Specialization and Diversification in Agricultural Transformation: The Case of West Punjab, 1903-1992

Takashi KUROSAKI *

September 2002 (Previous versions: March 2002, August 2001)

*Institute of Economic Research, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8603 JAPAN. Phone: 81-42-580-8363; Fax: 81-42-580-8333; E-mail: kurosaki@ier.hit-u.ac.jp.

Specialization and Diversification in Agricultural Transformation: The Case of West Punjab, 1903-1992

Takashi Kurosaki

Keywords:

agricultural transformation, comparative advantage, diversification, growth accounting, structural change.

Abstract:

In this article, the role of crop specialization and diversification in agricultural transformation is investigated empirically. Changes in aggregate land productivity are associated structurally with inter-crop and inter-district reallocation of land use. Results from a region with the oldest history of agricultural commercialization in developing countries show that cropping patterns of subsistence agriculture changed substantially, with rising concentration of crop acreage in districts with higher and growing productivity. Rapid specialization in crop production was observed at the district level recently, after a phase with sporadic specialization. These changes reflected comparative advantage and contributed to the improvement in aggregate land productivity.

Takashi Kurosaki is Associate Professor, the Institute of Economic Research, Hitotsubashi University, Tokyo.

The author would like to thank two anonymous reviewers and an editor of this journal for their constructive comments and suggestions. He is also grateful for helpful comments to S. Were Omamo, Tetushi Sonobe, Kei Otsuka, Jonna Estudillo, Kaliappa Kalirajan, Odaka Konosuke, Kyoji Fukao, and seminar participants at the Foundation for Advanced Studies on International Development, Indira Gandhi Institute of Development Research, the Institute of Developing Economies, and the Annual Meeting of the Agricultural Economics Society of Japan. Funding was provided through the Asian Historical Statistics Project by the Japanese Ministry of Education. Any remaining errors are the author's.

Specialization and Diversification in Agricultural Transformation: The Case of West Punjab, 1903-1992

In this article, the role of crop specialization and diversification in the process of agricultural transformation is investigated empirically for the case of West Punjab. Changes in aggregate land productivity are associated structurally with inter-crop and inter-district reallocation of land use. This structural association enables us to characterize the nature of market development and agricultural transformation in a specific region. In the initial period of agricultural transformation, the diversity of a traditional and subsistence agriculture might go down because of crop shifts reflecting comparative advantage (Timmer, 1997). The actual pattern of specialization and diversification might differ according to aggregation levels and region/time specific factors. This article is the first attempt to quantify these patterns using actual data from a developing country.

The existing studies with similar motivations include Huffman and Evenson, who demonstrate the importance of crop specialization for productivity growth in U.S. agriculture, Sonobe and Otsuka, who quantify the role of subsectoral resource reallocation in productivity growth in Taiwanese manufacturing industries, and Timmer and Szirmai, who investigate the effects of labor shifts across the subsectors of manufacturing on labor productivity in Asia. This article is different from these studies in that resource reallocation over space is investigated explicitly using agricultural data that directly correspond to different aggregation levels.

The empirical application is to the case of West Punjab over the period 1903-1992, which roughly corresponds to the area of Pakistan Punjab today. Datasets are newly compiled by the author at more aggregate (West Punjab) and less aggregate (district) levels. This case is ideal for the objective of this article because Punjab has experienced a rapid growth of agricultural production during this period with extensive statistics available. Thanks to agricultural development, the Punjab Province is the richest among the four provinces of Pakistan, although the absolute income level in the province still remains at the level of low income countries on a global scale.

To put the scope of this article in a different way, it is an attempt to incorporate into the growth accounting literature an idea borrowed from the literature on equilibrium models for agricultural households. Regarding the former, historical records have shown that agricultural productivity has been growing due to the introduction of modern technologies, commercialization of agriculture, capital deepening, factor shifts from agriculture to nonagricultural sectors, etc. This whole process could be called 'agricultural transformation,' to which the contribution of each of these factors has been quantified in the existing literature on growth accounting using macro data (Timmer, 1988).

Since the macro statistics reflect the aggregated behavior of micro agents, it is desirable to associate quantitative changes at the macro level with those at more disaggregated level. The current article is an attempt in this direction, focusing on the role of crop specialization and diversification. The significance of crop shifts in the process of agricultural transformation can be understood through development of rural markets. If all producers choose crops on the principle of comparative advantage, and all producers face the same relative prices, land reallocation occurs only when technology or relative prices change. In agriculture, however, the assumption that all producers face the same relative prices is not justifiable because spatial dimensions and transportation costs are important in crop production (Takayama and Judge; Baulch). With substantial transportation costs, farmers may optimally choose a crop mix that *does not* maximize expected profits evaluated at market prices but that *does* maximize expected profits evaluated at farm-gate prices after adjusting for transaction costs (Omamo, 1998a; 1998b), and farmers' supply response curve has kinked or discontinuous points (Key, Sadoulet, and de Janvry).[1]

Subjective equilibrium models for agricultural households provide other reasons why decision prices for farmers may diverge from market prices. Households need to be self-sufficient in farm labor if labor markets are missing (de Janvry, Fafchamps, and Sadoulet). The objective function of farmers may include considerations for production and consumption risk and/or domestic needs for family if insurance markets are incomplete (Kurosaki and Fafchamps). In these cases, their production choices can be expressed as a subjective equilibrium evaluated at household-level shadow prices.

During the initial phase of agricultural transformation, it is likely, therefore, that diversification levels are similar between different aggregation levels because each region has to grow crops its residents want to consume due to the absence of well-developed agricultural produce markets. As rural markets develop, however, the discrepancy between the market price of a commodity and its decision price at the farm level is reduced. In other words, the development of rural markets is a process which allows farmers to adopt production choices that reflect their comparative advantages more closely, contributing to productivity improvement at the aggregate level evaluated at common, market prices. If this development occurs, production at a less aggregate level could be less diverse than that at a more aggregate level.[2] Initially, when some produce markets are thin with volatile prices and insurance markets are incomplete, farm households may participate in produce markets only marginally. As their constraints on consumption smoothing are eliminated, however, they may increase their production of lucrative crops (Kurosaki and Fafchamps). Similarly, development of rural labor markets enables farmers to grow more market-oriented crops through reduction of constraints on family labor endowments (de Janvry, Fafchamps, and Sadoulet).

The article is organized as follows. Data used in this article are explained briefly in the next section. An overview of agricultural transformation in West Punjab is presented in the second section, with a discussion of the timing of structural change in land productivity. The effects of crop specialization on agricultural growth are quantified in the third section, followed by a section in which the dynamics of crop diversification are analyzed at the district level. Findings and implications of the investigation are summarized in the final section.

Data

A dataset for this type of investigation should cover a sufficiently long period that contains the initial phase of low market development as well as the following phase of dynamic agricultural transformation. At the same time, data quality should be comparable as much as possible throughout the period. The case of West Punjab since the early twentieth century is ideal for investigation, the region having experienced rapid agricultural growth during this period, especially famous for the Green Revolution since the late 1960s. It still remains, however, as a low income, developing area with substantial potential for future development. Furthermore, among developing countries, the Indian Subcontinent is exceptional in the availability of long term agricultural statistics collected by colonial and post-independence governments. Data for this article are estimated and compiled from these sources.[3]

Punjab was partitioned between India and Pakistan in August 1947 when the two countries achieved independence. It is thus impossible to define a geographical space that *both* corresponds to the contemporary national borders *and* is time-invariant throughout the study period. Since one of the purposes of this article is to associate changes in aggregate land productivity with inter-crop and inter-district reallocation of land use, the latter condition is more important. Therefore, the western half of the

British Province of Punjab is chosen as the more aggregate level in this article and called 'West Punjab.'[4] This area accounted for approximately 90% of farm land in Pakistan Punjab in 1981.[5]

Considering the data availability, a 'district' is chosen as the less aggregate level in this article. It is a basic unit of local administration. Since district boundaries were changed occasionally, the boundaries during 1905-1919 are chosen as the reference, resulting in twenty-eight districts for the British Punjab. When two or more districts were merged and re-divided into several districts, crop areas are divided proportionally if more detailed data are not available. In 1947, the Indian Empire was partitioned into India and Pakistan. Among the twenty-eight districts of the British Punjab, fifteen districts in the west belonged to Pakistan and the rest belonged to India. Therefore, West Punjab in this article is divided into fifteen districts, whose boundaries are adjusted by the author to make them time-invariant throughout the study period.

For each of the fifteen districts, data for cropped area and output are compiled for twelve major crops — rice, wheat, barley, sorghum, pearl millet, maize, chickpea, rape and mustard, sesame, sugarcane, tobacco, and cotton. The first seven crops are foodgrains and the rest are cash crops including oilseeds. These crops comprise the 'major crops' subsector in the national income accounting of Pakistan today. The major crops subsector occupies about 70% of value-added from crops and about 40% of value-added from agriculture, and its share was higher in the colonial period. Among the twelve crops, wheat, rice, sugarcane, and cotton are the most important, accounting for 76% of the total output value from the twelve crops before independence and 87% after independence. The gross output values from these crops are aggregated using 1960 prices. Ideally, the sum of value-added evaluated at current prices and then deflated using some price index would be a better measure, but the sum of gross output values at constant prices is used as a proxy due to the absence of reliable data on input prices and quantities before independence.[6] Therefore, the analyses in the subsequent sections are based on a balanced panel of fifteen districts over ninety years (1903-1992) covering twelve crops.

Agricultural Transformation and Crop Shifts in West Punjab

Agriculture in West Punjab has undergone major structural change over the study period. Aggregate farm output (Q_t) evaluated at 1960 prices was about 7.5 times larger in the 1990s than in the early twentieth century (figure 1). Partial productivity with respect to labor (L_t) and land (A_t) improved substantially as well, with levels in the 1990s more than double the levels in the early twentieth century.[7] All three indices in figure 1 show that the average annual growth rate was significantly higher after the mid 1950s. It should be noted that the acceleration in growth occurred *before* the introduction of Green Revolution technology that enhanced the land productivity of wheat and rice dramatically since the late 1960s.

[Insert figure 1]

Before independence in 1947 also, all three indices increased gradually, although the growth rates were lower than those after independence.[8] The decade from the mid-1940s is a period when farm output stagnated. This could be due to economic disruption during World War II and the turmoil caused by the partition of Punjab in 1947, to which we return in the fourth section.

Another aspect of agricultural transformation in West Punjab is a spatial shift of crops. A diversification index of $D \equiv 1 - H$ is a convenient measure to characterize spatial crop shifts, where H is the Hirschman-Herfindahl index. This can be defined in two ways. First, a diversification index of district shares under crop i can be defined as $D_{it} = 1 - H_{it} = 1 - \sum_h (A_{hit}/\sum_k A_{kit})^2$, where A_{hit} is an area under crop i in district h in year t compiled from the dataset explained in the previous section. Intuitively, it shows the probability of hitting different districts if two points are randomly chosen from the whole area under crop i in West Punjab. A small D_{it} implies that cultivation of the crop of concern is concentrated in a few districts. Second, from the same raw data, another diversification index can be calculated, defined as $D_{ht} = 1 - H_{ht} = 1 - \sum_i (A_{hit}/\sum_j A_{hjt})^2$. Index D_{ht} has the intuitive meaning of the probability of hitting different crops if two points are randomly chosen from the whole area under cultivation in district h. Therefore, it is a measure of diversification for a particular district.

D_{it} is plotted in figure 2 for the four important crops.[9] The most dynamic change is observed in cotton, the most important cash crop in West Punjab. D_{it} of cotton continuously decreased during the ninety year period. In the early twentieth century, its regional concentration was less than those of rice and sugarcane; cotton is the crop with the highest concentration in the 1990s. An interesting pattern is observed in rice, whose concentration increased rapidly during the colonial period and has decreased slightly since the early 1940s. Rice in West Punjab is cultivated as a cash crop also. Another cash crop, sugarcane, shows a gradual increase in the diversification index during the colonial period but then it has stabilized during the post independence period. In sharp contrast to these cash crops, D_{it} for wheat, a staple food for Punjabis, is stable throughout the twentieth century. The absolute level of D_{it} is much above those of the three cash crops.

[Insert figure 2]

To examine whether or not the direction of spatial shifts shown in figure 2 is consistent with comparative advantage, table 1 reports the ratio: (i) the sum of the areas under the crop of concern in districts that are ranked the top four in terms of comparative advantage in yield disparity ranking, divided by (ii) the sum of the areas under the crop of concern in districts that are ranked the bottom four.[10] The comparative advantage in the table is approximated by the per-acre gross revenue of each crop relative to the average revenue of the other three crops. If the ratio for crop i in table 1 increases over time, it implies that the area under the crop is being concentrated in districts that have comparative advantage in producing the crop over the other three crops. For rice and cotton, the pattern shown in table 1 is a mirror image of that in figure 2 — when the diversification index declines in figure 2, the ratio in table 1 increases. For sugarcane, the ratio in table 1 shows a gradual increase over time.[11] In sharp contrast to these cash crops, the ratio of wheat shows no trend. Therefore, these observations give support to the interpretation that spatial shifts of major crops were associated with potential differences in comparative advantage.

[Insert table 1]

The concentration of each commercial crop in a few districts is likely to affect crop diversification in each district. Figure 3 plots the median, top 25%, and bottom 25% of D_{ht} in each year for the fifteen districts, together with the same index at the province level for comparison. The exact districts ranked at these positions are not the same throughout the period, but the churning among districts is not large. The figure shows first that the crop mix is more diversified at the more aggregate level than at the district level throughout the period. Second, both D_{ht} and D_t show a decline in the latter half of the study period. Third, the difference between province and district-median indices widens in the latter half. Fourth, as shown by the top and bottom 25% plots, the dynamic paths differ widely from district to district. Thus, the acceleration of aggregate land productivity growth in figure 1 since the mid-1950s is associated with a decreasing diversification index in each district.

[Insert figure 3]

Combining these observations from figures 1-3, we hypothesize that the shift of cultivated areas from less lucrative to more lucrative crops and from less productive to more productive districts was an important source of agricultural growth in West Punjab. In the next section, therefore, its effect on land productivity is quantified in a growth accounting framework, and in the fourth section, the heterogeneity among districts is investigated further.

Before these, the period demarcation needs to be determined because we do not know exactly when structural change(s) occurred, although a casual look at figures 1-3 suggests that most of the time series show a turnaround around 1947, a natural candidate for a breakdate, when India and Pakistan achieved independence. Therefore, a series of tests are conducted for a structural change of unknown timing, following the procedure by Hansen.[12] Time series for the tests include Q_t/A_t, which is further analyzed in the next section, and H_t and H_{it}, which are the variables of concern in the fourth section.

First, a time series model with a deterministic trend is estimated by ordinary least squares (OLS),

$$\ln(Q_t/A_t) = a_0 + a_1 D_\tau + (b_0 + b_1 D_\tau)t + u_t, \tag{1}$$

where D_τ is a dummy variable for years after the structural change at time τ, and u_t is an error term with zero mean. For all candidate breakdates, Chow statistics for the null hypothesis of no structural change ($H_0 : a_1 = b_1 = 0$) are estimated and their sequence is plotted as a function of candidate breakdates in figure 4. The Chow test sequence reaches a high of 29.6 in 1951, which is the Quandt statistic, followed by 27.7 in the next year.

[Insert figure 4]

Under the assumption that the breakdate is unknown a priori, Bai and Perron provide the critical value at 1% for this case, which is 16.6. Since this is well below the test statistics in 1951 or 1952, the hypothesis of no structural break is rejected. Based on the breakdate estimate at 1951 or 1952, the sample is then split in two and the test is re-applied to each subsample, following Bai and Perron's sequential procedure. The hypothesis of two or more structural breaks is not supported by the data.

Second, a similar model is estimated for D_t using OLS, resulting in a Quandt statistic of 21.1 in 1952. The hypothesis of no structural break is rejected and that of two or more structural breaks is not supported. For D_{ht}, fifteen time series are stacked up and a similar model with deterministic trends is estimated by seemingly unrelated regression (SUR). Under the assumption of common breaks, the Chow statistic reaches its highest value of 29.3 in 1953, followed closely by 29.1 in 1952, allowing rejection of the hypothesis of no structural break. Here again, based on the breakdate estimate at 1952 or 1953, the sequential test cannot reject the hypothesis of only one breakdate.

Therefore, all three time series have one structural break in the early 1950s. The least squares estimates for the breakdate are 1951 for $\ln(Q_t/A_t)$, 1952 for D_t, and 1953 for D_{ht}. We adopt the agricultural year 1952 as the breakdate in the following analysis, because it is the median of the three and is associated with the second highest Chow statistic for $\ln(Q_t/A_t)$ and D_{ht} with a very small difference.

Crop Shifts as a Growth Source in West Punjab

In this section, the effects of crop shifts on land productivity are investigated. Conceptually, crop shifts may affect agricultural growth in two different ways. First, how much of the agricultural growth in West Punjab can be explained by an aggregate shift from low value-added crops to high value-added crops? This aspect, i.e., the contribution of *inter-crop* land reallocation to agricultural growth is quantified in the first subsection using province-level data. Since trade with areas outside West Punjab is critically important to changes in the aggregate share of each crop, this exercise sheds light on the role of trade on growth from an angle different from cross-country growth regressions or multi-sector growth accounting from the demand side (Chenery).

Second, using district-level data, the contribution of *inter-district* land reallocation to agricultural growth can be quantified. With developed rural markets, individual farmers or individual districts can specialize in crops in which they have comparative advantage, leading to productivity improvement at the aggregate level. This aspect corresponds to the role of agricultural commercialization within West Punjab.

Effect of Inter-Crop Land Reallocation on Land Productivity

Define $Y_t \equiv Q_t/A_t$, which shows land productivity at the aggregate level in terms of real output value per acre. By decomposing the total production into subsectors comprising various crops denoted by i, the agricultural growth from year 0 to year t can be decomposed as

$$\begin{aligned}(2)\quad \ln(Q_t/Q_0) &= \ln(A_t/A_0) + \ln(Y_t/Y_0) \approx \frac{A_t - A_0}{A_0} + \frac{Y_t - Y_0}{Y_0} \\ &= \frac{A_t - A_0}{A_0} + \frac{1}{Y_0}\left[\sum_i s_{i0}(Y_{it} - Y_{i0}) + \sum_i (s_{it} - s_{i0})Y_{i0} + \sum_i (s_{it} - s_{i0})(Y_{it} - Y_{i0})\right],\end{aligned}$$

where $s_{it} \equiv A_{it}/\sum_k A_{kt}$, which is the area share of crop i in year t. The first term of the last expression, i.e., $(A_t - A_0)/A_0$, shows area effects and the remaining block within a bracket shows land productivity effects. The land productivity term is further decomposed into three elements: aggregate crop yield effects, inter-crop shift effects, and a residual. Following the terminology of Timmer and Szirmai, the second term is called ‘static’ inter-crop shift effects and the third term is called ‘dynamic’ inter-crop shift effects in this article. The second term shows ‘static’ effects since it becomes more positive when the area under crops whose yields were initially high increases relatively.[13] In contrast, the third term shows ‘dynamic’ effects because it becomes more positive when the area under dynamic crops (i.e., crops whose yields are improving) increases relative to the area under non-dynamic crops.

The decomposition results following equation (2) can be interpreted in several ways. One is to focus on the contribution of crop shift effects relative to the total growth of land productivity. This is in line with the existing literature on growth accounting and is discussed below. Another is to calculate the contribution of crop shift effects relative to the maximum potential for the crop shift effects. The maximum potential for the static shift effects depends on the initial level of specialization/diversification and the range of feasible crop shares that can be grown under agronomic constraints, among which irrigation and rainfall are the most important. The maximum potential for the dynamic shift effects depends on the same factors as for the static effects and the response of Y_{it} to changes in s_{it}. Since the current dataset does not allow a precise estimation of these factors, this exercise is not attempted in this article.[14]

[Insert table 2]

Table 2 shows the decomposition results. In West Punjab, production increased at 2.0% per year in the first period until 1952 and the growth rate has accelerated to 5.6% since then. Area effects explain 71% of the first period growth, whereas land productivity effects account for 68% of the second period growth. These results are in line with those posited by Hayami and Ruttan in their international comparisons of the sources of agricultural growth.

Of this improvement in aggregate land productivity, growth in aggregate crop yield explains 57 and 78% respectively. The rest is explained both by static and dynamic shift effects. The absolute level of shift effects, both static and dynamic, is lower in the first period than in the second period. In the first period, dynamic shift effects are more important than static shift effects, whereas in the second period, the relative contribution of dynamic shift effects is similar to that of static shift effects. By decades, Kurosaki (2001) shows that the contribution of the static shift effects to productivity growth was the highest during the period 1952-62, surpassing the contribution from growth in aggregate crop yield, which supports the conjecture in Kurosaki (1999) that land reallocation toward high value crops was one of the main engines of agricultural growth, especially before the Green Revolution. These results show that the *inter-crop* land reallocation was an important source of land productivity growth in West Punjab.

Effect of Inter-District Crop Shifts on Aggregate Crop Yields

The aggregate crop yield effect for crop i in equation (2) can be further decomposed as

$$(3) \qquad Y_{it} - Y_{i0}$$
$$= \sum_h s_{hi0}(Y_{hit} - Y_{hi0}) + \sum_h (s_{hit} - s_{hi0})Y_{hi0} + \sum_h (s_{hit} - s_{hi0})(Y_{hit} - Y_{hi0}),$$

where $s_{hit} \equiv A_{hit}/\sum_k A_{kit}$, which is the share of district h in the cultivated area of crop i in year t. Three terms on the right hand side of equation (3) are interpreted similarly to terms in equation (2): the first term shows the effects of the average crop yields in the district (we label them 'District crop yield effects' in the table), the second term indicates 'Inter-district crop shift effects (static)'[15] and the third term shows 'Inter-district crop shift effects (dynamic).'

An important aspect of equation (3) is that the effects of factor reallocation over space are incorporated explicitly in the decomposition. In other words, so-called 'yield effects' in the existing literature based on macro data are often a mixture of pure yield effects (e.g., due to shifts in total factor productivity in producing individual crops) and spatial crop shift effects.

In interpreting the decomposition results from equation (3), two points should be noted in addition to the unit of normalization. First, for crop reallocation across districts to have a substantial impact on aggregate land productivity, agronomic conditions need to be heterogeneous and the extent to which land can be switched across crops without loss of yield should be high. Otherwise, specialization in production may not be possible or efficient. As regards the agronomic conditions, soil quality, irrigation, rainfall, temperature, etc. vary across districts in West Punjab. The fact that the ranking of the twelve crops in a district in terms of per-acre gross revenues differ substantially across districts is consistent with heterogeneous agronomic conditions.[16] Furthermore, the ranking of the fifteen districts in producing a crop in terms of per-acre gross revenues did not change much over the study period. This suggests that the actual land switching across crops within a district occurred without a substantial loss of yield.

The second point is spatial resolution of the data. In principle, the decomposition formula in (3) can be applicable to any spatial unit of crop production h, as long as the more aggregate unit of concern is exactly equal to the sum over h. Depending on data availability, the unit can be either a field plot, a farm household, a village, or a region. If one has data on a province divided into districts and into all farmers, the magnitude of crop shift effects is likely to larger when the province-level yield growth is decomposed into the farm-level growth than when it is decomposed into the district-level growth. For this reason, the district boundaries in this article are adjusted so that they are time-invariant and homogeneous in the geographical size. The discussion below is conditional on this particular choice of spatial resolution and therefore not comparable to other spatial resolutions.

[Insert table 3]

Table 3 shows the decomposition results. Due to data availability, it does not include the first decade covered in table 2, but this difference does not affect our results qualitatively.[17]

Wheat yields did not grow at all in the first period. During the second period, however, wheat yields grew at 2.9% per year on average. Table 3 shows that only 6% of this growth can be attributable to crop shifts across districts, implying that the change in the average wheat yield in the district is the dominant source of growth in this period. This does not imply that spatial crop shift effects are negligible since the table quantifies the *inter-district* crop shift effects only and it is possible that *intra-district*, *inter-household* crop shift contributed to the yield growth at the district level through a decomposition mechanism similar to equation (3). Nevertheless, considering the fact that wheat is a staple food for Punjabi farmers and it is grown all over Punjab, the results in table 3 seem to suggest that spatial reallocation is not very important in the case of wheat.

Results for rice show a striking contrast to wheat. In the first period, when aggregate rice yields grew at 0.27% (statistically significant at 1% level), the major source of yield growth is crop shifts across districts. More than 60% each of yield growth is attributable to static and dynamic crop shift effects respectively (the effects of rice yield growth at the district level are negative). During the second period, the rice yield growth at the district level is the only source of growth as the case of wheat.

In the case of sugarcane, inter-district crop shift effects are not substantial in both periods. The aggregate yield of sugarcane thus grew mainly through improvements in yields at the district level.

The case of cotton presents an interesting pattern. Dynamic crop shift effects are important in both periods. They explain more than one fourth of aggregate yield growth of cotton. On the other hand, static crop shift effects are nil. This implies that the improvement of cotton yields in the province was facilitated by moving cotton areas from districts with stagnating or decreasing productivity to districts with increasing productivity. Because those districts with more rapid growth in cotton yields were those with low yields in the initial years, the static effects did not contribute to the yield growth at the aggregate level. Expansion of cotton production for both domestic and foreign markets was the most important development of Punjab's agriculture during the colonial period. Pakistan's economy today is also heavily dependent on cotton production in Punjab. Our investigation has shown that the land productivity of cotton improved not only through improvements of crop yields at the district level but also through reallocation of cultivated land to districts whose cotton yields were improving rapidly.

The evidence in table 3 shows the importance of inter-district crop shifts for cash crops like rice and cotton in facilitating yield growth at the province level. Combining this finding with those in the previous section, we can conclude that the historical change in West Punjab in the study period is consistent with crop shifts reflecting static and dynamic comparative advantage.

Dynamics of Crop Diversification at the District Level

As shown in figure 3, the diversification indices declined over the ninety year period with acceleration in recent decades, but the inter-district variation is large. Table 4 shows the distribution of time trends during the two periods demarcated by the agricultural year 1952, estimated by a model of equation (1) with $\ln(Q_t/A_t)$ replaced by D_t or D_{ht}.

[Insert table 4]

At the aggregate level, the trend in D_t is moderately positive in the first period, while it is substantially negative in the second period. At the district level in the first period, some districts are associated with a positive trend in D_{ht}, some with no trend, and others with a negative trend. In sharp contrast, all but one districts are associated with a negative and significant trend in the second period. The mean of the estimates for the district-level time trend before 1952 is almost twice as large in its absolute value as the estimate for the aggregate level. These suggest that crop specialization occurred only sporadically among districts in the first period, whereas it was accelerated on average, affecting all the districts, in the second period.

What forces are behind these trends of D_{ht}? To approach this question, determinants of the dynamics of crop diversification indices at the district level are explored in this section. By clarifying the district attributes that explain the difference in trends across districts and between two periods, we can speculate on the historical development of markets behind these dynamics. With this motivation, we estimate a simple panel model for the determinants of D_{ht},

$$D_{ht} = a_h + (b_0 + \sum_k b_k X_{hk})t + u_{ht}, \quad (4)$$

where a_h is a district fixed effect that determines the level of diversification, t is a scalar that captures a time trend, X_{hk} is a district attribute that determines the growth rates of diversification, b_0 and b_k are coefficients to be estimated, and u_{ht} is an error term with zero mean. We implicitly assume that the *level* of diversification may also be a function of district attributes but we do not estimate them explicitly since district fixed effects absorb their effects completely. Since X_{hk} is multiplied by t, a positive value of b_k implies that a district with higher X_{hk} experienced a more rapid diversification, or a less rapid specialization.

This model is applied separately to the two periods, using fixed effects panel estimation. As potentially important determinants of the trend of diversification indices, three district attributes are included in X_{hk}: the initial diversification level, irrigation ratio, and road density.[18] As already discussed, the maximum potential for crop shift is determined by the initial level of specialization/diversification and the range of feasible crop shares that can be grown under agronomic constraints. The irrigation ratio represents such agronomic constraints. Road density represents market infrastructure because the development of rural markets is key to actual changes in cropping patterns.[19]

[Insert table 5]

The intercepts on a time trend are not significant in the first period but significantly negative in the second period. Their magnitudes are similar to the mean of the estimates for the district-level time trends in table 4, with the factor of 100 due to normalization.

The coefficient on the initial diversification level is negative in both periods but statistically significant at 1% only in the second period. The significant effect in the second period suggests that a district that was more diversified at the beginning is able to specialize more because it had more potential in that direction of crop shifts. In other words, this variable controls for the potential in specialization due to the initial difference.

The effects of the irrigation ratio is statistically significant at 1% in both periods but with the opposite signs. More irrigated districts are associated with more positive trends before 1952, whereas they are associated with more negative trends after 1952. In general, the irrigation development makes it technically feasible to grow diverse crops. During the colonial period, when large-scale irrigation was first introduced, crop shifts to high value-added crops implied a more diversified cropping pattern, thereby contributing to an increase in the diversification index. In contrast, after the early 1950s, when large-scale canal irrigation was becoming an obsolete technology, further shifts to high value-added crops in highly irrigated districts implied more specialization. This interpretation is also consistent with the shapes of figures 2 and 3. Crop shifts reflecting comparative advantage do not necessarily imply specialization at the regional level. When the agronomic conditions do not allow a large scale cultivation of high value-added crops, the removal of the agronomic constraints is likely to lead to an increase in the diversification index, as was the case of irrigation in West Punjab during the colonial period.

The coefficient on road density also takes the opposite signs — districts with better road networks are associated with moderately positive trends before 1952, whereas they are associated with substantially negative trends after 1952. The strongly negative effect in the second period seems to suggest that the road infrastructure helps specialization in agriculture when specialization occurs all over the districts. We can interpret this as the effect of reduced transportation costs in narrowing the gap between the market price of a commodity and its farm-gate price. The positive effect in the first period (although its significance level is only marginal) seems to suggest that the initial development of roads occurred only locally so that its impact on district-level specialization was weak. The evidence here is therefore mixed, but the clearer result for the second period is consistent with the story of market development facilitating inter-regional trade, which reflects comparative advantage of each region.

By combining results in table 5 with those in the previous sections, the following pattern of agricultural transformation is indicated for the case of West Punjab. First, land reallocation to more lucrative crops and to districts with more comparative advantages began during the colonial period with non-negliglble effects on aggregate land productivity. But its impact on specialization in each district differed widely — some districts with better irrigation diversified their cropping patterns while other districts experienced little trends in diversification indices. The reason for specialization not to occur at the provincial level and in some districts could also be attributed to external factors during the last decade of the first period in West Punjab. These factors include the disruption of world trade during World War II, the introduction of a nation-wide market control under the 'War Economy' in the British Empire of India, and the turmoil caused by the partition of Punjab between India and Pakistan in 1947. The disruption continued for a while even after 1947, as Hindu middlemen, who were major market agents in colonial Punjab, migrated to India and they were not replaced immediately. Under these conditions, which were not suitable for spatial market development, it is reasonable for the specialization process to be stagnant.

After the early 1950s, district-level diversification indices went down rapidly in almost every district. This could be attributed to geographical integration of agricultural produce markets across districts. As the Punjab economy was recovering from the chaotic Partition in the early 1950s, Muslim middlemen began to fill up the vacancy in agricultural marketing. Agricultural produce markets gradually became more efficient in the second period.[20] This integration drove each district to grow more crops in which it had comparative advantage, resulting in a rapid specialization at the district level.

Conclusion

In this article, the role of crop specialization and diversification in the process of agricultural transformation has been investigated empirically for the case of West Punjab. In the analysis, changes in aggregate land productivity are associated structurally with inter-crop and inter-district reallocation of land use. The empirical part is based on newly-compiled production data of Punjab's agriculture for the period 1903-1992.

Quantitative results show that, first, cropping patterns of a traditional and subsistence agriculture changed substantially over the period with rising concentration of crop acreage in districts with higher and growing productivity, which contributed to the improvement in land productivity at the aggregate level. This change is therefore consistent with crop shifts reflecting static and dynamic comparative advantage. The crop shifts toward high value-added crops are discernible throughout the study period.

Second, the crop diversification level also changed as the cropping patterns changed. The pace of the change in diversity varied over space and between time periods. In the first period, the time trend of the diversification indices was not significant at the province level and took both negative and positive signs at the district level. In contrast, in the second period, the diversity level went down at both levels, but at a lower pace at the aggregate level. The rapid specialization at the district level in the second period suggests that spatial market development is accelerated only after a phase with sporadic specialization. The case of West Punjab thus shows that, even in a region with the oldest history of commercialization of agriculture in developing countries, two phases could be distinguished in the specialization process.

Notes

1 In the growth accounting framework, a similar phenomenon is more often interpreted as a disequilibrium. If not all producers choose activities on the comparative advantage principle, there is a room for growth by reallocating resources in a way closer to maximize profits. Then output can increase without technological or price changes, yielding a so called 'disequilibrium' effect in the literature on inter-sectoral factor reallocation (Syrquin, 1984; 1988).

2 Timmer (1997) stylizes the contrast between the diversification of national food production and that of farm level production for Asian agriculture. He speculates that, as agricultural transformation continues further, the diversification level of national production might go up again because improved commercialization of agriculture allows ecologically diverse regions to pursue their comparative advantage. In this phase also, the diversification level at the farm level is likely to go down, since the principle of comparative advantage continues to work, driving each farm to specialize in the activity it can perform the best. See also Omamo (1998b) for an explicit quantification of the effects of transportation costs on the farm-level diversification.

3 District-level data before independence were drawn from various issues of *Report on the Season and Crops of the Punjab* beginning from 1901/02. Post independence data were compiled from government publications of agricultural statistics. See Kurosaki (2000) for details. All the data used in this article are available from the author on request.

4 In related studies based on the same data sources, Kurosaki (1999) analyzes agricultural growth and changes descriptively in areas that correspond to the current borders of India and Pakistan. Kurosaki (2000) presents the compiled datasets as well as the details of data compilation procedures for the two countries.

5 Time in this article is counted by an agricultural year in Pakistan Punjab. For

instance, the agricultural year 1981 is a period from July 1980 to June 1981. It is simply denoted as '1981' in the article.

6 Three points should be made about the choice of aggregation procedure. First, to infer possible bias from using output values instead of value-added figures, we investigated the value-added ratio to the sum of gross output values since the early 1950s, when Pakistan began to estimate national income statistics. The ratio was stable until the 1970s. During the 1970s, the ratio decreased gradually and was stabilized again in the early 1980s. Second, our procedure neglects changes in comparative advantage due to changes in real prices of competing crops and factors. To infer possible bias from neglecting long term price changes, other base years (e.g., 1939 and 1981) were also tried for aggregation weights (Kurosaki, 2000). The results reported in this article are insensitive to the choice of base years. Nevertheless, it is possible that short term price changes also affected comparative advantage, whose impacts are left for future study. Third, our method of calculating output values per acre implies that 'comparative advantage' discussed in the empirical part of this article is conditional on these prices. It may not reveal the comparative advantage evaluated at social shadow prices if government seriously distorts relative prices. It is likely that this bias is small because the results reported in this article are qualitatively similar to those based on 1939 prices, when there were little direct interventions into agricultural markets by the colonial government.

7 The number of agricultural workers (L_t) was estimated by interpolating census estimates for the sum of 'cultivators' and 'agricultural laborers.' For original data sources and intermediate estimates of the undivided India, the study by Ono and Saito was used as a reference. A_t is defined as the sum of areas under the twelve crops.

8 OLS regression of the log of data in figure 1 on a time trend yielded the following

estimates for growth rates:

	Before 1947	After 1947
Q_t	1.53%	3.69%
Q_t/L_t	1.05%	1.32%
Q_t/A_t	0.53%	1.57%

All of them were statistically significant at 1%.

9 To remove temporal variation due to weather shocks and others, figures 2 and 3 and tables 1-3 employ data after taking a moving average over three years (area data) or five years (production data). In other parts of the article, where regression methods are used to smooth out the temporal variation, the original data are used without taking the moving averages.

10 The first decade is not covered in table 1, because only the extrapolated data are available for output data before 1908. District-level yield data were not published in the original data sources until 1908. The extrapolated data are not used in the exercises for tables 1 and 3 since inter-district variation in crop yields is important. On the other hand, since province-level output figures are available even before 1908, other tables and figures cover the full period.

11 The ratios for sugarcane in earlier periods are abnormally small because sugarcane yields in minor districts were exceptionally high due to its limited cultivation on a very few progressive farms.

12 See Kurosaki (2001) for empirical results when the study period is divided into pre-1947 and post-1947 periods and into decades.

13 Note that $(s_{it} - s_{i0})Y_{i0} = (s_{it} - s_{i0})(Y_{i0} - \bar{Y_0})$, where $\bar{Y_0}$ is the crop average of per-acre output values in the initial period.

14 A mechanical way of calculating the maximum potential is to assume that the

range of feasible crop shares is defined by the maximum number of crops grown at 12, the minimum number at 1, $0 \leq s_i \leq 1$ for all i, and $\sum_i s_i = 1$, and to treat s_{i0}, Y_{i0}, and $Y_{it} - Y_{i0}$ as given. Results are available from the author when the hypothetical shares that maximize the crop shift effects under these assumptions are used to normalize the crop shift effects.

15 Note that $(s_{hit} - s_{hi0})Y_{hi0} = (s_{hit} - s_{hi0})(Y_{hi0} - \bar{Y}_{i0})$, where $\bar{Y}_{i0}$ is the district average of per-acre crop values in the initial period.

16 This ranking was calculated when table 1 was compiled.

17 See footnote 10.

18 All the explanatory variables of X_{hk} are normalized by their means and standard deviations. The irrigation ratio is a moving average over three years with mid year 1922, when the Indus irrigation canal network was almost completed. Since road conditions changed significantly before and after independence, road density is compiled separately for the two periods. Because the only statistics available are sporadic and based on different definitions, the road density for the first period is calculated from the total kilometers of metalled and unmetalled roads managed by provincial and district governments in 1905, and that for the second period is calculated from the total kilometers of roads maintained by government highway departments in 1985. To avoid potential endogeneity bias (roads being constructed in areas with larger marketed surplus, i.e., more specialization), the 1985 measure was instrumented with the 1905 one and other variables for the first period.

19 In the initial trials, other agronomic constraints such as soil quality index, rainfall, and their cross terms with the irrigation ratio and other market development indicators such as urban population ratio were also tried. However, since the number of districts is so small that we can incorporate only a limited number of explanatory variables in X_{hk}, we restricted our attention to models with only one variable each

or one principal component each from the three factors, i.e., the initial diversification level, agronomic constraints, and market development indicators. In the case of the former, a set of three variables was chosen from all the possible combinations under a criterion of the least squared residuals. Since the signs and significance levels of the result were similar to those obtained from the principal component model, the one-variable-each model in table 5 was chosen because it has an advantage of easy interpretation.

20 See Kurosaki (1996) for empirical evidence of spatial market integration of wheat markets in the 1980s.

References

Bai, J., and P. Perron. "Estimating and Testing Linear Models with Multiple Structural Changes." *Econometrica* 66 (January 1998): 47-78.

Baulch, B. "Transfer Costs, Spatial Arbitrage, and Testing Food Market Integration." *Amer. J. Agr. Econ.* 79 (May 1997): 477-487.

Chenery, H. "Growth and Transformation." *Industrialization and Growth: A Comparative Study.* H. Chenery, S. Robinson, and M. Syrquin, eds., pp.13-36. New York: Oxford University Press, 1986.

de Janvry, A., M. Fafchamps, and E. Sadoulet. "Peasant Household Behavior with Missing Markets: Some Paradox Explained." *Economic Journal* 101 (November 1991): 1400-17.

Hansen, B.E. "The New Econometrics of Structural Change: Dating Breaks in U.S. Labor Productivity." *J. of Econ. Perspectives* 15 (Fall 2001): 117-128.

Hayami, Y., and V.W. Ruttan. *Agricultural Development: An International Perspective*, revised and expanded ed. Baltimore: Johns Hopkins University Press, 1985.

Huffman, W.E., and R.E. Evenson. "Structural and Productivity Change in US Agriculture, 1950-1982." *Agr. Econ.* 24 (January 2001): 127-147.

Key, N., E. Sadoulet, and A. de Janvry. "Transactions Costs in the Agricultural Household Supply Response." *Amer. J. Agr. Econ.* 82 (May 2000): 245-259.

Kurosaki, T. "Government Interventions, Market Integration, and Price Risk in Pakistan's Punjab." *Pakistan Development Review* 35 (Summer 1996): 129-144.

——. "Agriculture in India and Pakistan, 1900-95: Productivity and Crop Mix." *Economic and Political Weekly* 34 (December 25, 1999): A160-A168.

——. "Compilation of Agricultural Production Data for India and Pakistan Areas, c.1900-1990." COE Discussion Paper No.D99-24, Hitotsubashi University, Tokyo, 2000.

——. "Specialization and Diversification in Agricultural Transformation: The Case of Rural Punjab, c.1900-1995." Discussion Paper Series A No.406, Hitotsubashi University, Tokyo, 2001.

Kurosaki, T., and M. Fafchamps. "Insurance Market Efficiency and Crop Choices in Pakistan." *J. of Dev. Econ.* 67 (April 2002): 419-453.

Omamo, S.W. "Transport Costs and Smallholder Cropping Choices: An Application to Siaya District, Kenya." *Amer. J. Agr. Econ.* 80 (February 1998a): 116-123.

——. "Farm-to-Market Transaction Costs and Specialisation in Small-Scale Agriculture: Explorations with a Non-separable Household Model." *J. of Dev. Studies* 35 (December 1998b): 152-163.

Ono, S., and O. Saito. "Senzen Indo no Yugyo Jinko Suikei," (Estimation of the Population of Actual Workers in Pre-War India). COE Discussion Paper No.D99-26, Hitotsubashi University, Tokyo, 2000.

Sonobe, T., and K. Otsuka. "A New Decomposition Approach to Growth Accounting: Derivation of the Formula and its Application to Prewar Japan," *Japan and the World Economy* 13 (January 2001): 1-14.

Syrquin, M. "Resource Reallocation and Productivity Growth." *Economic Structure* *and Performance: Essays in Honor of Hollis B. Chenery.* M. Syrquin, L. Taylor, and L.E. Westphal, eds., pp.75-101. New York: Academic Press, 1984.

——. "Patterns of Structural Change." *Handbook of Development Economics, Vol I.* H. Chenery and T.N. Srinivasan, eds., pp.203-273. Amsterdam: North-Holland (Elsevier Science B.V), 1988.

Takayama, A., and G. Judge. *Spatial and Temporal Price and Allocation Models.* Amsterdam: North Holland, 1971.

Timmer, C.P. "The Agricultural Transformation," *Handbook of Development Economics, Vol I.* H. Chenery and T.N. Srinivasan, eds., pp.275-331. Amsterdam: North-Holland (Elsevier Science B.V), 1988.

——. "Farmers and Markets: The Political Economy of New Paradigms." *Amer. J. Agr. Econ.* 79 (May 1997): 621-627.

Timmer, M.P., and A. Szirmai. "Productivity Growth in Asian Manufacturing: The Structural Bonus Hypothesis Examined." *Structural Change and Economic Dynamics* 11 (December 2000): 371-392.

Table 1. Concentration of Cropped Area in Districts with Comparative Advantage

	Wheat	Rice	Sugarcane	Cotton
1912	1.20	0.60	0.09	0.67
1922	1.73	1.45	0.13	1.83
1932	1.39	2.92	0.16	4.12
1942	1.40	3.46	0.99	5.80
1952	1.18	4.75	0.45	4.61
1962	1.25	3.39	1.01	16.10
1972	2.04	2.39	0.82	13.92
1982	1.22	2.30	0.88	13.28

Notes:

(1) The number in this table shows the following ratio:

(the sum of the areas under the crop of concern in districts that are ranked the top four in terms of comparative advantage in yield disparity ranking) / (the sum of the areas under the crop of concern in districts that are ranked the bottom four), where the comparative advantage is approximated by the per-acre gross revenue of each crop to the average revenue of the other three crops.

(2) Districts with a negligible area under crop of concern (less than 10ha) are excluded from the analysis.

Table 2. Decomposition of Agricultural Growth in West Punjab, 1903-1992

	Average annual growth rates (%)				
	Area effects	Land productivity effects			
		Sub-Total	Aggregate crop yield effects	Inter-crop shift effects (static)	Inter-crop shift effects (dynamic)
1903 to 1952	1.43	0.58	0.33	0.06	0.19
	(71.2)	(28.8)			
		(100.0)	(56.5)	(10.1)	(33.5)
1952 to 1992	1.81	3.83	3.01	0.41	0.42
	(32.1)	(67.9)			
		(100.0)	(78.4)	(10.7)	(10.9)

Notes:

(1) See equation (2) for the decomposition formula.

(2) Average annual growth rates are calculated by simple interest.

(3) Numbers in parenthesis show relative contribution to the growth (%).

Table 3. Contribution of Inter-District Crop Shifts to Growth in Aggregate Crops Yields

	Average annual growth rates (%)			
	District crop yield effects	Inter-district crop shift effects (static)	Inter-district crop shift effects (dynamic)	Total
A. Wheat				
1912 to 1952	-0.03	0.02	0.06	0.04
	(-76.8)	(40.2)	(136.7)	(100.0)
1952 to 1992	2.69	0.08	0.10	2.86
	(93.9)	(2.8)	(3.3)	(100.0)
B. Rice				
1912 to 1952	-0.08	0.18	0.17	0.27
	(-27.4)	(64.0)	(63.4)	(100.0)
1952 to 1992	0.70	-0.03	-0.04	0.63
	(110.2)	(-4.2)	(-6.0)	(100.0)
C. Sugarcane				
1912 to 1952	2.44	0.13	-0.04	2.53
	(96.4)	(5.0)	(-1.5)	(100.0)
1952 to 1992	0.52	0.01	0.00	0.53
	(98.2)	(2.0)	(-0.2)	(100.0)
D. Cotton				
1912 to 1952	1.26	-0.04	0.44	1.66
	(75.9)	(-2.5)	(26.7)	(100.0)
1952 to 1992	4.22	0.01	1.72	5.96
	(70.8)	(0.2)	(28.9)	(100.0)

Notes:

(1) See equation (3) for the decomposition formula.

(2) and (3) See Table 2.

Table 4. Time Trends of Diversification Indices (D_t and D_{ht})

	Period 1 (1903-1952)	Period 2 (1953-1992)
West Punjab (OLS estimates)		
Parameter estimates	0.0006	-0.0012
Standard error	0.0002	0.0001
Distribution of parameter estimates for 15 districts (SUR estimates)		
Mean	0.0001	-0.0024
Standard deviation	0.0013	0.0014
Number of districts with:		
Statistically significant and positive trend	6	0
Statistically non-significant trend at 5%	4	1
Statistically significant and negative trend	5	14

Table 5. Determinants of District-Level Diversification Indices (Fixed-effects panel estimation)

	Dependent variable = D_{ht}					
	Period 1 (1903-1952)			Period 2 (1953-1992)		
District fixed effects	(jointly significant at 1%)			(jointly significant at 1%)		
Effects on trend						
Intercept	0.0117	(1.49)		-0.2357	(27.80)	***
Initial diversification level	-0.0066	(0.78)		-0.0254	(2.74)	***
Irrigation ratio	0.0861	(9.21)	***	-0.0279	(3.05)	***
Road density	0.0155	(1.67)	*	-0.0739	(8.11)	***
R-squared	0.761			0.901		
Adjusted R-squared	0.755			0.898		

Notes:

(1) Absolute values of t-statistics are denoted in parentheses, with *** denoting significance at 1%, ** at 5%, and * at 1% (two-sided test).

(2) The three explanatory variables are normalized as $(X^*_{hk} = (X_{hk} - \bar{X}_k)/\sigma_{Xk} * 100$, where σ_{Xk} is a sample standard deviation of X_{hk} across h. Therefore, the parameter under "Intercept" shows the trend of D_{ht} when all the other three explanatory variables are set at their means.

(3) The number of observations is 750 in Period 1 (15 districts x 50 years) and 600 in Period 2 (15 districts x 40) in period 2.

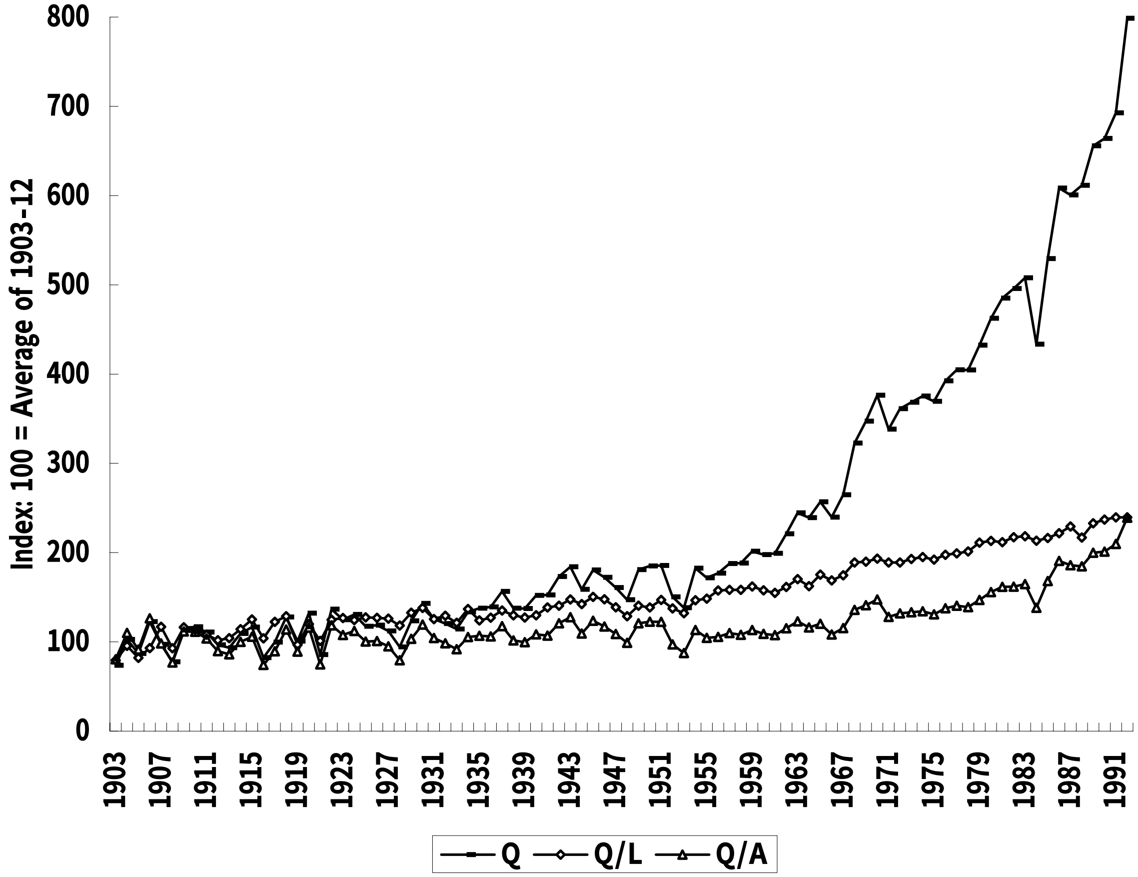

Figure 1. Agricultural output in West Punjab, 1903-1992

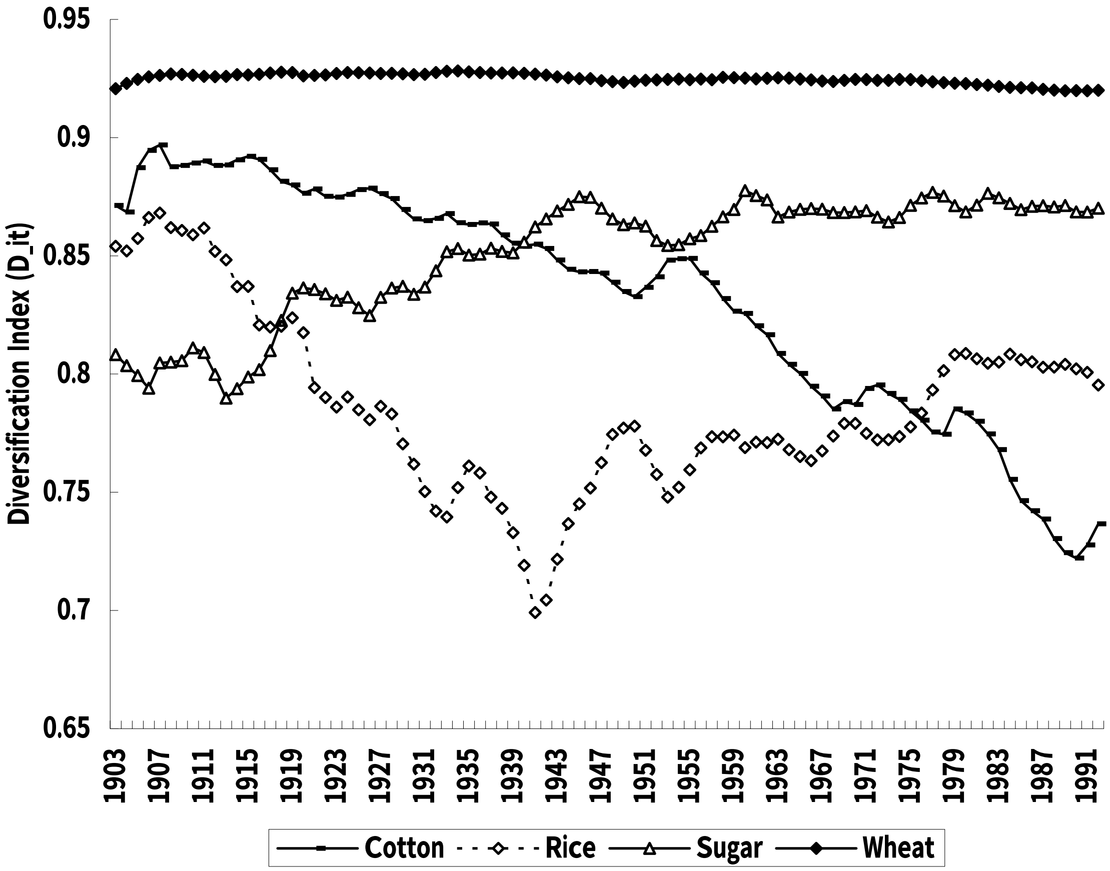

Figure 2. Concentration of production of the four important crops in West Punjab

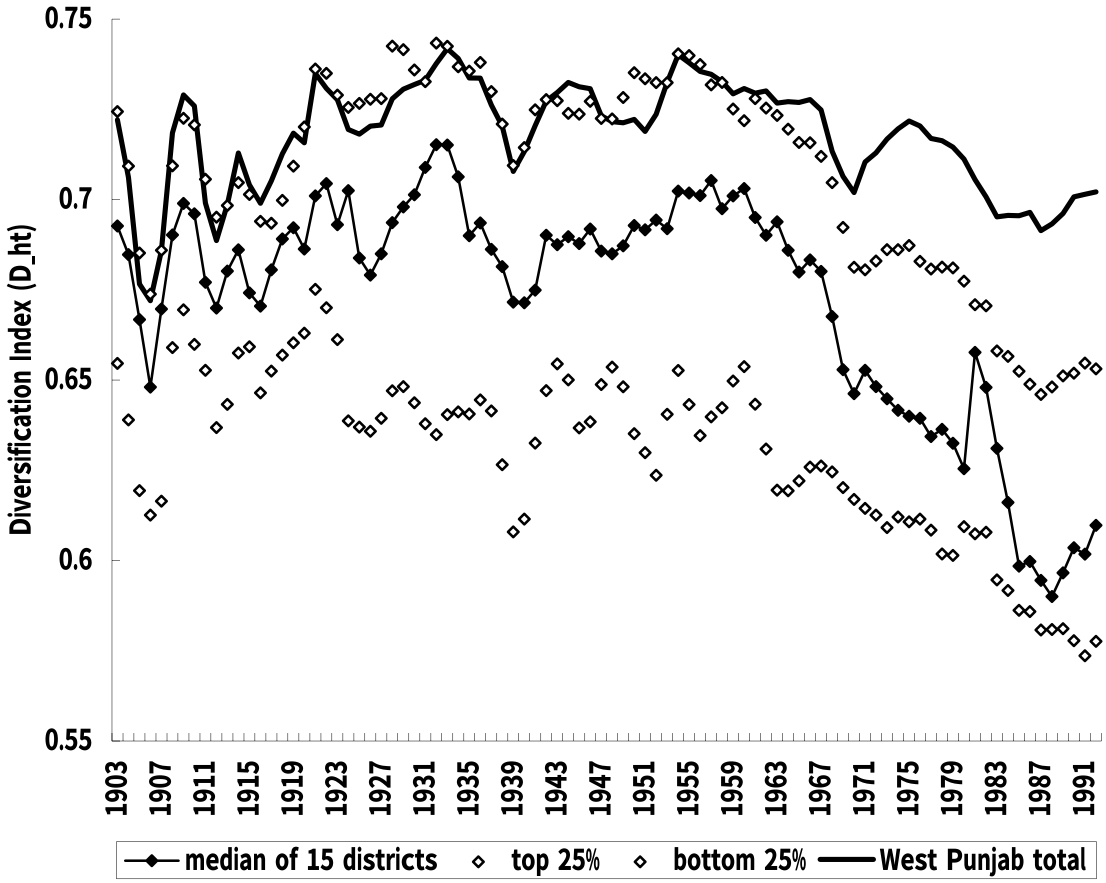

Figure 3. District and province level crop diversification in West Punjab

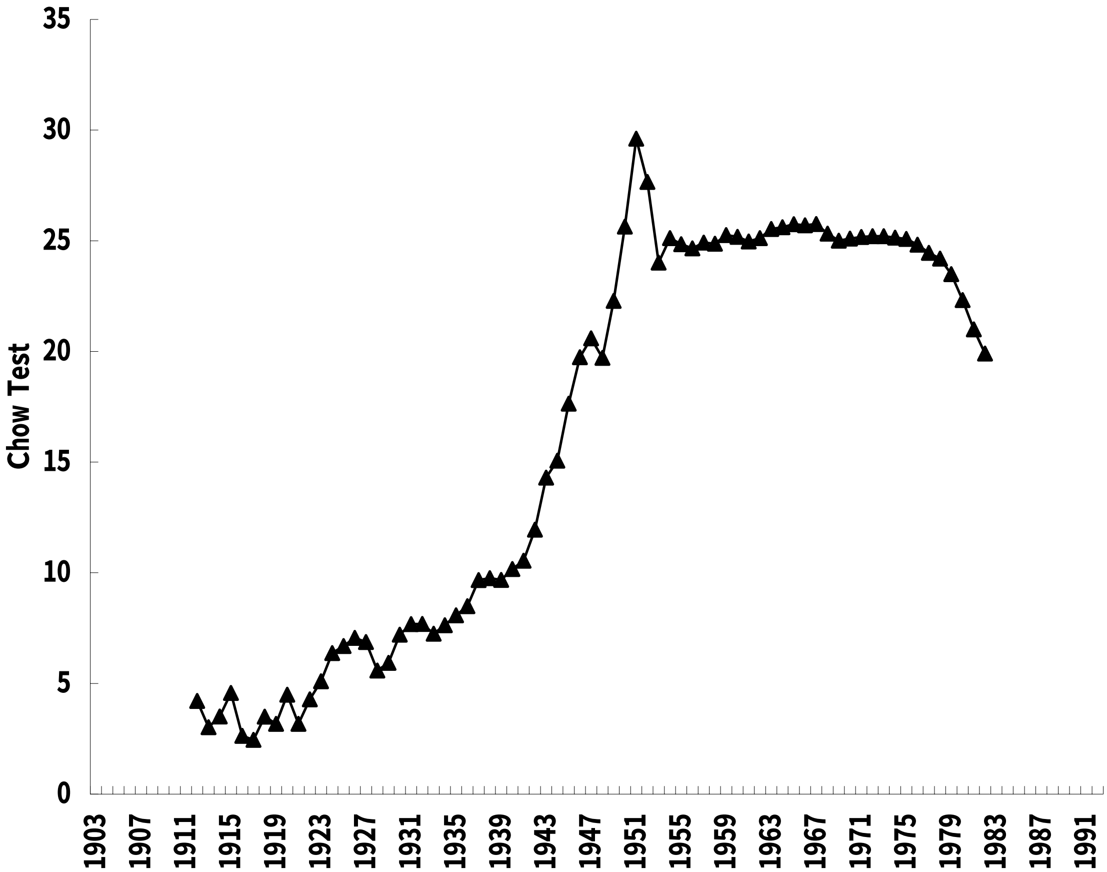

Figure 4. Testing for structural change of unknown timing:
Chow test sequence for Q/A