

A Methodological Review of Studies on Effects of Financial Aid on College Student Success

Jin Chen, Indiana University

Desiree Zerquera, Indiana University

Abstract

The nationally growing concerns on college student success have encouraged scholarly investigation in the effectiveness of financial aid policies that aim to narrow the achievement gaps between social groups. Studies on effects of financial aid though recognize the role of financial aid in increasing college access, choice and subsequent persistence, they disagree to a great extent on effects of specific types of financial aid (e.g. loan) when utilizing different analytical methods and/or dissimilar data sets. With no careful scrutiny on the soundness of the research design when adopting policy recommendations, the initiatives on closing the achievement gap would likely to be jeopardized or result in vain. In this paper, we first reviewed methodological issues critical to financial aid studies on college student success, including measures of financial aid, nature of outcome variables, longitudinal process and contexts of student success, differential aid effects across subgroups and the omitted variable bias (as well as self-selection bias). Both advantages and disadvantages adopted by researchers to account for these methodological challenges were discussed. We then proposed the limitations incurred by various data sources, and issues related to data availability, quality and reliability. Finally, directions for future research were suggested.

Introduction

The United States has made remarkable achievement in expanding college access for an increasingly large student population in recent decades. The national participation rate, one of the highest in the world (Tinto, 2005), reached 62% in 2006 (NCHEMS, 2009), and the undergraduate enrollment in U.S. colleges and universities increased by 32% between 1998 and 2008 (NCES, 2010). The access gap between income groups has been narrowed as the college enrollment of economically disadvantaged students has risen constantly (Tinto, 2005).

However, large disparities remain in patterns of attendance and success in college across income groups ((Tinto, 2005; Engle and Tinto, 2008). For low income families, how and where they attend higher education institutions are very much restricted by their financial constraints (Tinto, 2005). Economic stratification in participation in terms of institutional selectivity and enrollment intensity has been widely documented. Particularly, compared to their high-income counterparts, low-income students are less likely to enroll in four-year sector, or elite institutions, or as full-time (Cabrera, Burkurn, and La Nasa, 2005; Carnevale and Rose, 2003; Bowen, Kurzweil, and Tobin, 2005; NCES, 1999). More importantly, gaps in postsecondary educational attainment for historically underrepresented groups, such as the low-income, remain (Bedlla, 2010; Engle and Tinto, 2008). Specifically, as low-income youth disproportionately enroll in two-year colleges, they are less likely to achieve a bachelor's degree within six years (NCES, 2003-151, Table 2.1 A). Student socioeconomic background matters even after sector, school selectivity, as well as academic preparation (e.g. test scores) are controlled for (Tinto, 2005).

In addition to the economic stratification in patterns of college attendance, the relatively high price-sensitivity of low-income students also to some extent explains the disparity in academic achievement (Price, 2004; Tinto, 2004). That being said, it is clear and imperative that

we provide academically qualified and economically challenged students with the financial means that promote their college attendance and educational attainment (Tinto, 2005). For decades, a broad range of federal and state efforts have been made to encourage low-income students' participation and continuation in higher education, including the provision of various types of financial aid, the "most popular and least threatening" (p.38) fiscal mechanism that influences student success (Richardson Jr, and Ken, 2002). However, the escalating college cost along with the dramatic shifts from grant aid to loans, and from need-based aid to merit-based scholarships since the early 1990s has superseded gaps in college affordability and postsecondary educational attainment between income groups (Chen, 2008).

It is to the interest of both policy makers and educational researchers to understand the mechanism as well as the effectiveness of financial aid policies that target students in need. The bulk of financial aid studies indicate that financial aid in general is likely to increase college access, choice and subsequent persistence (Cabrera et al., 1993; Nora, 1990; St. John, Cabrera, Nora, and Asker, 2000; Hossler, et al., 2008); however, the adoption of different methodologies often times leads to scholarly disagreement on the significance, magnitude and/or direction of effects of certain types of financial aid. With no careful scrutiny on the soundness of research design when accepting policy recommendations, the initiative on closing the achievement gap is likely to be jeopardized. The purpose of this paper is to examine methodological challenges faced by and research strategies used in financial aid studies by reviewing the work done in this policy arena.

Methodological Challenges and Research Strategies

A methodology is a systematic way of solving the research problem (Kumar, 2005), and may also be defined as a generic framework established by the academia for acquiring new

knowledge via collecting and evaluating the existent knowledge (Sekaran, 2003). A methodology is of the unique importance to a research, since it identifies tools, strategies, process measurement and evaluative criteria for specified research aims (Sekaran, 2003). Therefore, a sound and appropriate methodology is critical to the success of a research.

For financial aid studies that aim to explore the mechanism and to evaluate the effectiveness of the policy, multidisciplinary perspectives and methodological preferences are well presented. Given that strengths and limitations exist in any methodology adopted, an investigation into methodological challenges faced by and research strategies employed in financial aid studies is likely to provide some insight into the development as well as the future direction of this research area. Specifically, this section discusses measures of financial aid, target population and sample, soundness and appropriateness of data, and techniques of analysis using quantitative deductive approach.

Measures of Financial Aid

Measures of financial aid (or more accurately, the financial aid policy) vary with specific research questions. The decision on how to quantify the financial aid policy not only has significant implications for policy analysis but also shapes the overall design of the research. A number of studies take into account the total amount of aid students receive each year (Dynarski, 2003), neglecting different effects associated with different types of aid. Another form of aggregation is the utilization of financial aid status, i.e. whether received financial aid or not, (St. John, et al., 2005), unduly assuming the homogeneity of financial aid. As Chen (2008) points out, “researchers usually use an aggregated variable of financial aid, without account for differences by subtypes” (217). This, to a greater extent, clouds the effects of different types and amounts of financial aid.

Considering the different policy initiatives, scholars differentiate need-based aid from non-need-based (or merit-based) aid in their analyses. For instance, Stater (2009) define need-based aid as the sum of all need-based grants and loans¹, and merit-based aid as the sum of state and institutional non-need-based scholarship when examining their effects on college GPA. Other works in line with Stater's include Heller (1999), Somers (1995), Herzog (2005) and Farrell (2007). The separation of these two major policy initiatives, though demonstrating improvement, still shows insufficient consideration of subtypes of aid. Take need-based financial aid as an example, grants, loan and work-study have been found to influence students college decisions via different mechanisms, the statement of which is supported by findings from a wide array of studies (e.g. Astin, 1975; Astin and Cross, 1979; St. John and Starkey, 1995; St. John, Kirshstein, and Noell, 1991; Voorhees, 1985; DesJardins, et al, 2002; St. John, 1991; Singell, 2002).

Among studies that examine different subtypes of financial aid, specifically those that focus on loans, have reported mixed findings (e.g. St. John, Kirshstein, and Noell, 1991; Voorhees, 1985; DesJardins, et al, 2002; Astin, 1975; Carroll, 1987; Peng & Fetters, 1978) and therefore warrant additional examination. Research indicates that the failure to distinguish between loan types, such as subsidized loans vs. unsubsidized-loans, is likely to contribute to misunderstandings of loan effects (Singell, 2002; Chen, 2008). For example, need-based loans such as the Perkins loans and Stanford subsidized loans, are likely to positively relate to students' persistence; while non-need-based (or unsubsidized) loans such as the Stanford Unsubsidized loans, are found to be trivial in predicting students' retention (Singell, 2002).

¹Stater's (2009) measure of need-based aid includes federal, state and institutional need-based grants; Federal Perkins Loans; and Federal Stafford Loans.

In addition, the fact that many students receive more than one type of financial aid intrigues a few scholars to make comparisons between effects of a single form of aid and that of aid packages (Astin, 1975; St. John, 1989; Murdock, 1990; Hu and St. John, 2001). While applauded for the initiatives on examining individual or combinations of aid type(s), these attempts failed to consider amounts of each type awarded to students (Heller, 1999).

Although differentiation of financial aid (policies) seems to acquire superiority over aggregated measures, the specific measure used in practice is largely determined by the research questions and the availability of data.

Population and Sample

The target population refers to the persons or group(s) “whose behavior and well-being are affected by (a) public policy” (Schneider and Ingram, 1993, p.334). To achieve different goals, higher education financial aid policies target different groups. For example, the Federal Pell grant targets low-income undergraduates and certain post-baccalaureate students in (U.S. Department of Education, 2010) ; the Indiana Twenty First Century Scholars program ensures college affordability for college enrollees from low and moderate-income Indiana families (State Student Assistance Commission of Indiana, n.d); the Georgia's Helping Outstanding Pupils Educationally program, a nationally recognized state funded merit-based financial aid policy, aims to reward degree seekers with satisfactory academic records (Georgia Student Finance Commission, 2011); and institutional financial aid are used strategically to recruit students for the purposes of maintaining educational quality, expanding applicant pool or maximizing institutional prestige (McPherson and Schapiro, 1991). Ideally, a study on the entire target population would provide most information on the policy effects, however, it is often times

impractical (e.g. given the dynamics of the population) or prohibitively expensive to conduct a census inquiry (Adèr, Mellenbergh, and Hand, 2008).

Instead, researchers select subsets of the target group to gain information about the whole population (Webster, 1985). The advantages of sampling include economy, timeliness, (wide) scope and accuracy of data (Cochran, 1977). Although sampling inevitably brings in errors, a representative sample would provide valid inferences about the entire group(s) of interest (Cameron, Gardner, Doherr and Wagner, 2008). Ideally, a simple random sample (SRS) of sufficient size is representative of the target population and is the most favorable to statistical inferences. However, in many cases, a list of members of the target population from which we can randomly select is not available; even if the randomness is met, a sufficient number of sampling units with certain characteristics are also essential for meaningful statistical inferences (Thomas and Heck, 2001). For most financial aid studies and higher education research in general, more complex sampling frames are used, among which stratified sampling and cluster sampling are usually employed (Thomas and Heck, 2001; e.g. NCES, 1995, 1996).

Albeit this paper does not discuss sampling frames and data collection, it is important to be aware of and control for complex sampling structures (e.g. intra-class correlation) and biases introduced during the process (e.g. over- or under-representation) in analyses. Common practices to account for complex sampling frames and representativeness biases include applying weights to observations that are over- or under-represented and statistical methods such as hierarchical modeling techniques to account for multistage clustering (Thomas and Heck, 2001; Hox, 1998).

Data Source and Quality

Given the fact that most studies in this field adopt quantitative and deductive approaches, in this section I will discuss the limitations incurred by secondary data available to researchers in this field, as well as issues related to data comprehensiveness, quality, and reliability. Given that the literature reviewed for this paper is exclusively dependent on secondary data, this section only discusses relevant issues in this regard.

In general, data for studies in this field come from various sources, which are housed by different educational entities at different levels, namely, national, state and institutional. Hossler and colleagues (2008) conducted a comprehensive review on the strengths and limitations of these three levels of datasets (see table 1). To select datasets from these three levels, the key tradeoff is between richness of the data and the generalizability of the study. Simply put, national level datasets are more generalizable yet less comprehensive.

Table 1 Data Strengths and Limitations by Source

Source	Exemplars	Strengths	Limitations
National	NPSAS, BPS, HSB, NLSY	Robust set of student background variables Standard definitions of aid Longitudinal	Insufficient samples size for assessing state aid Impossible to assess institutional aid programs Lack college experience measures
State	Indiana , Georgia	Appropriate for examining state aid programs Track in-state transfers over time	Lack institutional aid elements Lack college experience measures Can't track out-of-state transfers
Institutional		Academic and social integration measures Merit- and need-based aid	Can't track enrollment patterns beyond the institution Can't examine aid effects beyond the institution

Source: Table adapted from Hossler, et al. (2008), p. 395-397.

Additional challenges include difficulties as well as risks associated with integrating datasets from different sources/levels, unavailability of certain information, and reliability and

validity of survey data. A comprehensive design of study usually incorporates different dimensions of information, some of which require multifaceted or multilevel data, e.g. student level data, institutional level data and state level data. Integrating data from different sources is challenging, because it requires understanding of different definitions of the same constructs and craft in dealing with complex survey designs.

Furthermore, the availability and measurability of certain factors often propose challenges to financial aid studies. Information on family income for financial aid non-applicants is usually unavailable; so are measures on student high school performance (St. John, 2004). Social factors such as emotional health and peer support (Pritchard and Wilson, 2003; McGrath and Braunstein, 1997), which play an unelectable role in student success, are neither readily available nor measurable. Limited data collections/observations on the unit of analysis put extra threat to longitudinal analyses (Chen, 2008).

Finally, the reliability and the validity of a particular survey constrain the use of data in exploring effects of aid. Given that the self-reported nature of most surveys, information such as aid type and amount offered may neither be correctly recalled nor recorded (St. John, 1990). The nonrandom sample attrition in surveys such as NPSAS-87 leads to non-representative sample of college students (Boatman and Long, 2009; St. John, et al., 2005), therefore the generalizability of the study is restricted. And the extent to which data are missing put additional threat to the study.

Techniques of Analysis

Once research questions, target population, sample and data source are determined, appropriate techniques of analysis are to be applied to either test theoretical hypotheses or provide new understanding of a policy. Selection of analytical techniques ought to take into

account the nature of outcome variables, the temporal dimension of financial aid and student success, the policy context, and the subgroup differences. Particularly, techniques to correct self-selection or omitted variable bias will be discussed.

Natures of Outcome Variables

As stated in previous sections, quantifiable indicators of student success include academic performance (or college GPA), persistence to graduation and degree attainment or completion. Natures of these outcome variables vary. Specifically, GPA, usually measured on a four point scale, is a continuous variable; persistence, either year-to-year or within-year, is a dichotomous measure of continuous enrollment; degree attainment or completion, is either dichotomous (completed a degree or not) or multinomial (what types of degree) or continuous (time-to-degree) based on specific research questions being asked. When different outcome variables are assessed, corresponding method ought to be used. For example, Stater's (2009) exploration in the relationship between financial aid and student academic performance represents the application of linear regression models in this line of studies. When persistence indicator is the outcome variable of interest, unlike their predecessors, such as Pascarella & Terenzini (1980), who employed linear models, scholars (e.g. Cabrera et al., 1990; St. John et al, 2000) started to make use of logistic regression analysis which "captures the probabilistic distribution embedded in dichotomized distributions" and "avoids violating the assumptions of homoscedasticity and functional specification" (Chen, 2008, p.220). More recently, some studies (St. John and Chung, 2006; Yi, 2008) used multinomial logistic regression to predict probabilities of receiving different types of degree. Although linear regression models, binary logit models and multinomial logit models take good consideration of the nature of outcome

variables, these models fail to address the dynamic aspect of student success and the changing nature of financial aid over time.

Temporal Dimension

“Student success is a longitudinal process” (Perna, and Thomas, 2006, p8.). As Hossler and colleagues (2008) conclude in their review of persistence studies, “the temporal nature of persistence is implicitly recognized in the extant literature on educational attainment”; yet, except for a few studies (St. John, 1991; Chen and Desjardins, 2008; DesJardins et al., 1999, 2002; Ishitani and Desjardins, 2003) that addressed the time-varying characteristics of both financial aid and student behavior, most researchers approached this analysis either with cross-sectional perspectives or by incorporating only two points in time (e.g. Tinto, 1982), ignoring the fact that “changes over time in financial aid packages can influence students’ academic and social integration processes, as well as their subsequent persistence decisions” (St. John et al., 2000, p.41; as cited in Hossler, et al., 2008).

Since the early 1990s, scholars in this field have started to address this time-dependent nature of student success via different analytical techniques. St. John and colleagues (1991) were among the first who applied sequential regression analyses to student persistence/departure study. By running logistic regressions on samples from each time period, the sequential analysis recognized the longitudinal aspect of student departure/persistence, yet its limitations, like what Chen (2008) points out, “lies in the fact that the impact of time on the student outcome was not fully explored and the effects of factors in previous time periods could not be controlled for in the estimation of subsequent outcomes” (220).

Only recently, have persistence/departure studies introduced more advanced techniques developed in other fields, such as economics and sociology, to address the need for controlling

time or demonstrating the temporal dynamics of financial aid and student success. Among the most commonly used is the event history analysis or survival analysis (Chen, 2008; Desjardins et al., 1994, 2002, 2003; Doyle, 2006; Gross and Torres, 2010), which is used for predicting occurrence and timing of events with a set of covariates. Survival models surpass other regression models in two aspects: a) they are capable of dealing with censored observations, for which only partial information on timing of the event is available; and b) the functional forms take into account perceived values of both time-varying and time-invariant covariates (Allison, 1984; Yamaguchi, K. 1991). The latter feature makes it possible to conduct analysis in a dynamic manner. Given the discrete time points of observation (by year or by semester), discrete-time models as an approximation for continuous-time models are often used in most educational researches (e.g. Chen and DesJardins, 2008; Gross and Torres, 2010). In spite of the widely acknowledged advantages of event history analysis, one weakness of this approach as I can see, is the insufficient consideration of influences from contexts or environments. In fact, individuals from the same institution, classroom, field of study, etc., tend to be more homogenous, i.e. sharing certain characteristics, which violate the assumptions of independence (among observations) for most regression analyses. As a result, Ordinary Least Squares (OLS) tends to bias the standard errors downward and the null hypothesis (i.e. the effect is not significant) is more likely to be rejected (Osborne, 2008). Although educational scholars are aware and capable of controlling for nested effects by adding additional levels of analysis to their empirical models (e.g. Titus, 2004; Titus, 2006), these attempts are limited to linear regression, binary logit models, and multinomial logistic regression.

Policy Context

“Student success is shaped by multiple levels of context” (NPEC, 2006, P.9). To address contextual influences, scholars incorporate institutional characteristics or policy environment in their theoretical and analytical framework (e.g. Bergen-Milem, 2000; Bean, 1990). In practices, environmental variables are usually directly incorporated into statistical models without being adjusted for the nested effects. Like what is stated in the previous paragraph, failure to consider nested effects would result in biased estimates. Only a few studies of student success (e.g. Kim et al., 2003; Titus, 2004) involve both student and institution as units of analysis, and even fewer studies (e.g. Titus, 2006) add additional level (e.g. state) to analyses when public policy context is framed into research questions.

Multilevel modeling techniques in this regard are appropriate for considering different layers of contexts. For example, Titus (2004) employs hierarchical generalized linear modeling (HGLM) to explore effects of institutional and individual characteristics on student persistence. As Titus (2004; 2006) suggested, HGLM outperforms other methods in three aspects: a) It allows for comprehensive analysis on influences from higher level factors after taking into account lower level variables; b) it takes into account the hierarchical/clustered nature of data, for which single-level technique leads to underestimated standard errors; and c) the use of maximum likelihood estimation as computing algorithm usually results in “robust, asymptotically efficient, and consistent parameter estimates when used with large samples with unequal group sizes” (Titus, 2004, p. 684). Additional to its complexity, this approach is also challenged by what is common in most social research--self-selection. Failure to adjust for probability of self-selection will lead to biased estimates (DesJardins, et al., 2002).

Differential Aid Effects on Subgroups

Effects of financial aid vary across socioeconomic groups (Chen, 2008; Hossler, et al, 2008), due to different price- and aid-sensitivities (Paulsen and St. John, 2002). The fact that students with lower SES and therefore more price-sensitive are more likely to enroll in community colleges (Dynarski, 2000) raises serious concerns for studies which fail to disaggregate students from different types of institutions (Hossler, et al, 2008).

Recognizing the differential effects of aid associated with student socioeconomic status, scholars (Paulsen and St. John, 2002; Walpole, 2003) began to compare aid responsiveness by running separate regression models on income groups. Conclusions with regard to effects of a particular type of aid across income groups are made, however, the significance of difference in aid effects could not be inferred (Chen, 2008). One way of addressing this issue is to include interaction terms of aid and income groups and estimate the model on the full sample (e.g. Dowd 2004). This approach not only enables researchers to test whether effects of aid is significantly different for different income groups, but also improves the model specification, because exclusion of significant interaction terms would result in biased estimates (Singer and Willett, 2003). Additionally, interaction terms between aid type and ethnicity, between aid type and time should be included and tested in recognizing the different aid responsiveness across ethnic groups as well as the time-varying nature of financial aid effects (Chen, 2008). In spite of all tempting benefits of including interaction terms, it needs to be cautioned that a) too many interaction terms will result in loss of statistical power; and b) interpretations of interactions between continuous variables might be challenging.

Omitted Variable Bias and Self-selection

Studies on policy effects are often criticized for lack of rigor in determining causality. The key challenges faced by financial aid researchers include ways to control for omitted variable bias and the related issue of self-selection. Omitted variable bias appears when the model specification is poor due to the left-out of important independent variables (Greene, 1993). Self-selection occurs when individuals or other entities choose whether to adopt a policy or participate in a program, etc, based on different characteristics, observable or not (Cellini, 2008).

In most cases, studies on policy effects involve comparing program participants to nonparticipants by controlling personal characteristics (Bailey, 2006). As long as students enroll in programs voluntarily, there remains a strong possibility that the two groups of students differ with respect to characteristics that might influence the outcomes of the policy (Bailey, 2006). The fact that students are nonrandomly assigned to or enrolled in particular programs suggests that certain individual characteristics, observed or unobserved, measured or unmeasured, are likely to affect the observed relationship between financial aid policy and students behavior (Chen, 2008). Specifically, financial aid eligibility or receipt is influenced or sometimes determined by factors such as students' race/ethnicity, family SES, cultural values, aspiration, and motivation, which also affect student's academic performance, persistence and degree attainment (Hossler, et al., 2008); therefore aid recipients may differ significantly in such aspects from non-recipients (Boatman & Long, 2009). In this regard, studies assuming the exogeneity of financial aid eligibility or receipt by using single-stage (logit) models or a dummy variable indicating aid status are severely biased (Hossler, et al., 2008). Realizing the impact of self-selection bias on policy analysis (Alon, 2005; DesJardins, 2005; Boatman and Long, 2009), scholars start to apply

methodological fixes for this problem, including experimental design, regression discontinuity, instrumental variable techniques, propensity score matching and panel data techniques.

Random assignment or controlled experiment is perhaps the ideal way of determining causality (Shadish, Cook, and Campbell, 2002; Weiss, 1998). However, random assignment in social context is extremely challenging. Cellini (2008) discusses four problems associated with random assignment experiments for financial aid policies, including a) high cost of time and money, b) low practicality of implementation and outcome measurement, c) jeopardized social equity, d) and the lack of external validity.

Financial aid policy is in fact never distributed in a random manner². Randomization is usually “politically unfeasible and morally unjustifiable, because some of those who are most in need (or most capable) will be eliminated through random selection” (Dunn, 2008, p. 322). Approximating controlled experiments, financial aid policy researchers innovate quasi-experiment designs to examine the treatment effects, in other words, to identify the causal relationships between a particular aid policy/program and the target group behavior. Regression discontinuity (RD) was designed for particular social experiments where scarce resources are provided for only proportion of needy participants (Dunn, 2008). RD techniques identify “treatment” and “control” groups based on an exogenously determined cutoff, assuming that students/participants’ just below and beyond the cutoff do not differ systematically in observable and unobservable characteristics. The only mean difference would be the “treatment”, financial aid receipt for example. Some higher education researchers (Kane, 2003; Van der Klaauw, 2002; Bettinger, 2004) draw on this approach to assess effects of financial aid policies. The credibility of RD lies primarily in its less biased estimates via eliminating selection based on participants’

² Except for purposeful experiments, such as the “Opening Doors Community College Program” and several other programs in K-12 arena.

own willingness and its simple implementation. Nevertheless, the challenges often lie in its requirement for relatively large sample size (Dunn, 2008) and detailed individual level information from which the exogenous discontinuity could be identified.

Besides, instrumental variable techniques are an econometric approach to control for endogeneity. The IV approach requires finding a (set) of instrumental variable(s) that are uncorrelated with the error term but highly correlated to the endogenous variable (Woodridge, 2002). Alon (2005) does an exceptional work that not only creates a conceptual framework for remedying endogeneity of financial aid, but also provides a concrete example that uses instrumental variable probit. Stater (2009) sets another example which uses census variables³ that describe the student's home zip code as instrumental variables for endogenous need-based financial aid and merit-based financial aid, because these IVs are outside of students' control and unrelated to unobserved factors impacting students' GPA. However, challenges of using IV also exist (Baum, 2009), including the difficulty of finding valid instruments, poor performance of IV in small samples, and the lower precision of IV estimates with weak instruments⁴.

Although it was first proposed by Rosenbaum and Rubin in 1983, only until recently did higher educational researchers (Reynolds and DesJardins, 2009; Herzog, 2007) employ the propensity score matching approach to make more rigorous inferential statements. The propensity score matching is an approach to match participants and non-participants based on the probability of treatment rather than on individual characteristics themselves. Either logit or probit regression models are used to estimate the propensity score for each observed entity, based on their observable characteristics (Reynolds and DesJardins, 2009). Take Herzog's (2007) study

³ These variables include median household income in zip code, unemployment rate in zip code, percent home zip code urbanized, percent home zip with bachelor's, percent home zip foreign born, percent home zip White or Asian, percent housing owner occupied, Percent home zip high school age, and distance from home state.

⁴ Weak instruments are excluded instruments that are only weakly correlated with included endogenous regressors.

of aid effects on freshmen retention as an example, the propensity score was regressed on demographic, pre-college, and first-year university experience variables, and three groups—grant recipients, loan recipients, and non-aid recipients were matched on the nearest propensity score. In short, this approach allows researchers to adjust for selection bias, and is more advantageous than conventional matching techniques by avoiding matching on many observed variables—the “curse of dimensionality” (Reynolds and DesJardins, 2009; Herzog, 2007). Nonetheless, a) it may not fully control for endogeneity of ineligibility (Rosenbaum and Rubin, 1984; Titus, 2006; Rubin, 2004), due to unobserved heterogeneity⁵; and b) it requires large samples in which group overlap must be substantial (Shadish, Cook, and Campbell, 2002).

The last approach to be discussed here is the panel data techniques. Essentially this approach deals with and takes advantage of longitudinal data, for which the same units are observed for multiple times. Studies (Ehrenberg, Zhang, and Levin, 2006; Heller, 1999; Kane, 2004) often resort to fixed-effects models which treat time-invariant unobserved heterogeneity as unit-specific intercept, which captures individual characteristics that are not included in the model. As Zhang (2010) suggests, two-way fixed-effects models by adding period-specific error terms could control additional heterogeneity that is period-specific (i.e. affecting all units during the same time period) and unmeasured. Although “individual fixed-effects would theoretically provide the best control for omitted variable bias” (Cellini, 2008, p.341), this approach still suffers from three aspects: a) it does not control for time-varying heterogeneity, such as social attitude and occupational aspiration; b) it does not make inferences beyond the sample⁶; and c) it requires counterfactual information to make better causality inference (Cellini, 2008).

⁵ Unobserved heterogeneity here refers to unobserved factors that are correlated with outcome variables yet not correlated with the propensity score-matching regressors.

⁶ In this sense, random-effects model which makes distributional assumption of the time-invariant unit-specific error does make inference for the population rather than the sample of analysis.

Conclusions

A sound methodology is vital for conducting policy analyses and making policy recommendations. The adoption of a particular methodology should always be informed by and based on theories that explain various phenomena of student success in the higher education arena. Equally importantly, the comprehensiveness of a methodology is contributed by but not limited to valid measures of financial aid policies, precise identification of target populations, understanding of complex sampling frames (if using secondary data), awareness of advantages and disadvantages of data sources, and appropriate techniques of analysis that take into account natures of outcome variables, the temporal dimension of student success, policy context, the differential effect of financial aid on subgroups, and the endogeneity of financial aid status. Although this literature review is by no means thorough or conclusive, it does bring up the significance of sound methodologies in financial aid policy analyses and perhaps recommend a research agenda that focuses on improving various aspects of methodologies applied to this line of studies.

Directions for Future Research

The imperative for narrowing educational attainment gaps between income groups calls for policy initiatives that well address the needs of those economically disadvantaged. To assess the effectiveness of the current financial aid policies while recommending courses of action, researchers in this field are responsible for making claims based on theoretically legitimate, methodologically sound, and practically feasible studies and making explicit the limitations of each study. The methodological issues, such as measures of financial aid, longitudinal process of

student success, influence from the context, and inference about the causality, all present both challenges and opportunities for scholarly investigations in this field.

Therefore, future research on how financial aids affect college student success is likely to benefit from integrating the following methodological perspectives:

- Specify as clearly as possible the financial aid policy of interest in research questions, which provide basic rationale for determining measures of financial aid;
- Use statistical controls to account for complex sampling frames and non-representativeness of large-scale secondary data;
- When necessary, integrate multiple data sources to counterbalance limitations of individual data sets, and more importantly to allow a fuller range of view;
- When data permit, incorporate temporal dimension and policy context into analysis to account for confounding factors other than policy itself.
- Differentiate policy effects by adding interaction terms between policy and group indicator(s);
- Achieve more robust estimates of policy effects by applying advanced statistical methods such as regression discontinuity, instrumental variable techniques, propensity score matching, and panel data analysis, to control for omitted variable bias or self-selection bias.

References

- Adèr, H. J., Mellenbergh, G. J., & Hand, D. J. (2008). *Advising on research methods: A consultant's companion*. Huizen, The Netherlands: Johannes van Kessel Publishing.
- Allison, P. (1984), *Event History Analysis-Regression for Longitudinal Event Data*, Beverly Hills, CA: Sage Publications.
- Alon, S. (2005). Model mis-specification in assessing the impact of financial aid on academic outcomes. *Research in Higher Education*, 46(1), 109–125.
- Astin, A. W. (1975). *Preventing students from dropping out*. San Francisco: Jossey-Bass
- Astin, H. S., & Cross, P. H. (1979). *Student financial aid and persistence in college*. Los Angeles: Higher Education Research Institute.
- Bailey, T. R. (2006). Research on institution level practice for postsecondary student success. In *National Symposium on Postsecondary Student Success*, Washington, DC.
- Baum, C.F.(2009). Instrumental variables and panel data methods in economics and finance. Retrieved on Feb. 15, 2011, from <http://fmwww.bc.edu/GStat/docs/StataIV.pdf>
- Bean, J. P. (1990). Why students leave: Insights from research. In: D. Hossler, and J. P. Bean (eds.), *The Strategic Management of College Enrollments*. Jossey-Bass: San Francisco
- Bedolla, L. G. (2010). Good Ideas Are Not Enough: Considering the Politics Underlying Students' Postsecondary Transitions. *Journal of Education for Students Placed at Risk (JESPAR)*, 15(1), 9–26.
- Berger, J. B., and Milem, J. F. (2000). Organizational behavior in higher education and student outcomes. In: J. C. Smart (ed.), *Higher Education: Handbook of Theory and Research*. Agathon Press: New York, Vol. XV, pp. 268-338.
- Bettinger, E. (2004). How financial aid affects persistence. *NBER Working Paper*.
- Boatman, A. & Long, B. T. (2009). Does Financial Aid Impact College Student Engagement? The Effects of the Gates Millennium Scholars Program. Retrieved Feb.14, 2011, from http://gseacademic.harvard.edu/~longbr/Boatman_Long_Does_Aid_Impact_Student_Engagement_6-09.pdf.
- Bowen, W.G., Kurzweil, M. A., & Tobin, E.M. (2005). *Equity and excellence in American higher education*. Charlottesville, University of Virginia Press.
- Cabrera, A. F., Stampen, J. O., & Hansen, W. L. (1990). Exploring the effects of ability to pay on persistence in college. *Review in Higher Education*. 13(3), 303–336.
- Cabrera, A. F., Nora, A., & Castaneda, M. B. (1993). College Persistence: Structural Equations Modeling Test of an Integrated Model of Student Retention. *Journal of Higher Education*, 64(2).

- Cabrera, A. F., Burkum, K.R. & La Nasa, S. M. (2005). Pathways to a four-year degree: Determinants of transfer and degree completion. In Bowen, W.G., Kurzweil, M. A., & Tobin, E.M. (2005). *Equity and excellence in American higher education*. Charlottesville, University of Virginia Press. A. Seidman (Ed.). *College student retention: A formula for student success*. ACE/Praeger series on Higher Education.
- Cameron, A., Gardner, I., Doherr M.G., & Wagner, B. (2003). Sampling considerations in surveys and monitoring and surveillance systems. In *Animal disease surveillance and survey systems – methods and applications*. Edited by: Salman MD. Ames: Iowa State Press. 47-66.
- Carnevale, A.P. & S. J. Rose, (2003). Socioeconomic status, race/ethnicity and selective college admissions. The Century Foundation, New York. National Center for Education Statistics (1999).
- Carroll, C. D. (1987). Student financial assistance and its consequences. Presented for the 147th Annual Meeting of the American Statistical Association, San Francisco.
- Cellini, S.R. (2008). Causal inference and omitted variable bias in financial aid research: Assessing Solutions. *Review of Higher Education*, 31(3), Spring 2008: 329-354.
- Chen, R., & DesJardins, S. L. (2008). Exploring the effects of financial aid on the gap in student dropout risks by income level. *Research in Higher Education*, 49(1), 1–18.
- Cochran, W.G. (1977). *Sampling Techniques*, 3rd ed. John Wiley and Sons, USA.
- DesJardins, S. L. (2005). Investigating the efficacy of using selection modeling in research of the Gates Millennium Scholars Program. *Paper Presented to Gates Millennium Scholars Research Advisory Council*.
- DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (1994, May 29-June 1). *Studying the determinants of student stopout: Identifying "true" from spurious time-varying effects*. Paper presented at the 34th annual forum of the Association for Institutional Research, New Orleans, LA.
- DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (1999). An event history model of student departure. *Economics of Education Review*, 18(3), 375–390.
- DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (2002). Simulating the longitudinal effects of changes in financial aid on student departure from college. *Journal of Human Resources*, 37(3), 653–679.
- DesJardins, S. L., Kim, D. O., & Rzonca, C. S. (2003). A nested analysis of factors affecting bachelor's degree completion. *Journal of College Student Retention: Research, Theory and Practice*, 4(4), 407–435.
- Dowd, A. C. (2004). Income and financial aid effects on persistence and degree attainment in public colleges. *Educational Policy Analysis Archives*, 12(21).
- Dunn, W.N. (2008). *Public Policy Analysis – An Introduction, 4rd ed*, Harlow: Pearson Prentice Hall.
- Dynarski, S. M. (2003). Does aid matter? Measuring the effect of student aid on college attendance and completion. *American Economic Review*, 93(1), 279–288.

- Ehrenberg, R. G., Zhang, L., & Levin, J. (2006). Crafting a class: The tradeoff between merit scholarships and enrolling lower-income students. *The Review of Higher Education*, 29(2), 195–211.
- Engle, J., & Tinto, V. (2008). Moving Beyond Access: College Success for Low-Income, First-Generation Students. *Pell Institute for the Study of Opportunity in Higher Education*, 38.
- Farrell, E. (2007, February 2). High-income students get bulk of merit aid. *Chronicle of Higher Education*. Retrieved Feb. 13, 2011, from <http://chronicle.com/article/High-Income-Students-Get-Bulk/4219>
- Georgia Student Finance Commission. (2011). Georgia's HOPE program. Retrieved on Feb 20, 2011, <http://www.gsfc.org/gsfnew/HopeProgram.CFM?sec=3>.
- Greene, W. H. 1993. *Econometric analysis (2nd ed.)*. New York: Macmillan.
- Heller, D. E. (1999). The effects of tuition and state financial aid on public college enrollment. *Review of Higher Education*, 23, 65–90.
- Gross, J.P., & Torres, V. (2010). Competing risks or different pathways? An event history analysis of the relationship between financial aid and educational outcomes for Latinos. *Final Grant Report to the Center for Enrollment Research, Policy, and Practice*. Retrieved Feb 13, 2011, from <http://www.usc.edu/programs/cerpp/docs/Gross-Torres-SummaryReport.pdf>.
- Herzog, S. (2005). Measuring determinants of student return vs. dropout/stopout vs. transfer: A first-to-second year analysis of new freshmen. *Research in Higher Education*, 46(8), 883–928.
- Herzog, S. (2007). Estimating the influence of financial aid on student retention.
- Fayetteville, AR: Education Working Paper Archive. Retrieved February 15, 2011, from http://www.cis.unr.edu/IA_Web/RMAIR07Sherzog.pdf
- Hossler, D., Ziskin, M., Sooyeon, K., Osman, C., & Gross, J. (2008). Student aid and its role in encouraging persistence. *The effectiveness of student aid policies: What the research tells us* (pp. 101–116). Washington, DC: College Board.
- Hu, S., & St. John, E.P. (2001). Student persistence in a public higher education system. *The Journal of Higher Education*, 72(3), 265–286.
- Ishitani, T., & DesJardins, S. L. (2002). A longitudinal investigation of dropout from college in the United States. *Journal of College Student Retention*, 4(2), 173–201.
- Kane, T. J. (2003). A quasi-experimental estimate of the impact of financial aid on college-going. *NBER Working Paper, National Bureau of Economic Research*.
- Kane, T. J. (2004). *Evaluating the impact of the D.C. tuition assistance grant program*. National Bureau of Economic Research, Working Paper No. 10658. Cambridge, MA: National Bureau of Economic Research.

- Kim, M. M., Rhoades, G., and Woodard, D. B. (2003). Sponsored research versus graduating students: Intervening variables and unanticipated findings in public research universities. *Research in Higher Education* 44(1): 51-81.
- Kumar, R. (2005). *Research Methodology*. 2nd Edition. London: Sage Publications Ltd.
- McGrath, M., & Braunstein, A. (1997). The prediction of freshman attrition: An examination of the importance of certain demographic, academic, financial aid and social factors. *College Student Journal*, 31, 396-408
- McPherson, M., Schapiro, M. (1991). *Keeping College Affordable: Government and Educational Opportunities*. Washington, D.C.: Brookings Institution.
- Murdock, T. A. (1990). Financial Aid and Persistence: An Integrative Review of the Literature. *NASPA Journal*, 27(3), 213–21.
- National Center for Education Statistics. (1995). *Methodology Report for the 1993 National Postsecondary Student Aid Study*. Washington, DC: Author.
- National Center for Education Statistics. (1996). *Baccalaureate and Beyond Longitudinal Study: 1993/94 First Follow-up Methodology Report*. Washington, DC: Author.
- National Center for Education Statistics (1999). *Descriptive summary of 1995-96 Beginning Postsecondary Students*. National Center for Education Statistics, Statistical Analysis Report 1990-030. Washington DC: U.S. Department of Education, Office of Educational Research and Improvement.
- Nora, A., & Rendon, L. I. (1990). Determinants of predisposition to transfer among community college students: A structural model. *Research in Higher Education*, 31(3), 235–255.
- Osborne JW, editor.(2008). *Best Practices in Quantitative Methods*. Los Angeles: Sage Publications Inc.
- Pascarella, E. T., & Terenzini, P. T. (1980). Predicting freshman persistence and voluntary dropout decisions from a theoretical model. *The Journal of Higher Education*, 51(1), 60–75.
- Paulsen, M. B., & St. John, E. P. (2002). Social class and college costs: Examining the financial nexus between college choice and persistence. *Journal of Higher Education*, 73(2), 189–236.
- Peng, S. S., & Fetters, W. B. (1978). Variables involved in withdrawal during the first two years of college: Preliminary findings from the National Longitudinal Study of the High School Class of 1972. *American Educational Research Journal*, 15, 361-372.
- Perna, L. W., & Thomas, S. L. (2006). *A framework for reducing the college success gap and promoting success for all*. Washington, DC: National Postsecondary Education Cooperative.
- Price, D. V. (2004). *Borrowing inequality: Race, class, and student loans*. Colorado: Lynne Rienner Publishers, Inc.

- Pritchard, M. E., & Wilson, G. S. (2003). Using emotional and social factors to predict student success. *Journal of College Student Development*, 44(1), 18–28.
- Reynolds, C. L., & DesJardins, S. L. (2009). The use of matching methods in higher education research: Answering whether attendance at a 2-year institution results in differences in educational attainment. *Higher education: Handbook of theory and research*, 47–97.
- Richcardson Jr, R. C., & Kent, R. (2002). *Government policies and higher education performance in the U.S. and Mexico: report of a comparative study*. New York City: New York University & University of Puebla.
- Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, 516–524.
- Rubin, D. B. (2004). On principles for modeling propensity scores in medical research. *Pharmacoepidemiology and Drug Safety* 13: 855-857.
- Schneider, A., & Ingram, H. (1993). Social construction of target populations: Implications for political and policy. *The American Political Science Review*. 87(2): 334-347.
- Sekaran, U. (2003). *Research Methods for Business: A Skill Building Approach*, Southern Illinois University.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
- Singell Jr, L. D. (2002). Merit, need, and student self selection: is there discretion in the packaging of aid at a large public university? *Economics of Education Review*, 21(5), 445–454.
- Singer, J. D., & Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford/New York: Oxford University Press.
- Somers, P. (1995). A comprehensive model for examining the impact of financial aid on enrollment and persistence. *Journal of Student Financial Aid*, 25(1), 13–27.
- St. John, E. P. (1989). The evolving influence of student financial aid on persistence. *The Journal of Student Financial Aid*, Fall 1989, 19(3), 52-68.
- St. John, E. P. (1990). Price response in enrollment decisions: An analysis of the High School and Beyond sophomore cohort. *Research in Higher Education*, 31(2), 161-176.
- St John, E. P. (1991). The Effects of Student Financial Aid on Persistence: A Sequential Analysis. *Review of Higher Education*, 14(3), 383–406.
- St. John, E. P. (2004, August). *The impact of information and student aid on persistence: A review of research and discussion of experiments*. Paper prepared for David Mundel.
- St. John, E.P., Cabrera, A.F., Nora, A., & Asker, E.H. (2000). Economic influences on persistence reconsidered: How can finance research inform the reconceptualization of persistence models? In J.

- M. Braxton (Ed.), *Reworking the student departure puzzle : New theory and research on college student retention* (29–47). Nashville: Vanderbilt University Press.
- St. John, E., & Chung, A. (2006). Attainment. *Education and the Public Interest*, 193–213.
- St. John, E. P., Kirshstein, R. J., & Noell, J. (1991). The impact of student financial aid on persistence: A sequential analysis. *Review of Higher Education*, 14, 383-406.
- St. John, E. P., Paulsen, M. B., & Carter, D. F. (2005). Diversity, college costs, and postsecondary opportunity: An examination of the financial nexus between college choice and persistence for African Americans and Whites. *Journal of Higher Education*, 76(5), 545–569.
- St. John, E. P., & Starkey, J. B. (1995). An alternative to net price: Assessing the influence of prices and subsidies on within-year persistence. *Journal of Higher Education*, 66, 156-186.
- St. John, E. P., & Tuttle, T. J. (2004). *Financial aid and postsecondary opportunity for nontraditional-age, precollege students: The roles of information and the education delivery systems*. Indianapolis: The Education Research Institute and the Lumina Foundation.
- State Student Assistance Commission of Indiana. (n.d.). Twenty-first century scholars program. Retrieved on Feb 20, 2011, <http://www.in.gov/ssaci/2345.htm> .
- Stater, M. (2009). The impact of financial aid on college GPA at three flagship public institutions. *American Educational Research Journal*, 46(3), 782.
- Thomas, S.L. & Heck, R.H. (2001). Analysis of large-scale secondary data in higher education research: Potential peril associated with complex sampling design. *Research in Higher Education*. 42(5): 517-540.
- Tinto, V. (1982). Limits of theory and practice in student attrition. *Journal of Higher Education*, 53(6), 687–700.
- Tinto, V. (2004). *Student retention and graduation: Facing the truth, living with the consequences*. Washington, DC: The Pell Institute for the Study of Opportunity in Higher Education.
- Tinto, V. 2005. Moving beyond access: Closing the achievement gap in higher education. <http://faculty.soe.syr.edu/vtinto/Files/Harvard%202005%20Speech.pdf>, last accessed Jan 23, 2011.
- Titus, M. A. (2004). An examination of the influence of institutional context on student persistence at 4-year colleges and universities: A multilevel approach. *Research in Higher Education*, 45(7), 673–699.
- Titus, M. A. (2006). No College Student Left Behind: The Influence of Financial Aspects of a State's Higher Education Policy on College Completion. *The Review of Higher Education*, 29(3), 293–317.
- U.S. Department of Education. (2009). Federal Pell Grant Program. Retrieved on Feb 20, 2011, <http://www2.ed.gov/programs/fpg/index.html>

- Van der Klaauw, W. (2002). Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression–Discontinuity Approach*. *International Economic Review*, 43(4), 1249–1287.
- Voorhees, R. A. (1985). Student finances and campus-based financial aid: A structural model analysis of the persistence of high need freshmen. *Research in Higher Education*, 22, 65-92.
- Walpole, M. B. (2003). Socioeconomic Status and College. *The Review of Higher Education*, 27(1), 45–73.
- Weiss, C. (1998). *Evaluation*. Upper Saddle River, NJ: Prentice-Hall.
- Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA.
- Yamaguchi, Kazuo. 1991. *Event History Analysis*. Newbury Park, Calif.: Sage.
- Yi, P. (2008). Institutional climate and student departure: A multinomial multilevel modeling approach. *Review of Higher Education*, 31(2), 161.
- Zhang, L. (2010). The Use of Panel Data Models in Higher Education Policy Studies. *Higher Education: Handbook of Theory and Research*, 307–349.