An Econometric Study of the Factors Influencing Participation in Urban and Community Forestry Programs in the United States

Brandon W. Wall, Thomas J. Straka, and Stephen E. Miller

Abstract. Public participation in urban and community forestry (U&CF) programs in the contiguous 48 states of the United States can be explained by several different factors using econometric methods. The state’s percent of working population, income level, percent of forested land, dominant political affiliation, state government expenditures on education, and the number of communities participating in U&CF programs help explain the public participation rates in the programs. These factors accounted for 52% ($R^2 = 0.5218$) of the variability in public participation in U&CF programs in 2003. Knowing the factors that influence public participation in U&CF programs is vital to the success of the program and can assist federal, state, and local decision-makers in planning efforts.

Key Words. Econometric model; participation; public participation; urban and community forestry; urban forestry.

Urban and community forestry (U&CF) programs gained a major stimulus in 1991 when the USDA Forest Service began a new national program to improve the awareness, management, conservation, and care of the tree resources in and around communities. The program is funded by the USDA Forest Service and administered by state forestry or natural resources agencies. It is designed to “enhance the technical skills of individuals involved in the planning, development and maintenance of urban and community forests; improve understanding of the value of preserving existing tree cover; provide educational programs and technical assistance to state and local organizations; and establish demonstration projects to illustrate the value of forest and tree cover” (South Carolina Forestry Commission 2005; Urban Forestry South Expo 2005).

Over three-fourths of Americans live in urban areas and this program recognizes the importance of urban and community forests to the average American (Straka et al. 2005). The current program originated with the amendment of the Cooperative Forest Assistance Act of 1978 to authorize financial, technical, and related assistance to state forestry agencies in support of cooperative efforts in U&CF (Cubbage et al. 1993). From 1960 to 1997, the nation’s urban area increased from 10.2 to 26.7 million ha (25–66 million ac) (Vesterby and Krupa 2001). The trend is not expected to abate. Urbanization is linked to population growth, and by 2050, another 16.2 million ha (40 million ac) is expected to be converted to urban and other development uses (Alig et al. 2003).

The urban forest produces substantial benefits. It has a beneficial impact on the physical environment, increasing property value, and reducing energy costs and water runoff. It also improves the social environment by improving community health and well-being, consumer behavior, and recreational opportunities, reducing noise levels, and creating buffer zones. Urban forests provide environmental benefits by decreasing soil erosion and improving wildlife habitat and air and water quality.

Financial assistance includes U&CF planning, recreational development, air and water quality improvement programs, stormwater management, urban wildlife management, and economic, urban development and conservation management plans. In addition, technical assistance and information dissemination are program components. Typical program recipients are local governments, policymakers and elected officials, builders and developers, civic and community groups, neighborhood associations, nonprofit groups, local businesses, and urban forest councils (USDA Forest Service 2004; Straka et al. 2005).

The South Carolina Forestry Commission has a representative U&CF program (South Carolina Forestry Commission 2005). South Carolina divides eligible proposals into three categories: community forestry program development or improvement (e.g., development of tree ordinances, tree inventory software, student internships, and greenspace inventories), information and education (Arbor Day activities, work-
shops, and public service materials), and nonprofit administration (volunteer training, workshops, and temporary staffing).

Public volunteerism and active citizen participation are the keys to program success (Wates 2000). Participation can be measured in statistics like participating communities, number of days of volunteer assistance, and number of Tree City USA communities. These data are available on a state-by-state basis and show participation levels vary tremendously across the country, even when adjusted for area and population (USDA Forest Service 2004). For example, number of volunteer days per capita varies by over 300% between adjacent states that one would expect to be similar (for example, between Alabama and Mississippi or Indiana and Illinois). Therefore, knowing and understanding the factors that explain this volunteerism and citizen participation will provide invaluable information for U&CF decision-makers and planners and will improve the effectiveness of these programs (Thindwa and Reuben 2003).

The purpose of this study was to provide a perspective of the characteristics that are associated with participation in U&CF programs. Econometric methods were used to identify various factors that are correlated with participation. This insight enables state, local, and nonprofit organizations to identify what does and does not motivate individuals, communities, county and municipal decision-makers, various industries/professionals, and others to participate in U&CF programs or to embrace and adopt the principles and approaches that improve urban and community forests.

**STUDY METHODS**

Econometrics is the branch of economics concerned with the quantitative analysis of economic and social behavior. It involves the specification of a model that forecasts or explains this behavior. Most often, the model is based on regression analysis. We developed an econometric or regression model to explain participation in U&CF programs. The model is cross-sectional (i.e., an analysis based on one or more variables collected at a single point in time as opposed to looking at the variables over time). That is, this study was performed to explain the factors which determine U&CF program participation, not to predict future participation in the program.

Data collection was attempted for the contiguous 48 American states for the year 2003. Data for six states were incomplete; this lowered the number of viable observations to 42 ($n = 42$). The variable being explained (dependent or left-hand side variable) in the model is denoted as participation in an individual state. It represents the number of days of public volunteer assistance and participation in the program in a specific state (USDA Forest Service 2004).

Table 1 lists the type of factors that were considered for inclusion in the model. They include educational variables (like percent high school graduates), economic variables (like per capital income), and demographic variables (like percent urban population). Sources included the US Census Bureau (2004, 2005), USDA Forest Service (2005), USDA Forest Service Forest Inventory Analysis (2005), Internet sites such as 50states.com (2005), Economagic.com (2005), Demographia.com (2005), National Center for Education Statistics (2003), and the USDA Forest Service U&CF accomplishment report for fiscal year 2003 (USDA Forest Service 2004).

All variables were computed on either a percentage of population or total land area or a per capita basis. This was done to “even the playing field” because with cross-sectional data, it is not appropriate to compare large states such as Texas and California to smaller states such as Delaware and Rhode Island.

The statistical package SAS was used in conjunction with STATA to run several regressions to model volunteer assistance days as a function of the independent variables. The most effective model in explaining participation was chosen, evaluated, and interpreted. SAS was used to run the regressions and STATA was used to evaluate the models for multicollinearity (high correlation between independent variables), heteroscedasticity (nonconstant variance), and autocorrelation (correlation between successive errors between fitted and actual values). Economic and social data tend to be more interrelated than biologic and physical data. For example, educational level and per capital income are likely correlated. If both these variables are chosen for the final model, a basic assumption of regression models will be violated. Econometric modeling is particularly prone to these three possible assumption violations; thus, each potential problem is analyzed in detail.

**RESULTS**

Econometric modeling involves a series of regression models that test the validity of economic or behavioral theory as to why something happens like the factors that explain participation. The variables first thought to be significant in explaining public volunteer assistance days in U&CF programs were Communities, Income, Forestland, WorkingPop, and High School. These variable definitions and the rationale for including them in the model are discussed subsequently. We discuss the expected impact (positive or negative) of each independent variable before running the regression model. Hopefully, the variable coefficients (positive or negative) will conform to these expectations or further model development would be required.

Participation was defined as the number of days of volunteer assistance in U&CF programs per million people in an individual state. This was the first of many potential models that were evaluated. The econometric process involves running a series of regressions until an optimal explanatory model is found. Each model is evaluated for fit and regression assumption violations until the final version is identified.
The variable expected to be most significant in explaining participation was the number of communities per million people in each state with U&CF programs. Community participation should positively impact program participation (Robinson 1993). A larger number of participating communities should indicate there are more diverse programs and projects in the state that will be able to engage more diverse groups of people and provide a better chance of the program reaching willing participants.

States with a higher percentage of affluent (above average per capital disposable income) and more educated people (the percent of people in the state 25 years of age or older with a high school diploma) were believed likely to have higher participation rates. Educated people may be more aware of the importance of trees and urban forests and should be more active in participating in projects that enhance urban areas. An educational U&CF program is more likely to attract them as a result of their higher “thirst” for knowledge. States with a higher average per capita income were also expected to have higher participation rates because these people may have more leisure time and opportunity to participate in U&CF programs.

The amount of forested land in the state was expected to be negatively correlated with public participation in U&CF programs. People in states such as Nebraska, Kansas, and the Dakotas may have a greater desire for urban forests and therefore may participate in the programs more than people in states with high amounts of forested land such as Maine, West Virginia, and New Hampshire. An earlier study (Straka et al. 2005) found that most participants in U&CF programs were working population aged. A final dependent variable, WorkingPop (percent of state population between ages 18 and 65), was included in the model because the authors proposed that state age structure will affect participation rates. WorkingPop was expected to have a positive impact on participation.

This first regression produced an $R^2$ of 0.3830 and the adjusted $R^2$ of 0.2973. This relatively large difference was expected as a result of the low number of observations ($n = 42$), low degrees of freedom, and the high number of parameters ($[k-1] = 5$) relative to the number of observations. The two significant variables in the model were Communities and Forestland. Communities was statistically significant at the alpha = 0.05 level ($P = 0.0010$) and Forestland was significant at the alpha = 0.10 level ($P = 0.0635$). Income and HighSchool were very insignificant ($P$ value = 0.4443 and 0.7251, respectively).

Econometrics is used to evaluate regression models that economic theory suggests should explain causal relationships; thus, we describe the model before the regression was run, even going so far as to predict the sign of the coefficient based on theory. Once the regression results are known, the results and theory are evaluated and a (hopefully) stronger model is proposed in terms of more significant variables and a higher $R^2$. One might ask, “Why not just use stepwise regression and identify the ‘best variables’?” The idea is to use economic theory to improve the explanatory power of the model not to merely “mine” the variables.

Theory says that education should be positively related to participation (Verba et al. 1995), but all education variables evaluated were very insignificant. A related variable, per capita state government expenditure on education, did seem to have an impact. Dropping HighSchool from the model and adding the percent of government expenditures on education (%Education$) increased the $R^2$ and retained an education variable in the model.

Other income variables were evaluated such as disposable income per capita and average household income. Although neither of these variables were significant by themselves in the regression, when disposable income was divided by household income, the new variable became significant. Income per capita was dropped from the model and replaced with the new income variable Income* defined as disposable income as a percentage of household income. This was ef-

---

**Table 1. Variables considered for inclusion in the econometric model.**

<table>
<thead>
<tr>
<th>Education data</th>
<th>Demographic and other data</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Percent of people over 25 with a high school degree</td>
<td>• Percent of urban land</td>
</tr>
<tr>
<td>• Percent of people over 25 with a bachelor’s degree or higher</td>
<td>• Percent of urban population</td>
</tr>
<tr>
<td>• Elementary and secondary school enrollment</td>
<td>• Race (percent minority)</td>
</tr>
<tr>
<td>• Average state SAT scores</td>
<td>• Sex (percent male)</td>
</tr>
<tr>
<td>• Percent enrollment in high school</td>
<td>• Number of ISA Certified Arborists in the state</td>
</tr>
<tr>
<td>• Number of degree granting institutions and branches</td>
<td>• Region (South, Northeast, Midwest, and so on)</td>
</tr>
</tbody>
</table>
| • Average pupil/teacher ratio in elementary schools | • Number of certified Tree City USA communities

*Tree City USA is sponsored by National Arbor Day in cooperation with the USDA Forest Service and the National Association of State Foresters. This certification is a means of providing direction, technical assistance, public attention, and national recognition for urban and community forestry programs (The National Arbor Day Foundation 2006).
fective in increasing the $R^2$ while keeping an economic variable in the model.

Political affiliation of the state was also evaluated. A dummy variable with 1 = Republican and 0 = Democrat was added to the model. The variable represented the direction the state voted in the 2004 presidential election. One could argue this is a “conservative” or “liberal” variable or that one party is more “progovernment” and community action-oriented.

Another common way to reevaluate the data are to add higher order forms of variables such as squares, cubes, logs, square roots, and reciprocals. For the most part, the best and simplest form for this model was linear. However, a simple graph of variable shape showed that the dependent variable, participation, was exponentially shaped and the independent variable Communities was slightly exponential in shape. Transforming this independent variable by squaring it better matched participation and alone raised the $R^2$ from 0.4467 to 0.5218 and the adjusted $R^2$ from 0.3519 to 0.4398. All other variables showed very little of a nonlinear trend when graphed. Other functional transformations or models such as a semilog or log-linear model also proved to be poor fits. The improved regression equation is shown subsequently and the variables are described in Table 2. The improved form of regression with standard errors and t-values below coefficients is shown below.

$$\text{Participation}_i = 79698 + 0.23875 \text{Communities}^2_i - 652.7581 \text{WorkingPop}_i - 56.9062 \text{Forestland}_i$$

$se = (26462) \quad (0.0436) \quad (343.585) \quad (25.870)$

$\text{t} = (3.01) \quad (5.47) \quad (-1.90) \quad (-2.20)$

$$419.4388 \text{Income}^*_i + 3536.4228 \text{Political}_i - 364.7847 \%\text{EducationS}_i + e_i$$

$$(174.709) \quad (1489.95) \quad (152.634)$$

$$(−2.40) \quad (2.37) \quad (−2.39)$$

$$R^2 = 0.5218$$

$$\text{Adj-R}^2 = 0.4398$$

In the improved regression, the $R^2$ increased to 0.5218, and the adjusted $R^2$ increased to 0.4398. The new model satisfied several criteria. It maximized the $R^2$ and adjusted $R^2$ while keeping the model as simple as possible. All of the independent variables were significant at the alpha = 0.10 level. All previously determined theoretical criteria were satisfied with an education variable, economic variable, age structure variable, participating communities variable, and forested land variable.

The impact of political affiliation on public participation produced an interesting result. Holding all other variables constant, the states that voted Republican in the 2004 presidential election have on average, 3,536 more days of public volunteer assistance per million people than states that voted for the Democratic candidate. This could be explained by something as simple as “conservative” versus “liberal,” but it is likely much more complex than that. It could be the result of wealth (amount of free time), concern for the environment, attitude toward the government, or any of a multitude of factors that interact with that variable.

It is also interesting to note that the coefficients for WorkingPop with that Forestland, Income,* and %Education$ were all negative. Holding the other variables constant, if the percent of people in the state between the ages of 18 and 65 increased by one percentage point, then on average, the days of volunteer assistance will decrease by approximately 653 days per million people. This possibly indicates that throughout the United States, middle-aged people, those who must work and provide for themselves participate less in U&CF programs. It is logical to think that many projects, especially those that require outdoor activity such as planting trees in parks, engage children and adolescents. Also, older people may have more time and therefore may participate more in technical, educational, or administrative projects than do middle-aged people.

Forestland (percent of the state that is forested) seems to serve as an “equalizing” or “weighting” variable in the model constant, if the percentage of forested land in the state increased by one percentage point, then on average, the number of days of volunteer assistance will decrease by approximately 57 days per million people. This agrees with the previous hypothesis that forest land and participation will be negatively correlated. States with fewer forests tend to participate more in U&CF programs than states with more forests. People in states that have less forest land may place more importance on their urban forests and trees than people who live in states that are more heavily forested.

Income* was slightly more complicated to interpret because it was defined as disposable income as a percentage of household income. As the proportion of disposable income to...
household income increases by one percentage point, while holding the other variables in the model constant, participation in U&CF programs decreases on average by 419 days per million people. Does increased disposable income provide other recreational opportunities that compete with less costly U&CF program recreational opportunities?

It is notable that the coefficient for the percent of government expenditures on education was also negative. Holding the other variables in the model constant, if the percentage of state government expenditures on education increases by one percentage point, then on average, the number of days of volunteer assistance decreases by approximately 365 days per million people. Federal funds for the U&CF program are typically matched one to one by state grants. Because both education and available matching funds are tied to state budgets, the relationship between the two may account for negative correlation.

Communities was the most significant variable (P = 0.0048). The fact that the variable is a square means that the impact of one additional community is larger as the total number of participating communities increases. There is some sort of interaction between communities or “bandwagon effect” that causes much greater volunteerism as the total number of participating communities increases. Holding the other variables in the model constant, if the number of communities participating in U&CF programs increases from nine to ten, then on average, the number of days of volunteer assistance will increase by five days per million people; if it increases from 99 to 100 communities, the number of days of volunteer assistance will increase by 48 days per million people; and if it increases from 999 to 1000 communities, the number of days of volunteer assistance will increase by 477 days per million people. This was expected because more participating communities may result in the programs reaching more people willing to volunteer their time.

After the model was specified, the three major pitfalls of econometrics were addressed: multicollinearity, autocorrelation, and heteroscedasticity. Models based on social and economic data are prone to these problems that can produce erroneous t ratios and affect test of hypotheses and confidence intervals.

Multicollinearity is defined as highly correlated right-hand side variables. It was not suspected to be a problem in this model for the reasons 1) the R² are of the regression was not unusually high, 2) all of the right-hand side variables were statistically significant at the alpha = 0.10 level, 3) auxiliary regressions showed that three variables (Communities, WorkingPop, and Forestland) were not significantly explained by the other independent variables, 4) coefficients had the predicted signs, 5) the correlation coefficients were not very high for any of the independent variables, and 6) the variance inflation factor (VIF), calculated using STATA, was less than ten for all of the variables.

Autocorrelation can be defined as the correlation between members of observations ordered in time (like in time-series data) or space (like in cross-sectional data). The classic linear regression model assumes that correlation does not exist in the disturbances terms. Autocorrelation means that the disturbance term relating to any observation is related to or influenced by the disturbance term relating to any other observation (Gujarati 1999). Because autocorrelation typically exists in time-series data, it was not suspected in this model.

An important assumption in regression models is that the disturbance terms entering the population regression function are homoscedasticity; that is, they all have the same variance (Gujarati 1999). The result of heteroscedasticity is inefficient ordinary least squares estimators. Heteroscedasticity is prevalent in models using cross-sectional data (like this model). A graph of the residuals on the fitted value of y (ŷ) showed a strong negative relationship and this suggests heteroscedasticity.

A formal test for heteroscedasticity reaffirmed the suspicion. The Cook-Weisberg test for heteroscedasticity reported a χ² test statistic of 110.76 with a P value of < 0.0001 and the White test for heteroscedasticity reported a χ² test statistic of 41.45 and a P value of 0.0279. The null hypothesis for both tests (Ho: constant variance) was rejected, indicating heteroscedasticity.

To solve the problem, the “Robust” option in STATA was used. This command makes the standard errors asymptotically efficient. The coefficients remained the same, whereas...
the standard errors of Communities, WorkingPop, and %Education decreased and the standard errors of the constant, Forestland, Income, and Political increased. All variables became less significant, but all were still significant at the alpha = 0.15 level. Table 3 presents the descriptive statistics for the variables included in the final model. The final form of regression with standard errors and t-values below coefficients is shown below:

\[
\text{Participation}_i = 79698 + 0.23875 \text{Communities}^2_i - 652.7581 \text{WorkingPop}_i - 56.9062 \text{Forestland}_i \\
\text{Income}_i = 443.95 \text{Political}_i - 364.7847 \%\text{Education}_i + \epsilon_i
\]

State expenditures on education are highly correlated with educational quality and economic growth at the state level (Sims 2004). This study suggests that the public participation in U&CF programs in a highly developed country is higher in states with better education, stronger economies, lower percent of forested land, and in states which engage more communities to participate in U&CF projects or programs.

**CONCLUSION**

The purpose of this study was to explain public volunteer participation in U&CF programs in the contiguous 48 states. The model presented here indicates what factors are likely to explain public participation in U&CF programs. The insight provided by this study may be useful to the USDA Forest Service as well as state and local forestry agencies in improving participation rates in the programs. Because the number of participating communities is the most significant variable, the program directors should strive to engage more communities to encourage more public participation. Also, the directors of the programs may want to concentrate efforts on the “middle-aged” population, as the proportion of middle-aged people in the state increased, participation declined. Conversely, special attention may be paid to those younger than 18 and older than 65 who appear to participate in the program. The program directors may also want to focus on promoting U&CF programs more in heavily forested states, “blue states” (states that lean toward the Democratic party), and states that allocate a larger percent of government expenditures to education, because these states tend to participate less in the programs.

There should be other, more detailed studies done to evaluate the effect of state demographic, social, cultural, and geographic characteristics on public participation in U&CF programs. Other factors could be examined such as the psychological factors that drive people to volunteer or what allows for their “free” time to dedicate to a cause such as urban forestry.

Table 3. Simple descriptive statistics for the variables included in the final model.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participation</td>
<td>42</td>
<td>1744.38</td>
<td>4469.08</td>
<td>38.03400000</td>
<td>28627.04</td>
</tr>
<tr>
<td>Communities$^2$</td>
<td>42</td>
<td>7363.73</td>
<td>13165.90</td>
<td>2.72000000</td>
<td>64815.01</td>
</tr>
<tr>
<td>WorkingPop</td>
<td>42</td>
<td>62.0857143</td>
<td>1.5576047</td>
<td>59.60000000</td>
<td>68.4000000</td>
</tr>
<tr>
<td>Forestland</td>
<td>42</td>
<td>44.7080952</td>
<td>2.2919189</td>
<td>2.95000000</td>
<td>89.6000000</td>
</tr>
<tr>
<td>Income$^*$</td>
<td>42</td>
<td>64.8329762</td>
<td>3.9783032</td>
<td>56.44300000</td>
<td>73.3700000</td>
</tr>
<tr>
<td>Political</td>
<td>42</td>
<td>0.5952381</td>
<td>0.4967958</td>
<td>0.00000000</td>
<td>1.0000000</td>
</tr>
<tr>
<td>%Education$^*$</td>
<td>42</td>
<td>31.6670881</td>
<td>5.5399067</td>
<td>20.70660000</td>
<td>47.2396000</td>
</tr>
</tbody>
</table>

©2006 International Society of Arboriculture
Also, adjustments or improvements could be made to some of the variables in the model. For example, the income variable could be manipulated to represent an economic variable, which is more applicable to the study and easier to interpret and education could be evaluated more extensively. Theory suggests that state education level should be tied to U&CF Program participation (Sims 2004), but this study found no evidence of this. In the future, plans can be made to improve the factors that decrease participation and exploit the factors that increase it. Improved planning and better focus can increase public participation in U&CF programs and lead to the enhancement and enrichment of America’s urban areas.

This study provides valuable insights into the factors that influence U&CF programs. Planners and policymakers will find these results have significant implications to the success of these programs.

LITERATURE CITED

Brandon W. Wall
Graduate Research Assistant
Department of Forestry and Natural Resources
Clemson University
Box 340317
Clemson, SC 29634-0317, U.S.
wallback@clemson.edu

Thomas J. Straka (corresponding author)
Professor
Department of Forestry and Natural Resources
Clemson University
Box 340317
Clemson, SC 29634-0317, U.S.
tstraka@clemson.edu

Resumen. La participación pública en los programas forestales urbanos y comunales (U&CF, por sus siglas en inglés) en 48 estados en los Estados Unidos puede ser explicada por diferentes factores usando métodos econométricos. El por ciento estatal de población trabajando, nivel de ingresos, por ciento de terreno forestados, afiliación política dominante, gastos en educación del gobierno estatal y el número de comunidades participantes en los programas U&CF ayudan a explicar la tasa de participación pública en los programas. Estos factores respondieron por el 52% ($R^2 = 0.5218$) de la variabilidad en participación pública en programas U&CF en el 2003. El conocimiento de los factores con influencia en la participación pública en los programas U&CF es vital para su éxito y puede asistir a los tomadores de decisiones federal, estatal y local en los esfuerzos de planeación.