Coordination among Small Localities and Returns to Scale in Property Value Assessment: Evidence from New York State

Yusun Kim
Department of Public Policy
University of Connecticut, Hartford, Connecticut, 06103
yusun.kim@uconn.edu

Yilin Hou
Department of Public Administration and International Affairs
Syracuse University, Syracuse, New York, 13244-1020
yihou@syr.edu

John Yinger
Departments of Public Administration and International Affairs and of Economics
Syracuse University, Syracuse, New York, 13244-1020
jyinger@maxwell.syr.edu

Abstract

This paper takes advantage of cooperative agreements among small tax-assessing jurisdictions to explore economies of scale in property valuation. New York State incentivizes neighboring cities and towns to unify their assessment function while maintaining separate taxing authority. With administrative data for years 2003-2014, we estimate cost function models with control function and instrumental variable estimators to examine returns to scale in property assessment. We use information on jurisdictions' border intersection and prior experience in sharing public service with neighbor localities to address potential bias in selecting coordination partners. Our results show consistent evidence of economies of size. Our elasticity estimate of total assessment cost with respect to parcel count is -0.9. We also suggest that expanding an assessing jurisdiction's size may save personnel, operational and contractual costs.

JEL codes: H21, H71, R51

Key words: Property tax, Value assessment, Local financial administration, Economies of scale, Cost function

1. Introduction

This paper examines the technology for value assessment, a key dimension in property tax administration. The property tax is the main revenue source for local governments and the financial base of local autonomy in the United States. The administration of the property tax – registration, assessment, appeal, levy, collection, and enforcement – is mostly a local function. The Assessor's Office of each tax-assessing jurisdiction (or "unit") is charged with determining the taxable value of each and every parcel, which historically was and usually is a varying ratio of, and based on, an estimate of the parcel's market value.

Local tax-assessing units all started small; over time, the trend has been to consolidate the assessing function from small to larger jurisdictions that are often at a higher level, for example, from village to town to county. This paper estimates the returns to scale in property assessment to provide empirical evidence as support for consolidation. As noted by Duncombe & Yinger (1993), the notion of scale contains several dimensions. In the context of this paper, "scale" has two dimensions: one is jurisdiction size measured by the number of parcels; the other is the quality or outcome of assessment. We focus on the first dimension in order to identify economies of size in property assessment, which are said to exist when the assessment cost per parcel falls as jurisdiction size increases, holding assessment quality constant. These scale estimates will cast real and practicable implications for the establishment, consolidation or division of local jurisdictions at least for optimal administration of the property tax.

There is a continuous line of prior research examining returns to scale in public service provision. These include studies of returns to population scale in education (Tholkes, 1991; Pratten, 1991; Duncombe & Yinger, 2007) and in police and fire services (Wasylenko, 1977; Duncombe & Yinger, 1993). This body of literature offers a valuable guide for us in estimating the appropriate cost function in order to capture economies or diseconomies of scale in different population ranges; this literature also reminds us that changes in scale may lead to adjustment costs (Kenny, 1982; Cotton, 1996; Howley, 1996; Duncombe & Yinger, 2007). In contrast, the literature on the economies of scale in property assessment (Wicks & Killworth, 1967; Sjoquist & Walker, 1999) stopped by the turn of the century, leaving a gap. This study is designed to fill the gap.

Since the property tax is administered by local governments, most of which are small, many U.S. states incentivize small, neighboring tax-assessing localities to consolidate or centralize their assessment function to the county, under the assumption that larger assessing units will achieve economies of size. Since the 1980s, New York state has been encouraging small assessing units to combine their tax assessors' offices while each preserves autonomy in budgeting and taxation. This inter-municipal collaboration is much more operable administratively and politically than consolidation that takes a long procedure. We take advantage of this program of NY to see whether merging the property assessing function indeed generates economies of size. We estimate the magnitude of returns to size using a unique panel data set of NY assessing jurisdictions from 2003 to 2014, with a cost function framework and instrumental variable strategy.

The results of this paper provides empirical evidence and carry practicable implications not only for possible rearrangement of local tax-assessing jurisdictions in American states but also for the design, consolidation and division of local governments in other countries. In a broader sense, this paper extends the perennial discussion on equity and efficiency in public service provision to the study of property tax administration.

2. Literature Review

The literature on returns to scale in the provision of public services shows mixed evidence of economies of size. Early studies of fire, roads, policing, and public libraries (for example, Ahlbrandt, 1973; Walzer, 1972; Deller et al., 1988) resorted to *ad hoc* functional forms rather than drawing from economic theories to model costs. Subsequent research started adopting Bradford et al.'s (1969) cost function framework that adapted the theory of cost minimization at the firm level to the public sector (Fox, 1981). Duncombe and Yinger (1993) estimated a general cost function in fire protection and delineated four dimensions of returns to scale.

Empirical results in the education finance literature often point to a school district enrollment size that minimizes costs (Ratcliffe et al., 1990; Downes & Pogue, 1994; Duncombe et al. 1995, 1996; Imazeki & Reschovsky, 1997). Tholkes (1991) and Pratten (1991) summarize the potential sources of economies of size in education that may apply to other public services. The sources are indivisibility of labor input, increased dimensions conducive to sharing capital or technology, specialization, price benefits of size from input purchases, lower cost of innovation, and positive learning spillover effects.[1] While most studies rely on cross-sectional variation in size, Duncombe and Yinger (2007) base their estimation on enrollment changes that accompany school district consolidation, using the cost function framework to identify returns to size in assessment. Andrews et al. (2002) underlines the methodological challenges, such as the measurement of performance and efficiency, and highlights the importance of addressing simultaneity and omitted variable biases in cost function models.

Few studies have estimated economies of size in property tax administration. To our knowledge, Netzer (1966) and Wicks and Killworth (1967) were the first to provide jurisdiction-specific cost estimates, which average 1.5 percent of property tax revenue. This line of research stopped for over three decades. Sjoquist and Walker (1999), using a sample of 138 county-level assessment offices in Georgia and a trans-log cost function model, found evidence of significant economies of size, with an assessment volume elasticity of 0.3. This result indicates that consolidation of small assessing units would significantly reduce the cost per parcel. However, this article does not address potential endogeneity in performance measures or consolidation decisions. Then, this line of literature stopped again. The current study aims to fill in the gap in the literature with most recent empirical evidence and more solid methodology.

3. Property Assessment in New York State

As a strong home-rule state, NY's property tax system is highly decentralized, with the assessment function exercised by a large number of small jurisdictions, as many as 1,546 assessing units in 1983. Despite reform efforts in the ensuing decades, the number in 2017 still stood at 994 (ORPTS, 2018), including 932 towns (equivalent to townships in other states) and 62 cities.[2] Many small, rural towns face financial, technical and other challenges to conduct reappraisal on a regular basis.[3] Moreover, jurisdictions vary considerably in their assessment practices as well as the qualifications and status of their assessors. For example, 524 jurisdictions employ part-time assessors. By a recent survey, 180 assessors each serve two to three jurisdictions (ORPTS, 2011).

As noted in the academic and professional literature on property assessment, appointed assessors are better insulated from political pressure than their elected counterparts (Bowman & Mikesell, 1989; 1990b). Appointment of assessors is heralded as professional standard, whereas election of assessors is a tradition in strong home-rule states. In NY state, most jurisdictions now appoint their assessors, which reflects and is in line with a nationwide trend toward professionalism;

on the other hand, up to 6 percent of jurisdictions still cling to the tradition of electing a sole or a board of 3 assessors (ORPTS, 2018).[4]

The NY *Real Property Tax Law* does not provide a uniform assessment cycle or assessment ratio as the standard; neither does the *Law* prescribe the county as the level of government for the assessment function. Instead, the State has funded three financial aid programs since the late 1970s to promote best practices in property assessment. The first program, *Cyclical Reassessment Aid*, incentivizes reassessment at shorter intervals.[5] The second, the *Consolidated, Coordinated and County Assessment Program*, encourages small assessing units to merge their assessment to the county. The more recent *Coordination Assessment Program* (CAP), introduced in 1994, is an alternative to consolidation, to assist moving the assessment function up to an intermediate level. It is designed to avoid the full merger of two or more local governments which is a long process that faces high political hurdles and may incur adjustment costs. CAP participants receive a one-time lump-sum state aid in the year they unify their assessing function. The maximum state aid for each municipality was capped at $140,000 until 2005, then limited to $100,000 starting in 2006.

The main requirements a CAP imposes on participating jurisdictions are employing a single assessor, assessing at a uniform percentage of market value, and using the same assessment calendar. The sole assessor is employed by one of the participants of a CAP, with the total assessment cost, including assessor compensation, shared among CAP participants based on their respective parcel counts. Each year the state Department of Taxation calculates an equalization rate to translate assessed values into market values. This rate is identical for jurisdictions in the same CAP. Each participating jurisdiction sets its own tax rate, maintains its own assessment appeal procedure as well as separate assessment rolls, and can file its own appeal about the equalization rate (ORPTS, 2011).

Each CAP agreement requires a majority vote of the participating Town or City Boards for approval. Between 2000 and 2010, a total of 101 jurisdictions formed 51 CAPs. In the same period, 17 CAPs were dissolved. We translate the variation in CAP arrangements into changes in the size of assessing units to help estimate the returns to size.

4. Analytical Framework and Models

4.1. The Cost Function Framework

We follow the standard cost function approach, developed by Bradford et al. (1969), for estimating returns to scale in public production. This approach is derived from the economic theory of cost minimization in the provision of public services; it uses a modified version of the standard private sector cost function at the firm level. We start with education finance literature on educational costs (Duncombe & Yinger, 1993, 2011; Duncombe et al., 1995; Downes & Pogue, 1994; Imazeki & Reschovsky, 2004) and adapt it to the context of property tax assessment. We draw heavily on the framework in Duncombe & Yinger (2007), which uses enrollment changes from school consolidation to estimate economies of size in education.

In comparison to the provision of other public services that involve different types of resources and multiple actors with vested interests, property assessment is straightforward and more of a technical nature. Nevertheless, assessment also requires inputs (I_i in i-th jurisdiction) such as labor costs for assessor and staff, capital costs and contractual expenses if external contractors are hired to conduct field visits. These inputs translate into an intermediate output which we call property assessment services, $G_i = g(I_i)$, such as data bases. The second stage of the Bradford et al. (1969) framework translates G_i into a final output, S_i, which is what voters care about. The most plausible measure of S_i is assessment uniformity – the inverse of variation in assessment ratio across properties, measured as the coefficient of dispersion (COD). This final output is a function of the intermediate output (G_i), physical factors like the number of parcels of property (N_i) that measures jurisdiction size, and other environment variables at the jurisdiction level, (D_i), such as population density that affect assessing costs. The second-stage production function can be written as:

$$S_i = h(N_i, G_i, D_i) \text{ or } h'(N_i, g(I_i), D_i).$$

The first-stage cost function indicates the minimum cost for a given G_i and given input prices (W_i). This cost is not directly observed, but often measured by spending, E_i. Spending is also affected by inefficiencies in assessment, however; so a complete model requires controls for inefficiency, F_i. Solving the second-stage production function for G and substituting into the first-stage yields the following second-stage:

$$E_i = c[h^{-1}(N_i, S_i, D_i), W_i, F_i].$$

We specify this function in multiplicative form, which is consistent with the Cobb-Douglas production technology and leads to a log-log estimating equation. To focus on assessment cost per parcel, we divide all cost factors by N_i. Introducing year subscripts plus year and jurisdiction fixed effects leads to our basic estimating equation:

$$\ln(E_{it}/N_{it}) = \beta_0 + \beta_1 ln(N_{it}) + \beta_2[ln(N_{it})]^2 + \beta_3 ln(S_{it}) + \beta_4 ln(W_{it}) + \beta_5 \ln(F_{it}) + \Pi \boldsymbol{X}_{it} + \theta_t + \delta_i + \varepsilon_{it} \quad (1)$$

Economies of size exist if the elasticity of $\frac{E_{it}}{N_{it}}$ with respect to N_{it} is negative. We include a quadratic term of $ln(N_i)$ to test whether economies of size depend on size.

To estimate this equation, we measure W_{it} county level average wage of relevant occupations in the private sector that are comparable to local assessors. Following Sjoquist and Walker (1999), we assume that the price of capital input for property assessment is the same across jurisdictions within the same state. We account for inefficiency, F_{it}, using the share of parcels that are wholly tax exempt, because the identification and verification of exempt parcels requires staff time and resources that could be spent on conducting assessment instead.[6] We also assume that other sources of inefficiency do not vary over time, so that their impact is absorbed by the jurisdiction fixed effects, δ.[7] Vector $\boldsymbol{X}_{it}$ includes other observable jurisdiction characteristics that vary over time and affect assessment costs. These include environmental determinants of costs often used in the cost function literature, such as the population growth rate and the median income share. We also include the share of agricultural properties as a proxy for rural parcels, as well as the share of commercial and industrial properties, and the annual growth rate of the median market value. The county level unemployment rate and housing price index are also included in the model to account for confounding effects of wealth and economic fluctuations in the county area.

4.2. Accounting for the Decision to Join a CAP

The main variation in jurisdiction size over time occurs when jurisdictions enter a CAP. Joining a CAP might incur adjustment costs as well as achieve economies, or diseconomies, of size. Hence, we include a binary variable to indicate whether a jurisdiction is in a CAP and we treat this status as endogenous. Jurisdictions can choose to merge their assessing functions with their neighbors, with the intention to reduce their high assessment costs, or choose not to enter a CAP when they can manage the costs. Stylized facts in Table 2 suggest that CAP participants have lower COD, higher equalization rates, and higher assessment ratios for residential property, which are all strong indicators of assessment outcome, than non-CAP units. Should there be a systematic pattern of jurisdictions joining CAPs due to high assessment costs, transformations in jurisdiction size that reflect this choice would become endogenous. Then, using a simple two way fixed effect model to estimate the cost function models would lead to a simultaneity bias. Therefore, we model the decision to join a CAP, using a rich set of exogenous variables as well as instrumental variables to estimate the cost function models.

The endogeneity problem we face is that a jurisdiction's decision to enter a CAP may depend on assessment costs (which is the dependent variable in our model), the number of parcels and the change in parcel count from joining a CAP. To illustrate, let N_{it} be the number of parcels in a given jurisdiction, i, and ${N^{CAP}}_{it}$ be the total parcel count when the jurisdiction joins a CAP (thus $C_{it} = 1$).

$$\begin{cases} {N^{CAP}}_{it} = N_{it} + N_{jt} & \text{if } C_{it} = 1 \text{ and } C_{jt} = 1 \\ {N^{CAP}}_{it} = N_{it} & \text{if } C_{it} = 0 \end{cases}$$

Simply regressing assessment expenditure on the actual ${N^{CAP}}_{it}$ may lead to biased estimates, however, because ${N^{CAP}}_{it}$ depends on the decision to join a CAP, which is endogenous. We use two different approaches, control function and instrumental approach as our main identification strategies to address this issue.

Instrumental Variable Approach I

First, we construct three instrumental variables for ${N^{CAP}}_{it}$, based on the intersection of geographic borders and the history of inter-municipal cooperation in providing public services. The first instrument ($Z_{1,i,t-1}$) is the cumulative number of CAPs formed in a county until the previous year, excluding the CAP that *i* is in, which reflects the exposure of a jurisdiction to an environment of collaboration among neighboring tax assessing jurisdictions. We hypothesize that the more CAPs are formed in a county, the higher the probability for a given jurisdiction to enter a CAP. Indeed, we observe a positive correlation between this instrument and the predicted probability of joining a CAP. The second instrumental variable ($Z_{2,i,t-1}$) is the ratio of adjoining jurisdictions that had experience in sharing other public services among of all other jurisdictions in the same county. The third instrumental variable ($Z_{3,i,t-1}$) is inspired by the "shift-share" (Bartik, 1991) instrument[8]. We use county-level variation of experience in inter-municipal collaboration of public service provision and interact it with jurisdiction level spatial variation that does not vary over time and is exogenous to a jurisdiction's spending decisions in year *t.* We construct $Z_{2,i,t-1}$ as the mean count of counties that a given county c shares public services with, multiplied by the relative size of neighboring jurisdiction whose borders are contiguous to that of a given town/city, within the same county. We report the conditional first-stage F-test statistics and Hansen J-test statistics in the result section, to check for the validity of these instrumental variables.

Instrumental Variable Approach II

As an alternative approach, we construct a separate prediction of each jurisdiction's CAP decision for each year, based on exogenous jurisdiction level traits. Then we use the predicted CAP decision as an instrumental variable for $N^{CAP}{}_{it}$. In order to generate an exogenous prediction of the change in parcel count from forming a CAP, we assume that jurisdictions will look at towns and cities nearby to find a potential partner when they are deliberating whether to participate in a CAP. We also define the parcel count of a potential partner, N_{jt} in various ways, using several different criteria of choosing a potential partner. First, we assume that jurisdictions will choose the largest neighbor in the same county as its partner. We define $N'{}_{it}$ as the sum of N_{it} and the number of parcels in a chosen partner, j, that is the largest other district within the same county, regardless of the partner's CAP experience. The size of jurisdiction j would be exogenous to jurisdiction i. Second, we limit the pool of potential partner to jurisdictions that never joined a CAP during the sample period, and choose one that has the largest parcel count in the same county.

Finally, we assume that the potential partner will be an adjoining jurisdiction that is closest in size and located in the same county. We then use a logit model of the decision to join a CAP based on observable and exogenous traits of both a given jurisdiction and its potential partner which can be expressed as follows:

$$CAP_{it} = \alpha_0 + \alpha_1 N_{i,t-1} + \alpha_2 N_{j,t-1} + \alpha_3 \boldsymbol{Z}_{i,t-1} + \Upsilon\, \boldsymbol{Y}_{i,t-1} + \Upsilon\, \boldsymbol{Y}_{j,t-1} + \varepsilon_{i,t-1} \qquad (2)$$

$$\begin{cases} C^*{}_{it} = 1 \ \text{ if } \hat{C}_{it} > 0.5 \\ \qquad = 0 \ \text{ if } \hat{C}_{it} \le 0.5 \end{cases}$$

where $\boldsymbol{Z}_{i,t-1}$ is a vector of three instrumental variables used to predict CAP decision, and vector $\boldsymbol{Y}_{i,t-1}$ includes all exogenous traits of jurisdiction *i* included in equation (1) as well as the annual growth rate of housing price between the 65th and 95th percentile. The rationale of including the latter is that wealthier jurisdictions with less frequent reassessment might choose longer intervals due to strong resistance from owners of expensive properties that would not benefit from frequently updating property values. Vector $\boldsymbol{Y}_{j,t-1}$ includes the same set of predetermined exogenous characteristics of a potential partner jurisdiction.

Using predicted values of CAP_{it} from model (2), which we will note as $\hat{C}_{it}$, we then define a dummy variable of whether a jurisdiction is more likely to be in a CAP in year *t*, which will be associated with this latent variable:

$$\begin{cases} C^*{}_{it} = 1 \ \text{ if } \hat{C}_{it} > 0.5 \\ \qquad = 0 \ \text{ if } \hat{C}_{it} \le 0.5 \end{cases}$$

To estimate the economies of size among jurisdictions that became larger by joining a CAP or smaller upon leaving one, we construct a new measure of size to reflect the CAP status, defined as:

$$N^*{}_{it} = N_{it} + N_{jt}\, C^*{}_{it}$$

where the parcel count of the potential partner, N_{jt}, is exogenous to that of jurisdiction i. Thus, predicted parcel count ${N^*}_{it}$, is estimated solely based on exogenous information and will be used as an instrumental variable for ${N^{CAP}}_{it}$ in the second stage regression. The size of a jurisdiction that seems more likely to join a CAP with its potential partner j will become the sum of the jurisdiction's own parcel count and that of the potential partner, only in years where ${C^*}_{it}$ equals one, and otherwise remain the same as N_{it}. This allows us to exploit the meaningful variation in jurisdiction size using exogenous information, rather than introducing an endogeneity problem into our final expenditure model.

The second stage equation (3) is a modified version of equation (1),

$$\ln\left[\frac{E_{it}}{{N^{CAP}}_{it}}\right] = \beta_0 + \beta_1\, ln(\widehat{N^{CAP}}_{it}) + \beta_2[ln(\widehat{N^{CAP}}_{it})]^2 + \beta_3 K_{1,it} + \beta_4 K_{2,it} + \Pi \boldsymbol{X}_{it} + \theta_t + \delta_i + \varepsilon_{it} \quad (3)$$

where $K_{1,it}$ and $K_{2,it}$ are dummy indicators of predicted participation in a CAP, to pick up adjustment costs. Variable $K_{1,it}$ equals one when a given jurisdiction is predicted to join a CAP this year but did not the previous year, while $K_{2,it}$ is an indicator of a jurisdiction being predicted to join a CAP for two consecutive years since the previous year[9]. Estimates of β_3 and β_4 will both tell us whether there are (dis)economies of size in the short run, specifically during the first one or two years after a jurisdiction joins a CAP. Meanwhile, the estimates of β_1 and β_2 will tell us the overall long-run effect of size on assessment cost using variation in size across all assessing units. Vector $\boldsymbol{X}_{it}$ includes the same exogenous jurisdiction-level covariates from equation (1) that may affect assessment spending decisions.

CAPs may also affect assessment cost through channels other than parcel count. One channel is state aid, since CAPs receive one-time, lump-sum grant from the state, A_{it}. The direction of potential omitted variable bias in our elasticity estimate by excluding A_{it} is ambiguous:

Although significantly and positively correlated with CAP by construction, state aid often comes alongside technical assistance that is earmarked for assessment, which may help reduce costs or lead to an increase in assessment expenditure. As an alternative to directly controlling for state aid by including it as an endogenous covariate, we adjust our measure of own-source assessment expenditure as the dependent variable by subtracting state aid received for reassessment purposes from total assessment expenditure.

Another channel is through the requirement that jurisdictions in a CAP adopt the same assessment ratio (assessed value as percent of market value). We control for the interval (in years) since the last assessment, R_{it}, to account for the possibility that jurisdictions in a CAP may be able to cut costs by conducting reassessments more frequently with their partners so that each assessment is done at a lower cost. If entering a CAP leads to a rise in the assessment ratio due to more frequent assessments, R_{it} may partly capture the associated impact on costs.[10] We instrument for R_{it} by either drawing on the copycat theory or using an average predicted value of reassessment frequency between jurisdiction *i* and its potential partner *j*. We use the same approach to instrument for S_{it}, or COD which measures assessment outcome that median voters care about.

Control Function Approach

One of the main limitations of using instrumental variables is that the coefficient estimates of the size variable will be only informative of the local average treatment effect, which only applies to compliers (as shown in appendix table A2) and limits external validity of the findings. For instance, the LATE effect identified by our second instrumental variable approach is the cost effect of CAP induced change in jurisdiction size, *if* and *only if* jurisdictions were predicted to enter a CAP and joined a CAP. This effect will be different from the average cost effects from actual change in jurisdiction size among a more general study of the population of assessing units.

As an alternative approach, we employ a control-function approach to account for the unobservable factors in the error term by modeling them directly and including the fitted values in the main model.[11] Control-function estimators are known to be more efficient and precise than instrumental variable approach using 2SLS estimators, which have stricter identifying restrictions (Heckman & Richard, 1985; Imbens & Wooldridge, 2007). With the control function, we can conduct simple and robust tests of exogeneity of other covariates (Wooldridge, 2015). We run separate first stage regressions for each of the endogenous covariates on a set of exogenous variables, then retrieve the residuals from each regression. The second stage regressions include the residuals as additional regressors instead of the endogenous variables, as in equation (4).

$$\ln\left[\frac{E_{it}}{{N^{CAP}}_{it}}\right] = \beta_0 + \beta_1\, ln({N^{CAP}}_{it}) + \beta_2[ln({N^{CAP}}_{it})]^2 + \beta_3 K_{1,it} + \beta_4 K_{2,it} + \Pi\boldsymbol{X}_{it} + u_{1,it} + u_{2,it} + u_{3,it} + \theta_t + \delta_i + \varepsilon_{it} \quad (4)$$

where $u_{1,it}$, $u_{2,it}$ and $u_{3,it}$ each indicate predicted residuals from first stage regressions for CAP decision as well as the endogenous covariates.

5. Data

Our sample is comprised of 760 tax assessing towns and cities in New York State, 78 of which participated in 38 CAPs between years 2003 and 2014.[12] Appendix Table A1 lists the CAPs by their year of formation and dissolution. Annual expenditure and revenue data are from the New York State Comptroller's Office and are measured as 2014 US dollars. We match these data to jurisdiction-level administrative information related to assessment practices and the assessing environment from the New York Office of Real Property Tax Services (ORPTS). Data provided by ORPTS include total assessed value of exempt parcels, parcel count by property class, number of exempt parcels per property class, assessment ratio, year of reassessment, and annual records of state aid receipt for property assessment by program. All financial variables are expressed in 2003 dollars.

We construct the COD variable for each assessing unit with parcel-level sales data from the annual *New York Market Value Survey,*[13] using only arms-length sales. We focus on three major classes of residential property, which account for 95% of the sample: one-family year-round residence, rural residence with acreage, and two-family year-round residence. We develop two CODs for use in the tests, one for all three classes and the other for single-family year-round residences alone. We express these CODs as natural logarithms. For ease of interpretation, we multiply them by -1, so that positive coefficients on S_{it} indicate an improvement in assessment uniformity.

Data on inter-municipal cooperation are from the Division of Local Government and School Accountability in the State Comptroller's Office. Our sample begins in 2005, which is when state aid for inter-municipal cooperation was initiated. We calculate border contingency among jurisdictions using the ArcMap10 software and the civil boundaries shape file provided by the NYS GIS Clearinghouse (NYS GIS Program Office, 2018).

We construct median tax share with median housing prices from the *Market Value Survey* series and growth rate in top tier home value using Zillow Home Value Index, which measures typical home values for residential properties in the 65th to 95th percentile range. Population density is from the *American Housing Survey*, population growth from State Comptroller's Office. We also include several county level covariates including housing price index from the Federal Housing Finance Agency. The county-level unemployment rate is from the Bureau of Labor

Statistics, and the county-level private sector wage is from New York Department of Labor. Tables 1 lists the variables and their data sources, while Table 2 displays the mean values of the key variables among CAP participants and non-participants in two periods.

[Tables 1 & 2 here]

Table 3 provides stylized facts concerning jurisdiction size and assessment costs. Panel A shows the summary statistics of the parcel-count distribution in quartile among CAP participants, pre- and post-CAP entrance, and non-participants. The upper half shows the parcel counts of jurisdictions as the basis of the context; the lower half shows changes in parcel count, the source of the variation we exploit in this paper. Both participants pre-CAP and non-participants are typically small, with similar mean counts and the lower two quartiles; whereas CAPs (i.e., participants post-CAP) are substantively larger in the mean and the two lower quartiles than individual participants pre-CAP – the mean is even larger than the third quartile of pre-CAP participants. The average change in parcel count among participants pre-CAP and non-participants indicates the trend growth of real estate. This count is almost the same for these two groups; the count for participants post-CAP is over 16 times larger than for non-participants and participant pre-CAP. Thus, it is clear that CAP is an effective mechanism to increase the size of assessing units. This paper captures the effect of the change in size from CAP formation.

Panel B of Table 3 provides a description of average assessment costs among different groups of jurisdictions. The upper half shows that the assessment costs are higher among jurisdictions that never participated in a CAP as well as the changes in assessment costs among the 78 jurisdictions that merged their assessment functions with their neighbors. The total assessment costs are lower, on average, during years they participated in a CAP. Most of this reduction in total costs seems to be driven by savings in operational and personnel cost, while contractual expenses increase. The rows under predicted collaborations show the average assessment cost when each jurisdiction is predicted to collaborate with their largest neighbor, and that of jurisdictions in years they are predicted not to collaborate.

[Table 3 here]

Figure 1 illustrates the distribution of parcel count among both jurisdictions that joined CAPs and those that never did, as well as how parcel count changes on average, before and after participants entered CAPs[14].

[Figure 1 here]

6. Results

In order to examine the returns to scale in property assessment, our empirical tasks are twofold. The primary task is to construct an alternative measure of assessing unit's size to address the selection problem when using actual parcel counts as the size variable. We start by modeling the likelihood of an assessing unit combining its assessment function with the largest neighbor in the same county. We model this decision based on observable jurisdiction characteristics that are exogenous to the jurisdiction's decision-making process. Then we predict a new size variable, assuming that assessing units that are more likely to merge their assessing function with others in a given year would collaborate with the largest neighbor in the same county, expecting to benefit from economies of scale. Table 4 reports the estimation results that is used to generate predicted parcel count for the IV estimation as well as predicted residual for the control function estimation. Figure 2 depicts the significant overlap between the distributions of actual and predicted parcel counts, estimated using equation (2). The second task is to estimate the two parameters, β_1 and β_2 in equation (4), using the predicted estimates of parcel count and error term from the first stage.

[Table 4 here]

[Figure 2 here]

We find evidence in support of economies of size in property assessment. Our estimate of the elasticity of total assessment cost with respect to parcel count is approximately -0.9. This estimate is consistent across different estimation approaches and model specifications. Table 5 shows the fixed effect estimates that assumes all explanatory variables as exogenous.

[Table 5 here]

Should jurisdiction's decision to join a CAP be an endogenous decision, the fixed effect estimates would be biased. However, the control function and IV models employed in this study all yield similar elasticity estimates that range between -0.9 and -1.1. The baseline control function estimates in Table 6 as well as alternative control function estimates using different assumptions of potential CAP partner, as reported in Tables 7 to 9 all suggest increasing returns to scale. One noticeable difference is that the short-run increase in adjustment cost appears to be significant in the fixed effect estimates, which is not supported by the control function estimates.

[Table 6-9 here]

[Figure 3]

Figure 3 summarizes these findings. Panel A in Figure 3 illustrates the predicted assessment costs across the distribution of parcel counts between 500 and 30,000. This range of parcel count represents 98% of the total population of local assessing units in the State of New York. We observe increasing returns to scale throughout this distribution of parcel count. The control function estimates that relies on the assumption that a given jurisdiction will merge its assessment function with the largest jurisdiction suggest that jurisdictions that have more than 20,000 parcel counts tend to have closer to constant returns to scale. Nonetheless, given that the average parcel count among NY assessing unit is 2,328 and 95% of all assessing units in NY have less than 10,000 parcel, our results show that merging assessing functions lead to reduction in assessment costs. Figure 4 plots the assessment cost curve and elasticity estimates based on various control function estimates. To account for potential confounding effects of time trend in assessment cost that may be led by technology development, we also include jurisdiction-specific linear time trend in alternative model specifications. The elasticity estimates remain significant even after controlling for linear time trend (baseline control function estimate is -0.895 which is statistically significant at 5% level).

[Figure 4]

We also find economies of scale in each subcategory of assessment costs. Estimates under columns (5) to (7) in Table 6 as well as Appendix Table A8 suggest increasing returns to scale in all subcategories. On average, personnel and operational costs each account for 40 to 45 percent of total assessment costs, while the share of contractual costs have a wider range between zero (in-house assessment) to one (contracted-out assessment). Figure 5 shows the heterogeneity in returns to scale based on how much each jurisdiction spend on contractual expense as a share of their total assessment cost. The underlying intuition is that the returns to scale may be different among assessing units contract out their assessment function from those that rely fully on in-house resources to provide assessment services. Due to data limitations, we cannot observe whether each jurisdiction contract out mass reappraisal activities each year. However, this exercise allows us to indirectly check for differences between jurisdictions that are more likely to contract out their assessment functions and those that are not. The cost curve in Figure 5 becomes flatter at a lower part of the parcel count distribution, among jurisdictions that spend more than 30 percent of their total assessment cost on contractual expense. This suggest that the benefits from positive economies of scale may be smaller for relatively larger jurisdictions that contract out assessment activities.

[Figure 5]

Meanwhile, our finding is consistent with Sjoquist and Walker (1999) in that overall there are substantial economies of size. Our results demonstrate stronger evidence of economies of size than the findings of Sjoquist and Walker (1999), where they interpreted their elasticity estimate of 0.323 as evidence of economies of scale[15]. Our cost saving estimates are apparently larger than that of Sjoquist and Walker (1999) that estimated that hypothetical size expansion of assessment units in Georgia that have outputs less than the median volume may reduce costs by about 20 percent. However, one must note that tax-assessing units in New York are typically smaller than those in Georgia, where assessment is conducted at the county level.

7. Conclusion

This paper examines whether tax assessing units can lower their assessment costs by expanding parcel count, when controlling for assessment quality and other assessment practices. We take advantage of a New York State CAP program that incentivizes expansion in the size of assessing units (in parcel count) when they combine their assessment function with neighbors. We use the cost function approach to conduct estimation and address potential selection bias in jurisdiction's decision to enter a CAP. Using selection-on-observables approach, we model each assessing unit's decision to collaborate with a CAP partner, by using spatial intersection of jurisdiction boundaries and history of inter-municipal cooperation for providing local services.

Overall, we find increasing returns to scale in property assessment among NY assessing units. Our elasticity estimate of total assessment cost with respect to parcel count is approximately -0.9, which is consistent across different choice of methodology and model specifications. Our findings demonstrate that local assessing units with more than 500 parcels and less than 30,000 parcels may be able benefit from economies of size. We also find suggestive evidence that an average sized assessing unit with 4,000 parcels, may save approximately 25 percent of costs by doubling its size. Such findings are consistent with those from a much earlier study (Sjoquist & Walker, 1999), confirming the existence of substantial economies of size in property assessment.

Given that all personnel, operational and contractual costs show significant economies of scale, it is difficult to dissect the detailed mechanism of cost savings induced by CAPs with currently available data. Potential mechanisms may include better equipment, well-trained staff as well as more regular or frequent reappraisals. The positive coefficient estimates on reassessment cycle variable, which measures "years since last reappraisal" suggest that infrequent assessment may increase assessment costs. Conversely, if collaborating with a partner jurisdiction leads to more frequent or annual reappraisal, such jurisdictions may be able to reduce assessment costs. In our sample we observe that CAP participants tend to reassess at shorter intervals than non-CAP participants.

This paper has explored an alternative policy for small tax assessing jurisdictions beside politically costly consolidation. The presence of economies or diseconomies of size has important policy implications on the design of local property tax systems and for collaborative governance among localities.

References

Ahlbrandt, R. (1973). Efficiency in the provision of fire services. *Public Choice*, 16(1), 1-15.

Altonji, Joseph G. and David Card, "The Effects of Immigration on the Labor Market Outcomes of Less Skilled Natives," in *Immigration, Trade and the Labor Market*, John Abowd and Richard. Freeman, eds. (Chicago, University of Chicago Press: 1991).

Andrews, M., Duncombe, W., & Yinger, J. (2002). Revisiting economies of size in American education: are we any closer to a consensus? *Economics of Education Review*. 21(3), 245-262.

Bick, Robert & New York Association of Assessors, New York. (2016). *Assessment Staffing Survey*. Unpublished raw data.

Bick, Robert. (2017, September 27). Personal interview with the Assessor of Town of Clay, New York.

Bartik, Timothy. (1991). *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute.

Borland, Melvin & Steven Lile. (1980). "The Property Tax Rate and Assessment Uniformity." *National Tax Journal.* 33(1): 99–102.

Bowman, John H., & William A. Butcher. (1986). "Institutional Remedies and the Uniform Assessment of Property: An Update and Extension." *National Tax Journal.* 39(2): 157–69.

Bowman, John H., & Mikesell, John L. (1989). "Elected versus Appointed Assessors and the Achievement of Assessment Uniformity." *National Tax Journal*. 42:181-189.

_______ (1990a). "Assessment Uniformity: The Standard and Its Attainment." *Property Tax Journal* 9: 219–33.

_______ (1990b). "Improving Administration of the Property Tax: A Review of Prescriptions and Their Impacts." *Public Budgeting and Financial Management*. 10(2): 151-176.

Bradford, D. F., Malt, R. A., & Oates, W. E. (1969). The rising cost of local public services: Some evidence and reflections. *National Tax Journal.* 22(2): 185-202.

Card, David. (2001). "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*. 19: 22–64.

Case, Anne C., James R. Hines, Jr., & Harvey S. Rosen. (1993). Budget spillovers and fiscal policy interdependence: Evidence from the states. *Journal of Public* Economics. 53(3): 285–307.

Chicoine, D. L., & Giertz, J. F. (1988). Uniformity in a dual assessment system. *National Tax Journal*, 43(3): 247-256.

Cornia, G. C., & Slade, B. A. (2006). Horizontal inequity in the property taxation of apartment, industrial, office, and retail properties. *National Tax Journal*. 59(1): 33-55.

Cotton, K., 1996. Affective and Social Benefits of Small-Scale Schooling. *ERIC Digest*.

Cotton, K. (1996). Affective and social benefits of small-scale schooling. *Eric Digest*. EDO-RC-96-5.

Deller, Steven C., David L. Chicoine & Norman Walzer. (1988). "Economics of Size and Scope in Rural Low-Volume Roads." *Review of Economics and Statistics.* 70(3): 459–65.

Dornfest, Alan. (2010). "In search of an optimal revaluation policy: benefits and pitfalls" in Bahl et al eds. 2010. *Challenging the Conventional Wisdom of the Property Ta*x.

Downes, T. A., & Pogue, T. F. (1994). Adjusting school aid formulas for the higher cost of educating disadvantaged students. *National Tax Journal*. 47(1): 89-110.

Duncombe, W., & Yinger, J. 1993. "An Analysis of Returns to Scale in Public Production, with an Application to Fire Protection." *Journal of Public Economics.* 52(1): 49–72.

Duncombe, W. D., Miner, J., & Ruggiero, J. (1995). Potential cost savings from school district consolidation: A case study of New York. *Economics of Education Revi*ew. 14, 265–284.

Duncombe, W. D., Ruggiero, J., & Yinger, J. M. (1996). Alternative approaches to measuring the cost of education. In H. F. Ladd, *Holding schools accountable: Performance-based reform in education* (pp. 327–356). Washington, DC: The Brookings Institution.

Duncombe, W., & Yinger, J. (2007). Does school district consolidation cut costs*? Education Finance and Policy.* 2(4): 341-375.

Ferguson, R. F. (1991). Paying for public education: New evidence on how and why money matters. *Harvard Journal of Legislation*. 28: 466–498.

Ferguson, R. F., & Ladd, H. F. (1996). Additional evidence on how and why money matters: A production function analysis of Alabama schools. In H. F. Ladd, *Holding schools accountable: Performance-based reform in education* (pp. 265–298). Washington, DC: The Brookings Institution.

Fox, William. (1981). "Reviewing Economies of Size in Education." *Journal of Education Fin*ance. 6. 273–96.

Geraci, V. J., & Plourde, J. L. (1976). The determinants of uniform property tax assessment. *Assessors Journal*. 11(4): 235-253.

Geraci, V. J. (1977). Measuring the benefits from property tax assessment reform. *National Tax Journal*. 195-205.

Harris, Robert. (2017, April 21). Personal interview with the Assessor of Flat Creek in Montgomery County, New York.

Heckman, J. & Robb, Richard (1985). "Alternative methods for evaluating the impact of interventions". Journal of Econometrics. Elsevier BV. 30 (1–2): 239–267

Howley, C. (1996). Compounding disadvantage: The effects of school and district size on student achievement in West Virginia. *Journal of Research in Rural Education*. 12, pp.25-32.

Imazeki, J. & Reschovsky, A. (1997). The development of school finance formulas to guarantee the provision of adequate education to low-income students. In W. J. Fowler Jr, *Developments in school finance* (pp. 123–147). Washington, DC: US Department of Education

Liu, G. (2018). The effect of sale methods on the interest rate of municipal bonds: A heterogeneous endogenous treatment estimation. *Public Budgeting & Finance*. 38(2). 81-110.

Netzer, Dick (1966). *Economics of the Property Tax*. Washington, D.C.: The Brookings Institution.

Netzer, D., & Berne, R. (1995). Discrepancies between ideal characteristics of a property tax system and current practice in New York. *Journal of Education Finance*. 21(1), 38-56.

New York State GIS Program Office. (2018). New York State Civil Boundaries Data. Albany, NY.

New York State Office of Tax Policy Analysis. (2011). *2011 Report on Effectiveness of State Technical and Financial Assistance Programs for Property Tax Administration*. Albany, NY.

New York State Office of Tax Policy Analysis. (2018). *Assessment Equity in New York: Results from the 2017 Market Value Survey*. Albany, NY.

Office of the New York State Comptroller. (Various Years). *New York State Open Book: Local Government and School Accountability Financial Data*. Albany, NY.

______. (Various Years) *Comprehensive Annual Financial Report.* Albany, NY.

______. (Various Years) *Annual Market Value Survey.* Albany, NY.

Pratten, C. (1991). *The Competitiveness of Small Firms*. New York: Cambridge University Press.

Ratcliffe, K., Riddle, B., & Yinger, J. M. (1990). The fiscal condition of school districts in Nebraska: Is small beautiful? *Economics of Education Review*. 9: 81–99.

Roehr, William. (2017, April 21). Personal interview with the Managing Principal of Montgomery County, New York.

Sanderson, E., & Windmeijer, F. (2016). A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of Econometrics*. 190(2): 212-221.

Sjoquist, David L. & M. B. Walker. (1999). "Economies of Scale in Property Tax Assessment." *National Tax Journal,* (June): 207-220.

Skidmore, M., C. L. Ballard & T. R. Hodge (2010). "Property Value Assessment Growth Limits and Redistribution of Property Tax Payments: Evidence from Michigan." *National Tax Journal*. 63(3): 509-538.

Stock, James H. and Motohiro Yogo. (2005). "Testing for Weak Instruments in Linear IV Regression," in Donald W. K. Andrews and James H. Stock, eds., *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge University Press, 2005, pp. 80–108.

Tholkes, R. J. (1991). Economies of scale in rural school district reorganization. *Journal of Education Finance*. 16, 497–514.

Walzer, N. (1972). Economies of scale and municipal police services: The Illinois experience. *The Review of Economics and Statistics*, 431-438.

Wasylenko, M.J. (1977). Some evidence of the elasticity of supply of policemen and firefighters. *Urban Affairs Quarterly.* 12(3):365-382.

Wicks, John H., & Michael N. Killworth. (1967). "Administrative and Compliance Costs of State and Local Taxes." *National Tax Journal.* 20(3): 309–15.

Wooldridge (2000). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Mass.: MIT Press.

Yinger, John H. (2012). "Four Flaws in New York States Property Taxes and How to Fix Them: Small Assessment Units." See his personal website.

Notes

[1] A related literature that estimates education production functions at the student or classroom level does not find evidence that school district size affects student performance (Deller & Rudnicki, 1993; Walberg & Fowler, 1987; Ferguson, 1991; Lee & Smith, 1977). A comparison of methods is provided in Duncombe and Yinger (2011).

[2] In addition, 133 villages have their own town assessor for the purposes of village taxes only. These assessors are not considered in this study (Yinger, 2012).

[3] As of 2010, there were 238 assessing units that had not conducted any reappraisals in the prior sixteen-year period. (ORPTS, 2011)

[4] Whether appointed or elected, sole assessors serve six-year terms; elected board of assessors serve four-year terms.

[5] This program was initially introduced as Attainment Aid in 1977 and replaced with the current program in 2010.

[6] In interviews with us and in our survey through the New York Association of Assessors, several town assessors list administering large quantities and categories of property tax exemptions as one of their major challenges. (Robert Bick, Assessor, Town of Clay, New York, Personal interview, September 27, 2017; Robert Harris, Assessor, Flat Creek of Montgomery County, New York & William F. Roehr, Managing Principal, Montgomery County, New York, April 21, 2017)

[7] This two-part strategy (fixed effects plus measures correlated with inefficiency) has been used to account for inefficiency in spending to promote student performance in education (Eom et al. 2014).

[8] The third instrument is inspired by Bartik (1991)'s instrument that utilizes the interaction between variation in nationwide inflow of immigrants and the geographic distribution of immigrants in the past at city level to identify a short-run causal effect of migration on various outcomes

[9] We define the join indicator as follows:

$$J_{1,it} = 1 \text{ if } C^*_{i,t-1} = 0 \text{ and } C^*_{it} = 1$$

$$J_{2,it} = 0 \text{ if } C^*_{i,t-2} = 0 \text{ and } C^*_{i,t-1} = 1 \text{ and } C^*_{it} = 1$$

[10] As a robustness check, we test whether our elasticity estimate is sensitive to the inclusion of a level-of-assessment variable, as measured by the state equalization rate.

[11] We adopt an approach that is similar to Todd (2006) and Duncombe & Yinger (2007). One downside of the control function approach is that it can be less robust than the standard instrumental variable approach when dealing with models with nonlinear functions of endogenous regressors.

[12] We exclude Tompkins and Nassau Counties, where assessing is conducted by the county itself. A third county, Montgomery, centralized assessment to the county level in 2018, which is outside our sample period.

[13] The New York state Office of Real Property Tax Services (ORPTS) only reports COD for a sample of assessing units that have not conduced revaluation over the three years prior to the market value survey year.

[14] In addition, appendix Tables A4 and A5 list the total as well as personnel and contractual costs of each CAP participant, respectively.

[15] While our main approach is similar to that of Sjoquist and Walker (1999), one must note that we use different model specifications and estimation strategy. Also, while Sjoquist and Walker (1999) uses total assessment cost as part of the dependent variable, our dependent variable is a measure of own source assessment cost that excludes state aid for reassessment.