Market Channel Analysis of Ornamental Plants using Clustering Procedures

Mahesh Pandit, Graduate Student
Department of Agricultural Economics and Agribusiness
101 Martin D. Woodin Hall
Louisiana State University and LSU AgCenter
Baton Rouge, LA 70803
Office: (225) 578-2728
Fax: (225) 578-2716
Email: mpandi2@tigers.lsu.edu

Krishna P. Paudel, Associate Professor
Department of Agricultural Economics and Agribusiness
225 Martin D. Woodin Hall
Louisiana State University and LSU AgCenter
Baton Rouge, LA 70803
Office: (225) 578-7363
Fax: (225) 578-2716
Email: kpaudel1@lsu.edu

Roger Hinson, Professor
Department of Agricultural Economics and Agribusiness
101 Martin D. Woodin Hall,
Louisiana State University and LSU AgCenter
Office: (225) 578-2753
Fax: (225) 578-2716
Email: rhinson@aagcenter.lsu.edu

Selected Paper prepared for presentation at the Southern Agricultural Economics Association (SAEA) Annual Meeting, Orlando, Florida, 3-5 February 2013

Market Channel Analysis of Ornamental Plants using Clustering Procedures

Abstract

Market channel alternatives that include garden centers, landscapers, mass merchandisers and re-wholesalers have contributed to the growth of ornamental crops sales in the United States (U. S.). The homogenous subpopulation of the U.S. nursery producer was clustered using mixture of export method and found that there exist four homogenous group of Nursery Producer. The impact of growers' business characteristics on shares of sales to these channels by firm size was estimated using multivariate fractional regression and nonparametric model. Important explanatory variables were regions of the U.S., sales of plant groups, kinds of contract sales, and promotional expenses, and their effect varies by cluster. We also found that in some cases nonparametric estimation procedure is better than parametric estimates.

***Key words*:** garden centers; landscapers; market channels; mass merchandisers; ornamental plants; re-wholesalers; multivariate fractional regression model, nonparametric

JEL Classifications: M31, L14

Market Channel Analysis of Ornamental Plants using Clustering Procedures

Farm-level sales of greenhouse and nursery crops in the U. S. increased about 18% between 2000 and 2006 (USDA 2007). One of the factors that encouraged this growth was the availability of sales opportunities within an existing set of market channels. Major market channel alternatives that have contributed to the growth of ornamental crops sales in the United States are mass merchandiser, home centers, single location garden centers, multiple location garden centers, landscape firms and re-wholesalers. The leading channels are the landscaper channels followed by re-wholesalers (Hall et al., 2011). Among these market channels, re-wholesaler channels become most popular in these days since it add value in the industry. For example, Florida is among the top producers of nursery products and has a large consumer base, and the share of sales to re-wholesalers was 51% compared to the national average of 21% (Hall et al., 2011). The wholesales channels consists of production, transportation, storage and marketing function until it the product reached to the consumer. Further, horticultural distribution centers (HDC) may be businesses positioned as supply chain middlemen, providing products and services including accumulation and sale of plant materials at locations convenient to the trade (Hinson et al., 2012).

Conceptually, market conditions and grower characteristics determine the grower's channel. Advantages and disadvantages of channel alternatives affect the ability of growers to achieve strategy and profit objectives. Growers are assumed to maximize profit by selling to customer who are willing to pay more for higher plant quality or other distinctive characteristics.

Previous literatures have shown that nursery producer is less sensitivity to price relative to quality and services and prefer to sell to independent garden centers (Hampton, 2001). Landscapers were also similar behavior as plant materials are tend to be smaller shares of a bid than are value –adding services(Hinson et al., 2012). Mass merchandisers offered high volume of sales, but their low price strategy and concessions, including minimum quantities and logistics services, were barriers for many sellers (Hampton, 2001)

Hampton (2001) used grower characteristics to estimated impacts of factors affecting channel choice and found that larger nurseries attended more and larger trade shows, used more aggressive sales tactics, and sold larger shares of output to mass merchandisers and re-wholesalers. A recent paper by Hinson et al. (2012) studied the impact of growers' business characteristics on share of sales to these channels and found that there exist strong effects of regions of the U.S, sales of plant groups, kind of contract sales. They mainly found that re-wholesaler channel is stronger and preferred channel for ornamental plants sales.

Since, the U.S. nursery industry is large industry; we expect that there is presence of homogenous subpopulation of nursery producer. If we group nursery producer with their characteristics, and do analysis in each cluster, the result obtained from the analysis are more reliable. Given the available clustering procedure, a model based cluster analysis performs better. The literature discuss on previous paragraph are based on the parametric results and old data, and each channel coefficients are estimated using univariate estimation procedure although there exist four channels. Hence, the estimated parameter may not be consistent if they are correlated. A multivariate estimation procedure is required since we have four different market channels and they are used simultaneously. Further, if the distribution assumption used in the parametric regression procedure is wrong, we the parameter are inconsistent and less efficient. Hence, we can use nonparametric estimation procedure if the distribution assumption failed to capture the nonlinearity or true nature of variable effect on the determination of share of market channels. Hence, we used cluster analysis to identify homogeneous groups of producers using model based clustering procedure, and then explore the impact of growers' business characteristics on shares of sales to these channels by farm size using multivariate fractional regression model in each cluster. We will also compare parametric results with nonparametric results.

Method

Cluster analysis

We expect that a heterogeneous population of growers consists of a collection of homogeneous subpopulations, but these subpopulations are unknown and will be characterized by appropriate clustering methods. We anticipate that the cluster-based mixed effect model will show that covariates have significant influence on the pattern of choice of market channels by U.S. ornamental plant growers. We conduct a cluster analysis suggested by Gormley and Murphy (2010) and (Gormley and Murphy, 2008), which are based on explanatory variables. They generated mixture of expert mode (MoE) for the rank order data; which cannot be applied directly in this case, however we can modify it so that we can use it for fractional dependent variable. A mixture of experts model (Jacobs et al., 1991; Jordan and Jacobs, 1994) which combine the idea of mixture models McLachlan and Peel (2000) and generalized linear models (Gormley and Murphy, 2008) works well for fractional data.

As usual, MoE models gives the relationship between a set of response and explanatory variables, but this models assume that the conditional distribution of the response given the explanatory variables is a finite mixture distribution (Gormley and Murphy, 2008). Let K be the number of homogeneous subpopulation of a heterogeneous population also known as expert networks. *Gating network coefficient* is the probability of i^{th} nursery producer belongs to subpopulation k is π_{ik} . The probability distribution for nursery producer y_i from subpopulation k is $p(y_i|\theta_k)$, where θ represents the parameters of the model for subpopulation k. Let x_i represents associated explanatory variables for i^{th} homeowners. Than the conditional probability of nursery producer i's choice y_i, given their associated explanatory variables x_i is

$$P(y_i|x_i) = \sum_{k=1}^{K} \pi_{ik} P(y_i|\beta_k) \quad (1)$$

The gating network coefficients are weighting probabilities constrained such that they are nonnegative and sum to one for each homeowner. The probability of nursery producer i's choice according to the expert networks in the mixture model are blended by the gating network coefficients to produce overall probability. Thus the probability of nursery producer i's preference is a convex combination of the output probabilities from the expert networks.

The gating network coefficients are assumed to be function of the nursery producer explanatory variables. These explanatory variables determine their preference bloc membership. The gating network coefficients in the MoE can be estimated using multinomial logistic function, since probability of homeowners belonging to each of K expert network can be viewed as success probabilities from a generalized linear model (Gormley and Murphy, 2008). Then nursery produceer i's gating network coefficients $\pi_i = (\pi_{i1}\pi_{i2}, \ldots, \pi_{iK})$ are modeled by a logistic function of their P explanatory variables $x_i = (x_{i1}, x_{i2}, \ldots, x_{iP})$. Then, the multinomial logistic model takes the following forms

$$\log\left(\frac{\pi_{ik}}{\pi_{i1}}\right) = \theta_{k0} + \theta_{k1}x_{i1} + \theta_{k2}x_{i2} + \cdots + \theta_{kL}x_{iL} \quad (2)$$

where expert network 1 is used as the baseline expert network and θ are the gating network parameter estimates.

Fractional regression model

Within the context market channel data, each expert network need to appropriate modeled. Let's assume a nursery producer chooses different market channels that make the total portfolio of market shares. These different market channels are mass merchandiser, garden center, landscape firm, and re-wholesaler. A nursery producer therefore chooses to use a fraction of market channels that makes total of 100%. Let $Y = (y_1, y_2, \ldots, y_M)$ represent fraction of market channels of M different market channels methods. Since the values associated with these variables are in fraction so they are limited to close interval[0,1]. An appropriate model should adjust the nature of fractional variables. A solution to deal with this type of variables is to use a nonlinear function satisfying $0 \leq g(.) \leq 1$, where $g(.)$ is nonlinear model proposed by Papke and Wooldridge (1996). Hence, the conditional mean of the dependent variable can be expressed as

$$E(y|x) = g(x\beta) \quad (3)$$

with x as a matrix of independent variables and βas a vector of parameters. A fractional model is specified using logistic link with Bernoulli distribution. We estimate β by maximizing Bernoulli log-likelihood function given by

$$LL(\beta) = \sum_{i=1}^{N} y_i \log[g(x_i\beta] + (1 + y_i) \log[1 - g(x_i\beta)] \quad (4)$$

with N being the number of nursery producers. The estimated parameter is consistent and asymptotically normal provided that $E(y|x)$ is correctly specified. Different approaches are discussed in previous literature for univariate cases (Hinson et al., 2012; Papke and Wooldridge, 2008; Ramalho et al., 2011). These authors have proposed a fractional regression model on the basis of quasi-likelihood and logit conditional mean functions.

In our problem, we estimate the model simultaneously using multivariate specification. A recent manuscript by Murteira et al. (2012) has proposed generalization of a univariate specification shown in equation (1) to a multivariate specification with multinomial logit link and multivariate Bournoulli distributions[1].

Let $E(Y|X) = G(X;\beta) = [G_1(X,\beta], \ldots, G_M(X,\beta)]'$ be the m vector of conditional mean function with its components $E(y_m|X), m = 1, \ldots, M$, with $G_m = G_m(X,\beta)$. Here the the conditional mean $0 < G_m < 1$ for all m and $\sum_m^M G_m = 1$.We use multinomial logit specification expressed as:

$$G_m = \frac{\exp(X_m\beta)}{\sum_{l=1}^{M} \exp(X_l\beta)}, m = 1, \ldots, M \quad (5)$$

[1] An alternative to logit link function and Bernoulli distribution is to use a beta distribution in which density values lies between 0 and 1, however this is less common compared to the quasi-likelihood maximum likelihood estimation. A recent paper by Ramalho et al (2011) illustrated a different models and estimation procedure that can be used for multivariate fractional response variables with test procedure to check methodology and validity.

Let y_m be the fraction of m^{th} component market channels used by a nursery producer, then it follows multivariate bernouli (MB) distribution (Murteira et al., 2012). So the individual contribution to the log-likelihood can be expressed as:

$$\log L_i(\beta) = \sum_{m=1}^{M} y_{im} \log G_{im} = \sum_{m=1}^{M-1} \log \frac{G_{im}}{G_{iM}} + \log G_{iM} \tag{6}$$

Here, $G_{iM} = 1 - \sum_{m=1}^{M-1} G_{im}$. Then the quasi-maximum likelihood (QML) estimator is estimated by maximizing log-likelihood of all nursery producers (N) as given below:

$$LL(\beta) = \sum_{i=1}^{N} \log L_i(\beta) \tag{7}$$

The estimated parameter $\hat{\beta}$ is consistent and asymptotically normal regardless of the true conditional distribution$y$, provided that G is correctly specified.

Nonparametric Method

Use of a generalized kernel estimation in a nonparametric method allows incorporation of both continuous and categorical variables (Racine and Li (2004)). We consider share of sales through different channels (y_i) is affected by both continuous (Z^c) and categorical (Z^d) explanatory variables entering in a nonparametric fashion as shown in

$$y_i = g\left(Z_i^c, Z_i^d\right) + u_i \tag{8}$$

here $g(.)$ has an unknown functional form. We also assume that $u_i$has mean zero and variance$Var(u_i|Z_i) = \sigma^2(Z_i)$. A details procedure to estimate equation 8 is available on Li and Racine (2007)

Model specification test

We use the method suggested by Hsiao et al. (2007) to determine correct specification of models (parametric vs. nonparametric). This test is compatible with both continuous and categorical explanatory variables. Assume that the parametric model is correctly specified, and then the null and alternative hypotheses are:

H_0: Fractional regression model

H_1: Nonparametric model

The test statistics proposed by Hsiao et al. (2007) is

$$\hat{J}_n = \frac{n(\hat{h}_1 \dots \hat{h}_q)^{\frac{1}{2}} \hat{I}_n}{\sqrt{\Omega}} \tag{9}$$

where

$$\hat{I}_n = n^{-2} \sum_i \sum_{i \neq j} \hat{u}_i \hat{u}_j K_{\gamma,ij}$$

$$\Omega = \frac{2(\hat{h}_1 \dots \hat{h}_q)}{n^2} \sum_i \sum_{i \neq j} \hat{u}_i \hat{u}_j W_{h,ij}^2 \ L_{\lambda,y}^2$$

$$\hat{u}_i = y_i - g(Z_i, \hat{\beta})$$

Here, K is kernel function, W and L represents product kernel for continuous and discrete variable respectively. Details expressions for them are given in Appendix A1 and A2. J_n is distributed with $N(0,1)$ under the null hypothesis. J_n test diverges to $-\infty$ if H_0 is false. Thus, we reject the null hypothesis for large values of the test statistics.

Data, Variables Used and Justification

Data for this analysis were obtained from the National Nursery Survey, 2009[2]. Data about sales, employment, product types and forms, market channels, production and marketing practices, regional trade, and other influencing factors were collected for the year 2008 using mail and e-mail surveys in 50 U.S. states. A list of nursery plant producers was taken from sources that included National Plant Health Board, departments of agriculture in each state, grower associations, and business databases. Dillman (2000) protocol was used for design and implementation of surveys. The survey was sent to 15,000 producers by regular mail and to 1,900 producers in 12 states by email. A total of 3,044 valid responses was received for a 17% response rate. Of these responses, 312 were from the e-mail survey. Descriptive statistics of e-mail and mail survey respondents are similar[3] in nature so we analyzed the data that combines both email and mail survey respondents. Based on information available we merged market channels as: 1. Re-wholesalers (rw), 2. Mass Merchandise + Home Centers (mm), 3. Single and Multiple Location Garden Centers (gc), and 4. Landscape Firms (ls). Figure 1 shows that landscape channel is the most common use channel with 43.72% of respondents indicated that they use landscape firms as market channels. Second most used market channel is re-wholesalers with 38.16% followed by garden center with 35.61% as the third major market channels. The least common use market channels is mass merchandiser and home center with only 5.66% respondents indicated that they use it. A more details about market channels can be found in Hodges et al. (2009).

---

[2] The 2009 National Nursery Survey was conducted by the Green Industry Research Consortium of University Horticulturists and Economists, organized as a multi-state project (S-1021) under the National Institute of Food and Agriculture at the U.S. Department of Agriculture.

[3] We use two sample mean comparison t-test for each variable and found that they are not significant.

The share of market channels are available in percentage in raw data and modified in proportion for analysis propose. So they are varying between 0 and 1. For each producer sum of fraction of market channels is equal to one. Hodges et al. (2008) found that significant regional differences existed with respect to use of market channels. Based on the number of firms, we divided growing regions into five categories (Midwest, Northeast, Pacific, Southeast, and others) as defined in Table 2[4].

We expect that type of plant is an important factor for the share of market channels. Nursery plant producers use computers for functions such as accounting/cost analysis, inventory, financial investment analysis, and digital imaging for disease diagnosis (Hodges et al., 2010). Computer technology enables nursery producers to evaluate benefits of different market channels; hence the use of computer can be an important factor that determines the fraction of market channels.

Firms specialize for production of different categories of plants (vine, annuals, trees etc) so the sales volume of these plant category effect choice of market channels, so plant types are used as explanatory variable in the model. The literature suggests that operator's age is important in many economic choice decision (Pandit et al., 2012; Paudel et al., 2013). Operator age was not available, so age of the firm was included.

The emergence of mass merchandisers as a form of retail for nursery plants has led to a higher share of plants reported as sold through contracts. These contracts reduce risk in some senses, but also may increase risk. In the contract relationship, growers are expected to have the agreed upon number of plants available for shipment. However, if

[4] The states corresponding to the regions categories are Midwest: IA, IL, MI, MN, MO, OH, and WI; Northeest: CT, DE, MA, ME, NH, NJ, NY, PA, RI, and VT; Pacific: AK, CA, HI, OR, and WA; Southeast: AL, FL, GA, MS, and SC; Others: AR, AZ, CO, ID, KS, KY, LA, MT, NC, ND, NE, NM, NV, OK, SD, TN, TX, UT, VA, WV, and WY.

retail demand is not sufficient, the retailer does not order shipment and the grower must find other outlets for the material. Still, the contract relationship typically does encourage or require the grower to expand and to control costs diligently. We expect that firms selling through contracts are more likely to choose appropriate share of market channels to maximize production. Further, we expect that choice of market channels are associated with type of contracts, such as contract to other producers, contract to garden center and contract to mass merchandisers.

We expect that different kinds of promotion increase sales of product of a firm. As sales increase, we expect that the chance of optimal share of these market channels increased due to higher profitability linked to increased demand. We used number of trade shows attended in 2008 (trade), web site promotion expenses and trade show promotion expenses as variables impacting the share of market channels. The choice of these explanatory variables is consistent with the study by Pandit et al. (2012) and Hinson et al. (2012).

We expect farm size as measured by total sales volume to have effect on choice of market channels by nursery producers. Previous literature has shown that small and large nursery plant producers behave differently (Hinson et al., 2012). In order to address effect of farm size we use farm income dummy in the model. A farm income variable is used to define farm size with annual sale volume above $500K per year equal to 1, and 0 otherwise. A summary statistics and definition of dependent and independent variables are provided in Table 1.

Results and Discussion

As described in method section we make cluster analysis using the method suggested by (Gormley and Murphy, 2008). Our result shows that we have four homogeneous populations in the nursery producer data. We found that number of observation in each cluster as follows: cluster1 (579), cluster2 (441), cluster3 (1609) and cluster4 (57). Multivariate fractional and nonparametric model are estimated for each cluster.

Cluster 1

Parametric estimates

The parametric estimates for market channels for cluster 1 is given in Table 2 and their marginal effect is given in Table 3. Average marginal effect shows that nursery producer who lives in Midwest, Northeast and other regions use 52.73%, 56.62% and 57.844% more share of mass merchandiser channels compared to nursery producer in southeast region[5]. Our result states that the nursery producers located in Midwest region use highest share of (68%) of market channels form garden center compared to nursery producer in southeast region. We found that cotton producer in Pacific use landscape as least share of market channels compared to cotton producer in southeast region. Further, results shows that share of market channels by nursery producer located in the Midwest region from re-wholesalers is less by 22% compared to nursery producer located in Southeast region. Higher sales of Other plants (*pg5*) means the nursery producer uses higher share of mass merchandiser. Similarly higher sales of foliage plants imply that nursery producer uses higher share of landscape firm by 0.122.

---

[5] In multinomial logit model the sign of the marginal effect (ME) may not be the same sign of regression coefficients. $\frac{\partial p_{ij}}{\partial x_j} = p_{ij}(\beta_j - \bar{\beta}_i$, and $\bar{\beta}_i = \sum_l p_{il}\beta_l$. For a variable$x$, the ME is positive if $\beta_j > \bar{\beta}_i$.

An increase in total sale under contract (*ctcts*) implies that nursery producer uses less share of mass merchandiser by 0.032. Higher number of trade show attended (*trade*) indicates that nursery producer uses higher share of garden center by 0.014.Results indicate that higher website promotion expenses (*pawsss*) means nursery producer uses higher share of f landscape firm. Older farmer (*agef*) is likely to use less share of of re-wholesalers.

Nonparametric estimates

The partial regression plots from nonparametric estimate for cluster1 are given in figure from 2 to 5 for mass merchandiser, garden center, landscape firm and re-wholesalers respectively. For mass merchandiser this figures shows that trade show attend(trade) and trade show promotion expenses has nonlinear relationship with higher sales through mass merchandiser who attend trade 5-10 and trade show expenses 65-85. Figure 3 shows that higher sales of shurbs/trees and other plants, trade show promotion expenses means that nursery producer uses higher share of garden center. Age of firm has quadratic relation share of garden center. The nonparametric estimate for landscape firm is given in Figure 4. This figure shows that higher sales of trees/shurbs, number of IPM practices adoption indicate that nursery producer use higher sharer of landscape firm. There is no distinct pattern number of trade show attended and website promotion expenses. Larger firm has higher rate of share of market channels through landscape firm. In case of re-wholesalers nonparametric estimates shows that there is increase in sales through re-wholesalers for higher sales for trees/shrubs. Nursery producer who use computer for farm management aids has higher sales through re-wholesalers.

Cluster 2

Parametric estimates

The parametric estimates for market channels for cluster 2 are given in Table 2 and their marginal effect is given in Table 4 . Average marginal effect shows that nursery producer who located in Midwest, Northeast and Pacific regions use 0.60%, 0.61% and1.09 % less share of mass merchandiser compared to nursery producer in southeast region. Result states that the nursery producers in Pacific region use13.96% more share of market channels form garden center compared to southeast region.

Our result shows that higher sale of trees/shrub plants (*pg1*) means nursery producer uses less landscape firms by 0.002. We also found that higher sales of vines (*pg3*) mean that nursery producer uses less share of mass merchandiser. In contrast, higher sales of vines (pg3) and foliage (pg4) indicates that nursery producer uses higher share of landscape by 0.001 and 0.013 respectively. We found that if the nursery producer is contracted to mass merchandiser (*tcmm*) sales through mass merchandiser is less by 0.6% compared to the producer who do not contracted. Resuts indicates that higher expenses on website promotion (*pawsss*) suggest that nursery producer uses more share of mass merchandiser, and trade show expenses (*patss*) suggest that nursery producer uses more share of garden center. Further, we found that nursery producer who use computer for farm management use mass merchandiser more by 1.3% then who do not use computer for their farm management. If the nursery producer adopts more IPM practices the share of market channels through mass merchandiser is less by 0.001, but for garden center is more by 0.006. Further we found that large farm uses mass merchandiser as a market channels more by 6.62% compared to small firm.

Nonparametric estimates

The non-parametric estimates for cluster 2 are given from figure 6 to 10. Figure 6 shows that there is increase in sales through mass merchandiser with increase in sales of trees/shurbs upto certain point and then decreases. Further, we found that there exists regional difference on share of sales through garden center. This figures show that higher sales of trees/shurbs and number of IPM adoption has higher share of sales through garden center. We found that a higher sale of trees/shurbs, tradeshow promotion expenses has negative impact on sales through landscape firm. In contrast we found that increase in sales of foliage and vines trees, and age of firm increases sales through landscape firm up to a certain threshold and then decreases. Nursery producer who use computer, and have large firm share more garden center than who do not use computer and small firm respectively. We also found that there exists regional difference on sales through re-wholesalers. Further, higher sales of trees/shrub, total sales under contract, contracted to mass merchandiser, trade show promotion expenses, age of firm, and use of computer imply that nursery producer sales higher amount through re-wholesalers. For number of IPM practices adoption we found quadratic type of relation.

Cluster 3

Parametric estimates

The parametric estimates for market channels for cluster 3 are given in Table 2 and their marginal effect is given in Table 5. Average marginal effect shows that nursery producer who lives in Midwest, Northeast, Pacific and Other regions use 14.35%, 15.27%, 10.54% and13.80% more share of garden center compared to nursery producer in southeast region. Result states that the nursery producers in Pacific region use 12.97% less share of market channels form pacific compared to southeast region. Further results shows that nursery producer who lives in Midwest, Northeast, and Other region use 14.37%, 11.63%, and 10.25% use less re-wholesalers compared southeast region.

Higher sales of bedding plants (pg2) mean that nursery producer sales higher amount of nursery product through garden center. We also find that higher sale of foliage (pg4) and other (pg5) plants means nursery producer sales more through re-wholesalers. Higher total sales under contract imply that nursery producer sales higher amount of nursery producer through re-wholesalers. If the cotton producer are contracted to garden center sales through garden center is more by 31.03%, land scape firm and re-wholesalers are less by 16.44% and 13.20% respectively compared to nursery producer who do not contracted.

We found that higher trade show promotion expenses (*patss*) suggest that higher sale through landscape but less sales through re-wholesalers. In contrast we found that if the nursery producer uses computer sales through landscape is 10.23 more and sales through re-wholesalers is less by 8.79%. If the nursery producer adopt more IPM (*nipm*) practices, nursery producer sales less through land scape. If the farm size (farm-size) is large, sales through mass merchandiser, land scape are more by 3.01% and 6.73% respectively compared to small firm. In contrast we found opposite effect for garden center and re-wholesalers with marginal values 5.48% and 4.32% respectively.

Nonparametric estimates

The nonparametric estimates of cluster 3 for each market channels are given in figure 10-14. This figure show that higher sale of foliage, other plants suggest that nursery producer sale higher through mass merchandiser, however if nursery producer contracted to garden center or higher website promotion expenses implies that nursery producer sales less through mass merchandiser. Higher sales of bedding plants, contracted to garden center, and higher number of IPM practices adoption indicate that nursery producer sales more through garden center. In contrast, there is less use of garden center for more number of trade show attended. For landscape firm, we found that number of trade show attended has positive impact on sales through landscape, and large firm are more likely to sales through landscape firm. We also found regional difference for sales through re-wholesalers. The variables that have positive effect on it are sales of foliage plants, total sales under contract, website promotion expenses, and number of IPM adoption. Whereas, sale of bedding plants has negative effect and numbers of trade attend have negative effects.

Cluster 4

Parametric estimates

The parametric estimates for market channels for cluster 4 are given in Table 2 and their marginal effect is given in Table 6 . Average marginal effect shows that nursery producer who lives in Pacific and Other regions use 21.87% and 16.51 less share of mass merchandiser compared to nursery producer in southeast region. Further, Nursery producer in pacific region use 30.03% more garden center and in other region use 13.70% less re-wholesalers. Results indicate that higher sale of trees/shrub (pg1) imply that nursery producer use less mass merchandiser whereas they use more landscape firms. Higher sales of bedding plants (pg3) and foliage (pg4) imply that nursery producer use less garden center and landscape firm. Further, higher sales of Other plants indicate that nursery producer use more wholesalers.

If the nursery producer is contracted to garden center (*tcgc*) sales through landscape firms is 7.97% more compared to who are not contracted. Further if the they are contracted to mass merchandiser sales through mass merchandiser and land scape firms is 24.34% and 10.47% more compared to who are not contracted to mass merchandiser. On the other hand we found less use of garden center with average marginal effect of 37.34%. If the number of trade show attended increases, the sales through mass merchandiser increases by 0.255.

Higher website (*pawsss*) and trade show (*patss*) expenses means that nursery producer sales less through mass merchandiser where as they use more landscape firms and re-wholesalers. If the farm age (*agef*) is rises, nursery producers sale less through landscape firm. If the nursery producer uses computer (*dcomp*) for their farm management aids the sales through mass merchandiser and landscape firm is more by 18.53% and 8.27% respectively, If number of IPM practices increases the sales through landscape firm decreases. Further large farm size nursery producer are less likely to sell through mass merchandiser but more likely to sales through re-wholesalers.

Nonparametric estimates

Nonparametric estimates for each market channels for cluster 4 are given in Figure 14-17. There also exists regional difference on sales through mass merchandiser for cluster 4. We found that there is higher sales through mass merchandiser for higher in sales of bedding, foliage and other plants, contracted to garden center, higher number of trade show attended, and higher website promotion expenses. For higher sales through contract, nursery producer use more sales from mass merchandiser up to a threshold level after that they sales through mass merchandiser. We found that, use of computer for their farm management aids imply that nursery producer use less mass merchandiser. Higher sales of bedding plants, total sales under contract, contract to garden center, website promotion expenses, age of firm suggest nursery producer use less garden center. Further, we found that higher sales of trees/shrubs and higher number of IPM practices imply that nursery producer sales more nursery product through landscape firm. Nursery producer, who use computer for their farm management use less landscape firm than those who do not use computer. Finally, we found that higher trade shows expenses, number of IPM practices adoption means higher sales through re-wholesalers.

Model comparison test

We use method suggested by Hsiao et al. (2007) to compare parametric and nonparametric estimate as a model specification test. The model specification test is given in table 7. Table 7 shows that for cluster 1 garden center equation, nonparametric estimation is better than parametric estimation. For cluster 2, we found that except mass merchandiser, all equation nonparametric estimation is better than nonparametric estimation. In case of cluster 3, our result shows that landscape equation should be estimated using nonparametric model. Further, re-wholesalers and garden center are better if they are estimated nonparametric ally for cluster 4.

Conclusions

We used model based clustering procedure to find homogenous population of the U.S. nursery producer. Our results show that there exists four homogenous subpopulation of the U.S. nursery producer. We estimate multinomial fractional model (parametric) and nonparametric model to analyze market channels used by the U.S. nursery producer. Then, the fractional model and nonparametric model was compared using Hsiao et al. (2007) test. Our results show that three groups of variables - region, the plant groups, and the contract production group- played an important role in channel choice. Our results suggest that parametric mode is not specified in all cases so we need to use nonparametric model if parametric model is not appropriate.

We found that for cluster 1, sales of plant group foliage, total sale under contract has positive effect on sales through mass merchandiser; number of trade show attended has positive effect on garden center but negative effect for land scape farmers. Trade show promotion expenses have positive effect on landscape firm but negative effect on re-wholesalers and age of firm has negative impact on re-wholesalers.

For cluster 2, sales of vines plants, farm age and number of IPM adoption has negative effect for mass merchandiser whereas website promotion expenses and use of computer for farm management aid have positive effect. Further, trade show expenses and number of IPM adoption have positive impact on share of market channels through garden center. Similarly, sales of vines and foliage have positive effect on fraction of sales through landscaper firm. Further, we found that sale of other plant group has negative impact on re-wholesalers.

For cluster 3 we found that sales of trees/shrub has negative impact on sales of mass merchandiser and sales of bedding plants, and total sales under contract has positive effect on market channels through garden center. We also found that sales of foliage and other plants, total sales under contract, total sales under contract, trade show promotion expenses and number of IPM adoption has negative effect on landscape firm. For re-wholesaler, plant group trees/shrub, total sales under contract and use of computer for farm management has positive impact, whereas sale of vines, foliage and other plants, total sales under contract and tradeshow promotion expenses has positive effects.

For cluster 4, the variables that have positive effect on mass merchandiser are contracted to mass merchandiser, trade show attended, the variables that have positive effect on mass merchandiser are contracted to mass merchandiser, trade show attended, computer management aids and that have negative effects are sales of trees/shrubs plants, website and trade show promotion expenses. We also found that sales of vines and total sales under contract have negative effect on market channels through garden center. Sales of trees/shrubs, contract to garden center and mass merchandiser, website and trade show promotion expense and computer managements aids has positive effect on market channels through landscape firm, on other hand sales of vines and foliage plants, firm age and number of IPM practices adopted has negative effects. Finally, sales of other plants, tradeshow and website promotion expenses, and farm size have positive effect on market channels through re-wholesalers, whereas computer management aids has negative effects.

Although we found that most of the cases parametric model are specified in this case. However if there is non-linear relationship exist we should estimate them nonparametric ally. A flexible model which combine both parametric and nonparametric model can be used in future that are likely to be better specified. Overall, this study was intended to assist growers' understanding of opportunities, including identification of business characteristics associated with increased or reduced shares of wholesale sales made through four different market channels. Results from our study suggest progress in understanding the relationships between channels, but additional studies are needed to confirm these findings. A caution is that data for this analysis represent only growers' opinions and perceptions. Studies of other links in the supply chain, including re-wholesalers, retailers, and service providers such as landscapers, could provide other perspectives

References

Dillman, D. 2000. *Mail and Internet Surveys: The Tailored Design Method*. New York: John Wiley & Sons.

Gormley, C., and B. Murphy (2010) Clustering Ranked Preference Data Using Sociodemographic Covariates, ed. S. Hess, and A. Daly, 1st Edition, pp. 543–569.

Gormley, I.C., and T.B. Murphy. 2008. "A Mixture of Experts Model for Rank Data with Applications in Election Studies." *The Annals of Applied Statistics* 2(4):1452-1477.

Hall, C., A. Hodges, M. Palma, and A. Collart. 2011. "2008-09 National Nursery Survey: Preliminary Results." Texas A&M University.

Hampton, W.R. 2001. "Trade Flows and Marketing Practices of Louisiana and Gulf States Nurseries." M.S. thesis, Louisiana State University and Agricultural and Mechanical College.

Hinson, R.A., K.P. Paudel, and M. Velastegui. 2012. "Understanding Ornamental Plant Market Shares to Rewholesaler, Retailer, and Landscaper Channels." *Journal of Agricultural and Applied Economics* 44(2):173–189.

Hinson, R.A., K.P. Paudel, M. Velastegui, M.A. Marchant, and D.J. Bosch. 2012. "Understanding Ornamental Plant Market Shares to Rewholesaler, Retailer, and Landscaper Channels." *Journal of Agricultural and Applied Economics* 44(02):173–189.

Hodges, A., M. Palm, and C. Hall. 2009. "Trade Flows and Marketing Practices within the U.S. Nursery Industry, 2008."

Hodges, A., M. Palma, and C. Hall. 2010. "Trade Flows and Marketing Practices within the Us Nursery Industry, 2008." *Southern Cooperative Series Bulletin* 411:63.

Hodges, A.W., C.R. Hall, B.K. Behe, and J.H. Dennis. 2008. "Regional Analysis of Production Practices and Technology Use in the Us Nursery Industry." *HortScience* 43(6):1807-1812.

Hsiao, C., Q. Li, and J.S. Racine. 2007. "A Consistent Model Specification Test with Mixed Discrete and Continuous Data." *Journal of Econometrics* 140(2):802-826.

Jacobs, R.A., M.I. Jordan, S.J. Nowlan, and G.E. Hinton. 1991. "Adaptive Mixtures of Local Experts." *Neural Computation* 3(1):79-87.

Jordan, M.I., and R.A. Jacobs. 1994. "Hierarchical Mixtures of Experts and the Em Algorithm." *Neural Computation* 6(2):181-214.

Li, Q., and J.S. Racine. 2007. *Nonparametric Econometrics: Theory and Practice*. New Jersey: Princeton University Press.

Murteira, J., E. Ramalho, and J. Ramalho. 2012. "Regression Analysis of Multivariate Fractional Data." mimeo.

Pandit, M., K.P. Paudel, and R. Hinson. 2012. "Intensity of Integrated Pest Management (Ipm) Practices Adoption by Us Nursery Crop Producers." Selected paper presented at the Annual meeting of the Agricultural and Applied Economics Association, Seattle, Washington, August 12-14.

Papke, L.E., and J.M. Wooldridge. 1996. "Econometric Methods for Fractional Response Variables with an Application to 401(K) Plan Participation Rates." *Journal of Applied Econometrics Journal* 11(6):619–632.

Papke, L.E., and J.M. Wooldridge. 2008. "Panel Data Methods for Fractional Response Variables with an Application to Test Pass Rates." *Journal of Econometrics* 145(1):121-133.

Paudel, K., M. Pandit, and M. Dunn. 2013. "Using Spectral Analysis and Multinomial Logit Regression to Explain Households' Choice Patterns." *Empirical Economics*:1-22.

Racine, J., and Q. Li. 2004. "Nonparametric Estimation of Regression Functions with Both Categorical and Continuous Data." *Journal of Econometrics* 119(1):99-130.

Ramalho, E.A., J.J.S. Ramalho, and J.M.R. Murteira. 2011. "Alternative Estimating and Testing Empirical Strategies for Fractional Regression Models." *Journal of Economic Surveys* 25(1):19-68.

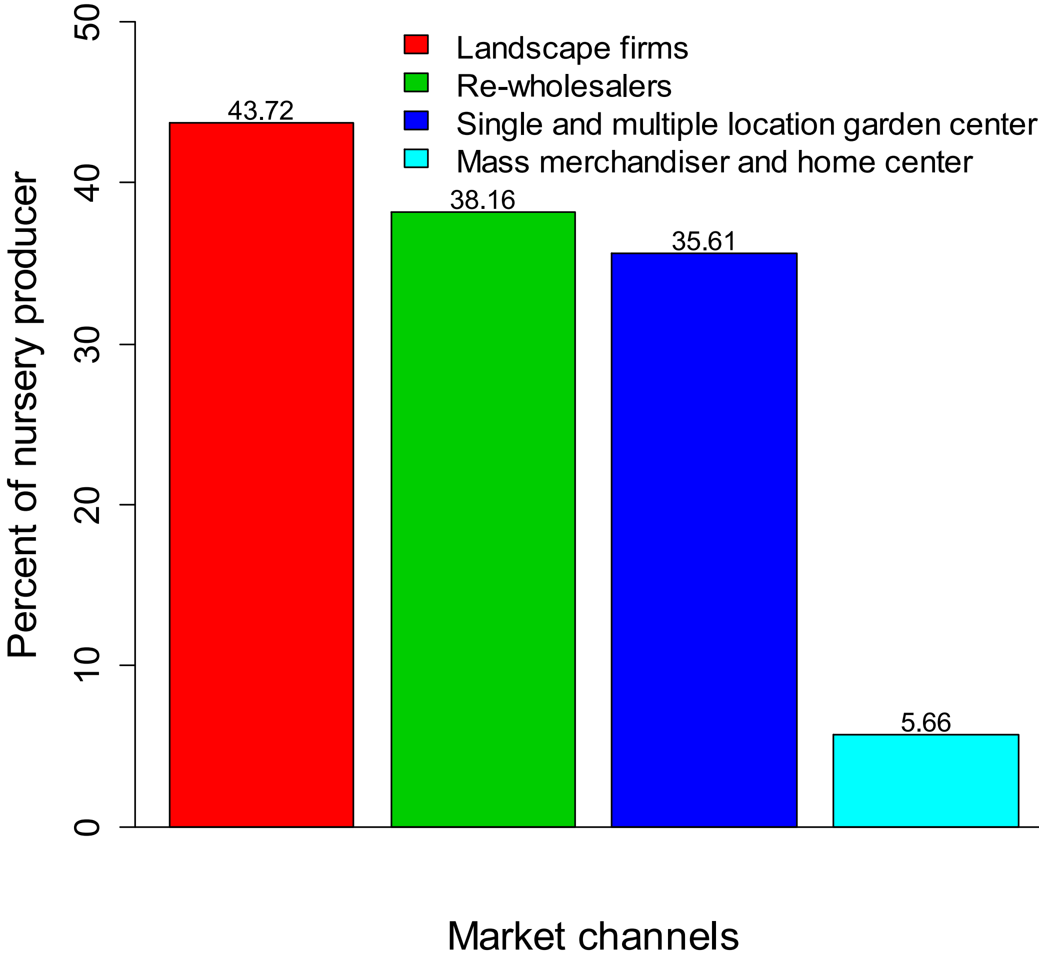

Figure 1: Market channels sales of nursery products in the U.S., 2008.

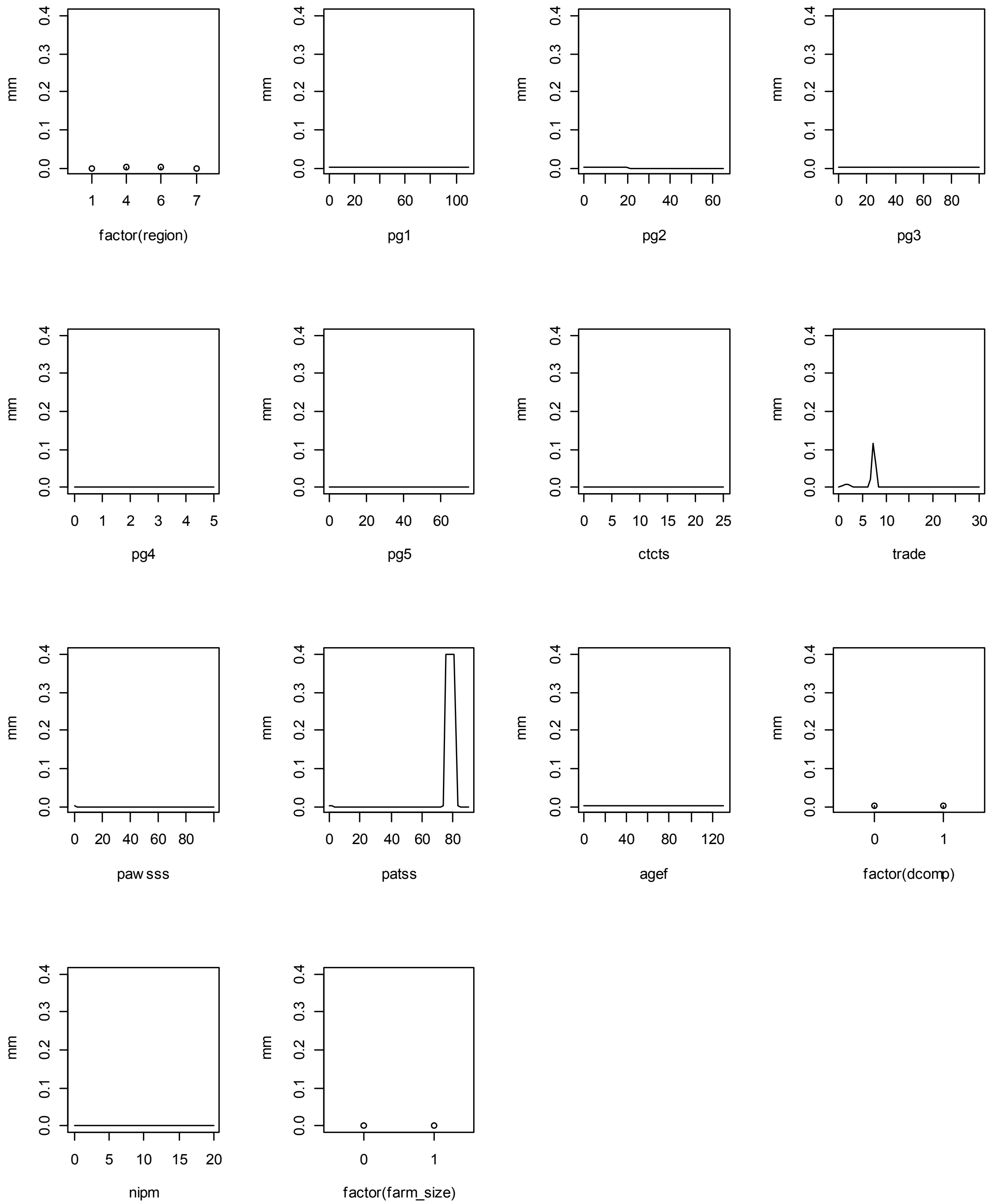

Figure 2: Partial regression plots obtained from nonparametric model of mass merchandiser (Cluster 1)

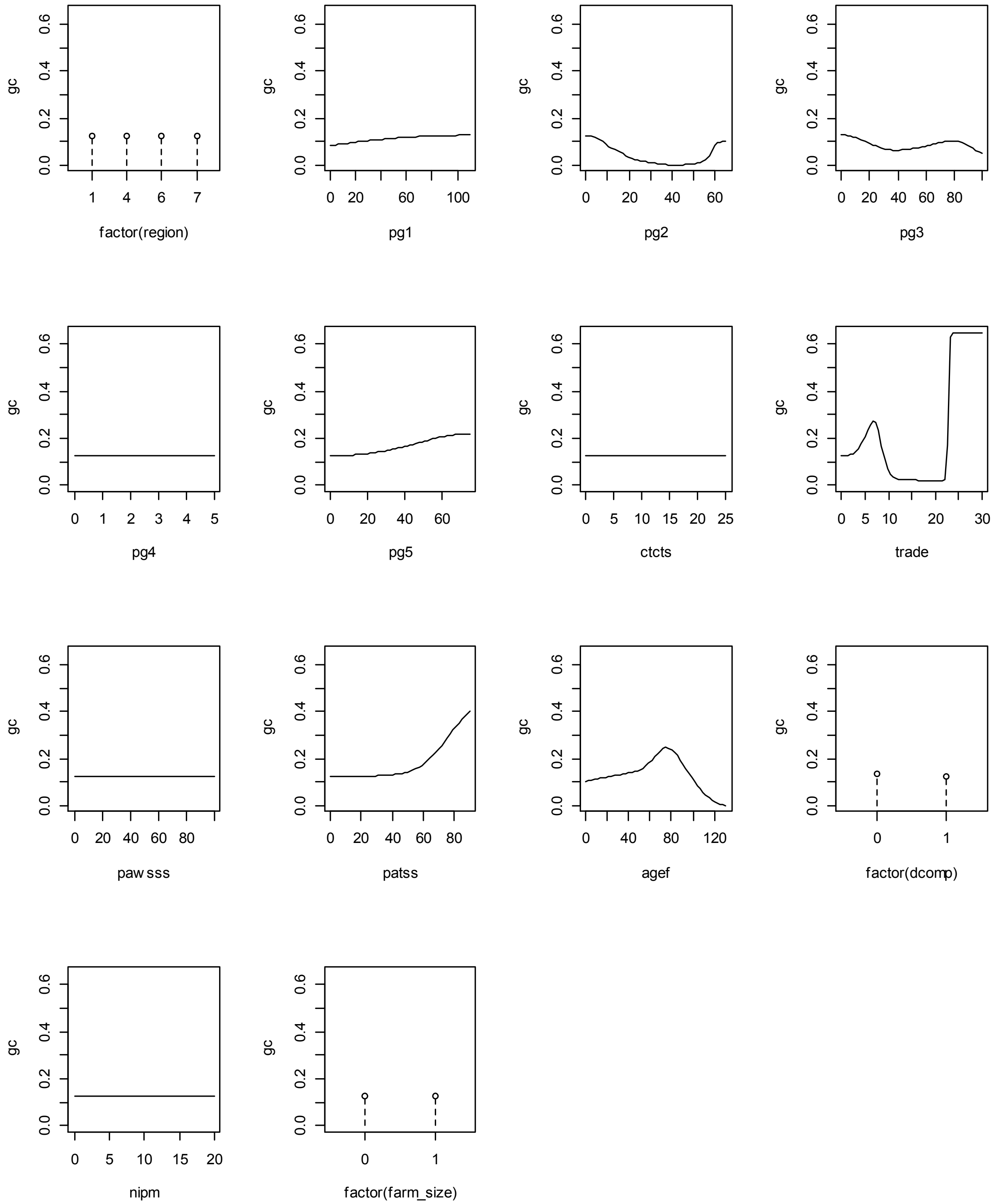

Figure 3: Partial regression plots obtained from nonparametric model of garden center (Cluster 1)

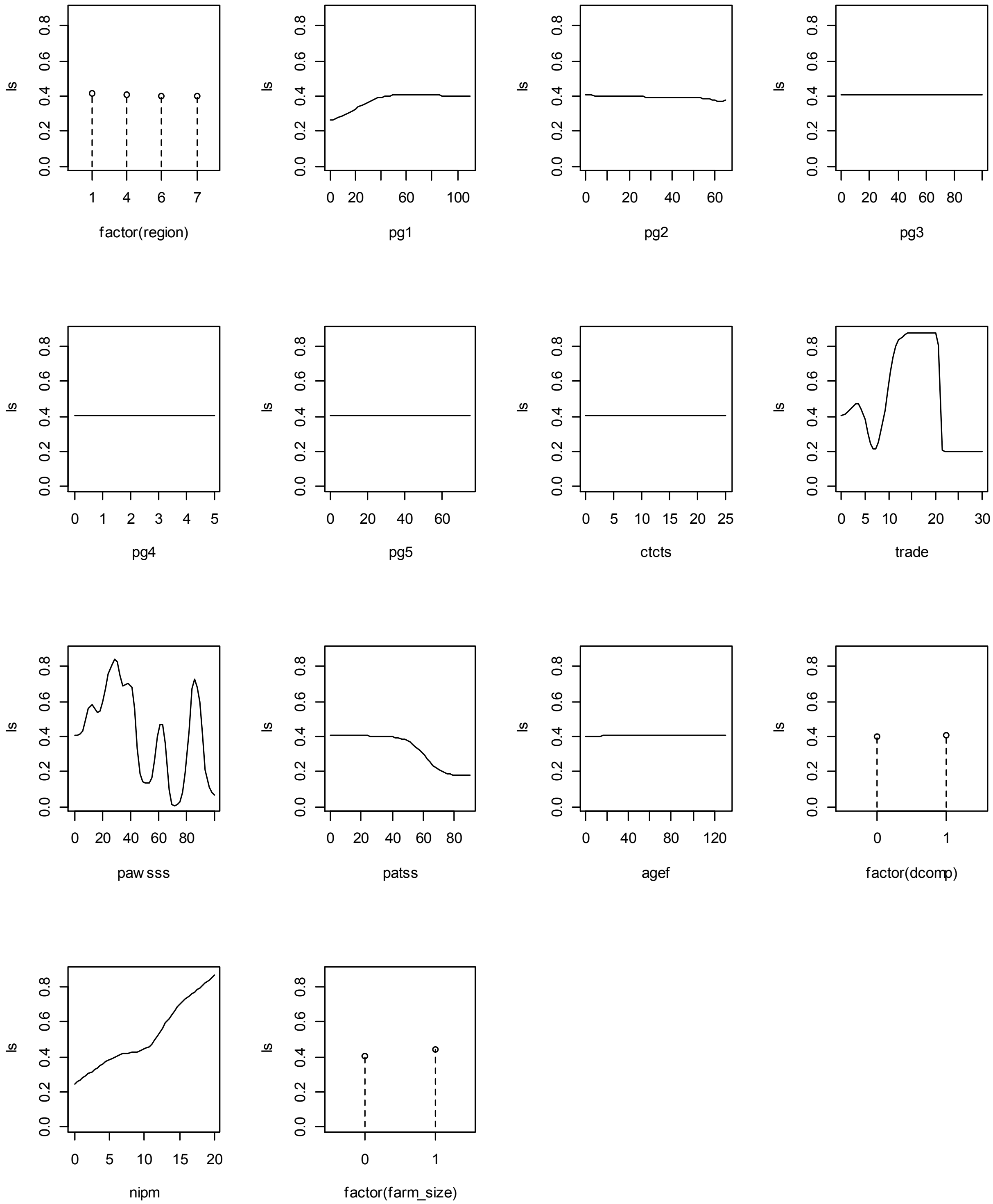

Figure 4: Partial regression plots obtained from nonparametric model of landscape firm (Cluster 1)

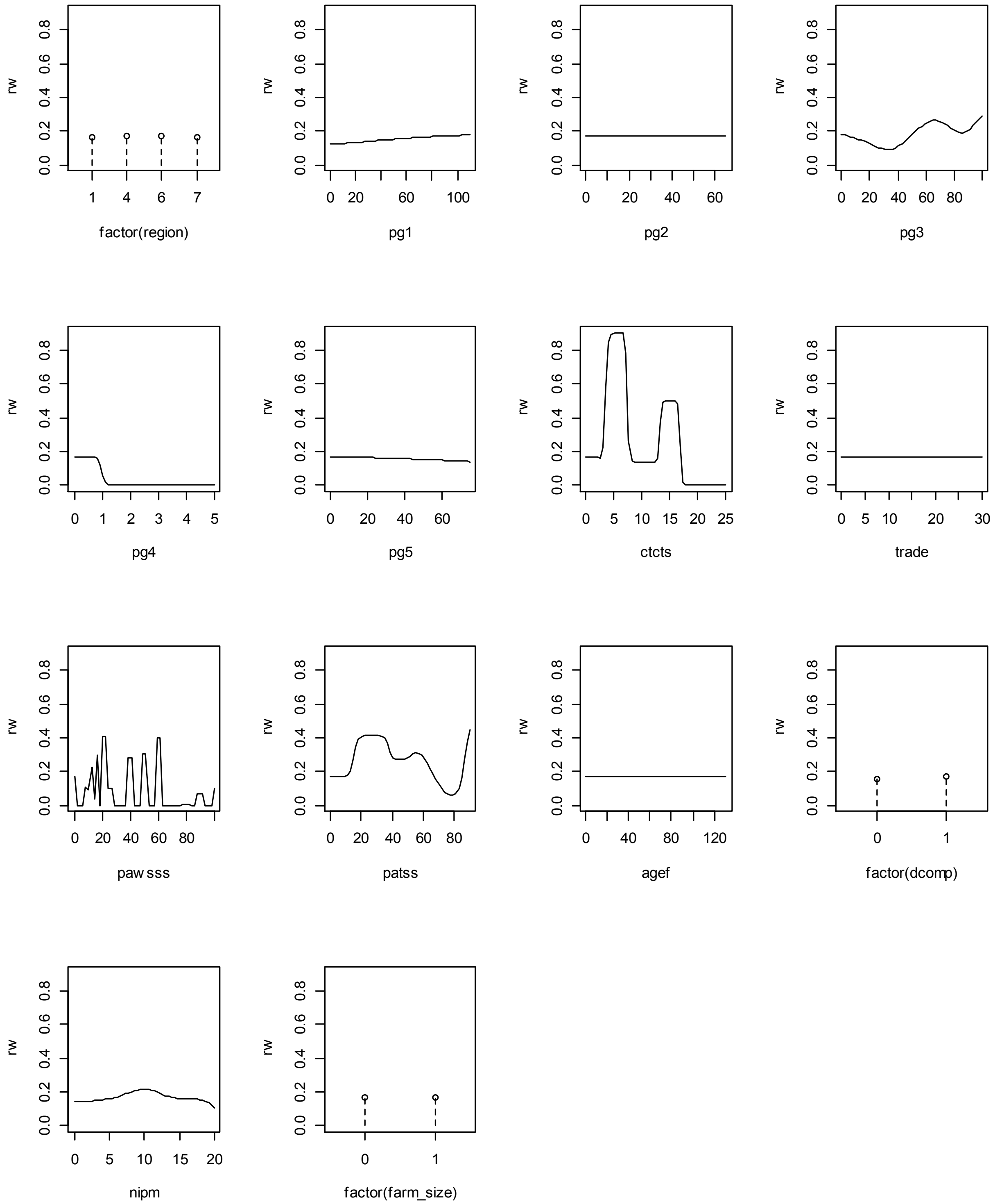

Figure 5: Partial regression plots obtained from nonparametric model of re-wholesalers (Cluster 1)

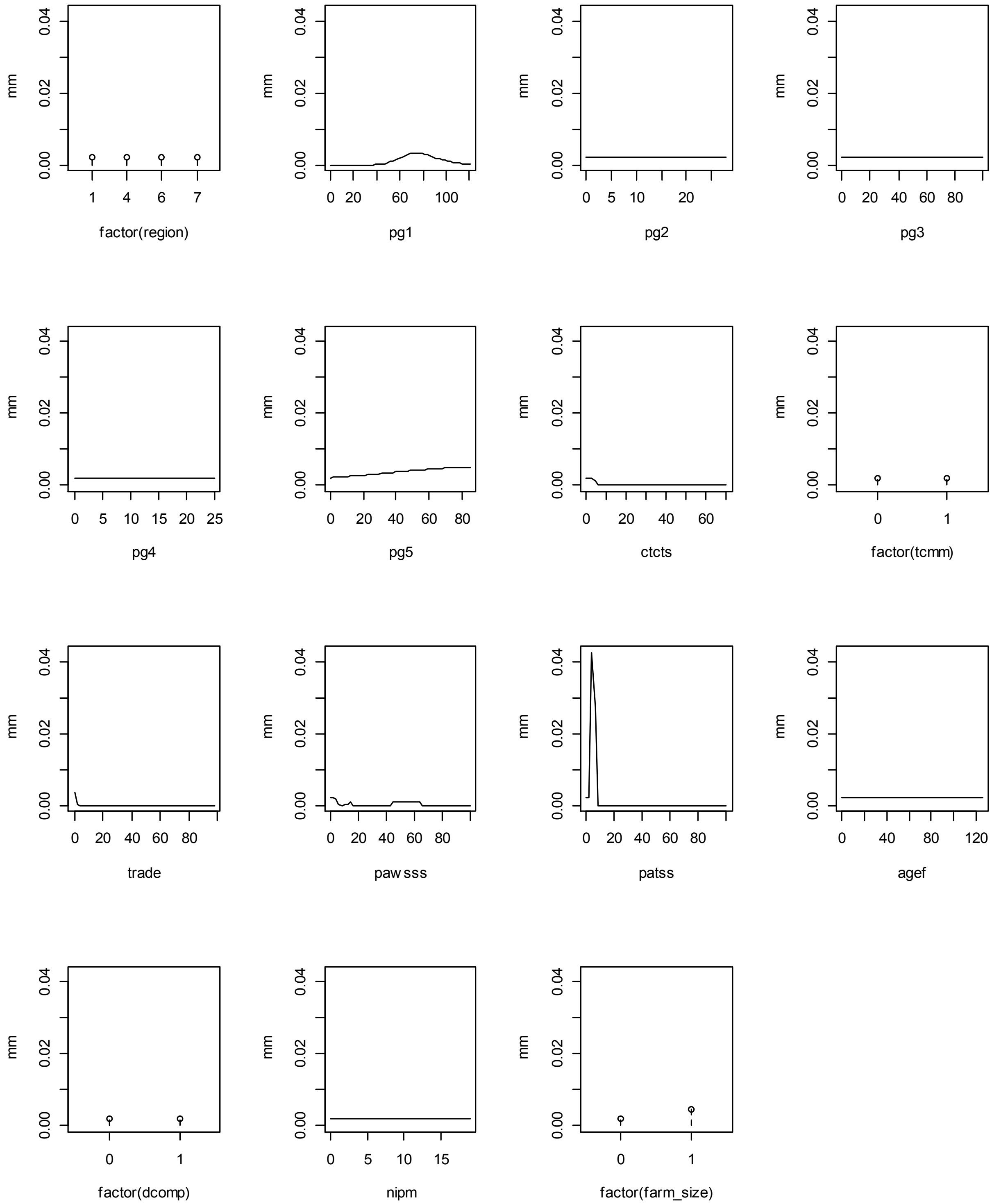

Figure 6: Partial regression plots obtained from nonparametric model of mass merchandiser (Cluster 2)

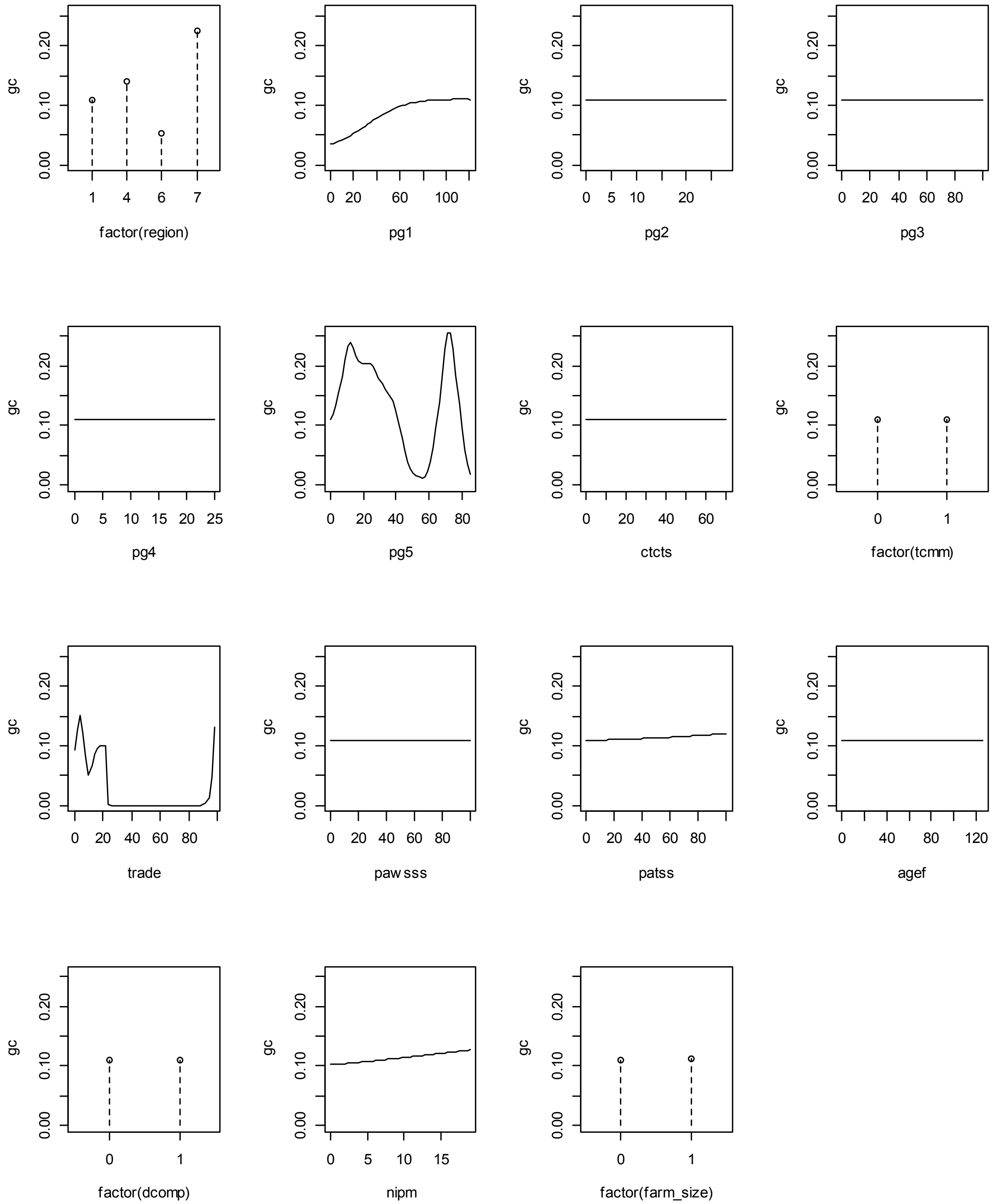

Figure 7: Partial regression plots obtained from nonparametric model of garden center (Cluster 2)

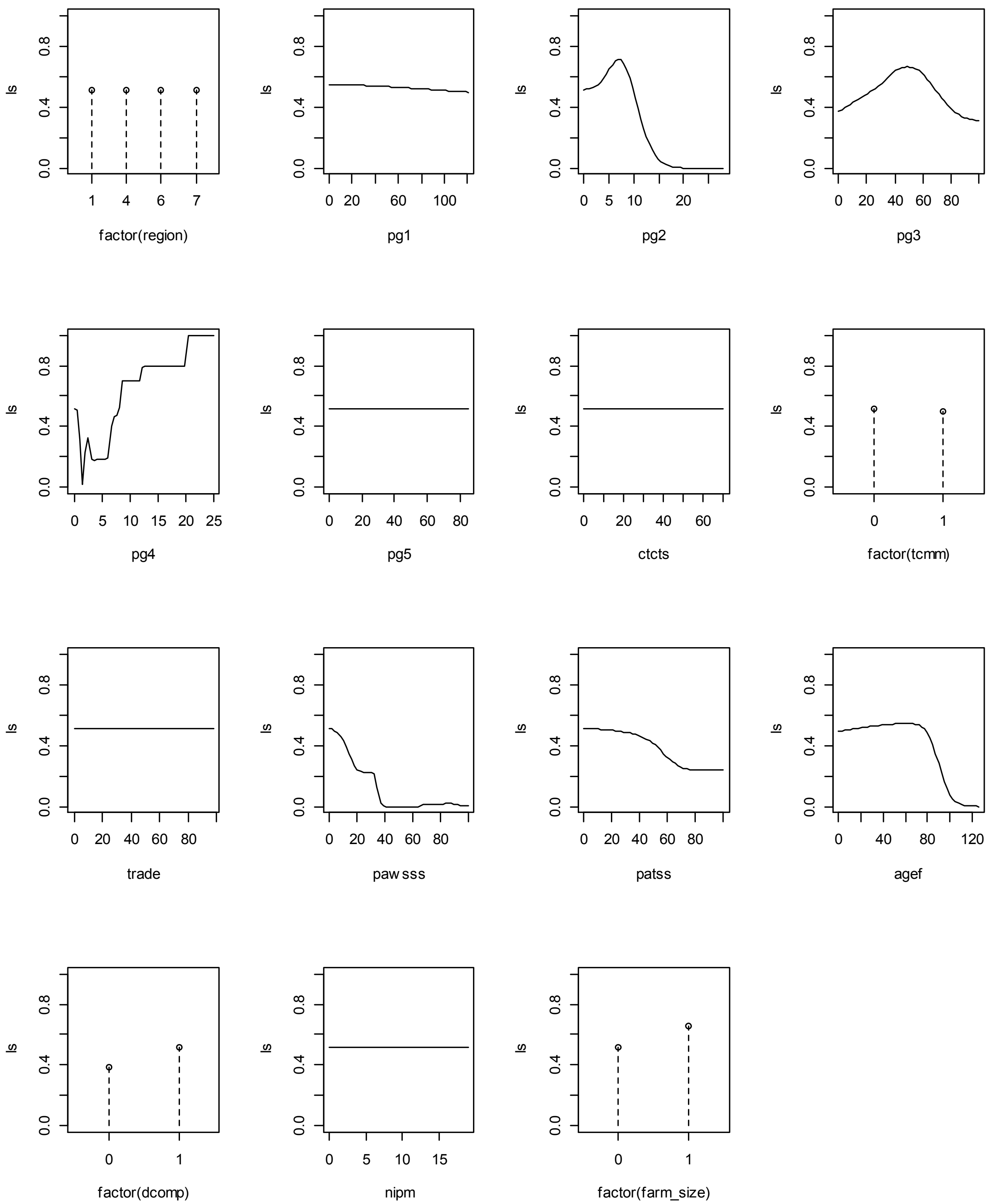

Figure 8: Partial regression plots obtained from nonparametric model of landscape firm (Cluster 2)

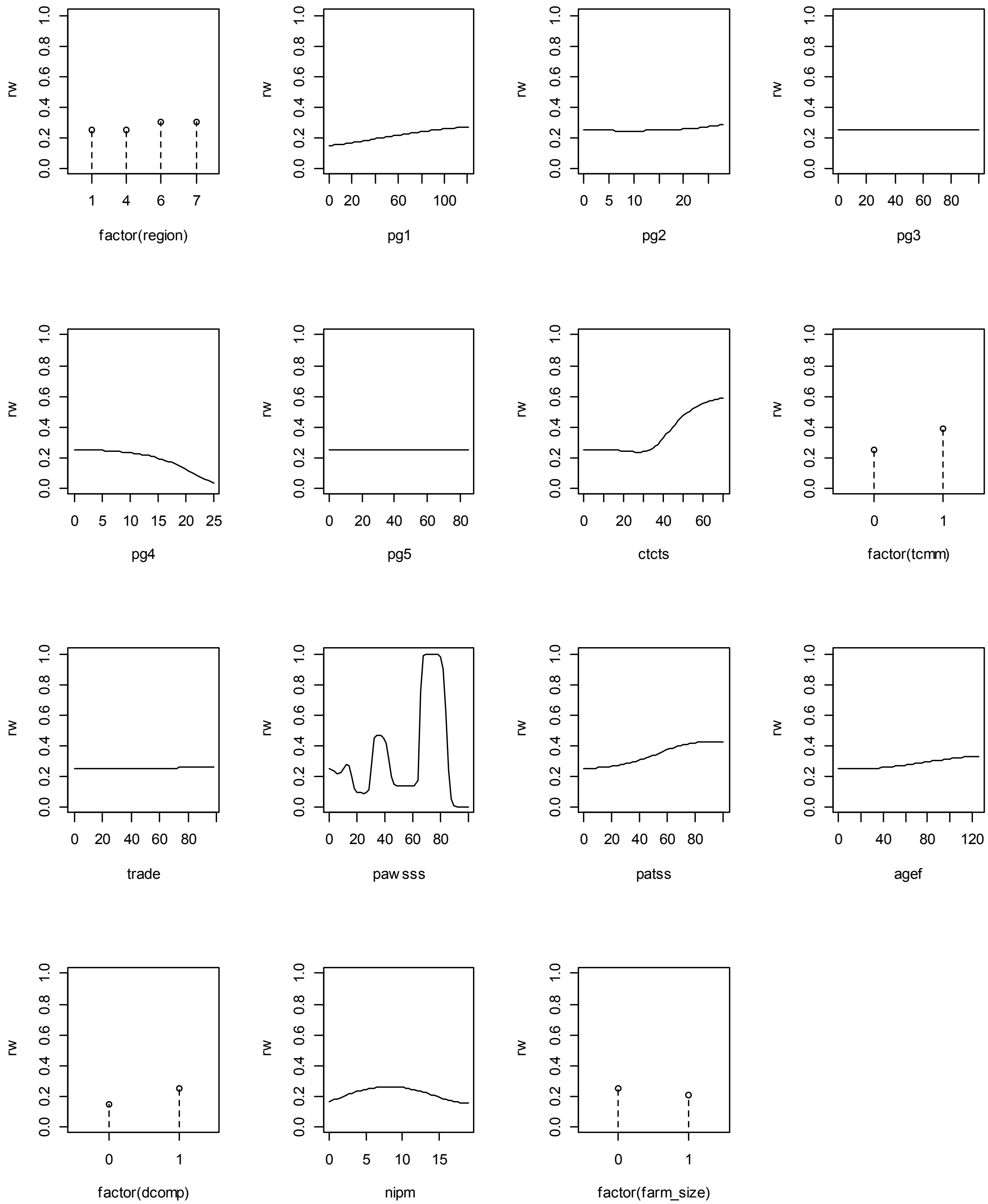

Figure 9: Partial regression plots obtained from nonparametric model of re-wholesalers (Cluster 2)

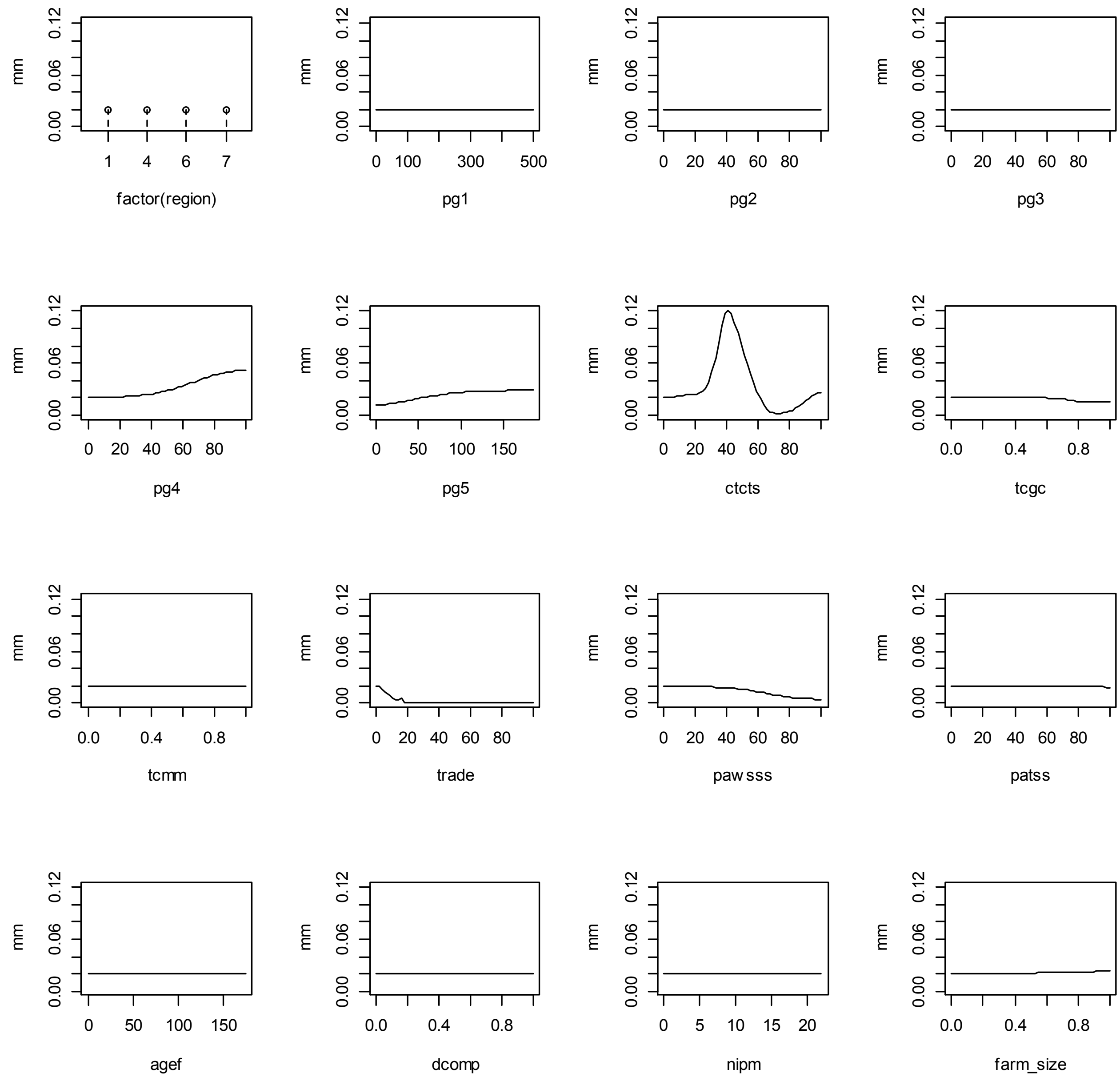

Figure 10: Partial regression plots obtained from nonparametric model of mass merchandiser (Cluster 3)

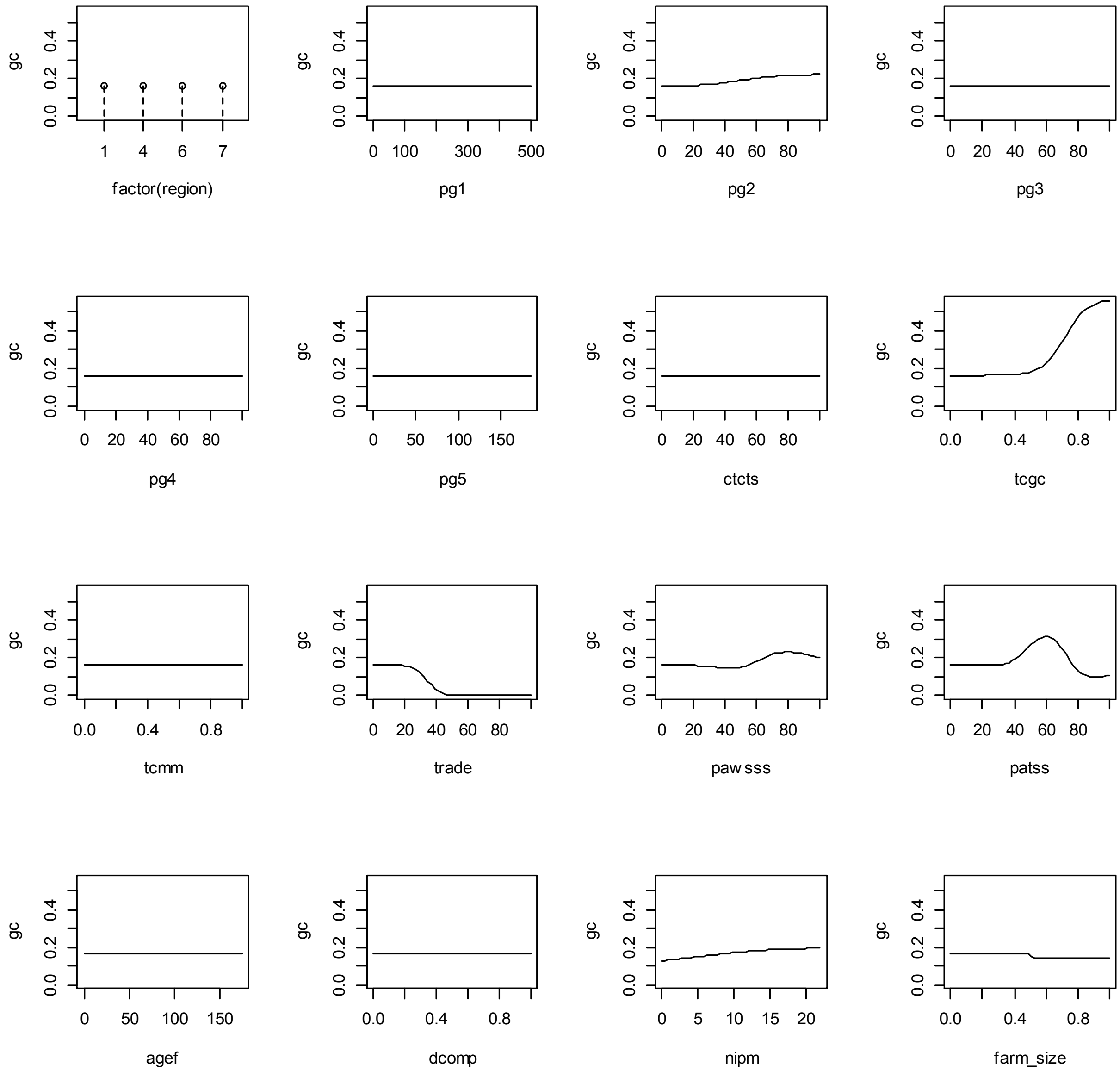

Figure 11: Partial regression plots obtained from nonparametric model of garden center (Cluster 3)

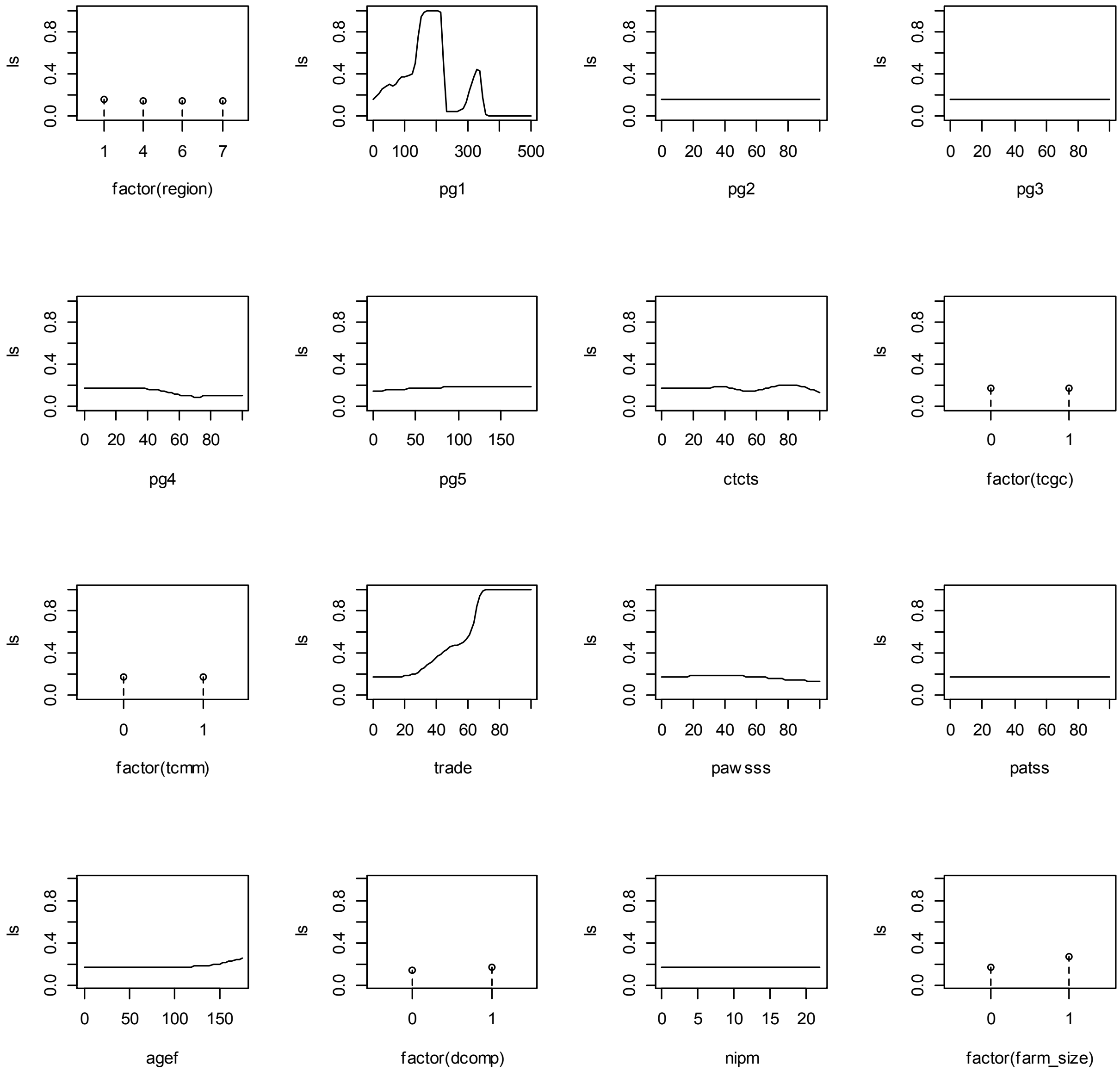

Figure 12: Partial regression plots obtained from nonparametric model of landscape firm (Cluster 3)

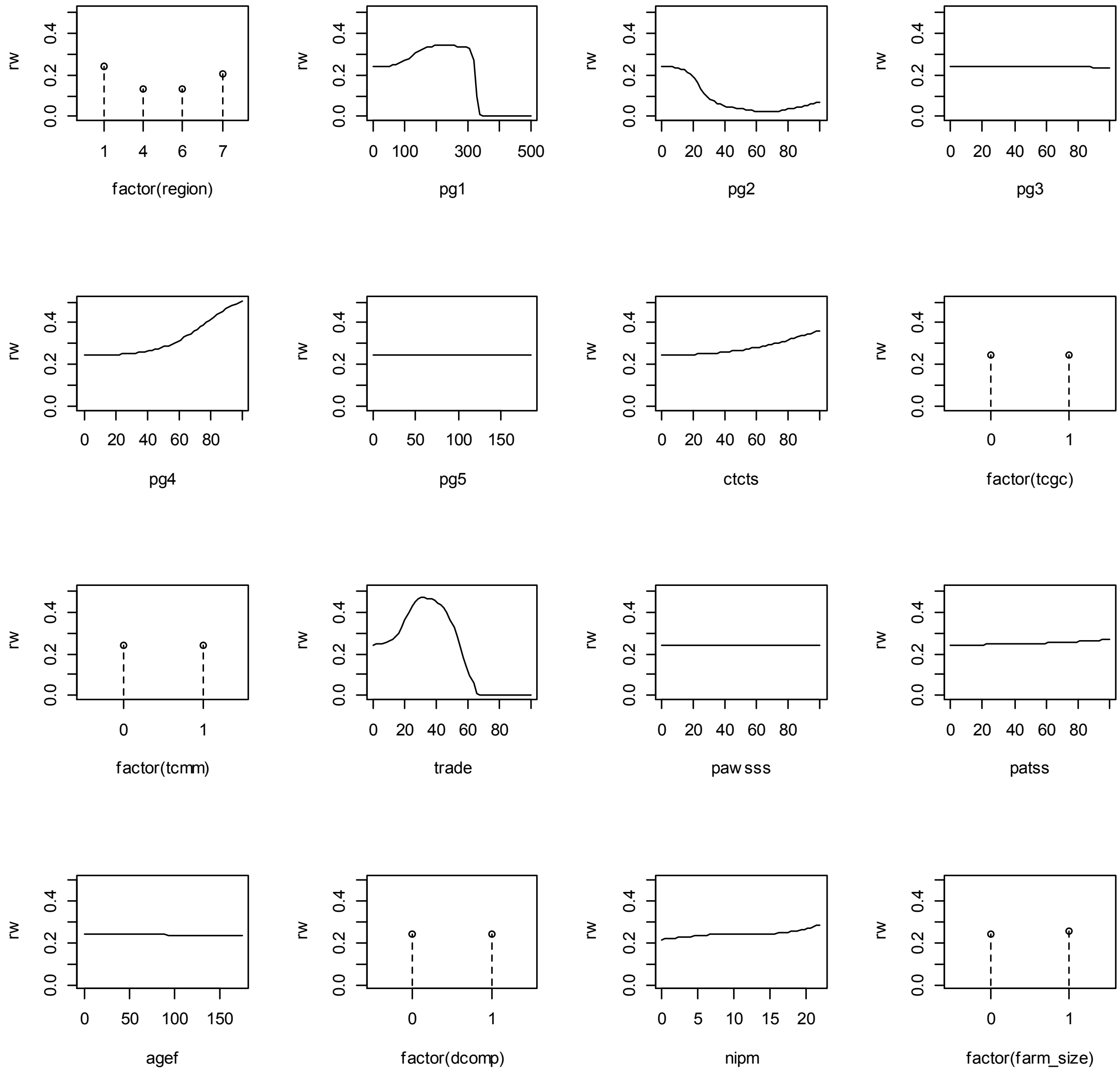

Figure 13: Partial regression plots obtained from nonparametric model of re-wholesalers (Cluster 3)

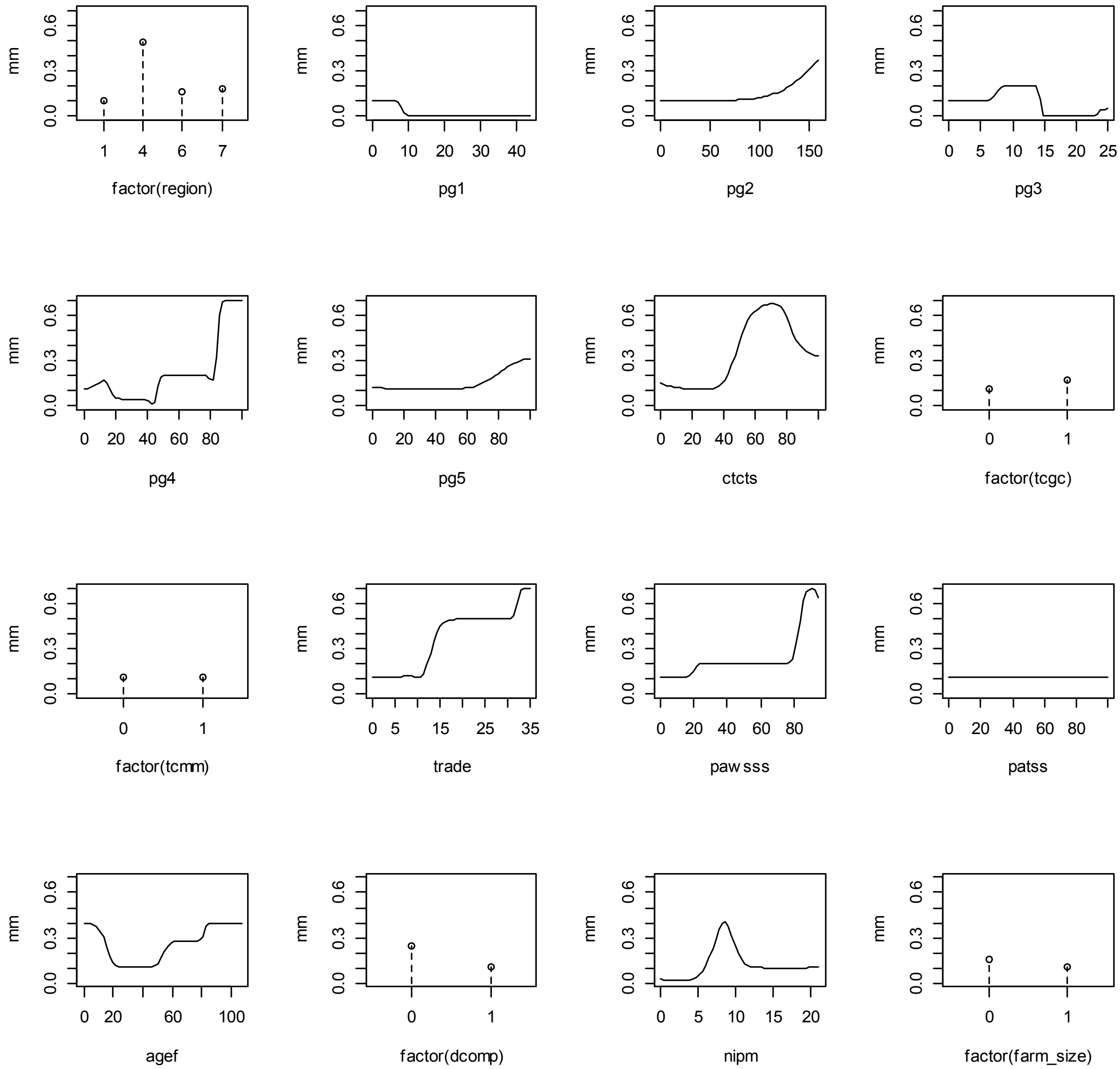

Figure 14: Partial regression plots obtained from nonparametric model of mass merchandiser (Cluster 4)

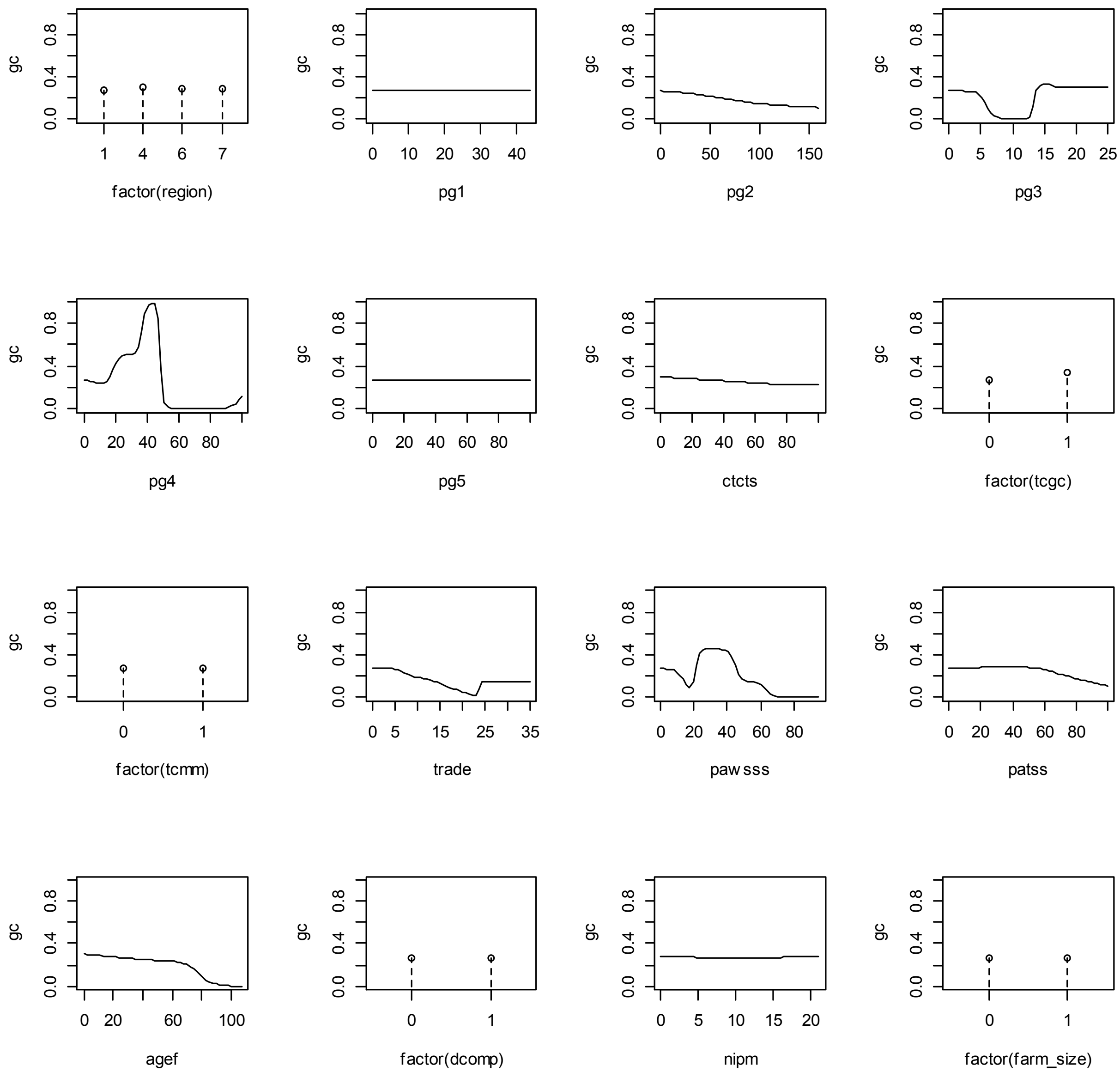

Figure 15: Partial regression plots obtained from nonparametric model of garden center (Cluster 4)

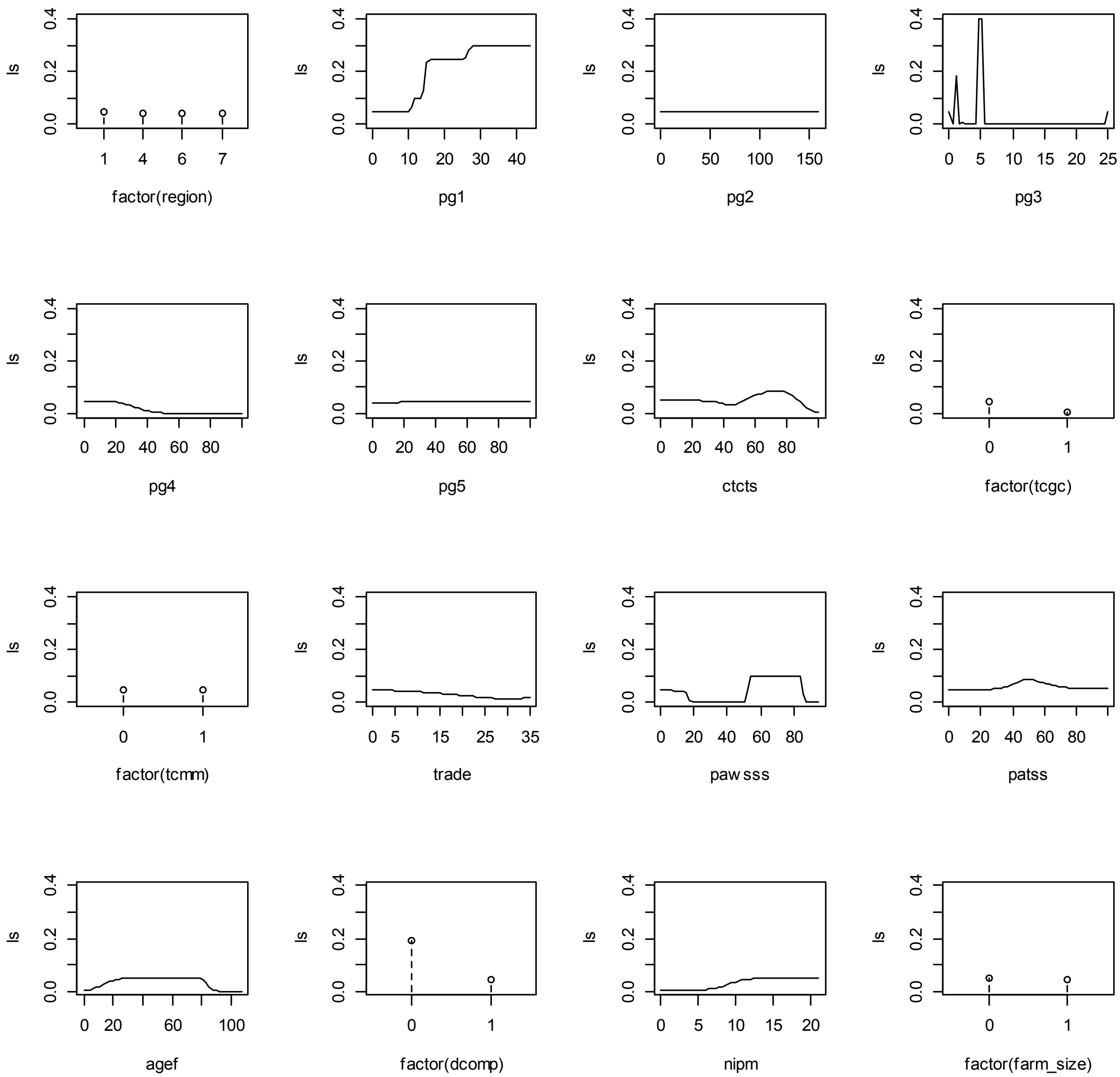

Figure 16: Partial regression plots obtained from nonparametric model of landscape firm (Cluster 4)

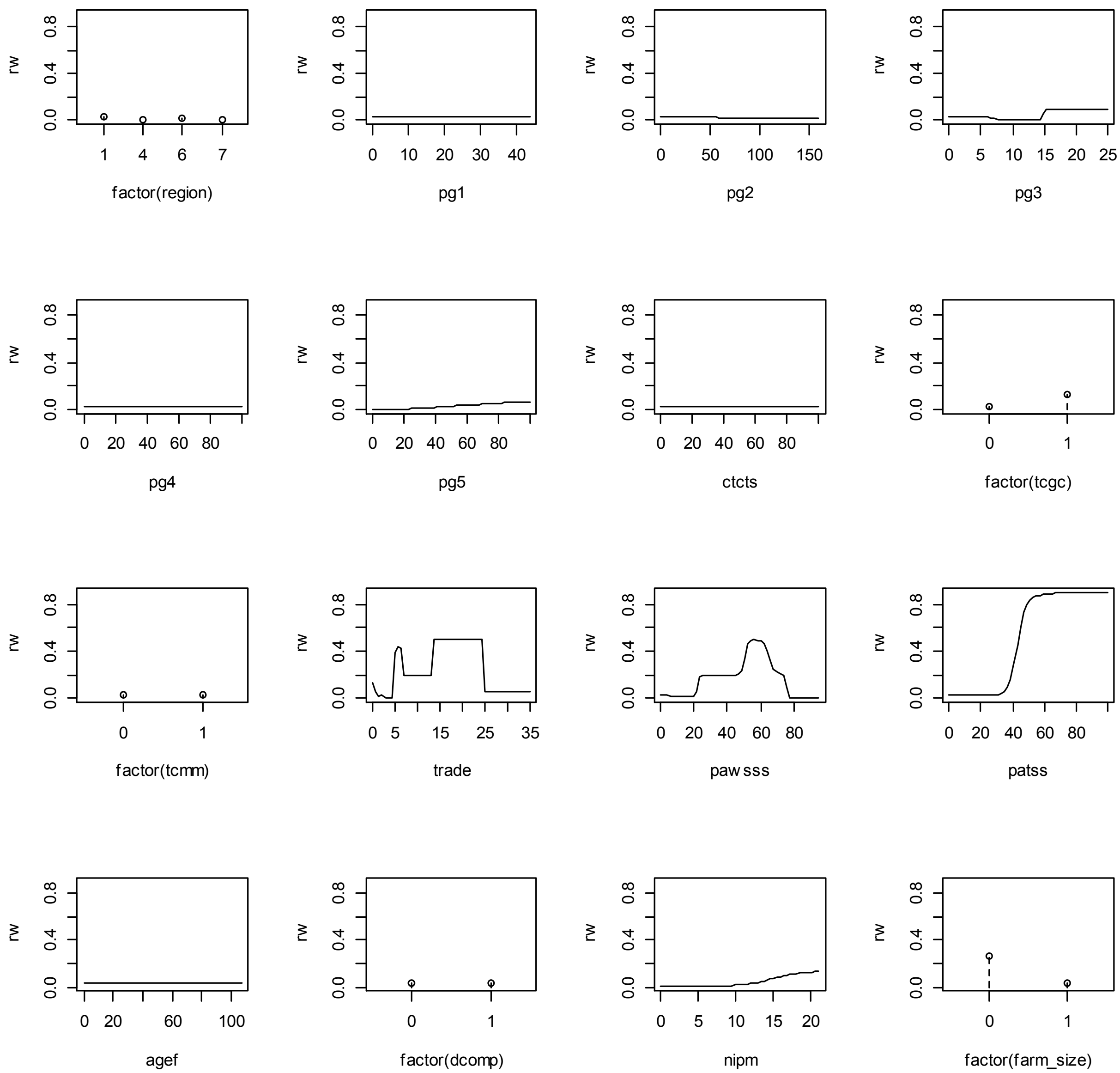

Figure 17: Partial regression plots obtained from nonparametric model of re-wholesaler (Cluster 4)

Table 1. Variable definition and summary statistics

Variables	Variable Definition	Cluster1		Cluster 2		Cluster 3		Cluster 4	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
<u>Market Channels</u>									
mm	Mass merchandiser	0.0027	0.0264	0.0053	0.0437	0.0277	0.1366	0.2507	0.2710
gc	Gardend Center	0.1183	0.2415	0.1316	0.2314	0.2001	0.3401	0.2455	0.2733
ls	Landscape firms	0.3895	0.4160	0.4259	0.3846	0.1884	0.3284	0.0927	0.1734
rw	Re-wholesalers	0.1488	0.2748	0.2053	0.2824	0.2038	0.3479	0.1970	0.2449
<u>Regions</u>									
Midwest	Equals 1 if Midwest, otherwise 0	0.3333	0.4718	0.0839	0.2776	0.1200	0.3250	0.1404	0.3504
Northeast	Equals 1 if Northeast, otherwise 0	0.3247	0.4687	0.0635	0.2441	0.2094	0.4070	0.0702	0.2577
Pacific	Equals 1 if Pacific, otherwise 0	0.0138	0.1168	0.0862	0.2809	0.2088	0.4066	0.2456	0.4343
Southeast	Equals 1 if Southeast, otherwise 0	0.0294	0.1690	0.5805	0.4940	0.2418	0.4283	0.2807	0.4533
Other	Equals 1 if Southeast, otherwise 0	0.2988	0.4581	0.1859	0.3895	0.2200	0.4144	0.2632	0.4443
<u>Plant Group</u>									
pg1	Sales of trees/shurbs($00,000)	68.8559	36.0595	72.9113	35.4249	16.1571	33.5701	4.2982	9.8034
pg2	Sales of bedding plants($00,000)	2.3252	7.0442	0.778	3.121	18.2896	31.5795	28.7895	41.7722
pg3	Sales of vines($00,000)	14.9776	20.1517	37.4803	36.1311	4.6100	14.1269	1.3509	4.6847
pg4	Sales of foliage($00,000)	0.0881	0.5431	0.3583	1.9883	6.1403	21.3119	15.3684	30.1316
pg5	Sales of other($00,000)	10.8263	15.4196	9.6564	16.267	55.9774	40.4580	47.7544	39.6778
<u>Contracted production</u>									
ctcts	Total sales under contract ($00,000)	0.6356	2.8480	3.5500	9.9010	12.8513	27.1200	40.5789	34.3204
tcgc	Contract to garden centers (1 if positive, otherwise 0)	-	-	-	-	0.0733	0.2608	0.4561	0.5025
tcmm	Contract to mass merchandisers (1 if positive, otherwise 0)	-	-	0.0045	0.0673	0.0062	0.0786	0.8772	0.3311
<u>Kinds of promotions</u>									
trade	Number of trade shows attended in 2008	1.0466	2.0962	1.8299	6.868	1.1827	3.5634	3.0877	5.2620
pawsss	Website promotion expenses ($000)	7.1628	19.3333	7.5295	20.2403	9.1681	22.9587	10.0526	21.3796
patss	Trade show promotion expenses ($000)	3.1027	11.0482	10.3091	22.3367	6.8421	20.5811	18.8246	30.8846
<u>Others</u>									
nipm	Number of integrated pest management used	6.3834	4.2212	6.6984	4.1388	7.7272	4.5863	11.7544	5.5236
agef	Firm age (2008 minus year established)	28.0104	21.5227	20.3447	18.2424	21.5196	20.0474	26.3158	19.8406
dcomp	Computer management aids	0.6546	0.4759	0.6304	0.4832	0.5979	0.4905	0.8772	0.3311
farm_size	Equal 1 if total sale >500K, 0 otherwise	0.2884	0.4534	0.39	0.4883	0.2474	0.4316	0.6491	0.4815

Note: variable *tcgc* and *tcmm* are invariant in cluster1, variable *tcgc* is invariant in cluster 2 and they are not use in model as well.

Table 2. Parametric results for fractional regression model

Variables	Cluster 1 mm	Cluster 1 gc	Cluster 1 ls	cluster2 mm	cluster2 gc	cluster2 ls
Constant	-21.53260	-2.91014	-0.29224	-35.09154***	-0.70813	1.28676***
	(0.998)	(0.336)	(0.800)	(0.000)	(0.224)	(0.001)
Midwest	15.64667	2.46805	0.82523	-9.27076***	0.27800	0.05205
	(0.999)	(0.408)	(0.446)	(0.000)	(0.580)	(0.901)
Northeast	14.53238	2.38935	0.52476	-8.84431***	-0.28594	-0.07109
	(0.999)	(0.421)	(0.622)	(0.000)	(0.499)	(0.846)
Pacific	-2.56987	3.38171	-0.56825	-16.83781***	0.74591*	-0.14752
	(1.000)	(0.282)	(0.727)	(0.000)	(0.086)	(0.661)
Other	15.08574	2.57051	0.51107	3.71975*	0.01496	-0.29684
	(0.999)	(0.386)	(0.630)	(0.063)	(0.964)	(0.262)
pg1	0.01249	-0.00006	0.00009	0.05212	-0.00107	-0.00705*
	(0.820)	(0.992)	(0.986)	(0.135)	(0.852)	(0.078)
pg2	0.00600	0.01434	0.06630*	0.08955	0.08348	0.07102
	(0.987)	(0.780)	(0.062)	(0.255)	(0.200)	(0.258)
pg3	-0.02298	-0.00569	0.00512	-0.05707***	0.00189	0.00573
	(0.703)	(0.570)	(0.474)	(0.003)	(0.689)	(0.155)
pg4	-9.85169	-0.41093	0.28408	0.16936	0.06103**	0.08199***
	(0.997)	(0.621)	(0.510)	(0.658)	(0.048)	(0.007)
pg5	0.03662	0.01370	0.00457	0.04778	0.01147	0.00569
	(0.567)	(0.237)	(0.639)	(0.118)	(0.179)	(0.412)
ctcts	-8.43405	-0.02877	-0.01767	-0.03577	-0.02123	-0.00975
	(0.988)	(0.673)	(0.724)	(0.389)	(0.112)	(0.348)
tcmm	-	-	-	-12.23338***	0.80327	0.18010
				(0.000)	(0.566)	(0.790)
trade	0.09487	0.04949	-0.07309	0.03720	-0.00337	0.00691
	(0.705)	(0.554)	(0.351)	(0.236)	(0.587)	(0.341)
pawsss	-0.00443	0.00702	0.00812	0.04104***	-0.01163	-0.00497
	(0.949)	(0.466)	(0.313)	(0.008)	(0.178)	(0.356)
patss	0.00381	-0.01552	-0.02118*	-0.09189*	0.00746	-0.00387
	(0.927)	(0.257)	(0.056)	(0.063)	(0.109)	(0.349)
agef	0.02065	0.01105	0.00995	0.00158	0.00190	0.00280
	(0.610)	(0.216)	(0.168)	(0.921)	(0.771)	(0.634)
dcomp	1.12568	-0.40825	0.07690	15.71868***	-0.15707	-0.08155
	(0.705)	(0.329)	(0.820)	(0.000)	(0.616)	(0.744)
nipm	-0.09419	0.01138	0.01649	-0.32514**	0.02532	-0.02796
	(0.735)	(0.812)	(0.662)	(0.028)	(0.418)	(0.257)
farm_size	0.20657	0.02749	0.24654	15.82764***	-0.16785	0.10826
	(0.925)	(0.948)	(0.458)	(0.000)	(0.479)	(0.590)

Table 2. Contd.

Variables	cluster3			Cluster 4		
	mm	gc	ls	mm	gc	ls
Constant	-1.81425**	-0.25563	0.87720**	1.38353	3.24864	-0.46370
	(0.031)	(0.382)	(0.011)	(0.549)	(0.157)	(0.869)
Midwest	0.73170	1.08169***	0.64363**	1.62898**	1.54646*	1.69280
	(0.213)	(0.000)	(0.033)	(0.047)	(0.054)	(0.102)
Northeast	0.75189	0.94548***	0.31581	1.36441	1.22360	-0.03494
	(0.106)	(0.000)	(0.164)	(0.182)	(0.441)	(0.976)
Pacific	0.04478	0.26230	-0.56651***	-0.81077	1.33142*	0.56443
	(0.915)	(0.194)	(0.005)	(0.286)	(0.076)	(0.521)
Other	0.69841*	0.84798***	0.25490	-0.15236	0.84172	3.08296***
	(0.083)	(0.000)	(0.193)	(0.834)	(0.287)	(0.000)
pg1	-0.02369***	0.00183	0.00350	-0.04089	-0.02227	0.12824***
	(0.009)	(0.396)	(0.221)	(0.266)	(0.589)	(0.002)
pg2	0.00698	0.01456***	0.00279	-0.01287	-0.02055**	-0.03028*
	(0.410)	(0.000)	(0.497)	(0.180)	(0.015)	(0.095)
pg3	0.00261	-0.01870**	-0.00670	0.06334	0.04127	-0.15063
	(0.871)	(0.013)	(0.106)	(0.373)	(0.450)	(0.162)
pg4	-0.00363	-0.00857**	-0.02574***	-0.01568	-0.01042	-0.04028**
	(0.659)	(0.013)	(0.000)	(0.196)	(0.415)	(0.030)
pg5	-0.00298	-0.00458*	-0.01167***	-0.05585***	-0.03747***	-0.04441**
	(0.684)	(0.052)	(0.000)	(0.000)	(0.001)	(0.025)
ctcts	-0.00963*	-0.01511***	-0.01662***	0.00846	-0.00095	-0.00117
	(0.088)	(0.000)	(0.000)	(0.301)	(0.896)	(0.882)
tcgc	0.26728	1.38579***	-0.10461	0.00983	0.31933	1.12116**
	(0.535)	(0.000)	(0.658)	(0.982)	(0.533)	(0.034)
tcmm	1.82997*	-0.10795	0.00688	1.60501	-1.09668	1.82392**
	(0.081)	(0.891)	(0.992)	(0.205)	(0.171)	(0.034)
trade	-0.03277	-0.01822	0.01090	0.20517**	0.04239	0.09175
	(0.410)	(0.334)	(0.395)	(0.010)	(0.420)	(0.155)
pawsss	-0.00540	0.00094	-0.00177	-0.03679***	-0.02805**	0.00629
	(0.409)	(0.759)	(0.553)	(0.000)	(0.044)	(0.585)
patss	0.00096	-0.00529*	-0.00890***	-0.03056***	-0.01606**	-0.00232
	(0.864)	(0.057)	(0.004)	(0.000)	(0.014)	(0.716)
agef	0.00351	-0.00122	-0.00281	-0.00448	0.00039	-0.04066*
	(0.618)	(0.757)	(0.494)	(0.744)	(0.982)	(0.099)
dcomp	0.29809	0.28614*	0.67823***	2.22287**	0.95909	2.34197**
	(0.357)	(0.086)	(0.000)	(0.016)	(0.260)	(0.033)
nipm	-0.05580*	-0.01026	-0.03786**	-0.01165	-0.02348	-0.14169**
	(0.087)	(0.548)	(0.022)	(0.807)	(0.627)	(0.027)
farm_size	0.81561***	-0.02210	0.38026**	-1.61630**	-0.41028	-0.75578
	(0.005)	(0.888)	(0.018)	(0.027)	(0.569)	(0.259)

Note: Value given in parenthesis are P-value. *, **, and *** represent parameters are significant at 0.10,0.05, and 0.01 level of significance.

Table 3. Marginal effect for cluster 1.

Variables	mm	gc	ls	rw
Midwest	0.52732***	-0.04353*	-0.29760***	-0.22051***
	(0.000)	(0.092)	(0.000)	(0.002)
Northeast	0.56622***	-0.03806*	-0.34870***	-0.19623***
	(0.000)	(0.094)	(0.000)	(0.003)
Pacific	-0.00407***	0.68703***	-0.50165***	-0.18182***
	(0.001)	(0.000)	(0.000)	(0.000)
Other	0.57844***	-0.02249	-0.36890***	-0.19519***
	(0.000)	(0.377)	(0.000)	(0.003)
pg1	0.00005	-0.00003	0.00000	-0.00002
	(0.522)	(0.962)	(0.997)	(0.977)
pg2	-0.00013	-0.00451	0.01351***	-0.00889**
	(0.809)	(0.203)	(0.001)	(0.022)
pg3	-0.00009	-0.00130	0.00176*	-0.00037
	(0.450)	(0.157)	(0.095)	(0.688)
pg4	-0.03632***	-0.07888*	0.12244***	-0.00883
	(0.006)	(0.078)	(0.008)	(0.814)
pg5	0.00015**	0.00148	-0.00039	-0.00119
	(0.014)	(0.129)	(0.770)	(0.299)
ctcts	-0.03269***	0.00438	0.01412	0.01103
	(0.005)	(0.641)	(0.393)	(0.251)
trade	0.00051	0.01421***	-0.02153***	0.00689
	(0.187)	(0.002)	(0.008)	(0.321)
pawsss	-0.00004	0.00020	0.00114	-0.00131
	(0.676)	(0.792)	(0.296)	(0.136)
patss	0.00008	-0.00013	-0.00325***	0.00331***
	(0.207)	(0.904)	(0.010)	(0.001)
agef	0.00005	0.00058	0.00111	-0.00174**
	(0.389)	(0.447)	(0.253)	(0.030)
dcomp	0.00345	-0.06884	0.05767	0.00797
	(0.196)	(0.107)	(0.247)	(0.854)
nipm	-0.00054	0.00007	0.00276	-0.00246
	(0.360)	(0.986)	(0.604)	(0.586)
farm_size	-0.00034	-0.02061	0.05207	-0.03210
	(0.924)	(0.527)	(0.213)	(0.366)

Note: Value given in parenthesis are P-value. *, **, and *** represent parameters are significant at 0.10,0.05, and 0.01 level of significance.

Table 4. Marginal effect for cluster 2

Variables	mm	gc	ls	rw
Midwest	-0.00608***	0.03520	-0.01109	-0.01803
	(0.000)	(0.562)	(0.890)	(0.803)
Northeast	-0.00611***	-0.02883	0.01104	0.02391
	(0.000)	(0.492)	(0.882)	(0.726)
Pacific	-0.01093***	0.13966**	-0.11110	-0.01763
	(0.000)	(0.047)	(0.109)	(0.762)
Other	0.03478	0.02377	-0.09152	0.03296
	(0.184)	(0.577)	(0.120)	(0.487)
pg1	0.00025	0.00046	-0.00171**	0.00100
	(0.102)	(0.491)	(0.044)	(0.152)
pg2	0.00016	0.00461	0.00905	-0.01383
	(0.498)	(0.343)	(0.374)	(0.224)
pg3	-0.00026**	-0.00023	0.00133*	-0.00084
	(0.011)	(0.655)	(0.094)	(0.233)
pg4	0.00049	0.00059	0.01348*	-0.01456***
	(0.767)	(0.897)	(0.064)	(0.004)
pg5	0.00019	0.00098	0.00019	-0.00136
	(0.144)	(0.299)	(0.891)	(0.274)
ctcts	-0.00012	-0.00192	-0.00030	0.00234
	(0.528)	(0.229)	(0.891)	(0.197)
tcmm	-0.00608***	0.11077	-0.04504	-0.05965
	(0.000)	(0.541)	(0.466)	(0.677)
trade	0.00015	-0.00109	0.00185	-0.00090
	(0.340)	(0.359)	(0.341)	(0.372)
pawsss	0.00020***	-0.00112	-0.00023	0.00115
	(0.000)	(0.290)	(0.849)	(0.234)
patss	-0.00039	0.00140**	-0.00136	0.00035
	(0.142)	(0.016)	(0.133)	(0.608)
agef	-0.00000	-0.00000	0.00049	-0.00048
	(0.981)	(0.998)	(0.661)	(0.643)
dcomp	0.01359***	-0.01538	-0.01315	0.01494
	(0.001)	(0.667)	(0.791)	(0.736)
nipm	-0.00136**	0.00609*	-0.00808	0.00335
	(0.045)	(0.081)	(0.100)	(0.439)
farm_size	0.06620***	-0.04140	0.02478	-0.02669
	(0.000)	(0.130)	(0.545)	(0.438)

Note: Value given in parenthesis are P-value. *, **, and *** represent parameters are significant at 0.10,0.05, and 0.01 level of significance.

Table 5. Marginal effect for cluster 3

Variables	mm	gc	ls	rw
Midwest	0.00421	0.14351***	-0.00398	-0.14374***
	(0.865)	(0.004)	(0.930)	(0.000)
Northeast	0.01330	0.15272***	-0.04967	-0.11636***
	(0.553)	(0.000)	(0.164)	(0.000)
Pacific	0.00433	0.10548***	-0.12973***	0.01991
	(0.812)	(0.005)	(0.000)	(0.520)
Other	0.01348	0.13801***	-0.04890	-0.10259***
	(0.480)	(0.000)	(0.127)	(0.000)
pg1	-0.00109***	0.00037	0.00085	-0.00012
	(0.010)	(0.425)	(0.127)	(0.763)
pg2	0.00005	0.00253***	-0.00100	-0.00158**
	(0.895)	(0.000)	(0.133)	(0.010)
pg3	0.00048	-0.00311**	0.00049	0.00213**
	(0.483)	(0.024)	(0.580)	(0.020)
pg4	0.00032	0.00093	-0.00429***	0.00304***
	(0.337)	(0.167)	(0.000)	(0.000)
pg5	0.00010	0.00030	-0.00186***	0.00145***
	(0.742)	(0.499)	(0.001)	(0.001)
ctcts	0.00004	-0.00121**	-0.00171***	0.00288***
	(0.869)	(0.015)	(0.001)	(0.000)
tcgc	-0.01388	0.31035***	-0.16443***	-0.13203***
	(0.302)	(0.000)	(0.000)	(0.000)
tcmm	0.17664	-0.07949	-0.04577	-0.05137
	(0.298)	(0.601)	(0.726)	(0.583)
trade	-0.00128	-0.00426	0.00448	0.00106
	(0.446)	(0.230)	(0.101)	(0.661)
pawsss	-0.00022	0.00045	-0.00038	0.00015
	(0.417)	(0.380)	(0.458)	(0.753)
patss	0.00024	-0.00018	-0.00128**	0.00121***
	(0.306)	(0.726)	(0.025)	(0.007)
agef	0.00021	-0.00001	-0.00049	0.00030
	(0.454)	(0.984)	(0.463)	(0.650)
dcomp	-0.00046	-0.01395	0.10232***	-0.08791***
	(0.972)	(0.617)	(0.000)	(0.001)
nipm	-0.00173	0.00258	-0.00585**	0.00500*
	(0.211)	(0.353)	(0.035)	(0.063)
farm_size	0.03015**	-0.05481**	0.06732**	-0.04325*
	(0.013)	(0.043)	(0.023)	(0.084)

Note: Value given in parenthesis are P-value. *, **, and *** represent parameters are significant at 0.10,0.05, and 0.01 level of significance.

Table 6. Marginal effect for cluster 4

Variables	mm	gc	ls	rw
Midwest	0.07498	0.07024	0.03871	-0.18393***
	(0.597)	(0.569)	(0.609)	(0.001)
Northeast	0.09877	0.08534	-0.05405	-0.13006
	(0.584)	(0.726)	(0.204)	(0.225)
Pacific	-0.21878**	0.30034**	0.01099	-0.09254
	(0.015)	(0.018)	(0.851)	(0.220)
Other	-0.16510*	0.03013	0.27205***	-0.13708**
	(0.070)	(0.789)	(0.000)	(0.019)
pg1	-0.00745*	-0.00387	0.01093***	0.00039
	(0.050)	(0.429)	(0.000)	(0.935)
pg2	0.00033	-0.00172**	-0.00141	0.00280**
	(0.771)	(0.046)	(0.243)	(0.017)
pg3	0.00991	0.00593	-0.01358*	-0.00226
	(0.227)	(0.477)	(0.077)	(0.768)
pg4	-0.00063	0.00045	-0.00233*	0.00251*
	(0.696)	(0.797)	(0.070)	(0.092)
pg5	-0.00453***	-0.00095	-0.00102	0.00650***
	(0.000)	(0.401)	(0.349)	(0.000)
ctcts	0.00137	-0.00077	-0.00026	-0.00034
	(0.266)	(0.481)	(0.601)	(0.702)
tcgc	-0.05170	0.02411	0.07970**	-0.05211
	(0.364)	(0.735)	(0.033)	(0.362)
tcmm	0.24347***	-0.37349***	0.10473***	0.02529
	(0.005)	(0.007)	(0.000)	(0.789)
trade	0.02556***	-0.01079	0.00090	-0.01568*
	(0.004)	(0.156)	(0.755)	(0.052)
pawsss	-0.00355***	-0.00209	0.00202***	0.00362***
	(0.007)	(0.286)	(0.004)	(0.005)
patss	-0.00333***	-0.00032	0.00094**	0.00271***
	(0.000)	(0.700)	(0.014)	(0.000)
agef	0.00024	0.00143	-0.00289*	0.00122
	(0.875)	(0.487)	(0.068)	(0.548)
dcomp	0.18533***	-0.01936	0.08271**	-0.24868*
	(0.006)	(0.869)	(0.020)	(0.065)
nipm	0.00333	0.00049	-0.00955**	0.00573
	(0.600)	(0.932)	(0.029)	(0.314)
farm_size	-0.19481**	0.07294	-0.00508	0.12948*
	(0.026)	(0.446)	(0.911)	(0.095)

Note: Value given in parenthesis are P-value. *, **, and *** represent parameters are significant at 0.10,0.05, and 0.01 level of significance.

Table 7. Model consistent test

Model	cluster 1	Cluster 2	Cluster 3	Cluster 4
mm	0.6949	0.0994	-0.7145	-0.0149
	(0.10)	(0.60)	(0.30)	(0.10)
gc	-0.6566***	-0.4675***	0.5867	1.5661***
	(0.00)	(0.00)	(0.10)	(0.00)
ls	0.6102	1.8182***	2.6805***	1.2100
	(0.10)	(0.00)	(0.00)	(0.10)
rw	-0.8809	0.5699***	-0.1829	2.5804
	(0.40)	(0.00)	0.20	(0.00)

Note: Value given in parenthesis are P-value. *, **, and *** represent parameters are significant at 0.10,0.05, and 0.01 level of significance.