CREB Working Paper No. 02-16

Relative Wage Variation and Industry Location in Punjab

Zunia Saif Tirmazee

Centre for Research in Economics and Business
Lahore School of Economics

Centre for Research in Economics and Business (CREB)

Lahore School of Economics

Intersection Main Boulevard Phase VI, DHA and Burki Road
Lahore 53200, Pakistan
Tel: 042-36561230; 042-36560936
Email: creb@lahoreschool.edu.pk

CREB Working Paper No. 02-16

Relative Wage Variation and Industry Location in Punjab

Zunia Saif Tirmazee
Teaching Fellow, Lahore School of Economics

First printing March 2016.

Lahore School of Economics
Intersection of Main Boulevard, Phase VI, DHA, and Burki Road
Lahore 53200, Pakistan
Tel.: +92 42 3656 1230
www.creb.org.pk
creb@lahoreschool.edu.pk

Price: PRs100

Preface

The Centre for Research in Economics and Business (CREB) was established in 2007 to conduct policy-oriented research with a rigorous academic perspective on key development issues facing Pakistan. In addition, CREB (i) facilitates and coordinates research by faculty at the Lahore School of Economics, (ii) hosts visiting international scholars undertaking research on Pakistan, and (iii) administers the Lahore School's postgraduate program leading to the MPhil and PhD degrees.

An important goal of CREB is to promote public debate on policy issues through conferences, seminars, and publications. In this connection, CREB organizes the Lahore School's Annual Conference on the Management of the Pakistan Economy, the proceedings of which are published in a special issue of the Lahore Journal of Economics.

The CREB Working Paper Series was initiated in 2008 to bring to a wider audience the research being carried out at the Centre. It is hoped that these papers will promote discussion on the subject and contribute to a better understanding of economic and business processes and development issues in Pakistan. Comments and feedback on these papers are welcome.

Abstract

This study tests the wage differential between nonproduction (white-collar) and production (blue-collar) workers across districts of Punjab; it also seeks to determine whether these variations affect the industrial structure of these districts for the periods 2000/01 and 2005/06. Such an analysis is important because (i) it yields information on the income convergence in a country and (ii) it points out that policies targeted at regional development must take into account data on factor prices. For instance, when addressing equity concerns, the government must focus on enhancing the skills of workers in regions where skill levels are low – giving them a better chance of finding jobs – rather than setting up blue-collar-intensive industries here since any wage differential would be arbitraged away when labor is perfectly mobile.

The findings from the first stage of the analysis suggest there is evidence of relative wage inequality in Punjab: generally, nonproduction workers earn higher relative wages in central Punjab and lower relative wages in southern and western Punjab. These findings are consistent over the five-year period 2001–06. The second stage of the analysis shows that the industrial mix also varies across Punjab as districts differ in terms of the wages they offer both nonproduction and production workers.

Relative Wage Variation and Industry Location in Punjab

1. Introduction

The aim of this study is to determine whether relative factor price inequality leads to dissimilarities in the industries located across selected regions. Bernard, Redding, Schott and Simpson (2002) give four reasons for why – even with trade – relative factor prices should vary across regions: (i) multiple cones of diversification, (ii) region-industry technology differences, (iii) agglomeration, and (iv) increasing returns to scale. This paper deals specifically with variations in factor intensity (multiple cones of diversification) as an explanation for relative factor price equality (RFPE). Regions characterized by RFPE have more industries in common.

This paper tests the proposition that the larger the difference in the relative wages of two regions, the fewer industries they will have in common. This geographic variation in workers' relative wages – and, therefore, in industrial production – is rationalized by the neoclassical theory of trade, which suggests that regions with an abundance of a particular factor have more industries that use this factor intensively than regions where it is scarce.

The literature indicates that there is no single determinant of industry location. Factor supply and price determine a region's comparative advantage in production and, in turn, where industries decide to locate. Importantly, industries cannot exist in isolation because this would disregard the connection between markets and industries. These linkages are another important determinant of industry location, so much so that an entire subdiscipline of economics, new economic geography, is devoted to understanding how industries choose to locate.

Testing for RFPE is problematic because one needs to account for unobserved variations in factor quality, especially with regard to labor. Workers' quality across regions can vary due to differences in the education system or availability of training programs for workers, etc.

These differences in turn affect the productivity of workers and hence the relative wages of (nonproduction) workers.

Following the methodology used by Bernard, Redding, Schott and Simpson (2008), this study looks at variations in the wages of nonproduction or white-collar workers relative to production or blue-collar workers across districts of Punjab; it then examines how these variations affect the industries located in these districts for the periods 2000/01 and 2005/06. The key advantage of this methodology is its potential to control for unobservables such as factor quality.

Such an analysis is important for two reasons. First, it yields information on income convergence in a country. Second, it has public policy implications, as Bernard et al. (2008) suggest. Policies targeting regional development need to take into account information on factor prices and hence the regional comparative advantage because relocating white-collar-intensive industries to lagging regions could result in a comparative cost disadvantage to such firms (as the neoclassical theory of trade predicts). Other policies that aim to boost workers' skills in lagging regions through education and training are better alternatives in the long run. Additionally, policies that are directed at regions rather than individuals generally prove futile, especially when labor is perfectly mobile.

We test for RFPE in Punjab using a simple ordinary least squares (OLS) model whereby the ratio of the relative wage bill of nonproduction workers is regressed on 35 district dummies. The estimated coefficients thus obtained provide a measure of relative factor price inequality. Next, the RFPE estimates generated from the first regression are used to test whether factor price inequality leads to differences in the set of industries that regions (in our case, districts) produce. This is done by regressing an index of industrial similarity between two districts on the RFPE estimates.

The study's findings in the first stage suggest there is evidence of factor price inequality in Punjab: nonproduction workers generally earn higher relative wages in central Punjab and lower relative wages in southern and western Punjab. These findings are consistent over the five-year period 2001–06. The second stage of the analysis shows that the industrial mix does vary across Punjab as regions differ in terms of the wages they offer both nonproduction and production workers.

Moreover, this finding is robust to various econometric techniques and across the five-year period 2001–06.

2. Literature Review

What determines industry location in a particular area? Brülhart (2001) summarizes three schools of thought in responding to this question. The first, the neoclassical trade model, deals solely with factor proportions, factor endowments, and the technology stock of individual regions. Thus, whether industries locate in a particular place depends on their comparative advantage. The second school of thought on new trade theory maintains that industry structure is determined by the "first-nature" (Krugman, 1993) characteristics of a region – such as land and differences in accessibility – and by "second-nature" characteristics or 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. Finally, the third school of thought, new economic geography, takes into account only the second-nature characteristics of new trade theory.

This paper focuses on all those determinants that relate to the labor market. In the case of neoclassical trade theory, we are concerned with determinants such as relative factor endowment or the factor intensity of production. Where new economic geography and new trade theory models are concerned, we take into account factor endowments (first-nature characteristics) and labor market pooling (second-nature characteristics).

2.1. Factor Price Equalization

One of the most important results of the neoclassical theory of international trade is the factor price equalization (FPE) theorem. This suggests that all regions producing a similar product mix, using similar technologies and with the same product prices, will have the same factor prices if certain assumptions hold (see Lerner, 1952). These stylized assumptions are as follows. First, regions have access to identical technologies, which is another way of saying that they have the same production function. Second, they produce the same products and do not specialize entirely in any one product, that is, their factor intensities are the same. Third, we assume constant returns to scale.

Fourth, we assume perfect competition.[1] Finally, the regions are ranked by factor intensity in the same order at all factor prices, or more technically, there no factor intensity reversals.

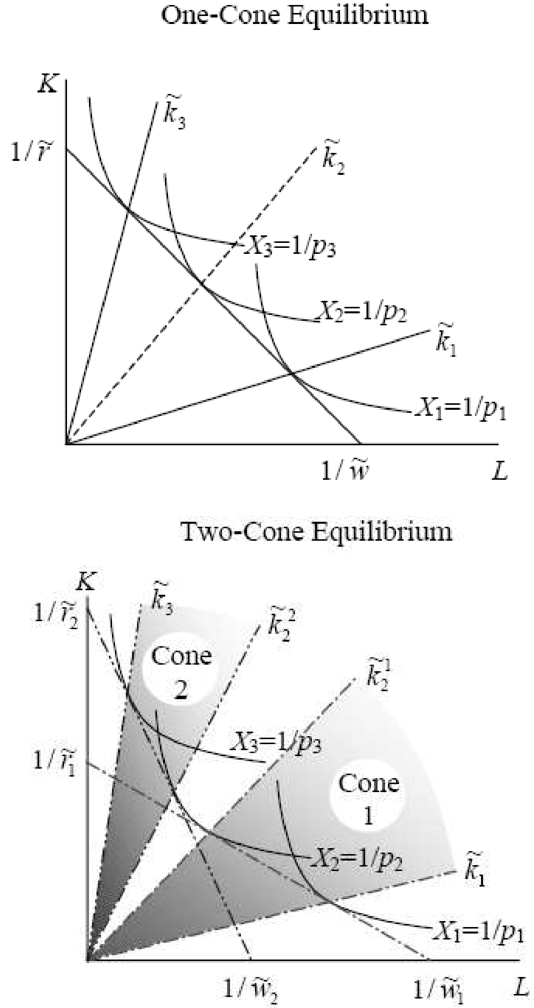
Beyond these given conditions, the role of factor endowments must not be overlooked. Samuelson (1949) notes that, apart from similar factor intensities and commodity prices, regions must have similar factor endowments to qualify for FPE. For factor endowments that are sufficiently dissimilar, it is impossible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

The literature on trade explains the Heckscher–Ohlin model through the framework of cones of diversification, where a "cone" is the range of endowment vectors that select the same set of industries (Figure 1). Deardorff (2002) uses the standard Lerner (1952) diagram to explain the interaction between diversification and factor prices under varying factor endowments. The idea is explained on an x–y plane where each axis represents units of a factor of production. Deardorff shows a transition from complete factor price equality to factor price inequality through two diagrams, one showing a one-cone equilibrium (the case for RFPE) and the other a two-cone equilibrium (the case for relative factor price inequality).

In the analysis that follows, let K represent the number of nonproduction (white-collar) workers and L the number of production (blue-collar) workers. The term r is the wage offered to nonproduction workers and w is the wage offered to production workers. The intercept $1/r$ represents the number of nonproduction workers needed to produce a good if only nonproduction workers were employed. The intercept $1/w$ represents the corresponding case for production workers.

[1] Kemp and Okawa (1997) show that FPE can occur under certain conditions, even in the presence of imperfect competition.

Figure 1: Cones of diversification



Source: Deardorff (2002).

As one moves toward the y-axis, goods become more white-collar-intensive; the farther one moves from the y-axis, the less white-collar-intensive (and the more blue-collar-intensive) they become. The lines that emanate from the origin represent various factor intensities: the steeper they are, the greater is the skill-intensity ratio (the number of nonproduction / the number of production workers). For instance, $\widetilde{k_3}$ represents higher skill intensity than $\widetilde{k_2}$.

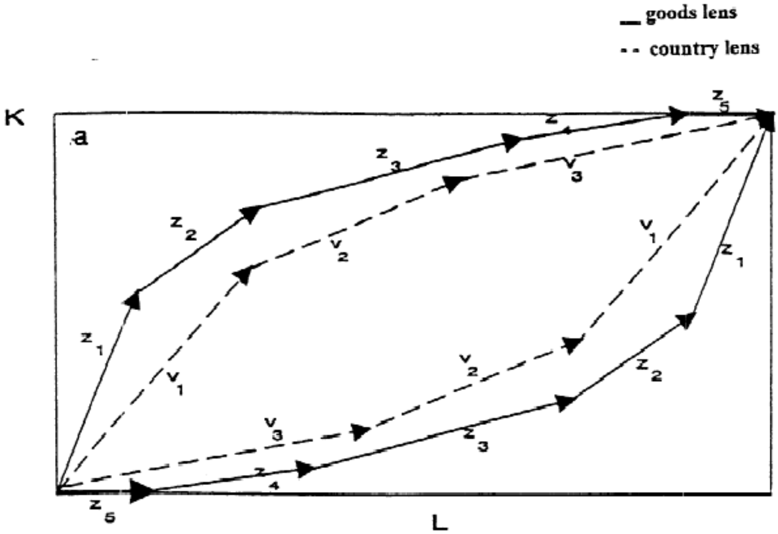
The panel for a one-cone equilibrium represents a world in which the production technologies for each good are all aligned perfectly with a single isocost line, that is, nonproduction workers earn the same relative wages throughout. Given the composition of their existing workforce, all regions hold to a unique relative wage for nonproduction workers for their respective skill-intensity ratios. Thus, the entire mix of goods lies within a single cone of diversification.

In the panel for a two-cone equilibrium, the production method for good 2 is not identical to that for the other two goods. In such circumstances, there is no unique relative wage for nonproduction workers that leads to all three goods being produced optimally. Instead, two relative wage rates emerge (for nonproduction workers) that lead to the optimal (cost-minimizing) production of good 1 and good 2, or good 2 and good 3 separately.

Since all 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. While there is FPE within each cone, it does not hold as we move to the next cone. Thus, the idea that the number of common industries that regions produce increases with a fall in the relative wage differential lies in the Lerner diagram.

Empirical analyses use various ways of testing for FPE. Under Deardorff's (1994) "lens condition," for example, the variation in factor intensities (goods lens) across regions must be greater than the cross-regional variation in factor endowments (country lens). The lens condition can be visualized in the shape of a box diagram, the dimensions of which represent the factor endowments of a country (Figure 2). For the lens condition to hold, the goods lens must lie outside the country lens to enable FPE. This would suggest that the endowments of various regions are similar enough for the same goods to be produced. Numerous studies use the lens condition to test for FPE, notably Bernard, Robertson, and Schott (2010), Demiroglu and Yun (1999), and Debaere and Demiroglu (2003).

Figure 2: Lens condition satisfied



Source: Deardorff (1994).

Another method of testing for FPE is the cointegration approach pioneered by Burgman and Geppert (1993), who find proof of FPE in the long-run stationary cointegrating relationship among stationary series of unit labor costs for six major industrialized countries (Canada, France, Germany, Japan, the UK, and the US). Berger and Westermann (2001) – who use the same dataset, but control for sample bias and use real instead of nominal data on unit labor costs – find only limited evidence of factor price co-movement.

Originally, tests for FPE looked at cross-country evidence. Another, more recent, line of research emerges in Bernard et al. (2008) (for the UK), Bernard, Redding and Schott (2005) (for the US), and Bernard et al. (2010) (for Mexico). The procedure followed is fairly simple and involves regressing the ratio of the wage bills for nonproduction and production 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 the regional dummies then determines whether RFPE holds. If the coefficient is significantly different from 0, one can reject the null hypothesis of RFPE. All the studies cited above show that RFPE does not hold.

FPE has implications for production patterns in the context of the neoclassical model of trade, which posits that relative factor price inequality leads to regions producing different industries. All those regions that differ with respect to their factor prices will also differ in terms of the industries (goods) they produce. Put more technically by 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 seeks to answer the following questions:

- What factors support the phenomenon that a regional similarity in patterns of production is determined by relative factor prices?
- Assuming the hypothesis above holds true, what forces sustain this phenomenon, that is, why are industries able to benefit from the relative factor price differential (in this case, the relative wage differential)? Why is this differential not arbitraged away?

2.2. Impact of Relative Wage Differential on Patterns of Production

The simplest answer to the first question is that, under the neoclassical model of trade, regions that offer white-collar workers low relative wages have a higher concentration of industries that employ white-collar workers more intensively than regions offering them high relative wages (Bernard et al., 2008). Simply put, factor prices vary across regions due to differences in their respective factor endowments. The abundant factor in a particular region will have a lower relative factor price than other factors. The endowment – and, therefore, the relative price – of a factor thus determine the pattern of production across regions.

Few studies have put forward direct evidence of the impact of the relative wage differential on patterns of production. For the most part, the literature finds a relationship between input abundance and input intensity, which is simply another way of addressing the question at hand: a similar input intensity across regions also suggests that they produce common industries.

Bernard, Redding, Schott, and Simpson (2004) provide direct evidence of this in explaining the phenomenon for a two-factor (skilled and unskilled workers) and three-good (computers, machinery, and textiles) economy on the Lerner diagram. The study shows that large differences in the factor endowments of skilled and unskilled labor lead regions to fall within separate cones of diversification. In examining the two cones of diversification (one for relatively skilled labor-intensive goods and the other for relatively unskilled labor-intensive goods), they find 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 the Heckscher-Ohlin model.

An empirical analysis of the effect of the relative wage differential between skilled and unskilled workers on production patterns is carried out using data from the 1970s to the 1990s for 209 four-digit manufacturing industries across 63 counties and Scottish regions in the UK. The analysis itself is a simple OLS regression where the number of common industries between two regions is regressed on the estimates of the relative wage bill gap. These, in turn, are the coefficient estimates of a linear regression of the relative wage bill estimates of skilled and unskilled workers on the regional dummies.

Overall, the study provides strong evidence to support the hypothesis that the UK's regional production structure reflects a multiple-cone general equilibrium trade model. Bernard et al. (2005) and Bernard et al. (2010) carry out similar analyses of the US and Mexico, respectively, both studies confirming the multiple-cone general equilibrium hypothesis.

Hanson and Slaughter's (2002) empirical analysis shows that FPE across states in the US is consistent with similar production techniques or factor intensities. Using data on the unit factor requirements of 14 large states and 40 sectors across all industries for 1980 and 1990, they find that production techniques are very similar across large US states, especially among neighboring states or those with similar relative labor endowments.

While their analysis determines whether the unit factor requirements are similar across states, FPE can also occur in regions with two different production techniques due to "scale effects, externalities, or differences in underlying technology." The authors thus test for productivity-adjusted FPE using fixed effects to check for each state pair, labor type, and year combination. A joint test of significance for 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 argue that, since "industry-neutral state productivity differences allow for nominal state wage differences," they look for evidence of FPE using data on state industry production techniques rather than direct data on state wages. Therefore, their implicit null hypothesis is that productivity-adjusted wages are equalized across the sample states.

Debaere and Demiroglu (2003) look for variations in factor endowments (labor and capital) in a country and their effect on diversification in production. Using data for 28 manufacturing sectors and 28 countries, they construct a lens of factor endowments and factor intensities, and conclude that developed and developing countries do not lie within the same cone of diversification. This implies the absence of FPE because most developing countries are concentrated in the corner of the lens, which violates the lens condition, whereas most OECD countries lie in the same cone. In a lens diagram, all countries whose endowment vectors lie within the lens of factor intensities lie in the same cone of diversification, indicating FPE.

The authors check for the robustness of their results by adjusting the data for productivity differences. However, even after correcting for productivity differences, the sample of developing countries continues to violate the lens condition, thus confirming the finding that developed and developing countries do not lie in the same cone of diversification. Moreover, when the analysis is applied to data for skilled and unskilled labor in selected OECD countries, the results are similar: most OECD countries still lie within the same cone.

Reeve (2006) demonstrates the importance of factor endowments in explaining the structure of production, and estimates to what extent factor accumulation determines changes in industrial structure. The study uses a cross-sectional dataset on the 15 largest manufacturing industries in 20 OECD countries for the period 1970–90. Its findings have policy implications, in particular for investment in physical and human capital. Such factor accumulation affects patterns of production in the same way as one would expect in the case of sector-specific trade and industrial policies.

The author estimates a model of factor proportions, using data on five factors: arable land, capital stock, and three categories of educated labor (high, medium, and low). The relative endowment of each factor is measured by the ratio of its share of the sample to its share of GDP. In order to account for compositional changes in industrial production, the study uses the percentage change in industry output as a fraction of total manufactures.

Regressing compositional changes in industrial production on relative factor endowments, Reeve (2006) interprets the estimated coefficients – estimated using one-step and iterated generalized least squares – in terms of the comparative advantage of a factor in an industry. In order to determine the extent to which changes in production structure are caused by changes in factor endowments, the author tests the null hypothesis of constant coefficients of factor endowments over time (using the Wald test). Any significant changes in techniques of production – implying that the null hypothesis is rejected – as measured by the coefficients of factor endowments over time are potentially important sources of change in industrial structure.

Harrigan (1997) tests the significance of factor endowments in explaining industrial structure. The study estimates three equations to predict industry output as a function of factor endowments. Two of these regressions cannot be used to forecast changes in industrial production over time: country fixed effects with constant coefficients and parameters that vary over time following a random walk. The author concludes that, overall, the model does not do a very good job of determining industry location.

In a similar study, Harrigan and Zakrajsek (2000) test the Heckscher-Ohlin model's prediction that regions abundant in a particular factor tend to specialize in production that uses that factor intensively. Their empirical analysis employs a panel dataset on industrial specialization and factor endowment differences for a broad sample of developed and developing countries for 1970–92. Allowing for technological differences, their results clearly show the importance of factor endowments in specialization. Bernstein and Weinstein (2002) use a cross-section of international and regional data to demonstrate that factor supplies affect gross output and that this effect is more evident internationally than intra-nationally.

Moroney and Walker (1966) use an entirely different methodology to test the Heckscher-Ohlin model at a regional level. They test the hypothesis that ordering commodities by their factor intensity ratios is equivalent to ordering regions by their comparative advantage. What this implies is that the varying factor intensities of a region's production are indicative of the varying comparative cost advantage. Thus, industries that require a relatively low capital–labor ratio for manufacturing are concentrated in regions with a lower capital endowment.

The authors test this hypothesis using industry-level data on capital–labor ratios and location quotients[2] for the southern US (the east-south central, southern Atlantic, and west-south central census regions), which is relatively labor-abundant, and the non-South, which is relatively capital-abundant. The results of a rank correlation test indicate a correlation between the industry capital–labor ratio and location quotients. After some experimentation with the data, the study establishes a negative correlation between the two measures, confirming the Heckscher-Ohlin hypothesis for these regions. Overall, the South – being labor-abundant – specializes in the production of goods that are labor-intensive relative to the non-South.

Most studies address the question of a persistent wage differential in terms of the perfect or imperfect mobility of workers. Based on a sample of workers in Mexico, Bernard et al. (2010), citing Hanson (2004), find there is insufficient migration to arbitrage away the persistent wage differential. Bernard et al. (2008) argue that the stability of their estimates for relative factor price inequality is due partly to imperfect labor mobility – a hypothesis that finds empirical support in Hughes and McCormick (1994). Moreover, even under perfect mobility, the relative wage differential of skilled and unskilled workers seems to persist because workers are reluctant to migrate in the face of "factor-region-specific amenities or living costs." If the cost of migration is greater than its expected gains, then workers will choose not to migrate in order to benefit from the wage differential.

Deardorff (1994) puts forward a similar argument, emphasizing the importance of "something else" other than the wage differential that affects where workers decide to live. He cites the "prices of non-traded goods and utility-relevant facilities" as factors that compel mobile labor to prefer one region to another.

---

[2] A measure of regional output concentration by industry.

Kerr (1954) has explained this idea by identifying five hurdles to free labor mobility: the preferences of individual workers and individual employers, and the actions of the community of workers, employers, and the government. In the context of this study, only the first two are relevant where, under free labor mobility, workers may choose not to migrate because of personal preferences, the loss of their kinship network, the higher cost of living in an industrial hub, and the direct costs of migration. This implies choosing not to take advantage of the relative wage differential between regions.

What this study seeks to establish is that, if two regions face the same good prices and their technologies and factor endowments are similar enough, then they will lie in the same diversification cone and have the same factor prices. Alternatively, if their factor prices are similar enough, they should also lie in the same diversification cone.

2.3. New Economic Geography

It is important to establish that industries do not exist in isolation. Rather, there are interconnecting forces that become significant in determining industry location. Such models offer an alternative to the neoclassical framework: they assume increasing returns to scale and imperfect competition, and unlike the neoclassical theory of trade, allow for factor mobility.

To account for the determinants of industry location in the context of labor markets as predicted by new trade theory and new economic geography, we focus on labor market pooling. This implies that industries are more likely to locate in areas where they will find workers with the desired skills, and that workers will move to areas that require their skills. Labor market pooling leads to industry clusters, thereby favoring the creation of pools of specialized workers who acquire the cluster-specific skills valued by firms.

Marshall (1890) has emphasized that industries prefer to locate specifically in areas that offer a constant market for skills. Dumais, Ellison and Glaeser (2002) test the importance of Marshallian externalities on a firm's decision to locate, which they find is driven more by labor market pooling than any other explanatory variable, including knowledge spillovers and forward and backward linkages.

Overman and Puga (2010) use establishment-level data from the UK's Annual Respondent Database to empirically assess the importance of labor market pooling as a source of agglomeration economies. By regressing a measure of spatial concentration on a measure of the potential for labor pooling,[3] the study examines whether sectors with a higher potential for labor pooling experience greater spatial concentration even after controlling for other industry characteristics (including the importance of localized intermediate suppliers). The authors conclude that labor pooling does determine where industries choose to locate.

Gabe and Abel (2009) use census data on 468 occupations across 33 knowledge areas in the US to examine how agglomeration activity is affected by labor pooling (measured by specialized knowledge),[4] among other controls, including the use of specialized machinery and knowledge spillovers. When the importance of labor market pooling doubles, the agglomeration index (locational Gini coefficient) rises by about 40 percent. In a separate regression, they find that occupations with similar knowledge requirements tend to co-locate. This is measured by regressing dissimilar knowledge on occupational agglomeration. Such behavior confirms Marshall's (1890) notion of a "constant market for skill." Both sets of results provide evidence for the importance of labor pooling as a determinant of industry location.

3. Methodology

In order to determine any departures from FPE, we follow the methodology adopted by Bernard et al. (2005). This serves two purposes. First, it determines if there is any inequality in the relative wages of white-collar (nonproduction) and blue-collar (production) workers across districts of Punjab. Second, it establishes how relative factor prices and their geographic variations are associated with industrial structure.

[3] The average of the difference between the percentage change in the firm's employment and the percentage change in the sector's employment.

[4] The extent to which an occupation's knowledge profile differs from that of the average US job.

3.1. Relative Demand for Nonproduction and Production Workers

The test we employ assumes "cost minimization, constant returns to scale, Hicks-neutral technology differences and weak separability of the production technology in white-collar and blue-collar workers" (Bernard et al., 2005). This methodology controls for unobserved factor-region-industry heterogeneity in the quality of factors.

Consider a production technology defined over two factors of production: white-collar or nonproduction workers (*N*) and blue-collar or production workers (*A*). Although this analysis considers only two factors of production and many industries, it can be generalized to as many industries and factors of production as needed. The production function for region *r* and industry *j* is:

$$Y_{rj} = \mathrm{T}_{rj} f_j\left(A_{rj}; N_{rj}\right)$$

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 production and nonproduction workers, respectively, in region *r* and industry *j*. The production function f_j is industry-specific but is similar for a particular industry for all regions. Technology differences across disaggregated products within industries are estimated by the differences in A_{rj}.

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

$$\underset{A_{rj}:N_{rj}}{min} .: w_A A_{rj} + w_N N_{rj} \tag{1}$$

such that

$$Y_{rj} = \mathrm{T}_{rj} f_j\left(A_{rj}; N_{rj}\right) \tag{2}$$

where w_A and w_N are the quality-adjusted wages of production and nonproduction workers, respectively. Assuming Hicks-neutral technology differences by region and constant returns to scale, the minimized unit cost function is defined as:

$$C = \mathrm{T}_{rj}^{-1} \lambda_j (w_r^A, w_r^N) Y_{rj} \tag{3}$$

λ_j is a function homogenous of degree 1 that varies across industries. Cost minimization means that firms in a perfectly competitive market take prices as given or, under imperfect competition, decide prices according to their downward-sloping demand curve. This analysis assumes constant returns to scale but "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" (Bernard et al., 2008).

The quality-adjusted employment level and factor wages in region *r* equal the quality adjuster times the observed variable:

$$z_{rj} = \theta^z_{rj}\widetilde{z_{rj}} \text{ and } w^z_{rj} = \widetilde{w^z_{rj}}\theta^z_{rj} \tag{4}$$

where $z \in (A, N)$ and indexes the factors of production, $\widetilde{z_{rj}}$ denotes region- and industry-specific observed quantities of the factors of production unadjusted for quality, $\widetilde{w^z_{rj}}$ represents the region- and industry-specific observed wages unadjusted for quality, and θ^z_{rj} is a quality adjustment parameter for industry *j*, region *r* and factor *z*, allowing for unobserved variations in quality and equal to 1.

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

$$z_{rj} = T^{-1}_{rj}\frac{Y_{rj}\partial\lambda_j(.)}{\partial w^z_r} \tag{5}$$

Dividing the two first-order conditions obtained from Shephard's lemma (one for production workers and one for nonproduction workers), the relative demand for the quality-adjusted quantities of production and nonproduction workers is given by

$$\frac{N_{rj}}{A_{rj}} = \frac{\partial\lambda_j(.)\big/\partial w^N_r}{\partial\lambda_j(.)\big/\partial w^A_r} \tag{6}$$

The term for region-industry productivity A_{rj} does not appear in the equation above because Hicks-neutral technology differences affect the marginal revenue product identically for every factor.

Similarly, using Shephard's lemma and equation (4), the observed relative demand for nonproduction and production workers is given by

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

3.2. Null Hypothesis for RFPE

The null hypothesis states that all relative factor prices are equalized across regions. Thus, the quality-adjusted relative wages and factor use across regions *r* and *s* must be equal such that

$$\frac{w_r^N}{w_r^A} = \frac{w_r^N}{w_r^A} \text{ and } \frac{N_{rj}}{A_{rj}} = \frac{N_{sj}}{A_{rj}} \tag{8}$$

Therefore, to reject RFPE, we would need to ascertain that the relative wages of white-collar and blue-collar workers differ across regions by using the data available for observed nonproduction and production workers' wages. However, any unobserved variations in factor quality might suggest relative factor price inequality even if the true quality-adjusted relative wages are equal across regions.

Under H_0 for RFPE, observed relative wages and observed factor employment (not adjusted for quality) across regions are given by:

$$\frac{\widetilde{w_r^N}}{\widetilde{w_r^A}} = \frac{\theta_{rj}^{N}\widetilde{w_s^N}}{\theta_{rj}^{A}\widetilde{W_s^A}} \text{ and } \frac{\widetilde{N_{rJ}}}{\widetilde{A_{rJ}}} = \frac{N_{sj}\big/\theta_{rj}^{N}}{\widetilde{A_{sJ}}\big/\theta_{rj}^{A}} \tag{9}$$

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

Multiplying wages and employment in this way eliminates the region-industry-factor quality adjustment parameters. As a result, the observed relative wage bills are equal. Using these, the null hypothesis now becomes

$$\frac{\widetilde{N_{rJ}}\widetilde{w_{rJ}^N}}{\widetilde{A_{rJ}}\widetilde{w_{rJ}^A}} = \frac{\widetilde{N_{sJ}}\widetilde{w_{sJ}^N}}{\widetilde{A_{sJ}}\widetilde{w_{sJ}^A}} \tag{10}$$

This implies that the relative wage bills are equal across regions and that differences in unobserved factor quality cause relative employment and relative wages to vary across regions.

Under the alternative hypothesis of relative factor price inequality, observed relative wage bills vary across regions due to unobserved differences in factor quality and differences in quality-adjusted wages. The latter causes the relative unit factor input requirements across regions to differ. The alternative hypothesis of non-RFPE states that the quality-adjusted relative wage (w^N/w^A) differs across regions *r* and *s* by a factor γ_{rs}^{NA}, which denotes the difference in quality-adjusted relative wages:

$$\frac{w_r^N}{w_r^A} = \gamma_{rs}^{NA}\frac{w_r^N}{w_r^A} \tag{11}$$

Region s is the benchmark region, and $\gamma_{rs}^{NA} = \gamma_r^{NA}/\gamma_s^{NA}$ and $\gamma_s^{NA} = 1$. Observed relative wages differ across regions due to unobserved variations in factor quality and variations in quality-adjusted wages. This could also be a result of sampling errors or irregularities in the response rate in different regions.

$$\frac{\widetilde{w_r^N}}{\widetilde{w_r^A}} = \gamma_{rs}^{NA}\frac{\theta_{rj}^N\widetilde{w_s^N}}{\theta_{rj}^A\widetilde{w_s^A}} \tag{12}$$

Similarly, observed factor employment now differs across regions because of differences in factor quality and factor demand caused by the variation in relative wages adjusted for quality:

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

Again, combining relative factor use and relative wages by multiplying equations (12) and (13), the terms for unobserved factor quality cancel out. However, the relative wage bills now differ across regions due to unobserved variations in factor quality:

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

where

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

and 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 within the braces in equation (15) captures differences in unit factor input requirements. From equation (14), finding that $\Phi_{rsj}^{NA} \neq 1$ is enough to reject the null hypothesis of RFPE.

3.3. Econometric Specification for RFPE

Under the null of RFPE, within an industry across regions, the ratio of the wage bills for white-collar to blue-collar workers is the same. This implies that, for industry *j*, the relative wage bill for region *r* equals the value of that for any base region *s*. In testing for RFPE by allowing every individual region to be the base region, we can avoid any bias in the outcome that might otherwise result from our choice of region. The equation to be estimated is:

$$ln\frac{RWB_{rj}^{NA}}{RWB_{sj}^{NA}} = \sum\alpha_{rs}^{NA}d_r + \varepsilon_{rsj}^{NA} \tag{16}$$

The dependent variable $\frac{RWB_{rj}^{NA}}{RWB_{sj}^{NA}} = \frac{\widetilde{N_{rJ}}\widetilde{w_{rJ}^{N}}}{\widetilde{A_{rJ}}\widetilde{w_{rJ}^{A}}}\Big/\frac{\widetilde{N_{sJ}}\widetilde{w_{sJ}^{N}}}{\tilde{A}_{sj}\tilde{w}_{sj}^{A}}$ is the ratio of the relative wage bill for nonproduction workers in region *r* with respect to a base district *s*, d_r is a set of regional dummies, α_{rs}^{NA} denotes the coefficients of the regional dummies where *s* is the base region, and ε_{rsj}^{NA} is the stochastic error term. Using equation (16), we can check for bilateral RFPE between districts by comparing regions with one another.

4. Data

The study uses plant-level data from the Census of Manufacturing Industries (CMI) for Punjab for 2000/01 and 2005/06. Conducted jointly by the Pakistan Bureau of Statistics, the provincial Directorate of Industries, and the Bureau of Statistics, the CMI covers 2-digit, 3-digit, 4-digit and 5-digit industries under the Pakistan Standard Industrial Classification. The census provides data on the quantity and value of inputs and outputs, 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 nonproduction workers, and the district and industry to which they belong. The CMI data on the number and wages of production and nonproduction employees in each firm allows us to calculate the wage bills. The analysis also requires information on the industries produced by each region, which the CMI provides according to geographic subdivision at the district, province and national levels.

Table 1 shows how Punjab is organized from an administrative perspective. The zonal distribution reflects Cheema, Khalid, and Patnam's (2008) analysis of the geography of poverty in Punjab. According to this distribution, Punjab can be divided into four zones: northern, southern, central and western Punjab, with central Punjab being the largest. Based on this administrative breakdown, the analysis is carried out at two levels: the smaller district level and the larger zonal level. For 2001, this analysis covers 34 districts and 592 4-digit industries. For 2005/06, it covers 35 districts – including the 34 districts from 2001 and the newly created district of Nankana Sahib – and 540 4-digit industries.

Table 1: Administrative regions and districts

Zone	District	Zone	District
Northern	Attock	Western	Bhakkar
	Chakwal		Dera Ghazi Khan
	Jhelum		Khushab
	Rawalpindi		Layyah
			Mianwali
Central	Gujranwala		Muzaffargarh
	Gujrat		Rajanpur
	Hafizabad		
	Jhang	Southern	Bahawalnagar
	Kasur		Bahawalpur
	Lahore		Khanewal
	Mandi Bahauddin		Lodhran
	Nankana Sahib		Multan
	Narowal		Rahimyar Khan
	Okara		
	Pakpattan		
	Sahiwal		
	Sargodha		
	Sheikhupura		
	Sialkot		
	Toba Tek Singh		
	Vehari		

Note: Given that Nankana Sahib was made a separate district in 2006, the analysis for 2001 does not treat it as a separate district.

5. Empirical Results for Bilateral RFPE

In order to test for bilateral RFPE between districts by comparing regions with one another, we estimate equation (16) allowing every individual region to be the base region. Since the analysis spans 35 districts, the equation is estimated a total of 35 times for both years (2001 and 2006), with each individual district as the base region.

5.1. Testing for Bilateral RFPE

An F-test to determine the joint significance of α_{rs}^{NA} in equation (16) is performed for each district and zone for both years. This indicates any

RFPE between the base district and all other districts. The results are given in Table 2.

Table 2: Bilateral rejections by base district

	Percentage of rejections		Distribution of rejections across base regions		
	Level of significance				
	5%	10%	Min.	Mean	Max.
Zones					
2000/01	75.0	75.0	1	1	3
2005/06	25.0	75.0	0	1	2
Districts					
2000/01	23.5	26.5	0	3	14
2005/06	14.3	20.0	0	4	13

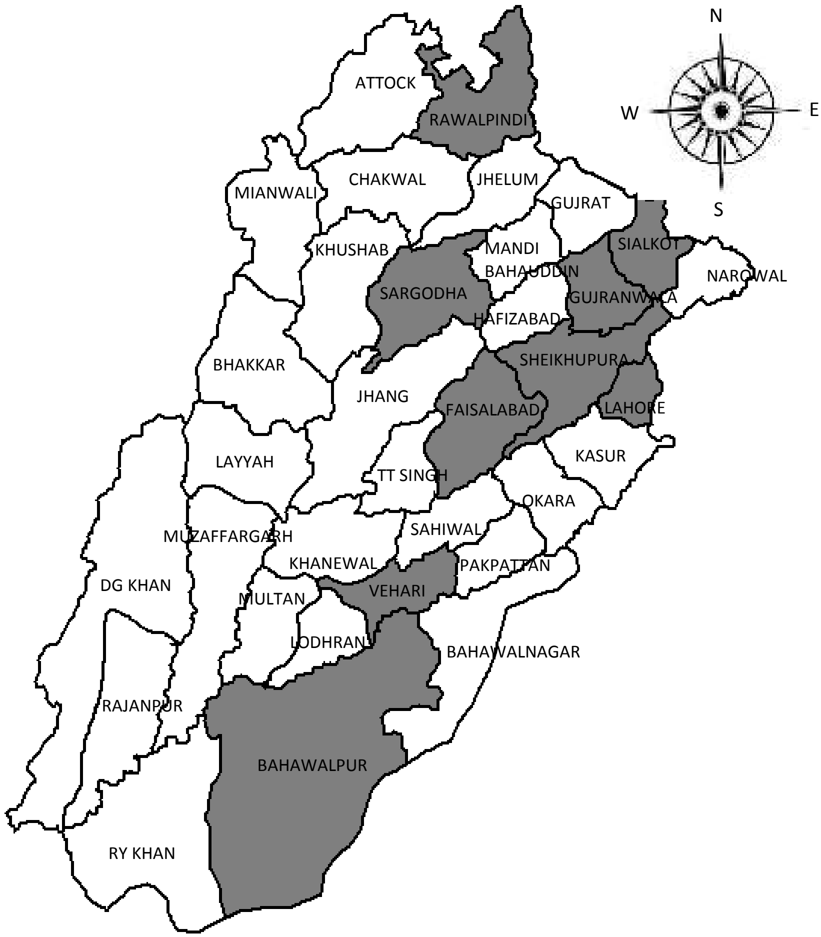
Note: The bilateral rejections of RFPE are based on our estimation of equation (16) for all possible base regions.
Source: Author's calculations based on data from CMI.

The percentages given in the table indicate the proportion of base districts for which we are able to reject the null for RFPE. For 2001, out of 34 regressions for each base district, we can reject the null hypothesis for RFPE (at a 5 percent level of significance) for nearly 23 percent. This proportion includes eight districts, that is, when the base district is Rawalpindi, Sargodha, Faisalabad, Sheikhupura, Lahore, Gujranwala, Vehari and Sialkot. However, at a 10 percent level of significance, we are able to reject the null for RFPE for 26 percent of the districts, including Bahawalpur.

For 2006, out of 35 regressions, we can reject the null for RFPE (at a 5 percent level of significance) for 14 percent of the districts. This is the case when the base districts include Lahore, Faisalabad, Gujrat, Sialkot and Dera Ghazi Khan. At the 10 percent level of significance, we are able to reject the null for RFPE for 20 percent of the districts, including Multan and Nankana Sahib. It is worth noting that most of these districts are located in central Punjab (Figures 3 to 6), where the industrial mix of districts is very wide and well diversified. This point is illustrated in the next section. For now, we have established empirically that the center contributes the most to factor price inequality in Punjab. This means that the relative wages of nonproduction workers are not constant across

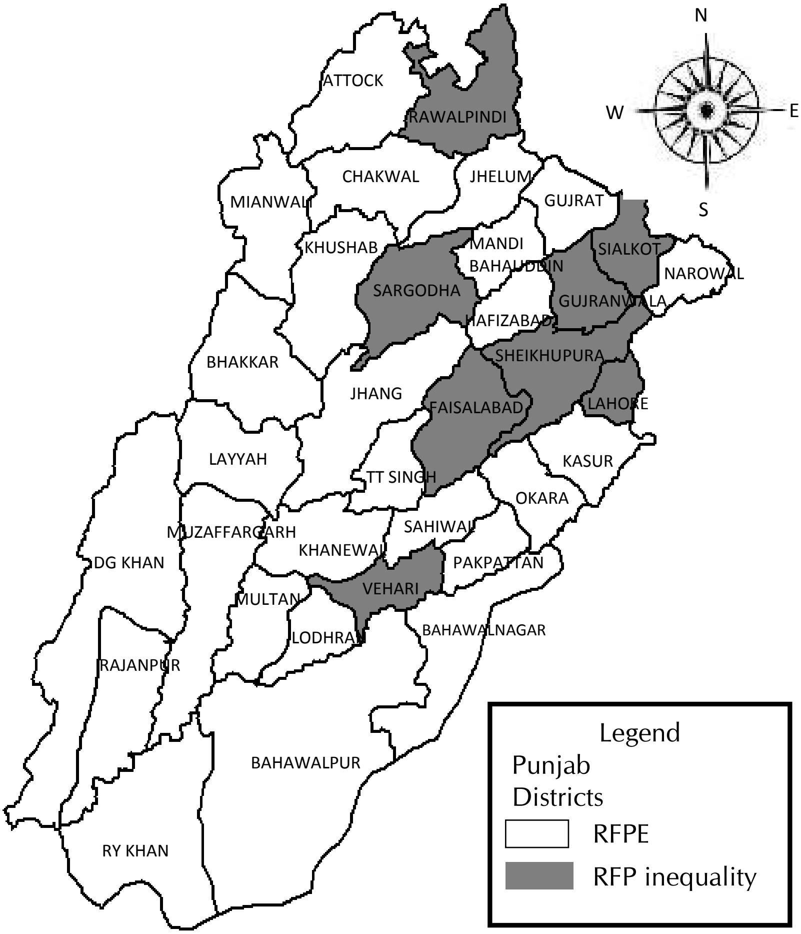
Punjab and, therefore, we need to test whether this differential affects its industrial landscape.

Figure 3: Prevalence of relative factor price inequality at district level, 2001



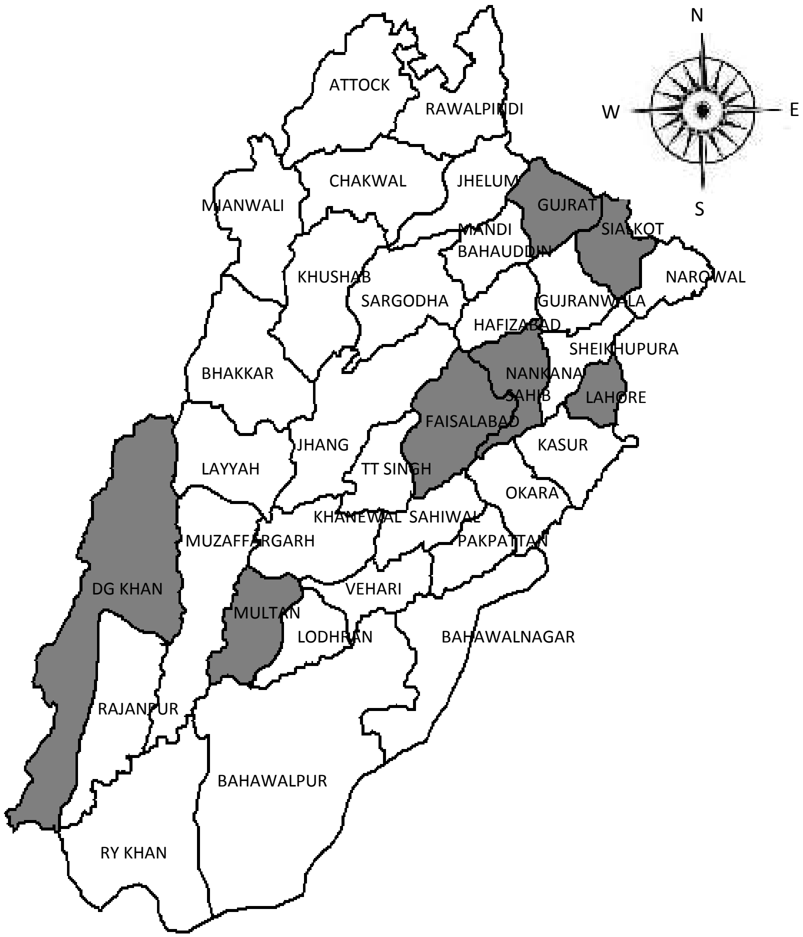
Note: 10 percent significance level.

Figure 4: Prevalence of relative factor price inequality at district level, 2001



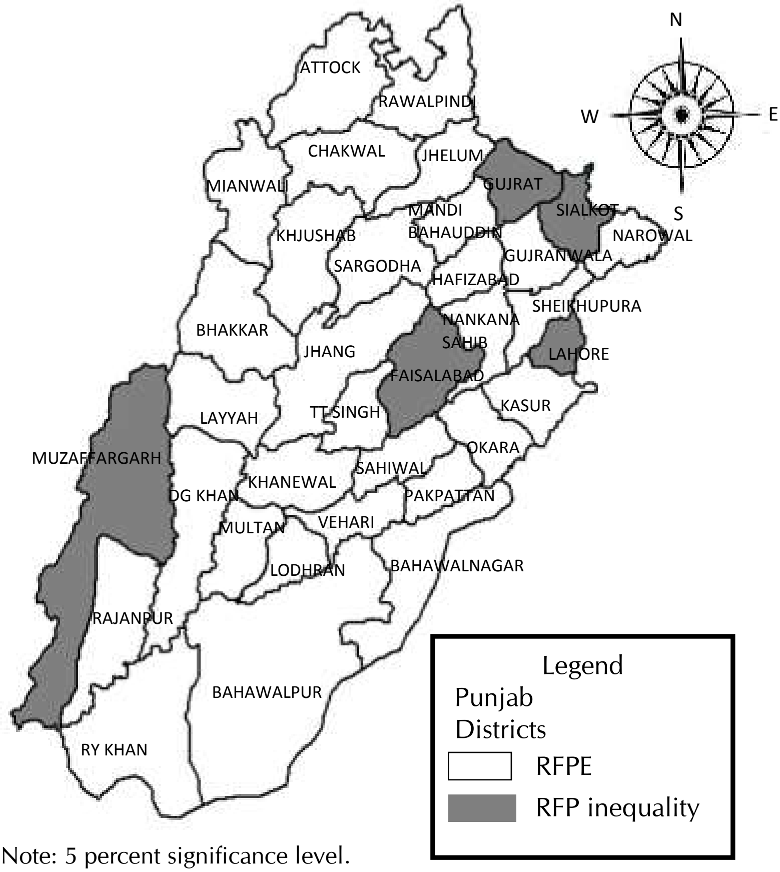
Note: 5 percent significance level.

Figure 5: Prevalence of relative factor price inequality at district level, 2005



Note: 10 percent significance level.

Figure 6: Prevalence of relative factor price inequality at district level, 2005



Note: 5 percent significance level.

5.2. Relating Factor Intensity and Factor Prices

Under the Heckscher–Ohlin theorem, districts abundant in nonproduction labor should have lower relative wages for nonproduction labor. Thus, in the first stage of the analysis, for each regression of the form equation (16), if a district is significantly positive it is likely to have a higher relative wage bill for nonproduction workers than the base district, suggesting that it is less abundant in nonproduction labor. Table 3 shows how many times a district's coefficient is statistically significant in the 35 regressions in which it appears (where each regression is done for a particular base district).

These results are based on testing the individual coefficients of the district dummies in equation (16). Depending on its sign, a significant coefficient indicates that the relative wages of nonproduction workers in that particular district are higher or lower than in the base district. Table 3 shows that the coefficient for Lahore is significant the most number of times. More interestingly, for each of the 11 regressions in which it appears significant, its coefficient remains positive, strongly suggesting that the relative wages of nonproduction workers are higher in Lahore than in the respective base district. These base districts include Rawalpindi, Gujrat, Sialkot, Mianwali, Multan and Nankana Sahib.

On closer examination of the data in Table 4, Lahore appears to be rich in industries characterized as relatively blue-collar-intensive: of the 86 industries produced in Lahore, 17 use production workers intensively. The skill intensity of industries for this study was measured by estimating the ratio of nonproduction workers to production workers for each industry within a district. For the top 25 percent of industries classified as skill-intensive, the ratio of nonproduction to production labor was greater than 0.406 for 2006 and greater than 0.441 for 2001. In Lahore, 83 percent of the industries are blue-collar labor-intensive, indicating the abundance of such labor in the district. In comparison to the base region in each of these regressions, Lahore has the highest percentage of blue-collar-intensive industries, barring Sialkot. Multan, for example, accounts for 72 percent, Gujrat for 84 percent and Mianwali for 40 percent of its industries being classified as blue-collar-intensive.

Table 3: Bilateral rejections by base district, 2005/06

Base district	Total 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 Bahauddin	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
Toba Tek Singh	3	3	2	0	0	3	2
Dera Ghazi Khan	3	3	2	3	2	0	0
Muzaffargarh	3	3	1	1	0	2	1
Rahimyar 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 data from CMI 2005/06.

Table 4: Industrial structure

Zone/district	2005/06 Number of industries	2005/06 % Low white-collar intensity industries	2005/06 % High white-collar intensity industries	2000/01 Number of industries	2000/01 % Low white-collar intensity industries	2000/01 % 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
Toba Tek Singh	11	64	36	10	70	30
Hafizabad	9	89	11	6	67	33
Nankana Sahib	7	71	29	-	-	-
Jhang	6	50	50	19	79	21
Mandi Bahauddin	5	60	40	3	33	67
Vehari	5	40	60	3	33	67
Pakpattan	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
Dera Ghazi 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
Rahimyar 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 on the ratio of nonproduction to production workers in an industry in a particular district. High white-collar intensity industries constitute the top 25 percent of industries based on this ratio.
Source: Author's calculations.

The results for Faisalabad are comparable to those for Lahore, given their similar industrial composition: 78 percent of the industries in Faisalabad are blue-collar-intensive. In the six instances that the district is significant (in six different regressions), it has a positive sign, suggesting that the relative wages of white-collar workers are higher in Faisalabad than in the base districts Rawalpindi, Gujrat, Mianwali and Multan. All these base districts also have a lower percentage of blue-collar-intensive industries, which explains why their relative wage bills for white-collar workers are lower. Again, Sheikhupura, which has a similar industrial composition to that of Faisalabad and Lahore, yields comparable results.

Rajanpur is significant in 16 regressions out of 35; of these 16 regressions, the district is significantly negative in 13 cases. This suggests an abundance of white-collar-intensive industries in Rajanpur and, therefore, a lower relative wage bill for white-collar workers compared to the base districts. These results make sense, especially given Rajanpur's industrial mix in a sample comprising industries that are classified as white-collar-intensive.

Sialkot is unique in that it is significantly negative in seven out of 35 regressions. This implies that the relative wages of white-collar workers are significantly lower in Sialkot than in Rawalpindi, Gujranwala, Gujrat, Chakwal, Faisalabad, Lahore and Vehari. This result is difficult to explain because 88 percent of Sialkot's industries are classified as blue-collar-intensive – a higher proportion than for any of the base districts except Gujranwala (Table 4). While this should lead to a higher relative wage bill for nonproduction workers, as in Lahore, the sign of the Sialkot coefficient in each of these regressions suggests otherwise.

Multan lies in between the two extremes, Lahore and Sialkot. Out of 35 regressions, it is significant in seven – three times below and four times above 0. The district has lower relative wages for skilled workers than Lahore, Faisalabad and Gujranwala, which is borne out by its higher percentage (28 percent) of white-collar-intensive industries relative to the base districts (Lahore, 17 percent; Faisalabad, 22 percent; Gujranwala, 10 percent). Multan also has higher relative wages for nonproduction workers than Sahiwal, Hafizabad, Mianwali and Gujrat.

Mianwali (60 percent) is the only district with a higher percentage of white-collar-intensive industries than Multan and, therefore, has lower relative wages for white-collar workers. However, given that Gujrat (16 percent), Sahiwal (24 percent) and Hafizabad (11 percent) have lower percentages of white-collar-intensive industries than Multan, the latter should, in theory, have lower relative wages for nonproduction workers – which is not the case. The reason for this differential in the opposite direction remains unclear.

Gujrat also lies midway between the two extremes and has significantly lower relative wages for nonproduction workers than do Lahore, Sheikhupura, Gujranwala and Faisalabad. All these base districts, barring Faisalabad, have a higher or equal percentage of blue-collar-intensive industries. Gujrat has higher relative wages for nonproduction workers than Sialkot and Bahawalpur, where Sialkot has a higher percentage of blue-collar-intensive industries and Bahawalpur has a higher percentage (38 percent) of white-collar-intensive industries compared to Gujrat.

What emerges from these results is that the relative wage bill for nonproduction workers is higher in districts situated in central Punjab, that is, Lahore, Sheikhupura, Gujranwala and Faisalabad. These districts have several common features: they form the industrial hub of Punjab and produce a large number of industries, most of which are blue-collar-intensive (see Table 4) and, therefore, have higher relative wages for white-collar workers (and lower relative wages for blue-collar workers).

Table 5 shows that these districts account for a very small share of manufacturing employment in white-collar-intensive industries, reinforcing the argument that they are abundant in production labor and, under the Heckscher–Ohlin theorem, produce industries that are relatively blue-collar-intensive. However, the quality of human capital in these regions is far higher and would, therefore, suggest otherwise. The quality of human capital across various districts of Punjab is given by the human development index in Table A1 in the Appendix, which shows that districts in the center (the industrial hub of Punjab) rank much higher than districts in either the south or west.

Table 5: Distribution of manufacturing employment across industries of varying factor intensity, by district

Zone/district	2005/06			2000/01		
		% Manufacturing employment in			% Manufacturing employment in	
	Number of industries	Low white-collar intensity indus.	High white-collar intensity indus.	Number of industries	Low white-collar intensity indus.	High white-collar intensity indus.
North						
Chakwal	6	99.13	0.87	4	100.00	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	77.40	22.60
Sahiwal	17	76.05	23.95	17	93.98	6.02
Nankana Sahib	7	68.88	31.12	-	-	-
Okara	14	61.05	38.95	3	29.16	70.84
Sargodha	14	53.80	46.20	14	54.58	45.42
Jhang	6	36.94	63.06	19	57.73	42.27
Mandi Bahauddin	5	32.90	67.10	3	45.97	54.03
Toba Tek Singh	11	29.77	70.23	10	39.31	60.69
Pakpattan	4	0.33	99.67	4	1.87	98.13
Narowal	3	0.00	100.00	2	0.00	100.00
West						
Muzaffargarh	10	91.84	8.16	6	95.65	4.35
Dera Ghazi Khan	6	91.05	8.95	7	95.32	4.68
Khushab	12	87.39	12.61	7	84.60	15.40
Bhakkar	7	56.45	43.55	2	56.10	43.90
Mianwali	5	45.35	54.65	4	40.40	59.60
Rajanpur	3	0.00	100.00	1	100.00	0.00
Layyah	3	0.00	100.00	3	0.00	100.00
South						
Bahawalnagar	5	100.00	0.00	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	17.40	82.60
Rahimyar Khan	9	8.42	91.58	10	30.65	69.35
Lodhran	3	3.05	96.95	4	27.31	72.69

Source: Author's calculations.

Before declaring that these districts are scarce in nonproduction labor, we need to establish that Punjab as a whole does not specialize in white-collar-intensive manufacturing industries to the extent that blue-collar workers are not substitutes for white-collar workers if need be. Thus, the diversification cones for individual districts are not as diverse after all because the factor intensity ratios and relative factor prices for many districts are not very distinct.

In about 75 percent of Punjab's industries, the ratio of nonproduction workers to production workers is less than 0.5. Specifically, this ratio was less than 0.406 as of 2005/06. Of this proportion, half the industries have a ratio of less than 0.257. This implies that, on the whole, Punjab's industries do not demand a very high skill level.[5]

5.3. Alternative Explanations for Counterintuitive Results

The evident scarcity of skills in districts that comprise the industrial hub of Punjab may be a result of nonproduction workers in these areas being employed in sectors other than manufacturing (such as the services sector) to which their skills are more suited. The scarcity might, therefore, be artificial: as more nonproduction workers find jobs outside the manufacturing sector, industries are compelled to attract other nonproduction workers at very high wages. On the other hand, districts that offer nonproduction workers significantly lower relative wages are located in the south or west of Punjab – mainly Bahawalpur, Rajanpur and Multan. Under the Heckscher–Ohlin theorem, this should suggest an abundance of nonproduction workers in these areas.

A common feature of these districts is that, while they do not have as many industries as Lahore, Faisalabad, Sheikhupura or Gujranwala, they have a higher proportion of white-collar-intensive industries and, therefore, lower relative wages for nonproduction labor. Moreover, their percentage share of manufacturing employment in Punjab's white-collar-intensive industries is very high compared to districts in central Punjab (see Table 5). Again, the districts in question are less developed than districts such as Lahore, Faisalabad, Sheikhupura and Gujranwala. While

[5] The industries that are classified as white-collar-intensive based on this ratio (i.e., where the ratio of nonproduction to production workers is greater than 1) include: vegetable and inedible animal oils and fats, starch and its products, petroleum products, lime, plaster and its products, and iron and steel (basic). The share of total district employment in these industries is very small (see Table 5).

the lower relative wages for nonproduction labor in these areas suggest an abundance of white-collar workers, the socio-development statistics indicate otherwise.

What pushes up the relative wages of nonproduction workers in Punjab's more developed districts? As discussed earlier, one explanation is that, on average, Punjab's manufacturing industries do not require a very high level of skills. Such skills are abundant in the west and south (less developed areas), while opportunities for these semi-skilled workers to improve their capacity are limited. Thus, the kinds of skills required in manufacturing reflect what people here can easily achieve, given the resources and opportunities available.

The limited range of both opportunities for employment as well as skills means that manufacturing industries in south and west Punjab face an excess supply of semi-skilled labor, which they are willing to hire only at lower relative wages. Thus, the higher relative wages for nonproduction workers in central Punjab are due to the increase in relative demand for white-collar labor in these areas. In the less developed areas, the excess supply of semi-skilled workers may be responsible for driving down relative wages.

Some alternative explanations for these counterintuitive results are discussed below. First, manufacturers are able to offer different wage rates across Punjab when certain factors render labor immobile. People may be reluctant to migrate because it would mean losing the support of their kinship network, while living in an industrial hub might entail a higher cost of living in addition to the direct cost of migration. If the costs associated with migration are greater than the expected gains, then workers may choose not to migrate despite the wage differential. This is likely to occur if real wages are equalized across regions while the cost of living varies significantly across districts, giving workers little incentive to migrate. Thus, the difference in real wages is significantly smaller than that in nominal wages (see Moretti, 2011).

Second, we need to take into account the quality of labor. This is done by incorporating information on the quality of the labor force into the observed factor prices. The differential in the relative wages of nonproduction workers implies that the quality of skills is higher in districts such as Lahore, Faisalabad and Gujranwala. This would suggest that even blue-collar workers in these districts are more efficient than white-collar workers in the south and west.

Since the commodities produced across Punjab do not require distinct factor intensity vectors, whatever is being produced can be manufactured in the larger districts using blue-collar labor more intensively because the overall quality of labor is better here than in the south and west where more resources are needed to produce the same output. This is similar to the Leontieff paradox in the US, where the country was found to be exporting labor-intensive products despite being capital-intensive. On further examination, it was discovered that the original analysis had not accounted for the role of human capital: the US labor force is three times as efficient as that of its trading partners.

In the same way, the blue-collar-intensive products manufactured by Punjab's industrial hub are produced by a labor force that is, on average, more efficient. Moreover, the elasticity of substitution of production workers for nonproduction workers appears to be very high since almost all industries fall under the head of consumer goods. Nominal wages differ across districts due to variations in the marginal product of labor; the differences in nominal wages should reflect exactly the differences in the marginal product of labor in Punjab's industries.

Finally, these results could also be due to a demand bias. Per capita income is higher in districts in the center, and rising consumption demand in this part of Punjab might explain why industries manufacturing consumer goods have chosen to locate here. This demand bias, coupled with the high elasticity of substitution of factors, may have encouraged blue-collar-intensive industries to locate in the relatively more developed part of Punjab.

The finding that more developed districts such as Lahore, Faisalabad and Gujranwala have higher relative wages for nonproduction workers might also have to do with the CMI's sample of firms. For the less developed districts that appear to have a higher share of white-collar-intensive industries, it may be that the CMI is simply missing many of the blue-collar-intensive firms located here either because they were too small to be captured by the census or because the census response rate was lower in those districts. Thus, either sampling errors or irregularities in the response rate could explain these unexpected results.

The two exceptions are Gujrat and Sialkot, which have significantly lower relative wages for nonproduction workers than the base districts (which are more white-collar-intensive than either). The relatively white-collar-intensive industries in these two districts include grain milling, light engineering, pharmaceuticals and beverages. However, these industries account for only 2.7 percent of the workforce for Gujrat and a negligible 0.34 percent of the workforce for Sialkot. The lower relative wages for nonproduction workers is what one would expect as these districts are home to a large number of industries and also have a fairly skilled labor force.

Table 6 gives similar results for 2001. Most of the districts that have significantly higher relative wages for nonproduction workers are concentrated in central Punjab and include Lahore, Faisalabad, Gujranwala and Sialkot. Other districts are Dera Ghazi Khan, Lodhran and Vehari. As Table 4 shows, a common feature of all these districts is that they have a relatively high percentage of blue-collar-intensive industries. They also account for a very low share of employment in Punjab's white-collar-intensive industries and thus a very high share of employment in the province's blue-collar-intensive industries (see Table 5). This indicates a relative abundance of nonproduction labor and explains why these districts have higher relative wages for nonproduction workers than the base districts.

Table 6: Bilateral rejections by base district, 2000/01

Base district	Total 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
Dera Ghazi 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
Toba Tek 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 Bahauddin	1	1	1	0	0	1	1
Sahiwal	1	1	1	1	1	0	0
Bahawalnagar	1	1	0	0	0	1	0
Rahimyar 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

Source: Author's calculations based on data from CMI.

On the other hand, districts with significantly lower relative wages for nonproduction workers compared to the base districts include Jhelum, Pakpattan, Sargodha and Muzaffargarh. Again, this suggests a relative abundance of nonproduction labor in areas of Punjab that are poorly developed, such as Pakpattan, Lodhran, Vehari and Muzaffargarh. That the more developed districts such as Gujranwala, Faisalabad, Lahore and Sialkot have higher relative wages for nonproduction workers implies a dearth of nonproduction labor.

5.4. Testing for RFPE at the Zonal Level

Table 1 divides Punjab into four zones: north, center, west and south. Table 2 analyzes factor price equalization at the larger zonal level. An F-test to check for the joint significance of α_{rs}^{NA} in equation (16) is performed for each of the base zones for both years to determine RFPE between a base zone and all other zones of Punjab. Of the four regressions (one for each base zone) for 2000/01, we can reject the null of factor price equality in 75 percent of these cases, both at the 5 and 10 percent level of significance. For 2005/06, out of the four regressions, 25 percent (one zone) exhibit factor price inequality at the 5 percent level of significance. This particular zone is central Punjab, which reinforces our earlier results. On the other hand, 75 percent of these zones indicate factor price inequality if we use a 10 percent level of significance for 2005/06.

In testing for the individual βs in equation (16), Table 7 shows that, for 2001, there is at least one rejection for each base zone except when the center is taken as the base. This suggests that one particular zone has significantly different relative factor prices from the base zone in each regression. Not surprisingly, the zone in question is the center, which has a positive significant coefficient in all the regressions, suggesting that it has higher relative wages for nonproduction workers. When the center is the base zone, there are three rejections: the remaining three zones have significantly lower relative wages for nonproduction workers than the center at both a 5 and 10 percent level of significance. Thus, on the whole, the center has higher relative wages for nonproduction workers, which reinforces the results of the district-level analysis.

Table 7: Bilateral rejections by base zone, 2000/01

Base zone	Total rejections	Signif. level		Positive signif. coeff.		Negative signif. coeff.	
		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 calculations based on data from CMI.

For 2005/06, Table 8 shows that the north zone does not have significantly different relative wages compared to the other zones at any significance level, whereas the center has significantly different relative wages for nonproduction workers compared to the west and south. Both these zones have significant negative coefficients, suggesting that they have lower relative wages for nonproduction workers. When the west is the base zone, the center and north appear to have significantly higher relative wages for nonproduction workers.

Table 8: Bilateral rejections by base zone, 2005/06

Base zone	Total rejections	Signif. level		Positive signif. coeff.		Negative signif. coeff.	
		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 calculations based on data from CMI.

There is only one rejection – in this case, the center – for the south zone. The center has a significant negative coefficient, suggesting that it has significantly higher relative wages for nonproduction workers compared to the south. Overall, therefore, the results show that relative wages for nonproduction workers in the center are higher than in every other zone. The center and the west contribute more to factor price inequality in Punjab than the other two zones. We can see from Table 4 that the districts located in the center and west house an extensive industrial setup, making it likely that the high level of manufacturing activity in these two zones has led to the relative wage variation.

6. Relative Wage Differences and Industrial Structure

This section investigates whether departures from RFPE are associated with differences in the set of industries that regions produce. We estimate the following OLS model:

$$Z_{rs} = \beta_o + \beta_1|\alpha_{rs}^{NA}| + \beta_2 I_r + \beta_3 I_s + \mu_{rs} \quad (17)$$

where Z_{rs} measures the similarity of industrial structure (the number of common industries produced by two regions), α_{rs}^{NA} represents the coefficient estimates of the dummies in equation (16), which account for the bilateral wage bill differentials between two regions, I_s is the total number of industries produced by region *s* and I_r is the total number of industries produced by region *r*.

β_1 needs to be interpreted cautiously because it measures the responsiveness of one endogenous variable to the other. While we cannot, therefore, comment on its magnitude, we can interpret the direction of the relationship: this needs to concur with neoclassical trade theory, which holds that regions producing the same set of industries are active in the same cone of diversification and thus characterized by RFPE. Therefore, β_1 in equation (17) should have a negative sign. A fall in the relative wage differential between two regions must be associated with those regions producing a greater number of common industries.

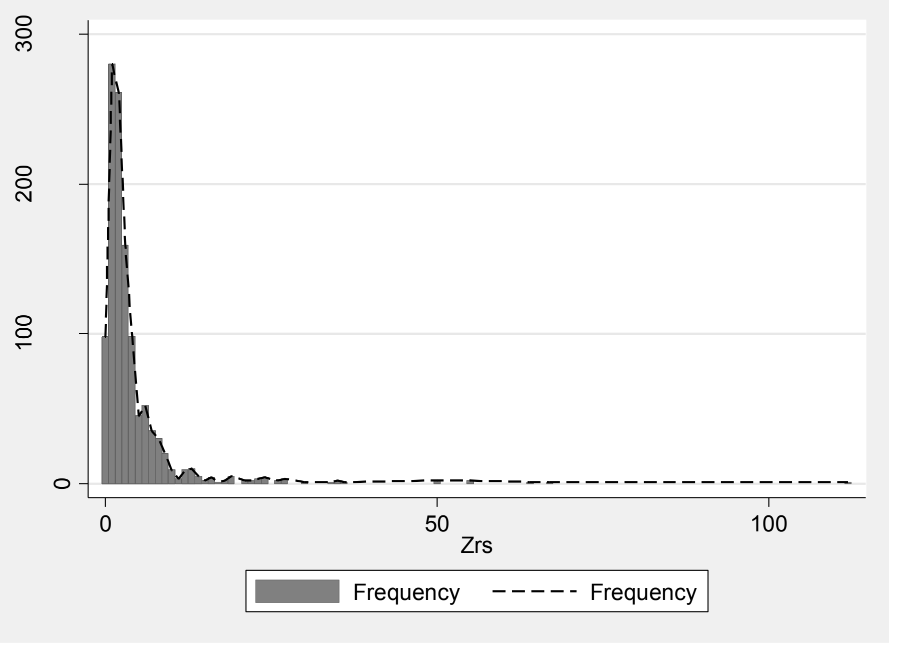
To determine whether our results are driven by the choice of dependent variable, we test equation (17) and its predictions using another dependent variable – Krugman's specialization index. This is an alternative measure of industrial similarity, which it predicts through the employment structure of two regions. Mathematically, the index is written as follows:

$$Z_{rs} = \sum_j \left|\left(\frac{L_{rj}}{L_r} - \frac{L_{sj}}{L_s}\right)\right| \quad (18)$$

The index is the sum of the absolute difference between industries' specific share of employment in each district: in district *r*, L_{rj}/L_r is the employment share of industry *j* in the total employment of district *r*; in district *s*, L_{sj}/L_s is the employment share of industry *j* in the total employment of district *s*. The value of the index ranges from 0 to 2. The industrial similarity between two regions decreases as the index approaches 2 and increases as it approaches 0. Therefore, β_1 is expected to have a positive sign when using Krugman's index as a measure of industrial similarity.

As the shape of the frequency polygon for Z_{rs} shows in Figures 7 and 8, the dependent variable is discrete, nonnegative and highly skewed. Z_{rs} is positively skewed for both periods of analysis. In the case of a nonnegative, skewed, discrete dependent variable such as this, simple OLS produces noninteger values. OLS can also predict negative values of the dependent variable. This means carrying out either a negative binomial regression or a Poisson regression to test if our results are driven by the choice of regression analysis.

Figure 7: Frequency polygon and histogram for Z_{rs}, 2000/01

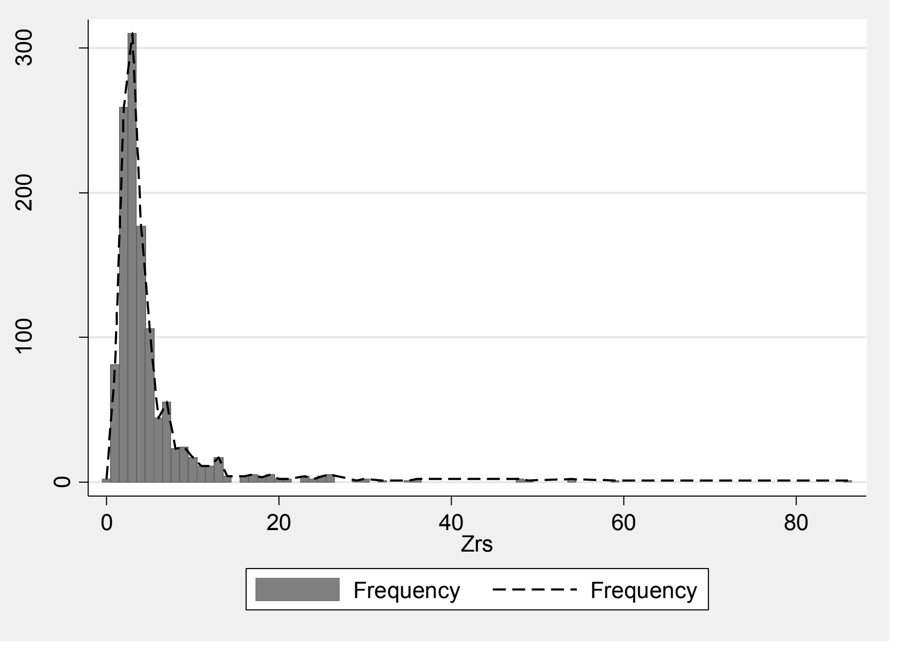


The deciding factor as to which regression analysis is more suitable is a likelihood ratio test to determine if the dispersion parameter is equal to 0 (in which case we would apply the Poisson regression) or is significantly greater than 0 (in which case we would apply the negative binomial regression). This parameter is the amount by which the variance of a variable differs from the mean. In a negative binomial regression, we would have

$$Variance = mean + kmean^2$$

If $k = 0$, then the mean is equal to the variance, implying that there is no dispersion and the Poisson regression is more suitable. If k is significantly greater than 0, then the variance is greater than the mean, which shows that the variable is over-dispersed (and under-dispersed otherwise). In this case, a negative binomial regression would fit the data better.

Figure 8: Frequency polygon and histogram for Z_{rs}, 2005/06



If the likelihood ratio test is carried out to determine whether the dispersion parameter is equal to 0, then our hypothesis is:

$$H_0: k = 0$$

$$H_1: k > 0$$

If the p-value of the chi-squared statistic obtained from the negative binomial regression is equal to 0, this is sufficient evidence of *k* being greater than 0. However, it is important to note that the negative binomial regression transforms the dependent variable into a log of expected counts. Its coefficient estimates would, therefore, indicate the difference in the log of expected counts in response to a one-unit change in the predictor variable. This makes it necessary to calculate the true coefficient estimates by means of the marginal effects after estimating the negative binomial regression.

Table 9 summarizes industrial specialization across districts. The first indicator of industrial specialization is "regions per industry," which shows the extent to which an industry is dispersed across Punjab or, more simply, in how many districts a particular industry is located. We see from the first row of the table that some industries are highly localized and found in as few as one district as of 2005/06. Some examples could be industries that manufacture motor vehicle bodies or fit out caravans and mobile homes, both of which are located only in Lahore.

Table 9: Industrial specialization across districts

Districts	Year	Min.	Median	Max.
Regions per industry	2005/06	1.00	4.00	34.00
As % of all regions		2.86	11.43	97.14
Regions per industry	2000/01	1.00	3.00	27.00
As % of all regions		2.94	8.82	79.41
Industries per region	2005/06	3.00	9.00	86.00
As % of all industries		3.06	9.18	87.76
Industries per region	2000/01	1.00	10.00	112.00
As % of all industries		0.65	6.45	72.26
Bilateral overlap as a % of larger region's industries	2005/06	2.32	22.22	100.00
	2000/01	0.00	14.29	75.00
Krugman's specialization index	2005/06	0.034	0.995	1.97
	2000/01	0.070	0.995	1.99

Source: Author's calculations based on data from CMI.

Similarly, as of 2005/06, ethyl alcohol and spirits were manufactured only in Nankana Sahib; in 2000/01, industries such as motorcycle and rickshaw manufacturing were confined to Sheikhupura. Other industries, however, such as rice husking and rice and grain milling, were produced in almost every district apart from Bahawalnagar in 2005/06. In 2001, wheat and grain milling was spread across 27 out of 34 districts (excluding Dera Ghazi Khan, Rajanpur, Mandi Bahauddin, Chakwal, Bhakkar, Khushab and Mianwali). The indicator shows that, over the years, certain industries have penetrated a greater range of districts – up to 79 percent of districts in 2000/01 and up to 97 percent in 2005/06.

The next two rows of Table 9 give indicators of industrial diversification at the district level. The indicator "industries per region" gives the number of industries active in a particular district. As of 2005/06, certain districts account for only 3 percent of the industries active in Punjab. Specifically, Narowal accounts for only three industries out of 98 in the province, including the manufacture of taps, valves and vacuum pumps, rice husking and milling, and animal or vegetable oils and fats. Lahore is the most diversified district, housing 86 (87 percent) industries out of 98 in 2006. Similarly, in 2000/01, Rajanpur accounted for only one industry (leather footwear) out of a total of 155, while Lahore accounted for 112 or 72 percent of Punjab's active industries. Thus, the indicator shows an increase in industrial diversification at the district level over this period.

The next two rows of the table give indicators of industrial similarity. The first is the bilateral overlap ratio, which is the ratio of industries common to two given districts, measured by the number of common industries as a percentage of the region's industries. Over the given period, the maximum proportion of industries common to any two districts has gone up from 75 percent in 2001 to 100 percent in 2006, while the minimum proportion has risen from 0 percent to 2 percent. This is indicative of greater industrial diversification at the district level over time.

The last row of Table 9 gives Krugman's specialization index, which trends downward from 0.07 in 2001 to 0.03 in 2006, indicating increasing industrial similarity over time. The maximum value of the index falls slightly from 1.99 to 1.97; this also points to growing industrial similarity over the sample period.

7. Empirical Results for Industrial Diversification

The estimation results for equation (17) are given in the first columns of Tables 10 (for 2005/06) and 11 (for 2000/01), and are in line with neoclassical trade theory. The statistically significant negative sign of the coefficient of the absolute relative wage bill gap ($|\alpha_{rs}^{NA}|$) in column 1 of both tables suggests that, as the difference in a district's relative wages compared to the base district goes up, the number of common industries produced goes down. Technically, the two districts will lie in the same cone of diversification or be active within the same set of industries if they have similar relative factor prices (in this case, relative wages).

This relationship remains significant over time. The second columns in Tables 10 and 11 give the results for the same regression for 2005/06 and 2000/01, respectively, but controlling for district pair dummies. The equilibrium relationship between the absolute relative wage bill gap and the number of common industries between two districts still holds and is statistically significant for both years.

7.1. Robustness Checks via Alternative Regression Technique

The third columns in Tables 10 and 11 give the marginal effects for the negative binomial regression conducted for both periods (see Tables A2 and A3 in the Appendix for the complete results). A Poisson regression is not carried out because Z_{rs} is over-dispersed as shown by the likelihood ratio test.[6] The coefficient estimates indicate similar results, i.e., a statistically significant negative relationship between the absolute relative wage bill gap and industrial similarity, marking the fact that these results are not driven by the choice of regression technique. The fourth columns in both tables suggest that the results of the marginal effects are robust to the inclusion of region pair dummies as the coefficient of the absolute relative wage bill gap retains its statistical significance and direction.

---

[6] The p-value of the chi-bar-square is equivalent to 0 for both years (Tables A2 and A3 in the Appendix), implying that the alpha term (dispersion parameter) is significantly greater than 0. This suggests that a negative binomial regression fits the data better.

Table 10: Relative wage differential and industrial structure, 2005/06

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS		Negative binomial		OLS	
	Number of common industries				Krugman's specialization index	
Absolute relative wage bill gap	-1.402*** (0.341)	-1.159*** (0.387)	-1.505 *** (0.207)	-0.868*** (0.182)	0.024 (0.028)	0.093*** (0.026)
Number of industries in *r*	0.164*** (0.007)	0.140*** (0.013)	0.0834*** (0.003)	0.688*** (0.242)	0.009*** (0.001)	0.006*** (0.001)
Number of industries in *s*	0.159*** (0.007)	0.103* -1.159***	0.0810*** (0.003)	0.695*** (0.240)	-0.002*** (0.001)	-0.001 (0.001)
Constant	0.517* (0.267)	-	-	-	0.903*** (0.022)	-
Scale factor for marginal effects	-	-	4.094	3.717	-	-
Regional dummies	No	Yes	No	Yes	No	Yes
R2	46.06	49.50	-	-	14.90	44.70
Observations	1,191	1,191	1,191	1,191	1,225	1,225

Note: The coefficient estimates in cols. 3 and 4 are marginal effects and not the estimated negative binomial regression coefficients (which are given in Tables A2 and A3 in the Appendix). Standard errors are given in parentheses.
Source: Author's calculations.

Table 11: Relative wage differential and industrial structure, 2000/01

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS		Negative binomial		OLS	
	Number of common industries				Krugman's specialization index	
Absolute relative wage bill gap	-2.007*** (0.348)	-1.763*** (0.372)	-1.612*** (0.183)	-1.114*** (0.142)	0.096*** (0.020)	0.103*** (0.019)
Number of industries in *r*	0.129*** (0.007)	0.107*** (0.012)	0.056*** (0.003)	-0.013 (0.0497)	0.006*** (0.000)	0.004*** (0.001)
Number of industries in *s*	0.123*** (0.007)	0.112*** (0.012)	.053*** (0.003)	0.310*** (.051)	-0.001*** (0.000)	0.008*** (0.001)
Constant	0.537* (0.289)	-	-	-	0.851*** (0.017)	-
Scale factor for marginal effects	-	-	3.087	2.524		
Regional dummies	No	Yes	No	Yes	No	Yes
R2	38.88	57.50			19.90	93.60
Observations	1,122	1,122	1,122	1,122	1,122	1,122

Note: The coefficient estimates in cols. 3 and 4 are marginal effects and not the estimated negative binomial regression coefficients (which are given in Tables A2 and A3 in the Appendix). Standard errors are given in parentheses.
Source: Author's calculations.

7.2. Robustness Checks via Alternative Dependent Variable

The fifth columns in Tables 10 and 11 give estimates for equation (18) using Krugman's specialization index as the dependent variable. As expected, the sign of the absolute relative wage bill gap is now positive. In Table 11, this coefficient is also statistically significant, indicating that the industrial similarity between two districts decreases as the relative wage differential rises.

For 2005/06, we are unable to obtain a statistically significant coefficient for the relative wage bill gap. However, once the region pair dummies are controlled for, there is a negative relationship between the relative wage differential and the industrial similarity between two districts. Column 6 in Table 11 reinforces the results for the alternative dependent variable for 2001 by showing that these are robust to the inclusion of region pair dummies. The results retain their direction as well as their statistical significance.

Using various specifications, this section demonstrates that the equilibrium relationship between the absolute relative wage bill gap and the industrial similarity between two given regions in Punjab is in accord with neoclassical trade theory. Based on two periods, 2005/06 and 2000/01, there is substantial evidence to support the view that industrial similarity falls as the relative wage gap between two districts goes up. In other words, regions characterized by similar relative factor prices tend to lie within the same cone of diversification.

8. Conclusion

This empirical analysis tests for bilateral FPE in Punjab at the district as well as zonal level. It also tests the neoclassical trade theory proposition that departures from RFPE lead to differences in the mix of industries that regions produce. Theoretically, regions abundant in a particular factor exhibit a lower relative price for that factor than regions in which it is scarce. Regions with a lower relative price for a particular factor have a higher concentration of industries using that factor intensively than regions with a higher relative factor price.

In terms of the FPE hypothesis, the study shows that, as of 2000/01, Punjab was characterized by factor price inequality prevailing in the central and western zones. The number of bilateral rejections is very high when districts from these two zones are taken as base districts. For 2005/06, the study yields consistent results at the district level with most bilateral rejections occurring in cases where the base district is located in central Punjab. This suggests that the center contributes the most to relative factor price inequality over the sample period, which should be the case because the zone houses a large number of industries. The bulk of industrial activity in Pakistan is concentrated in Punjab, which therefore attracts workers with skills suited to manufacturing activities. Workers will likely want to move to such thick labor markets because there is a higher probability of finding employment (and, therefore, a lower probability of unemployment). Denser markets are more productive[7] and are able to offer a higher wage premium, which might explain why the relative wages of nonproduction workers are higher in central Punjab.

At the zonal level, the results are the same for both years as we can reject the null for RFPE for three out of four zones at the 10 percent level of significance. The results also suggest that the relative wages of nonproduction workers are generally higher in the center, indicating that labor skills are scarce in these particular districts. However, this scarcity does not mean that districts such as Lahore, Gujranwala and Faisalabad, which have better indicators of human capital than other districts in Punjab, necessarily have lower skill levels. It could just mean that an artificial scarcity is created when nonproduction labor is employed in other sectors such as services. Thus, the manufacturing sector hires nonproduction labor at higher relative wages in the larger districts that comprise Punjab's industrial hub.

The higher wages of nonproduction workers in this case also reflect the better quality of human capital -- the returns to a higher level of education are obviously going to be greater. Additionally, Punjab's manufacturing industries do not require a very high level of skill; the kinds of skills (low) in demand are those that are easily available in the larger districts. Moreover, there is a demand bias that has attracted these industries manufacturing consumer goods to locate all over Punjab irrespective of the factor supply.

[7] Recent work by Henderson (2003) and Moretti (2004) seeks to provide more direct evidence by testing whether total factor productivity at the firm level is higher in denser areas.

In testing the second proposition, there is strong evidence to show that the number of industries common to any two regions falls as the absolute relative wage bill gap between these regions increases. This hypothesis is robust to various econometric specifications.

Such an analysis is important not only because it yields information on the income convergence within a country, but also because it has public policy implications. The regional policies typically adopted by governments range from direct subsidies and tax incentives for firms, to subsidized loans, technology transfer programs, export assistance and providing the necessary infrastructure and workforce training. However, policies targeting regional development need to take into account any information on factor prices and hence on the regional comparative advantage. So, for example, producing more white-collar-intensive industries in central Punjab could result in a comparative cost disadvantage for such firms because white-collar labor is more expensive here.

While encouraging white-collar-intensive industries to locate in lagging regions is an option, this must be complemented by other policies that boost workers' skills in such regions through education and training. As Moretti (2011) points out, policies that target regions rather than individuals generally prove futile, especially when labor is perfectly mobile. Government efforts to subsidize blue-collar-intensive firms that choose to locate in a lagging area would encourage blue-collar workers to move here from the center to take advantage of the higher relative wages. This would arbitrage away any benefits from such a policy. In order to produce a more equitable outcome, the government must focus on improving and enhancing the skills of workers in lagging regions so that they have a better chance of finding good job opportunities other than those in manufacturing, for example.

References

Ahmed, H. (2012). The impact of public school enrolment on child labor in Punjab, Pakistan. *Lahore Journal of Economics*, *17*, 1–34.

Berger, H., & Westermann, F. (2001). *Factor price equalization? The cointegration approach revisited* (Working Paper No. 471). Munich: Ifo Institute, Center for Economic Studies.

Bernard, A. B., Redding, S., & Schott, P. K. (2005). *Factor price equality and the economies of the United States* (Discussion Paper No. 696). London: Centre for Economic Performance.

Bernard, A. B., Redding, S., Schott, P. K., & Simpson, H. (2002). *Factor price equalization in the UK* (Working Paper No. 02/11). London: Institute for Fiscal Studies.

Bernard, A. B., Redding, S., Schott, P. K., & Simpson, H. (2004). *Relative wage variation and industry location* (Discussion Paper No. 614). London: Centre for Economic Performance.

Bernard, A. B., Redding, S., Schott, P. K., & Simpson, H. (2008). Relative wage variation and industry location in the UK. *Oxford Bulletin of Economics and Statistics*, *70*(4), 431–459.

Bernard, A. B., Robertson, R., & Schott, P. K. (2010). Is Mexico a lumpy country? *Review of International Economics*, *18*(5), 937–950.

Bernstein, J. R., & Weinstein, D. E. (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*, *137*(2), 215–243.

Burgman, T. A., & Geppert, J. M. (1993). Factor price equalization: A cointegration approach. *Review of World Economics*, *129*(3), 472–487.

Cheema, A., Khalid, L., & Patnam, M. (2008). The geography of poverty: Evidence from the Punjab [Special edition]. *Lahore Journal of Economics*, *13*, 163–188.

Courant, P. N., & Deardorff, A. V. (1992). International trade with lumpy countries. *Journal of Political Economy*, *100*, 198–210.

Courant, P. N., & Deardorff, A. V. (1993). Amenities, nontraded goods, and the trade of lumpy countries. *Journal of Urban Economics*, *34*(2), 299–317.

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* [Course handout]. Retrieved from http://www.econ.lsa.umich.edu/_alandear/courses/441/handouts/]

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

Demiroglu, U., & Yun, K. K. (1999). The lens condition for factor price equalization. *Journal of International Economics*, *47*, 449–456.

Dumais, G., Ellison, G., & Glaeser, E. L. (2002) Geographic concentration as a dynamic process. *Review of Economics and Statistics*, *84*(2), 193–204.

Gabe, T. M., & Abel, J. R. (2009). *Labor market pooling and occupational agglomeration* (Staff Report No. 392). New York, NY: Federal Reserve Bank of New York.

Gallaway, L. E. (1963). The North–South wage differential. *Review of Economics and Statistics*, *45*(3), 264–272.

Hanson, G. H. (1997). Increasing returns, trade and the regional structure of wages. *The Economic Journal*, *107*, 113–133.

Hanson, G. H. (2004). What has happened to wages in Mexico since NAFTA? Implications for hemispheric free trade. In A. Estevadeordal, D. Rodrik, A. Taylor, & A. Velasco (Eds.), *Integrating the Americas: FTAA and beyond* (chap. 15). Cambridge, MA: Harvard University Press.

Hanson, G. H., & Slaughter, M. J. (2002). Labor market adjustment in open economies: Evidence from US states. *Journal of International Economics*, *57*, 3–29.

Harrigan, J. (1997). Technology, factor supplies, and international specialization: Estimating the neoclassical model. *American Economic Review*, *87*(4), 475–494.

Harrigan, J., & Zakrajsek, E. (2000). *Factor supplies and specialization in the world economy* (Finance and Economics Discussion Paper No. 43). Washington, DC: Board of Governors of the Federal Reserve System.

Henderson, J. V. (2003). Marshall's scale economies. *Journal of Urban Economics*, *53*, 1–28.

Hughes, G., & McCormick, B. (1994). Did migration in the 1980s narrow the North–South divide? *Economica*, *61*(244), 509–527.

Kemp, M. C., & Okawa, M. (1997). Factor price equalization under imperfect competition. *Review of International Economics*, *5*(2), 153–178.

Kerr, C. (1954). The Balkanization of labor markets. In E. W. Bakke (Ed.), *Labor mobility and economic opportunity* (pp. 92–110). Cambridge, MA: MIT Press.

Krugman, P. (1993). First nature, second nature, and metropolitan location. *Journal of Regional Science*, *33*(2), 129–144.

Leamer, E. E. (1995). *The Heckscher–Ohlin model in theory and practice* (Princeton Studies in International Finance No. 77). Princeton, NJ: Princeton University.

Lerner, A. P. (1952). Factor prices and international trade. *Economica*, *19*(73), 1–15.

Marshall, A. (1890). *Principles of economics*. London: Macmillan.

Moretti, E. (2004). Workers' education, spillovers and productivity: Evidence from plant-level production functions. *American Economic Review*, *94*(3), 656–690.

Moretti, E. (2011). Local labor markets. In D. Card & O. Ashenfelter (Eds.), *Handbook of labor economics* (vol. 4B, pp. 1237–1313). Philadelphia, PA: Elsevier.

Moroney, J. R., & Walker, J. M. (1966). A regional test of the Heckscher–Ohlin hypothesis. *Journal of Political Economy*, *74*(6), 573–586.

Overman, H. G., & Puga, D. (2010). Labor pooling as a source of agglomeration: An empirical investigation. In E. L. Glaeser (Ed.), *Agglomeration economics* (pp. 133–150). Chicago, IL: University of Chicago Press.

Reeve, T. A. (2006). Factor endowments and industrial structure. *Review of International Economics*, *14*(1), 30–53.

Samuelson, P. A. (1949). International factor-price equalization once again. *The Economic Journal*, *59*(234), 181–197.

Appendix

Table A1: Human development index for 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
Toba Tek 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
Dera Ghazi Khan	30	0.611
Muzaffargarh	31	0.602
Lodhran	32	0.600
Bahawalpur	33	0.596
Rahimyar Khan	34	0.594
Rajanpur	35	0.555

Source: CREB, Lahore School of Economics, 2006.

Table A2: Negative binomial regression, 2005/06

	Number of common industries
Absolute relative wage bill gap	-0.4489***
	(0.043)
Number of industries in *r*	0.019***
	-0.0008
Number of industries in *s*	0.018***
	-0.0007
Constant	1.026*
	-0.289
Ln alpha	-2.53
Alpha	0.0949
Regional dummies	Yes
Pseudo-R^2	17.45
P-value of chi-bar statistic	0.000
Observations	1,170

Source: Author's calculations.

Table A3: Negative binomial regression, 2000/01

	Number of common industries
Absolute relative wage bill gap	-0.899***
	(0.0690)
Number of industries in *r*	0.016***
	(0.0008)
Number of industries in *s*	0.016***
	(0.0008)
Constant	0.980***
	(0.0424)
Ln alpha	-1.666
Alpha	0.1889
Regional dummies	Yes
Pseudo-R^2	16.77
P-value of chi-bar statistic	0.000
Observations	1,040

Source: Author's calculations.

Lahore School of Economics

Centre for Research in Economics & Business

Recent Working Papers

No. 01-16
Education, Health Knowledge, and Maternal Health in Pakistan
Shandana Dar and Uzma Afzal

No. 03-15
The Impact of Remittances and Parental Absence on Children's Wellbeing in Rural Punjab
Nida Jamil

No. 02-15
Natural Disasters and Social Capital: Evidence from a Field Experiment in Pakistan
Uzma Afzal, Ginger Turner and Farah Said

No. 01-15
Caste, Social Exclusion, and Opportunities for Education in Rural Punjab
Tayyaba Tamim, Marriam Musa, Hana Tariq and Shahzeb Rathore

No. 04-14
Education, Employment, and Women's Say in Household Decision-Making in Pakistan
Duryab Fatima

No. 03-14
The Impact of Parents' Subjective Aspirations on Schooling Investment in Rural Punjab
Anam Ashraf

No. 02-14
Attitudes Towards Risk in the Wake of a Rare Event: Evidence from Pakistan
Farah Said, Uzma Afzal and Ginger Turner

No. 01-14
Resource Misallocation and Aggregate Productivity in Punjab
Muhammad Haseeb and Theresa Thompson Chaudhry

No. 04-13
Labor Pooling as a Determinant of Industrial Agglomeration
Najam uz Zehra Gardezi

No. 03-13
The Effects of Agglomeration on the Formation and Scale of Operation of New Firms
Maryiam Haroon

No. 02-13
Agglomeration and Firm Turnover
Marjan Nasir

No. 01-13
Determinants of School Choice: Evidence from Rural Punjab, Pakistan
Hamna Ahmed, Sahar Amjad, Masooma Habib and Syed Ahsan Shah

No. 03-12
The Effects of External Migration on Enrolments, Accumulated Schooling, and Dropouts in Punjab
Rabia Arif and Azam Chaudhry

No. 02-12
The Determinants of Child Health and Nutritional Status in Punjab: An Economic Analysis
Uzma Afzal

Policy Papers

No. 01-11
Pakistan and Lessons from East Asia: Growth, Equity, and Governance
Khalid Ikram

No. 01-10
A Strategy for Reversing Pakistan's Dismal Export Performance
Hamna Ahmed, Mahreen Mahmud, Naved Hamid and Talal-Ur-Rahim

These papers can be accessed at: www.creb.org.pk

The Lahore School of Economics (established in 1993) is one of Pakistan's leading centres of learning for teaching and research in economics, finance and business administration. Its objectives are (i) to train young Pakistanis as professional economists, finance managers, accountants, financial analysts, bankers, and business executives, and (ii) to undertake research in economics, management, finance, and banking to deepen the understanding of major facts, issues, and policies.

The Centre for Research in Economics and Business (CREB) is an independent research centre at the Lahore School of Economics. CREB's mission is to conduct and facilitate research, coordinate and manage the Lahore School's postgraduate program, and promote discussion on policy issues facing Pakistan. The research focus at CREB is on the management of the Pakistan economy, income distribution and poverty, and the role of the modern services sector in the area of economics; and financial markets in the area of business management.

The Lahore School's publication program comprises the Lahore Journal of Economics, Lahore Journal of Policy Studies, Lahore Journal of Business, a Text Book Series, Lahore School Case Study Journal, the CREB Working Paper Series, and CREB Policy Paper Series. The program encourages both in-house and external contributors.

Lahore School of Economics
Intersection Main Boulevard Phase VI, DHA and Burki Road
Burki Lahore 53200, Pakistan