RELATIVE WAGE VARIATION AND INDUSTRY LOCATION WITHIN DISTRICTS OF PUNJAB

Zunia Saif Tirmazee

Research Fellow
Lahore School of Economics

This dissertation is submitted for the degree of
M.Phil Economics
2012

Supervised by:
Dr. Theresa Chaudhry

ABSTRACT

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

INTRODUCTION

The aim of this paper primarily is to see whether departures from relative factor price equality leads to differences in the mix of industries that regions produce. Bernard et al (2002) suggests four reasons as to why, even with trade, relative factor prices should vary across regions, namely: i) multiple cones of diversification, ii) region-industry technology differences, iii) agglomeration and iv) increasing returns to scale. However this paper specifically deals with cones of diversification as an explanation of relative factor price equality. Regions characterized by relative factor price equalization have greater number of industries that they produce in common, or to be more precise, this paper aims to test the proposition that that larger the differences in the relative wages of two regions the smaller is the number of industries that they produce in common. This geographic variation in relative wages of workers and therefore the industrial production finds its justification in the basic predictions of the neoclassical theory of trade which suggests that regions with an abundance of a factor have a higher concentration of industries that use that factor intensively than regions with a scarcity of that factor.

Testing for factor price equalization is problematic because one needs to account for unobserved variations in factor quality especially in case the of workforce quality. Workers' quality across regions could vary due to differences in the education system or availability of training programs for workers. These differences in turn affect the productivity of workers and hence the relative wages of skilled workers. My thesis will follow the methodology used by Bernard, Redding Schott and Simpson (2008) to look at geographic variation in relative wages and industry mix but for Punjab over the period 2000-01 and 2005–2006. The key advantage of this methodology is its ability to control for unobserved variations in factor quality.

Such an analysis is important because firstly it yields information about income convergence within a country. Secondly it has public policy implications as well, as suggested by Bernard et al (2008). It points to the fact that policies targeted at regional development need to take into account the information regarding factor prices and hence the regional comparative advantage. Because relocating skill intensive industries to lagging regions may result in comparative cost disadvantage for such firms as the neoclassical theory of trade would predict. However other policies that concern boosting the skill level of workforce in the lagging regions through education and worker training are better alternatives promising success in the long run.

LITERATURE REVIEW

What determines industry location in a particular area? This question perhaps has no single answer. When it comes to finding out the determinants of industry location the answer can be best summarized in three schools of thought as have been laid out in Brulhart(2001). Firstly the *neo-classical trade models* deal entirely with factor proportions, factor endowments and technology of individual regions. Thus, whether industries locate at a particular place depends on their comparative advantage. The second school of thought relates to the models in the*New Trade Theory*. These models maintain that industry structure is determined firstly due to “first nature”(Krugman, 1993) characteristics of regions. These includes features like land characteristics and differences in accessibility. Also these models look at the “second nature” characteristics. These mainly refer to the Marshallian externalities which include intellectual spillovers, labor market pooling and forward(large input markets) and backward linkages(large product markets) with endogenous market size effects. The third school of thought is based on models on the *New Economic Geography* models that take into account only the second nature characteristics.

This paper however focuses on all those determinants that relate to labor market. In the case of the neoclassical theory of trade we are concerned with determinants like relative factor endowments or factor intensity of production. And in case of models in New Economic Geography and New Trade Theory we specifically talk about factor endowments (first nature characteristics) and labor market pooling(second nature characteristics).

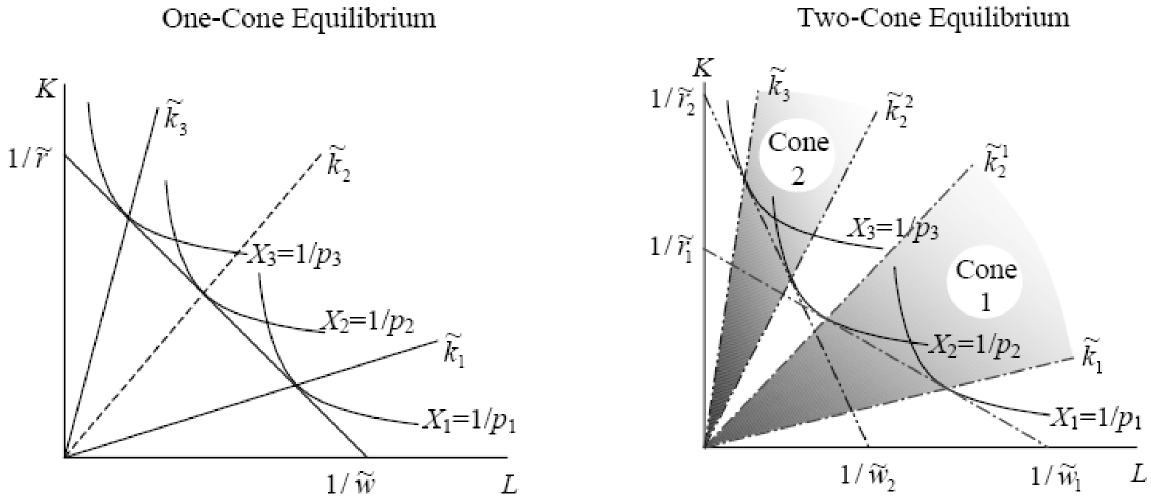
One of the most important results from the neoclassical theory of international trade is the factor price equalization theorem (FPE). The FPE theorem suggests that all regions producing a similar product mix with similar technologies and same product prices will have the same factor prices, if certain assumptions hold as have been laid down in Lerner (1952). These stylized assumptions are that, firstly, regions are to have access to identical technologies (which is just another way of saying that they have the same production functions). Secondly, they produce the same products and they do not fully specialize in any products (factor intensities are the same). Thirdly it assumes constant returns to scale. Another assumption is that of perfect competition[1]. Lastly the ranking of regions according to their factor intensity remains the same at all factor prices (or more technically there no factor intensity reversals).

Beyond these given conditions, the role of factor endowments must also not be overlooked. Samuelson (1949) notes that apart from similar factor intensities and commodity prices regions have to have to have similar factor endowments as well to qualify for factor price equalization. For factor endowments dissimilar enough, it is impossible for regions to produce same commodities and therefore unlikely that factor price equalization will hold.

The Heckscher Ohlin model (HO model) purports a relationship between regional production structure and relative factor prices. This idea has been explained in the literature on trade through the framework of **cones of diversification** where the word 'cone' refers to the range of endowment vectors that select the same set of industries. A notable name in this respect is of Alan V. Deardorff. Deardorff (2002) uses the standard Lerner (1952) diagram to explain the interaction between diversification and factor prices under varying factor endowments. The whole idea is explained on an x-y plane where each axis represents units of factor employed for production. Deardorff (2002) shows a transition from complete factor price equality to factor price inequality through two diagrams, one showing a one cone equilibrium (the case for relative factor price equality) and the other for a two cone equilibrium (the case for relative factor price inequality).

[1] Kemp and Okawa (1997) show that, there might be factor price equalization under certain conditions even in the presence of imperfect competition.

Figure 1: Cones of Diversification.

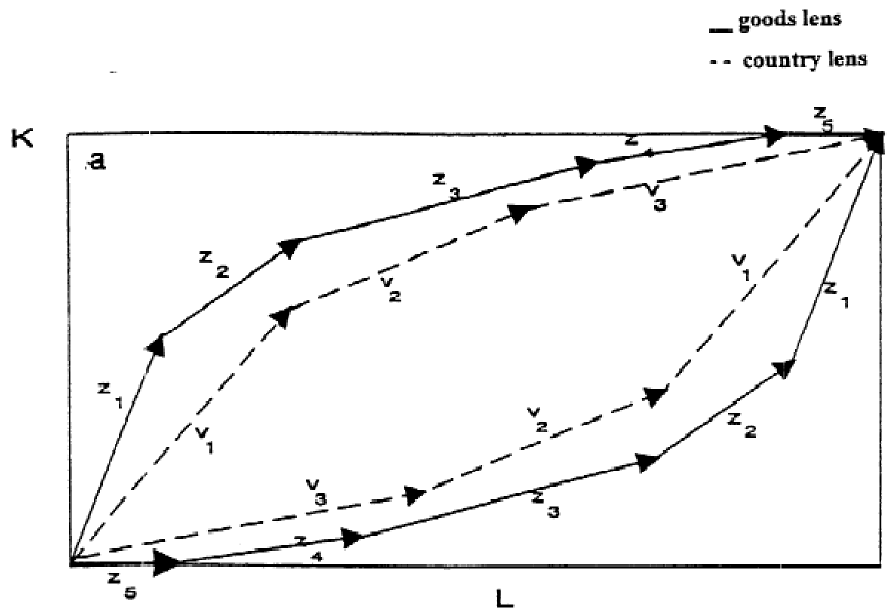


Source: Deardorff (2002)

The panel for one cone equilibrium shows a world where production technologies for each good are such that they all align perfectly with a single isocost line (same relative factor prices throughout). So given their factor endowments all the regions cling to a unique relative factor price at their respective factor intensity ratios. So the entire mix of goods lies within a single cone of diversification. However the panel for the two cone equilibrium shows the case where the production technique of good 2 is not identical to the other two goods. There could be any reason for it, but the reason that Deardorff puts up is the higher price of good 2 now as compared to under complete RFPE. Under such a circumstance, there does not exist a unique relative factor price that leads to all the goods to be produced optimally. Thus under such a case then there emerge two relative factor prices that lead to optimal (cost minimizing) production of good 1 and good 2 or good 2 and good 3 separately. Now all of the goods cannot be produced at a single factor price. This shows that there are now two cones of diversification each defined by a range of factor intensity ratios and their own unique relative factor prices. Within each cone there is factor price equality, but this equality does not hold as we move to the other cone. Thus the idea that the number of common industries that regions produce increases with the fall in relative wage differential finds its explanation in the Lerner diagram.

Empirical analyses have employed various ways to test for factor price equalization. Deardorff (1994) has developed a condition called the "lens condition" to test for factor price equalization. This condition basically requires the variation in factor intensities across regions to be greater than cross-region variations in factor endowments. The lens condition can be visualized on a box diagram whose dimensions represent the factor endowments of a country and the top right and bottom left are country origins (see figure 1). Starting from both the country origins, if one keeps placing endowment vectors for factors head to head in increasing order for various regions there emerges a polyhedral shape which is called the 'country lens'(dashed lines in figure 1 below). Similarly, placing the sectoral factor use vectors head to head in increasing order from both the regions gives another polyhedaral figure called the 'goods lens' (solid lines in the figure below). Now for the lens condition to hold goods lens must lie outside the country lens, and only then will there be factor price equalization. Because then this suggests that endowments of various regions are similar enough for goods produced to be the same. The lens condition has been used by researchers as a test for proving FPE. Notable among these are Bernard et al (2010), Demiroglu (1999) and Debeare and Demiroglu (2003).

Figure 2: Lens Condition satisfied.



Source: Deardorff (1994)

The cointegration approach is another way to check for FPE. Pioneers of this approach are Burgman and Geppert (1993) who have found proof of FPE in long run stationary cointegrating relationship between stationary series of unit labor costs of six major industrialized countries Canada, France, Germany, Japan, Great Britain, and the U.S. Later using the same data set that Burgman had used, Berger and Westerman (2001) after controlling for the sample bias and using real data instead of nominal figures for unit labor cost, find only a limited evidence of factor price co-movement.

Originally, the tests of factor price equalization looked at cross-country evidence. Another line of research that has more recently emerged is that of Bernard and Schott who have tested for relative factor price equalization (RFPE) *within* countries, for the UK (2008), United States (2005) and Mexico (2010). The procedure followed is fairly simple and involves regressing a ratio of wage bills of skilled and unskilled workers for each industry in a region on a set of regional dummies thus controlling for unobserved variations in factor quality and production technology. The coefficient of regional dummies then decides if RFPE prevails or not. If the coefficient is significantly different from zero it leads one to rejecting the null hypothesis of Relative Factor Price equality. All of the studies cited above have shown that relative factor price equalization does not hold.

Factor price equalization has implications for production patterns, related to the predictions of the neoclassical model of trade. The neoclassical model of trade predicts that departures from relative factor price equality are associated with differences in the set of industries that regions produce. All those regions that differ with respect to their factor prices differ also in terms of industries (goods) that they produce. A more technical way of saying that would be in words of Leamer (1995), "While factor prices might differ between countries that find themselves within different *cones of diversification*, they should be the same when factor endowments are such that all countries select the same range of goods to produce."

The literature review in this respect is thus done to answer mainly the following two questions:

1. What are the reasons or the theoretical underpinnings that might support the phenomenon that regional similarity in the patterns of production is determined by the relative factor prices?

2. Secondly, if this is the case then what are the possible forces that allow this phenomenon to sustain i.e., why are industries able to benefit from the relative factor price differential (in this case relative wage differential)) as in why would this differential not be arbitraged away?

As far as the first question is concerned the simplest answer would be the predictions of the neoclassical model of trade that "regions with low relative wages of white-collar workers have a higher concentration of industries that use white-collar workers intensively than regions with higher relative wages of white-collar workers"(Bernard et al, 2008). Simply put, factor prices vary across regions due to differences in the factor endowments of each region. Regions will have a lower relative factor price of a factor that they have an abundance of, as compared to other factors. The endowment and therefore the relative price of a factor then determine the pattern of production across regions.

There is a dearth of literature on direct evidence of impact of relative wage differential on the pattern of production. Mostly the literature finds a relationship between input abundance and input intensity, which is just another way of addressing the research question at hand,since similarity in input intensity across regions is also suggestive of industries being produced in common across regions.

The most direct evidence is provided by Bernard et al. (2004), which attempts to explain the same phenomenon for a two factor (skilled and unskilled workers) and three goods (computer, machinery and textiles) economy on the Lerner diagram. The study shows that large differences in factor endowments of skilled and unskilled labor lead regions to fall in separate cones of diversification and that as one looks across the two cones of diversification (one for relatively skilled labor intensive goods and the other for relatively unskilled labor intensive goods), there is a decrease in the relative wages of unskilled labor (or an increase in the relative wages of skilled labor) which is in line with the predictions of Hecksher-Ohlin model. The empirical analysis for studying the effect of relative wage differentials of skilled and unskilled workers on production patterns has been done using data from 1970s to 1990s for the UK for 209 four-digit manufacturing industries across 63 counties and Scottish Regions. The analysis itself is a simple OLS regression where the number of industries common in two regions is regressed on the estimates of relative factor price equality (these in turn were the coefficient estimates of a linear regression of relative wage bill estimates of skilled and unskilled workers on the regional dummies). Overall the study provides strong evidence in favor of the hypothesis that the regional production structure of UK is in line with a multiple cone general equilibrium trade model. The same kind of analysis has also been done for United States in Bernard et al. (2004) and Mexico in Bernard et al. (2010) both studies confirming the multiple cones general equilibrium hypothesis.

Hanson and Slaughter (2002) have done an empirical analysis to show that factor price equalization across states is consistent with similar production techniques (factor intensities). Using data for unit factor requirements for 14 large US states and 40 sectors, spanning all industries, for 1980 and 1990 they find that production techniques are very similar across large US states, especially for neighboring states or states with similar relative labor supplies. The analysis conducted consists of checking if the unit factor requirements are similar across states. However, regions can have factor price equality yet have two different production techniques due to scale effects, externalities, or differences in underlying technology.

Therefore they test for productivity-adjusted FPE by for each state-pair, labor-type, and year combination through fixed effects. A joint test of significance of all the fixed effects coefficients is equivalent to testing the null hypothesis that there are no factor-specific, industry-neutral productivity differences across states. The authors believe that since "industry-neutral state productivity differences allow for nominal state wage differences," that is why they look for evidence of FPE using data on state industry production techniques instead of direct data on state wages. Therefore their implicit null hypothesis is that productivity adjusted wages are equalized across sample states.

Similarly Demiroglu (2003) looks for variation in factor endowments (labor and capital) within a country and its effect upon diversification in production. Using data for 28 manufacturing sectors and 28 countries they construct a lens of factor endowments and factor intensities and are able to conclude that developed and developing countries do not lie in the same cone of diversification and hence do not have factor prices equalized (as most developing countries are bunched near the corner of the lens which leads to violation of the lens condition) whereas most OECD countries do. On a lens, all the countries lie in the same cone of diversification (factor price equalization) whose endowment vectors lie within the lens of factor intensities. The authors also check for the robustness of their results by adjusting the data for productivity differences. However even after correcting for the productivity differences developing countries continue to violate the lens condition thus confirming the results that developed and developing countries do not lie in the same cone of diversification. Moreover they perform the analysis using data for skilled and unskilled labor for some selected OECD countries and find similar results that most OECD countries lie within the same cone.

Reeve(2002) demonstrates the importance of factor endowments in explaining the structure of production independent of industry idiosyncrasies and quantifies to which extent factor accumulation determine shifts in industrial structure in a cross section of 15 largest manufacturing industries of 20 OECD countries for the years 1970-1990. This finding has policy implications in particular concerning investments in physical and human capital. Such accumulation of factors impacts patterns of production in the same way as one would expect in case of sector-specific trade and industrial policies. The author reaches this conclusion empirically by estimating a model of factor proportions. His endowment data includes five factors: capital stock, arable land, high-educated labor, medium-educated labor, and low-educated labor. The relative endowment of factor is measured by taking its share in sample to its share in GDP. Moreover percentage changes in industry output as a share of total manufactures are used to account for the compositional changes in the industrial production. Regressing compositional changes in industrial production on relative factor endowments he interprets the estimated coefficients (estimated through one step GLS and iterated GLS) as comparative advantage of a factor in an industry. Then to see to what extent are changes in the structure of production driven by changes in factor endowments he tests the null hypothesis of constant coefficients of factor endowment over time using Wald tests. For nearly every factor, the null of constant coefficients can be rejected at standard levels of significance. Thus it implies that changes in the techniques of production, as measured by the coefficients of factor endowments over time, are potentially important sources of change in industrial structure. Similarly Harrigan(1995) was the first one to do such an analysis that tests the significance of factor endowments in explaining industrial structure. His study consists of estimating three equations to predict industry output as a function of factor endowments. Two of these specifications use country fixed effects with constant coefficients and thus cannot be used to predict changes in industrial production over time where the third one allows parameters to vary over time following a random walk. He concludes that overall the model does not do a very good job in determining the industry location. Similar work is done by Harrigan and Zakrajsek (2000) where they test for the predictions of HO model that regions tend to specialize in production that require intensively the factor that this region is abundant in. Their empirical analysis consists of using a panel data technique on data for industrial specialization and factor endowment differences fora broad sample of rich and developing countries for 1970-92. Allowing for technological differences their results clearly show the importance of factorendowments on specialization. Bernstein and Weinstein (2002) using a cross-section of international and regional data demonstrates that factor supplies affect gross output and that this effect is more evident internationally than intranationally.

Moroney and Walker (1996) however through an entirely different methodology havecarried out a regional test for Hecksher-Ohlin model. The hypothesis that they attempt to test is that ordering commodities in the order of their factor intensity ratios is equivalent to ordering regions according to their comparative advantage. All that this hypothesis implies is that varying factor intensities of a region's production is an indication of varying comparative cost advantage of these regions. So industries requiring a relatively lower capital-labor ratio for their manufacturing are concentrated in regions that have a lower endowment of capital. Using industry level data on capital-labor ratios and the location quotients (measure of regional output concentration by industry) for the United States' South (that is, the East South Central, South Atlantic, and West South Central census regions) which is relatively labor abundant, and the non-South which is relatively capital abundant, the authors test the hypothesis by doing a rank-correlation test. Correlation was found between the industry capital-labor ratio and the location quotients and it was seen that after some experimentation with the data they were able to prove that there exists a negative correlation between the two measures confirming the Heckscher Ohlin hypothesis for regions within the US. Overall, they conclude that the South being labor abundant specializes in the production of goods that are labor intensive compared to that of the non-South.

As far as the second question is concerned, most literature provides explanation of the persistent wage differential by commenting on the mobility of workers both perfect and imperfect. Bernard et al. (2010) to provide a reason for the persistent wage differentials quote Hanson (2004), arguing that workers within Mexico do not migrate enough to arbitrage away the persistent wage differential. Bernard et al. (2008) in order to explain the stability of their estimates for relative factor price inequality quote two reasons. Firstly they believe that it is partly due to the imperfect mobility in labor markets. This imperfect labor mobility gains its empirical support in the analysis of Hughes and McCormick (1994). Secondly they believe that even under perfect mobility, relative wage differentials of skilled and unskilled workers seems to persist due to resistance on the part of workers to migrate in of face of factor-region-specific amenities or living costs. So essentially if the costs associated with migration are greater than the expected gains from it then workers would rather not choose to migrate to benefit from the wage differential. A further explanation of the same argument could be in the words of Deardorff (1993) where he emphasizes the importance of “something else” other than the wage differential itself that affects the decisions of workers as to decide where to live. Elaborating, he quotes “prices of non-traded goods and utility relevant facilities” as being relevant factors that lead mobile labor to prefer one region over another. Analysis on similar lines has been done by Kerr (1954). In his essay he summarizes the idea well by purporting five hurdles to free labor mobility. These barriers are the preferences of individual workers and individual employers, the actions of the community of workers, employers and the government. However only the first two for us at this stage are meaningful, where under free labor mobility people may find it hard to leave home and living elsewhere because of their preferences, loss of the support of the kinship network, higher cost of living in industrial hubs, in addition to the direct costs associated with migrating. So individuals by exercising free choice isolate themselves from taking advantage of relative wage differentials prevalent across regions.

In summary, what we need to establish is the fact that if two regions face the same prices of goods and if their technologies and their factor endowments are sufficiently similar then they lie in the same diversification cone, and will have the same factor prices; or for factor prices similar enough regions lie in the same diversification cone.

Before proceeding to the other two models of industrial location we need to establish that industries do not exist in isolation, but rather they are interconnected through a variety of forces. These interconnecting forces than for us become significant in determining the industry location. These models offer an alternative to the neoclassical framework and assume increasing returns to scale, imperfect competition, and also they allow for factor mobility in contrast to neo-classical theory of trade. To account for the determinants of industry location in context of the labor markets as predicted by the models of New Trade theory and New Economic Geography our focus mainly is labor market pooling. Labor market pooling implies that industries are more likely to locate in areas where they are likely to find workers with desired skills and that workers move to areas where their skills are required. Labor market pooling leads to industries forming clusters thereby favoring the creation of pools of specialized workers who acquire cluster-specific skills valuable to the firms. Alfred Marshall(1890) emphasized that industries prefer locating specifically in areas that offer constant market for skills.

Dumais et al.(1997) in an attempt to test the importance of Marshallian externalities on a firm’s decision to locate found that decisions regarding industrial location are far more driven by labor market pooling than any other explanatory variable including measures for intellectual spillovers and forward and backward linkages. Overman and Puga (2009) assess the importance of labor market pooling as a source of agglomeration economies empirically using establishment level data from the United Kingdom's Annual Respondent Database (ARD). By regressing a measure of spatial concentration on a measure for potential for labor pooling(which is the average of the difference between the percentage change in the establishment's employment and the percentage change in the sector's employment) check whether sectors where establishments experience a higher potential for labor pooling experience greater spatial concentration even after controlling for other industry characteristics that include measure of the importance of localized intermediate suppliers. They conclude that indeed labor pooling does determine where industries choose to locate. Todd and Abel (2009) using 2000 U.S Census data on occupational employment for 468 occupations in 33 knowledge areas examine that labor pooling measured by specialized knowledge (the extent to which an occupation's knowledge profile differs from the average U.S. job) among other controls includinguse of specialized machineryand the importance of knowledge spillovers exhibit the greatest impact on agglomeration activity. A doubling of the importance of labor market pooling is associated with about a 40 percent increase in the agglomeration index (locational Gini coefficient). Further in a separate regression they found that occupations with similar knowledge requirementstend to co-locate. This is measured by regressing dissimilar knowledge on occupational agglomeration. Such a behavior indeed ensures Marshall's (1890) "constant market for skill". Both results provide evidence for the importance of labor pooling as a determinant of industry location.

A review of the literature shows that there is no single determinant of industry location. Factor supplies of a region and factor prices determine the comparative advantage in production and hence serve as an important determinant for industries to decide where to locate. What else is important is the realization of the fact that industries cannot exist in isolation. Analyzing industries in isolation ignores the linkages existing among industries and markets. These linkages and interconnecting forces between industries are another important determinant of industry location so much so that we have an entire discipline of economics “New economic geography” devoted to understanding how industries choose to locate.

METHODOLOGY

The methodology for finding out any departures from Factor Price Equality is the one developed by Bernard et al (2005). The methodology adopted is expected to serve two purposes. Firstly it will determine if there is inequality in the relative wages of white (skilled) and blue collared (unskilled) workers across regions. Secondly, we wish to determine the manner in which spatial variation in relative factor prices is linked to industrial structure. The test holds under a range of assumptions about production, factors and markets, and is based upon the identifying assumptions of cost minimization, constant returns to scale, Hicks-neutral technology differences and weak separability of the production technology in white-collar and blue-collar workers. This methodology is robust to unobserved factor-region-industry heterogeneity in the quality or composition of factors.

Consider a production technology defined over two factors of production, including white-collar or non-production or skilled workers (N) and blue-collar or unskilled or production workers (A). For the purpose of our analysis the only two factors of production and many industries have been considered, but the analysis can be generalized to the case of as many numbers of factors of production and industries. The value-added production function for industry j and region r is assumed to take the following form:

$$Y_{rj} = T_{rj} f_j(A_{rj}; N_{rj})$$

Where T_{rj} is a Hicks-neutral productivity shift parameter that allows technology to vary across regions and industries. A_{rj} and N_{rj} are the quality adjusted inputs of unskilled and skilled workers respectively in region r and industry j. Every factor enters production through the function f_j , which can vary across industries but is the same across regions within an industry. The technology differences across diaggregated products within industries are estimated by differences in the Hicks-neutral productivity shifter A_{rj} .

Firms in region r and industry j minimize their costs. The cost minimization problem of a firm is thus given by:

$$\min_{A_{rj}:N_{rj}}.: \ w_A A_{rj} + w_N N_{rj} \quad (1)$$

Such that

$$Y_{rj} = T_{rj} f_j(A_{rj}; N_{rj}) \quad (2)$$

Where w^A and w^N are the quality adjusted wage of unskilled workers and skilled workers respectively. Now under the assumptions of constant returns to scale and Hicks-neutral technology differences across regions, the minimized unit cost function is defined as:

$$C = T_{rj}^{-1} \lambda_j (w_r^A, w_r^N) Yrj \quad (3)$$

λ_j is the homogenous of degree one function that varies across industries. Cost minimization induces firms act either as price-takers in product markets (perfect competition) or choose prices subject on a downward sloping demand curve (imperfect competition). The analysis assumes constant returns to scale, however it can also be extended to allow for internal and external increasing returns to scale and to incorporate labor market imperfections, as long as employment continues to be chosen to minimize costs given factor prices.

The quality-adjusted employment level and wage of factors in a region r equals the observed variable scaled by the quality adjuster,

$$z_{rj} = \theta_{rj}^z \widetilde{z_{rj}} \text{ and } w_{rj}^z = \widetilde{w_{rj}^z} \theta_{rj}^z \quad (4)$$

where z Є (A,N) it indexes factors of production, $\widetilde{z_{rj}}$denotes region and industry specific observed quantities of factors of production unadjusted for quality ,$\widetilde{w_{rj}^{z}}$ are region and industry specific observed wages of workers unadjusted for quality and θ_{rj}^{z} denotes a quality adjustor for industry j, region r and factor z; it allows for unobserved variation in quality. Without loss of generality, units in which to measure the quality of each factor of production are chosen such that factor quality in a reference or base region (s) is equal to one i.e., θ_{rj}^{z}=1.

Using Shephard's lemma ($\partial C/\partial w^{z}$) the demand for quality-adjusted factor z is

$$z_{rj} = T_{rj}^{-1} \frac{Y_{rj}\, \partial \lambda_{j}(.)}{\partial w_{r}^{z}} \tag{5}$$

Dividing the two first order conditions obtained by Shephard's Lemma (one for production workers and one for non-production workers) the relative demand for quality-adjusted quantities of production workers and non-production workers can be written as follows:

$$\frac{N_{rj}}{A_{rj}} = \frac{\partial \lambda_{j}(.) / \partial w_{r}^{N}}{\partial \lambda_{j}(.) / \partial w_{r}^{A}} \tag{6}$$

The term for region-industry productivity, A_{rj} , does not appear in equation above. This is because of the property of Hicks-neutral technology differences to affect the marginal revenue product identically for every factor. In contrast, if technology differences across regions within industries were non-neutral, they will affect the relative marginal revenues of skilled and unskilled labor, and induce variation in the relative demand for these factors of production.

Similarly using Shephard's Lemma and (4) (the relationship between observed and quality adjusted employment), the observed relative demand for white collar and blue collar workers is given as follows:

$$\frac{\widetilde{N_{rj}}}{\widetilde{A_{rj}}} = \frac{\theta_{rj}^{A}}{\theta_{rj}^{N}} \frac{\partial\lambda_j(.)/\partial w_r^N}{\partial\lambda_j(.)/\partial w_r^A} \qquad (7)$$

Null Hypothesis of Relative Factor Price Equality (RFPE):

The null hypothesis states that all relative factor prices are equalized (RFPE) across regions. So quality-adjusted relative wages and factor usage across regions r and s must be equal,

$$\frac{w_r^N}{w_r^A} = \frac{w_r^N}{w_r^A} \; and \; \frac{N_{rj}}{A_{rj}} = \frac{N_{sj}}{A_{rj}} \qquad (8)$$

Therefore, to reject RFPE, it is sufficient to establish that relative wages of white-collar and blue-collar workers are different across regions by simply using data on observed white collar and blue-collar wages. But a major concern in that case would be the unobserved variation in factor quality. Such unobserved variation in factor quality might suggest relative factor price inequality even though true quality-adjusted relative wages are equalized across regions.

Therefore under this null hypothesis of RFPE, observed relative wages and observed factor usage (not adjusted for quality) across regions are given by:

$$\frac{\widetilde{w_r^N}}{\widetilde{w_r^A}} = \frac{\theta_{rj}^{N}\widetilde{w_s^N}}{\theta_{rj}^{A}\widetilde{W_s^A}} \& \frac{\widetilde{N_{rj}}}{\widetilde{A_{rj}}} = \frac{\widetilde{N_{sj}}/\theta_{rj}^{N}}{\widetilde{A_{sj}}/\theta_{rj}^{A}} \qquad (9)$$

Where $\theta_{rj}^{N} \neq 1$and$\theta_{rj}^{A} \neq 1$ or wages across regions differ due to unobserved variation. This methodology solves the problem of unobserved variations by combining observed wages and employment (9) into wage bills, where the wage bill= $z_{rj}\, w^{z}_{rj} = \widetilde{z_{rj}}\widetilde{w_{rj}^{z}}$.

Multiplying wages and employment in this way causes region-industry-factor quality adjustors to drop out. As a result, observed relative wage bills, which are generally available to empirical researchers, are equal. Using the wage-bills, the **null hypothesis** now turns out to be

$$\frac{\widetilde{N_{rj}}\widetilde{w^{N}_{rj}}}{\widetilde{A_{rj}}\widetilde{w^{A}_{rj}}} = \frac{\widetilde{N_{sj}}\widetilde{w^{N}_{sj}}}{\widetilde{A_{sj}}\widetilde{w^{A}_{sj}}} \tag{10}$$

This result implies that cost minimization and RFPE together call for the relative unit factor input requirements for quality-adjusted factors of production to be the same across regions and relative wages and relative employment vary across regions due to differences in unobserved factor quality.

Under the *alternative hypothesis* of non-RFPE, observed relative wages vary across regions due to unobserved differences in factor quality and differences in quality-adjusted factor prices. These differences in quality-adjusted factor prices themselves then induce relative unit factor input requirements across regions to differ. The alternative hypothesis of non-RFPE states that the quality-adjusted relative wages (w^N /w^A) differs across regions r and s by a factor, γ^{NA}_{rs}which denotes difference in quality-adjusted relative wages:

$$\frac{w^{N}_{r}}{w^{A}_{r}} = \gamma^{NA}_{rs}\frac{w^{N}_{r}}{w^{A}_{r}} \tag{11}$$

Where region s is the benchmark region, and γ^{NA}_{rs}= γ^{NA}_{r}/ γ^{NA}_{s} and γ^{NA}_{s}=1. Observed relative wages differ across regions, because of variation in factor quality and because of variation in true wages,

$$\frac{\widetilde{w^{N}_{r}}}{\widetilde{w^{A}_{r}}} = \gamma^{NA}_{rs}\frac{\theta^{N}_{rj}\widetilde{w^{N}_{s}}}{\theta^{A}_{rj}\widetilde{w^{A}_{s}}} \tag{12}$$

And similarly observed factor usage now varies across regions because of both differences in factor quality and differences in factor demand driven by the variation in quality-adjusted relative factor prices:

$$\frac{\widetilde{N_{rj}}}{\widetilde{A_{rj}}} = \frac{\theta_{rj}^{N}}{\theta_{rj}^{A}} \frac{\partial\lambda_j(.)/\partial w_r^N}{\partial\lambda_j(.)/\partial w_r^A} \frac{\partial\lambda_j(.)/\partial w_s^N}{\partial\lambda_j(.)/\partial w_s^A} \frac{\widetilde{N_{sj}}}{\widetilde{A_{sj}}} \tag{13}$$

Now again by combining relative factor usage and relative wages by multiplying the expressions for observed relative factor wages (12) and observed relative employment (13), the terms for unobserved factor quality cancel out. However, relative wage bills now differ across regions because of variations in factor prices and in factor usage,

$$\frac{\widetilde{N_{rj}}\widetilde{w_{rj}^N}}{\widetilde{A_{rj}}\widetilde{w_{rj}^A}} = \Phi_{rsj}^{NA} \frac{\widetilde{N_{sj}}\widetilde{w_{sj}^N}}{\widetilde{A}_{sj}\widetilde{w}_{sj}^A} \tag{14}$$

Where

$$\Phi_{rsj}^{NA} = \gamma_{rs}^{NA} \frac{\partial\lambda_j(.)/\partial w_r^N}{\partial\lambda_j(.)/\partial w_r^A} \frac{\partial\lambda_j(.)/\partial w_s^A}{\partial\lambda_j(.)/\partial w_s^N} \tag{15}$$

where $\gamma_{rs}^{NA} \neq 1$ such that

$$\frac{w_r^N}{w_r^A} = \gamma_{rs}^{NA} \frac{w_r^N}{w_r^A}$$

The term inside the braces in (15) captures differences in unit factor input requirements and it depends in general on the difference in quality-adjusted relative wages and differences in the relative prices of other factors of production. Analyzing (10) and (14), the null hypothesis of

RFPE can be tested in the presence of unobserved differences in factor quality by testing whether relative wage bills are equalized across regions. Finding that $\Phi_{rsj}^{NA} \neq 1$ is sufficient to reject the null hypothesis of RFPE.

Econometric specification:

Relative Factor Price Equalization

Under the null of RFPE, the ratio of the white collar worker's wage bill to blue collar workers' wage bill is the same across regions within an industry. This implies that, for an industry *j*, each region's relative wage bill equals the value for any base region *s* and, in particular, for the aggregate Punjab (punj),

$$\frac{\widetilde{N_{rj}}\widetilde{w_{rj}^{N}}}{\widetilde{A_{rj}}\widetilde{w_{rj}^{A}}} = \frac{\widetilde{N_{sj}}\widetilde{w_{sj}^{N}}}{\widetilde{A_{sj}}\widetilde{w_{sj}^{A}}} = \frac{\widetilde{N_{punj}}\widetilde{w_{punj}^{N}}}{\widetilde{A_{punj}}\widetilde{w_{punj}^{A}}}$$

The simplest way to test the null is to regress the ratio of wage bills for region *r* relative to the ratio for the aggregate Punjab on a set of region dummies d_r ,

$$ln\frac{RWB_{rj}^{NA}}{RWB_{punj}^{NA}} = \sum \alpha_r^{NA} d_r + \varepsilon_{rj}^{NA} \qquad (16)$$

Where $RWB_{rj}^{NA} = \widetilde{N_{rj}}\widetilde{w_{rj}^{N}} / \widetilde{A_{rj}}\widetilde{w_{rj}^{A}}$ and α_{rs}^{NA} are coefficients of regional dummies. ε_{rsj}^{NA} are the error term.

The *F*-test of whether the α_{rs}^{NA} are jointly equal to zero provides a test of the null hypothesis of RFPE across regions. Rejecting $\alpha_{rs}^{NA}=0$ is sufficient to reject the null hypothesis of RFPE. For any pair of regions *r* and *ŕ,* $\alpha_{rs}^{NA} = \alpha_{ŕs}^{NA}$ implies that those regions face the same relative factor prices.

Similarly testing RFPE by allowing every individual region to be the base region to avoid having results driven by the choice of region. So the equation to be estimated is

$$ln\frac{RWB_{rj}^{NA}}{RWB_{sj}^{NA}} = \sum \alpha_{rs}^{NA} d_r + \varepsilon_{rsj}^{NA} \quad (17)$$

Using equation (17), we can check for bilateral relative factor price equalization between districts, by comparing regions to each other, rather than comparing each district to only Punjab as a whole as in (16).

DATA

To carry out the analysis, plant level data taken from the *Census of Manufacturing Industries* (CMI) for Punjab for 2000-01 and 2005-06 has been employed. The CMI is conducted as a joint effort by the Federal Bureau of Statistics (FBS) Government of Pakistan and Provincial Directorates of Industries and Bureaus of Statistics (BOS). The CMI provides data for 2-digit, 3-digit, 4-digit and 5-digit classifications under the Pakistan Standard Industrial Classification (PSIC) according to geographic subdivision at the district, province and national levels. CMI provides data on quantities and values of inputs and outputs, census value added, contribution to GDP, fixed assets, stocks, employment, labor cost and industrial taxes.

The analysis in this paper requires information on the wage bills of production and non-production workers and the district and industry to which these workers belong. In turn, to calculate wage bills information is required on the number of production and non-production employees in a firm and their respective wage. This information is available in the CMI. The analysis also requires information on the industries produced by a region. This can be retrieved very easily as CMI has information according to geographic subdivision at the district, province and national levels.

The administrative organisation of the province of Punjab is provided in table 1. The zonal distribution is in accordance with what Cheema (2008) has proposed in his analysis on geography of poverty in Punjab. According to his distribution the province of Punjab can be divided into four zones namely North, South, Centre and West with the Centre being the largest.

Based on this administrative breakdown we have performed our analysis on two levels: firstly at the smaller district level and then at the larger zonal level. For 2001 this analysis was done for 34 districts and 592 4-digit industries and for 2005-06 this analysis was done for 35 districts (includes 34 districts from 2001 and newly created Nakanasahib additionally in 2005-06) and 540 4-digit industries.

Table 1: Administrative Regions and Districts

Zone	District	Zone	District
North	Attock	West	Bhakkar
	Chakwal		D.G. Khan
	Jhelum		Khushab
	Rawalpindi		Layyah
			Mianwali
Centre	Gujranwala		Muzaffargarh
	Gujrat		Rajanpur
	Hafizabad		
	Jhang	South	Bahawalnagar
	Kasur		Bahawalpur
	Lahore		Khanewal
	Mandi Bahauddin		Lodhran
	Nankanasahib*		Multan
	Narowal		R.Y. Khan
	Okara		
	Pakpattan		
	Sahiwal		
	Sargodha		
	Sheikhupura		
	Sialkot		
	TT Singh		
	Vehari		

* Nankanasahib was made a separate district in 2006. So the analysis for 2001 does not treat Nankanasahib as a separate district.

5. Empirical Results

In order to test for bilateral relative factor price equalization between districts, by comparing regions to each other, equation (17) was estimated by allowing every individual region to be the base region. Since the analysis is being done for thirty five districts of Punjab this equation was estimated for a total of thirty five times for both the years i.e, 2001 and 2006, keeping each individual district as the base region.

An F-test to check for the joint significance of α_{rs}^{NA} in equation (17) was performed for each of the districts and zones for both years in order to check for relative factor price equality between a base district and all other districts of Punjab, The results for the F-test of equation 17 are shown in table 2 for both 2001 and 2006. The table is a representation of bilateral rejections by base region.

Table 2: Bilateral rejections by base Districts.

	Percentage of rejections		*Distribution of rejections across base regions*		
	5% level of significance	*10% level of significance*	*Minimum*	*Mean*	*Maximum*
Zones					
2000-01	75%	75%	1	1	3
2005-06	25%	75%	0	1	2
Districts					
2000-01	23.5	26.5	0	3	14
2005-06	14.28	20	0	4	13

Notes: Bilateral rejections of relative factor price equality are based on estimation of equation (17) for all possible base regions.
Source: Author's own calculations using CMI.

Looking closely at these results we see that for 2001 out of 34 regressions done for each base district, for nearly 23 percent of these regressions were we able to reject the null hypothesis for relative factor price equality at the 5% level of significance. To be precise these 23 percent include eight districts namely when the base districts are Rawalpindi, Sargodha, Faisalabad, Sheikhupura, Lahore, Gujranwala, Vehari and Sialkot. However at 10% level of significance we were able to reject the null for relative factor price equality for 26 % of the districts including Bahawalpur. Using the 2006 data, looking at Table 2 we see that out of 35 regressions, for 14% of the districts are we now able to reject the null for relative factor price equality at the 5% level of significance. This is the case when the base districts include Lahore, Faisalabad, Gujrat, Sialkot and DG Khan. At 10% level of significance we are now able to reject the null for relative factor price equality for 20% of the districts including Multan and Nankanasahib. One noticeable thing in these results is the fact that most of these districts are located in the the Centre. This can be viewed from Figure 3, 4, 5, and 6. The industrial mix of districts in the Centre is very wide and diversified. This point will be illustrated later in the next section. However for now it has been established empirically that centre is and has been contributing most towards the factor price inequality within Punjab. So the relative wages of skilled workers are not constant across Punjab and therefore there is a need to test if this differential also impacts the industrial landscape across Punjab or not.

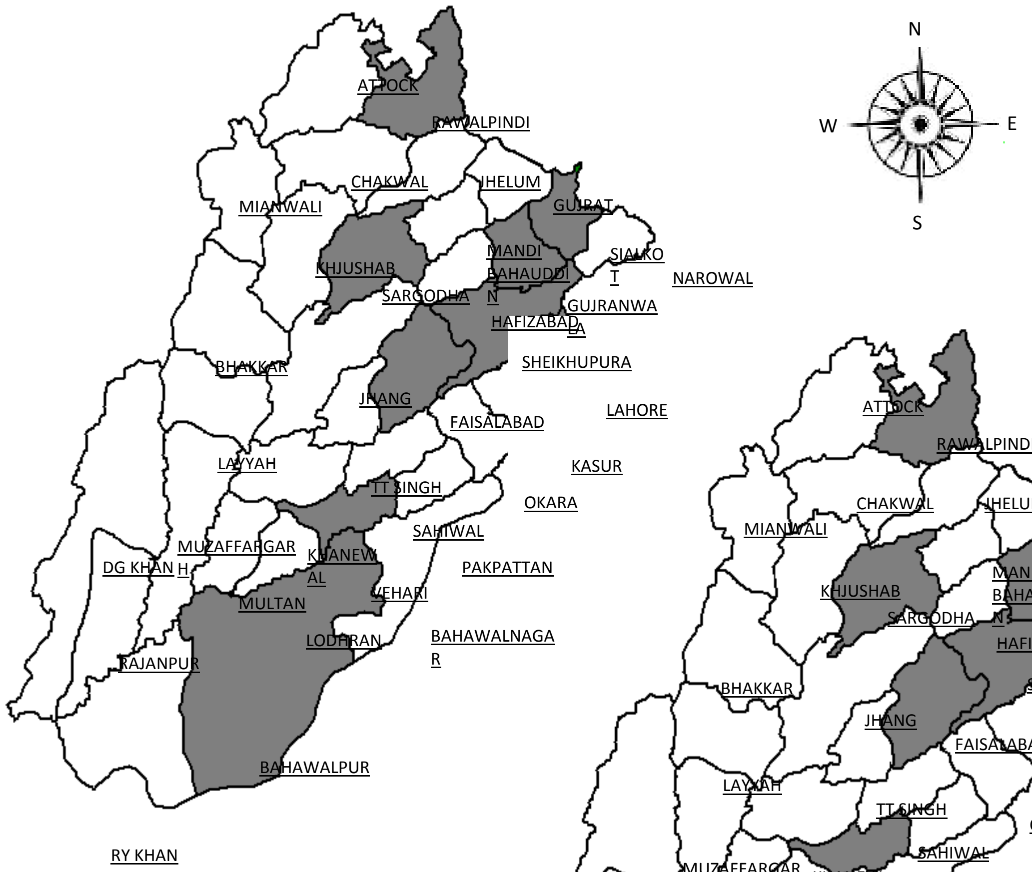


Figure 3: Prevalence of Relative Factor Price Inequality at District level at 10% significance level
Year: 2001.

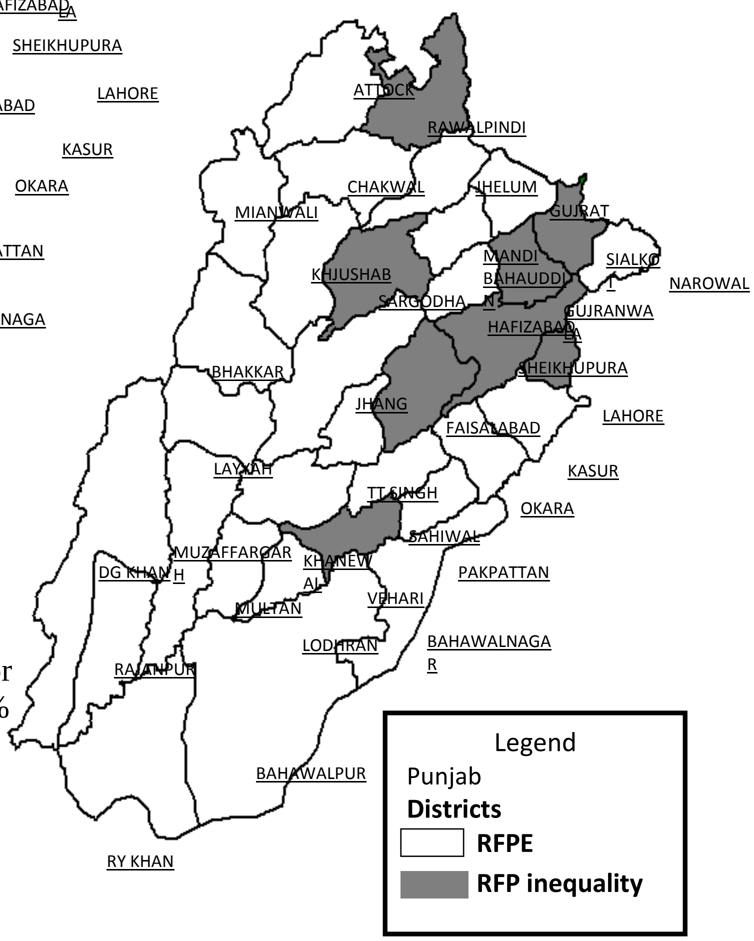


Figure 4: Prevalence of Relative Factor Price Inequality at District level at 5% significance level.

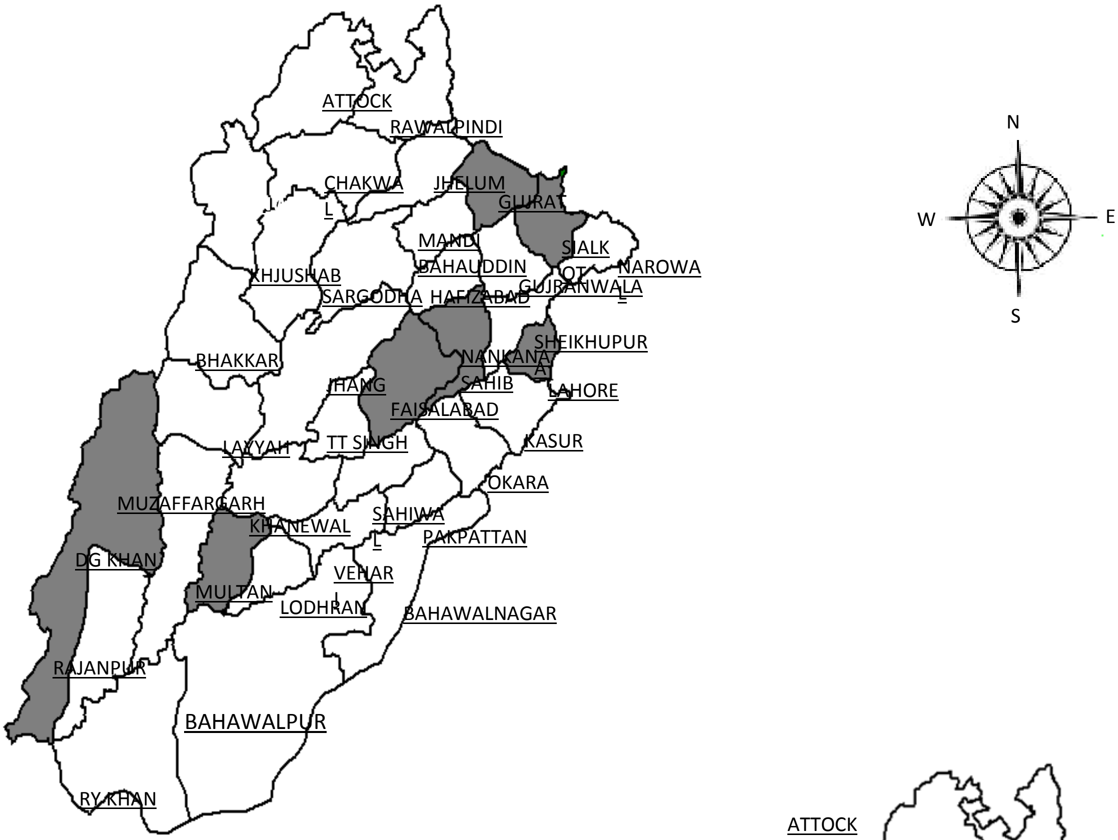


Figure 5: Prevalence of Relative Factor Price Inequality at District level at 10% significance level
Year: 2005.

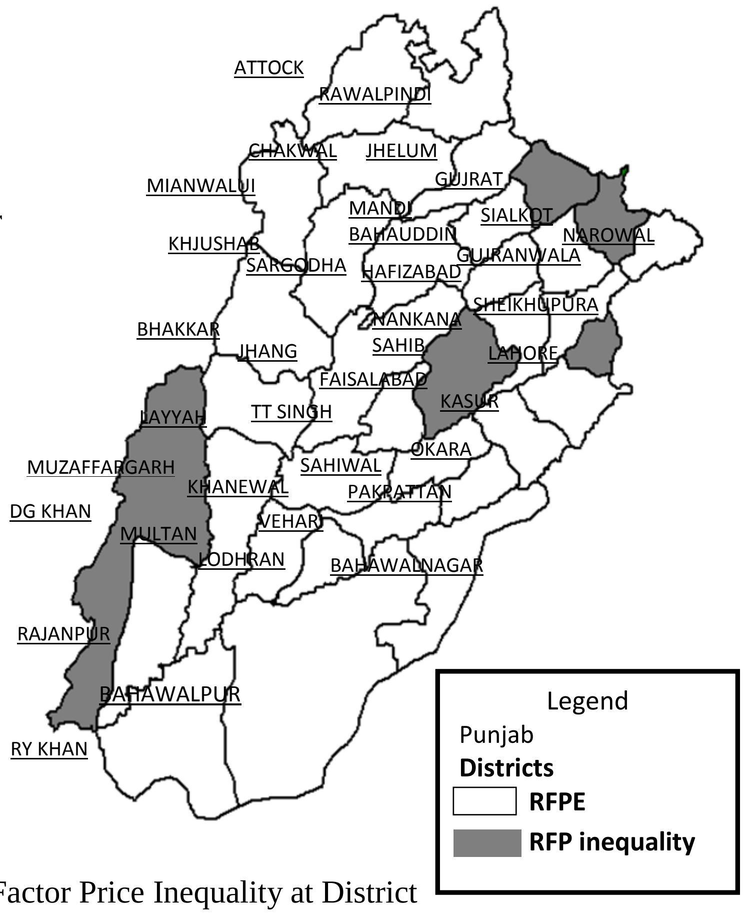


Figure 6: Prevalence of Relative Factor Price Inequality at District level at 5% significance level.
Year: 2005.

Relating Factor Intensity and Factor Prices:

According to the Heckscher-Ohlin theorem, districts abundant in skilled labor should have lower relative wages for skilled labor. So in the first stage of the analysis, for each regression of the form equation 17, if a district appears significantly positive it suggests this district has a higher relative wage bill for skilled workers than the base district suggesting that it is less abundant in skilled workforce. Table 3 gives results for the number of times a district's coefficient came out statistically significant in the thirty five regressions that it appeared in (where each regression was done for a particular base district). These results are based upon testing the individual coefficients of district dummies in equation 17. A significant coefficient indicates that the relative wages of skilled workers in that particular district are higher or lower than the base district depending upon the its sign. Looking at the table 3, the highest number of times a district appeared significant is Lahore. What is more interesting to note is the fact that for the 11 regressions in which it appeared significant its coefficient has stayed positive in each of these regressions, suggesting strongly the relative wages of skilled workers is higher in Lahore than the respective base district. These base districts mainly include Rawalpindi, Gujrat, Sialkot, Mianwali, Multan, and Nankanasahib.

On closer examination of the data, it appears that Lahore is rich in industries that can be characterised as being relatively more unskilled labor intensive. Out of the 86 industries that are produced in Lahore, 71 industries use unskilled labor intensively. The skill intensiveness of industries for this study was measured by estimating the ratio of the number of skilled to unskilled workers for each industry within a district. The top 25% of the industries based on this ratio have been classified as being skill intensive i.e. a ratio of skilled to unskilled labor greater than 0.406 for 2006 and greater than 0.441 for 2001. For Lahore, 83% of the industries are unskilled labor intensive pointing to an abundance of such labor in Lahore.

Moreover comparing it to the base regions in each of these regressions it is seen that of all these regions, Lahore has the highest percentage of unskilled labor intensive industries except Sialkot. Multan for example has 72%, Gujrat 84%, and Mianwali 40% of its industries being classified as being unskilled labor intensive. Faisalabad has results similar to that of Lahore since it has its industrial composition similar to that of Lahore.

Table 3: Bilateral Rejections by Base Districts (2005-06)

Base Districts	*2005-06*						
	Bilateral Rejections						
	Total no. of Rejections	*Significance Level*		*Positive Significant Coefficients*		*Negative Significant Coefficients*	
		10%	*5%*	*10%*	*5%*	*10%*	*5%*
RAJANPUR	16	16	12	3	1	13	11
LAHORE	11	11	10	11	10	0	0
SHEIKHUPURA	8	8	4	8	4	0	0
SIALKOT	7	7	4	0	0	7	4
MULTAN	7	7	2	4	0	3	2
SAHIWAL	7	7	3	3	1	4	2
GUJRAT	6	6	4	2	1	4	3
FAISALABAD	6	6	2	6	2	0	0
OKARA	6	6	3	0	0	6	3
GUJRANWALA	5	5	5	5	5	0	0
MANDI BAHA-UD-DIN	5	5	1	1	0	4	1
BAHAWALPUR	5	5	5	0	0	5	5
RAWALPINDI	4	4	1	1	0	3	1
KHUSHAB	4	4	1	4	1	0	0
MIANWALI	4	4	2	0	0	4	2
JHELUM	3	3	1	0	0	3	1
BHAKKAR	3	3	2	0	0	3	2
T.T. SINGH	3	3	2	0	0	3	2
D.G. KHAN	3	3	2	3	2	0	0
MUZAFFARGARH	3	3	1	1	0	2	1
R.Y. KHAN	3	3	0	3	0	0	0
CHAKWAL	2	2	2	1	1	1	1
HAFIZABAD	2	2	1	0	0	2	1
SARGODHA	2	2	1	0	0	2	1
KASUR	2	2	1	0	0	2	1
KHANEWAL	2	2	1	0	0	2	1
LODHRAN	2	2	1	2	1	0	0
VEHARI	2	2	0	1	0	1	0
PAKPATTAN	2	2	0	2	0	0	0
NANKANA SAHIB	2	2	2	0	0	2	2
ATTOCK	1	1	1	0	0	1	1
JHANG	1	1	1	0	0	1	1
BAHAWALNAGAR	1	1	0	1	0	0	0
NAROWAL	0	0	0	0	0	0	0
LAYYAH	0	0	0	0	0	0	0

Source: Author's calculations based on CMI 2005-06

Similar to Lahore, 78% of the industries in Faisalabad are unskilled labor intensive and therefore the six times that it has appeared significant in six different regressions it had a positive sign suggesting that the relative wages of skilled workers are higher in Faisalabad as compared to the base districts (Rawalpindi, Gujrat, Mianwali and Multan). All of these base districts have a lower percentage of unskilled labor intensive industries compared to Faisalabad (as is evident from table 4) therefore leading the relative wage bill of skilled workers to be lower in these districts. Sheikhupura likewise has results similar to those of Faisalabad and Lahore owing to the similarity in industrial composition of these three districts.

Rajanpur also has appeared significant in sixteen regressions out of 35 and out of these sixteen regressions thirteen times has it appeared significantly negative. This suggests an abundance of skill intensive industries in Rajanpur. Rajanpur has 100% of its industries classified as skill intensive and therefore the lower relative wage bill of skilled workers compared to the base districts.

Sialkot is one district that has appeared significantly negative in seven out of thirty five regressions. This suggests that relative wages of skilled workers is significantly lower in Sialkot than Rawalpindi, Gujranwala, Gujrat, Chakwal, Faisalabad, Lahore and Vehari. These results are difficult to explain since Sialkot has 88% of its industries classified as unskilled labor intensive. This proportion is higher than each of these base districts except Gujranwala as can be seen from table 4. Having a higher proportion of industries that are unskilled labor intensive should have caused the relative wages of skilled workers to be higher as in Lahore, but the sign on Sialkot in each of these regressions suggests otherwise.

Multan's is a case that lies in between the two extremes (Lahore and Sialkot). Out of 35 regressions, in seven of these it has appeared significant, thrice below zero and four times above zero. Multan has lower relative wages of skilled workers compared to Lahore, Faisalabad and Gujranwala. The data does confirm this prediction since Multan has a higher percentage (28%) of skill intensive industries than each of these base districts (Lahore 17%, Faisalabad 22% and Gujranwala 10%). Similarly Multan has higher relative wages for skilled workers compared to Sahiwal, Hafizabad, Mianwali and Gujrat). Only Mianwali

Table 4: Industrial Structure

Zone/District	2005-06			2000-01		
	No. of Industries	% of Low White Collar Intensity Industries	% of High White Collar Intensity Industries	No. of Industries	% of Low White Collar Intensity Industries	% of High White Collar Intensity Industries
North						
Rawalpindi	18	72%	28%	30	70%	30%
Attock	7	86%	14%	8	88%	13%
Chakwal	6	67%	33%	4	100%	0%
Jhelum	6	83%	17%	10	80%	20%
Centre						
Lahore	86	83%	17%	112	80%	20%
Gujranwala	59	90%	10%	64	86%	14%
Sheikhupura	49	84%	16%	67	78%	22%
Faisalabad	32	78%	22%	36	81%	19%
Kasur	29	72%	28%	27	78%	22%
Sialkot	26	88%	12%	23	87%	13%
Gujrat	19	84%	16%	12	83%	17%
Sahiwal	17	76%	24%	17	71%	29%
Okara	14	64%	36%	13	62%	38%
Sargodha	14	71%	29%	14	71%	29%
TT Singh	11	64%	36%	10	70%	30%
Hafizabad	9	89%	11%	6	67%	33%
Nansahib	7	71%	29%	-	-	-
Jhang	6	50%	50%	19	79%	21%
ManBah	5	60%	40%	3	33%	67%
Vehari	5	40%	60%	3	33%	67%
Pakpatan	4	25%	75%	4	25%	75%
Narowal	3	0%	100%	2	0%	100%
West						
Khushab	12	75%	25%	7	71%	29%
Muzaffargarh	10	70%	30%	6	50%	50%
Bhakkar	7	57%	43%	2	50%	50%
DG Khan	6	67%	33%	7	71%	29%
Mianwali	5	40%	60%	4	50%	50%
Rajanpur	3	0%	100%	1	0%	100%
Layyah	3	0%	100%	3	0%	100%
South						
Multan	25	72%	28%	34	74%	26%
Khanewal	12	58%	42%	11	73%	27%
Rahim Yar Khan	9	44%	56%	10	40%	60%
Bahawalpur	8	63%	38%	13	69%	31%
Bahawalnagar	5	60%	40%	6	50%	50%
Lodhran	3	33%	67%	4	50%	50%

*Note:*The degree of white collar intensity is based upon the ratio of skilled to unskilled workers in an industry in a particular district. High white collar intensity industries are top 25% of the industries based on this ratio.

(60%) has a higher percentage of skill intensive industries than Multan and therefore lower relative wage of skilled workers than it. But Gujrat (16%), Sahiwal (24%) and Hafizabad (11%) have lower percentages of skill intensive industries compared to Multan, so Multan in theory should have had lower relative wages of skilled workers which is not the case. So the reason for this differential in the opposite direction remains unclear.

Gujrat is another district that lies midway between the two extremes. It has significantly lower relative wages of skilled workers than that of Lahore, Sheikhupura, Gujranwala and Faisalabad. Of these base districts, all but Faisalabad have a higher or equal percentage of unskilled labor intensive industries. Though it has higher relative wages of skilled workers compared to Sialkot and Bahawalpur where the former has a higher percentage of unskilled labor intensive industries and Bahawalpur has a higher percentage (38%) of skill intensive industries compared to Gujrat.

Seeing all these results together the one thing that we notice is the fact that relative wage bill for skilled workers is higher in districts that are situated in the Centre namely Lahore, Sheikhupura, Gujranwala and Faisalabad, Some common features of these districts are that these districts are industrial hub of Punjab. They have a large number of industries that they produce and majority of these industries are unskilled labor intensive as can be seen by the table 4 and therefore have higher relative wages of skilled workers (lower relative wages of unskilled workers). Moreover the share of manufacturing employment of these districts in skill intensive industries is a very small figure as can be seen from the table 5 reinforcing the fact that these districts are abundant in unskilled labor and therefore produce industries that are relatively more unskilled labor intensive by Hecksher-Ohlin theorem. However the quality of human capital in these regions is much better and would therefore suggest otherwise. The quality of human capital across various districts of Punjab is depicted by the human development index in the table A in appendix 1. below. It can be seen that districts in the Center (the industrial hub of Punjab) have much higher ranks than districts in either South or west.

So before professing these districts as being skill scarce we need to establish the fact that Punjab as a whole does not specialize in manufacturing industries that are very skill intensive. So the diversification cones for individual districts are not very diverse after all as the factor intensity ratios and the relative prices of these factors for a lot of districts are not very distinct. About 75% of the industries that are being produced in Punjab have ratio of skilled to unskilled workers that is less than 0.5, to be precise 75% of the industries have this ratio less than 0.406 as of 2005-06. Out of these 75% industries, 50% of industries have this ratio less than 0.257. So the industries that are being produced in Punjab do not demand very high skill level on the whole.[2].

So the apparent skill scarcity in districts that are the industrial hub of Punjab might be due to the fact that skilled workforce in these areas get themselves employed in sectors other than manufacturing like the services sector where there skills are more suited. So the scarcity in these more developed regions of Punjab might only be artificial as more skilled workers find jobs other than those in manufacturing and therefore the industries have to attract the skilled workers here at very high wage rates. On the other hand districts that have significantly lower relative wages of skilled workers are located in the south or west of Punjab. These mainly include Bahawalpur, Rajanpur and Multan. Lower relative wages of skilled workers by the Heckscher Ohlin theorem should suggest an abundance of skilled workers in these areas. Some common features of these districts are that they do not have very large number of

[2] The industries that are being classified as very skill intensive based on this ratio i.e., ratio of skilled to unskilled workers is greater than one, include: industry for vegetable and inedible animal oil and fats, industry for starch and its products, petroleum products industry, lime, plaster and its products, and finally iron and steel basic industries. The share of total district employment employed in these industries is a very minuscule figure as can be viewed from table 5

industries that they are producing compared to districts like Lahore, Faisalabad, Sheikhupura or Gujranwala but the industries that they are producing have a higher proportion of skill intensive industries and therefore the lower relative wages of skilled workforce. Also the percentage share of manufacturing employment of these districts in the skill intensive industries of Punjab is very high compared to the districts in central Punjab as is evident from the table 5. Again all of these regions are less developed compared to the districts like Lahore, Faisalabad, Sheikhupra or Gujranwala. The lower relative wages of skilled workers in these areas suggests nothing but an abundance of skilled labor force here. But in reality this is not the case the socio-development statistics at least do not suggest this.

So one then needs to ask what is pushing relative wages of skilled workers down in the more developed districts of Punjab including Lahore, Faisalabad, Sheikhupura, and Gujranwala. One explanation could be the fact that the level of skill required in the manufacturing industries of Punjab as discussed earlier is not very high on average. This type of skill is found abundantly in these areas, and moreover the opportunities for these semi-skilled workers to get better and more advanced skills in these areas are limited. The kind of skills required in these manufacturing industries are what people here can easily achieve with the resources and opportunities that they have available to them. With not many opportunities of employment and not a very wide spectrum of skills to offer, the manufacturing industries in these areas face an excess supply of semi-skilled labor which they are willing to hire only at lower relative wages.

Some alternative explanations to these counterintuitive results are discussed below. The manufacturers are able to offer different wage rate across various regions of Punjab because of the factors that render labor immobile. People may find it hard to leave home and living elsewhere because of their preferences, loss of the support of the kinship network, higher cost of living in industrial hubs, in addition to the direct costs associated with migrating. So essentially if the costs associated with migration are greater than the expected gains from it then workers would rather not choose to migrate to benefit from the wage differential.

In addition, we need to take into account another factor which is the quality of labor. The methodology that we have followed controls for this factor and therefore believes that the information of quality of labor force is embedded in the observed factor prices. If we stick to this view, the differential in the relative wages of skilled workers points to the quality of skills being better in the big districts like Lahore, Faisalabad, or Gujranwala. Therefore we need to take into account the fact that even the relatively unskilled labor force in these districts is more efficient than the unskilled laborers of South and West. All that is being produced across Punjab is not such that it should require very distinct factor intensity vectors. Whatever is being produced can be manufactured in the big districts in a more unskilled labor intensive way since the overall quality of labor is better in these areas as opposed to South and West where more resources have to be put in to make the same. This is just like the Leontieff paradox where United States being capital intensive was found to be exporting labor intensive products, but on further examination it was discovered that the original analysis missed the role of human capital. United States was found to have its labor force three times more efficient than the US trading partners. Similarly in the case of Punjab, unskilled labor intensive products of Punjab in its industrial hub are being produced by a labor force that it is on average more efficient. Moreover the elasticity of substitution of unskilled labor for skilled labor seems to be very high here since almost all industries come under the head of consumer goods.

Finally these results could also be due to the demand bias. Since per capita income is higher in the districts in the centre, the consumption demand is on the rise here which could also be the cause for industries manufacturing consumer goods to be attracted here. So the demand bias coupled with the ease of elasticity of substitution of factors might have encouraged these unskilled labor intensive industries to come to the relatively more developed areas of Punjab.

As for other reasons why developed areas like Lahore, Faisalabad, Gujranwala should have higher relative skilled wages, it might have something to do with the sample of firms included in the CMI. For the less developed districts which appear to have a high share of skill-intensive industries, it may be the case that the CMI is missing a lot of the low-skill intensive firms in those districts, either because the firms are too small to be captured by the CMI, or because the response rate to the CMI census is lower in those districts. So either due to the sampling error or irregularities in the reponse rate might be leading to these unexpected results.

The two exceptions are Gujrat and Sialkot that have significantly lower relative wages of skilled workers compared to the base districts that are relatively more skill intensive than either of these. Analyzing the relatively skill intensive industries in these two districts one finds out that these are mainly grain milling, light engineering, pharmaceuticals and beverage manufacturing industries. However these industries are consuming only 2.7% of the workforce for Gujrat and only 0.34% of the workforce for Sialkot. The lower relative wages of skilled workforce is exactly what one would expect as these districts are home to a large number of industries and also have a fairly good quality of labor force in terms of its skills.

The results are similar for 2001, these results are presented in table 6. Most of the districts that have significantly higher relative wages of skilled workers are concentrated in the center of Punjab. These districts are mainly Lahore, Faisalabad, Gujranwala and this time Sialkot as well. Other districts are DG Khan, Lodhran and Vehari. Some common characteristic of these districts are that they have relatively higher percentages of industries for which the ratio of skilled to unskilled workers is very small, as can be seen in table 4. Also these districts have a very low share of employment in skill intensive industries of Punjab and therefore very high share of employment in unskilled labor intensive industries, this can be seen from table 5. These facts indicate that these districts are relatively more abundant in unskilled labor.

Table 5: District wise Distribution of Manufacturing Employment across Industries of varying Factor Intensity

Zone/ District	2005-06			2000-01		
	No. of Industries	% of Manufacturing Employment in Low White Collar Intensity Industries	% of Manufacturing Employment in High White Collar Intensity Industries	No. of Industries	% of Manufacturing Employment in Low White Collar Intensity Industries	% of Manufacturing Employment in High White Collar Intensity Industries
North						
Chakwal	6	99.13	0.87	4	100.0	0.00
Rawalpindi	18	95.42	4.58	30	97.87	2.13
Jhelum	6	95.35	4.65	10	96.19	3.81
Attock	7	73.24	26.76	8	82.99	17.01
Centre						
Sialkot	26	99.66	0.34	23	97.55	2.45
Gujrat	19	97.31	2.69	12	99.47	0.53
Kasur	29	94.66	5.34	27	96.01	3.99
Faisalabad	32	93.92	6.08	36	91.86	8.14
Hafizabad	9	93.68	6.32	6	80.55	19.45
Vehari	5	90.65	9.35	3	55.81	44.19
Sheikhupura	49	88.96	11.04	67	93.57	6.43
Lahore	86	85.33	14.67	112	86.79	13.21
gujranwala	59	78.62	21.38	64	77.4	22.6
Sahiwal	17	76.05	23.95	17	93.98	6.02
Nansahib	7	68.88	31.12	-	-	-
Okara	14	61.05	38.95	3	29.16	70.84
Sargodha	14	53.8	46.2	14	54.58	45.42
Jhang	6	36.94	63.06	19	57.73	42.27
ManBah	5	32.9	67.1	3	45.97	54.03
TTsingh	11	29.77	70.23	10	39.31	60.69
Pakpattan	4	0.33	99.67	4	1.87	98.13
Narowal	3	0	100	2	0	100
West						
Muzaffargarh	10	91.84	8.16	6	95.65	4.35
DG Khan	6	91.05	8.95	7	95.32	4.68
Khushab	12	87.39	12.61	7	84.6	15.4
Bhakkar	7	56.45	43.55	2	56.1	43.9
Mianwali	5	45.35	54.65	4	40.4	59.6
Rajanpur	3	0	100	1	100	0
Layyah	3	0	100	3	0	100
South						
Bahawalnagar	5	100	0	6	38.35	61.65
Khanewal	12	85.58	14.42	11	68.96	31.04
Multan	25	83.12	16.88	34	69.29	30.71
Bahawalpur	8	29.62	70.38	13	17.4	82.6
RyKhan	9	8.42	91.58	10	30.65	69.35
Lodhran	3	3.05	96.95	4	27.31	72.69

and therefore have relatively higher relative wages of skilled workers compared to the base districts. The only exception to this rule (as set by the Heckscher Ohlin theorem) in 2001 is Vehari that has a higher proportion of skill intensive industries (66%) yet has a higher relative wage bill of skilled workers compared to those of the base districts (that have lower proportion of skill intensive industries) like Gujranwala, Sargodha or Kasur. This is also evident from table 4. Likewise Lodhran is also an exception as it has 50% of skill intensive industries and 73% of the labor force employed in these industries indicating relatively higher skill intensity but has higher relative wages for skilled workers. However what is common between Lodhran and Vehari is that they both have very few number of industries. Lodhran has 4 industries and Vehari has only 3. Such a small number of industries might be the reason for this bias and that the Heckscher Ohlin effect is not being validated here.

On the other hand districts that have significantly relatively lower wages of skilled workers, compared to the base districts, are mainly Jhelum, Pakpattan, Sargodha and Muzaffargarh. Jhelum has lower relative wages compared to base districts like Gujranwala, Faisalabad, Sheikhupura, Lahore, Lodhran and DG Khan. Each of these base districts (except Sheikhupura, Lodhran and DG Khan) have a higher ratio of unskilled intensive industries and therefore higher relative wages of skilled workers than Jhelum. Pakpattan likewise has a very high percentage of skill intensive industries moreover about 98% of the labor force is employed in the skill intensive industries marking the fact that it is relatively skill abundant than the base districts and therefore the lower relative wages of skilled workers compared to the base districts.

Muzaffargarh and Sargodha also have very high percentages of skill intensive industries compared to the base districts and therefore have lower relative wages of skilled workers. But the sign in both the cases is opposite when Vehari is the base district. They both should have higher relative wages of skilled workers compared to Vehari owing to relatively higher percentage of skill intensive industries in Vehari, but this is not the case. The predictions of Heckscher Ohlin theorem have been reversed for Vehari, Lodhran, DG Khan and Jhelum. The reason once again could be the very small number of industries in these districts. Again in these results the one thing that is evident the fact that the results are suggesting more skilled labor force in

Table 6: Bilateral Rejections by Base Districts (2005-06)

Base Districts	*2000-01*						
	Bilateral Rejections						
	Total no. of Rejections	*Significance Level*		*Positive Significant Coefficients*		*Negative Significant Coefficients*	
		10%	*5%*	*10%*	*5%*	*10%*	*5%*
GUJRANWALA	10	10	6	10	6	0	0
SIALKOT	9	9	6	9	6	0	0
FAISALABAD	9	9	4	9	4	0	0
LODHRAN	7	7	1	7	1	0	0
D.G. KHAN	7	7	3	7	3	0	0
JHELUM	6	6	3	3	0	3	3
SARGODHA	6	6	5	0	0	6	5
BHAKKAR	6	6	3	2	0	4	3
VEHARI	6	6	3	6	3	0	0
JHANG	5	5	1	3	0	2	1
LAHORE	5	5	2	5	2	0	0
RAWALPINDI	4	4	2	2	0	2	2
KHUSHAB	4	4	1	3	1	1	0
MIANWALI	4	4	4	6	4	0	0
SHEIKHUPURA	4	4	3	2	1	2	1
MULTAN	4	4	3	0	0	4	3
KHANEWAL	4	4	4	3	3	1	1
PAKPATTAN	4	4	2	0	0	4	2
MUZAFFARGARH	4	4	2	0	0	4	2
HAFIZABAD	3	3	0	2	0	1	0
KASUR	3	3	2	0	0	3	2
CHAKWAL	2	2	0	1	0	1	0
T.T. SINGH	2	2	2	0	0	2	2
OKARA	2	2	2	0	0	2	1
BAHAWALPUR	2	2	1	0	0	2	1
ATTOCK	1	1	0	0	0	1	0
GUJRAT	1	1	0	0	0	1	0
MANDI BAHA-UD-DIN	1	1	1	0	0	1	1
SAHIWAL	1	1	1	1	1	0	0
BAHAWALNAGAR	1	1	0	0	0	1	0
R.Y. KHAN	1	1	1	0	0	1	1
NAROWAL	0	0	0	0	0	0	0
RAJANPUR	0	0	0	0	0	0	0
LAYYAH	0	0	0	0	0	0	0
NANKANA SAHIB	0	0	0	0	0	0	0

in all those areas of Punjab that rank the lowest in terms of socio economic development including Pakpattan, Lodhran, Vehari or Muzaffargarh where more developed regions like

Gujranwala, Faisalabad, Lahore or Sialkot have higher relative wages of skilled workers is suggestive of the fact the fact that they lack skilled workers.

Testing for Relative Factor Price Equality at Zonal Level of Punjab:

From table 1 we can see that Punjab can be divided into four zones i.e., North, Centre, West and South. Table 2 shows the analysis of factor price equalization at the larger zonal level as well. An F-test to check for joint significance of α_{rs}^{NA} in equation (17) was performed for each of the base zone for both years in order to check for relative factor price equality between a base zone and all other zones of Punjab. It can be seen from table 2 that for 2000-01 out of the four regressions (one for each of the base zone), in 75% of the regressions we were able to reject the null of factor price equality both at 5 and 10% level of significance. Whereas for 2005-06, out of the four regressions 25% (one zone) of the regressions are showing factor price inequality at 5% level of significance. This particular zone is Centre and therefore coincides with our results presented earlier. On the other hand 75% of the zones appear to reflect factor price inequality if we consider a 10% level of significance in the year 2005-06.

Table 7: Bilateral Rejections by Base Zones

	2000-01						
	Bilateral Rejections						
	Total no. of Rejections	*Significance Level*		*Positive Significant Coefficients*		*Negative Significant Coefficients*	
Base Districts		*10%*	*5%*	*10%*	*5%*	*10%*	*5%*
Centre	3	3	3	0	0	3	3
North	1	1	1	1	1	0	0
West	1	1	1	1	1	0	0
South	1	1	1	1	1	0	0

Source: Author's own calculations using CMI.

Similarly testing for the individual $\beta's$ in equation (17) we see from table 7 that for 2001 for each of the base zones except when Centre is taken as the base there is at least one rejection. This suggests that one particular zone has significantly different relative factor price from the base zone for each of the regressions. Not surprisingly, that particular zone in all the regressions is the Centre and in all these regressions it has positive significant coefficient suggesting that the Centre has higher relative wages of skilled workers. When the Centre is kept as the base zone there are three rejections, the remaining three zones have significantly lower relative wage of skilled workers from that of Centre at both the 5 and 10% levels of significance. Thus the Centre on the whole has higher relative wages for skilled workers. This coincides with the results of the district level analysis.

On the other hand for 2005-06 we can see from Table 8 that the North is not showing significantly different relative wages from any other zone at any significance level whereas the Centre has significantly different relative wages of skilled workers from two other zones, the West and South. Both of these zones have negative significant coefficients suggesting that they have lower relative wages of skilled workers. When the West is taken as the base we see that Centre and North appear to have significantly higher relative wages of skilled workers compared to the West.

Table 8: Bilateral Rejections by Base Zones

Base Districts	*2005-06*						
	Bilateral Rejections						
	Total no. of Rejections	*Significance Level*		*Positive Significant Coefficients*		*Negative Significant Coefficients*	
		10%	*5%*	*10%*	*5%*	*10%*	*5%*
Centre	2	2	2	0	0	2	2
North	0	0	0	0	0	0	0
West	2	2	1	2	1	0	0
South	1	1	0	1	1	0	0

Source: Author's own calculations using CMI.

For South there is only one rejection. This particular rejection is for Centre that has significant negative coefficient suggesting that Centre has significantly higher relative wages of skilled workers compared to South also. These results suggest that relative wages of skilled workers in the Centre are higher than every other zone. The Centre West, are contributing to factor price inequality in Punjab more than the other two zones. We can see from table 4 that districts located in the Centre and West are house to a very large industrial setup, so greater manufacturing activity in these two zones is probably what has contributed to relative wage variation.

Relative wage differences and industrial structure

This section serves to investigate whether departures from RFPE are associated with differences in the set of industries that regions produce. This calls for estimating the following ordinary least square (OLS) regression:

$$Z_{rs} = \beta_o + \beta_1|\alpha_{rs}^{NA}| + \beta_2 I_r + \beta_3 I_s + \mu_{rs} \quad (18)$$

Where Z_{rs} is a measure of similarity of the industrial structure. It is the count of the number of common industries produced by two regions. $\alpha_{rs}{}^{NA}$ are the coefficient estimates of dummies from (17) and they account for the bilateral wage bill differentials between two regions, I_s is the total number of industries produced by region *r* and similarly I_r is the count of total number of industries produces by region *s*. The β_1 however in this regression needs to be interpreted cautiously as it is the measure of responsiveness between two endogenous variables. We cannot therefore comment on the magnitude of β_1 but what we can interpret is the direction of relationship which needs to coincide with the neoclassical trade theory. It states that regions producing the same set of industries are active in the same cone of diversification and therefore are characterized by relative factor price equality. Therefore we would expect a negative sign of β_1 in (18). A fall in relative wage differential between two regions must be associated with those regions producing a greater number of industries in common.

To test if our results are driven by the choice of dependent variable we have also tested equation (18) and its predictions by using another dependent variable which is the Krugman's specialization index. This is an alternate measure of industrial similarity. This index predicts industrial similarity through the employment structure of two regions. It is mathematically represented as follows

$$Z_{rs} = \sum_j \left| \left(\frac{L_{rj}}{L_r} - \frac{L_{sj}}{L_s} \right) \right| \qquad (19)$$

It is the sum of the absolute difference between industries specific shares of employment within each district where $\frac{L_{rj}}{L_r}$ is the employment share of district r's industry j in total employment of district r and likewise $\frac{L_{sj}}{L_s}$ is the employment share of district s's industry j in total employment of district s. Its value ranges from 0 to 2. The industrial similarity between two regions is decreasing in this measure i.e., as this index approaches 2 the industrial similarity between two regions falls and vice versa. Therefore the sign of β_1 is expected to be positive when using Krugmann's index as a measure of industrial similarity.

The dependent variable is discrete, non-negative and highly skewed, as is also depicted by the shape of the frequency polygon of Z_{rs} in figure 7 and 8. Z_{rs} is positively skewed for both the periods of analysis. For such non-negative skewed discrete dependent variable simple OLS produces non-integer values. Also OLS can predict negative values of the dependent variable. Therefore either negative binomial regression or Poisson regression is required to be done to test if our results are being driven by the choice of the regression analysis. The deciding factor as to which regression analysis out of these two is more suitable is a likelihood ratio test of a dispersion parameter being equal to zero in which case we will apply the Poisson regression or this parameter being significantly greater than zero in which case we will apply the negative binomial regression. The dispersion parameter is the amount by which the variance of a variable differs from mean. So for negative binomial regression we have that

$$Variance = mean + kmean^2$$

If k=0, then mean =variance and in that case there is no dispersion and therefore the Poisson is more suited. But if we have that k is significantly greater than zero then variance>mean and we have the evidence of our variable being overdispersed (underdispersed otherwise) then negative binomial regression would fit the data well.

The likelihood ratio test is done to test if dispersion parameter is equal to zero. Our hypothesis then is

$$H_0: k = 0$$

$$H_1: k > 0$$

Figure 7: Frequency polygon and Frequency Histogram for Z_{rs}: 2000-01.

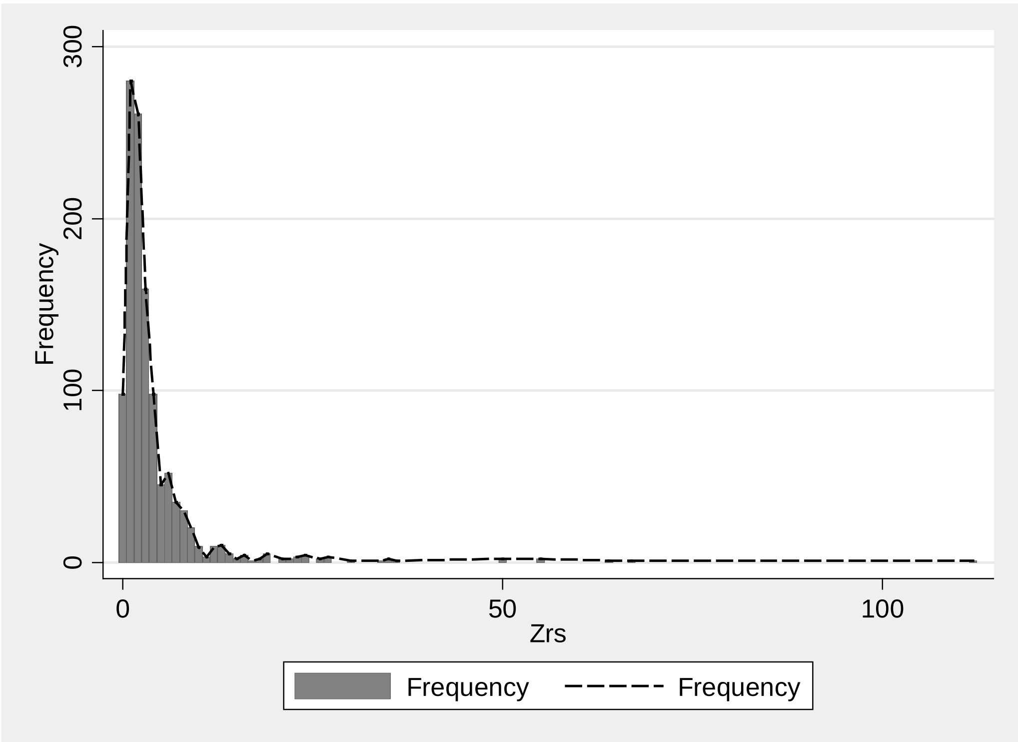
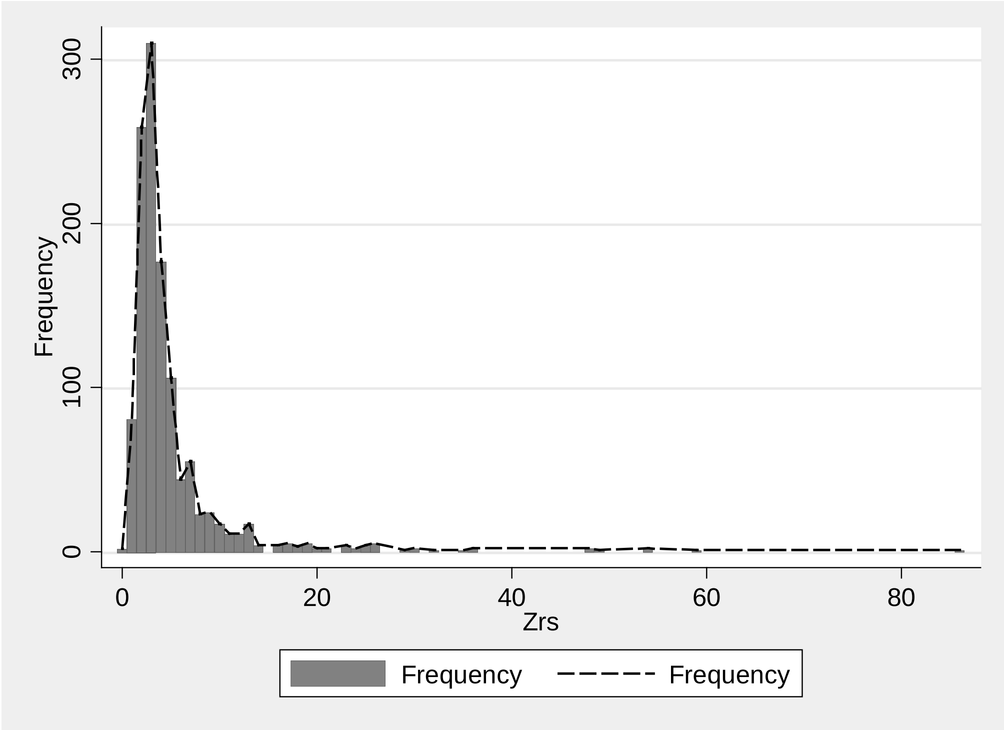


Figure 8: Frequency polygon and Frequency Histogram for Z_{rs} : 2005-06.



If the p-value of the chi-squared statistic obtained after the negative binomial regression is equal to zero then we have sufficient evidence of k being greater than zero.

However one important consideration regarding the negative binomial regression is that it transforms the dependent variable to a log of expected counts therefore the coefficient estimates of a negative binomial regression tell us the difference in the log of expected counts for a one unit change in the predictor variable. So we are required to calculate the true coefficient estimates by means of the marginal effects after the estimating negative binomial regression.

Table 9 summarizes industrial specialization across districts. The first indicator of industrial specialization is ‘regions per industry’. This particular indicator shows the extent of dispersion of an industry across Punjab or more simply, in how many districts is a particular industry located. We see from the first rows of table 9 that some of the industries are highly localized and are found in as few as just one district as of 2005-06. Some examples could be the industry for manufacturing of bodies for motor vehicles or fitting out of caravans and mobile homes that are located in Lahore only.

Table 9: Industrial Specialization across Districts

Districts	Year	Minimum	Median	Maximum
Regions per industry as per cent of all regions	2005-06	1 (2.86%)	4 (11.43%)	34 (97.14%)
	2000-01	1 (2.94%)	3 (8.82%)	27 (79.41%)
Industries per region as per cent of all Industries	2005-06	3 (3.06%)	9 (9.18%)	86 (87.76%)
	2000-01	1 (0.65%)	10 (6.45%)	112 (72.26%)
Bilateral overlap as a per cent of larger region's industries	2005-06	2.32%	22.22%	100%
	2000-01	0	14.29%	75%
Krugman's specialization index	2005-06	0.034	0.995	1.97
	2000-01	0.07	0.995	1.99

Source: Author's own calculations using CMI

Similarly ethyl alcohol and spirits were manufactured in Nakanasahib only as of 2005-06. Same is the case for 2000-01 where we see that some of the industries were being produced by just one district like industry for manufacturing of motorcycles and auto-rickshaws was localized to Sheikhupura only. Whereas on the other extreme there are some of the industries that were produced in almost entire Punjab as of year 2005-06. These include rice husking, rice and grain milling. This industry was found in all of the districts of Punjab except Bahawalnagar. Where in 2001 as well wheat and grain milling was one industry that was found in 27 out of 34 districts of Punjab. The excluded districts were DG Khan, Rajanpur, Mandibahauddin, Chakwal, Bhakkar, Khushab and Mianwali. This particular indicator is showing that over the years the industries have penetrated to a greater range of districts i.e., from a maximum of 79% of districts in 2000-01 to a maximum of 97% of districts in 2005-06.

The next two rows of table 9 display indicators of industrial diversification at the district level. The indicator 'industries per region' shows the number of industries that are active in a particular district. It can be seen that as of year 2005-06 some of the districts have only 3% of the industries that are active in entire Punjab. To be precise Narowal is only active in 3 industries out 98 industries of Punjab. This includes industries for manufacturing of taps, valves or, vacuum pumps, rice husking, rice milling, animal or vegetable oils and fats. On the other hand Lahore being most diversified was home to 86 (87%) industries out of 98 (for entire Punjab) in 2006. Similarly for 2000-01 some of the districts were active in as few as one industry out of 155. Rajanpur for example was active in only the leather footwear industry. On the other hand, some districts are as diversified as being active in 72% of industries. This particular district is Lahore which was active in 112 industries out of 155 in Punjab as a whole as of year 2000-01. This indicator is showing an increase in industrial diversification at the district level over the given period of time. The next two rows of table 9 are indicators for industrial similarity. The first indicator in this respect is the 'Bilateral overlap ratio' which just gives the ratio of industries that are common to two given districts. It is the number of common industries between two districts as a percent of the larger (in terms of number of industries) region's industries. Over the given period of time we have seen that the maximum proportion of industries that two regions share in common has gone up from 75% in 2001 to 100% in 2006 where the minimum of number of industries that any two districts in Punjab have in common has increased from no industries in common to 2% industrial similarity. This is an indication of greater industrial diversification at district level over time. The last row of table 9 is showing the trend in Krugman's specialization index over the given period of time. Krugman's specialization index similarly has shown a decreased from 0.07 in 2001 to 0.03 in 2006 which again is an indication of increasing industrial similarity over time. The maximum value of this index has also fallen though slightly from 1.99 to 1.97 which is yet another indication of growing industrial similarity over the period of five years.

EMPIRICAL RESULTS

The results for empirical estimation of equation 18 are shown in column (1) of table 10 for 2005-06 and column (1) of table 11 for 2000-01. The results we see do coincide with what neoclassical theory of trade predicts. A statistically significant negative sign on the coefficient of the absolute relative wage bill gap ($|\alpha_{rs}^{NA}|$) in column (1) of both the tables suggests that as the difference in the relative wages in a district compared to that of a base district goes up the number of industries that these two districts produce in common goes down or more technically these two districts will lie in the same cone of diversification or be active in the same set of industries if they have similar relative factor prices (relative wages in this case). Moreover over the course of time this relationship has retained its significance In column (2) of table 10 and table 11 the same regression has been performed for 2005-06 and 2000-01 respectively but now the district pair dummies have been controlled for and we see that our equilibrium relationship between the absolute relative wage bill gap and the number of industries produced in common between two districts still holds and is statistically significant as well for both the years.

The negative binomial regression analysis has also been performed in column (3) of table 10 and table 11 again for the two time periods. The Poisson regression has not been done because Z_{rs} is overdispersed as is evident from the likelihood ratio test.[3]

[3] The p-value of chibar square is equivalent to zero as is evident from table A and B of appendix 1 for both the years signifying that alpha (dispersion parameter) is significantly greater than zero thereby suggesting negative binomial regression as a better fit for the data at hand.

Table 10: Relative Wage Differential and Industrial structure, 2005-06

	(1)	(2)	(3)†	(4)†	(5)	(6)
	OLS		Negative Binomial		OLS	
	No. of Common Industries				Krugman's specialization index	
Absolute Relative	-1.402***	-1.159***	-1.505 ***	-0.868***	0.024	0.093***
Wage Bill Gap	(0.341)	(0.387)	(.207)	(.182)	(0.028)	(0.026)
Number of industries	0.164***	0.140***	0.0834***	0.688***	0.009***	0.006***
in r	(0.007)	(0.013)	(.003)	(.242)	(0.001)	(0.001)
Number of industries	0.159***	0.103*	0.0810***	0.695***	-0.002***	-0.001
in s	(0.007)	-1.159***	(.003)	(.240)	(0.001)	(0.001)
Constant	0.517*	-	-	-	0.903***	-
	(0.267)				(0.022)	
Scale Factor for Marginal Effects	-	-	4.094	3.717	-	-
Regional Dummies	No	Yes	No	Yes	No	Yes
R^2	46.06	49.5	-	-	14.9	44.7
Number of Observations	1191	1191	1191	1191	1225	1225

Notes: †The coefficient estimates are the marginal effects and not the estimated negative binomial regression coefficients. The estimates for Negative binomial regression are shown in Table B and C of Appendix 1.

Column (3) of table 10 and 11 provides the marginal effects for the negative binomial regression, and the complete results of negative binomial regression are shown in table B and C of appendix 1. The coefficient estimates of column (3) in table 10 and 11 show similar results i.e., a statistically significant negative relation between absolute relative wage bill gap and the industrial similarity, marking the fact that our results are not driven by the choice of regression technique. Column [4](4) again in both the tables 10 and 11 moreover also suggests that results of the marginal effects of negative binomial regression are robust to inclusion of region pair dummies as well, as the coefficient of absolute relative wage bill gap retains its statistical significance and direction.

44

Table 11: Relative Wage Differential and Industrial structure, 2000-01

	(1)	(2)	(3)†	(4)†	(5)	(6)
	OLS		Negative Binomial		OLS	
	No. of Common Industries				Krugmann's specialization index	
Absolute Relative	-2.007***	-1.763***	-1.612***	-1.114***	0.096***	0.103***
Wage Bill Gap	(0.348)†	(0.372)	(0.183)	(0.142)	(0.020)	(0.019)
Number of industries	0.129***	0.107***	0.056***	-0.013	0.006***	0.004***
in r	(0.007)	(0.012)	(0.003)	(0.0497)	(0.000)	(0.001)
Number of industries	0.123***	0.112***	.053***	0.310***	-0.001***	0.008***
S	(0.007)	(0.012)	(0.003)	(.051)	(0.000)	(0.001)
Constant	0.537*	-	-	-	0.851***	-
	(0.289)				(0.017)	
Scale Factor for Marginal Effects	-	-	3.087	2.524		
Regional Dummies	No	Yes	No	Yes	No	Yea
R^2	38.88	57.5			19.9	93.6
No. of Observations	1122	1122	1122	1122	1122	1122

Notes: †Standard errors in parentheses
†The coefficient estimates are the marginal effects and not the estimated negative binomial regression coefficients. The estimates for Negative binomial regression are shown in Table B and C of Appendix1.

Column (5) of table 10 and table 11 presents the estimate of equation (19) using Krugman’s specialization index as the dependent variable. As we had anticipated that the sign of the absolute relative wage bill gap using Krugman’s index as the dependent variable must be positive. The statistically significant positive sign of the absolute relative wage bill gap in column (5) of table 11 again is suggestive of the fact that industrial similarity between two districts is decreasing in the relative wage differential. For 2005-06 we are unable to get a statistically significant coefficient on the relative wage bill gap however when region pair dummies are controlled for, but we are able to show that negative relation does exist between relative wage differential and industrial similarity between two districts. Column (6) of table 11 strengthens the results of the alternate dependent variable for 2001 by showing that these are even robust to the inclusion of region pair dummies. The results retain their direction as well as their statistical significance.

Hence with the help of various specifications we have been able to demonstrate that the equilibrium relationship between absolute relative wage bill gap and the industrial similarity between two regions is in correspondence with the predictions of neo-classical theory of trade for the case of Punjab. Over two time periods i.e., 2005-06 and 2000-01 we have substantial evidence to support the view that industrial similarity does fall as the relative wage gap between two districts goes up, or in other words that regions characterized by similar relative factor prices tend to lie in the same cone of diversification.

CONCLUSION

This study has conducted an empirical analysis for Punjab to test firstly for bilateral factor price equalization in Punjab at the district as well as zonal level and secondly it also tests the proposition of the neoclassical trade theory that departures from relative factor price equality lead to differences in the mix of industries that regions produce. Ideally regions abundant in a factor exhibit a lower relative price of the factor (than regions scarce in that factor), and regions with a low relative price of a factor have a higher concentration of industries that use that factor intensively (than regions with higher relative price of that factor).

As for the factor price equality hypothesis, we have seen that as of 2000-01 there was factor price inequality in Punjab and that it mostly prevailed in the central and the western zone, as the number of bilateral rejections was very high when districts from these two zones are taken as base districts. For 2005-06 we get consistent results at the district level with most bilateral rejections in cases where districts were taken from the center as the base region. This suggests that the center has been contributing most towards the relative factor price inequality over the chosen period of time. At the zonal level results stay the same for both the years as three out of four zones reject the null for relative factor price equality at the 10% level of significance. Also results suggest that relative wages of skilled workers are mostly higher in the center suggesting skill scarcity in these particular districts. But this scarcity does not suggest that districts like Lahore, Gujranwala, Faisalabad, that have better indicators of human capital compared to the rest of the districts of Punjab have lower skill levels, it might just mean that this artificial scarcity is being created due to skilled labor force being employed in other sectors like the services sector. So the manufacturing sector hires the skilled labor at higher relative wages in these big districts that are the industrial hub of Punjab. Moreover the higher wages of skilled workers here is a reflection of the better quality of human capital so the returns to higher and better education would obviously be higher. Also the manufacturing industries of Punjab do not require very high level of skill in workers, the kind of skill (low) that they want is what they can find very easily in big districts. Moreover the demand bias is what has attracted these industries manufacturing consumer goods to settle all over Punjab irrespective of the factor supplies.

As for testing the second proposition that the results of the analysis have provided strong evidence in favor of the fact the number of common industries that two regions produce does fall down as the absolute relative wage bill gap between these two regions increases and that this hypothesis is robust to various econometric specifications.

BIBLIOGRAPHY

Berger, H. & Westermann, F. (2001). Factor Price Equalization? The Cointegration Approach Revisited.CESifo Working Paper Series 471, CESifo Group Munich.

Bernard, A. B., Robertson, R. S. and Schott, P. K. and Simpson, H.(2002). Factor price equalisation in the UK.IFS Working Papers W02/11, Institute for Fiscal Studies

Bernard, A. B., Robertson, R. S. and Schott, P. K. and Simpson, H.(2004). Relative Wage Variation and Industry Location.CEP Discussion Papers dp0614. Centre for Economic Performance, LSE.

Bernard, A. B., Redding, S. J. and Schott, P. K. (2005). Factor Price Equality and the Economies of the United States, CEPR Discussion Paper No. 5126.

Bernard et al. (2008). Relative Wage Variation and Industry Location in the UK. *Oxford Bulletin Of Economics And Statistics*. 70(4).

Bernard, A. B., Robertson, R. S. and Schott, P. K. (2010). Is Mexico A Lumpy Country?.*Review of International Economics.* 18(5). 937–950.

Bernstein, Jeffrey R. and David E. Weinstein (2002), "Do Endowments Predict the Location of Production? Evidence from National and International Data," *Journal of International Economics* 56: 55–76.

Brülhart, M. (2001). Evolving geographical concentration of European manufacturing industries. *Review of World Economics (Weltwirtschaftliches Archiv)* 137(2). 215-243.

Burgman, Todd A., and John M. Geppert. (1993). Factor Price Equalization: A Cointegration Approach. *Weltwirtschaftliches Archiv* 129(3). 472-87.

Cheema, A. Khalid, L.& Patnam, M.(2008). The Geography of Poverty: Evidence from the Punjab. *Lahore Journal of Economics.* Special Edition, 163-188.

Clark Kerr. (I954). The Balkanization of Labor, Markets," in Labor Mobility and Economic opportunity. Technology Press of M. I. T. and John Wiley and Sons.

Courant, Paul N. and AlanV. Deardorff. (1992). International Trade with Lumpy Countries. *Journal of Political Economy*100 .198–210.

Courant Paul N. & Deardorff Alan V., 1993. Amenities, Nontraded Goods, and the Trade of Lumpy Countries," *Journal of Urban Economics, Elsevier.* Vol. 34(2). 299-317.

CREB. (2006). Human Development Index for districts of Punjab. Lahore School of Economics.

Debaere, P. and Demiroglu, U. (2003). On the similarity of country endowments and factor price equalization. *Journal of International Economics* 59. 101–136.

Deardorff, A. V. (1994). The Possibility of Factor Price Equalization, Revisited. *Journal of International Economics 36* . 167-175

Deardorff, A. V. (2002). Introduction to Two-Cone HO Equilibrium. Handout for students. Retrieved from http://www.econ.lsa.umich.edu/_alandear/courses/441/handouts/]. University of Michigan.

Demiroglu, U., Yun, K.K. (1999). The lens condition for factor price equalization. *Journal of International Economics* 1.1–10.

Dumais, G., Ellison, G., and Glaeser Edward, L. (2002) Geographic Concentration as a Dynamic Process. *The Review of Economics and Statistics* 84(2).193-204.

Hanson, G. and Slaughter, M. (2002). Labour market adjustment in open economies: evidence from U.S. states. *Journal of International Economics* 57. 3–30.

Hanson, G. (2004). What has Happened to Wages in Mexico since NAFTA?. in Toni Estevadeordal, Dani Rodrick, Alan Taylor, and Andres Velasco (eds), *FTAA and Beyond: Prospects for Integration in the Americas*, Cambridge, MA. Harvard University Press.

Hanson, Gordon. (1997). Increasing Returns, Trade and the Regional Structure of Wages". *Economic Journal* 107. 113–33.

Harrigan, J. (1997). Technology, Factor Supplies, and International Specialization:

Estimating the Neoclassical Model. *American Economic Review* 87(4). 475–94.

Harrigan, J. and Zakrajsek, E. (2000), Factor Supplies and Specialization in the World Economy. Finance and Economics Discussion Series 2000-43, Federal Reserve Board of Governors.

Henry G. Overman & Diego Puga. (2010). Labor Pooling as a Source of Agglomeration: An Empirical Investigation. *Agglomeration Economics.* Ed. Edward C. Glaeser.133-150.

Hughes, G. and McCormick, B. (1994). Did migration in the 1980s narrow the North–South divide. *Economica* 61. 509–527.

John R. Moroney and James M. Walker. (1966). A Regional Test of the Heckscher-Ohlin Hypothesis. *Journal of Political Economy.* Vol. 74, No. 6. pp. 573-586.

Kemp, M. C. and Okawa, M. (1997). Factor Price Equalization under Imperfect Competition. *Review of International Economics* 5(2).153-78.

Krugman, P. (1993). First Nature, Second Nature, and Metropolitan Location. *Journal of Regional Science* 33.129-144.

Leamer, Edward, E. (1995). "The Heckscher-Ohlin Model in Theory and Practice," *Princeton Studies in International Finance* 77. 1-50.

Lerner, Abba P. (1952). Factor Prices and International Trade. *Economica.* 19.

Lowell E. Gallaway. (1963).The North-South Wage Differential. *The Review of Economics and Statistics* 45(3). 264-272.

Marshall, Alfred. 1890. *Principles of Economics*. London. Macmillan.

Reeve, T. A. (2006). Factor Endowments and Industrial Structure. *Review of International Economics 14(1). 30–53.*

Todd, G. and Abel, J. (2009).Labor Market Pooling and Occupational Agglomeration. Federal Reserve Bank of New York, Staff Report 392.

Samuelson, Paul A. (1949). International Factor-Price Equalisation Once Again. *Economic Journal* 59. 181-197.

APPENDIX

Table A : Human Development Index of Districts of Punjab (2006)

District	District Rank	HDI
Rawalpindi	1	0.799
Lahore	2	0.785
Jhelum	3	0.762
Sialkot	4	0.751
Gujranwala	5	0.748
Gujrat	6	0.745
Chakwal	7	0.740
Mandi Bahauddin	8	0.718
Attock	9	0.718
Narowal	10	0.709
Faisalabad	11	0.707
TT Singh	12	0.695
Sargodha	13	0.694
Hafizabad	14	0.690
Sheikhupura	15	0.681
Mianwali	16	0.673
Nankana Sahib	17	0.673
Khushab	18	0.673
Multan	19	0.666
Sahiwal	20	0.663
Okara	21	0.653
Layyah	22	0.649
Kasur	23	0.647
Khanewal	24	0.642
Bhakkar	25	0.641
Jhang	26	0.638
Vehari	27	0.634
Bahawalnagar	28	0.628
Pakpattan	29	0.621
DG Khan	30	0.611
Muzaffargarh	31	0.602
Lodhran	32	0.600
Bahawalpur	33	0.596
RY Khan	34	0.594
Rajanpur	35	0.555

Source: CREB (2006), Lahore School of Economics

Table B: Negative Binomial Regresson for 2005-06

```
Negative binomial regression                      Number of obs   =       1170
                                                  LR chi2(3)      =    1065.32
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -2519.7364                       Pseudo R2       =     0.1745
```

zrs	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
absrs	-.4482479	.043996	-10.19	0.000	-.5344784	-.3620173
ir	.0193087	.0008033	24.04	0.000	.0177343	.0208831
is	.0189055	.0007976	23.70	0.000	.0173423	.0204688
_cons	1.029681	.0354512	29.04	0.000	.9601976	1.099164
/lnalpha	-2.353893	.1055842			-2.560834	-2.146952
alpha	.0949986	.0100303			.0772403	.1168397

Likelihood-ratio test of alpha=0: chibar2(01) = 221.52 Prob>=chibar2 = 0.000

Table C: Negative Binomial Regression for 2000-01.

```
Negative binomial regression                      Number of obs   =       1040
                                                  LR chi2(3)      =     891.03
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -2211.4956                       Pseudo R2       =     0.1677
```

zrs	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
absrs	-.8997453	.0690056	-13.04	0.000	-1.034994	-.7644969
ir	.0165968	.00083	20.00	0.000	.0149701	.0182235
is	.0163371	.0008301	19.68	0.000	.0147101	.0179642
_cons	.980034	.0424892	23.07	0.000	.8967567	1.063311
/lnalpha	-1.666446	.0897509			-1.842354	-1.490537
alpha	.1889173	.0169555			.158444	.2252516

Likelihood-ratio test of alpha=0: chibar2(01) = 429.16 Prob>=chibar2 = 0.000