



Destiny Unbound: A Look at How Far from Home Students Go to College

Robert K. Toutkoushian¹ · Stephen Mayfield² · Samantha M. R. Jelks³

Received: 24 June 2023 / Accepted: 2 April 2024 / Published online: 10 May 2024
© The Author(s), under exclusive licence to Springer Nature B.V. 2024

Abstract

One issue that has received little attention is how students factor distance from home into their decisions about college. In this study, we used data from the Education Longitudinal Survey of 2002 (ELS:02) to examine the distances between a student's home and the colleges to which they applied, and how far from home they enrolled. We focused on how demand- and supply-side factors were related to the distances applied and enrolled. We tested the sensitivity of our findings to alternative ways of measuring the supply of postsecondary education within commuting distance, and identified factors associated with differences between a student's application and enrollment distances. Finally, we used quantile regression analysis to determine if the associations between demand- and supply-side factors and distances applied and enrolled varied along the distance distributions.

Keywords Distance · College choice · Migration

Introduction

Researchers have long been interested in how students make decisions about whether and where to go to college. Students are faced with a multitude of choices with regard to postsecondary education. One of the first tasks for students is to narrow down the list of institutions that they will consider attending. Studies have shown that student decisions about where to go to college are influenced by the cost (Han, 2014; Lillis & Tian, 2008), and related to personal attributes such as academic ability, gender, and race/ethnicity (Dillon & Smith, 2017; Kim, 2004; Perna, 2006), family socioeconomic measures including parental education and income (Han, 2014; Niu, 2015; Perna, 2006), school/community attributes (Morton et al., 2018), and institutional characteristics.

One issue that has received relatively little attention is the role of distance from home in decisions about college. The distance from home to college is important because unlike

✉ Robert K. Toutkoushian
rtoutkou@uga.edu

¹ Institute of Higher Education, University of Georgia, Meigs Hall 114, Athens, GA 30602, USA

² RTI International, Durham, NC, USA

³ Institutional Research & Decision Support, Ellucian, Strategy & Planning, Managed Services, Atlanta, GA, USA

K-12 education, it is common for students to move to where a college is situated when they enroll. Students may select an institution close to where they live or hundreds or even thousands of miles away. We refer to the distance from home that a student considers going (application) or actually goes (enrollment) as the geographic span of their postsecondary market.

In theory, students make decisions about how far from home they are willing to travel as part of their college search process. As noted by Hossler and colleagues (1987, 1999), high school students progress through a series of stages where they form choice sets of colleges that they consider attending, which then influences where they enroll. It is important to understand how distance from home factors into someone's decision to include colleges in their early choice set because it bounds how far they actually go for college and can highlight differences across student types in their relative aspirational and real distances traveled for college.

Learning about how these distance decisions are made is important for several reasons. As the distance between a student's home and college increases, so will the monetary and non-monetary costs of going to college. The cost may rise substantially if the student has to move and/or if the student attends an out-of-state public institution where tuition rates are notably higher for out-of-state students. At the same time, students who are willing to attend college farther from home may benefit from having more college options from which to choose. Where a student lives may also matter for their geographic span because the number of local college options varies greatly across the US. Some regions of the US have relatively few college options (Dache-Gerbino, 2018; Hillman, 2016), and the lack of options may force students to consider colleges farther from home.

We have little information about how students make the decision of how far to travel from home to college. Most studies that collected information about distance such as Alm and Winters (2009), Frenette (2004, 2006), González Canché (2018), and McHugh and Morgan (1984) used it as an independent variable to help explain whether and where students enrolled in college. In contrast, work by Mortenson (1996, 1998), Pryor et al. (2006), and Mattern and Wyatt (2009) focused on distance to enrolled college as the main variable of interest. These studies found that the median distance that first-time, full-time students travel to attend a 4-year college was between 80 and 100 miles, and that the distribution of distances to college attended was highly skewed to the right.

As noted by Mattern and Wyatt (2009, p. 19), however, "very little is known about the impact of student characteristics on distance traveled to attend college." The aforementioned studies provided breakdowns of median distance to college attended by selected factors but did not use inferential methods such as regression analysis to estimate the relationships of these factors to distance and test for their statistical significance. Mortenson (1996, 1998), for example, showed how the median distance to college attended varied by family income, father's education, and race, but did not test for the significance of each of these associations after controlling for other factors.

In this study, we used data from the Education Longitudinal Survey of 2002 (ELS:02), which is a nationally representative survey of high school students, to examine the distances between a student's home and the colleges to which they applied and enrolled. We begin by quantifying the distances between a student's home and the institutions to which they applied and enrolled. The descriptive statistics from this part of the study add to what we know about distance by calculating alternative summary measures of the geographic span of where students applied to college as well as where they finally enrolled.

We then estimated the associations between a number of demand-side personal and family characteristics and the distances from home to the institutions where students applied

and where they enrolled. The analysis in this section allows us to identify the relationships for each variable holding constant the others in the model, and test for their statistical significance. In addition, we used data on both the number of 4-year institutions within commuting distance from their home and the relative number of students enrolled at these institutions to identify whether a low supply of spaces led to an increase in distance to college. By considering these two variables, we expand on how researchers can operationalize the opportunities that students have for attending college near their home.

Finally, we also applied quantile regression analysis to our models for distances applied and enrolled to test whether the results from our OLS models were uniform across the distance distributions. It is possible that demand and supply factors have different relationships with distances applied and enrolled along their distributions. For example, a positive relationship between coming from a highly-educated family and the mean distance to college could be concentrated among those students who go farther from home. Standard ordinary least squares (OLS) regression models focus on how changes in covariates are related to the mean of the distribution and may overlook important differences for students who apply and/or enroll closer to or farther from home.

Conceptual Frameworks

Student Choice Models

The framework of how students make decisions about going to college originates from the work of Hossler and colleagues (1987, 1999), and was extended by others including Perna (2006) and Toutkoushian and Paulsen (2016). These models posit that students progress through a series of stages when transitioning from high school to college, where students begin by forming the aspiration to go to college, and then successively narrow down their choices of institutions that they would want to attend. This process culminates in the student either selecting a single institution in which to enroll or opting not to attend college.

The student choice models give rise to alternative measures of demand for college, including whether the student submitted standardized test scores to an institution, applied for admission, or actually enrolled (DesJardins et al., 2006; Toutkoushian, 2001a). The demand for college may then be expressed as a function of personal characteristics of the student such as gender, race, age, and academic ability, characteristics of the student's family (typically parental education and income), and attributes of the institution. As noted by Cooke and Boyle (2011), the decision to consider attending a given college is also spatial in nature because students must either commute to and from the institution or move to where the institution is located.

Cost–Benefit Analysis

Within the student choice model, cost–benefit analysis is useful for examining the role that distance plays in the decisions students make about college. In cost–benefit analysis, individuals take into account the perceived private costs and benefits that they would incur from various consumptive decisions before making them. In particular, factors that drive up the cost of college would be predicted to reduce the demand for going to college. This idea can be applied to the process students use when deciding how far from home they would go for college.

The financial and non-financial costs to students associated with different colleges and universities will vary along with their proximity to the institution. The direct cost of attendance is affected by distance because public colleges charge lower tuition rates for resident students, and travel costs increase with distance. And as distance rises, there is a greater likelihood that a student will have to move to where the institution is located and possibly face higher room and board costs and a higher cost to travel between their college and family. There are also larger search costs that students incur as the distance increases. In addition, non-pecuniary costs may be greater for institutions that are farther from home as students must learn about a new location and adapt to not being in close proximity to family and friends. Accordingly, the cost–benefit framework predicts that students prefer institutions in closer proximity than those that are farther away.

Social and Cultural Capital

Social and cultural capital may also be part of the student choice model and help inform the decisions that students make about how far from home to go to college. Social capital refers to the relationships that individuals make with others that help them to succeed in society (Hill et al., 2015; Mishra, 2020). Although the concept of social capital dates back to the nineteenth century, it became popularized through the work of Bourdieu (1973, 1977, 1990) and Coleman (1988). Meanwhile, cultural capital refers to the social assets possessed by a person that enable them to successfully navigate through situations in society (Guillory, 2023; Marginson, 2019). Sandefur et al. (2006) posit that higher levels of social capital are significantly associated with college enrollment. Social and cultural capital may therefore shift the demand for higher education.

The two forms of capital may intertwine when an individual's social network allows for inherited and learned cultural capital (Bourdieu, 1973), allowing parents, teachers, and students' own educational performances over time to inform cultural capital. Factors like reading competency, for example, may be instilled and fostered by parents who perpetuate their own educational values (Evans et al., 2010), thus representing inherited cultural capital, while teachers reward students in an academic context, allowing for learned cultural capital (Jaeger & Breen, 2016). De Graaf et al. (2000) explain that because of these patterns in both the home and the school, cultural capital surrounding education is institutionalized, rewarding those students who possess higher levels of cultural capital and propelling them towards higher levels of education. Meanwhile, Tramonte and Willms (2010) acknowledge that students with less cultural capital and fewer college-educated individuals in their social networks may not have access to the same level of cultural competence related to higher education and therefore may not have the same tools to pursue education with such ease.

For this and other reasons, those with higher social and cultural capital may feel more confident in going farther from home to college. Despite this, higher social and cultural capital may not entirely negate the psychic distress of considering a college that, while it may otherwise be an option, remains in a geographic region farther from those initial social and cultural connections. Research suggests that the psychic costs of going to college will rise along with distance regardless of capital (Greenwood, 1975).

Higher Education Markets

The field of microeconomics focuses on the role of markets in facilitating transactions between buyers and sellers. In this framework, an industry consists of multiple markets where sellers compete for customers. This notion can be applied to the higher education industry in the United States, in that most students consider attending only a subset of institutions. Postsecondary markets are often local, state, or regional in nature because students are free to move to where the institution is located to use its services. Dill (1997) found that location has been a crucial factor for decades, and Stephenson et al. (2016) discussed how markets for higher education may depend on college visits by students and parents, which are easier to facilitate with institutions in the same state and region.

How students select institutions, and hence the geographic span of their markets, varies from student to student. Becker and Toutkoushian (2013) and Stephenson et al. (2016) stress that higher education markets vary by degree level, major, and type of student. Meanwhile, Turner (2017) posits that academic ability is also a factor in determining which markets students consider for their college education. Conversely, students from lower-income families have fewer resources to identify the higher education markets that suit their needs.

This study makes a number of contributions to the literature. First, the study measures distance to college using a national data on high school students who seek to pursue a postsecondary education. Second, the study considers both distance to colleges applied and enrolled. Third, the role of the supply of institutions within commuting distance of home is expanded by looking at alternative definitions of supply. And finally, the study uses both OLS and quantile regression analysis to identify how key factors are associated with distances applied and enrolled.

Literature Review

Proximity to College and Enrollment

There are two main strands of studies that have incorporated distances to college as a variable. The largest strand consists of studies that used distance to college as an independent variable to help explain the enrollment decisions of students (Do, 2004; Frenette, 2006; Griffith & Rothstein, 2009; Hirschl & Smith, 2020; Holloman & Nolen, 2008; Leppel, 1993; Spiess & Wrohlich, 2010; Tinto, 1973; Turley, 2009). These studies have shown that the distance students must travel to attend college is an important factor in their decisions about college. For example, Griffith and Rothstein (2009) observed that students who live closer to selective institutions were more likely to apply to selective institutions. Turley (2009) concluded that students who live close to larger numbers of colleges were more likely to apply, especially to 4-year institutions that were nearby. For these reasons, Hillman's (2016) work on higher education deserts becomes increasingly relevant when considering barriers to pursuing higher education. Studies by Byun et al. (2015), and Wells et al. (2019) examined whether living in a rural area was related to access to college.

Focusing on enrollment, Leppel (1993) found that distance and academic ability are two of the primary predictors of whether a student will attend an institution. Frenette (2006) and Spiess and Wrohlich (2010) noted a relationship between distance and the cost of attending college, suggesting that those who live closer to a college experience lower

financial costs of attendance. Hirschl and Smith (2020) argued that aside from pecuniary costs, there are fewer social and familial barriers to attendance when colleges are close to students' homes. Do (2004) found that in addition to lower costs of attendance, living near a good public university increased the quality of the college a student attended, particularly for those from lower-income families.

The distance from home to college, whether in or out of state, is connected to the likelihood and cost of attendance, pecuniary and otherwise, as well as student success. While Frenette (2004) suggests that larger distances from higher education institutions are a deterrent to prospective students due to increased travel costs, Garza and Fullerton (2018) determined that staying close to home may impede academic progress for first-generation students. Beyond college costs, distance from college can impact a student's major and career path (Denzler & Wolter, 2011; Klasik et al., 2018). Those students in rural settings with fewer nearby colleges may not only have fewer choices but may also be less likely to attend college at all.

Other studies within this strand investigated the migration of students across state borders (Baryla & Dotterweich, 2001; Cooke & Boyle, 2011; Ghatak et al., 1996; Greenwood, 1975; Ishitani, 2011; McHugh & Morgan, 1984; Mixon, 1992; Mixon & Hsing, 1994; Toutkoushian & Hillman, 2012; Toutkoushian, 2001b; Tuckman, 1970; Zhang & Ness, 2010). Tuckman (1970) and McHugh and Morgan (1984) found that states in close proximity will absorb more students who pursue higher education outside their home state. Although out-of-state migration is related to distance, as noted by McHugh and Morgan (1984) the concept of distance is not the same across states. Even within states, distance may vary considerably and influence the enrollment decisions of students (Alm & Winters, 2009). González Canché (2018), for example, argued that the distance some students travel to college within large states is comparable to going out of state in other locations, and thus the simple in-state versus out-of-state dichotomy does not fully capture distance to college.

Descriptive Studies of Distance to College

The second strand of literature focused on summary measures of how far from home students went to attend college (Eagan et al., 2016; Mattern & Wyatt, 2009; Mortenson, 1996, 1998; Pryor et al., 2006). Two main results emerged from these studies: (1) The average distance between a student's home and where they went to college was between 80 and 100 miles, and (2) students were much more likely to attend colleges that were close to home rather than far away (i.e., the distribution was skewed to the right).

These studies also provided some insights as to why distances to colleges might vary by student and family characteristics. Mortenson (1996), for example, showed that the median distances to college attended increased with family income and the father's educational attainment, and that Chicano and Asian students had lower median distances than students in other racial/ethnic categories. Mattern and Wyatt (2009) likewise showed similar patterns in median distance attended, and also found that the median distances attended were greater for students with higher academic credentials.

There are, however, limitations with the descriptive studies of distance. First, they did not conduct tests to determine whether the differences in average distance to college attended by these factors were due to random chance or in fact represent a relationship between them. Second, the observed differences in median distance by, say, family income could be due to other factors that are correlated with family income. Finally, these studies focused exclusively on

the distance to college attended and did not provide information on the distances to colleges applied.

Issues in Measuring Distance

Studies in both strands of the aforementioned literature have had to address several issues pertaining to the measurement of distance to college. How studies addressed these issues depends in large part on the data at their disposal.

The first issue is what is meant by distance to college. Studies that examined whether distance was related to enrolling in any college often used the distance to the nearest college. Alm and Winters (2009), for example, found that as the distance to the nearest college in Georgia increased, students were less likely to enroll at a public 4-year institution in the state. Frenette (2006) likewise examined whether the distance to the nearest college in Canada was related to the probability of a student going to college. Other studies including Mattern and Wyatt (2009) and DesJardins et al. (2006) focused on the distance between a student and the specific institution that they attended. Finally, a distinction could be made between distance to the college attended versus distance to a college applied. Most studies only consider distance to college attended, in part because data on colleges to which students applied is often not available.

The second issue for measuring distance is how distance should be calculated. The descriptive reports by Pryor et al. (2006), Eagan et al. (2016), and Mortenson (1996, 1998) relied on self-reported distances from surveys of freshmen, where they were asked how far they traveled to attend college. Another frequently-used option for measuring distance was to calculate the distance between each student's home ZIP code and the ZIP code of the institution (Mattern & Wyatt, 2009; Frenette, 2004, 2006; González Canché 2018). This option is useful when the specific home address of the student is unknown. A similar approach was used in studies of out of state migration based on the ZIP codes of population centers (Cooke & Boyle, 2011; McHugh & Morgan, 1984). When the specific home address of the student was available (DesJardins et al., 2006), studies could better identify the distance from each student's home to their institution.

Finally, should distance be based on geography or travel? It is most common for distance measures to be calculated as the straight-line geographic distance between two points. This approach, however, may underestimate the distance that students must travel to get to college due to barriers that inhibit straight-line travel (e.g., lakes, winding roads), or areas where there is more traffic congestion. Denzler and Wolter (2011) addressed this issue in their study of Switzerland by calculating the distance students would travel to the nearest university using public transportation. Likewise, Gibbons and Vignoles (2012) measured the "rail distance" between each student's home and the three nearest institutions in England. Alm and Winters (2009), however, argued for the use of straight-line distance on the grounds that it is easier to obtain and very highly correlated with alternative measures of travel time.

Data and Methods

Data

We relied on data from ELS:02 conducted by the National Center for Education Statistics. ELS:02 is a nationally representative survey of a cohort of more than 16,000 10th grade students in 2002 who were then followed up in 2004, 2006, and 2012. We selected ELS:02

for this study for several reasons. First, it is a nationally representative dataset, and thus allows us to generalize findings to the broader US student population. Second, the survey collected information from students during the college search process in high school and their entry into postsecondary education. Finally, the data had sufficient information needed to construct the necessary distance measures, and included data on all colleges to which a student applied.¹ We limited our analysis to students in the 12th grade sample who lived in the contiguous United States.² For the college search stage, we examined high school seniors who applied to at least one American 4-year college within 2 years of high school. Similarly, the college enrollment sample included high school seniors who enrolled in an American 4-year college within 2 years of high school. After dropping observations with missing data on the variables of interest or applications/enrollments outside of the United States, our final sample consisted of approximately 6540 students at the application stage and 5170 at the enrollment stage.³

Dependent Variables

The dependent variables in the models that follow were derived from estimates of the distance between where a student resided and the 4-year institutions to which they applied or enrolled. We first assigned each student a longitude and latitude based on the geometric center of their home ZIP code. Likewise, we assigned a longitude and latitude to each institution based on the geometric center of its ZIP code.⁴ We then created a measure of the straight-line distance in miles from the student's home to each institution.⁵

Given that students often applied to more than one institution, we used three different summary statistics (mean, median, maximum) to represent the geographic span of their applications. The mean and median distances applied represent the central tendency or average distance students applied to colleges. The median is particularly important in this context due to the skewness of application distances. Furthermore, the maximum distance applied provides an estimate of how far from home a student was willing to go to college.

Demand-Side Variables

We constructed a set of covariates that relate to the demand-side of postsecondary markets. These variables were grouped into categories for student demographic and academic characteristics, family characteristics, and their local area socioeconomic characteristics.

¹ Although it is a more recent survey, the High School Longitudinal Study of 2009 (HSL:09) was not used because it only provides information on up to three institutions to which a student applied, and does not contain the student's ZIP code during high school.

² Applications and enrollments from students from the mainland to colleges in Hawaii and Alaska were included.

³ All sample sizes are rounded per NCES requirements. The data used in all tables have been weighted using the ELS variable F2F1WT to account for the complex sampling design used in the survey.

⁴ The student ZIP code was collected in the 2004 ELS first Follow-up Survey. We also used this ZIP code to infer the students' county, which is in turn aggregated up to the commuting zone level. The ZIP codes for colleges were gathered from the 2004 IPEDS file. If no college ZIP code existed in IPEDS 2004, we used the ZIP codes as reported in IPEDS surveys for other years.

⁵ The straight-line distances were derived using Stata's Geodist command. We acknowledge that there are other ways to measure distance, including the commuting distance between two points in miles and time. The computation of these alternate measures would require connecting the computer with the restricted-use data to the internet, which is prohibited by the licensing agreement with IES.

The demographic characteristics included the student's gender, race/ethnicity, and age as of grade 12, as prior studies have shown that demographic characteristics are often related to choices that students make regarding college. The academic characteristics included the student's high school GPA and SAT score. With regard to family characteristics, we created variables for family income (five categories), highest parental education level (four categories), whether the student lived with two heads of household, and whether the student changed high schools between 10 and 12th grade.

Supply-Side Variables

We used another set of variables to test whether the distance to college was influenced by the supply of college opportunities close to where students live. Hillman (2016) argued that the availability of postsecondary options in the vicinity of a student was an important factor in facilitating access to higher education. He defined a set of “commuting zones” that were geographic areas for a student based on clusters of bordering counties.⁶ He then used the number of institutions (2- and 4-year) in each commuting zone as a measure of the availability of postsecondary education within proximity of each student.

We extended this idea by first updating the set of commuting zones to those created from 2010 US Census data and by creating four separate supply-side variables for each commuting zone. The first two variables were the number of 2- and 4-year institutions within a student's commuting zone. We also used data on the number of undergraduates enrolled at these institutions to construct two additional variables for the ratios of 2- and 4-year enrollments to the population of 18- to 24-year-olds within each commuting zone. Students may view commuting zones with higher enrollment ratios as having a greater relative chance of enrollment close to home, which in turn could reduce the distance that students travel to college. We separated institutions by level so that we could examine whether the distance students applied/enrolled to 4-year institutions was related differently to the supply of 2- versus 4-year colleges. The economic theory of markets would predict that the effects of enrollment ratios for 2- and 4-year institutions would be different if the students who wish to pursue a bachelor's degree view 4-year institutions as closer substitutes than 2-year institutions.

Locational Variables

Finally, we created control variables to capture the effects of locational attributes where a student resides on the distance that they travel for college. Variables that we identified at the student's commuting zone level include the poverty rate, the percent of adults with at least a bachelor's degree, the median family income, and the unemployment rate.⁷ These factors relate to the socioeconomic climate and labor market opportunities that students faced close to home, which in turn may influence how far from home they decided to go to college. We also created a set of dummy variables for the student's state of residence,

⁶ The clusters of counties for each commuting zone were created by the USDA Economic Research Service using “journey to work” data from the 2000 US Census.

⁷ Median family income was generated as a weighted average of the counties that comprise each commuting zone (i.e., median household income times the number of households in county). The unemployment rate per commuting zone was an average for the years 2004–2006.

which captures the net effect of state-specific traits that impact postsecondary opportunities, such as the average price of attendance, desirability of the state for college, and geographic size.

Methods

The next step was to estimate the relationships between key explanatory variables and the mean, median, and maximum distances applied, and the distance enrolled. Because of the skewness of the dependent variables, we used the natural log of distance measures in our models. In each case, we used and reported robust standard errors,⁸ and all models were weighted to take into account the stratified sampling design used by NCES for ELS:02. The primary models relied on multiple regression analysis. The first regression model focused on the determinants of the log of distance applied ($\ln(DISTAPP)$), as in:

$$\ln(DISTAPP) = \alpha + P\beta + F\gamma + A\delta + L\theta + S\rho + \varepsilon \quad (1)$$

where P =personal factors, F =family factors, A =academic factors, L =locational factors, S =supply-side factors (with parameters α , β , γ , δ , θ , and ρ respectively), and ε =error term. The sample for these three models included those students who had applied to at least one 4-year college or university ($N=6540$). As noted, we measured $DISTAPP$ in three ways to account for the different distance applied statistics that could be used (mean, median, or maximum).

We estimated a similar multiple regression model for the log of distance enrolled:

$$\ln(DISTENR) = \alpha + P\beta + F\gamma + A\delta + L\theta + S\rho + \varepsilon \quad (2)$$

The model was restricted to those students who had enrolled in a 4-year institution ($N=5170$). The parameters in Eqs. (1) and (2) were estimated with OLS regression analysis. Accordingly, these coefficients denote the approximate percentage change in mean distances associated with one-unit changes in each variable.

We also modified the distance enrolled Eq. (2) as follows:

$$\ln(DISTENR) = \alpha + P\beta + F\gamma + A\delta + L\theta + S\rho + \omega(\ln DISTAPP) + \varepsilon \quad (3)$$

to examine the relationship between distances applied and distances enrolled. If the coefficient on the variable $\ln(DISTAPP)$ is less than one, it would suggest that distances enrolled were on average less than the distances applied. Equation (3) also enabled us to test whether the variables in the model had different relationships with distances enrolled after taking into account the distances applied. For example, it may be that students from higher-income families were less likely than other students to reduce the distances between colleges where they applied and where they enrolled because ability to pay for college was less of a constraint for them.

We supplemented the primary OLS regression models for distance applied and enrolled with additional models using quantile regression analysis. The advantage of quantile regression analysis is that it allowed us to determine whether the associations between the

⁸ We also estimated all of the models using standard errors clustered at the state level and found no appreciable differences in the significance levels for the variables.

Table 1 Summary statistics for distance from home to college of 4-year college applicants and enrollees

Variables	Mean	SD	Median	Maximum
Distance by application	255	453	98	4850
Distance by student				
Mean distance applied	198	309	97	4850
Median distance applied	182	320	87	4850
Maximum distance applied	343	538	148	4850
Distance enrolled	192	376	74	4850

Distances are measured in miles. Sample includes approximately 19,720 applications, 6540 applicants, and 5170 enrollees (all numbers are rounded per NCES requirements). Includes only ELS senior cohort members living in the mainland United States who applied to at least one 4-year college in the United States. Data are weighted to take into account the stratified sampling design used in ELS

explanatory variables and distances applied and enrolled varied along the distributions of distances.⁹ The models we estimated were of the form:

$$\ln(DISTAPP)^q = \alpha^q + P\beta^q + F\gamma^q + A\delta^q + L\theta^q + S\rho^q + \varepsilon^q \tag{4}$$

$$\ln(DISTENR)^q = \alpha^q + P\beta^q + F\gamma^q + A\delta^q + L\theta^q + S\rho^q + \varepsilon^q \tag{5}$$

where the superscript *q* denotes the *q*-th percentile of the distance distribution. We estimated Eqs. (4) and (5) at the 10th, 25th, 50th, 75th, and 90th percentiles of the distance distributions. Quantile regression analysis provides a useful robustness check to determine if the findings from the OLS model equally apply to students along the distance distributions. It is quite possible, for example, that factors such as student academic ability, parental education, and the supply of spaces near a student’s home have different associations with distance along the distribution.

Results

Descriptive Statistics

Table 1 provides descriptive statistics for the different ways of measuring the distance between each student’s home and the colleges to which they applied and/or enrolled. The table is divided into five main groups. In the first group each *application* was used as the unit of analysis, and thus the figures represent summary statistics on the distances of applications to 4-year institutions. The next three groups used each *student* as the unit of analysis to take into account that a student may have multiple application distances. Finally, the last group examined distances for the 4-year institutions where students enrolled.

⁹ While ordinary least squares selects coefficients that minimize the sum of squared deviations, quantile regression selects coefficients that minimize the sum of weighted absolute deviations. The weights in quantile regression analysis allow separate parameters to be estimated for different percentiles of the distribution. More details on quantile regression analysis can be found in Koenker and Hallock (2001).

Table 2 Descriptive statistics of ELS:02 4-year college applicants

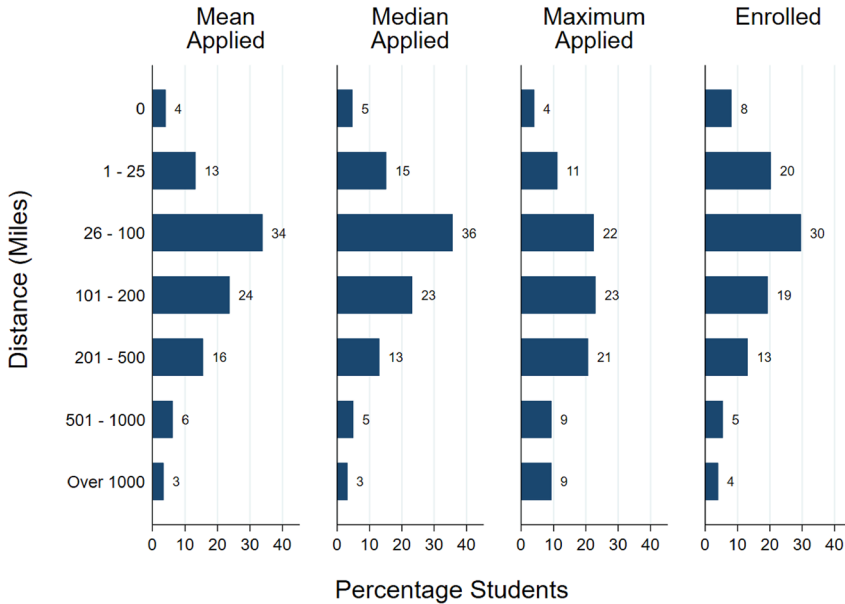
Variables	Mean	SD	Min	Max
Male	0.45	0.50	0.00	1.00
Age > 18 years	0.34	0.48	0.00	1.00
White	0.69	0.46	0.00	1.00
Black	0.12	0.33	0.00	1.00
Hispanic	0.09	0.29	0.00	1.00
Asian	0.05	0.22	0.00	1.00
Other race(s)	0.05	0.21	0.00	1.00
High school GPA	3.14	0.58	0.85	4.00
SAT composite score (50 pts.)	20.61	3.95	8.40	32.00
Attend same high school 10th and 12th	0.95	0.21	0.00	1.00
Family income: \$20K or less	0.08	0.27	0.00	1.00
Family income: \$20K–\$50K	0.32	0.47	0.00	1.00
Family income: \$50K–\$75K	0.23	0.42	0.00	1.00
Family income: \$75K–\$100K	0.17	0.38	0.00	1.00
Family income: \$100K or more	0.20	0.40	0.00	1.00
Parent education: high school	0.17	0.37	0.00	1.00
Parent education: associate's degree	0.20	0.40	0.00	1.00
Parent education: bachelor's degree	0.40	0.49	0.00	1.00
Parent education: graduate degree	0.24	0.43	0.00	1.00
Reside w/two heads of household	0.81	0.40	0.00	1.00
Commuting zone ^a				
Poverty rate (%)	13.46	4.09	6.00	35.62
Bachelor's or higher (%)	27.19	7.40	10.37	46.73
Median household income (\$1000) ^b	52.40	10.68	26.27	84.24
Unemployment rate (%)	5.13	1.13	2.62	15.59
Number of 4-year colleges	26.87	31.89	0.00	119.00
Ratio: 4-year enrollment to pop 18–24	0.32	0.16	0.00	2.54
Number of 2-year colleges	13.95	17.19	0.00	85.00
Ratio: 2-year enrollment to pop 18–24	0.24	0.13	0.00	0.94

Sample size = 6540 applicants (rounded per NCES requirements). Data are weighted to take into account the stratified sampling design used in ELS. Includes only ELS senior cohort members living in the mainland United States who applied to at least one 4-year college in the United States

^a2010 Commuting Zones. Socioeconomic data from the US Census 2009 5-Year American Community Survey and US Bureau of Labor Statistics Local Area Unemployment Statistics program, which was aggregated from county-level data to the commuting zone level. Postsecondary data is from the 2009 NCES IPEDS survey, which is the earliest collection of college county locations

^bWeighted average of the counties within the commuting zone

The skewness of distance data can be seen in Table 1, where the means were consistently higher than the medians. Starting with the first group, of the 19,720 applications the mean distance applied was 255 miles and the median was 98 miles. The average distances that students applied varied between 182 and 198 miles depending on the statistic used to represent average distance. The mean distance enrolled was very similar to the



Notes. $N = 6,540$ for mean, median, and maximum application distributions. $N = 5,170$ for enrollment distribution. Data are weighted to take into account the stratified sampling design used in ELS.

Fig. 1 Frequency distributions of the student distances applied and enrolled. $N=6540$ for mean, median, and maximum application distributions. $N=5170$ for enrollment distribution. Data are weighted to take into account the stratified sampling design used in ELS

mean distance applied; however, the median distance enrolled (74) was notably lower than the median distance applied (97).

The descriptive statistics for the explanatory variables in the study for the set of college applicants are shown in Table 2. Among the demographic factors, the majority of applicants were female, and about 69% of college applicants identified as white. Consistent with prior literature, college applicants were more likely to come from families where their parents were college educated. Finally, with regard to the supply-side constructs, there was a wide range of availability of 2- and 4-year college opportunities for students based on where they live.

Figure 1 provides a visual depiction of the distributions of the mean, median, and maximum distances students applied, and the distances enrolled. The figures illustrate the skewness in the distance data and the regional nature of postsecondary markets. More than half of the students applied to colleges that were, on average, within 100 miles of their home, and less than 10% of these students applied to colleges that were, on average, more than 500 miles from home. The distribution of distances enrolled was even more highly skewed to the right than were the distributions of applications.

OLS Regression Models

The results from our primary OLS regression models are shown in Table 3. The first three columns correspond to the log of mean, median, and maximum distances students applied,

Table 3 Determinants of log of student distance to colleges applied and enrolled

Variable	Log of distances applied using			Log of distance enrolled
	Mean	Median	Maximum	
Male	0.098*	0.099*	0.089	0.018
	(0.044)	(0.045)	(0.049)	(0.064)
Age >18 years	0.075	0.080	0.065	0.097
	(0.047)	(0.048)	(0.051)	(0.064)
Black	0.629***	0.578***	0.798***	0.479***
	(0.079)	(0.082)	(0.087)	(0.117)
Hispanic	-0.126	-0.219*	-0.057	-0.308*
	(0.104)	(0.106)	(0.116)	(0.146)
Asian	-0.153	-0.276***	-0.050	-0.469***
	(0.082)	(0.082)	(0.092)	(0.105)
Other race(s)	0.351**	0.302**	0.404***	0.017
	(0.107)	(0.114)	(0.116)	(0.167)
High school GPA	0.117*	0.153**	0.147*	0.074
	(0.053)	(0.055)	(0.060)	(0.081)
SAT composite score (50 pts.)	0.083***	0.075***	0.097***	0.069***
	(0.008)	(0.008)	(0.009)	(0.012)
Attend same high school 10th and 12th	-0.229*	-0.226*	-0.224*	-0.046
	(0.103)	(0.103)	(0.114)	(0.158)
Family income: \$20K or less	-0.143	-0.159	-0.162	-0.191
	(0.105)	(0.103)	(0.117)	(0.132)
Family income: \$20K–\$50K	0.022	0.009	0.021	0.034
	(0.063)	(0.064)	(0.070)	(0.088)
Family income: \$75K–\$100K	0.146*	0.134	0.155*	0.242**
	(0.070)	(0.071)	(0.077)	(0.093)
Family income: \$100K or more	0.353***	0.312***	0.396***	0.445***
	(0.066)	(0.068)	(0.073)	(0.095)
Parent Ed: associates	0.025	0.011	0.034	-0.068
	(0.076)	(0.076)	(0.085)	(0.108)
Parent Ed: bachelors	0.109	0.086	0.128	0.117
	(0.073)	(0.073)	(0.081)	(0.101)
Parent Ed: graduate	0.303***	0.293***	0.348***	0.247*
	(0.083)	(0.083)	(0.092)	(0.115)
Live w/two heads of household	-0.168**	-0.151*	-0.188**	-0.039
	(0.060)	(0.061)	(0.066)	(0.087)
CZ poverty rate (%)	0.020	0.023	0.030*	0.002
	(0.012)	(0.012)	(0.013)	(0.018)
CZ bachelor's or higher (%)	-0.029**	-0.028**	-0.024**	-0.040**
	(0.009)	(0.009)	(0.009)	(0.013)
CZ median income (\$1000s)	0.030***	0.031***	0.032***	0.036**
	(0.008)	(0.008)	(0.009)	(0.012)
CZ unemployment rate (%)	-0.046	-0.042	-0.053	0.024
	(0.031)	(0.031)	(0.034)	(0.041)

Table 3 (continued)

Variable	Log of distances applied using			Log of distance enrolled
	Mean	Median	Maximum	
CZ number of 4-year colleges (10s)	−0.028* (0.011)	−0.048*** (0.011)	−0.017 (0.012)	−0.037* (0.015)
CZ ratio of 4-year enroll to pop	−0.733*** (0.210)	−0.797*** (0.220)	−0.673** (0.210)	−0.901** (0.331)
Constant	1.656*** (0.466)	1.554*** (0.466)	1.277* (0.516)	1.223 (0.718)
R-squared	0.149	0.143	0.150	0.122

Sample size is 6540 applicants and 5170 enrollees (rounded per NCES requirements). Robust standard errors in parentheses. All data were weighted to account for the stratified sampling design used in ELS:02. Results for state covariates are withheld for brevity. Reference category for Race is White. Reference category for Family Income is \$50,000–\$70,000. Reference category for Parent's Highest Education is High School

CZ commuting zone

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

and the last column contains the results for the log of distance enrolled. Each model also included controls for state of residence.¹⁰

Beginning with applications, for most variables the results were comparable across the different summary statistics. Males applied to colleges that were about 10% farther from home than did females. There were sizable differences in median distances applied by race/ethnicity even after controlling for the other factors in the model. Black students, for example, applied to colleges that were about 60% farther from home than did white students. In contrast, the median distances applied by Hispanic and Asian students were about 20–30% closer to home than for white students, respectively. Students of higher academic ability, as measured by high school GPA and/or SAT scores, had significantly larger geographic spans than other students. For example, a 50-point increase in SAT scores was associated with a 7–10% increase in distance applied, and a 7% increase in distance enrolled. The results for the academic variables were consistent with the notion that students of higher academic ability may apply to colleges farther from home because they believe that they could gain admission to more institutions. In addition, students with high-income parents and/or highly educated parents also applied to colleges that tended to be farther from home. For example, students whose parents had graduate degrees applied to institutions that were about 30% farther from home compared to students whose parents did not earn a postsecondary degree. And students living with two parents applied to colleges that were about 15–20% closer to home.

Regarding the commuting zone variables, students who lived in areas that had higher median incomes tended to apply to colleges that were farther from home. In contrast, though, students who lived in areas where the educational attainment level was high were less likely to apply to colleges that were far from home. The data also showed that students

¹⁰ Due to space limitations, the results for the state variables in the models that follow are not shown but are available upon request.

Table 4 Determinants of log of student distance to college using alternative measures of postsecondary supply

Variable	Model 1	Model 2	Model 3
<i>Y = log of median distance applied</i>			
CZ number of 4-year colleges (10s)	−0.048*** (0.011)	−0.048*** (0.011)	−0.032 (0.025)
CZ ratio of 4-year enroll to pop	−0.797*** (0.220)	−0.774*** (0.234)	−0.774*** (0.233)
CZ number of 2-year colleges (10s)	–	–	−0.030 (0.043)
CZ ratio of 2-year enroll to pop	–	0.119 (0.225)	0.145 (0.227)
<i>Y = log of distance enrolled</i>			
CZ number of 4-year colleges (10s)	−0.037* (0.015)	−0.036* (0.015)	−0.018 (0.032)
CZ ratio of 4-year enroll to pop	−0.901** (0.331)	−0.821* (0.340)	−0.821* (0.340)
CZ number of 2-year colleges (10s)	–	–	−0.036 (0.059)
CZ ratio of 2-year enroll to pop	–	0.451 (0.313)	0.486 (0.319)

Sample size is 6540 applicants and 5170 enrollees (rounded per NCES requirements). Robust standard errors in parentheses. All data were weighted to account for the stratified sampling design used in ELS:02. Each model also included control variables for gender, age, race/ethnicity, high school GPA, SAT composite score, whether student attended the same high school in grades 10 and 12, family income, parental education, whether lived with two parents, commuting zone poverty rate, commuting zone educational attainment, commuting zone median family income, commuting zone unemployment rate, and state of residence
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

who lived in commuting zones with more 4-year postsecondary institutions nearby, or where there were more opportunities to enroll at these institutions, applied to colleges that were closer to home. These two findings support the work of Hillman (2016) who noted that the geographic differences in the supply opportunities of postsecondary education may affect college access.

Turning to the last column, we observed similar results for distances enrolled. For example, Black students enrolled farther from home and Hispanic and Asian students enrolled closer to home relative to white students. Likewise, students from high-income families and highly educated parents chose to enroll farther from home than other students. Although we found a positive association between a student's SAT score and distance enrolled, high school GPA was not statistically significant in the model for distance enrolled. Unlike the results for distance applied, the variables for gender, attending the same high school in grades 10 and 12, and living with two parents were not related to how far from home a student enrolled in college. Finally, students attended college closer to home when there were either more 4-year institutions or more seats per student within proximity to home.

In Table 4, we examined whether the results from the OLS models were affected by how we represented the supply of nearby postsecondary options. The previous models focused solely on the number of 4-year institutions and enrollments in these institutions within

commuting distance. The results for this model are repeated here as Model 1. However, it is possible that the distances students apply and enroll in 4-year institutions are also related to the supply of 2-year institutions. The relationships would be negative if students planning on a 4-year degree also considered enrolling at a 2-year institution, which would typically be in close proximity to their home. In Model 2, we added to the model a variable for the ratio of enrollments in 2-year institutions to population ages 18–24. Likewise, in Model 3 we added a variable for the number of 2-year institutions to the specification in Model 2.

We found that when we added a control variable for the ratio of 2-year enrollments to population to the model, the results for both the number of 4-year institutions and the ratio of 4-year enrollments to population remained negative and statistically significant. There was no similar negative association between the 2-year enrollment ratio and the median application distance or the enrollment distance. In contrast, adding the variable for the number of 2-year institutions led to the number of 4-year institutions no longer being statistically significant. Further investigation revealed that this was due to the very high correlation between the number of 2- and 4-year institutions in commuting zones ($r = +0.87$), which contributed to the substantial increase in the standard errors for these two variables.

Quantile Regression Models

Next, in Tables 5 and 6 we used quantile regression analysis to estimate the relationships of the covariates to the mean distances applied and distance enrolled at the 10th, 25th, 50th, 75th and 90th percentiles of the distance distributions.¹¹ Figure 2 presents selected findings from the quantile regression models for distances applied and enrolled, respectively. We found that the size and even the direction of the relationship between high school GPA and distances applied/enrolled varied by distance percentile, with positive relationships for those who stayed closer to home and negative relationships for those who went farther from home. A similar pattern for applications was seen for students from higher-income families, and yet the coefficients for high-family income did not vary for distances enrolled. The opposite pattern occurred for students whose parents had graduate degrees, and for the share of population enrolled in college within their commuting zone. For these two variables, the coefficients went from negative for shorter distances to positive for larger distances.

Finally, in Table 7 we examined the relationship between application and enrollment distances by adding the log of distance applied (mean, median, or maximum) to the list of explanatory variables for the distance enrolled model. The coefficient on each of the distance applied variables provides insight into whether students tended to enroll at institutions that were closer or farther from home than the institutions to which they applied.

The results in the first three rows show that distance enrolled was positively related to distance applied as expected, and the coefficients for each variable were significantly less than one. This suggests that as distance applied increased so did the distance enrolled but at a slower rate. The finding is consistent with the descriptive statistics showing that the geographic span of enrolled markets is smaller than the span of application markets. Not surprisingly, most of the other regressors in the model were not statistically significant, indicating that these factors are connected to application and enrollment distances in similar ways. In contrast, however, Asian students chose to enroll much closer to home even after

¹¹ The results were fairly robust when we used either the median or maximum distances applied.

Table 5 Quantile regressions of log of mean student distance applied

Variable	(1) OLS	(2) 10th	(3) 25th	(4) Median	(5) 75th	(6) 90th
Male	0.098* (0.044)	0.142* (0.059)	0.100** (0.036)	0.054 (0.031)	0.075* (0.033)	0.011 (0.038)
Black	0.629*** (0.079)	0.564*** (0.117)	0.677*** (0.085)	0.613*** (0.045)	0.609*** (0.053)	0.421*** (0.059)
Hispanic	-0.126 (0.104)	-0.540* (0.214)	-0.173 (0.128)	-0.063 (0.098)	0.131* (0.064)	0.092 (0.145)
Asian	-0.153 (0.082)	-0.146 (0.093)	-0.186** (0.071)	-0.087 (0.047)	-0.027 (0.066)	0.024 (0.084)
Other race(s)	0.351** (0.107)	0.079 (0.120)	0.091 (0.135)	0.212 (0.120)	0.394*** (0.110)	0.515*** (0.068)
High school GPA	0.117* (0.053)	0.281*** (0.068)	0.140** (0.047)	0.062 (0.036)	0.009 (0.033)	-0.094* (0.046)
SAT composite score (50s)	0.083*** (0.008)	0.081*** (0.011)	0.080*** (0.006)	0.077*** (0.006)	0.081*** (0.005)	0.071*** (0.007)
Attend same high school 10th and 12th	-0.229* (0.103)	-0.118 (0.125)	-0.071 (0.083)	-0.192 (0.104)	-0.395*** (0.051)	-0.316*** (0.065)
Family income: \$20K or less	-0.143 (0.105)	-0.821*** (0.124)	-0.158 (0.118)	-0.125 (0.069)	-0.097 (0.079)	0.053 (0.071)
Family income: \$20K–\$50K	0.022 (0.063)	0.006 (0.082)	0.026 (0.051)	-0.005 (0.035)	0.005 (0.043)	0.136* (0.062)
Family income: \$75K–\$100K	0.146* (0.070)	0.125 (0.090)	0.160* (0.066)	0.158** (0.051)	0.145*** (0.044)	0.145* (0.064)
Family income: \$100K or more	0.353*** (0.066)	0.492*** (0.085)	0.385*** (0.056)	0.313*** (0.047)	0.255*** (0.045)	0.183** (0.056)
Parent Ed: associates	0.025 (0.076)	-0.105 (0.126)	0.102 (0.057)	0.056 (0.046)	-0.086 (0.055)	-0.110 (0.067)
Parent Ed: bachelors	0.109 (0.073)	-0.028 (0.087)	0.174*** (0.048)	0.149** (0.048)	0.091 (0.056)	0.143* (0.064)
Parent Ed: graduate	0.303*** (0.083)	0.096 (0.111)	0.294*** (0.069)	0.293*** (0.059)	0.356*** (0.062)	0.303*** (0.064)
Reside w/two heads of household	-0.168** (0.060)	-0.335*** (0.076)	-0.153* (0.065)	-0.203*** (0.052)	-0.205*** (0.060)	-0.128** (0.048)
CZ number of 4-year colleges (10s)	-0.028* (0.011)	-0.035 (0.024)	-0.062*** (0.013)	-0.025* (0.012)	0.009 (0.009)	-0.019 (0.015)
CZ ratio of 4-year enroll to POP	-0.733*** (0.210)	-2.464*** (0.403)	-0.954*** (0.169)	-0.613*** (0.154)	-0.315 (0.172)	-0.360* (0.170)
Constant	1.656*** (0.466)	1.436 (1.061)	1.792*** (0.463)	1.941*** (0.389)	3.552*** (0.351)	3.221*** (0.457)
R-squared	0.149	0.134	0.103	0.088	0.112	0.135
Quantile value		2.567	3.752	4.585	5.313	6.182

Sample size is 6540 applicants (rounded per NCES requirements). Robust standard errors in parentheses. All data were weighted to account for the stratified sampling design used in ELS:02. Results for state and selected commuting zone covariates and age >18 are withheld for brevity. Reference category for Race is White. Reference category for Family Income is \$50,000–\$70,000. Reference category for Parent's Highest Education is High School

Table 5 (continued)

CZ commuting zone

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

taking into account their average distance applied. There was also some evidence that students from the highest-income families enrolled at distances that were farther from home than the average distances where they applied. Finally, the ratio of 4-year enrollments to population in the student's commuting zone was negative and significant, suggesting that a higher supply of spaces in proximity to students led them to be more likely to enroll closer to home conditional on how far from home they applied.

Summary and Discussion

Understanding how students make decisions about college remains an enduring topic of interest for higher education scholars, policymakers, and enrollment managers. Research has shown that these decisions are multifaceted, shaped in part by a combination of personal attributes of the student, the environment in which they were raised, their innate academic ability, where they live, and their capacity to pay the price of admission. In this study, we focused on one particular aspect of this decision: namely, how far from home students decide to go to college. This topic is important because unlike many other things that individuals purchase and consume, a 4-year college experience often requires students to move to where the college is located to take part in its services and thus distance is a factor in the decision. The decision to enroll close to or far from home has implications for the financial and non-financial costs of going to college, as well as the student's likelihood of success.

There are a few caveats to the study that warrant mention. First, due to data limitations we had to measure distances based on the ZIP codes for students and institutions. This approach will lead to some error—positive and negative—in estimating the actual distances based on a student's home and the institution, depending on where the student and institution are located within their respective ZIP codes. Second, the IES restricted-use data license prohibited us from also measuring distances based on commuting time and commuting miles. Future work can help test whether the results from our models are robust regarding how we measured distances. Finally, although the data from ELS were deemed to be the most appropriate for calculating distances, it would be important to determine if the findings in our study held for more recent cohorts of students. It is possible that the COVID-19 pandemic led to a reduction in the distances that students wanted to travel to college. In addition, the growing popularity of the common application form would reduce the search costs to students of applying to institutions, which may lead to increases in distances applied.

Our study adds to the literature on college choice in several ways. First, the study is unique by focusing on how students decide on the geographic span of their postsecondary markets at both the application and enrollment stages. Descriptive statistics revealed that students typically stay within 200 miles of their home when they go to college, and that the colleges to which they apply are on average farther from home than where they enroll. The difference could reflect the fact that price and ability to pay becomes more important when moving from the application to the enrollment stage of college choice.

Table 6 Quantile Regressions of Log of Mean Student Distance to College Enrolled

Variable	(1) OLS	(2) 10th	(3) 25th	(4) Median	(5) 75th	(6) 90th
Male	0.018 (0.064)	0.061 (0.075)	−0.044 (0.064)	−0.036 (0.041)	−0.049 (0.037)	0.068 (0.054)
Black	0.479*** (0.117)	−0.262* (0.126)	0.255 (0.234)	0.561*** (0.079)	0.578*** (0.062)	0.568*** (0.080)
Hispanic	−0.308* (0.146)	−0.730*** (0.121)	−0.389** (0.142)	−0.341* (0.146)	0.028 (0.136)	0.095 (0.085)
Asian	−0.469*** (0.105)	−0.559*** (0.092)	−0.587*** (0.106)	−0.492*** (0.106)	−0.235*** (0.068)	−0.292* (0.114)
Other race(s)	0.017 (0.167)	−0.190 (0.141)	−0.211 (0.265)	−0.097 (0.171)	0.046 (0.068)	0.638** (0.208)
High school GPA	0.074 (0.081)	0.404*** (0.096)	0.112 (0.089)	−0.038 (0.049)	−0.097* (0.046)	−0.149* (0.065)
SAT composite score (50s)	0.069*** (0.012)	0.036* (0.015)	0.059*** (0.012)	0.062*** (0.007)	0.076*** (0.007)	0.077*** (0.009)
Attend same high school 10th and 12th	−0.046 (0.158)	−0.132 (0.186)	0.035 (0.192)	0.057 (0.194)	−0.213** (0.077)	−0.312*** (0.078)
Family income: \$20K or less	−0.191 (0.132)	0.151 (0.178)	−0.021 (0.171)	−0.202 (0.115)	−0.277** (0.094)	−0.228* (0.105)
Family income: \$20K–\$50K	0.034 (0.088)	0.164 (0.119)	−0.071 (0.083)	0.024 (0.052)	0.086 (0.052)	0.187* (0.076)
Family income: \$75K–\$100K	0.242** (0.093)	0.358** (0.135)	0.197* (0.086)	0.209** (0.064)	0.158* (0.068)	0.221** (0.082)
Family income: \$100K or more	0.445*** (0.095)	0.553*** (0.134)	0.317** (0.110)	0.457*** (0.058)	0.352*** (0.048)	0.441*** (0.069)
Parent Ed: associates	−0.068 (0.108)	−0.080 (0.119)	−0.169 (0.094)	0.017 (0.064)	−0.020 (0.072)	−0.198 (0.104)
Parent Ed: bachelors	0.117 (0.101)	0.043 (0.106)	−0.011 (0.101)	0.181** (0.064)	0.189** (0.060)	0.080 (0.096)
Parent Ed: graduate	0.247* (0.115)	−0.043 (0.130)	0.143 (0.115)	0.378*** (0.076)	0.461*** (0.071)	0.265** (0.102)
Reside w/two heads of household	−0.039 (0.087)	0.037 (0.107)	−0.004 (0.085)	−0.035 (0.074)	−0.131* (0.051)	−0.146* (0.071)
CZ number of 4-year colleges (10s)	−0.037* (0.015)	0.021 (0.018)	−0.044** (0.014)	−0.069*** (0.012)	−0.024 (0.013)	0.012 (0.025)
CZ ratio of 4-year enroll to POP	−0.901** (0.331)	−3.102*** (0.519)	−1.814*** (0.405)	−0.790*** (0.204)	−0.282 (0.179)	−0.154 (0.271)
Constant	1.223 (0.718)	−1.538 (1.140)	1.288 (1.187)	2.524*** (0.502)	2.637*** (0.408)	3.554*** (0.739)
R-squared	0.122	0.170	0.091	0.079	0.083	0.118
Quantile value		1.805	3.038	4.318	5.195	6.094

Sample size is 5170 applicants (rounded per NCES requirements). Robust standard errors in parentheses. All data were weighted to account for the stratified sampling design used in ELS:02. Results for state and selected commuting zone covariates and age >18 are withheld for brevity. Reference category for Race is White. Reference category for Family Income is \$50,000–\$70,000. Reference category for Parent's Highest Education is High School

Table 6 (continued)

CZ commuting zone

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Second, we examined how the distances to college were shaped by a series of demand- and supply-side factors. The findings showed that student distances from home to college varied along with personal attributes such as race/ethnicity, their academic ability, and their family characteristics. Some of the results regarding student ability and parental socioeconomic status were to be expected. For example, the positive association between student SAT scores and average distance is consistent with the idea that higher-ability students cast a wider net when searching for colleges because they believe that they would be admissible to more institutions and/or eligible for more merit aid. Likewise, students whose parents had earned graduate degrees considered institutions that were, on average, about 30% farther from home. Other results, however, call for further investigation. In particular, we found large and statistically significant differences in average distances applied by race/ethnicity, and different patterns among racial/ethnic categories of students, with Black students applying to institutions that were 60% farther from home and Hispanic and Asian students opting to stay closer to home. These differences persisted even after taking into account academic performance, family characteristics, and locational attributes. Although the results from the statistical models cannot be interpreted as causal per se, the information about how student characteristics are associated with decisions about distance would certainly be useful to enrollment managers at institutions when assessing the likelihood of particular students applying to and attending their institutions.

We also expanded the notion of how to measure the supply of postsecondary options within proximity to a student by calculating the ratio of enrollments to college-age population and tested whether the supply of college options led to changes in the distance traveled to college. Our results showed that students who had more opportunities to enroll in a 4-year college close to home opted to apply and enroll in colleges that were closer to home. For example, for every increase of ten institutions within a student's commuting zone, there was a reduction of about 5% in distance applied. We further added to the literature by separating the variables for supply into 2-year and 4-year institutions and found that the negative association between geographic span and supply of institutions was mainly driven by the presence of 4-year institutions. Policymakers at the state level would find these results to be useful in assessing how postsecondary opportunities vary by region around their state and what it may mean for where residents go to college.

Finally, our use of quantile regression analysis is important because it demonstrates that the relationships between several variables in our study and distances applied and enrolled were not uniform across the distance distributions. Coming from a higher-income family, for example, had the largest impact on students who enrolled in colleges that were far from home, and in contrast the supply of postsecondary spaces close to home had a larger impact for those students who applied or enrolled closer to home. These and the other nonuniformities we observed are useful for researchers to better understand how these attributes inform decisions that students make about how far from home to go to college.

With regard to theory, our study highlights that our thinking about how students make decisions about college should be expanded to how they make decisions about their market or geographic scope. The results further showed that socioeconomic factors, measures of academic performance, and the availability of college opportunities

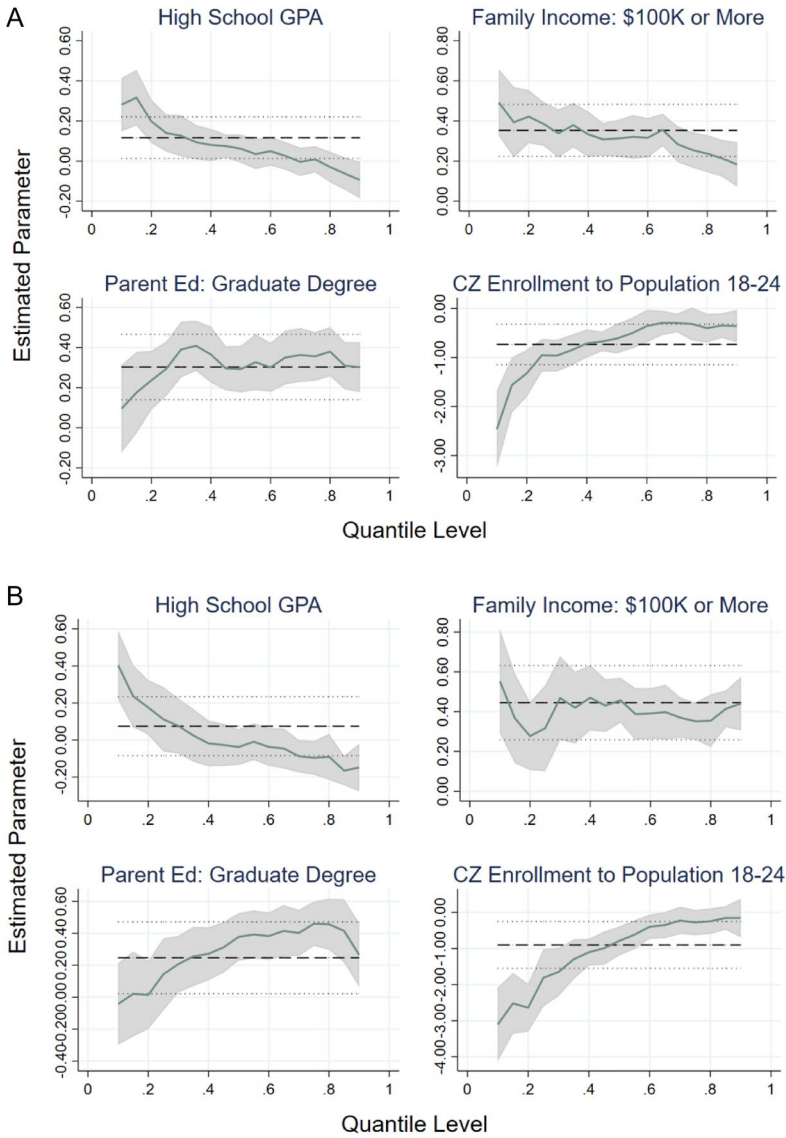


Fig. 2 Selected coefficients from quantile regression models. **A** Mean distance applied. Dashed lines represent OLS estimates. Solid lines represent quantile regression estimates. **B** Distance enrolled. Dashed lines represent OLS estimates. Solid lines represent quantile regression estimates

near home all play a role in how far from home students would like to go to college. The separate analyses of distances applied and enrolled can also be integrated into the student choice models at the application and enrollment stages as described by Toutkoushian and Paulsen (2016).

Table 7 Determinants of log of distance to colleges enrolled including distance of applications

Variable	<i>Y = log of distance enrolled</i>		
	Model 1	Model 2	Model 3
Ln(mean distance applied)	0.846*** (0.015)	–	–
Ln(median distance applied)	–	0.850*** (0.015)	–
Ln(maximum distance applied)	–	–	0.679*** (0.016)
Male	–0.036 (0.049)	–0.031 (0.048)	–0.022 (0.052)
Age > 18 years	0.034 (0.047)	0.028 (0.046)	0.059 (0.051)
Black	–0.024 (0.090)	0.011 (0.086)	–0.030 (0.097)
Hispanic	–0.207 (0.107)	–0.125 (0.094)	–0.270* (0.116)
Asian	–0.369*** (0.088)	–0.273** (0.084)	–0.463*** (0.093)
Other race(s)	–0.170 (0.127)	–0.120 (0.115)	–0.171 (0.140)
High school GPA	0.036 (0.067)	–0.019 (0.057)	0.038 (0.073)
SAT composite score (50 pts.)	–0.006 (0.009)	0.001 (0.008)	0.001 (0.010)
Attend same high school 10th and 12th	0.078 (0.119)	0.118 (0.123)	0.046 (0.127)
Family income: \$20K or less	–0.084 (0.102)	–0.061 (0.097)	–0.088 (0.111)
Family income: \$20K–\$50K	0.003 (0.065)	0.013 (0.062)	0.008 (0.070)
Family income: \$75K–\$100K	0.112 (0.068)	0.113 (0.068)	0.142 (0.073)
Family income: \$100K or more	0.146 (0.075)	0.168* (0.075)	0.182* (0.080)
Parent Ed: associates	–0.030 (0.076)	–0.024 (0.074)	–0.040 (0.084)
Parent Ed: bachelors	0.001 (0.072)	0.032 (0.069)	0.014 (0.078)
Parent Ed: graduate	–0.022 (0.084)	–0.014 (0.082)	0.006 (0.091)
Live w/two heads of household	0.114 (0.068)	0.092 (0.066)	0.105 (0.074)
CZ poverty rate (%)	–0.014 (0.014)	–0.015 (0.014)	–0.017 (0.015)

Table 7 (continued)

Variable	<i>Y = log of distance enrolled</i>		
	Model 1	Model 2	Model 3
CZ bachelor's or higher (%)	−0.017* (0.008)	−0.016 (0.008)	−0.025** (0.009)
CZ median income (\$1000s)	0.010 (0.008)	0.008 (0.008)	0.014 (0.009)
CZ unemployment rate (%)	0.035 (0.030)	0.029 (0.030)	0.039 (0.032)
CZ number of 4-year colleges (10s)	−0.013 (0.012)	0.005 (0.012)	−0.024 (0.013)
CZ ratio of 4-year enroll to pop	−0.380* (0.176)	−0.324 (0.166)	−0.524* (0.213)
Constant	−0.020 (0.566)	0.114 (0.559)	0.377 (0.601)
<i>R</i> -squared	0.517	0.543	0.437

Sample size is 5170 enrollees (rounded per NCES requirements). Robust standard errors in parentheses. All data were weighted to account for the stratified sampling design used in ELS:02. Results for state covariates are withheld for brevity. Reference category for Race is White. Reference category for Family Income is \$50,000–\$70,000. Reference category for Parent's Highest Education is High School

CZ commuting zone

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Taken together, the results from our study should be informative to higher education researchers and policymakers who seek to understand how students make critical decisions about pursuing a postsecondary education. State policymakers hoping to stem the brain drain from their locales, for example, can use the evidence here to advocate for expansions of postsecondary opportunities for students. Likewise, enrollment managers at 4-year institutions will find the results to be helpful when deciding on recruitment strategies for students by providing insights into the types of students who are more or less likely to be interested in traveling to their institution.

Acknowledgements We are grateful to the Spencer Foundation for financial support for this project (grant #201900227). An earlier version of this paper was presented at the annual meeting of the Southern Economic Association. We would like to thank participants at the Southern Economic Association conference and Steve DesJardins for comments on an earlier draft of this study.

Data availability The data used in this study can be obtained through a license with the National Center for Education Statistics.

Declarations

Conflict of interest I have no known conflict of interest to disclose.

References

- Alm, J., & Winters, J. (2009). Distance and intrastate college student migration. *Economics of Education Review*, 28, 728–738.
- Baryla, E., & Dotterweich, D. (2001). Student migration: Do significant factors vary by region? *Education Economics*, 9, 269–280.
- Becker, W., & Toutkoushian, R. (2013). On the meaning of markets in higher education. In M. Paulsen (Ed.), *Higher education: Handbook of theory and research* (Vol. XXVIII, pp. 323–376). Springer.
- Bourdieu, P. (1973). Cultural reproduction and social reproduction. In R. Brown (Ed.), *Knowledge, education, and cultural change*. Routledge.
- Bourdieu, P. (1977). *Outline of a theory of practice*. Cambridge University Press.
- Bourdieu, P. (1990). *The logic of practice*. Stanford University Press.
- Byun, S., Irvin, M., & Meece, J. (2015). Rural–nonrural differences in college attendance patterns. *Peabody Journal of Education*, 90, 263–279.
- Coleman, J. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, 95–120.
- Cooke, T., & Boyle, P. (2011). The migration of high school graduates to college. *Educational Evaluation and Policy Analysis*, 33, 202–213.
- Dache-Gerbino, A. (2018). College desert and oasis: A critical geographic analysis of local college access. *Journal of Diversity in Higher Education*, 11, 97–116.
- De Graaf, N., De Graaf, P., & Kraaykamp, G. (2000). Parental cultural capital and educational attainment in the Netherlands: A refinement of the cultural capital perspective. *Sociology of Education*, 73, 92–111.
- Denzler, S., & Wolter, S. (2011). *Too far to go? Does distance determine study choices?* IZA Discussion Paper No. 571. The Institute for the Study of Labor.
- DesJardins, S., Ahlburg, D., & McCall, B. (2006). An integrated model of application, admission, enrollment, and financial aid. *Journal of Higher Education*, 77, 381–429.
- Dill, D. (1997). Higher education markets and public policy. *Higher Education Policy*, 10, 167–185.
- Dillon, E., & Smith, J. (2017). Determinants of the match between student ability and college quality. *Journal of Labor Economics*, 35, 45–66.
- Do, C. (2004). The effects of local colleges on the quality of college attended. *Economics of Education Review*, 23, 249–257.
- Eagan, K., Stolzenberg, E., Ramirez, J., Aragon, M., Suchard, M., & Rios-Aguilar, C. (2016). *The American freshman: Fifty-year trends 1966–2015*. Higher Education Research Institute.
- Evans, M., Kelley, J., Sikora, J., & Treiman, D. J. (2010). Family scholarly culture and educational success: Books and schooling in 27 nations. *Research in Social Stratification and Mobility*, 28, 171–197.
- Frenette, M. (2004). Access to college and university: Does distance to school matter? *Canadian Public Policy*, 30, 427–443.
- Frenette, M. (2006). Too far to go on? Distance to school and university participation. *Education Economics*, 14, 31–58.
- Garza, A., & Fullerton, A. (2018). Staying close or going away: How distance to college impacts the educational attainment and academic performance of first-generation college students. *Sociological Perspectives*, 61, 164–185.
- Ghatak, S., Levine, P., & Price, S. (1996). Migration theories and evidence: An assessment. *Journal of Economic Surveys*, 10, 159–198.
- Gibbons, S., & Vignoles, A. (2012). Geography, choice and participation in higher education in England. *Regional Science and Urban Economics*, 42, 98–113.
- González Canché, M. (2018). Nearby college enrollment and geographical skills mismatch: (Re)conceptualizing student out-migration in the American higher education system. *The Journal of Higher Education*, 89, 892–934.
- Greenwood, M. (1975). Research on internal migration in the United States: A survey. *Journal of Economic Literature*, 13, 397–433.
- Griffith, A., & Rothstein, D. (2009). Can't get there from here: The decision to apply to a selective college. *Economics of Education Review*, 28, 620–628.
- Guillory, J. (2023). *Cultural capital: The problem of literary canon formation*. University of Chicago Press.
- Han, P. (2014). A literature review on college choice and marketing strategies for recruitment. *Family and Consumer Sciences Research Journal*, 43, 120–130.
- Hill, L., Bregman, A., & Andrade, F. (2015). Social capital for college: Network composition and access to selective institutions among urban high school students. *Urban Education*, 50, 316–345.
- Hillman, N. (2016). Geography of college opportunity: The case of education deserts. *American Educational Research Journal*, 53, 987–1021.

- Hirschl, N., & Smith, C. (2020). Well-placed: The geography of opportunity and high school effects on college attendance. *Research in Higher Education*, *61*, 567–587.
- Holloman, D., & Nolen, A. (2008). Recruiting place bound students: The influence of location on the college choice of high-school seniors. *Metropolitan Universities*, *19*, 41–53.
- Hossler, D., & Gallagher, K. (1987). Studying student college choice: A three-phase model and the implications for policymakers. *College and University*, *62*, 207–221.
- Hossler, D., Schmit, J., & Vesper, N. (1999). *Going to college: How social, economic, and educational factors influence the decisions students make*. The Johns Hopkins University Press.
- Ishitani, T. (2011). The determinants of out-migration among in-state college students in the United States. *Research in Higher Education*, *52*, 107–122.
- Jaeger, M., & Breen, R. (2016). A dynamic model of cultural reproduction. *American Journal of Sociology*, *121*, 1079–1115.
- Kim, D. (2004). The effect of financial aid on students' college choice: Differences by racial groups. *Research in Higher Education*, *45*, 43–70.
- Klasik, D., Blagg, K., & Pekor, Z. (2018). Out of the education desert: How limited local college options are associated with inequity in postsecondary opportunities. *Social Sciences*, *7*, 165–190.
- Koenker, R., & Hallock, K. (2001). Quantile regression. *Journal of Economic Perspectives*, *15*, 143–156.
- Leppel, K. (1993). Logit estimation of a gravity model of the college enrollment decision. *Research in Higher Education*, *34*, 387–398.
- Lillis, M., & Tian, R. (2008). The impact of cost on college choice: Beyond the means of the economically disadvantaged. *Journal of College Admission*, *200*, 4–14.
- Marginson, S. (2019). Limitations of human capital theory. *Studies in Higher Education*, *44*, 287–301.
- Mattern, K., & Wyatt, J. (2009). Student choice of college: How far do students go for an education? *Journal of College Admission*, *203*, 18–29.
- McHugh, R., & Morgan, J. (1984). The determinants of interstate student migration: A place-to-place analysis. *Economics of Education Review*, *3*, 269–278.
- Mishra, S. (2020). Social networks, social capital, social support and academic success in higher education: A systematic review with a special focus on 'underrepresented' students. *Educational Research Review*, *29*, 1–24.
- Mixon, F. (1992). Factors affecting college student migration across states. *International Journal of Manpower*, *13*, 25–32.
- Mixon, F., & Hsing, Y. (1994). The determinants of out-of-state enrollment in higher education: A tobit analysis. *Economics of Education Review*, *13*, 329–335.
- Mortenson, T. (1996). Freshmen enrolling in college farther from home, but who can afford to go so far away? *Postsecondary Education Opportunity*, *50*, 8–13.
- Mortenson, T. (1998). Interstate migration of college undergraduates. *Postsecondary Education Opportunity*, *70*, 1–12.
- Morton, T., Ramirez, N., Meece, J., Demetriou, C., & Panter, A. (2018). Perceived barriers, anxieties, and fears in prospective college students from rural high schools. *High School Journal*, *101*, 155–176.
- Niu, S. (2015). Leaving home state for college: Differences by race/ethnicity and parental education. *Research in Higher Education*, *56*, 325–359.
- Perna, L. (2006). Studying college access and choice: A proposed conceptual model. In J. Smart (Ed.), *Higher education: Handbook of theory and research*. (Vol. XXI). Kluwer Academic Publishers.
- Pryor, J., Hurtado, S., Saenz, V., Lindholm, J., Korn, W., & Mahoney, K. (2006). *The American freshman: National norms for Fall 2005*. Higher Education Research Institute. Retrieved July 9, 2013 from <http://heri.ucla.edu/PDFs/ResearchBrief05.PDF>
- Sandefur, G., Meier, A., & Campbell, M. (2006). Family resources, social capital, and college attendance. *Social Science Research*, *35*, 525–553.
- Spieß, C., & Wrohlich, K. (2010). Does distance determine who attends a university in Germany? *Economics of Education Review*, *29*, 470–479.
- Stephenson, A., Heckert, A., & Yerger, D. (2016). College choice and the university brand: Exploring the consumer decision framework. *Higher Education*, *71*, 489–503.
- Tinto, V. (1973). College proximity and rates of college attendance. *American Educational Research Journal*, *10*, 277–293.
- Toutkoushian, R. (2001a). Do parental income and educational attainment affect the initial choices of New Hampshire's college-bound students? *Economics of Education Review*, *20*, 245–262.
- Toutkoushian, R. (2001b). Changes in the demand for public land-grant universities: The use of panel datasets. *Research in Higher Education*, *42*, 679–707.
- Toutkoushian, R., & Hillman, N. (2012). The impact of state appropriations and grants on access to higher education and state outmigration. *The Review of Higher Education*, *36*, 51–90.

- Toutkoushian, R., & Paulsen, M. (2016). *Economics of higher education: Background, concepts, and applications*. Springer.
- Tramonte, L., & Willms, J. D. (2010). Cultural capital and its effects on education outcomes. *Economics of Education Review*, 29, 200–213.
- Tuckman, H. (1970). Determinants of college student migration. *Southern Economic Journal*, 37, 184–189.
- Turley, R. (2009). College proximity: Mapping access to opportunity. *Sociology of Education*, 82, 126–146.
- Turner, S. (2017). *Education markets: Forward-looking policy options*. Hutchins Center Working Paper #27. Brookings.
- Wells, R., Manly, C., Kommers, S., & Kimball, E. (2019). Narrowed gaps and persistent challenges: Examining rural–nonrural disparities in postsecondary outcomes over time. *American Journal of Education*, 126, 1–31.
- Zhang, L., & Ness, E. (2010). Does state merit-based aid STEM brain drain? *Educational Evaluation and Policy Analysis*, 32, 143–165.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.