The impact of ethnic diversity on the quality of exports: the case from China

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Journal of Economic Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>JES-08-2019-0367.R3</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Quality, Exports, Ethnic Diversity, Product Differentiation, China</td>
</tr>
</tbody>
</table>
The impact of ethnic diversity on the quality of exports: Evidence from China *

This Version: July 2020

Abstract

In this paper, I investigate the impact of ethnic diversity on the quality of exports from China. I employ the recent firm-level Chinese export data, merged with the Industrial Census and the 2000 National Population Census. My data shows that ethnically homogeneous provinces export products of 10 percent higher quality on average than ethnically heterogeneous provinces. More interestingly, this impact depends on the characteristics of the products. In particular, ethnic diversity has a negative impact on differentiated products but a positive impact on homogeneous products.

Keywords: Quality; Exports; Ethnic Diversity; Product Differentiation; China

JEL Classifications: F14, O1, R23.

*I would like to thank Volodymyr Lugovskyy, Bo Chen, Zhiyuan Li, Hong Ma, Miaojie Yu, Meredith Crowley, Gene Grossman, Sandra Poncet, Gabriel Felbermayr, Tony Venables, Peter Egger and participants in the seminars at Shanghai University of Finance and Economics, Ningbo Nottingham, the University of International Business and Administration, the GEP conference, the Austrasian Trade Workshop for helpful comments. All the remaining errors are my own.
1 Structured abstract

1.1 Purpose

The quality of exports is an important indicator of the economic development of a country. It is, therefore, important to understand its determinants. Existing literature reveals a number of determinants such as the position of a country, the distance to the trading partner, the size of the firm, and the productivity or the quality of inputs. In this study, I propose a new determinant, namely, ethnic diversity.

1.2 Methodology

In the first stage of my research, I estimate the quality of exports following Berry (1994) approach. I also address the endogeneity problem as suggested by Khandelwal (2010). My measure of ethnic diversity follows Easterly and Levine (1997). In the second stage, I regress the estimated quality on the measure of ethnic diversity, interacted by the degree of differentiation of the goods.

1.3 Findings

First, I find that ethnic diversity in general reduces the quality of exports. Second, I find that the impact varies with the degree of differentiation. In particular, while ethnic diversity decreases the quality of differentiated goods, it increases the quality of homogeneous goods.

1.4 Research limitations/implications

Because of data limitations, the research cannot investigate the impact of migration.

1.5 Practical/Social implications

The findings of my paper show how we can improve the quality of products based on the characteristics of the population and of the products.
1.6 Originality/value

In the last decade, there has been growing attention paid to the impact of ethnic diversity on economic performance. As industrial powerhouses such as China continue to urbanize, stitching together once-disparate ethnic groups, the role ethnic diversity plays in the economic life of a country must be examined from a variety of angles. My paper is the first to investigate the specific impact of that diversity on product quality and on how diversity interacts with the characteristics of products.
2 Introduction

The notion of quality is of long-standing interest in economics. According to the Linder hypothesis (Linder 1961), the quality of the traded goods is an indicator of a country’s development, which is supported by evidence across the world (Hummels and Klenow 2005; Hallak 2006; Schott 2004). It is, therefore, not surprising that many studies have investigated the determinants of quality, such as the “position” of a country (Hidalgo et al. 2007), the distance to the trading partner (Bastos and Silva 2010), the size of the firms (Verhoogen 2008), their productivity (Johnson 2012) or the quality of inputs (Manova and Zhang 2012).

Researchers are beginning to examine the role of ethnic diversity in economic performance. This factor has been shown to be a hindrance to economic growth (Easterly and Levine 1997) but has a pro-trade effect on the extensive margin via the transaction costs reduction (Dunlevy 2005; Herander and Saavedra 2005; Rauch and Trindade 2005) or the preference channel (Head and Ries 1998). Ethnic diversity also has strong effects on public goods provision (Alesina, Baqir and Easterly 1999), income inequality (Dincer and Hotard 2011; Dincer and Lambert 2012) and trust in others (Dincer 2011). My paper is the first to look at the specific impact of that diversity on product quality and how diversity interacts with the characteristics of the products. More specifically, I aim to answer the following questions: Does ethnic diversity have a statistically and economically significant impact on the quality of exports? Does this effect vary with the characteristics of the products?

Using the customs data from China, I find that ethnic diversity does have a significant impact on the quality of products. Exports from a hypothetical, completely ethnically heterogeneous province (i.e. where everyone belongs to a different ethnic group) in general shows that the quality was at least 10 percent lower as compared with a completely homogeneous province (where everyone belongs to the same ethnic group). But diversity also has positive effects. I find that its impact depends on the characteristics of the product. In particular, while
ethic diversity lowers the quality of the differentiated products, it can raise the quality of the homogeneous goods. This result helps shed light on how ethnic diversity affects the quality of products in the region.

The organization of the paper is as follows. In the next section, I explain why China offers a good case study, then I present the theoretical model in Section 4.2. Section 5 introduces the data and the measurement, while Section 6 presents the empirical specifications, results and robustness checks. Section 8 concludes.

3 Background: China as a case study

China is an ethnically diverse country; the mainland alone is composed of 56 ethnic groups of which 19 have at least one million members each. The ethnic minorities in China occupy several autonomous regions including Inner Mongolia, Guangxi Zhuang, Tibet, Ningxia Hui and Xingjian Uygur. There are also sub-provincial autonomous prefectures, as well as autonomous prefectures, counties, townships and villages scattered throughout China. Taken together, these factors show why China represents a good case study to examine the impact of diversity on its growing exports.

The literature suggests that ethnic diversity can hinder economic performance (Alesina and Ferrara 2005). Linguistic and cultural differences pose barriers that make business cooperation more challenging. Moreover, ethnic fractionalization implies a low level of governmental institution quality (Laporta et al. 1999) and of trust among people in the society (Dincer 2011). As a result, ethnic diversity has been the cause of less investment (Mauro 1995, Montalvo and Reynal-Querol 2005a) and more incidence of civil war (Montalvo and Reynal-Querol 2005b).

On the other hand, different ethnic groups have honed different skills, a specialization that in some cases dates back centuries. For instance, members of the Jino ethnic group, most of whom live in Yunnan province, are said to be great hunters. They are the experts in the use of traps and nooses to catch wild animals. The Hezhe group is skilled at carpentry, tanning and iron smelting. The women in the Daur group are renowned for their needlework, decorating
clothing with fine patterns. Perhaps the most famous example is the Mongol
group, whose former leader Genghis Khan established one of the largest empires
in the history. The Mongols are said to be very skilled in horse riding. They
can travel long distances carrying few supplies. An example of their skills is that
they are adept at drinking horse blood if water is in short supply. These stories
may explain the historic domination by certain ethnic groups of specific trades
(Churchill 2017).

4 Theoretical discussion

4.1 The impact of ethnic diversity on product quality

Admittedly, different ethnic groups may have different preferences and per-
spectives, and these may lead to disagreements and differing approaches. An
example is the language instructed at schools. While Hispanic parents might
prefer more resources devoted to bilingual education, Black parents believe that
will divert the resources away from the particular needs of their children. Asian
parents complain that their children do not receive sufficient resources and White
parents do not want to divert the school resources to any non-English activities
(Alesina, Baqir and Easterly 1999). A consequence of this disagreement is that
the level of capital investment is often at the suboptimal level (Mauro 1995;
Montalvo and Reynal-Querol 2005a), which has a negative impact on the prod-
uct quality. The disagreement among different ethnic groups may result in a
low level of trust (Dincer 2011) and social conflict (Montalvo and Reynal-Querol
2005b) which leads to a less desirable outcome of production, i.e. low quality
goods.

Another explanation comes from the capabilities required to produce quality.
According to Lazear (1999), people from different groups have disjoint informa-
tion sets that are possibly relevant to the job. People from different ethnicities,
especially local people, can bring their knowledge and experience to the team.
This is what Lazear (1999) calls “knowing the ropes”. For instance, a company
might want to hire local people because they best understand the local weather
and natural resources. Also, people from a particular ethnic group may possess the skills required for particular tasks, a phenomenon called “best practices” by Lazear (1999). A team composed of people with diverse backgrounds and skills is more likely to have the person best-suited to a job or task than a homogeneous team.

However, to realize gains from diversity, the information from different groups must be relevant and easily learned or transferred. “Knowing the ropes” and “best practices” are more likely to be relevant in homogenous sectors such as agriculture. Western provinces such as Sichuan, where many ethnic groups live, are well-known for their traditional food. In differentiated sectors such as manufacturing, local experience and culture are of less importance. Whether the disjoint information sets can easily be learned or transferred depends on how people communicate. People of different backgrounds and cultures face more difficulty when they converse and discuss. Indeed, according to the linguistic relativity principle, or the Sapir–Whorf hypothesis, speakers of different languages tend to think and behave differently depending on the language they use. A common object is therefore interpreted in different ways across different groups. As ideas become more complex, this divergence may be exacerbated. In my context, heterogeneous goods are more complex than homogeneous goods because they have different varieties and thus more characteristics than the latter. In the next section, I will put this idea into a formal framework.

**Claim 1** In multi-ethnic regions, the average quality of exported goods is lower.

### 4.2 The role of product differentiation

Following Grossman and Maggi (2000), I assume that the quality of a product is a function of the productivity of one highly productive worker (i.e. manager) and one relatively less skilled worker:

\[
q = x^\alpha y^{1-\alpha}
\]  

Without loss of generality, I can have \( x > y \). Admittedly the quality of
a product is a subject in the knowledge economy. As in Garicano and Rossi-Hansberg (2006), the manager will have to solve most of the complex problems. As a result, the productivity of the manager matters relatively more than that of the laborer in the quality function. Put another way, I assume that $\alpha > 0.5$.

This parameter $\alpha$ indicates the inverse degree of product differentiation. When $\alpha = 1$, the quality entirely depends on the productivity of the manager. The productivity of the worker is irrelevant, as in Lucas (1978). As a result, there is little variation in the quality, which is the definition of homogeneous goods. By contrast, when $\alpha$ is low the productivity of the worker is more relevant to the quality. The result is a wide range in quality whose level depends on the productivity of the workers. This is the feature of differentiated products.

How do I express diversity in the workforce? As in Grossman and Maggi (2000) in a mean preserving distribution, I represent it by the variance: high variance means more division. In this context (when there are only two people in the production team), this variance is proxied by the productivity of the manager $x$: the higher the $x$, the higher the variance. In other words, the value of $x$ represents the level of diversity.

**Proposition 1** High ethnic diversity is associated with high product quality in homogeneous goods and with low product quality in differentiated goods.

In my framework with only two people working together, a mean preserving distribution means the sum of their productivity is fixed, i.e., $x + y = M$ where $M$ is a constant. With this constraint, we can see that to maximize the quality in Equation (1) we have to choose $x = \alpha M$. For illustration purposes, I can take two extreme examples. When $\alpha = 0.5$ (i.e. the highest level of differentiation), the product quality is highest when $x = y = M/2$, which takes place in the least ethnically diverse society. When $\alpha = 1$ (the product is homogeneous), its quality is highest when $x = M$ and $y = 0$ which takes place in the most ethnically diverse society.
Table 1. Summary Statistics

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Export value (millions $USD)</td>
<td>0.095</td>
<td>1.5</td>
<td>0</td>
<td>565</td>
</tr>
<tr>
<td>Quantity (millions)</td>
<td>0.19</td>
<td>23.65</td>
<td>0</td>
<td>9,517</td>
</tr>
<tr>
<td>Population (millions)</td>
<td>50.93</td>
<td>27.46</td>
<td>2.616</td>
<td>91.24</td>
</tr>
<tr>
<td>(log) Total factor productivity</td>
<td>-0.0356</td>
<td>0.729</td>
<td>-7.146</td>
<td>4.428</td>
</tr>
</tbody>
</table>

Note: the minimum value of export value is 1$USD and the minimum value of export quantity is one unit.

5 Data and measurement

5.1 China customs data

In this project, I employ customs data on the universe of exporting firms in China in 2000, information provided by the China Customs Office. It records all types of trade, including processing trade, exchanges between international organizations, required materials and machines in an overseas contract, etc. However, as Dai, Maitra and Yu (2016) suggested, it is crucial to separate the processing trade from the other exporters in China. Indeed, they documented that, unlike other countries processing trade exporters, the Chinese processing trade firms are less productive and create less value added per worker than other industries. For this reason, I eliminate processing trade from my study. Moreover, I drop other forms of trade, such as gifts and exchanges, which are less likely to be decided by the production source. At the end, I put my focus on general trade, which accounted for 55 percent of the total export from China in 2000.

In order to limit my study to manufacturing firms, and also to incorporate the enterprises’ characteristics into my project, I merge this dataset with data from an industrial survey of manufacturing firms in China conducted by the National Bureau of Statistics. This survey covers all enterprises with annual revenue greater than CNY 5 million (or equivalently USD 800,000). This merged data accounts for 31 percent of the total export in 2000 and 10 percent of the companies in the industrial data. Table 1 reports the key variables that will be used to estimate the product quality in the paper.
5.2 National Population Census in 2000

I will use the population distribution, which is calculated from the National Population Census in China in 2000. This census has been carried out every decade since 1953. In this census, people have to declare their ethnic origin (minzu). The formation of minzu formally began in 1953 when the Communist Party of China (CPC) constructed a census to determine the national identity after the birth of the People’s Republic of China. More than 400 ethnicities were recognized. This was not a surprise given that the government drafted policies to award benefits to minorities, such as a seat in the National Peoples Congress for every ethnic group. In 1954, the government’s Yunnan Ethnic Classification Team was tasked with settling the question of how many truly distinct ethnic groups China contained. Based on the assumption that there is a strong correlation between language and ethnicity (Alesina et al. 2003), the team reduced the number of ethnicities to 56, which is still used today.

Ethnic groups can be defined or measured in a number of ways. The traditional measure is the ethnic fractionalization index. This index indicates the probability that two randomly selected individuals in the same region belong to different ethnic groups. In theory, the index reaches its maximum value of 1 when each person belongs to a different group. Alternatively, one can measure the polarization index, as proposed by Reynal-Querol (1998). This index has the maximum value when the population has two groups of equal size.

I employ the ethnic fractionalization index in the benchmark and then the polarization index in Section 6.4 to check the robustness of my results. The fractionalization index is calculated as follows:

$$ Div_p = 1 - \sum_k n_{pk}^2 $$

where $n_{pk}$ is the population share of group $k$ in province $p$. High $Div_p$ index indicates that province $p$ is ethnically diverse. This variable takes the value 1 when the province is completely heterogeneous and 0 when the province is
Table 2. Ethnic Diversity

<table>
<thead>
<tr>
<th>Province</th>
<th>ELF</th>
<th>PI</th>
<th>Province</th>
<th>ELF</th>
<th>PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jiangxi</td>
<td>0.00621</td>
<td>0.01434</td>
<td>Heilongjiang</td>
<td>0.09447</td>
<td>0.18254</td>
</tr>
<tr>
<td>Shanxi</td>
<td>0.00633</td>
<td>0.01269</td>
<td>Sichuan</td>
<td>0.09661</td>
<td>0.18814</td>
</tr>
<tr>
<td>Jiangsu</td>
<td>0.00710</td>
<td>0.01428</td>
<td>Tibet</td>
<td>0.13568</td>
<td>0.26943</td>
</tr>
<tr>
<td>Shaanxi</td>
<td>0.00993</td>
<td>0.01986</td>
<td>Gansu</td>
<td>0.16486</td>
<td>0.32206</td>
</tr>
<tr>
<td>Shanghai</td>
<td>0.01260</td>
<td>0.02522</td>
<td>Jilin</td>
<td>0.17143</td>
<td>0.31489</td>
</tr>
<tr>
<td>Anhui</td>
<td>0.01340</td>
<td>0.02680</td>
<td>Hunan</td>
<td>0.18939</td>
<td>0.33890</td>
</tr>
<tr>
<td>Shandong</td>
<td>0.01398</td>
<td>0.02799</td>
<td>Liaoning</td>
<td>0.27855</td>
<td>0.51502</td>
</tr>
<tr>
<td>Zhejiang</td>
<td>0.01713</td>
<td>0.03782</td>
<td>Hainan</td>
<td>0.29322</td>
<td>0.55650</td>
</tr>
<tr>
<td>Henan</td>
<td>0.02480</td>
<td>0.04939</td>
<td>Inner Mongolia</td>
<td>0.34329</td>
<td>0.62663</td>
</tr>
<tr>
<td>Guangdong</td>
<td>0.02950</td>
<td>0.05865</td>
<td>Ningxia</td>
<td>0.45643</td>
<td>0.89708</td>
</tr>
<tr>
<td>Fujian</td>
<td>0.03393</td>
<td>0.07733</td>
<td>Guangxi</td>
<td>0.51400</td>
<td>0.87828</td>
</tr>
<tr>
<td>Tianjin</td>
<td>0.05313</td>
<td>0.10413</td>
<td>Yunnan</td>
<td>0.53971</td>
<td>0.70491</td>
</tr>
<tr>
<td>Beijing</td>
<td>0.08378</td>
<td>0.16103</td>
<td>Guizhou</td>
<td>0.58795</td>
<td>0.72694</td>
</tr>
<tr>
<td>Hubei</td>
<td>0.08401</td>
<td>0.16496</td>
<td>Xinjiang</td>
<td>0.62428</td>
<td>0.88224</td>
</tr>
<tr>
<td>Hebei</td>
<td>0.08408</td>
<td>0.16358</td>
<td>Qinghai</td>
<td>0.63254</td>
<td>0.83549</td>
</tr>
</tbody>
</table>

Easterly and Levine (1997) used this index based on the ethno-linguistic classification that was constructed by the Soviet Union research team, printed in the *Atlas Norodov Mira* (Atlas of the Peoples of the World). Language, however, is just one of the many aspects of ethnicity. In some cases, different ethnic groups share a common language. To more fully capture ethnic differences, I then construct the fractionalization index based on ethnicity, which is taken from the Chinese National Population Census as discussed above. This index shows more fractionalization relative to the ethno-linguistic index (Alesina et al. 2003).

5.3 The quality of exports from China

I follow the suggestion by Berry (1994) that quality can be estimated as the excess sales after controlling for price, an idea that has been used widely (for instance, Hallak and Schott 2011; Khandelwal 2010). In particular I estimate the following demand function:

\[
q_{ijc} = \alpha + \beta p_{ijc} + \gamma pop_c + I_j + u_{ijc} \tag{2}
\]

In the above equation, the sales of a product \( j \) by company \( i \) to country \( c \)
depends on its price $p_{ijc}$, the market size (controlled by the country population $pop_c$) and the price index (controlled by the industry fixed effect $I_j$) which represents the business condition. The quality of the product is not observable and is treated as the error term.

A problem with this estimation is the endogeneity of the unit price $p_{ijc}$. Indeed, unit prices are often positively correlated with unobserved quality components, creating an upward bias. To correct for this problem, I need to determine causality with an instrumental variable (IV). Khandelwal (2010) suggests transportation costs should be included in the IV; but unfortunately, they are not available in China. I must then use two dummy variables: the country of destination dummy and another dummy that indicates whether the province where the firm is located has a major port. Intuitively, these two instruments capture how remote the firm is (indicated by the availability of a major port in the province where the firm is located) and how close to China the export market is. In other words, the two instruments capture the transport costs from the firm to the destination.

Before applying the instrumental variable approach, I need to check the relevance and the validity of the instruments. In particular, I need to check: (i) if the instruments are correlated with the endogenous variable (i.e. the unit price); and (ii) if they are uncorrelated with the error terms (the exogeneity condition). To be more precise, the error terms here are the demand shocks. It is clear that the distance between China and the destination country, as well as the availability of a major port in a province, are not related to the demand shocks. And the significance of the instrument coefficients in the first stage\(^1\) indicates the relevance of my instruments.

With these instruments at my disposal, I run the regression (2) for 94 of the total 98 HS two-digit level categories \(^2\). Out of the remaining 94 sectors, 9 have

\(^1\)Because I run the quality of estimation for each HS-2 digit sector, I cannot report all the first stage results here. These results are available upon request.

\(^2\)There are 4 sectors that have no observations or fewer than 10 observations. These sectors are: Live Animals; Pulp of Wood, Waste and Scrap of Paper; Aircraft, Spacecraft and Parts Thereof; Business Services, Health, Financial/Insurance Legal/Real Estate, Hotels, and Misc repair Business Services.
Table 3. Own Price Elasticity

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>First quartile</th>
<th>Third quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without IV</td>
<td>-0.74</td>
<td>-0.71</td>
<td>-0.87</td>
<td>-0.49</td>
</tr>
<tr>
<td>With IV</td>
<td>-1.04</td>
<td>-0.94</td>
<td>-1.30</td>
<td>-0.48</td>
</tr>
<tr>
<td>Khandelwal estimates</td>
<td>-1.28</td>
<td>-0.58</td>
<td>-1.44</td>
<td>-0.20</td>
</tr>
</tbody>
</table>

Note: My estimates are taken from Equation (2). To be consistent with Khandelwal (2010), the statistics are calculated conditional on negative own price elasticity positive own-price elasticity\(^3\).

I then only consider categories with negative own-price elasticity \(\beta\). In order to confirm my quality estimation, I compared my own-price elasticity statistics with other studies, in particular Khandelwal (2010) using U.S. data. Table 3 shows that my statistics do not differ significantly from the findings in Khandelwal (2010).

6 Empirical evidence

6.1 Empirical specifications

To investigate the impact of ethnic diversity on the quality of exports, I run the following reduced form estimation:

\[
Quality_{ijpc} = \text{Intercept} + \beta \text{Div}_p + \varsigma X + \gamma I + \delta_{ijpc} (3)
\]

\(Quality_{ijpc}\) is the estimated quality (in log terms) of a product \(j\), produced by the firm \(i\) that is located in the province \(p\) and exported to the country \(c\). The vector \(X = (X_i, X_p)\) controls the time-invariant characteristics. In particular, \(X_i\) is the vector of the firm’s characteristics such as its age, ownership, input expenses and productivity. I also control for the provincial factors by the vector \(X_p\) such as economic growth (GDP per capita), geography (coastal dummy), transportation, investment and urbanization (city population). The vector \(I = (I_s, I_c)\) controls

\(^3\)These sectors are: Sugars and Sugar Confectionery; Cocoa and Cocoa preparations; Photographic or Cinematographic Goods; Cork and Articles of Cork; Silk, Inc. Yarns and Woven Fabrics Thereof; Carpets and Other Textile Floor Coverings; Zinc and Articles Thereof; Tin and Articles Thereof; Ships, Boats, and Floating Structures.
for any sectoral effect ($I_s$) or destination effect ($I_c$, also known as the Alchian-Allen effect see Bastos and Silva 2010). The coefficient of interest is the coefficient of diversity (in log terms) $\beta$. As in the growth literature, I expect diversity to have a negative impact ($\beta < 0$) on the quality of exported goods.

In order to better gauge how ethnic diversity impacts quality, I investigate how the impact changes with the product characteristics. In particular, I include the interaction between the diversity index and the degree of product differentiation:

$$Quality_{ijpc} = \text{Intercept} + \phi Div_p + \theta Div_p * Diff_j + \varsigma X_i + \nu X_p + \delta_{ijpc} \quad (4)$$

Here I borrow the idea of a quality ladder as a means of product differentiation from Khandelwal (2010). In particular, I calculate the quality dispersion as a proxy of product differentiation. The marginal effect of diversity is then written as:

$$\frac{\delta Quality_{ijpc}}{\delta Div_p} = \phi + \theta Diff_j \quad (5)$$

Equation 5 implies that for differentiated products (when $Diff_j$ is high enough), the impact of diversity on product quality is the sign of the coefficient $\theta$. My conjecture is that $\theta$ is significantly negative. Moreover, the impact of diversity on homogeneous goods (when $Diff_j$ is close to 0) is the sign of the coefficient $\phi$. If $\phi$ is significantly positive, then ethnic diversity can have a positive impact on the quality of homogeneous goods.

6.2 Product quality is generally lower in multi-ethnic regions

Estimates of Regression 3 are reported in Columns 1 to 3 of Table 4. In Column 1, only the firm’s characteristics are controlled. In Column 2, I control for the firm’s and province’s characteristics. And in Column 3, I add the destination fixed effect to control for all destination-related characteristics. In all columns
the coefficient of interest $\beta$ is significantly negative, confirming my prediction in Section 4.1 that in provinces that are ethnically diverse, the quality of the exported products is generally lower. This result is consistent with the findings in the growth literature that ethnic diversity is, in general a hindrance to economic performance (Alesina and Ferrara 2005).

Figure 1 provides further supporting evidence to Claim 1. In this figure, I plot the kernel density estimates of the quality distribution in two provinces: Qinghai and Jiangxi. Recall that Qinghai has the highest ethnic diversity index and Jiangxi has the lowest ethnic diversity index. In other words, they represent the most- and the least- ethnically diverse provinces in China. Figure 1 shows that on average, there were more exported quality products from Jiangxi than from Qinghai. However, it is worth noting that all the exported products across different sectors were put together. A more precise comparison will be done within a sector, which is the subject of the next section.

6.3 The heterogeneous impact of ethnic diversity

My theory in Section 4.2 predicts that the impacts of ethnic diversity depend on the nature of the products. This section will lend evidence to support this theory.

Column 1 in Table 5 reveals some interesting results. The interaction term of
Diversity and Product Differentiation (coefficient $\theta$ in Equation 4) is significantly negative. It shows that the negative impact of diversity increases with the degree of product differentiation. However, the coefficient $\phi$ is insignificant. It prompts the result that ethnic diversity could have a positive impact on the quality of homogeneous goods. To verify this, I limit my sample data to products with a quality dispersion lower than 4.18 (10th percentile) and rerun the empirical exercise as in Equation 3. The impact of ethnic diversity (the coefficient $\beta$) is now significantly positive.

To provide further supporting evidence, I use two provinces that were discussed previously, namely Qinghai and Jiangxi. I look at the product quality in two sectors. The first one is an agricultural sector (Harmonized System code: 12). It encompasses the following products: oil seeds and oleaginous fruits; miscellaneous grains, seed and fruit; industrial or medicinal plants; straw and fodder. That sector is a representative case drawn from the homogeneous sectors.
Table 5. The impact of Diversity across products

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Div</td>
<td>-0.129</td>
<td>0.904***</td>
<td>0.495***</td>
<td>-0.00851</td>
<td>0.0136</td>
<td>0.603**</td>
</tr>
<tr>
<td></td>
<td>(0.316)</td>
<td>(0.311)</td>
<td>(0.186)</td>
<td>(0.130)</td>
<td>(0.122)</td>
<td>(0.238)</td>
</tr>
<tr>
<td>Div*Diff</td>
<td>-0.180***</td>
<td>-0.200***</td>
<td>-0.148***</td>
<td>-0.0972***</td>
<td>-0.0562***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0509)</td>
<td>(0.0502)</td>
<td>(0.0299)</td>
<td>(0.0340)</td>
<td>(0.0172)</td>
<td></td>
</tr>
<tr>
<td>Diff</td>
<td>0.124***</td>
<td>0.127***</td>
<td>-0.0288***</td>
<td>0.00897***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00479)</td>
<td>(0.00493)</td>
<td>(0.00310)</td>
<td>(0.00125)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Div*WorkSubs</td>
<td></td>
<td></td>
<td></td>
<td>-1.911***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.308)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WorkSubs</td>
<td></td>
<td></td>
<td></td>
<td>-0.958***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0324)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>147.245</td>
<td>147.245</td>
<td>147.245</td>
<td>147.245</td>
<td>146.327</td>
<td>147.245</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.014</td>
<td>0.014</td>
<td>0.013</td>
<td>0.02</td>
<td>0.02</td>
<td>0.01</td>
</tr>
</tbody>
</table>

NOTE: In all specifications, I include the firm’s characteristics such as the age, the firm’s status (foreign invested, state-owned) input expenses and productivity. The province characteristics such as GDP per capita, the amount of transported goods per kilometers, the number of investment projects, are also included. Div is the ethnic diversity index. In all columns except Column 3, I use Div as the ethno-linguistic fractionalization index whereas in Column 3 I use the polarization index. Diff measures the Product Differentiation. WorkSubs is the degree of substitutability among workers. Standard errors in parentheses *** p < 0.01, ** p < 0.05, * p < 0.1

The second sector is a manufacturing sector (Harmonized System code: 87). It encompasses the following products: vehicles other than railway or tramway rolling-stock, and parts and accessories thereof. That sector is a representative case drawn from the differentiated sectors. The mean quality of the first sector in Qinghai and in Jiangxi is 11.1 and 10.5, respectively. The mean quality of the second sector in Qinghai and in Jiangxi is 10.4 and 13.2, respectively. These results, again, support the conclusion that the quality of the homogeneous sector (the first sector) is higher in the more ethnically diverse region (Qinghai), while the quality of the differentiated sector (the second sector) is higher in the less ethnically diverse region (Jiangxi).

6.4 Robustness check

In the previous section, my regression results suggest that the negative impact of diversity on quality is robust to various specifications. In this section, I will check for robustness with other methods. First, in the benchmark regression, I apply the random effects estimator. While this estimator is more efficient, there is a worry about its inconsistency. Being aware of this concern, I will cross-reference with the fixed effects model. Column 4 in Table 4 shows that my results are still robust with this estimator, although the coefficient is slightly
smaller in absolute terms. Indeed, provinces that are ethnically homogeneous export products that are 5 percent higher in quality than multi-ethnic provinces. I can also use the fixed effects estimator when I check the impact of diversity across different products. Results reported in Column 2 in Table 5 confirm Claim 2 that ethnic diversity has a positive impact on homogeneous products, but its impact becomes negative with differentiated products.

Another concern is the measure of my independent variables, in particular the diversity index. In addition to the group fragmentation as the Div index provides us, I can also look at the polarization of the group. I then borrow the polarization index (PI) suggested by Reynal-Querol (1998), calculated as:

\[ PI_t = 1 - \sum_j \left( \frac{0.5 - n_{ij}}{0.5} \right)^2 n_{ij} \]

This index measures how polarized the group is. In other words, PI reaches its maximum value when there are two or more ethnic groups of equal size. The corresponding values of this index across provinces are shown in Table 2. Column 5 in Table 4 and Column 3 in Table 5 suggest that my results are robust with this measure of diversity\(^4\).

Finally, I want to test whether my results are robust to a different measure of differentiation. Instead of the quality dispersion, I use two alternative measures, the price dispersion and the elasticity of substitution taken from Broda, Greenfield and Weinstein (2006). While the dispersion of quality (and price) represent the vertical differentiation, the elasticity of substitution represents the horizontal differentiation. Columns 4 and 5 in Table 5 confirm that my results survive this test.

\(^4\)It is possible that migration influences my measures of diversity. But as Dincer and Wang (2011) reported, the index does not change significantly over the period of 1978 to 2002. This guarantees that the index is exogenous. Because this is a cross-sectional data, I can rule out the impact of migration: in any case, this is a snapshot of the impact of the distribution of ethnic diversity on quality of exports.
7 Further discussion

In addition to the product differentiation, the problem of ethnic diversity is also aggravated by the workers’ complementarity. This complementarity among workers requires all of them to perform their task well, which is more difficult when they cannot communicate efficiently. Another point I can make here is that people from different groups have less sympathy towards each other than do those within the same ethnic group. Again, if the workers are substitutable, this causes few problems to the team. But if they are complementary, the disharmony problem becomes more serious.

I then can check whether this hypothesis is correct, that when the workers are complementary, ethnic heterogeneity negatively affects the quality of products produced by the firm. I measure the degree of substitutability among workers by the wage dispersion across industries: the lower the wage dispersion the more substitutable the workers are or the less complementary the workers are. I then run the following regression:

\[ \text{Quality}_{ijpc} = \alpha + \beta \text{Div}_p + \delta \text{Div}_p \ast \text{Comp}_j + \zeta X_i + \nu X_p + \delta_{ijpc} \]  

(6)

Column 5 in Table 5 shows that the interaction term is negative, confirming my hypothesis. This result is consistent with Bombardini, Gallipoli and Pupato (2012), when they show that countries with a dispersed skill distribution specialize in products with reduced worker skill complementarity. Moreover, my data shows that wage dispersion is positively correlated with my two measures of differentiation, namely quality dispersion and price dispersion. Indeed, the correlations are 0.07 and 0.09 respectively. These results then explain the heterogeneous impact of ethnic diversity on quality as I find in the previous section.
8 Conclusion

Ethnic diversity is claimed to have a significant impact on economic growth. In this study, I investigate the impact of ethnic diversity on another dimension: the quality of products. I use customs data and the manufacturing survey in China to estimate the quality of exported goods from China in 2000. I conclude that products from a completely homogeneous province are more than 10 percent higher in quality than those from a completely heterogeneous province. While the impact of ethnic diversity is negative for differentiated sectors, it is positive for homogeneous sectors. This result allows us to propose a channel by which diversity influences quality. Indeed, workers in differentiated sectors are complementary, which means they need to work in tandem and communication is very important. That explains why my data shows that diverse provinces where people might have difficulty communicating do not produce differentiated goods of high quality. However, in homogeneous goods where experience and knowledge from ancestors can be relevant, diverse provinces can have an advantage in producing high-quality goods.

These results prompt a clear policy recommendation. Ethnic diversity should be a factor in designing industrial development policy. More precisely, homogeneous sectors should be promoted in regions with a high level of ethnic diversity, whereas differentiated sectors should be developed in regions with a low level of ethnic diversity. My paper contributes, therefore, to the understanding of the impact of diversity. It is exciting to follow this road, as others have shown that diversity can be a new source of comparative advantage.

References


50.


