



Panel Data Analysis: A Guide for Nonprofit Studies

Yuhao Ba¹ · Jessica Berrett² · Jason Coupet¹

Accepted: 24 February 2021 / Published online: 19 March 2021
© International Society for Third-Sector Research 2021

Abstract The growing push in nonprofit studies toward panel data necessitates a methodological guide tailored for nonprofit scholars and practitioners. Panel data analysis can be a robust tool in advancing the understanding of causal and/or more nuanced inferences that many nonprofit scholars seek. This study provides a walk-through of the assumptions and common modeling approaches in panel data analysis, as well as an empirical illustration of the models using data from the nonprofit housing sector. In addition, the paper compiles applications of panel data analysis by scholars in leading nonprofit journals for further reference.

Keywords Panel data analysis · Fixed effects · Random effects · Dynamic model · Arellano–Bond

Introduction

Research using panel data has been proliferating in numerous social science fields at an accelerated pace (Frees, 2004; Halaby, 2004; Zhu, 2013).¹ This is a trend that deserves attention and further facilitation in nonprofit scholarship. Here, a panel dataset is one that follows the same sample of subjects such as individuals, organizations, or communities over multiple time periods. Panel data can provide researchers with both time-series and cross-

sectional information on each subject in the sample (Hsiao, 2014).² For instance, using a sample of fifty nonprofits' service delivery records over five years, on the cross-sectional dimension, we can have comparative information about how these nonprofits are doing in their service delivery relative to their peers in a particular year. On the time-series dimension, we get to observe changes in variables such as numbers of fundraising events and community visits within individual nonprofits over the five-year time span.

Compared with cross-sectional approaches, panel data analysis can be a compelling approach to the generation of potentially more advanced information and knowledge (Zhu, 2013). This is because the time-series dimension of panel data can help reveal interesting and valuable findings that are otherwise unavailable in cross-sectional samples. Specifically, when applied properly, panel data analysis

¹ Regarding different types of data, in general, time-series data refers to a collection of observations made chronologically on one subject, for instance, historical stock prices of one firm. In this case, in order to have enough variation, the length of a time series is an important determinant of the usefulness of such data, panel data, as we mentioned in the paper, entail observations made on the same sample of subjects over time. The difference here is in the data structure. Compared with time-series data, panel data have more than one cross-sectional subject. Accordingly, both its length (i.e., number of time periods) and its width (i.e., number of cross-sectional subjects) jointly decide the usefulness of a panel dataset. Additionally, some studies also use the term time-series cross-sectional (TSCS) data to refer to panel data (Lewis-Beck et al., 2004), while some others suggest that TSCS data have comparatively fewer cross-sectional subjects than panel data (see Bell & Jones, 2015). As for the term longitudinal data, researchers often use it interchangeably with the term panel data (Frees, 2004).

² Here, a necessary yet challenging sampling strategy in panel data analysis is to keep track of the same sampled subjects until the completion of data collection, as attrition of subjects out of the sample could lead to incorrect inferences (Baltagi, 2008).

✉ Yuhao Ba
yba@ncsu.edu

¹ Department of Public Administration, North Carolina State University, Campus Box 8102, Raleigh, NC 27695, USA

² School of Public Affairs, University of Colorado–Colorado Springs, Colorado Springs, CO 80918, USA

can afford us the opportunity to observe within-subject changes in explanatory variables. In other words, since comparisons are made on the same subjects over time, panel data analysis can help us control for the unobservables that do not change over time but threaten the validity of the inferences we make about relationships that we are interested in. Further, panel data analysis can more convincingly model causal relationships (Hsiao, 2007), an advance from the preponderance of correlations that dominate the nonprofit literature.

Specifically, for nonprofit scholars and practitioners, panel data analysis can help address many theoretical and/or empirical questions. This is particularly the case when there are unobservables in the sampled subjects that rarely change over time but can lead to problematic conclusions. For instance, at the organizational level, when deciding the effect of public funding on the efficiency of nonprofits, as mentioned previously, with panel data, we are able to control for unobservables such as organizational design and culture that do not change over time. At the individual level (e.g., volunteers and donors), individual characteristics such as personality and values that are in many cases unobservable and yet stay relatively stable, can be controlled by panel data. Along this line, at the country/nation level, panel data can help researchers control for such unobservables including, for example, political systems, market structures, and cultures.

Despite its advantages, application of panel data analysis remains relatively uncommon in nonprofit studies. For instance, a review of all studies published in the *International Journal of Voluntary and Nonprofit Organizations (VOLUNTAS)*, *Nonprofit Management & Leadership (NML)*, and *Nonprofit and Voluntary Sector Quarterly (NVSQ)* between 2009 and 2018 shows that only 3.43% (52/1517) use panel data analysis.³ This can be attributed to the limited number of panel datasets built for nonprofit studies given the nascent state of the field. This is changing, however, as researchers and other advocates of nonprofit studies ramp up efforts to build representative nonprofit panel data (Faulk & Derrick-Mills, 2015).

Another important reason may be the lack of an accessible guide for panel data analysis tailored for nonprofit scholars and practitioners. For instance, in related fields including public administration and political science, a variety of panel datasets have been built covering topics such as budgeting and finance (see Diebold & Coggburn,

2018), public service and performance (see Jensen & Vestergaard, 2016), and policy analysis (see Yi & Chen, 2019) while only a few can be found in nonprofit studies, concentrating on topics such as financial management and performance (see Calabrese, 2011; Mayer et al., 2014) and cross-sector interaction and comparisons (see Amir-khanyan, Kim, & Lambright, 2008; Kim, 2015). In like manner, field-specific methodological guidance for panel data analysis has been developed in public administration (see Zhu, 2013) and political science (see Plümper et al., 2005), yet such efforts remain absent in the nonprofit literature. While knowledge and insights are deemed to be intersectional and mutualistic among these fields (Pandey & Johnson, 2019), we believe a nonprofit-specific guide is essential to help nonprofit scholars and practitioners better understand and apply panel data analysis with data and context that they are familiar with.

When researchers use panel data analysis in nonprofit studies (for example, Bromley et al., 2018; Coupet, 2018; Galaskiewicz & Bielefeld, 1998; Szper, 2013), their modeling approaches vary based on their respective research contexts and data. While they each provide theoretical and methodological guidance, the variance across these studies makes it difficult to compare and comprehend common panel data modeling approaches in a unified context. This is particularly the case for scholars and practitioners who are less experienced in statistical modeling. An accessible one-stop review of common panel data modeling approaches is thus critical for the consistency of future nonprofit research.

Application of panel data analysis varies based on one's research question(s) as well as the nature of the data. In general, two types of modeling approaches exist in panel data analysis: linear models and non-linear models, depending on the linearity in the regression parameters. In practice, the difference can be simplified by observing one's dependent variable, where linear models are used for continuous dependent variables and non-linear models for discrete ones such as binary or categorical responses. While some outcomes of interest in nonprofit studies are categorical (e.g., individuals' decision to engage in volunteer work), we decide to focus our discussion on linear panel models for two reasons. First, linear models have relatively broader application in social sciences (Rao & Toutenburg, 1995), and emphasizing linear models has the added advantage of illustrating different modeling approaches with a consistent model specification. Second, the intuition of panel data analysis in the context of nonprofit studies is explained by means of linear panel data models for heuristic purposes because they are relatively less complex than non-linear ones. Most of the arguments in favor of using panel data—such as the opportunity to control for time-invariant unobservables and to adjust for

³ Here, we acknowledge that it is common for nonprofit scholars to choose other publication outlets such as *Journal of Public Administration Research and Theory* (e.g., Cheng, 2018; de Wit & Bekkers, 2017) and *Public Performance & Management Review* (e.g., Pandey & Johnson, 2019). Given the broad coverage of these journals, however, we decided to narrow our focus on *VOLUNTAS*, *NML*, and *NVSQ*, which focus solely on nonprofit studies.

temporal dependence—apply to linear and non-linear models alike.

The article proceeds in three sections. The first section provides an accessible walk-through of the assumptions and common modeling approaches in linear panel data analysis. This is followed by an illustrative application of each model in a nonprofit housing setting using the IRS Form 990 data from 2009 to 2016 provided by GuideStar. In this example, we employ a “crowding out” theoretical model to analyze the effect of government grants and program service revenue on donations. The paper concludes with a discussion of the implications of the modeling approaches presented, an overview of non-linear panel data models as well as other designs for causal inferences, and suggestions for future research. In addition, we provide a compilation of applications of linear panel data analysis by nonprofit scholars from leading nonprofit journals as well as their datasets for further reference (see “Appendix” 1).

Linear Panel Data Models: A Walk-Through

Panel data analysis can arguably be seen as “a marriage of conventional regression analysis and time-series analysis” (Frees, 2004, p. 2), in which both cross-sectional (between subjects) and time-series (within subjects) variation can be addressed simultaneously. Depending on one’s research design, application of panel data analysis entails choosing from a series of specific modeling approaches. For instance, regarding controlling for the previously mentioned time-invariant unobservables that could lead to problematic claims for an argument, based on different assumptions, we could adopt different panel data modeling approaches such as random effects and fixed effects models. Understanding these choices requires a walk-through of the commonly used panel data modeling approaches. In this section, we present four linear panel data models: classical linear regression model, random effects model, fixed effects model, and dynamic panel modeling. Our walk-through includes their respective assumptions, applications, and limitations. The presentation is designed to be accessible with a minimum use of technical terms and equations (For equations, please see “Appendix” 2).

Classical Linear Regression Model

In linear regressions, ordinary least squares (OLS) regression is one of the most widely used modeling techniques and arguably the default approach to regression analysis for many scholars and practitioners. Even in studies that rely on more complex techniques, it often remains the departure point for the full analysis (Greene, 2003). Specifically,

OLS regression is used to model the linear relationship between a continuous dependent variable and one or more independent variables in a static setting that has no time-series variation. The logic behind OLS regression is to draw one best-fit regression line to capture the approximate linear relationship between the dependent variable and the independent variable(s) where the sum of squared residuals—the differences between observed values and predicted values of the dependent variable—can be minimized (Hutcheson & Sofroniou, 1999).

While OLS regression can be applied to many linear modeling situations, it is essential to review its key assumptions to avoid any misuses. First, an OLS model should be linear in parameters (i.e., no transformation in parameters). This is fundamental for your model specification. Violation of this assumption will make your estimated coefficients and standard errors unreliable. Second, no meaningful multicollinearity, which means that the independent variables specified in an OLS model should be relatively independent from each other and do not correlate to a significant level.⁴ Third, the expected value of the error terms should be zero across observations and should not be a function of, or correlated with, the independent variables specified. The presence of such correlation will make the specified independent variables endogenous and accordingly lead to problematic regression estimates. Fourth, homoscedasticity, which means that the error terms in an OLS regression should share the same variance across observations. This is important because when calculating coefficient estimates, the process assigns equal weight to individual error terms associated with each observation. If the variance of the error terms is not constant across observations, while the coefficient estimates will remain unbiased, their associated estimated standard errors will be biased, and hence the conclusions about their statistical significance will be incorrect. Fifth, the errors should be normally distributed. Here, in addition to assumptions 3 and 4, normality is assumed because if a model is properly specified, most predictions will be close to correct (i.e., errors approach 0) with a few high and low estimates. The error terms are thus assumed to form a normal distribution. This assumed normal distribution also helps to accommodate cases where explicit assumption of the error term is required such as in stochastic frontier analysis (see Coupet & Berrett, 2019) or for testing purposes (Greene, 2003). Lastly, random sampling of observations. This is a key

⁴ In practice, the presence of different levels of multicollinearity is common given the implicit and/or explicit interconnectedness of many variables in social sciences. The disadvantage of having multicollinearity is that it could lead to relatively large estimated standard errors for the coefficient estimates of independent variables that are correlated with others, which would undermine their statistical significance (Allen, 1997).

assumption when using samples to make inferences about their respective populations. That said, in practice, statistical estimates originating in non-randomly sampled data are also common and meaningful. For example, when using city-level information to analyze the relationship between city government support and nonprofit service delivery performance, the data collection would likely lead to a non-random sample given the lack of information on certain cities, particularly the smaller ones. Inferences made based on such samples can still be meaningful (e.g., generalizing to larger cities only) as long as the researchers acknowledge this limitation in their data. Taken together, violation of assumption 1 will lead to both problematic estimated coefficients and problematic standard errors, whereas violation of assumptions 2–5 will result in problematic standard errors. As for assumption 6, it makes sure that your estimated relationship can meaningfully generalize to the populations of interest.

Applying OLS regression to panel data creates a pooled model in which all observations from the sampled units at different time periods are pooled together and being treated as independent from one another. This pooling approach can lead to problematic inferences, however, as it ignores the internal structure of panel data (Bartels, 2015). That is, due to the two dimensions of information that panel data include (i.e., cross-sectional and time-series dimensions), panel data have a two-level hierarchical structure in which repeated measures observed at different time periods (i.e., level 1) can be considered nested in their respective sampled units (i.e., level 2; Bell & Jones, 2015). In this case, observations of a given unit at different time periods are likely to be related to each other. Inferences made based on this complete pooling approach without regard to the hierarchical structure of panel data can thus be problematic. This is because the effective sample size of a panel dataset can be much smaller than a complete pooling approach would assume, which would normally lead to underestimated standard errors (Bell & Jones, 2015).

More importantly, a complete pooling approach would assume no correlation between the error terms and the independent variables specified in a model (assumption 3). Yet in reality, such an assumption rarely holds given the potential omission of factors that can both account for the variation in the dependent variable and remain correlated with the independent variables observed (Greene, 2003; Wooldridge, 2013; Zhu, 2013). This is a pervasive issue in cross-sectional analysis and indeed a major motivation for using panel data analysis (Arellano, 2003). For instance, assume we are interested in the effect of employees' received monetary incentives on nonprofit service delivery outputs and we have a dataset that has such information on a group of nonprofits over multiple years. A pooling approach would lead us to incorrect conclusions because

we fail to control for these nonprofits' incentive structures that are not observed. Here, incentive structures can correlate with both nonprofits' service delivery outputs and their employees' received monetary incentives. Given this, such unobserved differences across sampled units that have correlations with both the dependent variable and the independent variable(s) can significantly misguide our analysis (Rosenbaum, 2005). In statistics and econometrics, these unobserved differences are termed unobserved heterogeneity (see Arellano, 2003).

Regarding these unobserved differences, the literature has identified three types of driving factors: (1) unit-specific time-invariant factors, (2) time-specific unit-invariant factors, and (3) unit- and time-varying factors (Hsiao, 2014; see Fig. 1). Here, unit-specific time-invariant factors are those that vary across sampled units but remain constant for the same units across time periods. Examples include demographic information such as gender and racial identity at the individual level, and culture and management style at the organizational level. Time-specific unit-invariant factors are those that are the same for all the sampled units in a given time period but vary across time periods. Examples include the broad political and economic environment such as governorship and presidency in a certain time period, and specific measures such as utility and tax rates. Lastly, unit- and time-varying factors are those that vary across both sampled units and time periods, such as individual and organizational physical and economic wellbeing.

In this case, variation driven by the first two types of factors concentrates on the cross-sectional dimension and the time-series dimension of panel data, respectively, while the last one varies on both dimensions. In practice, panel data modeling techniques primarily deal with the first two types of factors and assume that factors vary on both dimensions are independent from the variables specified in the model and will not bias the estimated results (Cameron & Trivedi, 2010; Frees, 2004; Hsiao, 2014). In other words, panel data modeling is better suited to control for unit-specific time-invariant or time-specific unit-invariant unobserved heterogeneity. Taking unit-specific time-invariant unobserved heterogeneity as an example, panel data modeling removes such unobserved heterogeneity and focuses on within-unit variation of the variables that we are interested in to provide explanations. In this case, within-unit variation of the key explanatory variable(s) of interest is essential to make valid inferences when using panel modeling to control for unit-specific time-invariant unobserved heterogeneity, and a lack of such variation will lead to non-generalizable estimations.

For instance, in our earlier example of the effect of employees' received monetary incentives on nonprofit service delivery outputs, assume the unit of analysis is at

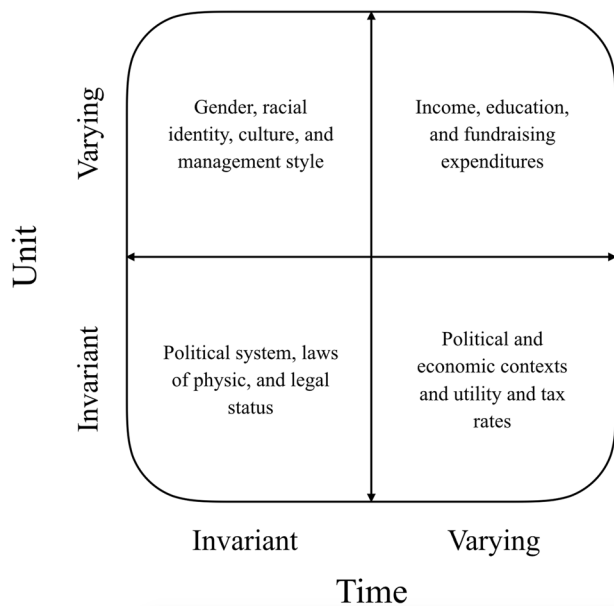


Fig. 1 A two-by-two matrix view of factors driving unobservable variation

the organizational level and we have a panel dataset that has each nonprofits' annual total employees' received monetary incentives and service delivery outputs for several years. If we introduce a nonprofit-specific term to control for their incentive structures, which we assume do not change over time, then the potential effect of incentive structures can be teased out from the original error term, and accordingly the risk of omitting such unit-specific time-invariant unobserved heterogeneity can be reduced. In this case, we can control for the incentive structures—assuming they do not change over time—without having data on them. Again, it is important to acknowledge that the within-unit variation of the key explanatory variable in this example—employees' received monetary incentives—is critical to make proper estimations of its effect on nonprofit service delivery outputs.

Thus, one major advantage of panel data models lies in their ability to control for unobservable differences without actually observing them (Arellano, 2003). Specifically, in dealing with those unobservable differences, or more precisely the potential correlation between the error term and the variables specified in a model, a correlation caused by some unit-specific time-invariant and/or time-specific unit-invariant factors, one common approach is to introduce a categorical term that is specific to individual units and/or time periods covered in the sample. In this case, the newly introduced categorical term(s) can function as an intercept(s) in our models to address those unobservable differences (Garson, 2013; Halaby, 2004). In practice, this can be simply understood as splitting the original error term into two parts. The first part remains random by absorbing

all the uncorrelated differences. The other part, which is specific to individual units or time periods, controls for the unobservable differences that are constant for the same units over time or the same for all units at a given time period but vary across periods, respectively.

Random Effects Model

In introducing this categorical term, two approaches exist based on the way we deal with it, namely, random effects and fixed effects. The differentiation can be challenging, in terms of both the nomenclature and the underlying theory. In general, random effects models assume that the newly introduced categorical term is not correlated with the independent variables specified while fixed effects models assume such correlation (Frees, 2004; Greene, 2003; Halaby, 2004; Hsiao, 2014). Here, while this newly introduced categorical term can represent variation caused by unit-specific time-invariant or time-specific unit-invariant factors, we focus on unit-specific time-invariant for the following walk-through and illustration. This is because in practice, researchers in nonprofit studies mostly are confronted with datasets that have a large number of cross-sectional units but a small number of time periods (i.e., wide panels). In this case, cross-sectional differences appear to be the major concern regarding the validity of the estimated coefficients. Examples include variation in volunteering behavior across 120 US metropolitan areas over three years (Rotolo et al., 2014), differences in types of noncash gifts to encourage charitable giving in 1,055,917 nonprofits across seven years (James III, 2018), and political participation and happiness of 5500 British households across 18 waves (Winters & Rundlett, 2015).

As aforementioned, in random effects (RE) models, the newly introduced unit-specific terms are assumed to be uncorrelated with the independent variable(s). That is, these terms, while formulated, can be considered independent from the cross-sectional units that are pooled in the sample in all time periods (Greene, 2003; Zhu, 2013). This is considered appropriate if the sampled cross-sectional units are drawn from a very large population (Greene, 2003), so the terms introduced can be assumed random across the sampled units yet still represent the characteristics of the underlying population. For instance, in their analysis of the determinants of the size of the nonprofit sector in the USA, Bae and Sohn (2018) use a series of RE models because those models use county-level data from the state of Indiana to approximate the average level in the U.S. The reasoning is that Indiana is ranked 28th in the number of nonprofits per 10,000 persons in all states in their analysis, and likewise can function as a proxy of the average USA level. In like manner, Yu (2016) uses RE models with a sample of 678 AIDS nonprofits to analyze

the growth of AIDS nonprofits in China, and the sample itself is considered to represent the entire population of the Chinese AIDS nonprofits.

The advantage of an RE model is that it could reduce the number of parameters to be estimated, since the categorical term introduced does not have to be specific to the cross-sectional units captured in the sample (Arellano, 2003). Accordingly, RE models can have an increased degree of freedom and a higher level of estimation efficiency (lower standard errors and more statistical significance), particularly with samples that have a large number of cross-sectional units. The drawback of RE models is that any undetected correlation between the unit-specific unobservable differences and the independent variables specified in the model can lead to inconsistent estimation, and by extension problematic inferences about the relationships that we are interested in. Therefore, maintaining the randomness of the random yet unit-specific categorical term is key to the successful application of RE models. In other words, prior to using RE models, substantive considerations regarding the relationship between a sample and its corresponding population, as well as its theoretical appropriateness, are recommended in addition to the statistical assumptions.

Fixed Effects Model

Unlike RE models, fixed effects (FE) models assume correlation between the newly introduced unit-specific term and the variables specified in the model. In other words, FE models relax the assumption of randomness of the unit-specific unobservable differences across the sampled units, an assumption required by RE models. In practice, a FE model deals with these non-random unit-specific unobservable differences by adding a full set of sampled unit-specific dummy variables to the model, given that the correlation between these differences and the variables specified in a model is assumed. For instance, in their analysis of donors' influence on nonprofit long-term product innovation, Ranucci and Lee (2019) use an FE approach with a sample of 247 nonprofit professional theaters in the USA from 2003 to 2013. In order to control for the unit-specific time-invariant unobservable differences, their FE model includes 247 theater-specific dummy variables in the calculating process to produce the estimated coefficients, though these dummy variables are not presented in the final results. Accordingly, the approach of including unit-specific dummy variables can be referred to as a least-squares dummy variable (LSDV) model (Hsiao, 2014).

Due to the relaxed assumption of no correlation between the unit-specific time-invariant unobservable variation and the variables specified in a model—for instance,

nonprofits' incentive structures and employees' received monetary incentives—FE models are widely applied in nonprofit studies. Examples include the impact of historical growth of civil society organizations on educational outcomes in a society (Bromley et al., 2018), the effects of nonprofits' capital campaigns on the fundraising performance of their peers in the same geographic region (Woronkiewicz & Nicholson-Crotty, 2017), and the influence of children's ages and life transitions on their parents' charitable giving and volunteering (Einolf, 2017). Here, while FE models are widely adopted, it is important to acknowledge that the introduction of dummy variables could reduce the degrees of freedom of the specified model, and thereby generate less efficient coefficients. Particularly in large N samples, this could lead to inadequate statistical power for the analysis and a higher risk of multicollinearity (Garson, 2013). That being said, scholars tend to prefer FE models as a relatively safer approach to account for the potential correlation between the unobserved individual effects and the variables specified in a model.

Random Effects Versus Fixed Effects

As mentioned previously, the fundamental distinction between RE and FE models is that FE models assume correlation between unit-specific time-invariant unobservable variation and the independent variables included in a model, while RE models assume no such correlation (Cameron & Trivedi, 2010). In this case, deciding between the two techniques is partially a matter of choosing between efficiency and consistency (Clark & Linzer, 2015), in addition to researchers' theoretical and/or practical considerations. For instance, RE models will be biased if correlation exists between those unobservable variation that is assumed random and the included independent variables. In practice, a simple solution to deciding which type of modeling should be utilized is to rely on the Hausman test. This test generates a statistic that examines the difference between the FE and RE estimators (Hausman, 1978). Specifically, it examines whether the RE estimators are significantly different from the FE estimators, which are known as consistent but less efficient (Wooldridge, 2013). A finding of significance indicates that the RE estimates are not consistent. In other words, there are unobserved variables correlated with the independent variables (Baltagi et al., 2003; Garson, 2013). In this case, FE estimates should be reported.

In the context of nonprofit studies, because unobserved heterogeneity is so likely, scholars will typically be best off using the FE approach. We acknowledge that some scholars may argue differently (Bell & Jones, 2015), but the notion that the individual data are unrelated to the

group or organizational effects captured by random effects is probably unrealistic. For instance, in an analysis involving a large sample of nonprofit organizations, an RE approach would assume that the nonprofit-specific controls (concepts such as organization culture, leadership, identity or size) are uncorrelated with the independent variables of interest. This is unlikely. Figure 2 illustrates a panel model where the dependent variable of interest is nonprofit performance and the independent variable of interest is the number of volunteers. In this simple example, researchers might observe a positive correlation between volunteers and size with an RE model. This relationship, though, might be driven by unobservable variables that are correlated with both the number of volunteers and nonprofit performance, such as size or the capacity to manage volunteers. An FE approach that appropriately controls for unobservable variables might set forth a completely different causal relationship.

Controlling for the relevant individual- or organization-specific characteristics is a common challenge for quantitative nonprofit scholarship, and likewise trying to fit these variables into models can be cumbersome for researchers (Shearer & Clark, 2016). An FE approach has the advantage of controlling for these individual-specific variables as unobservable variables, thus reducing challenges related to finding an appropriate way to specify nonprofit models to emphasize causal inferences.

Dynamic Panel Modeling

In addition to the aforementioned static approaches that highlight unit-specific time-invariant unobservable differences, another line of panel data analysis focuses on temporal dependence. Temporal dependence can be a threat to valid estimation in panel data analysis if independent variables are not contemporaneously associated with the dependent variable (Halaby, 2004; Zhu, 2013). In other

words, if current variation in the dependent variable can be partially explained by its variation in previous periods in a panel setting, focusing on current variation exclusively in model specification could lead to problematic inferences. In the literature, temporal dependence has been highlighted in various areas of research. For instance, organizational theories argue that variability in organizational characteristics can be partially explained by the history or inertia of individual organizations (Hannan & Freeman, 1984). Organizational memory or the “imprinting effects of the past” (Sydow et al., 2009, p. 689), can therefore be influential in shaping organizational behavior in current periods (Beckman & Burton, 2008). In this case, the aforementioned FE and RE models may suffer from biased or inconsistent estimations given the potential temporal dependence in their data (Bun & Sarafidis, 2013). To deal with this, dynamic panel modeling is introduced. Specifically, dynamic panel models assume that the relationship between the dependent variable and the independent variable(s) of interest in the current period partially relies on the values of the dependent variable in previous periods (Cameron & Trivedi, 2010, p. 293).

One common approach to dynamic panel models is to introduce a lagged dependent variable (LDV) as an independent variable to control for the imprinting or historical effects (Halaby, 2004; Zhu, 2013). The LDV approach is considered adequate in adjusting many dynamic processes (Bun & Sarafidis, 2013), including both the long-term and short-term historical effects (Wawro, 2002; Zhu, 2013). The relative effect sizes of the LDV and other independent variables can thus be used as measures to gauge the magnitudes of the historical and current effects of the variables of interest. In practice, the appropriateness of including an LDV largely depends on the question being asked. The primary consideration is whether or not past values of the dependent variable have predictive power over that of the current period. This is common in organizational research. An organization’s budget, for instance, might be a function of last year’s budget. In the literature, Weimar et al. (2015) use a dynamic panel model to examine the external determinants of membership numbers in German sport clubs from 1970 to 2011. Their choice of a LDV is justified by the dynamic nature of club membership growth in that membership growth from the previous year will likely have an effect on membership growth in the current year. Similarly, in their analysis of the impact of revenue diversification on fundraising efficiency, De Los Mozos et al. (2016) take the LDV approach to adjust the inertia in nonprofits’ fundraising efficiency. Here, it is also critical to acknowledge that inclusion of an LDV, though useful, is sometimes inadvisable because of its potential downsizing impact on the explanatory power of other independent variables (Keele & Kelly, 2006). A check on the theoretical

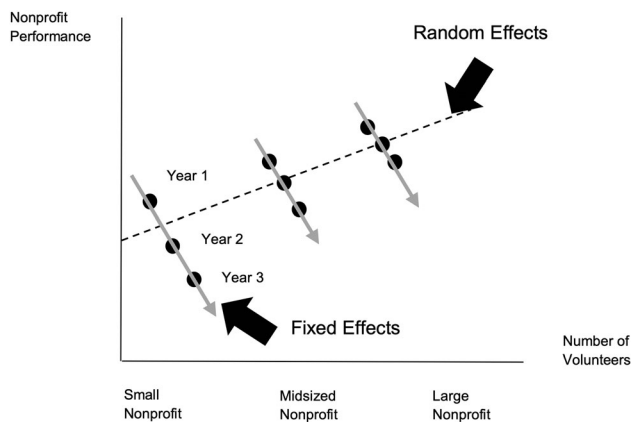


Fig. 2 Random effects versus fixed effects

$$\text{Donations} = F(\text{Government Revenue}, \text{Program Service Revenue}, \text{Fundraising Expenditures})$$

Fig. 3 Model for empirical illustration

and practical appropriateness of this approach is thus recommended as well.

When applying dynamic panel modeling with LDVs, and indeed all the models presented previously (i.e., pooled OLS, RE, and FE models), it is important to pay attention to the stationarity of the panel. This is because, when using data that have variation on the time-series dimension, stationarity ensures that statistical properties of the data are relatively invariant across all time periods. In other words, they do not depend on the time at which they are observed (Witt et al., 1998). In this case, stationarity can help us estimate meaningful relationships as it reduces the likelihood of observing a spurious correlation originating in long-term trends or some other reoccurring patterns in the data. In practice, one can use the augmented Dickey–Fuller (ADF) test to examine whether the panel is stationary, or if the lagged values are relevant to predict the variation over time (Maddala & Wu, 1999). Here, a finding of significance indicates that the panel is stationary and the inclusion of an LDV is statistically appropriate. Further, successful application of dynamic modeling requires that the sample should include enough time periods for model specification. For example, Weimar et al. (2015) have 42 periods in their data to ensure room for statistical analysis. This is particularly the case with fixed effects models, because the inclusion of individual-specific intercepts may lead to inconsistent estimations when the number of cross-sectional units is large and the number of time periods is relatively small (the so-called “Nickell bias”). This is because the demeaning of the dependent variables and the independent variables could lead to a correlation between the observables and the errors (Arellano, 2003; Nickell, 1981). In practice, researchers tend to run both models to compare the consistency of the results for further robustness (see Cheng, 2018).

Empirical Illustration

To illustrate the models presented previously, we develop a “crowding out” theoretical model to estimate the effect of government grants and program service revenue on donations. There have been many studies that attempt to tease out crowd-out effects. While some use experimental approaches (Jilke, Lu, Xu, & Shinohara, 2019), most use administrative data and the results can vary drastically. Some studies find that government grants decrease donations, while others argue that government grants increase donations, and still others find no significance (de Wit & Bekkers, 2017). Crowd-out studies have also used many

different approaches including OLS (Brooks, 2000), FE (Nikolova, 2015), and dynamic panel models (Heutel, 2014).

The difference in findings might be due in part to the difference in analytical approaches. To illustrate, we use a sample of Habitat for Humanity affiliates across the USA. Habitat for Humanity is a collection of nonprofit organizations that share the mission of “Seeking to put God’s love into action, Habitat for Humanity brings people together to build homes, communities, and hope” (Habitat for Humanity, 2020). The data come from the IRS Form 990 between 2009 and 2016. Our original sample included an unbalanced panel of 1120 Habitat for Humanity affiliates. Data missing from the key independent and dependent variables in our model were addressed with listwise deletion, resulting in a sample of 791 affiliates. The Arellano–Bond model uses an additional year of data to construct instruments, further reducing the sample size for Model 4. In this example, we estimate the effect of government revenue and program service revenue, which are alternative sources of nonprofit revenue, on donations (see Fig. 3). In our data, the dependent variable of total donations represents the total amount of donations, both cash and non-cash, received by one affiliate in a fiscal year. The independent variables include government revenue, program service revenue, and fundraising expenditures. Government revenue represents income streams from government grants and contracts, some of which are from the US Department of Housing and Urban Development (HUD) (Habitat for Humanity, 2014). Program service revenue is the income generated from services offered by Habitat affiliates, and fundraising expenditures represent the amount each affiliate spent on fundraising. All variables are in natural logs.

We estimate the theoretical model in Fig. 3 with the four approaches presented previously: OLS, RE, FE at the affiliate level (unit-specific), and a dynamic panel model.⁵

⁵ Here, it is important to acknowledge that directly adding an LDV as an independent variable could violate the OLS assumption that independent variables do not correlate with the error term. This is because the error term is supposed to encapsulate all the variation that is left in the dependent variable but are not explained by the independent variable(s). An LDV is thus likely to correlate with such remaining variation as it represents the variation in the DV that is from the previous period. This issue would be further complicated if the model specification includes panel-specific error terms such as FE terms (Zhu, 2013). In this case, we do not recommend directly adding an LDV to account for the dynamics process. We introduce this approach because it is our intention to provide the basic assumptions and modeling approaches that are involved in panel data analysis. To deal with such potential correlations that an LDV might induce, however, researchers could rely on instrumentation techniques such as a second-order LDV (the DV that is two periods before the current one; Anderson & Hsiao, 1981) or the generalized method of moments (GMM) estimator proposed by Arellano & Bond (1991).

Table 1 Comparison of pooled OLS, random effects, fixed effects, and dynamic model

Variables	(1) Pooled OLS	(2) Random effects	(3) Fixed effects	(4) Dynamic model
Contributions (lag)	–	–	–	0.129** (0.060)
Government revenue	0.001 (0.004)	– 0.009*** (0.003)	– 0.016*** (0.003)	– 0.012*** (0.003)
Program service revenue	0.285*** (0.025)	0.159*** (0.018)	0.044** (0.019)	0.029 (0.018)
Fundraising expenditures	0.385*** (0.020)	0.260*** (0.017)	0.113*** (0.0120)	0.047** (0.019)
Constant	5.108*** (0.254)	7.852*** (0.257)	10.980*** (0.317)	10.23*** (0.796)
R-squared	0.571	–	0.540	–
Year FE	Yes	Yes	Yes	Yes
Observations	3976	3976	3976	2431
N	791	791	791	648

Robust standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 1 displays the results of the four different estimation approaches. Recall that we are most interested in the effect of government and program service revenues on donations. In the OLS regression model, we clustered the standard errors by affiliate as recommended in pooled regression approaches. Note that government revenue seems to have no effect on donations, but that donors seem to reward service revenue. A 1% increase in program service revenues seems to result in a 0.28% increase in donations. Fundraising expenses also contribute to increasing donations.

As mentioned, pooled OLS is less appropriate in this case. This is because there are omitted variables that we know could affect donations, such as the organizational culture of the affiliate, which are unique to the organization and do not vary much over time. Comparing the OLS estimator with the RE and FE estimators sheds some light on just how pronounced the omitted variable bias is. When random effects are introduced, crowding out with regard to government revenues emerges, but the practical effect is low (less than a 0.01% marginal effect), and the coefficient for program service revenue almost halves ($\sim 0.16\%$). Our organization-level FE model does not change the coefficient on government revenues much, but the positive effect of service revenue almost drops out.

As mentioned previously, the assumption that an RE model makes about the nonprofit specific random effects term being uncorrelated to other covariates can bias the estimates. The difference between the FE and RE coefficients for program service revenue suggests that the bias in this case is quite pronounced. We tested this with a Hausman test ($\chi^2(9) = 180.47$, $p < 0.000$), rejecting the null hypothesis that there are no systemic differences between the coefficients. In the OLS and RE models, the coefficient of program service revenues is likely highly

correlated with the RE term (the difference between the mean number of donations and that of each affiliate). It is not surprising that these are correlated. Larger nonprofits, for example, would have both higher program revenues and higher donations. We posit that RE models are rarely, if ever, appropriate for organizational research. It is difficult to think of a scenario that an organization level random effects term would be uncorrelated with the covariates of interest. The presence of one influential omitted variable is all it takes to lead us to incorrect conclusions.

Lastly, when we add the lagged dependent variable, the difference in the program service revenue coefficient is decidedly less pronounced, but the statistical significance drops out. The crowd-in effect of program service revenues is now statistically indistinguishable from zero. Additionally, the model shows that current donations are, in part, predicted by past donations. There is still a very small, but statistically significant, crowd-out effect on government revenues. The effect of fundraising expenses decreases significantly as well.

Discussion and Conclusion

Panel data analysis is one of the most active and promising lines of research in econometrics owing to its advantage in enabling the development of more advanced estimation techniques and theoretical inferences (Greene, 2003). Accordingly, panel data modeling has been increasingly advocated and applied in various social science fields for enhanced quantitative inferences (see Frees, 2004; Halaby, 2004; Plümper et al., 2005; Zhu, 2013). In an effort to facilitate the application of panel data analysis tools in nonprofit studies, we offer this article as an accessible guide for both scholars and practitioners. This means that

we are not seeking to provide advancements in those techniques introduced, but to offer an approachable reference for novice panel data consumers. Specifically, we first contribute a simplified walk-through of four common linear panel data models with examples from leading nonprofit journals. This can be advantageous considering the growing availability of panel data as well as the continued proliferation of the nonprofit literature. The empirical illustration in the context of nonprofit financial management helps to further demonstrate the application of the linear panel models introduced. For researchers of quantitative nonprofit inquiry, our analysis presents a repertoire of linear panel data modeling techniques in broad brushstrokes. Additional information and extended discussion of the methods can be found in most of the cited articles and textbooks mentioned throughout this paper.

To summarize the process of working with linear panel data and the steps one should take in analyzing such data, we can think of this as a decision tree path (see “Appendix” 3 for an illustration). One could use a simple OLS regression model if the assumptions are met. However, in a case where we assume the errors are correlated with cross-sectional entities, we can address this concern by clustering these errors. Specifically, we can use pooled OLS with the assumption that temporal effects are absent and that the regression coefficients and the intercepts remain the same for each time period. Violation of the above assumption(s) leads one a step further, to utilize either a RE or FE model. Random effects modeling assumes that there is random unit-specific unobservable variation. However, in most nonprofit data we would assume that there is correlation between such variation and the specified independent variables. If this is assumed, then one could use fixed effects. As a final step, one would want to assess whether the dependent variable can be partially explained by the history, which indicates that the dependent variable(s) from the previous year(s) becomes an independent variable in the current year, particularly when there are enough time periods in the sample (and thus a relatively lower risk of Nickell bias). If the risk of Nickell bias is low and the research question permits, one would want to use a dynamic panel model.

For further reference, we offer a compilation of applications of linear panel models as well as their datasets between 2009 and 2018 in three leading nonprofit journals, including *VOLUNTAS*, *NML*, and *NVSQ*. The development of the compilation is deemed to be exhaustive to the extent possible and with rigor. In doing so, we follow a three-phase methodological design introduced by Maier et al. (2016), which consists of the process of reviewing existing studies in a more holistic manner (see “Appendix” 4). Out of 1,517 articles and research notes published, 30 apply linear panel data analysis (21 with fixed effects analysis, 8

with random effects analysis, and 7 with dynamic modeling; 6 use two approaches in one study). We hope this list can serve as the basis for further comprehension of fundamental principles of linear panel data analysis, since digesting existing research can be as powerful a means of advancing our collective knowledge as launching new studies (Cooper, 1989; David & Han, 2004; Light & Pillemer, 1984). The list can likewise facilitate the exploration of new research and/or application opportunities in nonprofit studies.

In addition to the four common linear panel models presented, it is important to acknowledge other panel analytical approaches. For instance, in nonprofit studies, researchers are likely to encounter situations where multi-level data and modeling are needed to deal with questions such as interaction between headquarters and local organizations in multisite nonprofit systems (Grossman & Rangan, 2001) as well as the impact of contextual-level factors such as social networks and generalized trust on individuals’ charitable giving and volunteering behavior (Glanville et al., 2015). In such cases, aside from the unit-specific time-invariant and time-specific unit-invariant models mentioned previously, which focus on one level, more advanced multi-level modeling such as spatial panel data models (Elhorst, 2003) should be utilized (Zhu, 2013). Furthermore, in RE models, we assume the newly introduced unit-specific terms are drawn from a random distribution. In other words, we introduce a random intercept in RE models. If both coefficients and intercepts are random, however, we can utilize random coefficients (RC) models. If sufficient variation presents in data, RC models can generate more efficient estimations (Wooldridge, 2005).

More broadly, while panel data analysis proves advantageous in various situations, it is important to acknowledge that it is not a panacea. There are downsides to panel data analysis. Operationally, building a valid panel data set can be costly due to the effort needed to cover both of the dimensions. Methodologically, the requirements and assumptions of different modeling approaches may complicate application of panel data analysis, which is why we develop this analysis as an accessible reference for how the common models can be utilized. In addition, prior to any modeling approaches, a check on the theoretical and substantive appropriateness of the research and its empirical design is always recommended.

Potentially fruitful points of departure for future analysis include a detailed articulation of other panel modeling approaches, such as multi-level modeling, RC models, difference-in-difference and group fixed effects, as well as working with interaction terms and lagged variables. A paralleled illustration of modeling techniques with categorical and limited dependent variables would benefit the methodological advancement of the field as well. Further,

building upon our short list of applications of panel data modeling in nonprofit research, a more systematic review of these studies is recommended to gauge and advance the state of the field, particularly regarding the common issues and biases in model selection and the consistency of results (see Halaby, 2004). Lastly, in addition to the STATA commands we provide in “Appendix” 5, a repertoire of other software or programming languages applying all the aforementioned modeling techniques should be prepared for scholars and practitioners with different tool preferences.

Funding This study was not funded by any funding/grants.

Declarations The authors comply with the journal’s ethical standards.

Conflict of interest The authors declare that they have no conflict of interest.

Appendix 1

See Table 2.

Table 2 Articles identified with linear panel data modeling

Article title	Author(s) and date	Journal	Method
Membership in nonprofit sport clubs: a dynamic panel analysis of external organizational factors	Weimar, D., Wicker, P., & Prinz, J. (2015)	<i>NVSQ</i>	Dynamic paneling
The accumulation of nonprofit profits: a dynamic analysis	Calabrese, T. D. (2011)	<i>NVSQ</i>	Dynamic paneling
Exploring the link between government funding and efficiency in nonprofit colleges	Coupet, J. (2018)	<i>NML</i>	Dynamic paneling
Recruitment and retention of referees in nonprofit sport organizations: the trickle-down effect of role models	Wicker, P., & Frick, B. (2016)	<i>VOLUNTAS</i>	Fixed effects + dynamic paneling
The impact of revenue diversification on expected revenue and volatility for nonprofit organizations	Mayer, W. J., Wang, H. C., Eggington, J. F., & Flint, H. S. (2014)	<i>NVSQ</i>	Dynamic fixed effects (Arellano–Bond)
The relationship of nonprofits’ financial health to program outcomes: Empirical evidence from nonprofit arts organizations	Kim, M. (2017)	<i>NVSQ</i>	Fixed effects
The effects of capital campaigns on other nonprofits’ fundraising	Woronkiewicz, J. & Nicholson-Crotty, J. (2017)	<i>NML</i>	Fixed and random effects
Organizing for education: a cross-national, longitudinal study of civil society organizations and education outcomes	Bromley, P., Schofer, E., & Longhofer, W. (2018)	<i>VOLUNTAS</i>	Fixed effects
Parents’ charitable giving and volunteering: are they influenced by their children’s ages and life transitions? Evidence from a longitudinal study in the united states	Einolf, C. J. (2017)	<i>NVSQ</i>	Fixed effects
Testing an economic model of nonprofit growth: analyzing the behaviors and decisions of nonprofit organizations, private donors, and government	Kim, Y. H., & Kim, S. E. (2016)	<i>VOLUNTAS</i>	Fixed effects
Principals and agents: an investigation of executive compensation in human service nonprofits	Nikolova, M. (2014)	<i>VOLUNTAS</i>	Fixed effects
Government funding of private voluntary organizations: is there a crowding-out effect?	Nikolova, M. (2015)	<i>NVSQ</i>	Fixed effects
The heterogeneity of competitive forces: the impact of competition for resources on United Way fundraising	Paarlberg, L. E., & Hwang, H. (2017)	<i>NVSQ</i>	Fixed effects
Volunteering in the United States in the aftermath of the foreclosure crisis	Rotolo, T., Wilson, J., & Dietz, N. (2015)	<i>NVSQ</i>	Fixed effects
The challenges of untangling the relationship between participation and happiness	Winters, M. S., & Rundlett, A. (2015)	<i>VOLUNTAS</i>	Fixed effects + dynamic paneling
Faith-based assumptions about performance: Does church affiliation matter for service quality and access?	Amirkhanyan, A. A., Kim, H. J., & Lambright, K. T. (2009)	<i>NVSQ</i>	Fixed effects
The relationship between confidence in charitable organizations and volunteering revisited	Bekkers, R., & Bowman, W. (2009)	<i>NVSQ</i>	Fixed and random effects
What determines the size of the nonprofit sector?: a cross-country analysis of the government failure theory	Matsunaga, Y., Yamauchi, N., & Okuyama, N. (2010)	<i>VOLUNTAS</i>	Fixed effects

Table 2 continued

Article title	Author(s) and date	Journal	Method
What accounts for the variations in nonprofit growth?: a cross-national panel study	Kim, Y. H., & Kim, S. E. (2018)	<i>VOLUNTAS</i>	Fixed effects
Understanding donor-advised funds: how grants flow during recessions	Heist, H. D., & Vance-McMullen, D. (2019)	<i>NVSQ</i>	Fixed effects
Cash is not king for fund-raising: gifts of noncash assets predict current and future contributions	James III, R. N. (2018)	<i>NML</i>	Fixed effects
Online or offline? Nonprofits' choice and use of social media in Hong Kong	Lam, W. F., & Nie, L. (2019)	<i>VOLUNTAS</i>	Fixed effects
Donor influence on long-term innovation within nonprofit organizations	Ranucci, R., & Lee, H. (2019)	<i>NVSQ</i>	Fixed effects
The effects of capital campaigns on local nonprofit ecologies	Woronkowicz, J. (2018)	<i>NVSQ</i>	Fixed effects
Doing well by returning to the origin: mission drift, outreach and financial performance of microfinance institutions	Pedrini, M., & Ferri, L. M. (2016)	<i>VOLUNTAS</i>	Fixed and random effects
The relationship between philanthropic foundation funding and state-level policy in the era of welfare reform	Mosley, J. E., & Galaskiewicz, J. (2015)	<i>NVSQ</i>	Random effects + dynamic paneling
Resource dependence in non-profit organizations: is it harder to fundraise if you diversify your revenue structure?	De Los Mozos, I. S. L., Duarte, A. R., & Ruiz, Ó. R. (2016)	<i>VOLUNTAS</i>	Random effects
The effects of resources, political opportunities and organisational ecology on the growth trajectories of AIDS NGOs in China	Yu, Z. (2016)	<i>VOLUNTAS</i>	Random effects
Playing to the test: organizational responses to third party ratings	Szper, R. (2013)	<i>VOLUNTAS</i>	Random effects
Factors contributing to the size of nonprofit sector: tests of government failure, interdependence, and social capital theory	Bae, K. B., & Sohn, H. (2018)	<i>VOLUNTAS</i>	Random effects

Appendix 2

See Table 3.

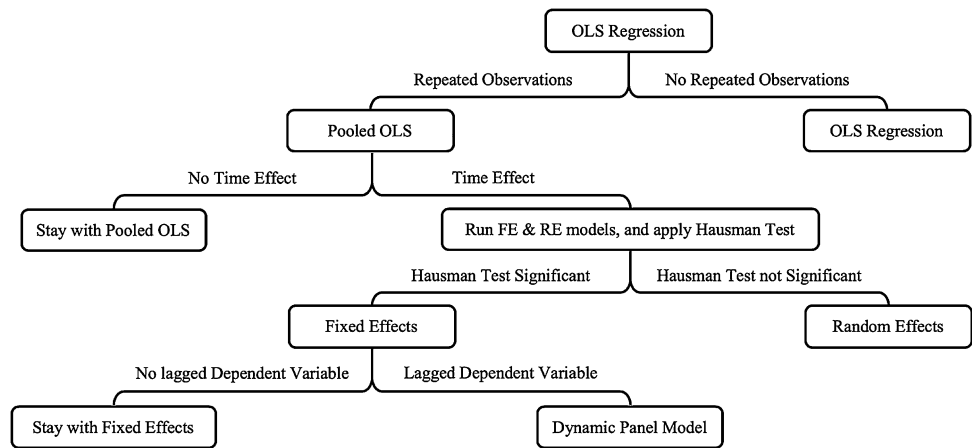
Table 3 List of regression equations

Model	Regression equation
OLS	$y_i = \alpha + x_i' \beta + \varepsilon_i$
Pooled OLS	$y_{it} = \alpha + x_{it}' \beta + \varepsilon_{it}$
Random effects	$y_{it} = \alpha + x_{it}' \beta + u_i + \varepsilon_{it}$
Fixed effects	$y_{it} = \alpha + x_{it}' \beta + f_i + \varepsilon_{it}$
Dynamic models	$y_{it} = \alpha + \delta y_{i,t-1} + x_{it}' \beta + e_{it}$

Appendix 3

See Fig. 4.

Fig. 4 Decision tree



Appendix 4

See Table 4.

Table 4 Review process

Phase	Step	Purpose	Techniques	Results
Scoping study	Define search scope	Specify search terms; journals (NML, NVSQ, and VOLUNTAS); search engine (ProQuest)	Close reading of five textbooks and discussion among authors	22 search terms and 3 leading nonprofit journals
First iteration	Identify potentially pertinent publications	Generate a pool of pertinent publications	Automated search in ProQuest Central	251 publications identified
Second iteration	Identify publications applying linear panel data modeling techniques	Filter the initial pool to identify publications with panel data modeling techniques	Manual coding and selection by the authors	24 publications identified
Third iteration	Identify publications that apply linear panel data modeling but did not get captured in ProQuest Central	Identify articles that may not be captured in ProQuest Central	Add articles that authors happen to come by in the research	3 additional publications identified

Table Adapted from Maier et al. (2016)

Appendix 5

See Table 5.

Table 5 Stata commands

Function	Stata command
OLS	reg DV IVs
Pooled OLS	reg DV IVs, vce(cluster id variable)
Set data as panel data set	xtset id variable time variable
Descriptive statistics of panel data set	xtsum DV IVs
Random effects	xtreg DV IVs, re
Fixed effects	xtreg DV IVs, fe
Hausman test	xtreg DV IVs, re est store RE xtreg DV IVs, fe est store FE hausman FE RE, sigmamore
Arellano–Bond estimator	xtabond DV IVs, lags (2) maxldep (3)

References

- Allen, M. P. (1997). The problem of multicollinearity. In *Understanding regression analysis*, pp. 176–180. Springer. Doi: https://doi.org/10.1007/978-0-585-25657-3_37
- Amirkhanyan, A. A., Kim, H. J., & Lambright, K. T. (2009). Faith-based assumptions about performance: Does church affiliation matter for service quality and access? *Nonprofit and Voluntary Sector Quarterly*, 38(3), 490–521. <https://doi.org/10.1177/0899764008320031>
- Anderson, T. W., & Hsiao, C. (1981). Estimation of dynamic models with error components. *Journal of the American Statistical Association*, 76(375), 598–606. <https://doi.org/10.2307/2287517>
- Arellano, M. (2003). *Panel data econometrics*. Oxford university press. https://books.google.com/books/about/Panel_Data_Econometrics.html?id=QBWO67VjsC
- Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277–297. <https://doi.org/10.2307/2297968>
- Bae, K. B., & Sohn, H. (2018). Factors contributing to the size of nonprofit sector: Tests of government failure, interdependence, and social capital theory. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 29(3), 470–480. <https://doi.org/10.1007/s11266-017-9888-3>
- Baltagi, B. H. (2008). *Econometric analysis of panel data*. John Wiley & Sons.
- Baltagi, B. H., Bresson, G., & Piroette, A. (2003). Fixed effects, random effects or Hausman–Taylor?: A pretest estimator. *Economics Letters*, 79(3), 361–369. [https://doi.org/10.1016/S0165-1765\(03\)00007-7](https://doi.org/10.1016/S0165-1765(03)00007-7)
- Bartels, B. (2015). Beyond ‘fixed versus random effects’: A framework for improving substantive and statistical analysis of panel, TSCS, and multilevel data. In R. Franzese (Ed.), *Quantitative research in political science*. SAGE Publications Ltd. <https://us.sagepub.com/en-us/nam/quantitative-research-in-political-science/book243299>
- Beckman, C. M., & Burton, M. D. (2008). Founding the future: path dependence in the evolution of top management teams from founding to IPO. *Organization Science*, 19(1), 3–24. <https://doi.org/10.1287/orsc.1070.0311>
- Bell, A., & Jones, K. (2015). Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, 3(1), 133–153.
- Bromley, P., Schofer, E., & Longhofer, W. (2018). Organizing for education: A cross-national, longitudinal study of civil society organizations and education outcomes. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 29(3), 526–540. <https://doi.org/10.1007/s11266-018-9979-9>
- Brooks, A. C. (2000). Public subsidies and charitable giving: Crowding out, crowding in, or both? *Journal of Policy Analysis and Management*, 19(3), 451–464. [https://doi.org/10.1002/1520-6688\(200022\)19:3%3c451::AID-PAM5%3e3.0.CO;2-E](https://doi.org/10.1002/1520-6688(200022)19:3%3c451::AID-PAM5%3e3.0.CO;2-E)
- Bun, M. J., & Sarafidis, V. (2013). Dynamic panel data models. UvA-Econometrics Discussion Paper, 1.
- Calabrese, T. D. (2011). The Accumulation of Nonprofit Profits: A Dynamic Analysis. *Nonprofit and Voluntary Sector Quarterly*, 41(2), 300–324. <https://doi.org/10.1177/0899764011404080>
- Cameron, A. C., & Trivedi, P. K. (2010). *Microeconometrics using stata*, Vol. 2. Stata press College Station.
- Cheng, Y. (2018). Nonprofit spending and government provision of public services: Testing theories of government-nonprofit relationships. *Journal of Public Administration Research and Theory*, 29(2), 238–254. <https://doi.org/10.1093/jopart/muy054>
- Clark, T. S., & Linzer, D. A. (2015). Should I use fixed or random effects? *Political Science Research and Methods*, 3(2), 399–408.
- Cooper, H. M. (1989). *Integrating research: A guide for literature reviews*. Sage Publications.
- Coupet, J. (2018). Exploring the link between government funding and efficiency in nonprofit colleges. *Nonprofit Management and Leadership*, 29(1), 65–81. <https://doi.org/10.1002/nml.21309>
- Coupet, J., & Berrett, J. L. (2019). Toward a valid approach to nonprofit efficiency measurement. *Nonprofit Management and Leadership*, 29(3), 299–320. <https://doi.org/10.1002/nml.21336>
- David, R. J., & Han, S. (2004). A systematic assessment of the empirical support for transaction cost economics. *Strategic Management Journal*, 25(1), 39–58.
- De Los Mozos, I. S. L., Rodríguez Duarte, A., & Rodríguez Ruiz, Ó. (2016). Resource dependence in non-profit organizations: Is it harder to fundraise if you diversify your revenue structure?

- VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 27(6), 2641–2665. <https://doi.org/10.1007/s11266-016-9738-8>
- de Wit, A., & Bekkers, R. (2017). Government support and charitable donations: A meta-analysis of the crowding-out hypothesis. *Journal of Public Administration Research and Theory*, 27(2), 301–319. <https://doi.org/10.1093/jopart/muw044>
- Diebold, J., & Cogburn, J. D. (2018). The determinants and opportunity costs of external management fees for state-administered pension plans. *Public Budgeting and Finance*, 38(4), 3–31. <https://doi.org/10.1111/pbaf.12207>
- Einolf, C. J. (2017). Parents' charitable giving and volunteering: Are they influenced by their children's ages and life transitions? Evidence from a longitudinal study in the United States. *Nonprofit and Voluntary Sector Quarterly*, 47(2), 395–416. <https://doi.org/10.1177/0899764017737870>
- Elhorst, J. P. (2003). Specification and estimation of spatial panel data models. *International Regional Science Review*, 26(3), 244–268. <https://doi.org/10.1177/0160017603253791>
- Faulk, L., & Derrick-Mills, T. (2015). *Shaping the Future of Nonprofit Data*. Shaping the Future of Nonprofit Data. <https://www.american.edu/spa/news/shaping-future-of-nonprofit-data-07062015.cfm>
- Frees, E. W. (2004). *Longitudinal and panel data: Analysis and applications in the social sciences*. Cambridge University Press.
- Galaskiewicz, J., & Bielefeld, W. (1998). *Nonprofit Organizations in an Age of Uncertainty: A Study of Organizational Change*. A. de Gruyter.
- Garson, G. D. (2013). *Longitudinal Analysis*. Statistical Associates Publishers.
- Glanville, J. L., Paxton, P., & Wang, Y. (2015). Social capital and generosity: A multilevel analysis. *Nonprofit and Voluntary Sector Quarterly*, 45(3), 526–547. <https://doi.org/10.1177/0899764015591366>
- Greene, W. H. (2003). *Econometric Analysis* (5th, illustrated ed.). Prentice Hall. <https://books.google.com/books?id=JJKWAQAAMAAJ>
- Grossman, A., & Rangan, V. K. (2001). Managing multisite nonprofits. *Nonprofit Management and Leadership*, 11(3), 321–337. <https://doi.org/10.1002/nml.11306>
- Habitat for Humanity. (2014, November 11). *HUD awards \$4.2 million capacity building grant to Habitat for Humanity*. <https://www.habitat.org/newsroom/11-11-2014-HUD-Capacity-Building>
- Habitat for Humanity. (2020). Our mission, vision, and principles. <https://www.habitat.org/about/mission-and-vision>
- Halaby, C. N. (2004). Panel models in sociological research: Theory into practice. *Annual Review of Sociology*, 30(1), 507–544. <https://doi.org/10.1146/annurev.soc.30.012703.110629>
- Hannan, M. T., & Freeman, J. (1984). Structural inertia and organizational change. *American Sociological Review*, 49(2), 149–164. <https://doi.org/10.2307/2095567>
- Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46(6), 1251–1271. <https://doi.org/10.2307/1913827>
- Heist, H. D., & Vance-McMullen, D. (2019). Understanding donor-advised funds: how grants flow during recessions. *Nonprofit and Voluntary Sector Quarterly*, 48(5), 1066–1093. <https://doi.org/10.1177/0899764019856118>
- Heutel, G. (2014). Crowding out and crowding in of private donations and government grants. *Public Finance Review*, 42(2), 143–175. <https://doi.org/10.1177/1091142112447525>
- Hsiao, C. (2007). Panel data analysis—Advantages and challenges. *TEST*, 16(1), 1–22. <https://doi.org/10.1007/s11749-007-0046-x>
- Hsiao, C. (2014). *Analysis of panel data*. Cambridge University Press.
- Hutcheson, G. D., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. SAGE Publications, Ltd. <http://methods.sagepub.com/book/the-multivariate-social-scientist>
- James, R. N., III. (2018). Cash is not king for fund-raising: Gifts of noncash assets predict current and future contributions growth. *Nonprofit Management and Leadership*, 29(2), 159–179. <https://doi.org/10.1002/nml.21334>
- Jaskyte, K. (2012). Boards of directors and innovation in nonprofit organizations. *Nonprofit Management and Leadership*, 22(4), 439–459. <https://doi.org/10.1002/nml.21039>
- Jensen, U. T., & Vestergaard, C. F. (2016). Public service motivation and public service behaviors: Testing the moderating effect of tenure. *Journal of Public Administration Research and Theory*, 27(1), 52–67. <https://doi.org/10.1093/jopart/muw045>
- Keele, L., & Kelly, N. J. (2006). Dynamic models for dynamic theories: The ins and outs of lagged dependent variables. *Political Analysis*, 14(2), 186–205. <https://doi.org/10.1093/pan/mpj006>
- Kim, M. (2015). Socioeconomic diversity, political engagement, and the density of nonprofit organizations in US counties. *The American Review of Public Administration*, 45(4), 402–416.
- Kim, Y. H., & Kim, S. E. (2018). What accounts for the variations in nonprofit growth? a cross-national panel study. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 29(3), 481–495. <https://doi.org/10.1007/s11266-016-9752-x>
- Light, R. J., & Pillemer, D. B. (1984). *Summing Up: The Science of Reviewing Research*. Harvard University Press. <https://books.google.com/books?id=5Y14EDR7UicC>
- Maddala, G. S., & Wu, S. (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics*, 61(S1), 631–652. <https://doi.org/10.1111/1468-0084.0610s1631>
- Maier, F., Meyer, M., & Steinbereithner, M. (2016). Nonprofit organizations becoming business-like: A systematic review. *Nonprofit and Voluntary Sector Quarterly*, 45(1), 64–86.
- Matsunaga, Y., Yamauchi, N., & Okuyama, N. (2010). What determines the size of the nonprofit sector?: a cross-country analysis of the government failure theory. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 21(2), 180–201. <https://doi.org/10.1007/s11266-010-9125-9>
- Mayer, W. J., Wang, H., Egginton, J. F., & Flint, H. S. (2014). The impact of revenue diversification on expected revenue and volatility for nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly*, 43(2), 374–392.
- Mosley, J. E., & Galaskiewicz, J. (2015). The relationship between philanthropic foundation funding and state-level policy in the era of welfare reform. *Nonprofit and Voluntary Sector Quarterly*, 44(6), 1225–1254. <https://doi.org/10.1177/0899764014558932>
- Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, pp. 1417–1426.
- Nikolova, M. (2015). Government funding of private voluntary organizations: Is there a crowding-out effect? *Nonprofit and Voluntary Sector Quarterly*, 44(3), 487–509. <https://doi.org/10.1177/0899764013520572>
- Paarlberg, L. E., & Hwang, H. (2017). The heterogeneity of competitive forces: the impact of competition for resources on united way fundraising. *Nonprofit and Voluntary Sector Quarterly*, 46(5), 897–921. <https://doi.org/10.1177/0899764017713874>
- Pandey, S. K., & Johnson, J. M. (2019). Nonprofit management, public administration, and public policy: Separate, subset, or intersectional domains of inquiry? *Public Performance and Management Review*, 42(1), 1–10. <https://doi.org/10.1080/15309576.2018.1557382>
- Pedrini, M., & Ferri, L. M. (2016). Doing well by returning to the origin. Mission drift, outreach and financial performance of

- microfinance institutions. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 27(6), 2576–2594. <https://doi.org/10.1007/s11266-016-9707-2>.
- Plümper, T., Troeger, V. E., & Manow, P. (2005). Panel data analysis in comparative politics: Linking method to theory. *European Journal of Political Research*, 44(2), 327–354. <https://doi.org/10.1111/j.1475-6765.2005.00230.x>
- Ranucci, R., & Lee, H. (2019). Donor influence on long-term innovation within nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly*, 48(5), 1045–1065. <https://doi.org/10.1177/0899764019843346>
- Rao, C. R., & Toutenburg, H. (1995). *Linear Models: Least Squares and Alternatives* (1st ed.). Springer. <https://doi.org/10.1007/978-1-4899-0024-1>
- Rosenbaum, P. R. (2005). Heterogeneity and Causality. *The American Statistician*, 59(2), 147–152. <https://doi.org/10.1198/000313005X42831>
- Rotolo, T., Wilson, J., & Dietz, N. (2014). Volunteering in the United States in the aftermath of the foreclosure crisis. *Nonprofit and Voluntary Sector Quarterly*, 44(5), 924–944. <https://doi.org/10.1177/0899764014546669>
- Shearer, R., & Clark, T. (2016). Are we drawing the correct conclusions? Regression analysis in the nonprofit literature. *Canadian Journal of Nonprofit and Social Economy Research*, 7(1). <https://doi.org/10.22230/cjnser.2016v7n1a213>
- Sydow, J., Schreyögg, G., & Koch, J. (2009). Organizational path dependence: Opening the black box. *The Academy of Management Review*, 34(4), 689–709.
- Szper, R. (2013). Playing to the test: Organizational responses to third party ratings. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 24(4), 935–952. <https://doi.org/10.1007/s11266-012-9290-0>
- Tsay, W.-J., & Chung, C.-F. (2000). The spurious regression of fractionally integrated processes. *Journal of Econometrics*, 96(1), 155–182. [https://doi.org/10.1016/S0304-4076\(99\)00056-1](https://doi.org/10.1016/S0304-4076(99)00056-1)
- Wawro, G. (2002). Estimating dynamic panel data models in political science. *Political Analysis*, 10(1), 25–48. <https://doi.org/10.1093/pan/10.1.25>
- Weimar, D., Wicker, P., & Prinz, J. (2015). Membership in nonprofit sport clubs: A dynamic panel analysis of external organizational factors. *Nonprofit and Voluntary Sector Quarterly*, 44(3), 417–436.
- Wicker, P., & Frick, B. (2016). Recruitment and retention of referees in nonprofit sport organizations: the trickle-down effect of role models. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 27(3), 1304–1322. <https://doi.org/10.1007/s11266-016-9705-4>.
- Winters, M. S., & Rundlett, A. (2015). The challenges of untangling the relationship between participation and happiness. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 26(1), 5–23. <https://doi.org/10.1007/s11266-014-9500-z>
- Witt, A., Kurths, J., & Pikovsky, A. (1998). Testing stationarity in time series. *Physical Review E*, 58(2), 1800–1810. <https://doi.org/10.1103/PhysRevE.58.1800>
- Wooldridge, J. M. (2005). Fixed-effects and related estimators for correlated random-coefficient and treatment-effect panel data models. *Review of Economics and Statistics*, 87(2), 385–390.
- Wooldridge, J.M. (2013). *Introductory Econometrics: A Modern Approach* (5th ed.). Cengage Learning. <https://books.google.com/books?id=C0KHwUKxys0C>
- Woronkiewicz, J., & Nicholson-Crotty, J. (2017). The effects of capital campaigns on other nonprofits' fundraising. *Nonprofit Management and Leadership*, 27(3), 371–387. <https://doi.org/10.1002/nml.21255>
- Yi, H., & Chen, W. (2019). Portable innovation, policy wormholes, and innovation diffusion. *Public Administration Review*, 79(5), 737–748. <https://doi.org/10.1111/puar.13090>
- Yu, Z. (2016). The effects of resources, political opportunities and organisational ecology on the growth trajectories of AIDS NGOs in China. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 27(5), 2252–2273. <https://doi.org/10.1007/s11266-016-9686-3>
- Zhu, L. (2013). Panel data analysis in public administration: Substantive and statistical considerations. *Journal of Public Administration Research and Theory*, 23(2), 395–428. <https://doi.org/10.1093/jopart/mus064>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.