# Appendix-1: Description of Propensity Score Matching Process

## Choice of Variables for Estimating for Propensity Matching Score Estimation

Since the original purpose of the study did not focus on the impact of the participation in government support program on planning accuracy, no firm characteristics were elicited for the time before the most recent investment in traceability. The following variables thus require a rationale for being selected for propensity score matching.

**Location:** It is reasonable to assume that location is not affected by participation and, within a rather large region, not varying over time. However it can affect participation, as regions differ in the types and amounts of resources they provide through support programs. We differentiated between being located in the northern region ‘Region Nord’ and the rest of the country.

**Firm size:** While firm size affects the decision to participate, e.g. through regional development objectives to support small and medium sized firms, it is certainly not fixed over time. However, based the relatively short time period of investigation, i.e. three years, we treated size as rather stable over time. Further, we would argue that it is unlikely that growth was triggered by participation, but rather through a pre-existing growth strategy that sought additional resources through participation. We measured size through the number of full time employees, because revenue brackets were not reported by a number of respondents.

**Involvement in primary production/harvest:** Also as part of regional development objectives, primary producers may be specifically targeted by some programs. Further, it is unlikely that program participation induced a processor to integrate upstream. Involvement in harvesting was elicited through a yes/no question.

**Quality management system certification:** Having been certified for certain quality management systems may reduce barriers to program participation and also affect realization of traceability benefits positively, if they already entail elements of traceability. We differentiated between firms that were certified according to one of ISO9000, International Food Standard (IFS) and British Retail Consortium (BRC) and those that were not certified according to either. All three standards have been available well before traceability and include additional requirements that go beyond mere product traceability, as required by the European Food Law.

**Strategic marketing:** Representing a strategic marketing focus, the decision to participate can be influenced by a firm’s branding effort. Although this can vary over time, it is not likely to change dramatically in a short period. Neither is it likely that participation in a government support program would change the branding focus. Instead, we would argue that a specific strategic marketing vision will enable the firm to determine what it would like to get out of participating in a support program and thus make the choice of an appropriate program easier and thus more likely. At the same, the presence of a clear strategic perspective should help achieve goals set out for traceability capacity and benefits. We differentiated three types of marketing: own brand, under license for another brand, and producing a no name product. Since there was no theoretical or empirical guidance on which one would be more appropriate for inclusion in the model we used each one for a separate model.

## Results of Propensity Score Matching Model Estimation: Goodness of Fit of Three Models

We chose a logistic regression model to estimate the propensity scores (probability of participation in a government support program) for three models that differed in regards to the strategic marketing variable included. The goodness of fit indicators for the entire model and the regression coefficients with their significance levels are shown in the table below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | **Model 1**  **(own brand)** | **Model 2 (licence)** | **Model 3**  **(no name)** |
| **Observations** |  | 56 | 56 | 56 |
| **Log likelihood** |  | 32.904 | 31.486 | 32.235 |
| **Cox and Snell R2** |  | 0.255 | 0.273 | 0.264 |
| **Nagelkerke R2** |  | 0.435 | 0.467 | 0.450 |
| **Constant** | β | -0.811 | -0.227 | -0.541 |
|  | S.E. | 1.074 | 1.159 | 1.065 |
|  | p(β≠0) | 0.450 | 0.844 | 0.612 |
| **regio\_NORD** | β | -1.402 | -0.911 | -1.169 |
|  | S.E. | 1.055 | 1.133 | 1.042 |
|  | p(β≠0) | 0.184 | 0.421 | 0.262 |
| **primary\_producer** | β | 3.498 | 3.418 | 3.871 |
|  | S.E. | 1.074 | 1.088 | 1.244 |
|  | p(β≠0) | 0.001 | 0.002 | 0.002 |
| **LAB\_FULL\_11\_25** | β | 1.540 | 1.503 | 1.616 |
|  | S.E. | 1.024 | 1.033 | 1.050 |
|  | p(β≠0) | 0.133 | 0.146 | 0.124 |
| **ISO9000\_IFS\_BRC** | β | 0.183 | 0.964 | 0.237 |
|  | S.E. | 0.997 | 1.740 | 0.933 |
|  | p(β≠0) | 0.854 | 0.580 | 0.800 |
| **Marketing Strategy** | β | 0.059 | -1.280 | -0.966 |
|  | S.E. | 1.709 | 1.110 | 1.226 |
|  | p(β≠0) | 0.972 | 0.249 | 0.431 |

β= Regression coefficient, S.E. = Standard error of the coefficient; p (β≠0) = Significance level.

Overall, the model fit is similar for all three models and satisfactory so that the assumption is justified that the included variables represent factors that impacted the participation decision. The regression coefficients, their standard errors and their inferred error probabilities, do not vary much either across models. Vertical integration involving primary production is highly significant (p < 0.01), while size, as measured by full time workforce is the only other variable that comes close to conventional significance levels (p < 0.15). All other variables are not significant. Since none of the models can be clearly identified as producing a significantly better fit than the other two, the matching process will select only those observations into the control group that have been selected in each of the three models. The matching algorithm ‘five nearest neighbors’ (5NN) with replacement was chosen to account for the small sample size and, as a consequence of that, for large stretches of the support region of the propensity score range being without observations. The table below shows the propensity scores for the support recipients and the matched observations or control group members for each of the three models, as well as the propensity score (as average of the three models’ scores) included in analysis.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Brand | Group a | Licence | Group a | No\_name | Group a | All\_three | Group a |
| 0.901 | R | 0.925 | R | 0.936 | R | 0.921 | R |
| 0.692 | R | 0.799 | R | 0.632 | R | 0.708 | R |
| 0.620 | R | 0.657 | R | 0.695 | R | 0.657 | R |
| 0.620 | R | 0.657 | R | 0.695 | R | 0.657 | R |
| 0.638 | R | 0.718 | R | 0.575 | R | 0.644 | R |
| 0.287 | R | 0.163 | R | 0.414 | R | 0.288 | R |
| 0.217 | R | 0.305 | R | 0.232 | R | 0.251 | R |
| 0.064 | R | 0.123 | R | 0.035 | R | 0.074 | R |
| 0.056 | R | 0.032 | R | 0.057 | R | 0.048 | R |
| 0.662 | M | 0.749 | M | 0.523 | M | 0.645 | M |
| 0.692 | M | 0.558 | M | 0.632 | M | 0.628 | M |
| 0.325 | M | 0.233 | M | 0.473 | M | 0.344 | M |
| 0.275 | M | 0.379 | M | 0.212 | M | 0.289 | M |
| 0.287 | M | 0.163 | M | 0.212 | M | 0.221 | M |
| 0.217 | M | 0.123 | M | 0.232 | M | 0.191 | M |
| 0.217 | M | 0.123 | M | 0.232 | M | 0.191 | M |
| 0.217 | M | 0.123 | M | 0.232 | M | 0.191 | M |
| 0.056 | M | 0.032 | M | 0.057 | M | 0.048 | M |
| 0.187 | M | 0.220 | M | 0.193 |  | 0.200 | RM |
| 0.187 | M | 0.220 | M | 0.193 |  | 0.200 | RM |
| 0.187 | M | 0.082 |  | 0.193 |  | 0.154 | RM |
| 0.064 | M | 0.123 | M | 0.086 |  | 0.091 | RM |
| 0.064 | M | 0.123 | M | 0.086 |  | 0.091 | RM |
| 0.056 | M | 0.095 |  | 0.057 | M | 0.069 | RM |
| 0.056 | M | 0.095 |  | 0.057 | M | 0.069 | RM |
| 0.056 | M | 0.095 |  | 0.057 | M | 0.069 | RM |
| 0.056 | M | 0.095 |  | 0.057 | M | 0.069 | RM |
| 0.054 |  | 0.083 |  | 0.069 |  | 0.069 | RM |
| 0.054 |  | 0.083 |  | 0.069 |  | 0.069 | RM |
| 0.064 | M | 0.043 |  | 0.086 |  | 0.064 | RM |
| 0.064 | M | 0.043 |  | 0.086 |  | 0.064 | RM |
| 0.051 |  | 0.083 |  | 0.027 |  | 0.054 | RM |
| 0.047 |  | 0.063 |  | 0.045 | 1 | 0.052 | RM |
| 0.047 |  | 0.063 |  | 0.045 | 1 | 0.052 | RM |
| 0.054 |  | 0.028 | M | 0.069 |  | 0.050 | RM |
| 0.064 | M | 0.043 |  | 0.035 | M | 0.047 | RM |
| 0.047 |  | 0.063 |  | 0.018 |  | 0.043 | RM |
| 0.047 |  | 0.063 |  | 0.018 |  | 0.043 | RM |
| 0.047 |  | 0.021 |  | 0.045 | M | 0.038 | RM |
| 0.047 |  | 0.021 |  | 0.045 | M | 0.038 | RM |
| 0.047 |  | 0.021 |  | 0.045 | M | 0.038 | RM |
| 0.056 | M | 0.032 | M | 0.022 |  | 0.037 | RM |
| 0.047 |  | 0.021 |  | 0.018 |  | 0.029 | RM |
| 0.014 |  | 0.033 | M | 0.007 |  | 0.018 | RM |
| 0.014 |  | 0.011 |  | 0.018 |  | 0.014 | RM |
| 0.014 |  | 0.011 |  | 0.018 |  | 0.014 | RM |
| 0.014 |  | 0.011 |  | 0.018 |  | 0.014 | RM |
| 0.014 |  | 0.011 |  | 0.018 |  | 0.014 | RM |
| 0.014 |  | 0.011 |  | 0.018 |  | 0.014 | RM |
| 0.012 |  | 0.021 |  | 0.006 |  | 0.013 | RM |
| 0.012 |  | 0.021 |  | 0.006 |  | 0.013 | RM |
| 0.011 |  | 0.021 |  | 0.006 |  | 0.013 | RM |
| 0.011 |  | 0.021 |  | 0.006 |  | 0.013 | RM |
| 0.012 |  | 0.007 |  | 0.006 |  | 0.008 | RM |
| 0.012 |  | 0.007 |  | 0.006 |  | 0.008 | RM |

R = Support recipients; M = Matched group, per model; C = Control group; RM = Remaining sample

**Appendix-2:** Paired Samples T-Test of Expected and Actual Implementation Cost Importance Scores

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Specific cost categories** | | **Received support (n=9)** | **Matched/Control (n=9)** | **Remaining sample (n=37)** |
| Purchase new equipment, hardware and software | Expected | 21.7 | 30.2 | 35.9 |
| Actual | 23.0 | 28.0 | 33.9 |
| *t-value* | *0.660* | *1.300* | *0.954* |
| *% Distribution a* | *22/56/22* | *0/44/56* | *19/46/35* |
| Certification, audit and external consultants | Expected | 16.7 | 26.9 | 17.0 |
| Actual | 19.3 | 24.9 | 18.9 |
| *t-value* | *1.619* | *0.715* | *1.717* |
| *% Distribution a* | *33/56/11* | *33/33/33* | *32/51/16* |
| Workforce, clerical and managerial staff time | Expected | 32.8 | 17.4 | 21.2 |
| Actual | 31.3 | 22.5 | 23.2 |
| *t-value* | *1.276* | *1.112* | *1.564* |
| *% Distribution a* | *11/56/33* | *44/33/22* | *41/38/22* |
| Training | Expected | 13.6 | 12.4 | 13.6 |
| Actual | 12.0 | 11.7 | 11.0 |
| *t-value* | *0.881* | *0.598* | *1.774* |
| *% Distribution a* | *11/56/33* | *33/33/33* | *16/38/46* |
| Material | Expected | 15.3 | 13.0 | 12.2 |
| Actual | 14.3 | 13.0 | 13.0 |
| *t-value* | *1.516* | *0.028* | *1.042* |
| *% Distribution a* | *0/67/33* | *22/44/33* | *30/54/16* |

aProportion (%) of firms with actual importance score higher than/equal to/lower than expected importance score.