Measurement of Crop Genetic Diversity in Economic Analysis

E. Meng[1], M. Smale[1] , Hu Ruifa[2], J.P. Brennan[3], D. Godden[4],

1 Economics Program, International Maize and Wheat Improvement Center, Apdo. Postal 6-640, 06600 Mexico, D.F., Mexico
2 Center for Chinese Agricultural Policy, Chinese Academy of Agricultural Sciences, 30 Baishiqiao Road, Beijing 100081, China
3 NSW Agriculture, Wagga Wagga Agricultural Institute, PMB, Wagga Wagga, NSW 2650
4 Department of Agricultural Economics, The University of Sydney, Sydney, NSW 2006

ABSTRACT

In recent years, output on genetic diversity in the economic literature has included conceptual pieces on the definition and measurement of crop genetic diversity, methodologies for estimating its value, and efforts to analyze its contribution to productivity and stability. However, because biological diversity refers in general to a broad area of scientific inquiry, the growing quantity of literature has also generated some confusion over the definition, measurement, and interpretation of genetic diversity in the context of economic analysis. This paper addresses some of the measurement issues encountered in incorporating genetic diversity into economic analysis by presenting a synthesis of several of the relevant concepts and tools. Using data collected in Australia and China, we compare the results from the application of some of the diversity concepts and discuss the implications of using them into future economic analysis.

Key Words: genetic/diversity/wheat/variety/measurement

Contributed paper presented at the 43rd Annual Conference of the Australian Agricultural and Resource Economics Society, Christchurch, NZ, January 1999. We wish to acknowledge the financial support for this work from the Australian Centre for International Agricultural Research. The views expressed in this paper are those of the authors and do not represent those of our organizations.

Discussion on the conservation and utilization of genetic resources continues to proliferate in both scholarly and popular arenas. Since economic-related questions of priority setting, resource allocation, property rights, and value lie at the heart of many of the issues currently being debated, many in the economics profession have also jumped into the fray. In recent years, their output has included conceptual pieces on the definition and measurement of crop genetic diversity, methodologies for estimating its value, and efforts to analyze its contribution to productivity and stability. Because biological diversity refers in general to a broad area of scientific inquiry encompassing all living organisms and their relationship to each other, it is not surprising that the growing quantity of literature has also generated some confusion over the definition, measurement, and interpretation of genetic diversity in the context of economic analysis.

Genetic diversity may be examined both in the broad setting of an integral system of the crop species or in the much more focused case of the variation within a single crop population. These are both valid, albeit different, concepts of genetic diversity, and require slightly different approaches in interpretation and analysis. This paper addresses some of the measurement issues encountered in incorporating genetic diversity into economic analysis by presenting a synthesis of several of the relevant concepts and tools. Our objective is to establish a framework within which these concepts and measures can be compared. Then, focusing specifically on crop genetic diversity and using data collected in Australia and China, we compare the results from the application of some of the diversity concepts. Finally, we discuss the implications of using them in future economic analysis.

Diversity Concepts and Indices in Economic Analysis

Ecologists, biologists, and geneticists have proposed a panoply of diversity estimators (e.g., Hawksworth 1995, Magurran 1991), many of which may be adapted for the study of crop genetic diversity. For applied economists, the major dilemma is how to relate diversity concepts defined on biological and genetic phenomena to the economic decisions of farmers and breeders in a way that is suitable for policy analysis. Farmers do not observe bands of DNA, but farmers' decisions will need ultimately to be addressed by policies designed to enhance crop genetic diversity in their fields.

For any economic analysis that includes crop genetic diversity to be useful, the concept of diversity utilized must be well defined and the measurement technique must be appropriate to the type of analysis and its objectives. For a given economic question, the use of more than one concept of genetic diversity may be feasible or even appropriate. Each concept will describe or classify diversity slightly differently, and none can be deemed correct or incorrect *a priori*. It is important to realize, however, that the appropriateness of the concept chosen is largely a function of the objectives of the study and of the level at which the analysis takes place. For example, the diversity concept for a study analyzing the determinants of on-farm crop diversity at the farm household level should be distinct from one used in a regional study of the relationship between the demand for and supply of modern varieties and crop diversity. Both of these concepts should in turn differ from one employed to study genetic resource utilization in a gene bank.

It is also important to distinguish between a concept of diversity and its measurement tool. The measurement tool is a mathematical construct, usually an index, that enables the appropriate concept of diversity to be expressed as a number or scalar varying over the units of observation and incorporated more conveniently in an analytical model. Many commonly used indices are drawn from ecological and agronomic literature and are based on some type of distance metric.

Classifying the Crop Population

Because it is a neutral mathematical construct, the same index may be used to express different diversity concepts. For example, crop populations can be classified by the names or criteria that farmers use to describe them, by their genealogies as recorded by plant breeders, or by the genetic identity that molecular analysis reveals. The use and interpretation of diversity indices requires caution since distinctions based on one method of classification, or taxonomy, may not be distinguishable when other criteria are used. Relying on named crop populations may overestimate diversity if populations identified by different names are similar or underestimate diversity if those identified by the same name possess important underlying genetic differences. The problem with the uniqueness of names is likely to be exacerbated for traditional varieties, or landraces, as opposed to improved varieties, since the same landrace is often be known by several names across villages or regions. Even with improved varieties, farmers often develop their own nomenclature.

Variation in plant characteristics and other types of descriptors can also serve as the basis of the taxonomy (Franco et al., 1998). Analysis based on specific characteristics and performance of plant populations decreases the likelihood of overlooking some of the differences that may not be picked up when relying on names. Morphological traits are physically observable descriptors often used in the crop science literature to describe plant populations and assess their diversity (Louette, Charrier, and Berthaud, 1997; Meng, 1997). These traits can be measured both quantitatively (e.g., height, spike (wheat) or ear (maize) length, thousand kernel or cob weight) and qualitatively (e.g., kernel or grain color, awn presence). Because observable variation in plant characteristics can result from either genetic differences or differences in the environment, precautions must also be taken to account for interactions between genotype and environment before drawing any conclusions regarding diversity levels. In certain crop populations, the presence of morphological differences may also mask the closeness of the actual genetic relationship (Dudley, 1994).

Applications of Diversity Concepts and Indices

Below, we discuss some of the means of conceptualizing and categorizing diversity, followed by a description of indices used to express the concepts and illustrations. It should be kept in mind that we are comparing several things; namely, measurements for one diversity concept using the same indices but different data sets, measurements for one diversity concept using different indices but the same data sets, and measurements for differing concepts using the same data sets.

Spatial Diversity

"Spatial diversity," probably the most recognized concept of diversity, refers to the amount of diversity found in a given geographical area. Although indices developed in the ecology literature are used primarily for analyzing species diversity, they can be adapted to the study of spatial diversity in crop populations. Magurran (1991) classifies ecological indices of species diversity by three criteria: (1) species richness, or the number of species encountered in a given sampling effort; (2) abundance, or the number of individuals associated with each of the species; and (3) "evenness." A count of species reported or collected in the area, although usually simplest to implement, assumes that all species at a site contribute equally to its biodiversity (Harper and Hawksworth, 1995). Since this is often not the case, frequency counts of individuals within a species provide more information. Indices reflecting abundance detect whether or not certain varieties or groups of varieties dominate others. The third category, "evenness," combines a measure of proportional representation with the number of species. Also called "equitability," it refers to the degree of equality in the abundance of the individuals, or the relative uniformity of their distribution across species. When all species in a sample are equally abundant, evenness reaches a maximum (Ludwig and Reynolds, 1993).

As measures of proportional abundance, indices such as the Shannon[1] or Simpson have been used to express both the concepts of evenness and richness (Magurran 1991). The Shannon index was originally used in information theory, but has been commonly applied to evaluate species diversity in ecological communities. It has also been widely used in the agronomic literature to transform qualitative traits into a scalar measure which can be compared over sets of varieties (Spagnoletti Zeuli and Qualset, 1987; Jain et al., 1975). The Simpson index is simply equal to one minus the Herfindahl index.

Meng (1997) has applied a Shannon index constructed with morphological data measured on seed samples collected from farmers in the economic analysis of diversity among Turkish wheat landraces. Economists have also used indicators of spatial diversity that are related to ecological indices. For example, Widawsky and Rozelle (1998) used the number of varieties accounting for a given percentage of cultivated area. The Herfindahl index, borrowed from the industrial economics literature on market shares and used by Pardey et al. (1996) and Hartell (1996), is defined as the sum of squared shares of total crop area planted to each unique variety, or one minus the Simpson index of area shares. Smale, Bellon, and Aguirre (1998) used a Simpson index constructed from area shares planted to races of maize.

Here, we have adapted and applied several of the indices used to represent spatial

---

[1] This index is also known as the Shannon-Wiener index.

diversity to data on wheat populations in China and Australia. Table 1 lists each index used by its name, category among spatial indices, and mathematical construction, with an accompanying explanation.

Table 1. Definition of spatial diversity indices used in this paper

Index	Category	Mathematical Construction	Explanation	Adaptation in this paper
Margalef	Richness	$D_{mg}=(S-1)/\ln N$	number of classes weighted by the logarithm of the total number of samples	number of crop populations per mill has
Berger-Parker	Abundance (Dominance)	$D=1/(N_{max}/N)$	the less dominant the most abundant class, the higher the value of the index	inverse of maximum area share occupied by any single crop population
Shannon	Richness and Evenness	$H'=-\Sigma p_i \ln p_i$	p_i is proportion, or abundance, of a class	p_i is area share occupied by ith crop population
J	Evenness	$J=H'/\ln S$	Shannon corrected by the logarithm of the number of classes	S is the total number of crop populations
Simpson	Richness and Evenness	$D=1-\Sigma p_i^2$	also represented in the form of $D=1/\Sigma p_i^2$	

Source: adapted from Aguirre, Bellon, and Smale, 1998. Mathematical construction as defined by Magurran (1991).

Richness, dominance, and evenness indices for seven major wheat-producing provinces of China are presented in Figures 1-6. By using the same two data sets for the calculation of the indices, we illustrate the potential for divergence in the results obtained for examining the concept of spatial diversity. The indices were constructed using two taxonomies, or means of distinguishing crop populations: (1) named varieties and (2) morphology-based groups of named varieties. The morphological groups were formed by combining maximum likelihood estimation with a clustering method to predict group membership statistically based on plant characteristics obtained from experimental trials (Franco et al., 1998). Specifically, the clustering is based on pairwise Gower distances among varieties measured on habit, resistance to stem rust, duration, height and kernel weight at time of release.

The implications of using different taxonomies to define crop populations are revealed when indices with identical construction, but wheat populations classified by cultivar name and morphology are compared. The relative rank of provinces by richness in both Figures 1 and 2 is not affected by the taxonomy employed. Falling in roughly descending order over the 1982-1997 period are Shanxi, Anhui, Hebei, Jiangsu, Shandong, Henan, and Sichuan. However, with the exception of Shanxi, divergence among all the provinces is reduced when varieties are grouped by similarities in morphological traits rather than by name.[2] Shanxi emerges as clearly superior in morphologically-based richness while Anhui and Hebei are less distinct. The range in index values (0.1 to 0.9 vs. 5 to 35) differs between the two definitions because there are several varieties per group. In general, however, although year-to-year changes visible for several of the provinces from the variety-based richness index appear to be larger than represented by the morphologically-based richness index, the overall pattern is similar for both during the time period.

Indices of cultivar dominance show cyclical patterns that reflect differences in the relative popularity of new varieties as they are released and adopted by farmers (Figure 3). When a number of varieties are available but no variety is clearly superior to others for the traits demanded by farmers, the index of dominance is likely to be low. Inadequate seed supplies relative to demand may also constrain the area planted to popular varieties. In general, the diffusion curves that underlay the dominance index illustrate the emergence and disappearance of popular varieties. Varieties may disappear either because they are replaced by varieties that are more attractive to farmers or because their seed sources diminish. The morphology-based index, as compared to the cultivar-based index, highlights the persistent dominance of a single group in Sichuan over the entire study period (Figure 4). The two dominance indices calculated exhibit a large amount of year to year variation; however, if we examine the relative order at two randomly-selected points in time, 1984 and 1995, we do find some similarities. In 1984, Hebei province is the least dominated, both in terms of individual varieties and morphological groups, while Henan and Sichuan provinces show the highest amount of domination. In 1995, we find that the least dominated provinces, regardless of crop population definition, are Hebei, Anhui, and Shanxi, although the order of the remaining provinces is not constant.

---

[2] The majority of wheat varieties included in our research are improved varieties.

While the data sets and taxonomies employed are the same for both richness and dominance indices, we see that the relatively high diversity levels attributed to Shanxi province based on the richness indicators are not as evident when using the dominance index. The diversity indices constructed from data on named wheat varieties, however, consistently place the provinces of Shanxi, Anhui, and Hebei at a higher level of diversity than the remaining provinces. Moreover, despite an increase in Sichuan's diversity levels in the early 1990's according to the dominance index, which is not captured by the richness index, both indices concur that Sichuan province is the least diversity of all the provinces included in the research. This finding also holds strongly when comparing the richness and dominance indices based on morphological groups, although these indices do not agree with respect to the relative order of the other provinces.

The contrast between Sichuan and the other provinces recurs in the evenness indices (Figures 5 and 6). Sichuan appears to be the least "rich" according to both ways of classifying crop populations, and the dominance of a single morphological group in that province is very pronounced. As illustrated in the evenness indices, the spatial distribution of wheat cultivars and morphology in Sichuan is relatively "poor" and "inequitable." While the evenness indices based on named cultivars for all provinces fall in a range between 0.5 and 0.9, the evenness index based on morphology separates Sichuan into the range below 0.5. Henan also appears to have a lower level of morphological evenness.

A comparison across the richness, dominance, and evenness indices calculated using the data set of named varieties shows, in general, Hebei and Shanxi to be the most diverse throughout the study period, while Henan and Sichuan provinces almost always rank among the least diverse. However, the relative order of the provinces changes considerably depending on the year as well as the category of spatial diversity being represented. A similar situation holds when using data on morphological groups. Henan and Sichuan provinces stand out as the least diverse regardless of the category of spatial diversity examined, but there is much less consistency in the conclusions regarding the remaining five provinces.

Turning to the situation in Australia, Figure 7 shows richness, evenness and dominance indices, as well as the Shannon and Simpson indices combining richness and evenness, for Australian wheat cultivars during the period from 1962 to 1993. The indices are represented on the same graph by relating each year's observation, for each index, to a base year (arbitrarily chosen as 1962). From 1965 to 1970, all indices of diversity except dominance dip in magnitude. Although there are fewer cultivars grown and less evenness among them, the dominant cultivar covers relatively less area than it does in the preceding period or the period that follows. According to both dominance and richness indices, diversity seems to improve from about 1980. Over the entire period the variation apparent in the evenness, Simpson and Shannon indices is far less pronounced than in the dominance and richness indices. This is not entirely unexpected, however, since evenness measures incorporate elements of both dominance and richness.

Temporal Diversity

The concept of temporal diversity is used to refer to the rate of change in the turnover of varieties.[3] Duvick (1984) has described it as "genetic diversity in time." The replacement of varieties is important for several reasons. Regular varietal turnover provides a possible response to the plateaus or reductions observed over time in yields by capitalizing on advances in knowledge and technology. Varietal turnover also reduces the potential exposure to disease epidemics resulting from pathogen mutations that overcome the genetic resistance in older varieties. This turnover is important for modern agriculture and in some ways substitutes for the spatial diversity that has been characteristic of traditional varieties (Apple, 1977; Plucknett and Smith, 1986). Because of the many elements of uncertainty involved, the breakdown of a variety's resistance to a pathogen is difficult to predict. Similarly, the future development and detrimental consequences of new pathogens are often difficult to foresee. The economically optimal rate of varietal turnover in a given area is jointly determined by a number of factors, including the rate of mutation of disease organisms, the structure of genetic resistance to disease in a variety, and the production environment (Heisey and Brennan, 1991).

Indices of temporal diversity, such as the average and weighted (by area) average age of varieties grown by farmers have been proposed, used, and reviewed by Brennan

---

[3] Temporal diversity, or the rate of change in varietal turnover, should not be confused with changes over time in other methods of describing diversity. For example, the number of available varieties would still be considered as a measure of spatial diversity.

and Byerlee (1991) and Brennan and Fox (1998). When variety ages are weighted by the percent of area they occupy in a given region, the index also expresses changes in their diffusion pattern across space.

The key data required to calculate an index of temporal diversity are the names of the varieties available within a given geographical area and their release dates, as well as information over time on the extent of their cultivated area. If landraces continue to maintain a presence in the geographical area, the absence of any official release dates for them must be considered in the calculation of an index reflecting temporal diversity. The reliance on varietal names for this index raises again the potential problems described earlier. Instead of named varieties, information on the development and incorporation of certain characteristics, such as genes to provide resistance to a specific disease or pest, could be used. Compiling this type of data requires considerably more information and expertise.

Apparent and Latent Diversity

By distinguishing between apparent and latent diversity, we acknowledge that genetic diversity is not necessarily observable. Apparent diversity refers to differences that are easily observed by farmers or scientists in the field. Latent diversity has been defined by Souza et al. (1994) as that diversity which is not observable until challenged by a new pathogen or a change in the growing environment.

Latent diversity for crop genetic resources has most frequently been assessed on the basis of two types of information: (1) genealogical characteristics for the relevant set of named varieties, and more recently, (2) genotype frequencies within and among the crop populations in a defined study site. The first was largely used as a proxy for the second with the assumption being that specific genes or gene combinations would be bestowed by parents on their offspring. However, advances in biotechnology have made possible new and more sophisticated molecular methods for detecting genetic variation at the DNA level. The use of biochemical or molecular markers can determine whether two populations are similar or not by offering a genetic fingerprint of each one. Each of these molecular methods involves an examination of the relationship between crop populations or individual materials from band patterns that reveal variation in DNA sequences

(Dudley, 1994). Since the information obtained by molecular methods provides a way of characterizing plant populations that is similar to the use of physical traits, it can also be used to construct various diversity indices.

A pedigree-based analysis of latent genetic diversity requires information as detailed and complete as possible on the ancestry of each variety. Since recorded genealogies exist only for varieties released by crop breeding programs, this type of analysis is applicable only to modern varieties. One disadvantage of the use of pedigree-based methods is the preclusion of landraces from the analysis.

Because varietal pedigrees are often quite complicated, many different methods of representing the information contained in the pedigree have been employed. Some of them focus on "pedigree complexity; that is, specific details from the varietal pedigrees, such as the numbers and origin of landraces in ancestry or numbers of breeding generations since the first cross (Gollin and Evenson, 1990). Numbers of distinct parental combinations and numbers of unique landrace ancestors per pedigree were used in Hartell (1996).

Another method of transforming pedigree information into a usable form for analysis is through the calculation of coefficients of diversity. The coefficient of diversity is equal to one minus the coefficient of parentage between any pair of varieties. The coefficient of parentage (COP) estimates the probability that a random allele taken from a random locus in a variety X is identical, by descent, to a random allele taken from the same locus in variety Y (Malecot, 1948). Values range from 0 to 1, with higher values indicating greater relatedness (for historical development and other examples of this concept, see Wright, 1922; Malecot, 1948; Kempthorne, 1969; and Cox et al. 1985).

Several disadvantages have been noted with regard to the use of the coefficient of parentage to construct indicators of genetic diversity. First, two alleles may be identical "by state," whether or not they share the same parentage. Second, coefficients of diversity derived from pedigrees are necessarily limited by the pedigree as it is recorded; the ancestors positioned at the outermost leaves of the genealogical tree are typically understood as unrelated. The outermost leaves are as recorded by scientists in the first plant breeding program in which they were used. The assumption that they are unrelated may be incorrect from a biological or ecological perspective and probably biases the COP downwards (Cox et al., 1985). Molecular methods now provide a means of testing the independence of ancestors.

Also potentially troublesome are the algorithms employed to calculate the coefficient of parentage. They impose the assumption that each parent contributes equally to offspring, despite the effects of recurrent selection and random genetic drift. This assumption is unlikely to be valid, especially for qualitative traits such as disease resistance and height. Although the COP between a cultivar and a selection from it must clearly be less than one, it value is usually assigned somewhat arbitrarily by the researcher. Finally, no information on actual genetic expression is conveyed directly through a COP-based index; instead, we depend on the complex relationship between the variety's pedigree and its expression, a relationship about which we are not necessarily very informed.

When weighted by the percent of area planted to the variety, the average coefficient of diversity expresses spatial diversity as well as the latent diversity conferred through ancestry. Figure 8 shows average and area-weighted coefficients of diversity (defined as 1 - the coefficient of parentage) for all named wheat varieties grown by farmers in Australia from 1962 to 1993. A point of reference for interpreting the magnitude of the coefficient of diversity is the value of 0.4375 assigned to varieties bred from the same parents, or sister lines. In all years, both the average and weighted-average coefficients of diversity are higher than this value. From the mid-1960's through the early 1970's, the area-weighted coefficient of diversity falls sharply. Although the average coefficient of diversity fluctuates around the same level over the time period of the study, the area-weighted coefficient of diversity shows a gradual increase in diversity beginning in the early 1970's. This finding suggests that while breeders have maintained a relatively constant level of diversity among the parents of the materials that are acceptable to farmers, the spatial distribution of these materials in farmers' fields has become increasingly uniform.

Figure 8 also illustrates an essential feature of the relationship between the average and area-weighted coefficients of diversity: the area-weighted coefficient of diversity must always be lower or, at best, equal to the average coefficient of diversity. Only in the case where the percentage distribution of wheat varieties by area shares is perfectly uniform does the area-weighted coefficient of diversity equal its average coefficient of diversity. In 1991, the two values approach each other closely, but then separate again in following years. Since then, the degree of divergence has remained much smaller than in previous decades.

Distance Indices

The taxonomic tree provides another means of measuring diversity among biological species. Weitzman (1992) proposed a distance measure to identify species for conservation that maximizes diversity among the surviving members of the set. Such a conservation goal would reflect the notion that the greater the distance among the members of the set, the less likely are the species to contain redundant characteristics. Solow et al. (1993) also discuss a measure which minimizes the distance between the surviving and extinct species in order to preserve the most representative sample of current species as possible. Diversity indices based on distance metrics and proposed by Weitzman and Solow and Polasky (1994) can be calculated using data that transforms differences in plant characteristics, genealogies, or specific alleles into distances. Distance metrics can thus be utilized to reflect concepts of apparent and latent diversity. By using area weights, they can also provide information on spatial diversity. Here, we use the pairwise measures of dissimilarity provided by coefficients of diversity for Australian varieties to calculate the index proposed by Solow and Polasky (1994). Following Solow and Polasky, we first transform the matrix of pairwise distances using an exponential function, $f(d) = \exp(-\Theta,d)$, $\Theta > 0$. This transformation ensures that certain desirable properties for a diversity measure hold (Weitzman, 1992). We use two parameter values for Θ, $\Theta=0.1$ and $\Theta=0.5$; a larger value of Θ places more weight on the number of species and less weight on their dissimilarity. Because the diversity index proposed by Weitzman (1992) does not give any weight to the number of species, the parameter value of 0.1 in the Solow-Polasky measure approximates the Weitzman index (Solow and Polasky, 1994) which we have not explicitly calculated here. Solow-Polasky diversity indices for both parameter values are shown in Figure 9. Both of these indices exhibit the drop in diversity levels in the late 1960's that most of the spatial diversity indices (Figure 7) as well as the area weighted COP measure (Figure 8) also showed.

However, the index with parameter value of 0.1 that attributes little weight to the number of crop populations shows a sharp increase in diversity beginning in1980. Note that the Solow-Polasky index as calculated here has not been weighted by variety area shares. Each point in Figure 9 therefore represents the genealogical distance among the members of the set of wheat cultivars grown by farmers in each year, without accounting for their relative abundance[4].

A Concept of Breeder-based Diversity

To examine the impact of diversity on productivity, we propose a new distance-based index that incorporates more directly the role of crop breeders. The index is area-weighted and is based on apparent morphological and performance characteristics.

There are two assumptions underlying the development of the index: (1) a primary goal of crop breeders is to increase the mean yield potential of new varieties through a continual assembly and use of genes and (2) improvement results primarily from a focus of their efforts on characteristics that are linked to specific genes or gene combinations. The characteristics contributing to improved yield potential in the set of released varieties are singled out here for particular attention. Figure 10 provides a graphical representation of the assumption underlying the index. If breeders are doing their job successfully, we would expect that each successive group of released varieties increases in yield potential. The proposed diversity index, the coefficient of focused genotype (CFG), reflects the amount of diversity arising from the variation observed in the yield-related characteristics targeted by breeders. Possible choices for breeder-targeted characteristics include physical traits such as height that are linked to yield performance as well as data scoring levels of disease and pest resistance. Figures 11 and 12 use experimental data on yield potential and height for a grouped set of 403 Chinese wheat varieties cultivated between 1982 and 1997. As expected, we see a gradual increase in mean group yield potential and a gradual decrease in mean group height.

Because the breeders' development of the variety terminates upon its release (no further manipulation of the variety's genetic composition by the breeder occurs) and the

---

[4] Modifying the index to incorporate spatial diversity concepts are probably feasible. See Widawsky and Rozelle (1998) for one example.

basic genetic structure determining the expression of the targeted characteristics remains constant, the index reflecting this focused effort will be unique by definition for each variety and does not change over time or space. The CFG score is nevertheless an outcome that is dependent on which other varieties appear in the entire set. Because the characteristics targeted by the breeder are visible, the CFG index also represents the concept of apparent diversity. Finally, by weighting the index for each variety by its area cultivated, the concept of spatial diversity is reflected in the index as well.

The calculation of the index uses a matrix of pairwise distances derived from morphological and performance-based traits that are specifically chosen to reflect the breeders' focused effort on genotype development. Since important agronomic traits are often categorical, we use Gower distances that are able to handle both continuous and discrete data. We begin by using the same clustering and maximum likelihood methods described above to group the Chinese provincial data for richness, abundance, and evenness measures (Franco et al., 1998). The final step of the methodology developed by Franco et al. consists of a canonical discriminant analysis (CDA) that maximizes the variability between groups while minimizing the variation within each group. A "score" for each variety is calculated from the canonical equation explaining the majority of the difference among groups; we interpret this score loosely as representing the contribution of the variety through its characteristics to the amount of between-group variation. Due to the nature of the index's construction, the individual scores for each variety must be interpreted relative to the group mean.

The CFG was calculated for the complete set of varieties cultivated in the seven major wheat-producing provinces of Anhui, Hebei, Henan, Jiangsu, Shandong, Shanxi, and Sichuan over the time period from 1982 to 1997. There were a total of 403 varieties cultivated in the relevant time period.[5] The ten traits originally chosen for the analysis were reduced to five variables after significant correlation was observed among several of the variables and after significant missing data problems could not be resolved. The variables used in the final analysis include three continuous variables, duration, thousand

---

[5] The Chinese data available are based on released varieties with a total cultivated area greater than 100,000 mu (**6,667** hectares). The data are thus a subset of all released varieties, although in most of these provinces, our data represent a large percentage of the cultivated area.

kernel weight, and height, and two discrete variables, stem rust resistance and growth habit.[6] All of these are key traits emphasized by wheat breeders in yield improvement.

Figure 13 illustrates the area-weighted values for the CFG in all seven provinces. With the exception of Shanxi, Hebei, and Anhui to a lesser extent, variation in CFG values over the time period do not fluctuate much. We noted from Figure 11 that overall progress in improvement of yield potential has been positive during this time period. Here, perhaps we see moderating effects on scientific progress coming from the area distribution of the varieties. The ranking of the provinces is also noteworthy. Once again, Shanxi and Hebei provinces exhibit relatively higher levels of diversity while Sichuan province once again falls in at the bottom.

Implications for Economic Analysis

Having laid out different concepts of diversity and methods to represent them, it is evident that a judicious choice of measurement tools is essential to applied economists conducting research on crop genetic diversity. Since by its definition, the diversity based on pedigrees or differences at the molecular level is not observable to farmers, it may not be appropriate to model latent diversity as an explicit choice variable in models of decision-making at the farm level or even in aggregate analysis of regional crop productivity and stability. Farmers select crop populations to cultivate based on characteristics or qualities that are observable rather than on genetic structure that is not. There is seldom a direct relationship between the presence of an individual gene and a specific, physical characteristic. Most economically important traits are determined by multiple genes, and the relationship is usually quite complicated and not yet scientifically understood. Consequently, the linkage between the economic decisions of farmers and genetic diversity measured at the molecular level is not straightforward to establish conceptually or empirically. The relationship between a specific ancestor and an observable physical characteristic is also more often than not unclear so that attempts to include pedigree-based measures of genetic diversity in production analyses have largely been inconclusive. With the exception of traits such as semi-dwarf height or certain

---

[6] Data were obtained from field trials conducted over time in different locations in China and published at the time of varietal release. Hence, observations on characteristics do not change over the time period of the study.

disease resistances that can be traced to specific parents, most characteristics are the result of the combination of genetic material contributed by multiple parents. However, this should not diminish the usefulness of indices of latent diversity for describing and comparing crop diversity levels in the field or breeding program.

Results from the calculations of all these diversity concepts are interesting to compare, but in order to lend an economic interpretation to the results, an understanding of the environment in which the changes took place is crucial. What factors caused the big swings and gradual changes? They must be the result of a combination of factors, including policy decisions; reactions to policy decisions on the part of farmers and scientists; developments in research; and input and output prices. By exploring more deeply past influences on diversity levels, we will develop a better understanding of what is likely to affect them in the future. Also, by clarifying the definitions of diversity and the ways to measure them, we can more effectively investigate the relationship of diversity to risk and changes in the level of productivity.

Diversity levels are not determined in isolation; instead variety area shares represent the simultaneous solution of supply and demand for different varieties. Using aggregate data on factors affecting variety supply and influencing farmer demand for varieties, we are exploring this simultaneous relationship in the context of spatial diversity in China and Australia (Brennan, Godden, Smale, and Meng, 1999). Ongoing work also continues attempts to improve upon the relationship between diversity levels and production-related issues at both aggregate and household levels. Previous studies have used a primal framework (Gollin and Evenson, 1998; Smale, Hartell, Heisey and Senauer, 1998); however, the inability to consider prices explicitly and the endogeneity of input choices are both drawbacks to this approach. To address these shortcomings and develop a fuller economic decision-making framework, we are estimating the cost function dual to the stochastic production function in order to shed more light on the "marginal cost" of diversity.

References

Aguirre, J.A. Gómez, M.R. Bellon, and M. Smale. 1998. *Regional Analysis of Maize Biological Diversity in the Southeast of Guanajuato, Mexico*. Draft Economics Program Working Paper. Mexico, D.F: International Maize and Wheat Improvement Center (CIMMYT).

Apple, J.L. 1977. The theory of disease management. In J.G. Horsfall (ed.), *Plant Disease: An Advanced Treatise.* New York: Academic Press.

Brennan, J.P. and Byerlee, D. 1991. The rate of crop varietal replacement on farms: Measures and empirical results for wheat', *Plant Varieties and Seeds* 4, 99-106.

Brennan, J.P. and Fox, P.N. 1998. Impact of CIMMYT varieties on the genetic diversity of wheat in Australia, 1973-1993", *Australian Journal of Agricultural Research* 49(2), 175-8.

Brennan, J.P., D. Godden, M. Smale, and E. Meng. 1999. Variety choice by Australian Wheat Growers and Implications for Genetic Diversity. Contributed paper, 43rd Annual Conference of the Australian Agricultural and Resource Economics Society, Christchurch, New Zealand, January.

Cox, T.S. Y.T. Kiang, M.B. Gorman, and D.M. Rodgers. 1985. Relationship between coefficient of parentage and genetic similarity indices in the soybean. *Crop Science* 25: (May-June): 529-532.

Dudley, J. 1994. Comparison of genetic distance estimators using molecular marker data. In American Society for Horticultural Science (ASHS) and Crop Science Society of America (CSSA) (eds.), *Analysis of Molecular Marker Data, Joint Plant Breeding Symposia Series.* Corvallis, ASHS, CSSA.

Duvick, D.N. 1994. Genetic diversity in major farm crops on the farm and in reserve. *Economic Botany* 38(2), 161-178.

Franco, J., J. Crossa, J. Villaseñor, S. Taba, and S.A. Eberhart. 1998. Classifying genetic resources by categorical and continuous variables. *Crop Science* 38: (in press).

Gollin, D. and R. Evenson. 1998. An application of hedonic pricing methods to value rice genetic resources in India. In R.E. Evenson, D. Gollin, and V. Santaniello (eds.), *Agricultural Values of Plant Genetic Resources.* Wallingford: CAB International.

Harper, J.L. and D. L. Hawksworth. 1995. Preface. In D. L. Hawksworth (ed.), *Biodiversity Measurement and Estimation.* The Royal Society and Chapman and Hall, London.

Hartell, J. 1996. The contribution of genetic resource diversity: the case of wheat productivity in the Punjab of Pakistan. M.S. thesis. University of Minnesota, St. Paul, Minnesota.

Hawksworth, D.L. (ed.). 1995. *Biodiversity Measurement and Estimation.* The Royal Society and Chapman and Hall, London.

Heisey, P.W., and J.P. Brennan. 1991. An analytical model of farmers' demand for replacement seed. *American Journal of Agricultural Economics* 73: 1044-1052.

Jain, S.K. et al., 1975, Geographical patterns of phenotypic diversity in a world collection of durum wheat, Crop Science 15, 700-704.

Kempthorne, O. 1969. *An Introduction to Genetic Statistics.* Ames: Iowa State University.

Louette, D., A. Charrier, and J. Berthaud. 1997. *In situ* conservation of maize in Mexico: Genetic diversity and maize seed management in a traditional community. *Economic Botany.* 51(1): 20-38.

Magurran, A., 1988, Ecological diversity and its measurement (Princeton: Princeton University Press).

Malecot, G. 1948. *Les mathematiques de l'Heredite.* Paris: Masson.

Meng, E. C. H., 1997. Land allocation decisions and *in situ* conservation of crop genetic resouces: the case of wheat landraces in Turkey. Ph. D. thesis. University of California, Davis, California.

Pardey, P.G., J. M. Alston, J.E. Christian, and S. Fan. 1996. *Hidden Harvest: US Benefits from International Research Aid.* Washington, D.C.: International Food Policy Research Institute (IFPRI).

Plucknett, D.L., and N.J.H. Smith. 1986. Sustaining agricultural yields. *BioScience* 36: 40-45.

Smale, M., J. Hartell, P. W. Heisey and B. Senauer. 1998. The contribution of genetic resources and diversity to wheat production in the Punjab of Pakistan. *American Journal of Agricultural Economics* 80 (August): 482-493.

Smale, M., M. Bellon, and A. Aguirre. 1998. *Variety Characteristics and Land Allocation in a Center of Maize Diversity.* Paper presented at the Annual Meetings of the American Agricultural Economics Association, Salt Lake City, Utah, August 1-3.

Spagnoletti Zeuli, P. L., and C.O. Qualset. 1987. Geographical diversity for quantitative spike characters in a world collection of durum wheat. *Crop Science* 27: 235-241.

Solow, A. and S. Polasky. 1994. *Measuring Biological Diversity*. Working Paper, Woods Hole Oceanographic Institution. Woods Hole: Woods Hole Oceanographic Institution.

Souza, E., P. Fox, D. Byerlee, and B. Skovmand. 1994. Spring wheat diversity in irrigated areas of two developing countries. *Crop Science* 34: 774-783.

Weitzman, M.L. On Diversity. *Quarterly Journal of Economics* 107, 363-405.

Widawsky, D. and S. Rozelle. 1998. Varietal diversity and yield variability in Chinese rice production. In M. Smale (ed.), *Farmers, Gene Banks and Crop Breeding: Economic Analyses of Diversity in Wheat, Maize, and Rice.* Norwell, MA: Kluwer Academic Publishers, Inc.

Wright, S. 1922. Coefficients of inbreeding and relationship. *American Naturalist.* 56: 330-338.

Figure 1. Wheat cultivar richness in China, by province, 1982 to 1997

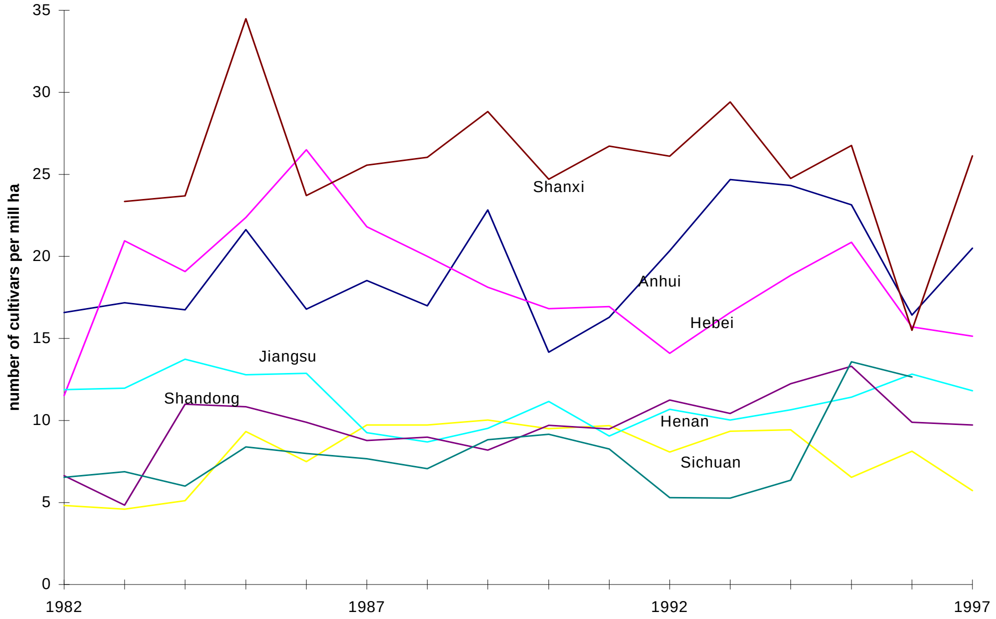

Figure 2. Richness in wheat morphological groups, by province, China, 1982 to 1997

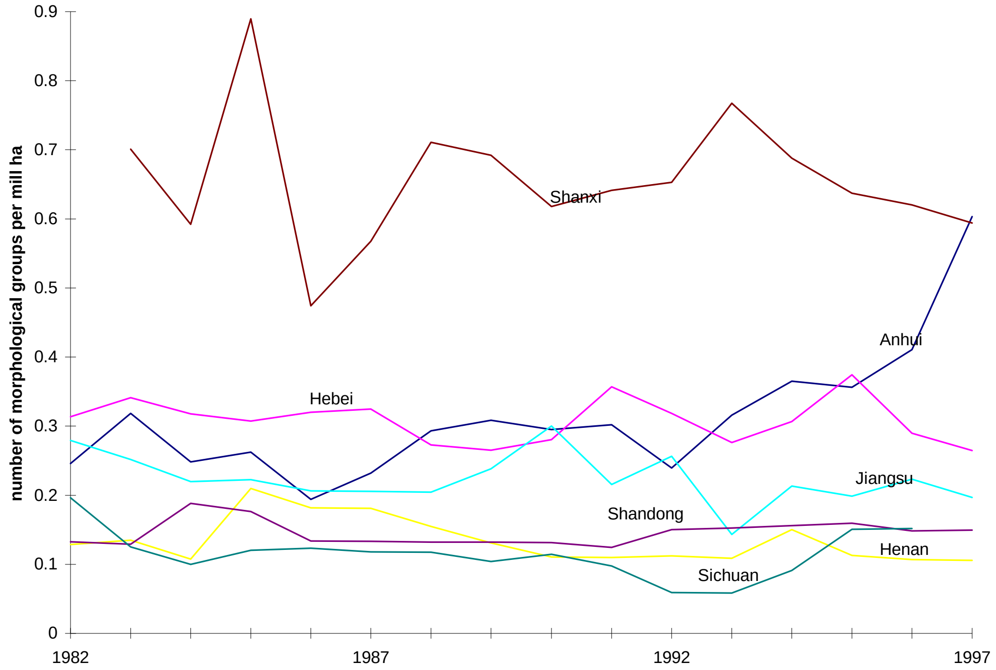

Figure 3. Wheat cultivar dominance in China, by province, 1982 to 1997

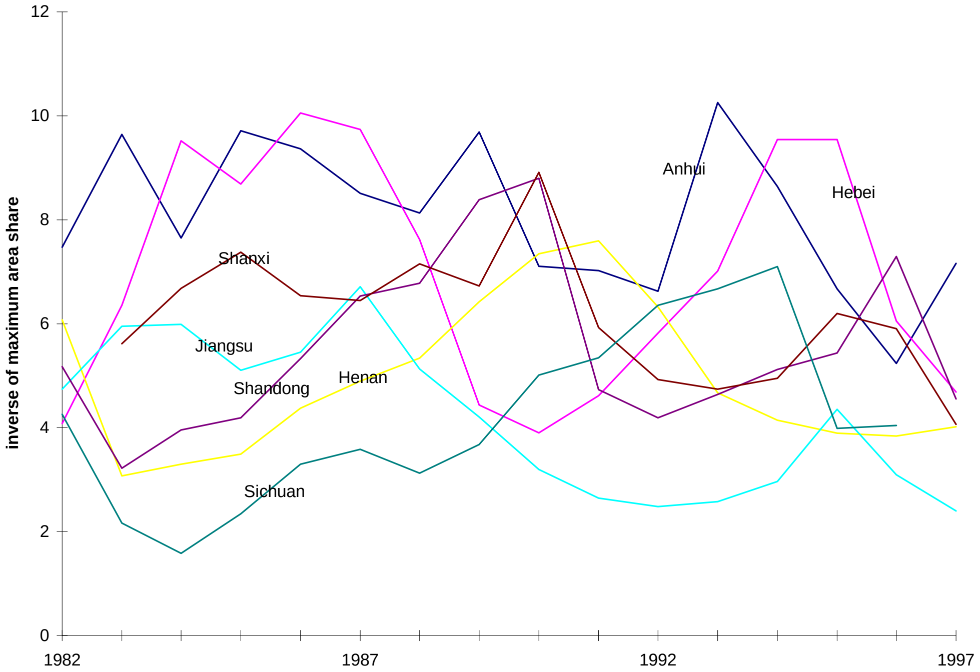

Figure 4. Dominance in wheat morphological groups, China, by province, 1982 to 1997

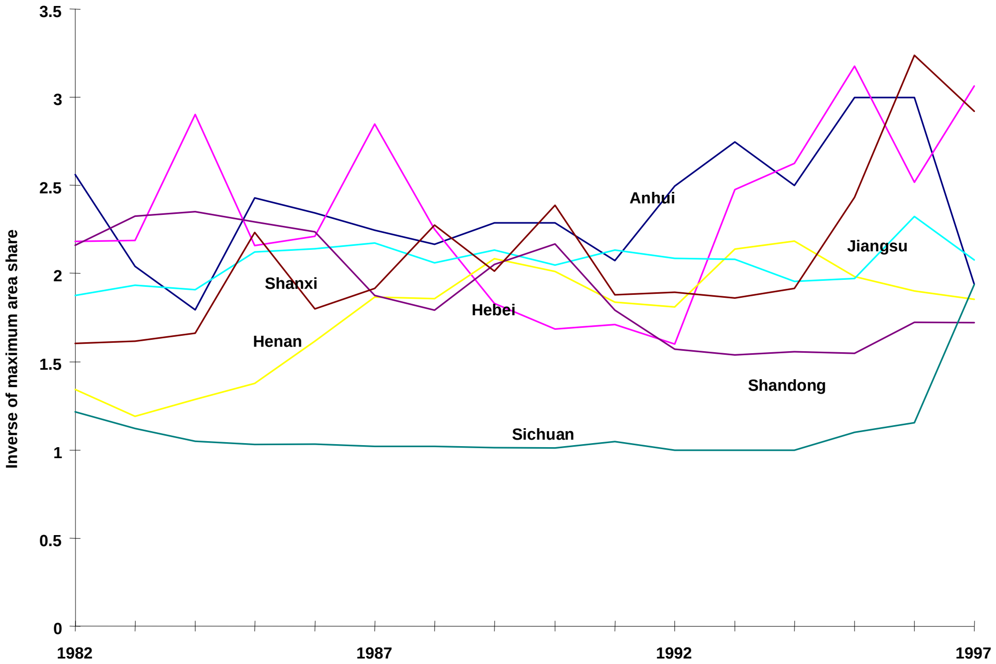

Figure 5. Wheat cultivar evenness in China, by province, 1982 to 1997

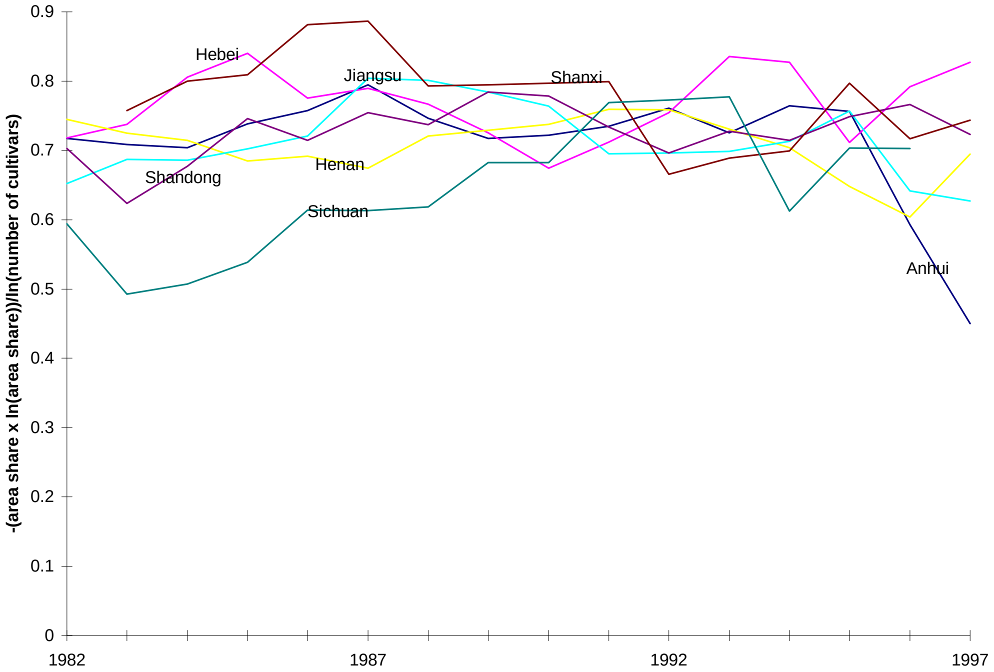

Figure 6. Evenness in wheat morphological groups, China, by province, 1982 to 1997

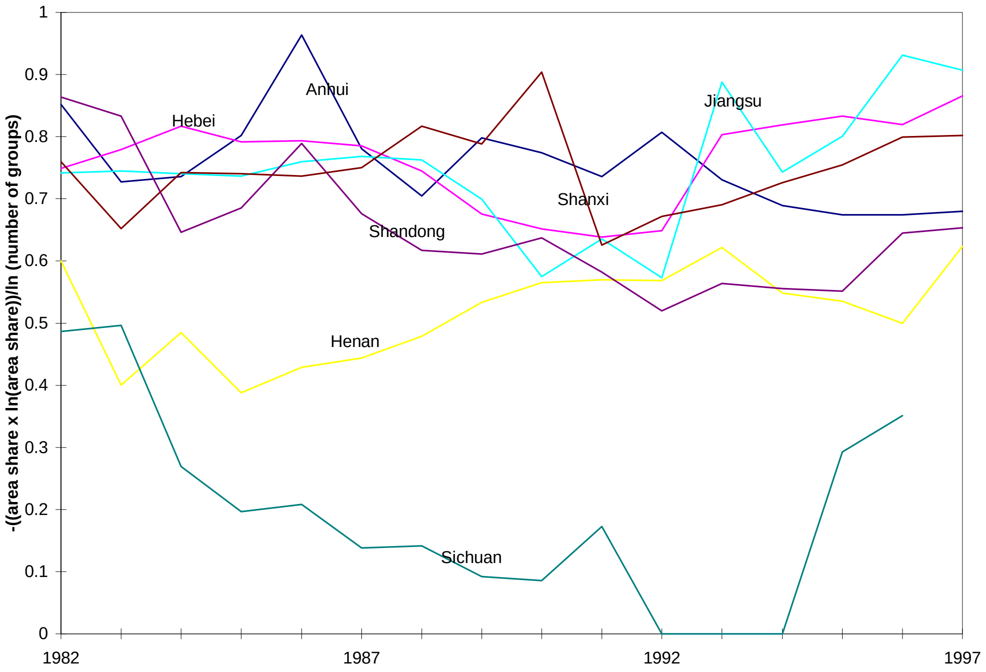

Figure 7. Richness, evenness, and dominance indices for wheats grown in Australia from 1969 to 1993

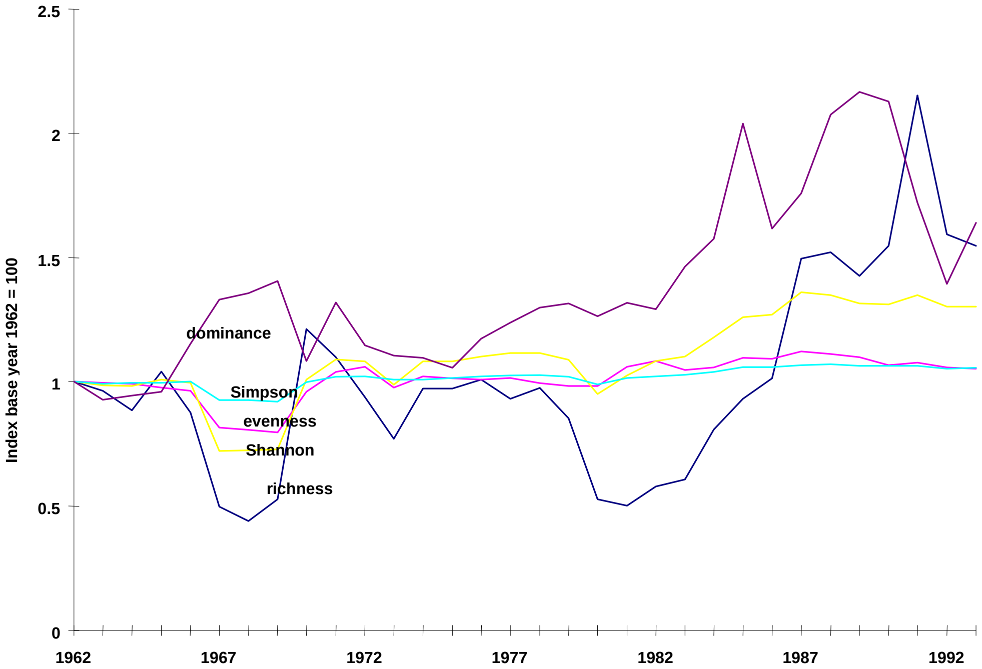

Figure 8. Unweighted and Area-Weighted of Coefficients Wheats Grown in Australia, 1962-1993

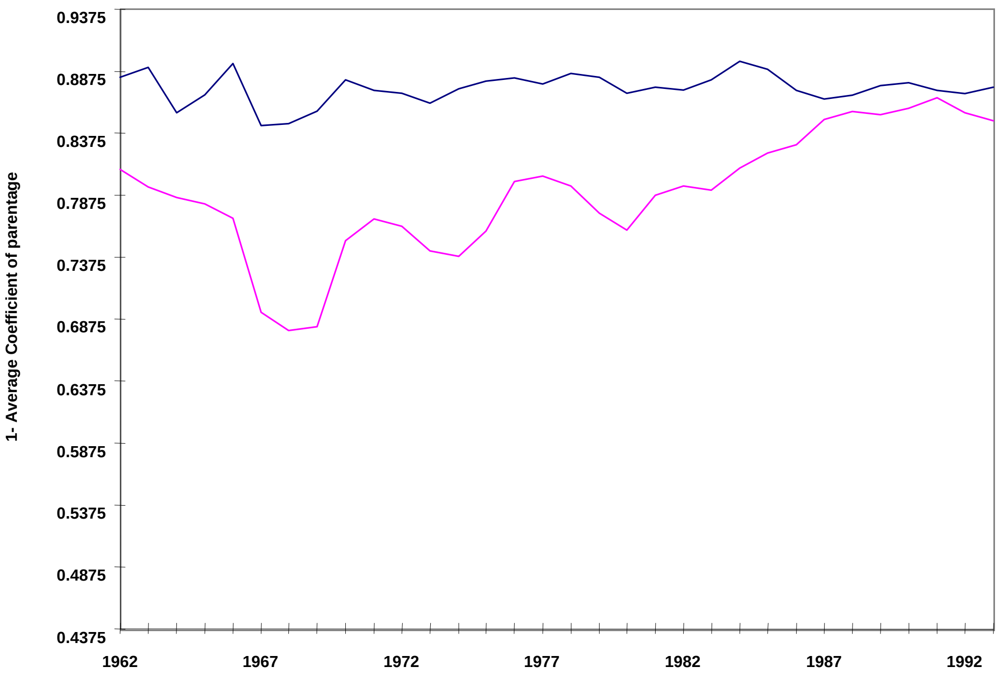

Figure 9. Solow Polasky Diversity Index, Australian Wheat, 1962-1993

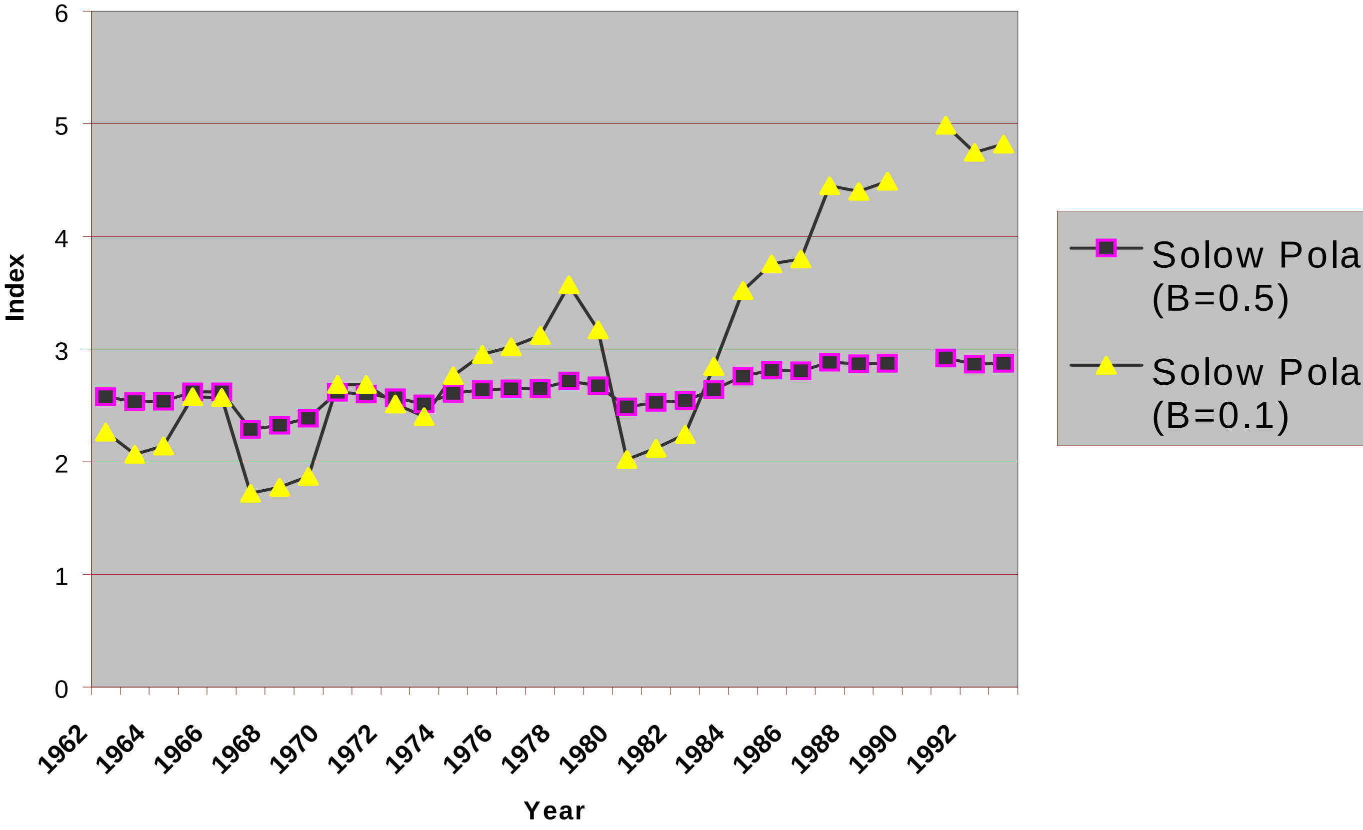

Figure 10.-

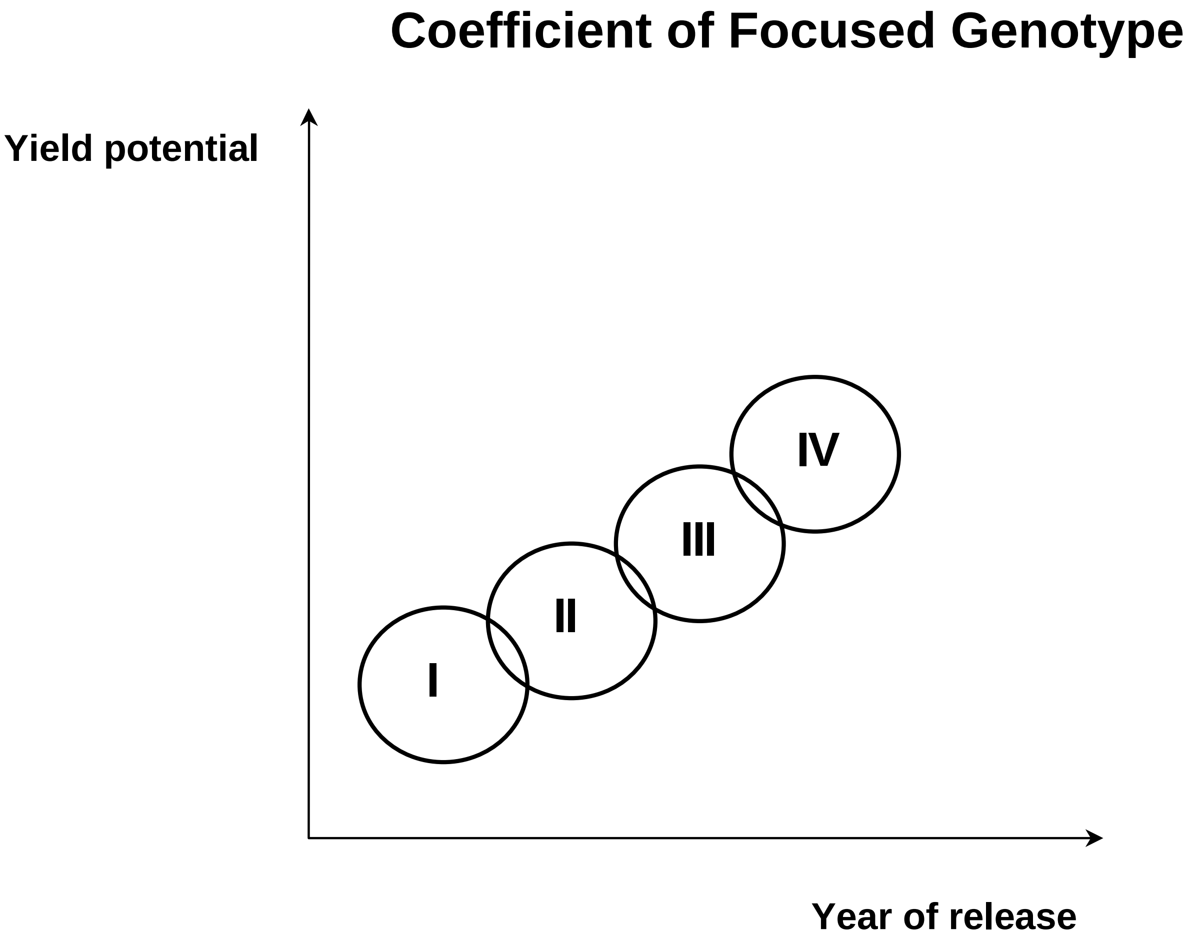
Coefficient of Focused Genotype
Yield potential
IV
III
II
I
Year of release

Figure 11

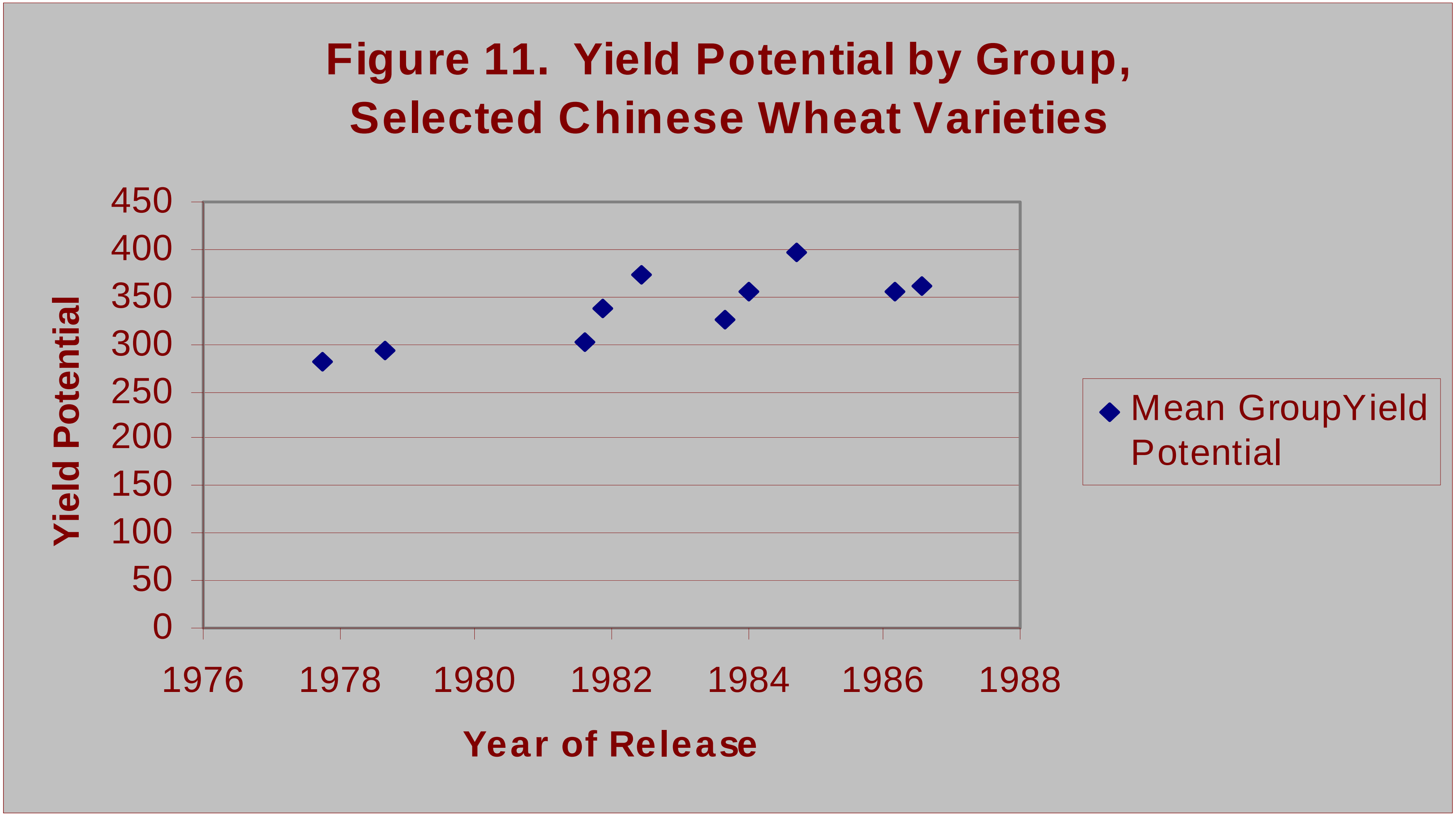
Figure 11. Yield Potential by Group,
Selected Chinese Wheat Varieties
Yield Potential
450
400
350
300
250
200
150
100
50
0
1976
1978
1980
1982
1984
1986
1988
Year of Release
Mean GroupYield Potential

Figure 12

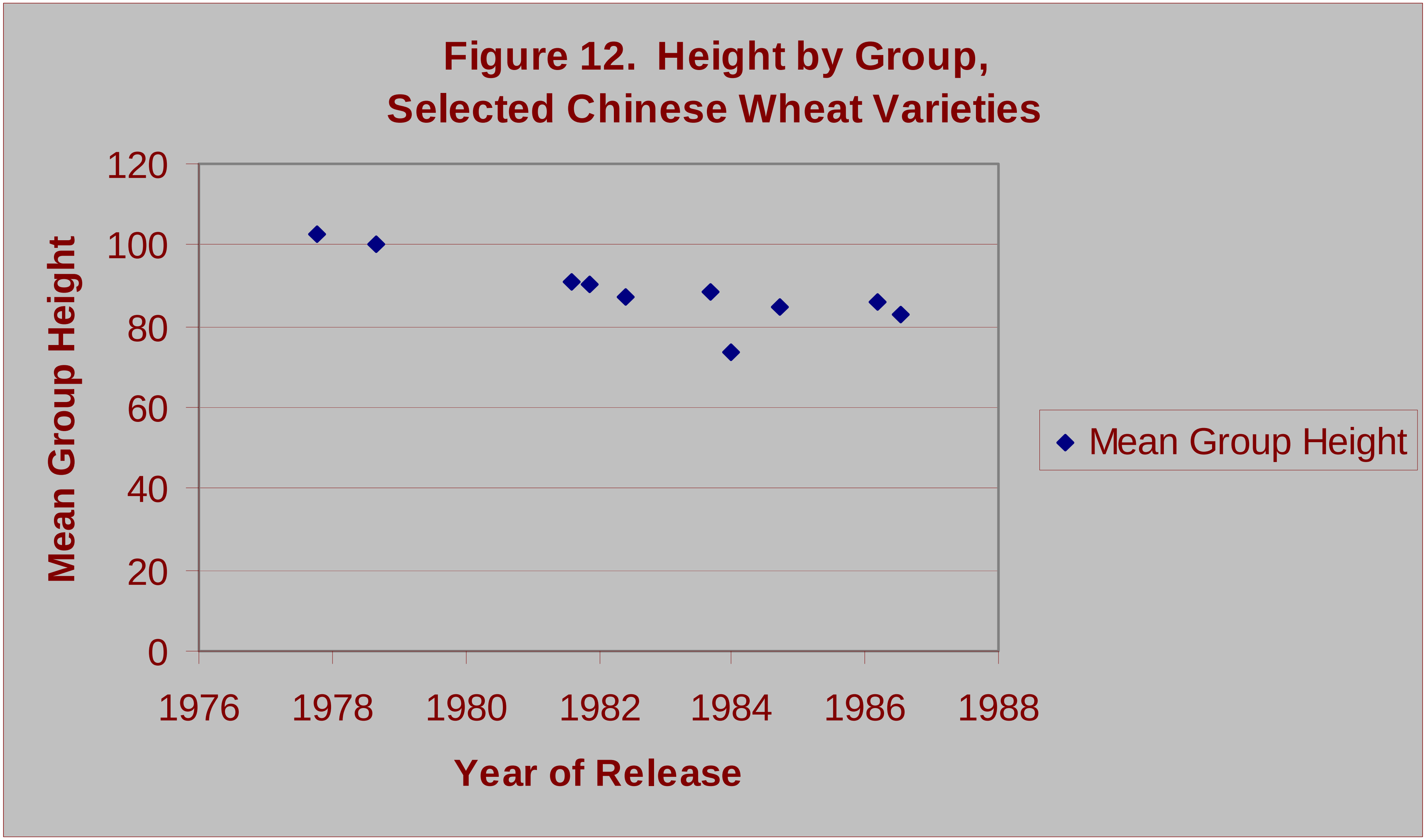
Figure 12. Height by Group,
Selected Chinese Wheat Varieties
Mean Group Height
120
100
80
60
40
20
0
1976
1978
1980
1982
1984
1986
1988
Year of Release
◆ Mean Group Height

Figure 13. CFG for Seven Chinese Provinces, 1982-1997