Estimating the Causal Effects of Dormitory Residence During College on Student Outcomes

C. Lockwood Reynolds
Department of Economics
Kent State University
P.O. Box 5190
Kent, OH 44242
Phone: (330) 672-1089
Email: creynol9@kent.edu

August 31, 2012

Abstract

Dormitory residence among first-year college students is an important source of heterogeneity in higher education in the United States. The small previous literature has found positive effects of dormitory residence on student outcomes such as grade point average and retention using regression methods controlling for selection on observable characteristics. However, selection on unobservable characteristics is likely to bias such estimates. This paper utilizes variation in the likelihood of being in a dormitory at a large, public university created by the university's rules about on-campus residency. In particular, the university uses a distance rule such that only students within a specific distance of the campus are allowed to opt out of dormitory residence during the first year. Using this source of variation in an instrumental variables methodology, this paper finds no evidence that dormitory residence during the first year has an effect on student retention and has only a modest effect on student grade point average. Further analysis suggests that any positive effects on grade point average are concentrated among the lowest ability students or students who live with higher ability roommates in the dormitory. The latter results are suggestive that any positive effects of dormitories operate through peer effects.

KEYWORDS: higher education, retention, dormitories
JEL CLASSIFICATION: I21, I23, J24

Acknowledgements: I would like to thank Jeffrey Smith, Eric Johnson, Shawn Rohlin, Nadia Greenhalgh-Stanley, and participants at the Southern Economics Association for their helpful comments and suggestions. All errors and omissions are my own.

1. Introduction

A prominent, but understudied, source of heterogeneity in the United States higher education system is the use of dormitories by students. Undergraduate degree-granting institutions in the United States had a total capacity of over 2.7 million beds in 2007. However, this is only enough capacity for approximately 20 percent of full-time equivalent undergraduate enrollment and capacity varies by sector. Private not-for-profit four-year institutions have total dormitory capacity for 42 percent of full-time equivalent undergraduates compared to 24 percent for public four-year institutions and only 2 percent for two-year institutions.[1] This variation in supply corresponds to differences in policies regarding on-campus residency, which is required among first-year students at many private institutions and some public institutions, as well as differences in institutional focus, with many public four-year and nearly all two-year institutions providing education to commuter students.

The primary question of this paper is what, if any, effect does dormitory residence during the first year of college have on student educational outcomes compared to living at home? Given the variation in access to this particular higher education input, it would seem important to understand exactly what role dormitories have on the production of human capital. Additionally, the role of dormitories in higher education may be changing. While some schools are considering expanding the use of dormitories, the higher education market is also moving towards increasing convenient access for students through the use of online and distance-learning classes that reduce the need for students to be on campus. Furthermore, while the public and the popular press appear to be increasingly concerned about the rising cost of college as measured by tuition and required fees, dormitories are expensive for both institutions and students. The average room and board in 2007 at four-year institutions was $7,903 compared to average tuition

[1] Author's calculations from the Integrated Postsecondary Education Data System (IPEDS).

and required fees of $11,459, equivalent to a 69 percent increase in the average cost of college to students who live on-campus.[2] Under the assumption that dormitory room and board rates are set to break-even, the total cost to institutions to provide on-campus residency can be approximated by simply multiplying the dormitory capacity by room and board across institutions. The results suggest a total expense of $20.4 billion in 2007, of which approximately half is in the public sector.[3]

Despite the potential importance of on-campus residency in higher education, there is a lack of empirical research in economics on the potential effects of dormitory residence on student outcomes. The most closely related research is the literature on the influence of peer ability on academic success during college (e.g. Sacerdote (2001), Zimmerman (2003), Foster (2006) and Stinebrickner and Stinebrickner (2006)). This literature generally finds small but positive effects of having a high ability roommate during college on academic performance, among students living in dormitories. These papers suggest a possible mechanism through which dormitories could impact students but because all of the students in these studies live in dormitories, as is common for the schools in question, it is not clear whether the results of such studies are generalizable to the larger question of the overall effect of living on-campus during college.[4]

---

[2] Obviously, these calculations ignore the fact that students still must have food and shelter living off-campus or at home but it is difficult to calculate appropriate measures of these alternative costs. At the very least the room and board charges represent a large increase in the sticker price of college attendance and likely are an increase in the pecuniary costs associated with attending college.

[3] Discussions with residence services suggests that room and board is priced to break-even and not to generate revenue for institutions. Whether or not this is true is difficult to know. However, even if the goal is to break-even on the provision of room and board, institutions may have financial incentives to encourage students to stay on campus in the form of increased purchases at university bookstores, on-campus independent restaurants, or sporting events.

[4] Dartmouth, Williams and Berea require all students to live on campus during at least the first year of college and roughly 90 percent of first-year students at the University of Maryland live in the dormitories. Stinebrickner and Stinebrickner (2006) also question whether peer effects should exist at these private or flagship institutions. The large, non-flagship public university considered in this study is arguably an environment where peer effects are more likely to be found because there may be greater variation in student type.

Students in dormitories often have better measured outcomes compared to students at the same university who live at home, and these differences in unconditional means have been used by institutions to justify dormitory expansions.[5] Dormitory residence has also been encouraged because of a large education literature on the theorized important link between a student's involvement in their education (Astin, 1984), and the integration of students into the academic or social systems of colleges (Tinto, 1975), or the strong correlations between student engagement and academic success (for example, Kuh et al., 2008). Students living off-campus are assumed, almost by definition, to be less engaged in their education and institution and therefore are less likely to be successful. This has led researchers to promote on-campus residency as critical to academic success. For example, Braxton and Hirschy (2005) argue that policies that allow students to live off-campus should be eliminated, if possible, because they hinder the social integration of the student.

While it is possible that students choosing to live off-campus are less likely to be successful, it is not clear that on-campus residency requirements would operationalize student engagement and therefore improve academic outcomes. Additionally, the decision to be in the dormitories in the first place may be based on characteristics of students, observed or unobserved by the researcher, that are also correlated with future higher education outcomes. Some education research using various regression methods controlling for observable characteristics of students have found that dormitory residence is associated with improved academic performance in the classroom and increased retention (e.g. Astin (1973), Chickering (1974), and Astin and Oseguera (2005)). However, these papers fail to account for selection on unobservable

[5] For example, at a 2008 Board of Trustees meeting at the Ohio State University, a proposal to require all second-year student to live on-campus was justified because the observed retention rates, as well as grade point averages, were higher among students in the residence halls than students living off campus.

characteristics and often include comparisons of students across different types of higher education institutions with different institutional focus and serving different student populations.

This paper attempts to identify the causal impact of dormitory residence on student outcomes in higher education by utilizing a unique dataset from a large, public four-year university in the Midwest. The initial data collected is the universe of students that matriculated to the university in the fall from 1999 to 2007. Application and financial aid data provide a host of background characteristics of the students and information on the high school the student attended was gathered by linking the student's high school to government databases. Additionally, the university provided measures of student academic performance in the form of grade point average, measured in the fall and overall during the first year, as well as retention measured in the second semester and second year.

Most importantly, students at this university are required to live on campus during the first year but are granted a commuter exemption if the student lives at home with his or her parents, provided that this residence is within a specified distance of the university. I use information on the location of the student's permanent residence to calculate whether the student is eligible to opt-out of the dormitories during their first year and then use this information, as well as information on the length of time of a potential commute, as instrumental variables to identify the causal effect of dormitory residence during the first-year on student higher education outcomes. I demonstrate that the instruments are strongly correlated with the likelihood that a student lives in the dormitory during their first year and an overidentification test fails to reject the null hypothesis that the variables represent valid instruments. I also demonstrate that my estimated results are robust to a variety of sample restrictions and choice of instrument sets.

I find that standard regression methods controlling for observable characteristics suggest that dormitory residence during the first year of college is associated with increases in retention to both the second semester and to the second year of college. Additionally, dormitory residence is associated with higher grade point averages during both the fall semester and the first year. However, the local average treatment effects from the instrumental variables approach suggest that students induced to be in the dormitories experience more modest effects. In particular, dormitory residence has no impact on retention measured in either period and only a small positive effect on student GPA. Furthermore, I find that the effect of dormitory residence on GPA decreases with ability with results concentrated among students from the bottom tercile of the ACT score distribution. However, dormitory residence has no effect on retention within any ability tercile.

Finally, I present evidence that any potential positive effects from dormitories may be due to the positive effects of peer ability as only those students induced to live in the dormitories and have higher ability roommates experience positive effects on GPA. These results both provide supportive evidence confirming prior estimates of peer effects and also expand our understanding of such effects by demonstrating that they exist in a different sector of the higher education sector than typically previously considered as well as among student induced to live in the dormitories. Overall, these results suggest that dormitory residence has little impact on institutional retention rates and only has positive effects on academic performance among lower ability students.

The remainder of the paper is organized as follows. Section 2 describes the data and methodological approaches used in the paper. Section 3 presents the primary results and robustness checks. Section 4 provides a discussion and concludes the paper.

2. Data and Methodology

2.1 Description of the Data

The analysis in this paper uses a unique dataset collected from a large, public four-year university in the Midwest.[6] The university is not the state flagship but is classified as a research-oriented university and offers degrees at both the undergraduate and graduate level (including PhDs). The school is considered to be at or above average within the state as well as nationally among public four-year institutions. For example, among public four-year institutions in the United States, the university is in the middle quintile for standardized test scores of entering students, the second-highest quintile for undergraduate enrollment and graduation rate, and just inside the top quintile for in-state tuition and required fees. The percent of students receiving any form of financial aid is in the second-highest quintile and the average amount of federal aid received is in the middle quintile.[7]

The data obtained from the university contains the universe of students matriculating in the fall semester for each year from 1999 to 2007, equaling approximately 33,000 students. For this analysis, I restrict the sample to those students who are in-state, and therefore are subject to the same tuition and required fees. I also focus on traditional college students by further restricting the sample to full-time students who are no older than 19 when they matriculate. These restrictions remove approximately 18 percent of the original sample. Additionally, because I am interested in studying dormitory residence versus living at home and commuting, as well as for purposes of the identification strategy described below, I restrict the geographic

[6] The use of the data is predicated on not revealing the name or specific location of the institution. Thus, I will refer the institution as "the university" throughout the paper.
[7] Author's calculations from the Integrated Postsecondary Education Data System (IPEDS).

area to those students whose parents live within 80 miles of the campus "as the crow flies."[8] This geographic area includes both urban and rural locations and many students at the university live in this region. Consequently, this sample restriction only removes an additional 12 percent of the original sample. In the estimation sample, 77 percent of students live on campus during the first year.

University application data provides basic demographic data on age, race and gender as well as extensive ability measures in the form of high school grade point average (GPA) and ACT and SAT scores. The majority of students at this university submit either ACT scores or submit both ACT and SAT scores. I impute ACT scores for those only submitting SAT scores based on an equipercentile method calculated from the subset of students supplying both test scores. Given the importance of the ability and demographic variables in the analysis, I restrict the sample to those students with valid information for all variables thereby removing an additional 5.6 percent of the students. The application data also contains information on the high school attended by the student which in most cases could be matched to data from the Common Core of Data (CCD) for public schools and the Private School Universe (PSU) for private schools to generate high school controls. From these datasets I constructed an indicator for whether the high school attended was public and I also calculated the average student to teacher ratio at the high school during the sample period. The 1.9 percent of cases for which high school data could not be constructed were removed from the sample.

The variables from the application data provide most of the covariates commonly used in higher education research with the notable exception of family background characteristics. Unfortunately, such information is not available from the application data. However, measures

[8] "As the crow flies" distances are the shortest distance between two points on the globe as measured by the great circle distance which accounts for the spherical nature of the planet.

of family background could be constructed from financial aid data for all students that submitted the Free Application for Federal Student Aid (FAFSA) form, which is 71 percent of students at this university. Given the potential importance of family background, in my primary analysis I restrict the estimation sample to those students submitting the FAFSA. Thus, my estimates should be interpreted as the effects of dormitory residence among the sample of students with financial need. Arguably this is a subsample of interest in its own right as such students may be more marginal in terms of attendance and likelihood of success, and thus could be impacted positively by dormitory residence. However, I will demonstrate that the main conclusions from the sample of financial aid students hold if I apply my methodology, minus the family background characteristics, to the full sample of students. Thus, the focus on financial aid students does not drive the results.

The family background information that I collect is real ($2008) family income and student dependency status. I also collect information on the marital status of parents as well as parental education. For parental education, I create an indicator for whether at least one of the parent's has completed at least a bachelor's degree. In addition to the student-level covariates on family background, I use information from the application data of the ZIP code of the student's residence to construct a variety of characteristics of the student's area of residence. From the 2000 Census, I calculated ZIP code measures of the distributions of educational attainment and race, the unemployment rate and poverty rate, per capital income and urbanicity. Combined, the FAFSA and ZIP code data provide a rich set of controls for the environment in which the student was raised.

The university also provided a set of outcomes to measure the effect of dormitory residence on student success. In particular, I am interested in studying what effect dormitory residence has on student retention due to either improved academic performance or to students being more engaged or integrated into their community, as hypothesized in the prior education literature. I investigate to what extent dormitory residence impacts student retention measured at two points in time: retention to the second (spring) semester and retention to the second year.[9] A non-retained student does not earn a degree from the university, which may be the primary concern of school administrators, but it is possible that the student transfers to another institution and completes a degree. However, students who leave their first institution often have a lower likelihood of completing a BA.[10] Since academic success is one channel through which retention can be affected, I also directly estimate the effect of dormitory residence on the grade point average of the student at both the end of the first semester as well as at the end of the first year. In addition to investigating the mean impact on grade point average, I also investigate how dormitory residence affects being in different GPA ranges. These outcomes provide a measure of student performance in the classroom.

Table 1 presents summary statistics for these outcomes for the entire estimation sample, and then separately by on-campus residency status during the first year. The last column presents the difference for each variable between the average for those students in the dormitories and those residing at home. On average, 88.8 percent of students return for the spring semester and 79.0 percent return for a second year. However, those students in the dormitories are much more likely than the students living at home to be retained to both periods.

---

[9] I use the university's definition of retention which is measured on the 15th day of the semester. This measure is designed to avoid incidental retention among dropouts caused by student pre-registration for classes in the prior semester.

[10] For example, reports of the 2004 cohort at the university in this paper suggest that of those students that fail to graduate from the university within 6 years, only 11.8% have completed a BA at another institution. Similarly, in the National Longitudinal Survey of Youth 1997 (NLSY97), approximately 68 percent of students who begin college at a four-year institution and earn at least 25 credits, approximately one year of courses, at this first institution go on to complete a bachelor's degree. In comparison, only 17 percent of those students that complete less than 25 credits at their first institution complete a BA (and only 25 percent complete a BA among students who earn less than 25 credits at their first institution but do earn additional credit later at another institution).

Approximately 90 percent of students in the dormitories return for the spring semester while only 83 percent of those students living at home return. Similarly, 80.4 percent of dormitory students are retained until the second year while only 74.2 percent likely to be retained. Similarly, the students in dormitories have higher GPAs during the fall semester and the overall during the first-year compared to students not in the dormitories. Furthermore, students in dormitories are 7.4 percent less likely to be failing (having a GPA<1.0) at the end of the first year. The primary question of this paper is how much of these positive differences can actually be ascribed to student residency in the dormitories instead of being the result of positive selection of students into dormitories based on observed or unobserved characteristics.

As a first pass at answering this question, the bottom panel of Table 1 presents summary statistics for selected observable characteristics. For the most part the differences in the mean values are small and economically insignificant, although the differences are typically statistically significant because of the large sample size. For example, those students in the dormitories during their first year have high school GPAs that are 0.03 points lower on a standard 4.0 point scale. Similarly, ACT scores are lower but by a third of a point or less on a 36 point scale. However, some substantial differences do exist. For example, dormitory students in this university are less likely to be male, less likely to be white and more likely to be black. Consistent with the concerns about possible selection into dormitories, students in dormitories are approximately 7 percent more likely to have parents with at least a bachelor's degree and have family incomes that are almost $8,000 higher, approximately 10 percent higher than those students choosing to live at home. Additionally, students in the dormitories are 5.4 percent less likely to have attended a public high school. The results suggest that families with higher resources are more likely to be in the dormitories during the first year of college.

To further explore the role of observable characteristics on the likelihood of students living in the dorms, instead of living at home, during their first year, I estimate a probit model of the on-campus indicator on the set of observable characteristics. The first column of Table 2 presents the marginal effects, estimated at the mean of all variables, of selected covariates on the likelihood that a student lives in the dorm during the first year of college. Conditional on the other covariates, men are 5.7 percentage points less likely to be in the dormitories during the first year while black and Hispanic students are more likely than white students to be in the dormitories. As was observed in the simple sample means, there is some evidence that students from better socio-economic backgrounds are more likely to be in the dormitory during their first year. For example, having a parent with at least a bachelor's degree increases the likelihood of being in the dormitories by 5.0 percentage points while a $10,000 increase in real family income increases the likelihood by 0.6 percentage points. Attending a public high school reduces the likelihood of being on-campus by 9.3 percentage points, possibly further evidence of differences in household resources. Conditional on these variables, student ability has little impact on dormitory residence with small marginal effects for both ACT scores and high school GPA.[11] Overall, there is some evidence that students select into dormitories based on socio-economic status.

2.2 Empirical Methodology

The basic question investigated in this paper is what is the causal effect of dormitory residence on student achievement and retention. As discussed, dormitory residence could affect achievement and retention but the results from standard regression methodologies controlling for observable characteristics of the student are confounded by the decisions made by students about

---

[11] As will be discussed, some students are not given the choice of being in the dormitories. However, the marginal effects on all variables are similar in magnitude and statistical significance if the sample is limited to those students who are allowed to choose whether they are in the dormitories during their first year.

dormitory residence based on unobservable characteristics. In particular, suppose that some student outcome Y is generated as

$$Y = f(D, X, U) \quad (1)$$

where D is an indicator for whether the student is in the dormitories during the first year, X is the set of characteristics observable to the researcher, and U is a set of unobserved covariates. The standard regression approach would be to estimate

$$Y = \alpha + \delta D + \beta X + \epsilon \, . \quad (2)$$

However, equation (2) only estimates the causal effect of dormitory residence on outcome Y if $E(\epsilon|D, X) = 0$, meaning that conditional on the observable characteristics there is no residual bias associated with unobserved characteristics. In the context of dormitory residence decisions this condition is unlikely to hold even with a rich set of observed covariates. For example, unobserved student motivation or interest in higher education, or even the student's inherent interest in being an "engaged" student, are likely to affect the dormitory decision as well as the measured educational outcomes.

The methodological approach in this paper attempts to identify the causal impact of dormitory residence on student success by exploiting variation in the probability of being in a dormitory during the first year induced by the university's rules governing on-campus residency. At this university, dormitory residency is a requirement for students during their first year unless they are given a commuter exemption. The commuter exemption rule states that a student can opt out of university housing if they will live at the home of their parent or guardian, provided that this residence is within the commuter exemption area which is defined as being within a 50 mile circle of the campus. In practice, students could input their address into an online mapping program to check their status but the university also provides a list of eligible ZIP codes with the application form. Thus, the commuter exemption rules change the likelihood of a student being on-campus based solely on the ZIP code in which the student resides. In addition to the commuter exemption rule, I also exploit variation in the costs of a potential commute by gathering data from online mapping sites about the expected time of a commute from each ZIP code to the campus. Students within the commuter exemption area may be more likely to opt out of university housing during their first year if the time costs of commuting are low regardless of how far away they live from campus "as the crow flies."

Importantly, these variables affect the likelihood that students are in the dormitories during their first year. Column (ii) of Table 2 presents the marginal effects from a probit of dormitory residence during the first year on student characteristics as in column (ii) but also including an indicator for whether the student's permanent residence is within the commuter exemption area, the potential commuting time and the interaction of the two commuting variables.[12] Note that these estimates also include the extensive set of ZIP code level covariates previously discussed. The marginal effects of student characteristics are similar to those presented in column (i) without the commuter variables. However, residing in the commuter exemption area reduces the likelihood of on-campus residency during the first-year by 20.6 percentage points while each additional 5 minutes of commuting time increases the likelihood of being in the dormitories by 1 percentage point. Furthermore, the interaction term is positive suggesting that the likelihood of being in the dormitories during the first year decreases with commuting time within the commuter exemption area.

[12] A linear probability model specification of this model corresponds to the first-stage regression in the instrumental variable approach described below. I report the marginal effects from a probit in Table 2 because of the ease of interpreting the coefficients, however the results of the linear probability model are similar and available upon request.

I use the commuter exemption rule, commuting time and the interaction of the two variables as the basis of an instrumental variable strategy.[13] I utilize these instruments in a two-stage residual inclusion (2SRI) framework (see Terza, Basu and Rathouz (2008) for a discussion as well as Hausman (1978) and Wooldridge (2002, chapter 12)) where I estimate the first-stage regression

$$D_i = \varphi + \gamma X_i + \theta Z_i + \mu_i \qquad (3)$$

where X_i is the set of covariates of the individual, their family and their locality described previously, and Z_i is the set of instruments. I also include in X cohort fixed effects to account for differences over time in macroeconomic and higher education trends. From equation (3), I calculate the residual $\widehat{u_i} = D_i - \widehat{D_i}$ and then include this first-stage residual in the second stage regression

$$Y_i = \alpha + \delta D_i + \beta X_i + \gamma \widehat{u_i} + \epsilon_i \ . \qquad (4)$$

For linear models, the estimated effect of dormitories on the outcome Y_i, $\hat{\delta}$, is identical between the 2SRI method and standard two-stage least squares (2SLS) estimation. However, in the case of nonlinear models, such as the probit and ordered probit models employed in this paper, 2SRI has better consistency in small samples compared to equivalent two-stage predictor substitution, the non-linear version of 2SLS (Terza, Basu, and Rathouz, 2008).[14]

Instrumental variable methods estimate local average treatment effects (LATE) (Angrist and Imbens, 1994) which is the impact of the treatment for those observations that are induced to be in the treatment group because of the instruments. In the case of this paper, the IV estimates

---

[13] An alternative specification would only include the commuting time for those students who are eligible to commute, meaning that they are inside the commuter exemption area. All estimates are robust to this specification or simply removing the interaction term.

[14] Two-stage least squares can be considered a special case of two-stage predictor substitution (2SPS) for linear models. In both methods, the researcher generates predicted values of $\widehat{D}$ from the first-stage equation (equation (3)) and substitutes those values in for D in the second-stage.

can be interpreted as the effect of dormitory residence on those students induced to be in the dormitories because of the commuter exemption rule and costs of commuting. This interpretation, as well as the robustness of the estimates to various sample restrictions and specifications of the instrument set, are discussed in Section 3.2 after the primary results in Section 3.1.

Estimates from IV methods are potentially subject to substantial bias in small samples if the instruments are only weakly correlated with the endogenous variable (Bound, Jaeger, and Regina, 1995), in this case dormitory residence in the first year. A common way to test the strength of the instruments is an F-test on the joint significance of the instrumental variables in the first stage. Using this method, the commuter exemption and driving time instruments are strongly correlated with dormitory residence with an F-statistic of 161.83 (p-value = 0.000). Additionally, because there are more instruments than endogenous variables, I can conduct an overidentification test to test the validity of the excluded instruments. The overidentification test fails to reject the null hypothesis that the commuter exemption rule and commuting costs, as well as their interaction, are valid instruments (p-value = 0.583). These results suggest that the instrumental variable strategy may be successful at identifying the causal effect of dormitory residence on student success.

3. Results

3.1 The Effect of Dormitory Residence on Retention and Academic Success

I begin by investigating the impact of dormitories on student retention in Table 3. Column (i) of Table 3 presents the marginal effects calculated at the mean of all variables of dormitory residence on retention to the second semester, estimated by a probit incorporating the observable characteristics but without utilizing the instrumental variables. The results suggest that, conditional on observable characteristics, residing the dormitories during the first year instead of living at home is associated with a 5.8 percentage point increase in the likelihood of returning for a second semester. The results in column (ii) suggest that on-campus residency is associated with a 6.8 percentage point increase in the likelihood of being retained to the second year. Additionally, the marginal effects are large relative to the effects of other observable characteristics. For example, the dormitory effect is similar in magnitude to a 0.5 increase in high school GPA and a substantial increase in family income. These results are similar in magnitude to those previously estimated (e.g. Astin (1973), Chickering (1974), and Astin and Oseguera (2005)).

As discussed in the previous section, these positive effects could be confounded by unobservable characteristics of the students. Columns (iii) and (iv) present the LATE estimates using the commuter exemption area and commuting times in the 2SRI methodology described previously. The IV method does not impact the marginal effects of the selected student characteristics but the positive effect of dormitories on both measures of student retention disappears. For retention measured in both the second semester and second year, the marginal effects are small and not statistically significant at conventional levels. The results suggest that, among students induced to live in the dormitories because of the commuter exemption rule and commuting times, dormitory residence has no impact on student success as measured by retention. This result stands in contrast to previous estimates and potentially suggests that there is little gain to students from on-campus residency requirements, possibly because student engagement is not strongly associated with student success or because forcing students to live on campus does not in and of itself operationalize student engagement or campus integration.

In Table 4, I present the estimates of dormitory residence on student academic performance as measured by grade point average during the first semester and overall grade point average during the first year. The results in column (i) suggest that, conditional on observable characteristics, dormitory residence increases GPA during the first semester by 0.187 points and increases first-year GPA by 0.225 points, all measured on a four point scale. These estimates are substantial compared to the effects of family income and parental education but are somewhat modest when compared to the average GPA. For example, the estimates suggest that dormitory residence is associated with a 7.3 percent ($= \frac{0.187}{2.561} \times 100$) increase in fall semester GPA and a 9.1 percent ($= \frac{0.225}{2.476} \times 100$) increase in first-year GPA. Interestingly, the coefficient on high school GPA suggests that there is nearly a one-for-one relationship between high school GPA and college GPA during the first year, while ACT scores have a comparatively small effect conditional on other factors. These results are suggestive that high school GPA may be a better predictor of college GPA, conditional on other factors, than standardized test scores, possibly because they represent a combination of the student ability captured in standardized test scores and the student's underlying motivation and effort placed into school.

Linear models may not the appropriate model for GPA since the standard four-point scale is both top- and bottom-censored and based on categorical grades. This is particularly problematic for GPA measured during the first semester or first year when GPA is determined by the grades in relatively few courses. As an alternative, I dichotomize GPA in both the first semester and the first year into four categories: 0 – 0.99, 1.0 – 1.99, 2.0 – 2.99, 3.0 +. I then used an ordered probit to estimate the effect of dormitory residence on the likelihood of being in these four parts of the grade distribution. The cofficients on dormitory residence, while not generally interpretable because of the nonlinear model, suggest that dormitory residence has a positive effect on both fall semester GPA in column (iii) and first-year GPA in column (iv). The lower panel presents the marginal effects estimated at the average of all covariates on the probability of being in the bottom category (GPA < 1.0, equivalent to less than a "D" average) or being in the top category (GPA ≥ 3.0, equivalent to having a "B" average or higher). The results suggest that dormitory residence is associated with a 2.2 percentage point decrease in the likelihood of having a fall semester GPA below 1.0 and an 8.3 percentage point increase in the probability of having a fall semester GPA at 3.0 or higher. The corresponding marginal effects for first-year GPA are a decrease of 3.1 percentage points for having a GPA below 1.0 and an 8.7 percentage point increase in the likelihood of have a GPA at or above 3.0. The results suggest that dormitory residence is strongly correlated with having a high GPA, above 3.0, but somewhat less strongly associated with avoiding a low or failing GPA.

The results from the IV approach in the linear model are presented for fall GPA in column (v) and for first-year GPA in column (vi). The IV results are smaller in magnitude than the OLS results, particularly for first-year GPA, but continue to show a modest positive and statistically significant effect of dormitory residence on student academic success. The estimated effects correspond to a 6.2 percent ($= \frac{0.159}{2.561} \times 100$) increase in fall semester GPA and a 6.5 percent ($= \frac{0.163}{2.476} \times 100$) increase in first-year GPA. Similarly, the coefficients and marginal effects in the ordered probit estimates utilizing the instruments are smaller, again more so for the first-year GPA. The results in column (viii) suggest that dormitory residence decreases the likelihood of having a first-year GPA below 1.0 by 1.9 percentage points, a statistically significant but not necessarily substantial effect. In comparison, dormitory residence increases the likelihood of having a first-year GPA above 3.0 by 5.6 percentage points.

Overall, the results in Tables 3 and 4 suggest that dormitory residence may have a positive effect on student academic success which does not translate into higher likelihood of retention. One possible reason is that student retention may be largely determined by factors other than in-class performance, for example financial concerns or personal interest in education. Thus, increasing the grade point average of students is not, in and of itself, a means of increasing student retention. However, the GPA effects estimated by IV methods are not substantial and therefore the change in academic performance may simply not be large enough to translate into student retention. An alternative explanation for why the GPA gains do not increase retention is that dormitory residence has only a moderate effect on preventing students from failing out of school. The results are suggestive that dormitory residence serves more to increase academic success at the top of the grade distribution rather than lowering the likelihood that a student is "unsuccessful." I find similar evidence from an ordered probit model that uses three categories of first-year GPA that correspond to being eligible for dismissal (GPA<1.5), students neither eligible for dismissal nor eligible for the Honors College (1.5≤GPA<3.3) and those students eligible for the Honors College (GPA>3.3).[15] The marginal effects from the IV approach suggest that dormitory residence reduces the likelihood of being eligible for dismissal by 2.5 percentage points while increasing the likelihood of being eligible for the Honors College by 3.2 percentage points.

3.2 Robustness Checks and Discussion

The results suggest that dormitories have no effect on student retention and only a moderate effect on student grade point average, among those students induced to live in the dormitories because of the commuting costs and commuter exemption rules. However, it is

[15] Students can apply directly to the Honor's College before matriculation but the 3.3 GPA rule would apply to any student deciding to apply once they have matriculated, and thus provides a useful benchmark for comparing student success.

possible that the commuting variables also affect the likelihood that a student appears in the sample. Some students who would have attended this university may choose not to do so because they do not want to be in the dormitories, either because of the cost, a need to live at home, or a lack of interest in living in residence halls. These students could, instead, choose to not attend college at all, or more likely, choose to attend a different institution.

Such changes in attendance behavior could lead to sample selection bias, although the exact direction of the bias depends on whether students are positively or negatively selected into the sample. However, it is unlikely that such changes are occurring. First, in the geographic area under consideration, the university studied in this paper is the top public institution and the local options students would turn to are either two-year colleges or four-year institutions with lower student test scores and graduation rates. Thus, there is no comparable local option for students to attend. Second, there is little evidence in the data of a differential change in the type of student attending college based on the instruments. Looking at students within 10 miles of the commuter exemption border, students eligible to commute have ACT scores of 21.68 and high school GPAs of 3.22 compared to ACT scores of 21.38 and high school GPAs of 3.22 for those students not eligible to commute. Furthermore, in the same geographic area there is no evidence of a difference in the likelihood of students being in the sample. I find that the total number of students who matriculated to the university during my sample period in each ZIP code, divided by the number of people in the ZIP code aged 25 or higher with a BA in 2000, is 0.032 on both sides of the commuter exemption border. Thus, there is no apparent difference in the likelihood of students appearing in my sample due to the commuter exemption rule.

The instrumental variable estimates previously reported are also robust to a variety of sample restrictions and specifications of the instrument set. For example, column (i) of Table 5 presents the IV results for each of the retention and GPA outcomes previously considered but estimated in the full sample, instead of the financial aid sample, without family background controls which are missing for the non-financial aid students. The results are similar for all outcomes to the previous IV estimates in the financial aid sample. There is no evidence that dormitory residence is associated with a change in the likelihood of the student being retained but dormitory residence does appear to have a moderate positive effect on the grade point average of students. Column (ii) presents the results from the financial aid sample but not limiting the sample to those students that live within 80 miles. Again, the results are largely consistent with the previous estimates.[16] Column (iii) presents the results using only the commuter exemption rule, and not commuting time, as the instrument for dormitory residents. While the precision of the estimates decreases, consistent with the weaker instrument set being used, the results are unchanged. Column (iv) includes information on whether the commute includes tolls which again has little impact. Overall, the IV estimates presented previously are robust to a variety of sample selections and the specification of the instrument set.

3.3 Results by Student Ability

While the IV estimates in the previous section suggest that dormitory residence has, on average, no impact on student retention and a moderate impact on academic success, the potential effects of on-campus residency could vary based on student characteristics. For example, one could argue that dormitories would have the largest impact among the more marginal students both because these students are more likely to need help or because the least marginal students are not in need of help. Alternatively, it is possible that the marginal students

[16] Note that as the distance from the university increases, students may have better local options for attending college. That the primary results do not change compared to the base sample provides additional evidence that the estimated effects of dormitory residence are not determined by the instruments changing college choices of potential students.

are so marginal that very little could improve their performance and therefore all potential positive effects are felt by less marginal students. To investigate the possibility of heterogeneous effects, I separately estimate the impacts of dormitory residence across student ability groups, as measured by terciles of student ACT scores.[17] The range of ACT scores of students in my sample spans the range of national scores and therefore provides potentially useful information about how dormitory may effect students across the ability distribution.

Table 6 presents the non-IV estimated effects of dormitory residences on all outcomes for students separately based on terciles of ACT scores in columns (i) – (iii). The top panel presents the marginal effects of dormitory residence on both measures of student retention, the middle panel presents the results from linear models of both measures of GPA, and the bottom panel presents the marginal effects, estimated from ordered probits, of dormitory residence on GPA < 1.0 and GPA ≥ 3.0 in both the fall semester and first-year. For nearly all outcomes, the positive association with dormitory residence decreases as student ability increases. For example, dormitory residence is associated with a 9.9 percentage point increase in retention to the second year in the bottom tercile of student ability, compared to a 6.3 and 3.2 percentage point increase in the middle and top terciles, respectively. Similarly, dormitory residence is associated with a 0.283 point increase in first-year GPA in the bottom tercile compared to a 0.228 and a 0.119 point increase in the middle and top terciles. The only exception to the decreasing positive effects of dormitory residence as ability increases is for the positive effect of dormitory residence on the likelihood of having a GPA above 3.0, which is similarly large in magnitude across ability terciles.

[17] The average ACT score at the university in the bottom tercile is 18, corresponding to the 34th percentile in the nation from 2009-2011 according to ACT, Inc. (http://www.actstudent.org/scores/norms.html). Similarly, the average scores in the middle and top terciles at the university are 22 and 26, corresponding to the 62nd and 84th percentiles of the national distribution.

Columns (iv) – (vi) of Table 6 present the IV estimates across student ability for all outcomes. The first-stage regression is estimated, and residuals are calculated for inclusion in the second-stage, separately within each ability tercile. Importantly, within each tercile, the instruments appear strong and the overidentification test is satisfied. The IV estimates suggest that dormitory residence has no statistically significant positive impact on either measure of retention across ability, and in fact may have a small negative impact among the highest ability students, possibly reflecting a higher likelihood of transfer among these students (Light and Strayer, 2004). At the very least, these results do not lend support to the idea that dormitory residence increases institutional retention.

In comparison, the IV estimates on all GPA outcomes are similar to the non-IV estimates among the lowest ability students, suggesting that dormitory may have a modest positive impact on academic success among these students. The IV estimates suggest smaller, but still positive, effects of dormitory residence on student GPA in the middle tercile and suggest that dormitory residence has no impact on GPA among the students in the highest ability tercile. Overall, the IV results in Table 5 suggest that while dormitory residence has no impact on student retention across student ability, on-campus residency may have a small but positive impact on GPA among the lower-ability students.[18]

3.4 Dormitory Residence and Peer Effects

One potential way in which dormitory residence may be expected to influence academic performance is through interactions with peers. Using quasi-random variation in rooming assignments, Sacerdote (2001) and Zimmerman (2003) find moderate evidence that the academic

[18] It is also possible that the effects of dormitory residence could vary across students based on family resources. I find little difference in the effects of dormitory residence across income groups in my sample. This could suggest that there is no variation across family income but more likely represents the fact that all of the students in the sample applied for financial aid and therefore have "financial need." Consequently, I do not report these results but they are available from the author upon request.

qualifications of roommates among students living in the dormitories impact educational outcomes of students during the first year. However, Foster (2006) raises questions about the underlying mechanisms through which peer effects could operate since the author finds no difference in the magnitude of the effect of peer ability between randomized roommates and "friends" suggesting that social attachments are not a channel through which peer effects operate. Stinebrickner and Stinebrickner (2006) argue that peer effects may operate through influence on behavior, such as study habits, that may not be captured by investigating peer ability. Similarly, it may be that simply having a roommate, instead of living at home, could have positive effects on academic success during the first year. For example, being in an environment where students are studying could nudge an individual towards studying more. Thus, it may be that dormitory residence, to the extent that it has any effect on academic outcomes at all, may operate through a more generalized form of peer effects or peer "exposure" than previously studied.

To test this hypothesis, I investigate whether the effects of on-campus residency has a differential impact based on whether the on-campus student lives alone, lives with a roommate(s) of lower ability than the student, or lives with a roommate(s) of higher ability than the student. Given the previously estimated strong correlation between high school and college GPA, I measure the quality of roommate based on the high school GPA, or average high school GPA for multiple roommates. Investigating this question is difficult because there are two levels of selection: first students may select into the dormitories and then students may select into rooming environments based on preferences over rooming assignments. The commuting instruments only solve the first problem, selection into on-campus residency, but do not solve the second selection concern. In fact, the commuting variables have no measurable impact on the likelihood that students living in the dormitories have a roommate or whether a roommate is higher or lower ability.[19] This is likely because the assignment of single rooms is highly supply-constrained. Conversations with residence services at this university suggest that many students request singles but that such requests far outstretch the available supply of single rooms each year. Thus, the majority of such requests cannot be met and only about twenty percent of students receive single rooms. The single room supply constraint also appears to be binding in the data as there is little evidence that observable characteristics, including family income and student ability, affect the likelihood of students having roommates (Table A-2).[20]

I use the same instrumental variable methodology as before in separate samples of students to investigate whether students induced to live on campus during the first year experience different effects based on the living arrangements they ultimately experience. Unfortunately, unlike the previous papers on peer effects, I do not have access to the residential hall room preferences submitted by the students which are used in the literature to control for the roommate assignment mechanism. Therefore, it is possible that some, though likely not all, of the estimated differences between those students with and without roommates or between different roommate abilities could be due to residual bias because of selection into rooming assignments due to unobservable characteristics. Thus, the results in this section should be interpreted as suggestive evidence of whether dormitory residence may be associated with some form of peer effects. First, I consider the sample of students living at home or living in dormitories without roommates, thus removing the students with roommates from the sample.

---

[19] For example, the marginal effects from a probit of having a roommate, among students in dormitories, are -0.029 (0.045) for being in the commuter exemption area, -0.000 (0.000) for commuting time, and 0.000 (0.001) for the interaction. Among students living with roommates, the marginal effects of having a lower quality roommate or roommates are -0.045 (0.091) for being in the commuter exemption area, -0.000 (0.001) for commuting time and 0.001 (0.001) for the interaction. Results from a multinomial logit across all three categories are similar.

[20] One might also be concerned about students who request to live with friends but Foster (2006) presents evidence that the effects of peer ability do not vary between friends and randomly assigned roommates. Consistent with this evidence, none of the results on peer ability are substantively affected when I drop the 22 percent of cases where students have a roommate who attended the same high school. These results are available upon request.

Second, I use the sample of students living at home or living in dormitories with lower-ability roommates. And third, I consider the sample of students living at home or living in the dormitories with higher-ability roommates.

Column (i) of Table 7 presents the non-IV effects of dormitory residence among the sample of students that either live off campus or on campus with no roommates during their first year. The non-IV estimated effects of dormitory residence among the sample of students that either live off campus or on campus with lower-ability roommates is presented in column (ii) and for student living off campus or on campus with a roommate of higher ability in column (iii). While the non-IV estimates are positive regardless of the living arrangements of students, the estimated effects of dormitory residence are larger for those students who end up living in the dormitories with roommates who had a higher high school GPA than the student.

The corresponding IV estimates in columns (iv) through (vi) suggest that only those students induced to live in the dormitory and have a higher-ability roommate experience positive effects on academic performance compared to living at home. Such students experience increases in average GPA in both the first semester and first year as well as relatively large likelihoods of having GPAs above 3.0. Additionally, such students experience a 3.8 percentage point increase in the likelihood of being retained to the second semester, the only statistically significant and positive effect estimated on retention using the instruments in any of the analyses thus far. However, there is no positive effect on retention to the second year, even among these students. In contrast, those students induced to live in the dormitories but who live alone or live with roommates whose high school GPA was lower than the student's see no economically significant effect on any measures of retention or grade point average.

Again, while the evidence in Table 7 should be taken as suggestive because the underlying assignment that determines roommates is unobserved, these results are broadly supportive of the previous literature on peer ability. Additionally, the results are consistent with the findings in Table 6 that positive dormitory effects decrease with ability, as higher ability students are less likely to have roommates with high school GPA that exceeds theirs. The results suggest that students induced to live in the dormitories experience increases in academic performance, provided that they are paired with roommates of higher ability. However, the results suggest no role for overall peer exposure as simply being on campus but without a higher ability roommate has no discernible effect on student performance compared to living at home. The results are striking because even students living alone or with lower-ability roommates would be expected to have access to high-ability peers in the same dorm or on the same floor but such access appears to have no effect. Additionally, while students who live alone may be less engaged in their college community than students with roommates, the lack of any positive effect of having a lower-quality roommate raises questions about whether student engagement can be operationalized through dormitory residence or whether engagement itself has a causal effect on student success.

4. Conclusion

This paper attempts to identify the causal effect of dormitory residence during the first-year of college on academic success and retention using unique data from a public four-year university with a commuter exemption rule that changes the likelihood that a student is in the dormitories based on location of their parent's residence. Consistent with the previous literature, standard regression methods relying on observable characteristics to solve the underlying selection problem of students into dormitories show positive impacts of dormitory residence on both student GPA and retention rates. In contrast, the estimates utilizing the commuter exemption rule and potential commuting costs as instruments suggest smaller effects of dormitory residence on student higher education outcomes. For those students induced to live in the dormitories due to the instruments, dormitory residence has no measurable effect on retention to the second semester or the second year. Additionally, the IV estimates suggest only modest effects on student GPA with the largest effects found at the top of the GPA distribution.

Further analysis suggests some heterogeneity across student ability in the estimated effects of dormitory residence on student academic performance with larger effects found for students from the lowest tercile of ACT scores and no effect found among those students in the top ACT tercile. However, there is no effect on retention for any ability group suggesting that on-campus residency requirements do not serve to increase the likelihood of students returning to the institution. Finally, I find suggestive evidence that any potential positive effects of dormitory residence on student college GPA may operate through the positive effects of peer ability previously found at institutions in other sectors of the higher education market. Those students who are induced to live in the dormitories and live with a roommate or roommates who had higher high school GPAs compared to the student show increases in overall GPA and large increases in the likelihood of having grade point averages over 3.0. This evidence is consistent with decreasing effects of dormitory residence with student ability as lower ability students are more likely to have roommates with higher high school GPAs. However, there is no evidence that those students induced to live in the dormitories, but who ultimately live alone or with lower-ability roommates, experience any change in their likelihood of retention or any change in any measure of GPA. Overall, these results suggest that simply inducing students to live in dormitories has little impact on student success for the average student but could have positive effects, depending on the rooming assignments, for lower ability students.

References

Angrist, Joshua D. and Guido W. Imbens, 1994. "Identification and estimation of local average treatment effects." *Econometrica*, 62(2), pp.467-475.

Astin, Alexander W., 1973. "The impact of dormitory living on students." *Educational Record*, 54(9), pp. 205-210.

Astin, Alexander W., 1984. "Student involvement: A developmental theory for higher education." *Journal of College Student Personnel*, 25(4), pp. 297-308.

Astin, Alexander W. and Leticia Oseguera, 2005. "Pre-college and institutional influences on degree attainment" in Alan Seidman, ed., *College Student Retention: Formula for Student Success*, Westport, CT: Praeger, pp. 297-308.

Bound, John, David A. Jaeger, and Regina M. Baker, 1995. "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak." *Journal of the American Statistical Association*, 90(430), pp 443-450.

Braxton, John M. and Amy S. Hirschy, 2005. "Theoretical developments in the study of college student departure." in Alan Seidman, ed., *College Student Retention: Formula for Student Success*, Westport, CT: Praeger, pp. 297-308.

Chickering, Arthur W., 1974. *Commuting versus resident students*. San Francisco, CA: Jossey-Bass.

Foster, Gigi, 2006. "It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism." *Journal of Public Economics*, 90 (8-9), pp. 1455-1475.

Hausman, J. A., 1978. "Specification tests in econometrics." *Econometrica*,46(6), pp. 1251-1271.

Kuh, George D., Ty M. Cruce, Rick Shoup, Jillian Kinzie, and Robert M. Gonyea, 2008. "Unmasking the Effects of Student Engagement on First-year College Grades and Persistence." *Journal of Higher Education*, 79(5), pp. 540-563.

Light and Strayer, 2004. "Who receives the college wage premium? Assessing the labor market returns to degrees and college transfer patterns." *Journal of Human Resources*, 39(3), pp. 746-773.

Official Proceedings of the One Thousand Four Hundred and Thirty-ninth Meeting of the Board of Trustees, Ohio State University, Columbus, OH, Feb. 1, 2008. http://trustees.osu.edu/assets/files/minutes/documents/020108_000.pdf

Sacerdote, Bruce, 2001. "Peer effects with random assignment: Results for Dartmouth roommates." *Quarterly Journal of Economics*, 116(2), pp. 705-746.

Stinebrickner, Todd R. and Ralph Stinebrickner, 2006. "What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds." *Journal of Public Economics*, 90(8), pp. 1435-1554.

Terza, Joseph V. Anirbush Basu, and Paul J. Rathouz, 2008. "Two-stage residual inclusion estimation: Addressing endogeneity in health economic modeling." *Journal of Health Economics*, 27(3), pp. 531-543.

Tinto, Vincent, 1975. "Drop from higher education: A theoretical synthesis of recent research." *Review of Educational Research*, 45(1): pp. 89-125.

Woodridge, Jeffrey M., 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA.

Zimmerman, David J., 2003. "Peer effects in academic outcomes: Evidence from a natural experiment." *The Review of Economics and Statistics*, 85(1), pp. 9-23.

Table 1: Summary Statistics for the Estimation Sample

	Full Sample		On-campus		Off-campus		
	Mean	St. Err.	Mean	St. Err.	Mean	St. Err.	Difference
On-campus	0.774	0.004					
Outcomes							
Second semester	0.888	0.003	0.905	0.003	0.830	0.007	0.074***
Second year	0.790	0.003	0.804	0.004	0.742	0.008	0.062***
GPA, fall	2.561	0.008	2.596	0.009	2.439	0.019	0.157***
GPA, first year	2.476	0.008	2.521	0.009	2.325	0.020	0.196***
GPA > 3.0, first year	0.336	0.004	0.342	0.005	0.317	0.008	0.024**
GPA < 1.0, first year	0.095	0.002	0.079	0.003	0.152	0.006	-0.074***
Variables							
Male	0.371	0.004	0.355	0.005	0.427	0.009	-0.072***
White	0.886	0.003	0.870	0.003	0.939	0.004	-0.069***
Black	0.087	0.002	0.103	0.003	0.032	0.003	0.071***
Hispanic	0.013	0.001	0.014	0.001	0.008	0.002	0.007***
Other minority	0.014	0.001	0.012	0.001	0.021	0.003	-0.008***
HS GPA	3.176	0.004	3.169	0.004	3.200	0.008	-0.031***
ACT score	21.467	0.031	21.403	0.035	21.685	0.064	-0.282***
HS, public	0.894	0.003	0.882	0.003	0.936	0.004	-0.054***
HS, student-teacher ratio	16.720	0.020	16.771	0.022	16.546	0.041	0.225***
Parental education, BA or higher	0.580	0.004	0.598	0.005	0.521	0.009	0.076***
Parent's married	0.724	0.004	0.720	0.004	0.737	0.008	-0.017*
Family income ($10,000)	8.110	0.047	8.285	0.055	7.510	0.087	0.775***
ZIP, % less than HS	0.142	0.001	0.147	0.001	0.127	0.001	0.019***
ZIP, % HS	0.359	0.001	0.360	0.001	0.357	0.002	0.004*
ZIP, % AA	0.056	0.000	0.057	0.000	0.054	0.000	0.003***
ZIP, % BA or higher	0.235	0.001	0.231	0.001	0.250	0.002	-0.019***
ZIP, unemployment rate	0.043	0.000	0.044	0.000	0.040	0.000	0.004***
ZIP, per capita income ($10,000)	2.614	0.006	2.606	0.007	2.639	0.012	-0.033**
ZIP, poverty rate	0.078	0.001	0.078	0.001	0.076	0.001	0.002*
ZIP, % urban	0.810	0.002	0.813	0.003	0.797	0.005	0.017***
ZIP, % minority	0.103	0.001	0.111	0.002	0.074	0.002	0.037***
Observations	14161		10964		3197		

Notes:
1) Other variables utilized in the analysis include age of the student, an indicator for dependency status of the student and an indicator for missing parental marriage information.

Table 2: Marginal Effects from Probits of Observable Characteristics on the Likelihood of Living On-campus During First-year

Variables	(i)	(ii)
Male	-0.057***	-0.050***
	(0.010)	(0.008)
Black	0.133***	0.104***
	(0.021)	(0.010)
Hispanic	0.087***	0.018
	(0.026)	(0.032)
Other minority	-0.059*	-0.041
	(0.033)	(0.034)
HS GPA	-0.029**	-0.008
	(0.012)	(0.008)
ACT score	0.001	0.002**
	(0.002)	(0.001)
HS, public	-0.093***	-0.040***
	(0.025)	(0.012)
HS, student-teacher ratio	0.012	-0.002
	(0.008)	(0.003)
Parental education, BA or higher	0.050***	0.029***
	(0.009)	(0.009)
Parent's married	-0.022***	-0.019**
	(0.008)	(0.008)
Family income ($10,000)	0.006***	0.005***
	(0.001)	(0.001)
Within commuter exemption area		-0.206***
		(0.017)
Driving time to campus		0.002**
		(0.001)
Exemption area * drive time		0.010***
		(0.001)
Joint significance of instruments (F-statistic)		161.83***
		[0.000]
Overidentification test (chi-square)		0.302
		[0.583]
Observations	14161	14161

Notes:
1) Standard errors are presented below marginal effects in parentheses while p-values for statistical tests are presented in square brackets beneath test statistics. Asterisks denote statistical significance at the 10% (*), 5% (**) and 1% (***) levels.
2) Regressions also include student's age, an indicator for parental marriage status missing, and an indicator for dependency status of the child, as well as ZIP code measures of percent with less than high school, percent high school graduate, percent with associate's degree, percent with at least a bachelor's degree, unemployment rate, per capital income, poverty rate, percent urban and percent minority. Regressions also include cohort fixed effects.

Table 3: Marginal Effects of Living On-campus During the First Year on Student Retention

	Probit		Probit, IV	
	Second semester	Second year	Second semester	Second year
	(i)	(ii)	(iii)	(iv)
On-campus	0.058***	0.068***	0.010	-0.014
	(0.008)	(0.012)	(0.010)	(0.018)
HS GPA	0.101***	0.188***	0.100***	0.187***
	(0.006)	(0.009)	(0.006)	(0.010)
ACT score	-0.001**	0.002**	-0.001**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
Family income ($10,000)	0.004***	0.006***	0.004***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Parental education, BA or higher	0.011**	0.023***	0.014***	0.027**
	(0.004)	(0.006)	(0.004)	(0.006)

Notes:
1) Standard errors are presented below marginal effects in parentheses. Asterisks denote statistical significance at the 10% (*), 5% (**) and 1% (***) levels.
2) Regressions also include indicators for the student's gender and race, age, indicators for whether parent's are married or for parental marriage status missing, and an indicator for dependency status of the child, indicators for whether the student attended a public high school, the teacher-student ratio at the student's high school, as well as ZIP code measures of percent with less than high school, percent high school graduate, percent with associate's degree, percent with at least a bachelor's degree, unemployment rate, per capital income, poverty rate, percent urban and percent minority. Regressions also include cohort fixed effects.
3) IV estimates are produced using two-stage residual inclusion (2SRI) where the instrumental variables in the first-stage regression are an indicator for whether the student resides in the commuter exemption zone, the commuting time to campus, and an interaction of the two.

Table 4: Effect of Living On-campus During the First Year on Student Grade Point Average

	Linear regression		Ordered Probit		IV		Ordered Probit, IV	
Coefficients	Fall GPA (i)	First-year GPA (ii)	Fall GPA (iii)	First-year GPA (iv)	Fall GPA (v)	First-year GPA (vi)	Fall GPA (vii)	First-year GPA (viii)
On-campus	0.187***	0.225***	0.223***	0.258***	0.159***	0.163***	0.199***	0.164***
	(0.026)	(0.025)	(0.031)	(0.031)	(0.041)	(0.040)	(0.055)	(0.056)
HS GPA	1.013***	1.066***	1.339***	1.403***	1.012***	1.064***	1.339***	1.401***
	(0.023)	(0.024)	(0.033)	(0.032)	(0.023)	(0.024)	(0.033)	(0.032)
ACT score	0.021***	0.021***	0.031***	0.033***	0.021***	0.021***	0.031***	0.033***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)
Family income ($10,000)	0.009***	0.011***	0.012***	0.015***	0.009***	0.011***	0.012***	0.015***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Parental education, BA or higher	0.055***	0.075***	0.073***	0.096***	0.057***	0.078***	0.074***	0.101***
	(0.013)	(0.014)	(0.019)	(0.020)	(0.013)	(0.014)	(0.019)	(0.020)
Marginal Effects of On-campus								
GPA < 1.0			-0.022***	-0.031***			-0.019***	-0.019***
			(0.003)	(0.004)			(0.006)	(0.007)
GPA > 3.0			0.083***	0.087***			0.074***	0.056***
			(0.011)	(0.010)			(0.020)	(0.019)

Notes:
1) Standard errors are presented below coefficients or marginal effects in parentheses. Asterisks denote statistical significance at the 10% (*), 5% (**) and 1% (***) levels.
2) Regressions also include indicators for the student's gender and race, age, indicators for whether parent's are married or for parental marriage status missing, and an indicator for dependency status of the child, indicators for whether the student attended a public high school, the teacher-student ratio at the student's high school, as well as ZIP code measures of percent with less than high school, percent high school graduate, percent with associate's degree, percent with at least a bachelor's degree, unemployment rate, per capital income, poverty rate, percent urban and percent minority. Regressions also include cohort fixed effects.
3) Ordered probit models are estimated on four categories of GPA: 0-0.99, 1.0-1.99, 2.0-2.99, 3.0+ .
4) IV estimates are produced using two-stage residual inclusion (2SRI) where the instrumental variables in the first-stage regression are an indicator for whether the student resides in the commuter exemption zone, the commuting time to campus, and an interaction of the two.

Table 5: Robustness Checks on Instrumental Variable Estimates

	Full sample (i)	No distance limit (ii)	Single instrument (iii)	Including tolls (iv)
Marginal effects from probit				
Second semester	0.017	0.003	-0.016	0.008
	(0.011)	(0.010)	(0.028)	(0.011)
Second year	0.000	-0.036**	-0.047	-0.018
	(0.017)	(0.017)	(0.044)	(0.019)
Linear regression				
Fall GPA	0.169***	0.146***	0.184*	0.170***
	(0.042)	(0.039)	(0.109)	(0.041)
First-year GPA	0.191***	0.146***	0.116	0.163***
	(0.041)	(0.038)	(0.113)	(0.041)
Marginal effects from ordered probit				
Fall GPA < 1.0	-0.025***	-0.017***	-0.023	-0.021***
	(0.007)	(0.005)	(0.016)	(0.006)
Fall GPA > 3.0	0.071***	0.069***	0.086	0.081***
	(0.019)	(0.019)	(0.054)	(0.020)
First-year GPA < 1.0	-0.030***	-0.015**	-0.013	-0.019***
	(0.009)	(0.006)	(0.019)	(0.007)
First-year GPA > 3.0	0.064***	0.048***	0.040	0.058***
	(0.017)	(0.018)	(0.053)	(0.018)
N	21802	16602	14161	14161
Joint significance of instruments (F-statistic)	172.79***	218.08	37.10***	75.18***
	[0.000]	[0.000]	[0.000]	[0.000]
Overidentification test (chi-square)	0.223	0.942		0.316
	[0.637]	[0.332]		[0.854]

Notes:
1) Standard errors are presented below coefficients or marginal effects in parentheses. Asterisks denote statistical significance at the 10% (*), 5% (**) and 1% (***) levels.
2) Regressions include indicators for the student's gender and race, age, high school GPA, ACT score, parental education indicators, real family income, indicators for whether parent's are married or for parental marriage status missing, and an indicator for dependency status of the child, indicators for whether the student attended a public high school, the teacher-student ratio at the student's high school, as well as ZIP code measures of percent with less than high school, percent high school graduate, percent with associate's degree, percent with at least a bachelor's degree, unemployment rate, per capital income, poverty rate, percent urban and percent minority. Regressions also include cohort fixed effects.
3) Ordered probit models are estimated on four categories of GPA: 0-0.99, 1.0-1.99, 2.0-2.99, 3.0+ .
4) IV estimates are produced using two-stage residual inclusion (2SRI) where the instrumental variables in the first-stage regression are an indicator for whether the student resides in the commuter exemption zone, the commuting time to campus, and an interaction of the two.

Table 6: Effect of Living On-campus During the First Year on Student Outcomes by ACT Score Terciles

	Non-IV			IV		
	Bottom Tercile	Middle Tercile	Top Tercile	Bottom Tercile	Middle Tercile	Top Tercile
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Marginal effects from probit						
Second semester	0.089***	0.055***	0.024**	0.021	0.011	-0.001
	(0.012)	(0.013)	(0.010)	(0.017)	(0.018)	(0.013)
Second year	0.099***	0.063***	0.032**	0.013	-0.018	-0.048***
	(0.018)	(0.019)	(0.014)	(0.027)	(0.026)	(0.018)
OLS						
Fall GPA	0.230***	0.187***	0.101***	0.220***	0.118*	0.066
	(0.034)	(0.037)	(0.033)	(0.059)	(0.060)	(0.052)
First-year GPA	0.283***	0.228***	0.119***	0.236***	0.131**	0.054
	(0.031)	(0.037)	(0.035)	(0.051)	(0.062)	(0.055)
Marginal effects from ordered probit						
Fall GPA < 1.0	-0.037***	-0.024***	-0.007**	-0.039***	-0.017*	-0.003
	(0.006)	(0.005)	(0.003)	(0.013)	(0.009)	(0.004)
Fall GPA > 3.0	0.068***	0.087***	0.048**	0.071***	0.063**	0.020
	(0.010)	(0.017)	(0.019)	(0.019)	(0.034)	(0.030)
First-year GPA < 1.0	-0.057***	-0.029***	-0.012***	-0.041***	-0.012	-0.007
	(0.008)	(0.006)	(0.004)	(0.014)	(0.010)	(0.006)
First-year GPA > 3.0	0.063***	0.084***	0.068***	0.048***	0.037	0.039
	(0.007)	(0.016)	(0.019)	(0.014)	(0.029)	(0.034)
N	5993	4348	3820	5993	4348	3820
Joint significance of instruments (F-statistic)				103.74***	132.81***	153.93***
				[0.000]	[0.000]	[0.000]
Overidentification test (chi-square)				0.031	1.134	0.06
				[0.861]	[0.287]	[0.937]

Notes:
1) Standard errors are presented below coefficients or marginal effects in parentheses. Asterisks denote statistical significance at the 10% (*), 5% (**) and 1% (***) levels.
2) Regressions include indicators for the student's gender and race, age, high school GPA, parental education indicators, real family income, indicators for whether parent's are married or for parental marriage status missing, and an indicator for dependency status of the child, indicators for whether the student attended a public high school, the teacher-student ratio at the student's high school, as well as ZIP code measures of percent with less than high school, percent high school graduate, percent with associate's degree, percent with at least a bachelor's degree, unemployment rate, per capital income, poverty rate, percent urban and percent minority. Regressions also include cohort fixed effects.
3) Ordered probit models are estimated on four categories of GPA: 0-0.99, 1.0-1.99, 2.0-2.99, 3.0+ .
4) IV estimates are produced using two-stage residual inclusion (2SRI) where the instrumental variables in the first-stage regression are an indicator for whether the student resides in the commuter exemption zone, the commuting time to campus, and an interaction of the two.

Table 7: Effect of Living On-campus During the First Year on Student Outcomes by Whether the Student Lives Alone, a Lower Ability Roommate or a Higher Ability Roommate

	Non-IV			IV		
	No roommate	Lower ability roommate	Higher Ability roommate	No roommate	Lower ability roommate	Higher ability roommate
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Marginal effects from probit						
Second semester	0.051***	0.046***	0.076***	0.009	0.004	0.038***
	(0.010)	(0.008)	(0.009)	(0.013)	(0.011)	(0.014)
Second year	0.058***	0.047***	0.097***	-0.007	-0.023	0.010
	(0.015)	(0.012)	(0.014)	(0.021)	(0.017)	(0.019)
OLS						
Fall GPA	0.156***	0.153***	0.248***	0.052	0.072*	0.236***
	(0.032)	(0.028)	(0.031)	(0.046)	(0.041)	(0.046)
First-year GPA	0.174***	0.189***	0.292***	0.070	0.098**	0.247***
	(0.032)	(0.028)	(0.029)	(0.046)	(0.041)	(0.046)
Marginal effects from ordered probit						
Fall GPA < 1.0	-0.020***	-0.013***	-0.034***	-0.009	-0.007	-0.036***
	(0.005)	(0.003)	(0.005)	(0.007)	(0.004)	(0.008)
Fall GPA > 3.0	0.066***	0.070***	0.098***	0.028	0.036	0.102***
	(0.015)	(0.014)	(0.013)	(0.021)	(0.022)	(0.020)
First-year GPA < 1.0	-0.024***	-0.020***	-0.047***	-0.009	-0.008	-0.039***
	(0.006)	(0.004)	(0.006)	(0.009)	(0.006)	(0.009)
First-year GPA > 3.0	0.059***	0.079***	0.098***	0.021	0.031	0.083***
	(0.014)	(0.013)	(0.011)	(0.021)	(0.024)	(0.018)
N	4932	6698	6444	4932	6698	6444
Joint significance of instruments (F-statistic)				763.64***	451.28***	510.72***
				(0.000)	(0.000)	(0.000)
Overidentification test (chi-square)				0.543	0.008	0.989
				(0.461)	(0.929)	(0.320)

Notes:
1) Standard errors are presented below coefficients or marginal effects in parentheses. Asterisks denote statistical significance at the 10% (*), 5% (**) and 1% (***) levels.
2) Regressions include indicators for the student's gender and race, age, high school GPA, ACT score, parental education indicators, real family income, indicators for whether parent's are married or for parental marriage status missing, and an indicator for dependency status of the child, indicators for whether the student attended a public high school, the teacher-student ratio at the student's high school, as well as ZIP code measures of percent with less than high school, percent high school graduate, percent with associate's degree, percent with at least a bachelor's degree, unemployment rate, per capital income, poverty rate, percent urban and percent minority. Regressions also include cohort fixed effects.
3) Ordered probit models are estimated on four categories of GPA: 0-0.99, 1.0-1.99, 2.0-2.99, 3.0+ .
4) IV estimates are produced using two-stage residual inclusion (2SRI) where the instrumental variables in the first-stage regression are an indicator for whether the student resides in the commuter exemption zone, the commuting time to campus, and an interaction of the two.

Table A-1: Summary Statistics in Full Sample

	Full Sample		On-campus		Off-campus		
	Mean	St. Err.	Mean	St. Err.	Mean	St. Err.	Difference
On-campus	0.750	0.003					
<u>Outcomes</u>							
Second semester	0.871	0.002	0.892	0.002	0.811	0.005	0.081***
Second year	0.740	0.003	0.760	0.003	0.680	0.006	0.080***
GPA, fall	2.479	0.007	2.526	0.008	2.335	0.016	0.191***
GPA, first-year	2.371	0.007	2.431	0.008	2.192	0.016	0.239***
GPA > 3.0, first year	0.314	0.003	0.322	0.004	0.293	0.006	0.028***
GPA < 1.0, first year	0.125	0.002	0.102	0.002	0.191	0.005	-0.089***
<u>Variables</u>							
Male	0.376	0.003	0.361	0.004	0.419	0.007	-0.057***
Black	0.082	0.002	0.097	0.002	0.038	0.003	0.059***
White	0.890	0.002	0.875	0.003	0.935	0.003	-0.059***
Hispanic	0.012	0.001	0.014	0.001	0.007	0.001	0.007***
Other minority	0.016	0.001	0.014	0.001	0.020	0.002	-0.006***
HS GPA	3.136	0.003	3.130	0.004	3.153	0.006	-0.023***
ACT score	21.355	0.025	21.307	0.029	21.499	0.049	-0.192***
HS, public	0.892	0.002	0.877	0.003	0.936	0.003	-0.060***
HS, student-teacher ratio	16.654	0.016	16.708	0.018	16.489	0.032	0.219***
Parental education, less than HS							
Parental education, HS							
Parental education, BA or higher							
Parent's married							
Family income ($10,000)							
ZIP, % less than HS	0.138	0.000	0.143	0.001	0.126	0.001	0.016***
ZIP, % HS	0.353	0.001	0.353	0.001	0.352	0.001	0.001
ZIP, % AA	0.056	0.000	0.057	0.000	0.054	0.000	0.003***
ZIP, % BA or higher	0.245	0.001	0.241	0.001	0.255	0.002	-0.014***
ZIP, unemployment rate	0.043	0.000	0.043	0.000	0.040	0.000	0.004***
ZIP, per capita income ($10,000)	2.707	0.005	2.705	0.006	2.712	0.010	-0.008
ZIP, poverty rate	0.076	0.000	0.077	0.000	0.076	0.001	0.001
ZIP, % urban	0.810	0.002	0.814	0.002	0.801	0.004	0.013***
ZIP, % minority	0.102	0.001	0.110	0.001	0.078	0.002	0.032***
Observations	21802		16359		5443		

Notes:
1) Other variables utilized in the analysis include age of the student, an indicator for dependency status of the student and an indicator for missing parental marriage information.

Table A-2: Marginal Effects from Probit of Observable Characteristics on Whether a Student in the Dormitory Has a Roommate

Variables	(i)
Male	-0.024**
	(0.010)
Black	-0.030
	(0.022)
Hispanic	0.044
	(0.031)
Other minority	-0.013
	(0.044)
HS GPA	0.025**
	(0.012)
ACT score	-0.003**
	(0.001)
HS, public	-0.049***
	(0.015)
HS, student-teacher ratio	-0.000
	(0.002)
Parental education, BA or higher	0.005
	(0.009)
Parent's married	0.019
	(0.012)
Family income ($10,000)	0.001
	(0.001)
Pseudo-R^2	0.011
Observations	9448

Notes:
1) Standard errors are presented below marginal effects in parentheses while p-values for statistical tests are presented in square brackets beneath test statistics. Asterisks denote statistical significance at the 10% (*), 5% (**) and 1% (***) levels.
2) Regressions also include student's age, an indicator for parental marriage status missing, and an indicator for dependency status of the child, as well as ZIP code measures of percent with less than high school, percent high school graduate, percent with associate's degree, percent with at least a bachelor's degree, unemployment rate, per capital income, poverty rate, percent urban and percent minority. Regressions also include cohort fixed effects.