013) seeks to identify the pure effects of language communication on economic performance. People speaking different Chinese dialects are not visually or ethnically different. Mandarin is the official language for reading and writing; therefore, a dialect only takes the form of spoken language and, thus, solely reflects the role of language in communication and social interaction. Because dialects are associated with regions, Chen et al. (2013) measured differences in dialects based on whether a person's *hukou* belongs to the same dialect region as the city in which they live. The measure that they adopt is an objective index of differences in dialects that avoids measurement errors in self-reported language ability. Because all people from the same region are assigned the same value of the measure, their dialect measurement is not related to individual ability.

Language not only conveys knowledge and information through mutual communication but also forms a channel for identifying whether a person is truly local. Those original local people share the same culture, habits, and ways of thinking, all of which form mutual trust among them; however, that trust score is lower among migrants insofar as the general public, local community, and government institutions are concerned (Wang et al., 2009). Thus, it is interesting to disentangle the "pure language returns" into communication effects and identity effects. In our study, we evaluate the ability of migrant workers to understand and speak the Shanghai dialect, and we instrument dialect ability by whether one's hometown is located in the *Wu* dialect region (denoted by a dummy *wuyu*), the distance between the individual's hometown and Shanghai, and an interaction term between *wuyu* and distance. We find that in OLS regressions, the returns to dialect are largely due to an endogeneity bias. After using instrumental variables (IV), dialect fluency is shown to significantly impact income in service jobs and in sales jobs in particular. In manufacturing and construction jobs, migrants with higher dialect fluency tend to be self-employed and earn more. By distinguishing listening ability from speaking ability, we find that listening ability does not significantly increase earning, whereas speaking ability does. Because local residents in Shanghai can understand Mandarin, migrants who can understand Shanghainese have little difficulty in exchanging information. Therefore, our results confirm that dialect is a channel through which people expose their identity. If local people who share the same values, culture, habits and ways of thinking do not trust migrants who cannot speak the local dialect, then speaking the local dialect becomes a way to integrate into the local society and reduce transaction costs in the labor market.

This paper is structured as follows: Section 2 reviews economics studies on language returns. Section 3 introduces the data we used. Section 4 presents the results of our data analysis, and Section 5 presents three groups of robustness checks. The last section concludes the paper.

2. Literature review

Language is a communication bridge that can generate economic returns. The question is, what type of communication? Language abilities yield returns through two channels, the first being the exchange of valuable information through communications in production. Marschak (1965) discussed the economics of uncertainty that characterize the problems of human information,

---

[4] *Hukou* is a registration identity that is determined according to the individual's mother's *hukou*, which is usually based on where she was from. *Hukou* is linked with locally financed social security and public services and results in discrimination against migrants, of whom only a limited number can change their *hukou*. It is especially difficult to change a hukou in big cities, such as Shanghai, Beijing and Guangzhou. Please see Chan and Buckingham (2008), Chan (2009) and Lu (2013) for details of the *hukou* system.

[5] Source: *The Data Report of Shanghai's* Sixth Population Census, available at the official website of Shanghai Bureau of Statistics, http://www.stats-sh.gov.cn/sjfb/201105/218819.html.

[6] In 2012, the urban population accounted for 52.6% of the total population in China, whereas secondary and tertiary industries accounted for 89.9% of the total GDP (NBS, 2013).