The Profit Puzzle: Legal Ownership and Dependence as Predictors of Acceptable Quality Performance in Contract Settings

Julia L. Carboni, MPA
PhD Candidate

School of Government and Public Policy
The University of Arizona
315 Social Science Building
PO Box 210027
Tucson, AZ 85721

jcarboni@email.arizona.edu

Draft- Please do not cite without permission from the author.

The United States government has a long history of contracting at all levels of government to compensate for what the government cannot or does not produce. However, pushes for efficiency and privatization in the 1980s led to a significant increase in contracting out government functions to the private sector, especially at the federal level (Pack 1987; Private Sector Survey on Cost Control 1984). This Hollow State trend (Milward and Provan 2000; Milward, Provan and Else 1993) continued into the 1990s, when there was increased bipartisan support for smaller, more efficient government and incorporation of market principles into government (Cooper 2003; Domberger 1994; Gore 1993; Osborne & Gaebler 1992). We continue to see increased government contracting for services (Cooper 2003; Green 1996).

Nonprofits are theoretically preferable trading partners for government funded human services because they are thought to be less prone to opportunism than their for-profit counterparts and will produce better quality services (Hansmann 1980; Light 2000; Wise 1990), particularly where outcomes are difficult to measure. As expected, nonprofits dominate the government funded human services field, but government increasingly relies on for-profits to provide services as well (Salamon 1993; GAO Report 2002). This brings to light questions of whether organizational form is a good indicator of quality. The guiding research questions in this paper are: does ownership form influence contract performance on quality dimensions, and under what conditions does ownership form matter most to produce quality performance? Based on social exchange theory, I predict that dependence at the dyadic and structural levels will also impact the likelihood of acceptable contract performance.

I study the Florida Department of Juvenile Justice (FDJJ) Residential Services Program. The FDJJ contracts with approximately 25 organizations to produce approximately 80 residential programs for juvenile delinquents each year. Programs do not engage in joint service production. Each program is systematically evaluated on quality annually. Though this paper focuses on a geographically isolated empirical example, these types of contracting arrangements (with a mix of for-profit and nonprofit
providers) are common in federal, state and local government settings. Examples include private prisons, foster care administration, housing voucher programs and others.

I have an unusual data set where many of the factors thought to influence contract performance are constant. The FDJJ invests heavily in systematic monitoring and managing of contracts and employs a professional staff with expertise in contract administration and program evaluation. All contracts are a fixed fee structure and sanctions are attached to performance and are systematically initiated and enforced. In addition, because of the nature of the population being served, providers cannot choose their clients, eliminating the potential for creaming the best and easiest to serve clients; clients are assigned to facilities by the FDJJ based on treatment needs and other factors deemed important by the FDJJ.

To evaluate what other factors might impact quality, I perform two types of analysis- a binary logistic regression and crisp set qualitative comparative analysis (QCA). The outcome variable of interest is acceptable quality performance, meaning that a program performs at a level that does not engender sanctions. Binary logistic regression is a maximum likelihood model estimate that estimates the probability that certain causal variables will predict an outcome of acceptable quality performance. QCA uses Boolean algebra to determine which causal conditions or sets of causal conditions may lead to the specified binary outcome. While both types of analysis focus on which causes may lead to a specified binary outcome, the methods are very different. I will compare the results of these two methods to illustrate why QCA may be of interest to public management researchers.

Organizational Form and Quality

Theoretical arguments from organization theory and economics predict that nonprofits will perform better than for-profit counterparts on quality dimensions where outcomes are difficult to measure and the possibility for opportunistic behavior exists. Arguments for nonprofit superiority are largely based on the notion that nonprofits do not need to pursue profit maximization and can thus
focus on delivering high quality services (Frumkin 2002). These arguments stem from the nondistribution constraint, which prohibits nonprofits from distributing profits to owners (Weisbrod 1988; Hansmann 1980). In addition to the nondistribution constraint, nonprofits enjoy tax exempt status, presumably allowing them to reinvest that otherwise foregone income into services. Nonprofits are also eligible to participate in certain government programs not available to their for-profit counterparts that may further reduce operating costs depending on the type of service being provided (e.g., National School Lunch Program). These benefits are thought to make nonprofits less focused on cost efficiency than their for-profit counterparts so they will be freer to use resources to support mission driven goals. Conversely, it is assumed that for-profits will behave opportunistically in situations with information asymmetry. In these cases, for-profits may reduce service quality in order to minimize costs and maximize profits (Hansmann 1980). Based on existing theory, I offer the following hypothesis:

\[ H1: \text{Nonprofits will be more likely to produce acceptable quality performance for contracts than their for-profit counterparts.} \]

Though there is considerable theoretical research on differences between nonprofits and for-profits, there is little empirical evidence on differences between nonprofit and for-profit contract performance. In public management, research on the effects of organizational form tends to focus on distinctions between public and private organizations rather than between types of private organizations (for-profit and nonprofit). Empirical studies that do compare nonprofit and for-profit organizations operating in the same sector tend to focus on the healthcare industry, but do not typically address government contracts directly. For an exception, see Heinrich (2000). Additionally, public management research on government contracting for human services typically focuses on relationships between government and nonprofits (e.g., Gazley 2008; Van Slyke 2007). The dearth of literature on how organizational form contributes to contract performance is likely due to the difficulty of measuring
outcomes for services, particularly where observability is low and multiple goals exist (Heinrich 2000; Kanter & Summers 1987; Salamon 1993; Weisbrod 1989).

**Other Organizational Attributes**

Though theory suggests organizational form may have an impact on quality, I argue that other factors more directly related to program goals may impact the likelihood that contractors will have acceptable program quality\(^1\). In particular, the quality of personnel may have a direct impact on contract performance, particularly along quality dimensions. Hart et. al. (1997) argue that quality of personnel may impact quality. Specifically, where contracts are incomplete in terms of required personnel qualifications, private agents may skimp on costs by hiring minimally qualified (or unqualified) staff to fulfill contractual obligations\(^2\). I offer the following hypothesis about personnel quality:

\[
H2: \text{The presence of quality personnel will improve the likelihood of contract performance.}
\]

By quality personnel, I mean those personnel who are educated and trained to deal with client populations and/or have prior experience working with the population. If staff are not adequately prepared to work with a particular client population, program quality may suffer (Hart et. al. 1997).

**Dependence as a Predictor of Performance**

To date, most literature on contract performance focuses on dyadic interactions between government and contractors with the contract as the unit of analysis. Factors thought to influence contract performance once contracts have been awarded include contract monitoring capacity; financial incentives; means for resolving disputes; level of task uncertainty; asset specificity and development of relational contracts (e.g. - Brown and Potoski 2003; Heinrich 2000; Fernandez 2007).

---

\(^1\) Anecdotal evidence suggests that some frontline workers are not even aware of ownership form.

\(^2\) Based on the arguments for nonprofit superiority, we might assume that nonprofits are more likely to focus on hiring well trained personnel. My particular data set does not show an association between organizational form and personnel training, but this is a topic worth exploring in greater detail.
While dyadic interactions are important in predicting contract performance, we must also incorporate the structure of exchange into analysis. The existing literature on contract performance is largely devoid of considerations of structure once contracts have been awarded. However, where multiple contractors are tasked with producing the same or similar service, structure of the supply market may impact contractor behavior. Granovetter (1985) argues that failure to take into account structure when considering individual behavior produces theoretically incomplete models. Actors “do not behave or decide as atoms outside a social context, nor do they adhere slavishly to a script written for them by the particular intersection of social categories they happen to occupy. Their attempts at purposive action are instead embedded in concrete, ongoing systems of social relations” (Granovetter 1985: 55). In essence, structure and agency both matter when considering actor behavior. Both should be incorporated into models that predict contract performance.

To consider how structure might impact contractor behavior, I turn to the social exchange theory (SET) literature. In SET, actor behavior is theorized to be largely a product of relation based power and dependence (Emersion 1972). Power and dependence are relational rather than actor attributes; power in a relationship is based on the amount of dependence actors have on each other. As one actor becomes more dependent on their exchange partner, that exchange partner becomes more powerful. In the absence of alternative exchange partners, dependent actors will be more likely to cede to the wishes of their more powerful partners in order to maintain beneficial exchange relationships (some utility derived) into the future. Demands of powerful partners may be counter to the preferences and goals of the dependent partner. In these relations, powerful actors may also be able to induce further performance of dependent partners through the use of rewards and sanctions (Molm 1997). Power and dependence are products of both dyadic and structural processes. Dyadic dependence concerns individual exchanges between actors while structural dependence is related to the presence of alternatives in the overarching structure of the exchange network (Emerson 1972).
Salamon (1981) argued that where third party governance involves contracting out services that are not well defined, government also cedes some discretion and authority to contractors. He asserts “instead of a hierarchical relationship between the federal government and its agents, therefore, what exists in practice is a far more complex bargaining relationship in which the federal agency often has the weaker hand” (260). Additionally, Cooper (2003) argues that government must manage horizontally (in collaborative relationships with contractors). These conceptions of contracting relationships lend credence to the idea that government is an interdependent partner in a collaborative process rather than a principal maintaining strict control through hierarchical arrangements.

Recent work on contract management and performance has focused on relational contracts to deal with situations of incomplete contracts where government cannot rely solely on hierarchical control. Relational contracts are characterized by ongoing, cooperative and flexible agreements rather than the discrete, arms length transactions described in principal-agent theory. Obligations and sanctions are typically diffuse, non-specific, and non-measurable (Macneil 1974, 1978; Baker, Gibbons, and Murphy 2002; Fernandez 2007). Contracted parties communicate frequently and engage in collaborative processes to address changing circumstances and gaps not found in the original agreement and use alternative means for dispute resolution (DeHoog 1990; Smith 1996; Sclar 2000; Fernandez 2007). In the absence of high contract specificity - collaborative processes, trust and shared norms facilitate adaptation to changing conditions.

I argue that dependence may serve as a substitute for development of relational contracts because it may shape the rational and calculative strategies of agents. Again, I use the social exchange conception of dependence which is the inverse of power. Where agents (contractors) are dependent on principals (government), they may be more likely to perform according to government expectations, particularly where goals are not aligned (i.e. – principal demands quality while the agent’s rational strategy is to pursue cost efficiency). Performance based on dependence is externally motivated.
Meeting principal demands becomes a rational/calculative strategy for maintaining the contracting relationship. If dependence reliably predicts agent behavior, government may be able to use resources more strategically to develop time and labor intensive relational contracts with agents not predicted to perform.

At the dyadic level, contractors will become increasingly dependent on government as they receive a smaller share of total funds allotted for services. Smaller funding shares (percentage of total funds) indicate that the government is less reliant on a particular contractor to provide a service. Because dependence is the inverse of power and is relational, I predict that as government becomes more powerful in these relations and contractors become less powerful, the contractors will be more willing to cede to the performance demands of government. As contractors become less dependent (more powerful) through receipt of increasing fund and market share they will be more likely to pursue individual organizational goals than to perform according to government demands. However, this only poses a problem where there is goal conflict between principals and agents.

**H3a**: Increased *dyadic dependence of contractors on government* will lead to increased likelihood of *contract performance*.

**H3b**: Under conditions of *misaligned goals between government and contractors*, decreased *dyadic dependence of contractors on government* will lead to decreased likelihood of *contract performance*.

At the structural level, the presence of alternatives for either exchange partner changes the nature of dependence in exchange relations (Kelley & Thibaut 1978). Alternatives reduce the dependence of one exchange partner, thereby reducing the power of the other exchange partner. Where alternatives exist, exchange partners may choose to limit or sever relationships with partners who do not meet their expectations.³

³ Sometimes, government must contract in nonoptimal situations, particularly where markets are thin (see Smith and Smyth 1996 for an example.)
In the contracting world, the presence of alternatives relates to competition. I focus on ex post (after the contract) competition. When government uses multiple contracts to produce a single service, it creates a competitive structure where contractors must compete\(^4\) for clients. As competition increases, government is less reliant on any one contractor and has less incentive to utilize contractors who underperform. Hart, et. al. (1997) propose that ex post competition among prisons for inmates might improve quality if judges (not prisoners) have information about quality and can assign inmates to prisons. This is based on the assumption that judges will send prisoners to higher quality prisons, thus giving prisons an incentive to maintain/improve quality.

In the FDJJ example, the FDJJ decides on where youth will be placed based on several characteristics including treatment needs. This creates complexity in ex post competition. Though the FDJJ appears to have multiple alternatives in its supply market, placement choice is constrained by a juvenile’s treatment needs. Underlying affiliations between exchange partners (contractors) may also increase the complexity of competition in a contractor supply market. Understanding the structure of the supply market is important for government to be a “smart buyer” (Kettl 1993) and to develop smart management strategies that make the most efficient use of limited resources. If dependence constrains behavior of contractors, government must understand the true nature of dependence in the supply market in order to use that dependence to its advantage.

In contractor supply markets with multiple providers, government funders appear to be the most powerful actor based on the structure of formal contracts. By choosing to exchange with multiple partners, government funders may decrease their dependence on any one exchange partner and decrease risk associated with non performance while also increasing the power to produce desired outcomes from contractors. It is not just a matter of choosing as many partners as possible though. For example, if government chooses five contractors controlled by the same parent organization, the

\(^4\) Competition may be active or passive but the presence of alternatives creates competition among producers/contractors for clients. When contracts are awarded, clients are not guaranteed.
structure of the supply market will be different rather than if government chooses five contractors each controlled by a different parent organization. Having multiple contractors controlled by the same parent organization reduces the structural dependence of any one contractor on government because it is affiliated with a powerful parent organization. In other words, as government is increasingly dependent on one parent organization for multiple programs, those programs may be more likely to maximize self interest by pursuing organizational goals that may or may not coincide with government goals. I should reiterate that in some circumstances, government faces thin markets and may not be able to structure a supply market so that it is competitive.

Government may also be legally constrained in its selection of providers. For example, in the Florida case, the FDJJ must send juvenile delinquents to facilities based on treatment needs (programs specialize in different treatment needs). This constrains government choice much more than if the FDJJ could send juveniles to any facility in the state. In sum, even though a government may create what seems to be a competitive supply market for services, competitiveness may be constrained by underlying affiliations among contractors through things like parent organizations and mandated treatment requirements.

These constraints will interact to produce a complex structure of exchange partners for government. It is important to analyze these constraints simultaneously rather than independently though. Considering the intersection of multiple affiliations will provide us with a better understanding of contractor position in the supply market than considering affiliations independently. Network theory and methods about multiple affiliations is based on Simmel’s (1955) theory about social circles: identity is created by the intersection of social circles. In the FDJJ setting, contractor position in a supply market is determined by its affiliation with a parent organization and its affiliation with a specialty. Fararo and Doreian (1984) extend the duality argument that examines interpenetration of actors and events (Breiger 1974; Wilson 1982) to situations with three distinct groups of nodes where actor affiliation with
two types of nodes provides a meaningful basis for understanding behavior. Their argument is based on
the following two axioms:

A1: There are three types of nodes.

A2: Links only exist between types of nodes. (modes)

Allowing only ties between types of nodes allows the researcher to examine overlaps and common
membership without the “noise” of person-to-person or group-to-group social relations. Tripartite
affiliation matrices can be conceptualized as follows where links exist between types of nodes but not
within a type of node.

<table>
<thead>
<tr>
<th></th>
<th>Node Type 1 (e.g.- program)</th>
<th>Node Type 3 (e.g.- parent organization)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node Type 1</td>
<td>--</td>
<td>B</td>
</tr>
<tr>
<td>(e.g.- program)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Node Type 2</td>
<td>A</td>
<td>C^5</td>
</tr>
<tr>
<td>(e.g.- circuit)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The tripartite matrix allows us to examine three modes (i.e. – contracted program, parent
organization, and specialty) simultaneously rather than examining two separate duality matrices.
Examining duality matrices of programs and parent organizations and programs and specialties gives us
a series of unconnected constellations that do not indicate interpenetration between all three types of
nodes and thus do not give a complete picture of the contracted program’s position in the supply
market. See Appendix A for a visual representation of bipartite and tripartite matrices derived from the
same data.

---

Fararo and Doreian (1984) examine both the restricted form of tripartite matrix and the generalized form of a
tripartite matrix. In the restricted form, links exist in a nested, hierarchical fashion so that “C” in the matrix above
is null. In the general form, links exist between all three types of nodes so that “C” is not null. This dissertation will
focus on a restricted form of the tripartite matrix and all further reference to tripartite matrices will refer to the
restricted form.
Because the SET conception of power is relational, government power over programs will be the inverse of program dependence on government and vice versa. In the case of parent organization affiliation, being tied to a powerful parent organization (having many programs) will increase the power of contracted programs. Being tied to a specialty with many programs will diminish the power of programs because increased competition means increased alternatives for government. These two factors will combine to produce a structural dependence score for each program. Combining these two factors results in the 2x2 matrix below; the matrix is a simple heuristic and in reality these variables interact along a continuum rather existing in rigid categories.

<table>
<thead>
<tr>
<th>Parent Organization</th>
<th>Few Programs</th>
<th>Many Programs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Many Programs</td>
<td>Abundant choice for government funder (power advantage to government)</td>
<td>Mutual Dependence (power balanced)</td>
</tr>
<tr>
<td>Few Programs</td>
<td>Mutual Dependence (power balanced)</td>
<td>Limited Choice for government funder (power advantage to programs)</td>
</tr>
</tbody>
</table>

If contracted programs are heavily dependent on government (cell 1), they may be more likely to comply with government demand even if the presence of other variables might predict otherwise. Conversely, government dependence on a particular program (reflected through parent organization affiliation and competitiveness of specialty) may lead to less compliance (cell 4). Based on the above arguments, I offer the following hypotheses:

*H4a: Increased structural dependence of contractors on government will lead to increased likelihood of contract performance.*

*H4b: Under conditions of misaligned goals between government and contractors, decreased structural dependence of contractors on government will lead to decreased likelihood of contract performance.*
In sum, I predict that dependence will play an important role in determining whether contractors will perform. Increased dependence on government at the dyadic and structural levels will lead to increased likelihood of contract performance, particularly where goals between principal and agent do not align.

RESEARCH METHODS

Research Setting

This study examines the Florida Department of Juvenile Justice Residential Services Program. Like most states, Florida traditionally employed a rehabilitative model of justice for juvenile delinquents. The state began a shift away from this model in 1994 with the creation of the Florida Department of Juvenile Justice (FDJJ). Legislation moved matters regarding juvenile delinquency from the Department of Health and Rehabilitative services to the newly created FDJJ. This shift led to few substantive changes to Florida’s juvenile justice program though.

In 2000, the Florida Legislature approved a comprehensive overhaul of juvenile justice programs referred to as the “Tough Love” plan. Under this plan, juvenile justice services shifted from rehabilitative, social service models to a more punitive criminal justice approach though treatment for offenders was still considered part of the program. This plan also changed the geographical structure of the FDJJ from one of districts to one that conformed to boundaries of Florida’s twenty judicial circuits. The FDJJ also organized itself into five largely autonomous offices: Administrative Services; Prevention and Victim Services; Detention Services; Probation and Community Intervention and Residential Services.

Residential Services programs provide a continuum of care for juveniles committed to the care and custody of FDJJ. Residential programs vary in size and scope but are more likely to be small, group home like settings rather than the large lock down institutions of detention programs. I focus on
residential programs rather than contractor parent organizations because the FDJJ decides on client placement and evaluates contracts at the program rather than the organizational level. In other words, policy is implemented at the program level, not the organizational level. Though programs share parent organizations, they are not engaged in joint service production.

Residential programs are utilized nationally for juvenile offenders. According to the most recent Federal data, in 2006 about 30% of juveniles placed in residential facilities nationally were placed in private (both for-profit and nonprofit) settings. That year, Florida placed 62.33% of its residential placements in private facilities, one of the highest rates for private placement of juvenile offenders. The range of private placement of juveniles was 1.4% in Missouri in to 72% in Pennsylvania (OJJDP).

[Insert Chart 1 about here]

Florida contracts with nonprofit and for-profit organizations to provide most of its residential services programs in addition to having several state run programs. In FY2008, Florida had 95 residential services programs. Thirty-six of these programs were operated by 12 nonprofit organizations; 44 of the programs were operated by 13 for-profit organizations. The remaining 15 programs were directly operated by the state. I examine only contracted programs in this paper. While comparison to government run programs is important, it is beyond the scope of this paper.

Once circuit judges determine placement type, the FDJJ assigns juveniles to programs based on treatment needs, offense and security concerns. The FDJJ structure is broken down into circuits. Florida Statute recommends that juveniles be placed in facilities in their home circuit if appropriate treatment options are available (FS, Title 47, Part 5, Chapter 982.02; FS, Title 47, Part 5, Chapter 985.601).

Data Collection

Data was collected primarily from government documents including Comprehensive Annual Reports, Quality Assurance Reports, state contract and budget documents and legislative statutes. Some of these documents were acquired through public records while others were provided by the FDJJ.
I have also conducted informal interviews with several key players at the FDJJ including the Director of Research, several contract administrators, and a Residential Services Manager. This data is part of a larger dataset that includes five years of program data.

Measures

**Dependent variable**

My dependent variable is “acceptable quality score” (coded 1 for acceptable, 0 for unacceptable). Acceptable quality scores are those scores that do not trigger sanctions according to FDJJ rules (scores equal to or greater than 70). Though I have continuous quality performance scores, I am more interested in the predicting the likelihood that a program will have acceptable performance than I am in predicting unit changes in the dependent variable for several reasons. First, government is ostensibly interested in acceptable performance. That some programs go above and beyond requirements is great but unnecessary according to the terms of the contract. Second, exceptional performance is not associated with rewards related to the contract.

Program quality is a multidimensional measure based on management accountability, residential community, treatment services, food services and safety/security. Though quality is comprised of a wide range of dimensions, the Cronbach’s alpha for these measures (continuous version) is high at 0.8386, indicating that they are measuring the same underlying construct- quality.

**Independent Variables**

I include four independent variables in my model. As previously discussed, variables thought to influence contract performance after the contract award such as monitoring and management, others are held constant in this setting. This provides a great opportunity to test whether dependence variables derived from social exchange theory influence the likelihood of contract performance. I also use organizational form and staff education as program level attributes that may impact contract performance.
Organizational Form- This is a categorical variable (coded 0=nonprofit, 1=for-profit). In the logistic regression, I predict that organizational form will not be a significant predictor of acceptable performance. In the QCA analysis, I predict that organizational form will be insufficient on its own to produce acceptable performance, but that it will combine with other variables to contribute to the outcome. This data was derived from company websites, financial statements and 990 tax forms of parent organizations.

Education- In best practice standards that the FDJJ recommends (but does not require), education of personnel is a proxy for qualified staff (but is not used to assess program quality). Specifically, best practices state that at least 70% of treatment staff should have at least a bachelor’s degree in a field related to juvenile justice treatment (i.e. – social work, counseling, psychology). However, nearly half of the contracted programs do not meet this recommendation. Many programs choose to employ staff without bachelor’s degrees or with degrees not related to assigned work (i.e. – business degree). Theory might predict that nonprofits will be more likely to meet this recommendation but I found no statistically significant association between meeting this recommendation and organizational form (see Table 1).

The variable is dichotomous (0= does not meet recommendation; 1= meets recommendation). This data was derived from FDJJ Quality Assurance reports. I acknowledge that the actual percentage of staff with appropriate education would have been a more useful variable. However, only a small number of reports contain this data. Reporting requirements are to include whether a program meets the standard or not. For some programs, language in reports was not clear about whether programs met the recommendation. Those cases were dropped from the analysis.

Dyadic Dependence (funds)- Following the social exchange theory perspective, this variable is the inverse of the share of total FDJJ Residential Services Program funds each program receives. I use the inverse because dependence is the inverse of power (Emerson 1972). A power measure would be
the percentage of total funds each program receives. The logic behind this variable is that as
government provides fewer funds to a particular program, it will become less dependent on that
program.

*Ex post competition*- This variable relates to the presence of alternatives. It is the percentage of
market share an individual program occupies in a specialty. For example if there are 10 programs in one
specialty, one program occupies 10% of the market share. There are six specialties in the FDJJ
Residential Services programs.

*Structural Power*- Again, following social exchange theory, the presence or absence of
alternatives constrains the choice of providers. If a program is attached to an organization that the FDJJ
is dependent on, it may be less likely to perform. Here I consider parent organization power and
specialty areas simultaneously. This variable is the number of programs a parent organization is
attached to divided by the number of specialties a parent organization participates in. Larger numbers
indicated decreased dependence of parent organizations on the FDJJ. Because dependence is a
relational attribute rather than an individual attribute, as organizations provide more programs to the
FDJJ, they will become more powerful because the FDJJ is more dependent on them to provide services.
This power will be mitigated by the number of specialties an organization participates in because FDJJ
choice is constrained by specialty. For example, if a parent organization has three programs in three
separate specialties it will have a power score of 1. A parent organization with three programs in one
specialty will have a power score of 3 because the state is more dependent on that organization for a
particular type of treatment.

**Control Variable**

*Size*- In the logistic regression, I control for the size of programs. Size varies from very small (10
beds) to very large (200 beds) though most programs have between 30 and 60 beds.
ANALYSIS AND FINDINGS

To shed light on why some contractors perform and others don’t, I use two types of analysis- a binary logistic regression and crisp set qualitative comparative analysis (QCA). Binary logistic regression is a maximum likelihood model estimate that estimates the probability that certain causal variables will predict an outcome. Crisp set QCA uses Boolean algebra to determine which causal conditions or sets of causal conditions may lead to a specified binary outcome. While both types of analysis focus on which causes may lead to a specified binary outcome, the methods are very different. In this section, I will explain each method and its potential usefulness for answering research questions.

QCA is a relatively new method that does not enjoy the stature afforded to more traditional methods based on correlational analysis. By comparing the results of a logistic regression with the results of QCA, I hope to illuminate some of the benefits and pitfalls of QCA for public management research. Comparing QCA results to a logistic regression will allow me to ground my explanations in an analytical tool familiar to public management scholars.

Binary logistic regression is a common form of regression familiar to most social scientists. It is an appropriate form of regression for dichotomous dependent variables. It uses maximum likelihood estimation to estimate the odds a certain event (coded as 1) will occur. The logistic function is bound by 0 and 1 and will not predict values beyond 1 or below zero.

Qualitative comparative analysis (QCA) was developed as a middle path between quantitative and qualitative analysis (Ragin 2008) based on set theoretic principles. QCA was originally geared toward small to medium N size research (5-50 cases) but recent work illustrates that QCA is also appropriate for large N datasets (Ragin 2008). QCA is appropriate for situations where there is causal complexity. Rather than focusing on the net effects of one causal variable, QCA identifies combinations of causal conditions that lead to a specified outcome. Of particular interests are INUS conditions, which are not sufficient on their own to predict an outcome but may combine with other conditions to predict
an outcome (Ragin 2008). For example, in the FDJJ dataset, legal ownership may not predict acceptable performance but legal ownership may combine with some other variable like organizational power to produce an outcome⁶.

Table 2 provides descriptive statistics, fuzzy set QCA calibration cutoffs and coding information for variables. Tables 3-5 provide frequencies of dichotomous variables. An important point is that there are more cases with acceptable performance than unacceptable performance. Because QCA examines the subset of an outcome (i.e. – acceptable performance), I focus on acceptable performance as an outcome. QCA is not symmetric, meaning that the inverse of what predicts acceptable performance predicts unacceptable performance.

For the logistic regression, I clustered the analysis on parent organization. This type of analysis provides robust standard errors which are more reliable in this case because programs are indirectly linked through a parent organization. Goodness of fit tests indicate the model is a good fit for the data. Wald tests for independent variables indicate that organizational form should not be in the model, but I left it in because it is of substantive interest. In the logistic regression, several variables are significant. Results of the regression (in logit terms) are reported in Table 6. The analysis indicates that dyadic dependence (funding) is highly significant (p>.01). Meeting the education requirement and parent organization power are moderately significant (p>.1). Because the N is small and standard errors are relatively large, I use caution when interpreting this model. I expect that future analysis on a larger sample (N approx 400) will alleviate some of these problems. However, it is interesting to note that organizational form did not come close to being significant in the regression model. Previous chi² analysis also indicates that there is no association between organizational form and acceptable program quality (see Table 7).

---

⁶ For a more comprehensive explanation of QCA, see Appendix B.
The QCA outcomes are a bit more complex. Analysis focuses on the outcome of acceptable quality score and does not examine unacceptable scores. Basically the analysis examines which conditions or combinations of causal conditions lead to the specified outcome of acceptable quality. There are no necessary conditions, meaning all conditions are a subset of the outcome (acceptable score).

The QCA analysis produced three paths to acceptable quality performance. The solution consistency is 0.91, meaning that 91% of the time when the conditions in the solution are present, the outcome (acceptable quality performance) will also be present. The solution coverage is 0.43, meaning that this solution explains 43% of cases with acceptable quality outcomes. See Table 8 for QCA output. At first glance, the solution coverage score seems unimpressive, but upon closer examination the results are striking because all three paths in the solution include for-profit status, thus excluding nonprofit programs from the solution coverage. This does not indicate that nonprofit programs cannot produce acceptable quality performance. The data clearly tell us otherwise. What it means is that for-profit status combines with other conditions to produce the specified outcome (acceptable performance). Nonprofit status does not consistently combine with other causal conditions to produce the outcome. This indicates that nonprofits will perform well under a variety of conditions and may not need the constraint of other factors. One interpretation of this is that nonprofit goals are aligned with principal goals or that nonprofits serve as stewards who are motivated to perform according to principal demands even where possibilities for opportunistic behavior exist (Davis, et. al. 1997; Van Slyke 2007).

The first path combined the following causal conditions- competitive specialty, education recommendation met, and for-profit status. This indicates that for-profits that exist in competitive specialties and meet the education recommendation are likely to produce acceptable quality outcomes. This path has a consistency of 0.92, meaning that 92% of the time this combination of conditions is present, the outcome of acceptable quality performance is also present.
The second path combines for-profit status with the absence of parent organization power, the presence of competition within a specialty, and the absence of dyadic dependence. The first two conditions align with my predictions, but the absence of dyadic dependence as part of this path is curious. This path has a consistency of 0.95 meaning that 95% of the time this combination of conditions is present, the outcome of acceptable quality performance is also present.

The third path combines for-profit status with the absence of structural power, the absence of dyadic dependence, and the presence of meeting the education requirement. Again, the absence of dyadic dependence is an interesting find. This path has a consistency of 0.997 meaning that 99.7% of the time this combination of conditions is present, the outcome of acceptable quality is also present.

**Hypotheses Support in Logistic Regression and QCA**

The outcomes of the regression and QCA do not disagree with each other- absence of dyadic dependence, absence of structural power, presence of education are important in both types of analysis, but QCA offers further insight as to how these conditions combine to produce the outcome.

*Hypothesis 1* (nonprofits will perform better than for-profits) was not supported by the logistic regression. In the QCA analysis, ownership form combines with other variables to produce causal paths that lead to the outcome. Specifically, for-profit status combines with other variables to produce the outcome.

*Hypothesis 2* (well trained personnel leads to acceptable contract performance) was supported by the logistic regression. This variable was part of one of the causal paths in the QCA analysis.

*Hypothesis 3* (increased dyadic dependence is associated with increased likelihood of performance and vice versa) is not supported by the logistic regression. The funding dependence variable is significant, but the coefficient goes in the wrong direction. The results indicate that decreased dyadic dependence based on funding is associated with the increased likelihood that contractors will have acceptable quality performance. QCA only assesses what leads to the specified
outcome (acceptable quality), but it also identifies decreased dyadic dependence based on funding as part of two paths that lead to acceptable program quality. Dyadic dependence based on competitiveness of the specialty is not significant in the logistic regression, but this variable is part of one path identified by QCA that lead to acceptable program quality.

*Hypothesis 4* (increased structural dependence is associated with increased likelihood of performance and vice versa) is supported by the logistic regression. This variable is also included in the QCA analysis. Specifically, the absence of structural power (increased dependence) is part of two paths that lead to acceptable program quality.

**CONCLUSION**

There are at least three big ideas in this paper. The first is that ownership form may matter, but perhaps not in the way the literature predicts. Nonprofits are not superior, but they may be more likely to perform better under a variety of conditions. For-profits on the other hand may be worthy trading partners, but only under certain conditions. In the face of increased contracting and thin markets, government may be forced to contract with for-profits. Hopefully, this research will lead to more nuanced theory about when for-profits will perform well as contractors. For government to be a “smart buyer” (Kettl 1993), it must look beyond organizational form distinctions.

The second big idea is that structure matters for performance. Structure may be more nuanced than a simple count of exchange partners though and we must delve deeper into potentially meaningful affiliations among contracts that may influence performance. Again, to be a “smart buyer” government must understand the complexity of its exchange network (supply market).

---

7 The coefficient is negative but the measure is based on power- larger numbers indicate increased power of parent organizations so a negative coefficient supports the idea that increased structural dependence (the inverse of power) is associated with the likelihood of acceptable program quality. Future analysis will use a measure of dependence (inverse of current measure in model).
The third big idea is that QCA may be a useful exploratory tool for complex questions. The results of the QCA analysis do not contradict the results of the logistic regression, so why might QCA be important? QCA may be an important tool to complement more traditional correlational analysis. In this case, QCA sheds light on how organizational form might combine with other variables to produce a specified outcome because QCA is focused on combinations of causal conditions rather than the net effects of independent variables. Identifying the combinations of causal outcomes, or “causal recipes” that lead to the presence or absence of an outcome will help us to think holistically about how and why outcomes occur rather than looking at the net effects of a single variable (Ragin 2008). Focus on causal recipes also leads us away from symmetrical reasoning in conventional quantitative research. The absence of combinations of causal conditions that lead to a specified outcome does not mean that the outcome will not occur; absence of an outcome may be caused by other combinations of causal conditions that are independent those that predict the presence of an outcome. QCA may provide us with a more nuanced understanding of our data than traditional analysis might.

That being said, QCA is not without its limitations. For instance, in fuzzy set analysis, the researcher must determine inclusion in the fuzzy set based on substantive knowledge of the data. Even in the best cases, this determination is subjective and may lead to problems in data analysis. Another huge limitation of QCA is that it does deal very well with longitudinal data. This is a problem for researcher who wants to assess outcomes over time.

Future Work

Future analysis will include a larger data set that covers program data for multiple years (N approx 400). Expanding the sample size will increase the power of regression models. Future analysis will also examine cost efficiency and program effectiveness performance outcomes in addition to quality performance outcomes. Preliminary analysis of data indicates that there is not a strong correlation between quality, cost efficiency and program effectiveness outcomes for contracted programs.
I should note that program quality is not a proxy for program effectiveness. Quality is a short term measure based on clearly specified processes and outputs while effectiveness is a longer term goal associated with recidivism reduction. Of real importance is whether program quality translates to longer term program effectiveness goals (Dias and Maynard-Moody 2006). The mission of the FDJJ is “to increase public safety by reducing juvenile delinquency through effective prevention, intervention and treatment services that strengthen families and turn around the lives of troubled youth”. The FDJJ believes that having high quality programs will support the accomplishment of this mission. However, analysis of program quality scores and program effectiveness scores demonstrates the two are not significantly related. Of the contracted programs studied in this paper, about a third of the programs with acceptable quality scores did not have acceptable effectiveness scores. A future direction of my work is to determine what accounts for program effectiveness and how these causal variables differ from those that explain program quality and cost efficiency (both short term measures). It is likely that program effectiveness has little to do with short term program outcomes and more to do with the situations juveniles are placed into once they leave the residential treatment facility.

In addition to examining multiple outcomes, future analysis will incorporate government run programs. These are program delivered directly by the FDJJ and thus subject to hierarchical control rather than more horizontal arrangements found in contract settings. Of particular interest is whether hierarchical arrangements are associated with the increased likelihood of acceptable performance on all measures. To date, empirical research on the success of contracting varies (Boyne 1998; Hirch 1995; Lavery 1999) but government continues to contract out. In some cases, contracting out appears to be driven by factors other than cost efficiency or meeting program goals and may reduce the capacity of governments to produce services (Hefetz and Warner 2004; Van Slyke 2003). Understanding how public and private programs interact in the same environment is of great importance.

---

8 Effectiveness is based on the rate of recidivism within one year of release from a Residential facility.
REFERENCES


Chart 1: Juvenile Residential Program Placement in Private Facilities

- United States - Private
- Florida - Private
Table 1: Organizational Form by education recommendation

<table>
<thead>
<tr>
<th></th>
<th>Nonprofit</th>
<th>For-profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does not meet</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td>Meets</td>
<td>19</td>
<td>22</td>
</tr>
<tr>
<td>Total</td>
<td>32</td>
<td>40</td>
</tr>
</tbody>
</table>

Pearson chi2(1) = 0.1388  Pr = 0.709
### Table 2: Coding, Descriptive Statistics, Fuzzy Set QCA Calibration Cutoffs

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coding</th>
<th>Data Source</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>QCA Fuzzy Calibration Cutoff Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>QA Score</td>
<td>0= unacceptable performance (quality score fell below 70) 1= acceptable performance (quality score is 70+)</td>
<td>FDJJ Quality Assurance Reports</td>
<td>0.713</td>
<td>0.455</td>
<td>0</td>
<td>1</td>
<td>N/A- crisp set</td>
</tr>
<tr>
<td>Organizationa l Form</td>
<td>0= nonprofit 1= for-profit</td>
<td>Organizational websites/financial reports; 990 forms</td>
<td>0.55</td>
<td>0.501</td>
<td>0</td>
<td>1</td>
<td>N/A- crisp set</td>
</tr>
<tr>
<td>Education</td>
<td>0= program does not meet the evidence based educational recommendations for treatment staff (&lt;70% of treatment staff have at least a bachelor’s degree in a helping field) 1= program does meet the evidence based educational recommendations for treatment staff</td>
<td>FDJJ Quality Assurance Reports (derived from Tier II, section 1.02)</td>
<td>0.569</td>
<td>0.4986</td>
<td>0</td>
<td>1</td>
<td>N/A- crisp set</td>
</tr>
<tr>
<td>Dyadic Dependence (funds)</td>
<td>This is the inverse of the share of total Residential Services Program funds individual programs receive from the FDJJ. Larger numbers indicate decreased dependence on the FDJJ.</td>
<td>FDJJ CAR Reports</td>
<td>154.634</td>
<td>9</td>
<td>102.663</td>
<td>2</td>
<td>21.443</td>
</tr>
<tr>
<td>Dyadic Dependence (competition)</td>
<td>This is the percentage of the specialty a program occupies (1 divided by the total number of programs in a specialty).</td>
<td>FDJJ Quality Assurance Reports; FDJJ CAR Reports</td>
<td>0.0496</td>
<td>0.03787</td>
<td>0.027</td>
<td>0.157</td>
<td>0.0357, 0.0714, .15 (fuzzy set for competition)</td>
</tr>
<tr>
<td>Structural Power</td>
<td>This is the number of programs attached to a parent organization divided by the number of specialties an organization participates in. Larger numbers indicate more power.</td>
<td>FDJJ Quality Assurance Reports; FDJJ CAR Reports</td>
<td>1.849</td>
<td>0.775</td>
<td>0.5</td>
<td>3.5</td>
<td>3, 2, 1 (fuzzy set for power)</td>
</tr>
<tr>
<td>Size</td>
<td>This is the number of beds</td>
<td>FDJJ Quality Assurance Reports</td>
<td>51.924</td>
<td>38.546</td>
<td>10</td>
<td>200</td>
<td>100, 50, 25 (fuzzy set for large programs)</td>
</tr>
</tbody>
</table>
Tables 3-5: Frequencies for Dichotomous Variables

**Table 3: QA Score - Dichotomous**

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Acceptable</td>
<td>23 28.75%</td>
</tr>
<tr>
<td>Acceptable</td>
<td>57 71.25%</td>
</tr>
<tr>
<td>Total</td>
<td>80 100%</td>
</tr>
</tbody>
</table>

**Table 4: Organizational Form**

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonprofit</td>
<td>36 45.00%</td>
</tr>
<tr>
<td>For-profit</td>
<td>44 55.00%</td>
</tr>
<tr>
<td>Total</td>
<td>80 100.00%</td>
</tr>
</tbody>
</table>

**Table 5: Meets Education Recommendations**

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does not meet</td>
<td>31 43.06%</td>
</tr>
<tr>
<td>Meets</td>
<td>41 56.94%</td>
</tr>
<tr>
<td>Total</td>
<td>72 100.00%</td>
</tr>
</tbody>
</table>
Table 6: Logistic Regression Results

Outcome: Acceptable Quality Score

<table>
<thead>
<tr>
<th></th>
<th>Coef</th>
<th>Robust Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organizational Form</td>
<td>0.1445212</td>
<td>(0.5613587)</td>
</tr>
<tr>
<td>Education</td>
<td>1.099746*</td>
<td>(0.6652501)</td>
</tr>
<tr>
<td>Dyadic Dependence (funds)</td>
<td>-0.0091925***</td>
<td>(0.0026562)</td>
</tr>
<tr>
<td>Dyadic Dependence (competition)</td>
<td>-5.594565</td>
<td>(6.280419)</td>
</tr>
<tr>
<td>Structural Power</td>
<td>-0.6315633*</td>
<td>(0.3804257)</td>
</tr>
<tr>
<td>Size</td>
<td>-0.0067677</td>
<td>(0.0083665)</td>
</tr>
</tbody>
</table>

N 70
Log likelihood -35.132
Pseudo R2 0.1611
Hosmer-Lemeshow chi2(8) = 7.38, Prob > chi2 = 0.4965

* Significant with $\alpha < 0.10$
** Significant with $\alpha < 0.05$
*** Significant with $\alpha < 0.01$

Table 7: Chi$^2$ Analysis of Organizational Form and Acceptable Program Score

<table>
<thead>
<tr>
<th></th>
<th>Nonprofit</th>
<th>For-profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Acceptable</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Acceptable</td>
<td>24</td>
<td>33</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>36</strong></td>
<td><strong>44</strong></td>
</tr>
</tbody>
</table>

Pearson chi2(1) = 0.6712  Pr = 0.413
Table 8: Results of Qualitative Comparative Analysis

<table>
<thead>
<tr>
<th>Causal Path</th>
<th>Raw Coverage</th>
<th>Unique Coverage</th>
<th>Consistency</th>
</tr>
</thead>
<tbody>
<tr>
<td>For-profit. Competitive Circuit. Education Recommendation Met</td>
<td>0.3282</td>
<td>0.1788</td>
<td>0.918813</td>
</tr>
<tr>
<td>For-profit. Competitive Circuit. Absence of Structural Power. Absence of Dyadic Dependence (funding)</td>
<td>0.231</td>
<td>0.0816</td>
<td>0.949055</td>
</tr>
<tr>
<td>For-profit. Absence of Structural Power. Absence of Dyadic Dependence (funding). Education Requirement Met</td>
<td>0.1714</td>
<td>0.022</td>
<td>0.996512</td>
</tr>
</tbody>
</table>

Solution Coverage: 0.431800  
Solution Consistency: 0.913669
Appendix A: Visual Representation of Bipartite and Tripartite Matrices Derived from the Same Data

Bipartite Matrices
Appendix B: Discussion of Qualitative Comparative Analysis

Qualitative comparative analysis (QCA) was developed as a middle path between quantitative and qualitative analysis (Ragin, 2008) based on set theoretic principles. QCA was originally geared toward small to medium N size research (5-50 cases) but recent work illustrates that QCA is also appropriate for large N datasets (Ragin, 2008).

QCA is based on analysis of set relations. Set relations can be conceptualized along a number of dimensions but the easiest way is to think about subsets. For example, nonprofit organizations are a subset of all types of organizations. We can also look at subset relations between variables that are causally connected. For example, a normative assumption in the literature on nonprofits is that nonprofit organizations make good human service contractors. In this example nonprofit organizations are a subset of good human service contractors. In this example, causal connections are driven by theory. For example, from an economic standpoint, nonprofits will be better providers of complex human services because they do not have a profit motivation and have less reason to undercut service.

Set theoretic analysis can be used to identify explicit connections between variables in two ways. First, one can examine whether cases with the same outcome share a set of causal conditions, making that causal condition necessary for the outcome to occur. In this case, the outcome variable is a subset of the causal condition. Second, it can also be used to determine whether a set of cases sharing causal conditions are a subset of cases with some outcome, making the causal condition sufficient, but not necessary to produce the specified outcome (Ragin, 2000). In this case, causal conditions are a subset of the outcome.

In quantitative research, sets are typically thought of as nominal scale variables that need to be converted into dummy variables to be used in correlational analyses. The premise of QCA is that data can be analyzed from a set theoretic perspective. Analysis of set relations is an important analytical tool that can be used independently of correlational analysis.

Set membership has typically been characterized as crisp sets with values of 0 or 1; cases belong to a set or they don’t. However, QCA has evolved to include fuzzy set membership scores in its models. Fuzzy sets assess the degree of membership in a set with values ranging from 0 to 1. Fuzzy sets allow for varying degrees of membership in a set and thus allow researchers to build more nuanced models than are possible with crisp sets (Ragin & Pennings, 2005). Fuzzy sets are calibrated so that degree of membership in a set conforms to theoretical or other substantive standards. Calibrating fuzzy sets requires that researchers have in depth knowledge of cases (Ragin, 2008).

One of the key differences between set theoretical analyses and correlational analysis is that set theoretic connections are not symmetrical (Ragin, 2008). In set theoretical analysis, the introduction of a new subset does not negate or diminish the importance of other subsets. In correlational analysis, the introduction of new variable with explanatory power may diminish the correlation between outcome and independent variables.

For example, there may be good service human contractors that are not nonprofit. This does not undermine the set theoretic claim that nonprofits make good human service contractors. Rather it demonstrates that there may be multiple subsets of good human service providers. These subsets need not overlap/be related. Conversely, the incorporation of other types of good human service providers in correlational analysis would lower the correlation between nonprofits and good human service contractors because of the assumption of symmetry in correlational analysis. A symmetrical argument driven by correlational analysis would be “nonprofit organizations are good human service contractors, other types of organizations are not good human service contractors”. Restated, the introduction of
good human service contractors who are not nonprofit reduces the correlation between nonprofit organizations and good human service providers.

Another feature of QCA is that it utilizes configurational thinking to develop “causal recipes” for specified outcomes. This differs from finding the net effects of independent variables, which is standard in conventional quantitative research. The assumption of “net effects” models is that each causal variable can influence the probability of the dependent variable, independent of the effects of other variables. The net effect is the unique contribution of the independent variable in production of the outcome variable. In this type of analysis, combined variables are treated as interactions that influence the effect of independent variables on dependent variables. These interactions make it difficult to assess net effects. (Ragin 2008)

In configurational thinking, net effects of variables are not particularly important. Rather the focus is on how variables combine to produce a specified outcome. Identifying the combinations of causal outcomes, or “causal recipes” that lead to the presence or absence of an outcome will help us to think holistically about how and why outcomes occur rather than looking at the net effects of a single variable (Ragin 2008). Focus on causal recipes also leads us away from symmetrical reasoning in conventional quantitative research. The absence of combinations of causal conditions that lead to a specified outcome does not mean that the outcome will not occur; it may be caused by other combinations of causal conditions that are independent of the first causal recipe.