Res High Educ (2009) 50:52–72
DOI 10.1007/s11162-008-9107-6

Impact of Increased Academic Intensity on Transfer Rates: An Application of Matching Estimators to Student-Unit Record Data

William R. Doyle

Received: 14 August 2007 / Published online: 24 September 2008
© Springer Science+Business Media, LLC 2008

Abstract The impact of increased academic intensity on transfer rates from community colleges to 4-year institutions has been estimated only from observational data, with the possibility of selection bias. This study uses matching estimators to overcome possible selection bias and estimate the causal impact of increased academic intensity on transfer rates. Using student unit record data from Tennessee for the years 1995 through 2004, I find that taking 12 or more credit hours increases the probability of transfer from between 11% and 15%.

Keywords Community college · Transfer · Matching · Policy

Among first time freshmen in the United States, 40% enter community colleges. Nearly 80% of these students intend to go on to attain a bachelor's degree. However, only 23% do so within 6 years (United States Department of Education 2005). These low attainment rates may be due to several factors, including lack of preparation, differing environments in community colleges and 4 year institutions, and lack of articulation between 4 year colleges and 2 year institutions.

Research done by Adelman (1999, 2004, 2006) shows one important contributor to eventual success in attaining a bachelor's degree for all students: high levels of academic intensity in the first year of higher education. Students who enter with and maintain a full, or close to full course load in their first year of higher education have degree completion rates that are nearly one third higher than their counterparts who do not (Adelman 2006).

This paper utilizes this finding from observational evidence in order to test its applicability as a policy intervention. To be applicable as an intervention, such a finding would need to be established as having a causal impact on transfer rates (Heckman et al. 1998a). This is essentially the goal of many experimental studies in educational evaluation.

W. R. Doyle (✉)
Peabody College of Vanderbilt University, 230 Appleton Pl., Nashville, TN 37203, USA
e-mail: w.doyle@vanderbilt.edu

Through randomized assignment, the causal effect of a policy intervention on a target population is identified (Heckman et al. 1997, 1998b).

This study does not rely on an experimental procedure, but rather on techniques for recovering causal estimates from observational data (Rubin 1974). In particular, I utilize matching methods to estimate the effect of increased academic intensity on transfer rates from community college students in Tennessee.

The plan of this paper is as follows. First, I briefly review some of the studies on community college transfers. I also highlight several of the findings from the series of studies completed by Adelman for the Department of Education. I then describe a counterfactual framework for estimating causal effects first described by Rubin (1974) and others (Holland 1986). The data and specific functional form to be used are next described, followed by results and conclusions. I find that the impact of increased academic intensity is to increase transfer rates, even after matching to ensure that the observed attributes of treated and control subjects are identical. This increase is not as large as would be predicted by traditional methods. I then conclude with policy implications.

Literature Review

There is a voluminous literature on transfer rates (Wassmer et al. 2004; Bradburn and Hurst 2001; Grubb 1991; Anderson et al. 2006; Dougherty and Kienzl 2006; Cheslock 2005; Shulock and Moore 2005; Light and Strayer 2004; Wassmer et al. 2004; Bailey and Weininger 2002; Surette 2001; Bradburn and Hurst 2001; Shaw and London 2001; Townsend 1995; Bauer and Bauer 1994; Lee et al. 1993; Prager 1993; Lee and Frank 1990). This study is primarily concerned with the impact of increased academic intensity in the community college on eventual transfer rates. This literature review briefly considers the literature that defines and describes transfer rates before going on to consider the literature that attempts to identify student and institutional characteristics that affect transfer rates. While this is a rich literature, with numerous studies, few have attempted to isolate that causal effect of increased academic intensity on transfer to a 4-year institution.

Transfer Studies

For the purposes of this literature review, the transfer studies under consideration are divided into two groups. The first group, descriptive studies, are those analyses which seek to define and measure one or several types of transfer rates. The problem with all transfer studies is defining the appropriate denominator. Should it be all students in community colleges? Those who seek a bachelor's degree? Those who make progress toward the associate's or other degree? Several of the important studies in this field are considered.

Next, I review the multivariate studies that have been done to predict and explain transfer rates. These studies have produced a number of important findings that will be used in estimating the effect of academic intensity on transfer rates in this study.

Descriptive Studies

Grubb, in his influential 1991 work on collegiate transfer rates using longitudinal data, identifies many of the key factors that may affect transfer rates (Grubb 1991). These

include the demographic backgrounds of students, high school achievement, the type of programs (academic or vocational) available in community colleges, and an increase in the number of students who take a small number of courses in the community college and then leave. Grubb refers to these students as "experimenters," and notes that these students are much less likely to transfer most likely due to insufficient high school counseling combined with a "*laissez faire*" approach to college counseling at the community colleges (Grubb 1991).

Other descriptive studies of transfer rates include Bradburn and Hurst (2001) study of transfer rates under alternative definitions of transfer. The authors find that the transfer rate rates are much higher among students who enrolled in an academic program, expected to complete a degree, and enrolled continuously through their first 2 years of higher education. Importantly for this study, the authors also find a higher enrollment rate among those students who enrolled for 12 or more credit hours in their first term within higher education.

Multivariate Studies

In addition to the studies that have been conducted to describe the transfer rate, multiple studies have been conducted to understand the factors that are associated with higher or lower transfer rates. These include studies that attempt to demonstrate the overall effect of community college enrollment on educational attainment, studies that attempt to estimate the influence of academic policies on transfer rates, and studies that estimate the influence of student characteristics on transfer rates (Kane 1999; Rouse 1998, 1995; Leigh and Gill 2003, 2004; Surette 2001). A smaller subset of studies investigates the impact of institutional characteristics on transfer rates (Wassmer et al. 2004; Shulock and Moore 2005).

Several studies seek to understand the overall effect of community colleges on educational attainment, either within individual states or within the United States as a whole. Rouse (1995) investigates what she terms the democratization and diversion effects of community colleges. She notes that while community colleges do increase access, they might divert students who would otherwise have completed a bachelor's degrees. While she does find a diversion effect of community colleges, the effect of democratization from these institutions more than offsets the losses from diversion. Leigh and Gill (2003) find a similar effect using a more recent data and a different strategy for the purposes of identification. Rouse (1998) using state-level data, again finds a possible diversion effect as a result of having more community college enrollment in a state, but an overall increase in attainment.

Several studies have taken up the impact of academic policies on transfer rates. Hilmer (1997) investigates the possibility that students may use community colleges as a cost-effective means to attend a high-quality 4-year institution. Hilmer finds that students who do transfer end up in institutions that are of higher quality than the students would have gone to directly from high school. Anderson et al. (2006) find that articulation agreements at the state level have little to no effect on the probability of transfer for individual students. There is evidence from other studies that articulation agreements are ill-formed and poorly structured even within individual institutions (Prager 1993).

Other studies have looked at the impact of student characteristics, such as race and gender on transfer rates. Surette (2001) finds that men are more likely than women to transfer from 2 year institutions to 4 year institutions. This finding is robust to multiple alternative specifications, including the effect of academic environments and labor-markets. Surette finds that marriage status and child-rearing responsibilities lower the probability of transfer for women. His findings are surprising in light of other results

demonstrating higher labor market payoffs for 4 year degrees for women than for men (Surette 2001).

Others have found that race and ethnicity may play a role in the community college transfer function. Lee and Frank (1990) find that students from African-American or Hispanic backgrounds are less likely to transfer from community colleges to 4 year institutions. However, their primary finding is that transfer rates primarily reflect the social class of individuals within the community colleges.

It is social disadvantage that impedes community college students from transferring, through the effect of social class on virtually all the academic behaviors associated with transferring. While social class is strongly and positively associated with almost all such behaviors, race/ethnicity is related only to a few, after social class is held constant.

Finally, some studies have investigated the impact of institutional characteristics on transfer rates. Wassmer et al. (2004) finds that institutions that enroll higher proportions of Latino or African-American students tend to have lower transfer rates, net of other factors. The impact of other factors, such as budget capacity and higher tuition and fees on transfer rates has also been explored (Shulock and Moore 2005). The evidence suggests that students may be discouraged from further enrollment by steep increases in fees.

This study will look at the impact of only one among the many possible factors that affect student completion: academic intensity upon initial enrollment.

Academic Intensity and College Completion

In an influential series of studies, Adelman (1999, 2004, 2006) studied the effect of pre-college and in-college experiences on college completion rates. Adelman's studies utilized an extraordinarily rich data set, with a combination of self-reported data from students and behavioral data from transcripts. In all of these studies, Adelman seeks to answer the question "What contributes most to bachelor's degree completion of students who attend 4-year colleges at any time in their careers?"(Adelman 1999, p. 1)

Adelman's primary findings in these studies is that the quality and amount of high school course-taking is the primary contributor to eventual attainment of the bachelor's degree. However, Adelman's most important finding for the purposes of this study has to do with the impact of increased academic intensity while enrolled in college (Adelman 1999, 2006).

Comparing students with more than 20 hours of completed coursework in their first year with those than less, Adelman finds that higher numbers of credit hours lead to increased college completion. He writes:

Earning less than 20 credits in the first calendar year following postsecondary entry is a distinct drag on degree completion… falling below the 20 credit threshold lessens the probability of completing a bachelor's degree by a *third* (Adelman 2006, p. 48-italics in original).

Adelman's findings provide substantial support for the idea that increased academic intensity in the first year of college can lead to increased completion. Adelman specifically recommends that students should "end their first calendar year of enrollment with 20 or more additive credits" (Adelman 2006, p. xxv). However, before acting on the findings of this study, the causal nature of the finding needs to be considered.

With the exception of the Rouse study, few studies have attempted to identify the causal impact of various post-matriculation interventions on transfer rates from the community college to the 4 year institution. The observational evidence available shows that students

who take more courses are substantially more likely to graduate. A natural conclusion from this literature would be to assume that policy interventions should be designed to encourage students to take more courses in their first year if at all possible. However, without more analysis, such an intervention might have unintended consequences.

If students who take more courses are more motivated or have better support from their community or family, then any intervention designed to encourage more coursetaking would have little to no impact on eventual transfer rates. On the other hand, if there is little difference between students with high levels of credits and those with low levels of credits, then the only plausible reason for the greater success of the former group is the credits themselves. This study attempts to isolate the effect of greater academic intensity on transfer rates. The next section provides a theoretical framework for counter factual inference and estimation that provides for the possibility of isolating causal effects from observational data.

Counter-Factual Framework for Estimation

The essential problem under study in this paper is one of causal inference. The observed association that has been extensively established is that students who take more course units in their first year of higher education are more likely to transfer and graduate than students who enroll in only a few courses (Adelman 1999, 2004, 2006).

This could be the result of a causal process. A hypothetical student taking more courses may end up being more intellectually engaged, while also becoming more integrated into campus life. If this same student had not taken more courses in her first year, then she may have decided not to pursue further postsecondary education, and failed to attain her goals (Tinto 1975; Braxton et al. 2000).

However, this association could also be the result of another, unobserved characteristic of students. A hypothetical student who takes more courses may be more motivated, have better family support, or have access to information about postsecondary success that students with lower course credits do not possess (Shaw and London 2001; Shaw and Coleman 2000).

For policymakers, this is a critically important problem. If the first condition holds, then a policy intervention designed to incentivize students to take more courses in their first year might be quite helpful in increasing transfer and graduation rates. If the second condition holds, policymakers would find that efforts to increase course credit attainment in the first year would not be effective in increasing transfer and graduation rates.

The purpose of this section is to lay out a counterfactual framework that will be used to make inferences from the analysis in the subsequent sections (Rubin 1974, 1976). For the purposes of this framework and subsequent analyses, the coursework component is reduced to a very simple intervention: did the student take a higher or lower number of credit hours in their first semester of enrollment?

The variable y_1 represents the transfer outcome for students who did in fact take a higher number of credit hours in their first semester of enrollment, while the variable y_o represents the transfer outcome for students who took a lower number of hours.[1] The

[1] "Higher" and "lower" numbers of hours are defined variably in this study. As explained in the methods section, I use cutoffs of 6, 9, and 12 h to define higher and lower numbers of credit hours.

impact of high levels of academic intensity for any individual (△) is represented by (Smith and Todd 2001):

$$\Delta = y_1 - y_0$$

Of course, we can not observe outcomes for any individual who simultaneously does and does not take a certain number of course credits. This is sometimes referred to as the fundamental problem of causal inference (Holland 1986). Instead, we observe different students, some of whom took more hours, others who did not. Let $z = 1$ denote students who took the higher number of credit hours, while $z = 0$ denote students who took the lower number of credit hours.

The students who take higher levels of credit hours are most likely different than their peers who took fewer credit hours. Other information beyond the number of course credit hours could be important in estimating the expected value of increased credit hours on transfer rates. This information can be summarized in the vector of student characteristics, $\boldsymbol{x}$.

The mean impact of treatment on the treated (TT) is typically the parameter of interest to be studied in evaluation literature. It estimates the effect of the treatment on those who received it, relative to what the outcome for those individuals would have been had they not received the treatment (Smith and Todd 2001). The effect of treatment on the treated can be estimated by:

$$\begin{aligned} TT &= E(\Delta|x, z = 1) = E(y_1 - y_0|x, z = 1) \\ &= E(y_1|x, z = 1) - E(y_0|x, z = 1) \end{aligned}$$

In this study, information about the average outcomes among the treated $E(y_1|\boldsymbol{x}, z = 1)$ is readily available—we can access information about transfer outcomes for every student who did in fact take the higher number of credit hours. However, information about the counterfactual outcome, $E(y_0|\boldsymbol{x}, z = 1)$ is not available. In a randomized study, this outcome information is contained within the control group, who were randomly assigned to not receive the treatment without regard to their characteristics $\boldsymbol{x}$ (Heckman et al. 1998a).

Given this problem with an observational data set, several possible approaches are possible.[2] Economists have relied on instrumental variables approaches in many causes, while psychologists and sociologists have similarly utilized simultaneous equation modeling. Other approaches include Heckman's model for accounting for selection into treatment or control groups (Heckman 1979). All of these models require quite strong assumptions regarding the parameters being estimated.

Natural experiments do not require strong assumptions for the purposes of identification. For instance, Fortin et al. (2004) utilize changes in the provision of welfare benefits in Quebec in order to estimate the effect of higher levels of benefits on the duration of welfare spells. However, in the current context, no such exogenous change in credit-taking occurred among students. However, future studies could exploit such a shift, possibly as a result of state policy, and estimate the effect of increased credit hours in the form of a natural experiment.

In addition to the more standard econometric approaches, alternative semi-parametric and non-parametric approaches have been proposed. One alternative method of estimating treatment effects in the current context would be the regression discontinuity design, which estimates the effect of assignment into treatment or control groups around either a sharp

[2] See Heckman and Robb (1985) for a complete discussion.

discontinuity (many times on the basis of an administrative rule) or several discontinuities in the treatment status (known as the "fuzzy" discontinuity design) (Hahn et al. 2001; Lemieux and Milligan 2006)

However, another, non-parametric approach is available: matching (Rubin 1974, 1976). As described above, matching overcomes the issue of selection bias by creating an equivalent control group for the treatment group (Dehejia and Wahba 1999, 2002; Morgan and Harding 2006; Vinha 2002). In education research, matching has been utilized in a variety of circumstances. For instance, Agodini and Dynarski (2004) estimate the impact of participation in the School Dropout Demonstration Assistance Program, comparing their matching estimators with results obtained from a true experimental design.

Using Matching Estimators to Eliminate Selection Bias

The use of matching to recover causal inferences from observational data is not new, going back to Rubin (1974) and even prior. However, its use has increased in the last decade or so. Matching seeks to overcome the fundamental problem of causal inference by identifying a region of common support for treatment and control groups, and estimating treatment effects only within this common range.

To accomplish this, the matching estimator α is identified by comparing the outcomes for the "treatment" group with the outcomes for the "control" group, given a common propensity for selecting into the treatment group, p (Smith and Todd 2001):

$$\begin{aligned}\alpha &= E(y_1 - y_0|z = 1)\\ &= E(y_1|z = 1) - E_{p|z=1}(y|z = 1, p)\\ &= E(y_1|z = 1) - E_{p|z=1}E_y(y|z = 0, p)\end{aligned}$$

The probability p, is defined as (Smith and Todd 2001):

$$\Pr(z = 1|x) < 1 \text{ for all } x$$

Given the condition above holds, and that there no values of $\boldsymbol{x}$ for which the probability of selection into the treatment group is 0, the analysis can proceed, and an estimate of the treatment effect can be calculated. A matching estimator for α can be defined as (Smith and Todd 2001):

$$\alpha_M = \frac{1}{N}\sum_{i \in I_i \cap s_p} [y_{1i} - \widehat{E}(y_{0i}|D = 1, p_i)]$$

In the above equation, $\hat{E}(Y_{0i}|D = 1, P_i)$ represents the matched outcome, and is equivalent to $\sum_{j \in I0} W(i,j)Y_0$. I_1 is the "treatment" group, or those who had high levels of credit hours. I_0 is the "control" group, or those who had low levels of credit hours. S_p represents the region of common support between the two groups.

The matching procedure finds a match for every treated unit i $I_1 \cap S_p$ among the control units, with a weighted average taken over all of the subclasses that are formed as a result of the matching procedure. Each comparison between treated and control units is made only when all of the units in the subclass have identical values for all of the selection variables. The effect estimated in each subclass is then aggregated into a weighted effect based on the total number of units in each subclass relative to the overall region of common support S_p.

There are several other options available to the analyst in creating matching treatment and control groups. Dehejia and Wahba (2002) describe the above method as follows.

One way to estimate this equation would be by matching units on their vector of covariates X_i. In principle, we could stratify the data into subgroups (or bins), each defined by a particular value of X; within each bin, this amounts to conditioning on X. The limitation of this method is that it relies on a sufficiently rich comparison group so that no bin containing a treated unit is without a comparison unit (Dehejia and Wahba 2002, p. 153).

To overcome the dimensionality issue described by Dehejia and Wahba (2002), many studies make use of propensity score matching. This method involves estimating the probability that $z = 1$ via a standard model for binary variables such as logit or probit models. The predicted value $p = pr(z = 1|\mathbf{x})$ is then used as an estimate of p as described above, and treatment and control groups are matched on the basis of their similarity for the values of $\mathbf{x}$. Many studies have demonstrated that, given the conditional independence assumption described below holds, that propensity score matching results in unbiased estimates of treatment effects (Dehejia and Wahba 2002; Vinha 2002; Smith and Todd 2001; Rubin and Neal 2000; D'Agostino 1998; Rosenbaum and Rubin 1983, 1985).

However, this study does not suffer from issues with dimensionality. Instead, given the very large sample size ($n > 10{,}000$) and the relatively small number of independent variables, utilizing direct matching is an effective method for identifying equivalent treatment and control groups. This is the same method utilized by Angrist (1998) and Card and Sullivan (1988) and discussed by Angrist and Krueger (1998) and Rubin (1976).

The bias in this estimator depends critically on whether there is enough information in $\boldsymbol{x}$ to support the analysis (Heckman et al. 1998a, b). To the extent possible, $\boldsymbol{x}$ must accurately reflect the decision on whether or not to undertake the treatment. If this information is sufficient, then the conditional independence assumption will hold:

$$(y_0 \perp z)|x$$

This assumption states that the information in $\boldsymbol{x}$ is sufficient to assume independence of the outcome y_0 and the treatment, z.

If this assumption holds, then the fundamental problem of causal inference can be overcome using matching. As stated, this depends critically on the availability of information that predicts whether or not individuals will choose the treatment or control groups. If this assumption is not met, then any inferences made as a result of the analysis will be biased in the same way that inferences from other methods are biased (Heckman et al. 1998a).

Data and Methods

This study utilizes data from the fall enrollment surveys taken on every public campus in the state of Tennessee for the years 1995 through 2004. In each year, a total of about 25,000 students enrolled in community colleges in the state. Each of these students is uniquely identified in this dataset.

The subset of the data used in this analysis concerns only those students who enrolled for regular credit hours in the years 1995 through 1998 as first time freshmen in community colleges in Tennessee. Descriptive statistics for each of the variables for each of the years in the dataset can be found in Table 1.

Table 1 Descriptive statistics for variables in analysis

	Mean/Proportion			
	1995	1996	1997	1998
No transfer	0.77	0.78	0.78	0.78
Transfer	0.23	0.22	0.22	0.22
Hours	7.7 (4.69)	7.84 (4.73)	8 (4.67)	7.93 (4.64)
Race: Asian/Pacific Islander	0.01	0.01	0.01	0.01
Race: black	0.07	0.07	0.08	0.08
Race: Hispanic	0.01	0.01	0.01	0.01
Race: other	0.05	0.01	0.01	0.01
Race: white	0.87	0.9	0.9	0.89
Sex: female	0.59	0.57	0.58	0.58
Sex: male	0.41	0.43	0.42	0.42
Age	22.17 (6.86)	21.87 (6.56)	22.13 (6.86)	22 (6.81)
Miles traveled	28.07 (38.88)	26.87 (37.26)	27.54 (38.88)	27.93 (39.15)
Major: undeclared	0.65	0.66	0.66	0.68
Resident	0.93	0.93	0.93	0.93
Population	10,774	11,371	11,957	11,664

Dependent Variable

The dependent variable in this study is a binary variable indicating whether a first-time student attending a community college took courses for credit in a public 4 year college in Tennessee at any point within 6 years of initial enrollment. This measure of the transfer rate meets the definition of students who had ever transferred used in Bradburn and Hurst 2001. This is a very open definition of transfer, as opposed to a more limited definition that would only include those students whose first transfer institution was a 4-year college, and who made the transition to a 4 year college within 2 years of initial enrollment. This variable is also limited in that it can only track transfer rates between public institutions, thereby omitting any students who may have transferred to a private institution in the state (private institutions enroll approximately 20% of students in the state) (U.S. Department of Education 2004).

I use this more open definition of transfer in order to capture every student who made the transition at one point or another from the community colleges to the 4 year institutions within the state, without regard for the timing or sequence of transfer patterns. Table 2 shows the proportion of students who ever transferred by various student characteristics. As expected, transfer rates are highest among those students who take more than 12 credit hours, with 33% of these students eventually transferring between 1998 and 2004 as compared to 18% of students who took fewer than 12 credit hours.

Independent Variable of Interest

The independent variable of interest for this study is the number of hours that the student takes in the first term of enrollment. As noted in the previous section on the counterfactual framework for estimation, a naive estimator of the causal effect of this variable on transfer

Table 2 Proportion of students who ever transferred

	1995	1996	1997	1998
Hours: <12	0.19	0.17	0.17	0.18
Hours: more than 12	0.34	0.35	0.35	0.33
Race: Asian/Pacific Islander	0.34	0.32	0.33	0.26
Race: black	0.16	0.15	0.16	0.17
Race: Hispanic	0.11	0.18	0.23	0.16
Race: other	0.23	0.18	0.17	0.25
Race: white	0.23	0.22	0.23	0.23
Sex: female	0.22	0.2	0.2	0.2
Sex: male	0.25	0.25	0.25	0.25
Age: <18	0.29	0.28	0.29	0.28
Age: 18–24	0.24	0.22	0.22	0.23
Age: 25–44	0.11	0.1	0.09	0.11
Age: 45–64	0.06	0.05	0.06	0.05
Age: 65+	0		0	0
Miles: 0–9	0.23	0.22	0.22	0.22
Miles: 10–19	0.22	0.22	0.22	0.22
Miles: 20–49	0.25	0.24	0.23	0.25
Miles: 50–99	0.26	0.24	0.22	0.26
Miles: 100+	0.12	0.14	0.1	0.1
Major: undeclared	0.25	0.22	0.22	0.23
Resident	0.22	0.21	0.21	0.22

rates would be biased because students self-select into higher or lower numbers of credit hours according to both observed and unobserved characteristics. This can be seen by inspecting the conditional distribution of this variable according to other characteristics of students. The distribution of student credit hours by institutional type is displayed in Fig. 1. As the figure shows, the median level of transfer rates, as well as the shape of the distribution, differs from one institution to another. For instance, at Chattanooga State, the average number of credit hours in the first semester of 1995 was 7.1, with a median of six, while at Motlow State, the average number of credit hours in the first semester was 8.4, with a median of 9. In the former case, the distribution is mostly right skewed, while in the latter, the distribution is left skewed.

The conditional distribution of credit hours taken also differs by race. Figure 2 shows the differing distribution of credit hours taken by students of different races in 1995. All of the distributions shown are more or less right-skewed, and the median number of hours taken is less than the average in all cases, indicating a situation where many students take few credit hours, while smaller number enroll for higher number of credit hours. The distribution for African-American and Hispanic students indicates lower overall credit hours taken with more skew. The average number of credit hours attempted in the first semester by African-American students in 1995 was 6.0, while among white students the average was 7.8. Among Hispanic students the average was actually highest, at 7.9, although these students constituted a small percent of the total enrollment at that time.

The matching techniques described in the next section are designed to account for the differing conditional distribution of the covariates to account for the selection bias that

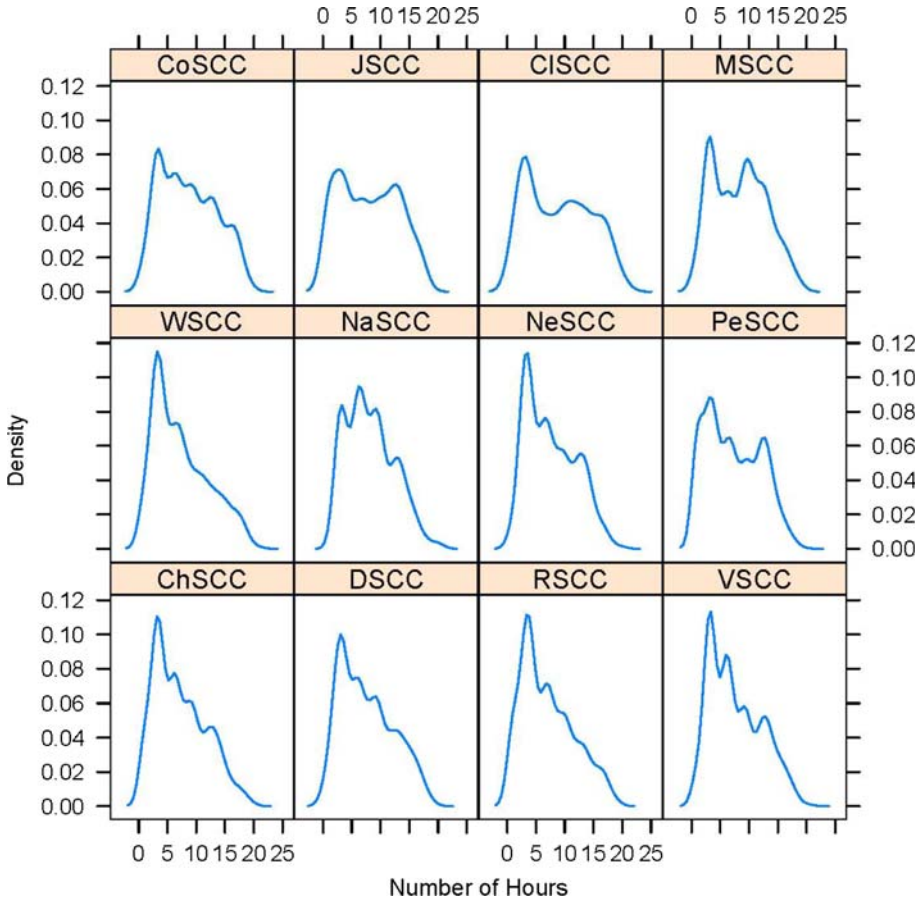

Fig. 1 Conditional distribution of credit hours by community college, 1995. *Note*: JSCC = Jackson State, NaSCC = Nashville State, ClSCC = Columbia State, MSCC = Motlow State, WSCC = Walters State, NeSCC Northeast State, PeSCC = Pellissippi State, CoSCC = Columbia State, ChSCC = Chattanooga State, DSCC = Dyersgurg State, RSCC = Roane State, VSCC = Volunteer State

would occur given the differing characteristics of students who take higher or lower numbers of credit hours in their first semester.

To make use of matching as a method for recovering the causal impact of increased credit hours, the number of credit hours must be recoded as a binary variable. I use three different possible definitions of academic intensity in this study. First, I use a definition that is based on 6 or more hours for students who take course credits as first time freshmen. The next definition uses 9 as the cutoff. Finally, following Adelman and others, I recode this variable to be 1 for students who take 12 or more hours of course credits as first time freshmen, and 0 for students who take fewer course credits in their first term. These variables will be considered the treatment variables for the purposes of this study, with those who took course credits above the specified cutoffs to be the "treatment" group and those who took fewer than the specified cutoff level of credit hours considered to be the "control" group.

I specify three different binary variables as the treatment variable in order to ensure that the study's findings are not influenced by a particular cutoff point, but instead show how

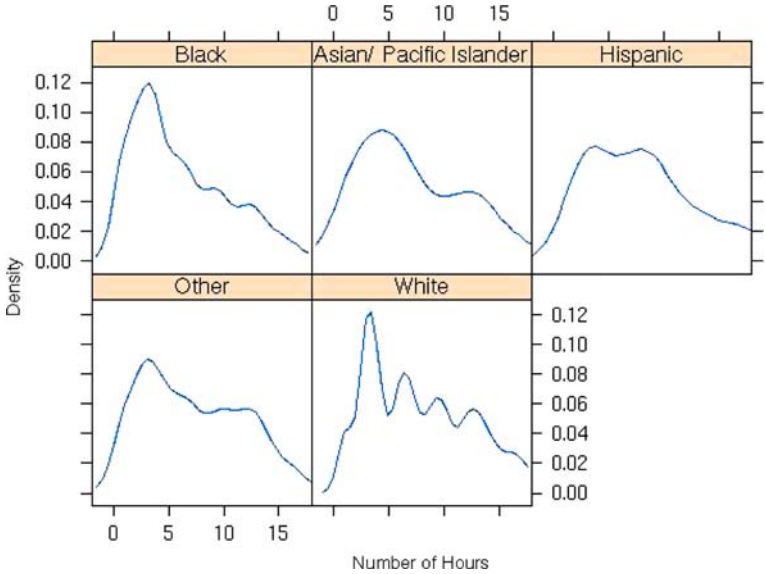

Fig. 2 Conditional distribution of credit hours by race/ethnicity, 1995

student transfer changes as students increase their coursetaking over the first semester of community college.

Selection Variables

As mentioned previously, the primary purpose of this analysis is to recover causal estimates of the effect of the treatment variable (6, 9, or 12 or more credit hours in the first semester of enrollment) on transfer rates. Any naive estimate of the effect of this variable would fail to take into account the fact that students self select into the treatment and control groups to account for this self selection, I use the richest possible set of selection variables from the data to account for the reasons students might self-select into the treatment or control groups.

First, following much of the previously cited research on collegiate transfer rates, I include both race and sex as selection variables (Lee et al. 1993; Lee and Frank 1990; Surette 2001). Race is coded into five categories: Black, Hispanic, White, and Other. Marginals for this variable are shown in Table 1.

While much of the available literature does not discuss the age of students, inspection of the data from Tennessee shows that older students are less likely to take higher numbers of credit hours, and are less likely to transfer from 2-year to 4 year institutions. I therefore include age of the student at matriculation into the community college as a selection variable.

I also utilize student residency to explain selection into the high- or low-credit hour groups. It may be more likely that students who move from out of state will be more likely to take courses full time. However, for most of the years in the analysis the vast majority of first time freshmen in community colleges are state residents.

I use the categorization of major to specify selection into treatment or control groups. It seems likely that students who enter higher education with a clear academic goal will be

Table 3 *T*-test for differences of means for students with course credits above and below 12 in the first semester

	1995		1996		1997		1998	
	Full	Matched	Full	Matched	Full	Matched	Full	Matched
Race: Asian/Pacific Islander	0.47	−0.16	−1.5		3.03	0.2	0.82	−0.73
Race: black	6.91	1.02	9.44	1.7	9.29	1.62	7.88	0.76
Race: Hispanic	0.32	−0.64	−0.36	−0.88	−0.17	−0.35	0.39	−0.99
Race: other	0.62	0.69	1.06	−0.31	−0.45	−0.94	2.5	−0.26
Race: white	−5.48	−1.05	−7.77	−1.15	−8.41	−1.13	−7.96	−0.18
Sex: male	−7.03	−1.89	−6.6	−1.85	−6.01	−2.9	−4.61	−2.12
Age	24.26	−0.77	21.36	0.74	23.61	0.8	22.23	0.74
Miles traveled	−2.39	−2.21	−4	−2.2	−4.09	−1.26	−3.11	0.09
Major: undeclared	−2.39	−1.53	−3	−3.25	−1.67	−2	−2.24	−1.69

T statistics are reported for differences in proportion or mean for each covariate and treatment status (hours >12)

more likely to both take more units and transfer in the future. As Table 1 shows, however, most students are classified as undeclared, meaning that this sole academic variable is not highly informative about transfer behavior.

Many studies have shown that student choice of higher education is influenced by the distance to the college that they must travel (Kane and Rouse 1995). It could also be the case that students who travel further are more likely to take more credits and transfer. I therefore use distance traveled as another selection variable.[3]

Last, I use the student's institution as a predictor of transfer rates. As many authors have shown, institutions differ critically with respect to their transfer rates (Shaw and London 2001; Wassmer et al. 2004). This holds true in Tennessee as well, where institutions differ dramatically in the percent of students who transfer from different community colleges to 4-year institutions.

Table 3 shows the results of *t*-tests for the differences of means for the treatment and control group and each of the above described selection variables in each year of the analysis. As the table shows, the large differences that are apparent in the unmatched samples are not evident in the matched sample. For instance, the T statistic for age is highly statistically significant in every year, indicating that students who take more credits are likely to be older. However, after matching, the statistic for differences in age is not significant, and is quite close to 0 in each year. Overall, the matching procedure provides a reasonably good balance between the treatment and control groups.

Matching Techniques

In this study I make use of a matching procedure that matches every treated unit only with those individuals who have precisely the same values on all of the selection variables. For instance, a black female student who is 22 years and a state resident and who enrolls at

[3] Distance traveled is measured as the distance from the student's county of origin to the institution's zip code area.

Motlow State Community College for more than 12 h per semester in her first year in college is compared only with those students who share all of those characteristics with the exception of the number of credit hours. This process is repeated for every single possible combination of covariate values in the dataset. If no control units can be found for a given treatment unit, then the treatment is discarded. Estimation was completed using the Match It library for the R statistical computing environment, as described in Ho et al. (2005).

As would be expected even given the small number of variables available, the number of subclasses produced by such a procedure is relatively large. In fact, the number of subclasses for just the twelve credit hour treatment from each year ranges from 467 in 1995 to 509 in 1998.

The treatment effect of taking more than 6, 9, or 12 credit hours is estimated via logistic regression.[4] In the case of the full data set, no weights are utilized, while in the case of the matched datasets, weights are based on the probability that a student is in a given subclass identified via the matching procedure.

Specifically, the weighting procedure works as follows. First, all treated students are given a weight of 1. The weights for matched control units are given by n_{ti}/n_{ci}, with n_{ti} representing the number of treatment units in the subclass, which is divided by the number of control units in the subclass, n_{ci}. The effect of the weighting within each subclassification group is to ensure that the control units as a group are given no more weight than the treatment units (Ho et al. 2005).

Results

The results indicate that there is a positive impact of increased credit hours on transfer rates, even after exactly matching on the selection variables described previously. However, this impact is less than would be expected compared to estimates from the full sample.

Tables 4–6 show the results from a logistic regression of transfer rates on the binary variable dividing students into more than the cutoff and less than the cutoff number of credit hours.

As Table 4 shows, students who took more than six credit hours are more likely to transfer than their peers. In 1998, the coefficient for the full sample for the binary variable of 6 h or more is 0.62, with a 95% confidence interval from 0.52 to 0.72 5. This indicates an increased probability of transfer of about 10%. In the matched sample, the predicted increase is smaller—about 8%. As Fig. 3 shows, a similar pattern holds across all years. There is a slight difference between those who take more than six and those who take fewer than six credit hours. Matching estimators are consistent with the full sample, but indicate an even smaller difference between these two groups.

Students who take more than nine credits are also more likely to transfer than their peers who take less than that amount. As Table 5 shows, the coefficient for more than nine credit hours is positive and significant for all years. This result is also apparent in Fig. 4, which shows the predicted difference in transfer rates between students who took fewer than nine credits and those who took more than nine. As with the results based on taking more than six credits, the effect is lower in the matched sample, and hovers around 10% for all years.

[4] Logistic regression is utilized for ease of comparison between these results and other studies of transfer. Results for simple z-tests for differences of proportions are identical.

Table 4 Maximum likelihood estimates from logistic regression, full and matched samples, all years treatment = 6 h or more

	1995		1996		1997		1998	
	Full	Matched	Full	Matched	Full	Matched	Full	Matched
Intercept	−1.64 (0.04)	−1.4 (0.05)	−1.79 (0.04)	−1.32 (0.05)	−1.79 (0.04)	−1.43 (0.05)	−1.67 (0.04)	−1.31 (0.05)
Hours: more than 6	0.64 (0.05)	0.54 (0.07)	0.75 (0.05)	0.44 (0.06)	0.75 (0.05)	0.57 (0.06)	0.62 (0.05)	0.44 (0.06)
Null log likelihood	11581.48	6616.63	11925.42	7591.06	12624.38	7409.93	12414.35	7636.16
Log likelihood	11415.96	6545.99	11699.6	7535.93	12391.88	7322.74	12251.76	7582.64
N	10,774	5,769	11,371	6,601	11,957	6,465	11,664	6,603

Table 5 Maximum likelihood estimates from logistic regression, full and matched samples, all years. Treatment = 9 h or more

	1995		1996		1997		1998	
	Full	Matched	Full	Matched	Full	Matched	Full	Matched
Intercept	−1.56 (0.03)	−1.34 (0.05)	−1.71 (0.03)	−1.33 (0.04)	−1.7 (0.03)	−1.38 (0.05)	−1.61 (0.03)	−1.3 (0.04)
Hours: more than 9	0.73 (0.05)	0.61 (0.06)	0.88 (0.05)	0.61 (0.06)	0.85 (0.05)	0.67 (0.06)	0.74 (0.05)	0.57 (0.06)
Null log likelihood	11581.48	6608.4	11925.42	7371.33	12624.38	7363.4	12414.35	7572.64
Log likelihood	11333.74	6507.19	11558.4	7259	12261.28	7228.94	12139.18	7469.55
N	10,774	5,707	11,371	6,323	11,957	6,333	11,664	6,483

Table 6 Maximum likelihood estimates from logistic regression, full and matched samples, all years. Treatment = 12 h or more

	1995		1996		1997		1998	
	Full	Matched	Full	Matched	Full	Matched	Full	Matched
Intercept	−1.45 (0.03)	−1.23 (0.04)	−1.58 (0.03)	−1.18 (0.04)	−1.6 (0.03)	−1.26 (0.04)	−1.5 (0.03)	−1.13 (0.04)
Hours: more than 12	0.81 (0.05)	0.67 (0.06)	0.95 (0.05)	0.65 (0.06)	1 (0.05)	0.78 (0.06)	0.82 (0.05)	0.55 (0.06)
Null log likelihood	11581.48	5919.38	11925.42	6856.64	12624.38	6826.83	12414.35	6872.39
Log likelihood	11321.25	5810.13	11542.01	6739.97	12166.88	6653.48	12118.81	6785.5
N	10,774	5,014	11,371	5,741	11,957	5,714	11,664	5,714

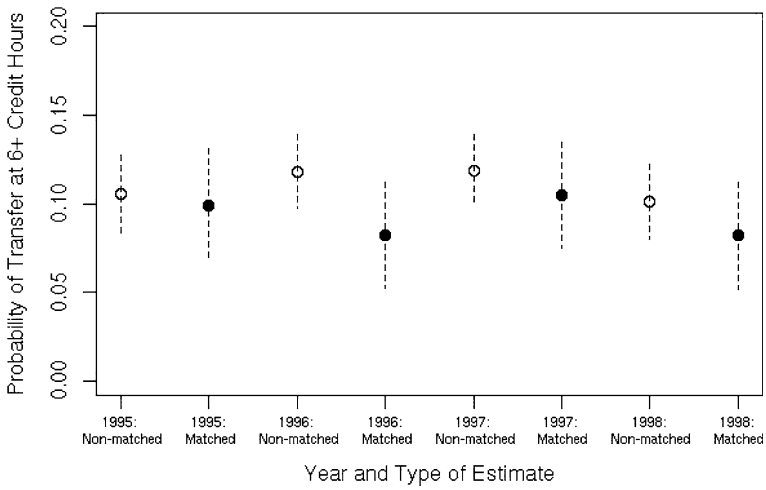

Fig. 3 Predicted difference between high and low course credits, full and matched samples, treatment = 6 h

The largest differences in subsequent transfer rates are found between students who take 12 or more hours and students who take fewer. Students who take 12 credits or more are much more likely to transfer, both in the matched and unmatched sample. Table 6 shows results from a logistic regression of transfer rates on a binary variable for students taking 12 h of course credits in their first year. Results for each year and for each sample are shown.

In 1995, the estimated coefficient for the full sample is 0.81, with a 95% confidence interval from 0.71 to 0.91. This translates into a predicted difference in transfer rates from low course credit students of about 15%, with a 95% confidence interval from 13% to 17%. Similarly for the matched sample, taking a higher number of course credits is associated with a higher predicted probability of transfer, but the predicted difference from the control group is not as large. Instead, the 95% confidence interval on the coefficient for higher course credits in the matched sample is bounded by [0.43, 0.67], with a midpoint at 0.55. This translates into a predicted increase in probability of transfer of about 11%, with a confidence interval from 9% to 14%.

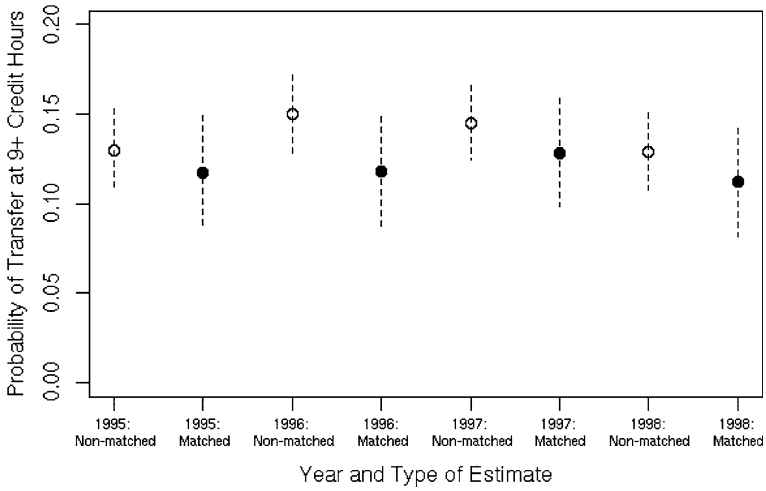

Fig. 4 Predicted difference between high and low course credits, full and matched samples, treatment = 9 h

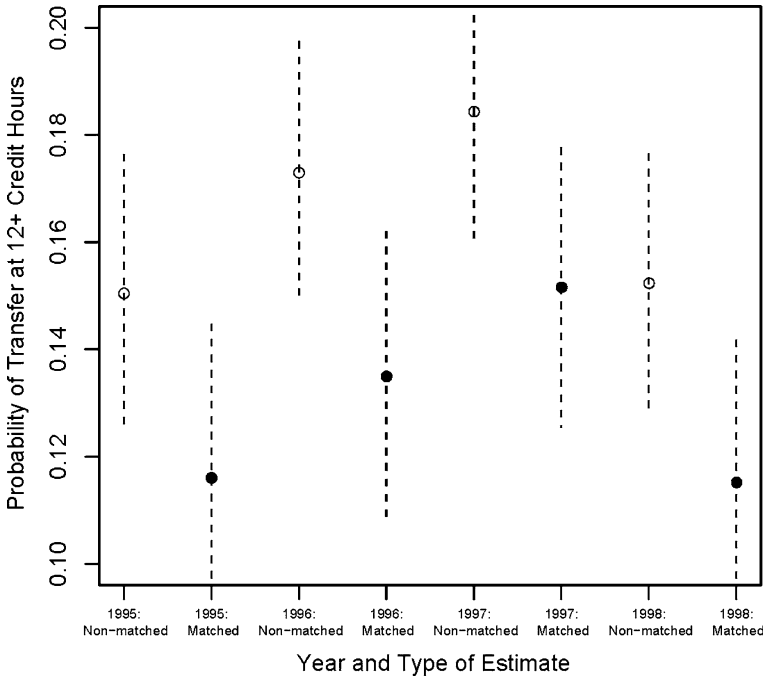

Fig. 5 Predicted difference between high and low course credits, full and matched samples, treatment = 12 h

A similar pattern holds in all of the years of the analysis, as shown in Fig. 5, which shows point estimates and 95% confidence intervals for the impact of increased academic intensity on transfer rates. In every year, the predicted increase in transfer rates is positive for the full and matched samples, but lower in the matched sample. The predicted increase

in probability of transfer in each year for the full sample range from around 15% to as high as 18%, while the range for the matched samples is from 11% to 15%.

From these results, it appears that there is an influence of increased credit hours on transfer rates for the students in Tennessee public institutions. This impact is conditioned on precisely the same covariates for race, sex, age, student residency, major, miles traveled, and institutional attendance for all students. The exact matching procedure used ensures that the control and treatment groups do not differ on any of these covariates. The next section will consider the limitations and implications of this finding.

Conclusions

This study has several limitations. First the data have several important limitations. The data only concern the population of students enrolled in Tennessee public institutions. This may limit the generalizability of the results to other states. The data itself is limited in many ways, particularly with regard to important student background behaviors that would help to understand student self-selection into higher or lower numbers of credit hours. In particular, data on student income, student high school performance, and whether students are first generation college goers would help enormously in predicting whether they will take more or fewer credit hours upon matriculation in the community colleges.

Second, the matching estimator described earlier is not the simple equivalent to a randomized study for the purposes of understanding the causal impact of a policy intervention. Instead, the matching estimator can only be said to be a measure of true causal impact when certain assumptions are met.

Heckman et al. (1997) review the essential features necessary to reduce bias in evaluation studies and summarize them as: (1) the treatment and control groups have identically distributed unobserved attributes; (2) observed attributes are identically distributed; (3) both treatment and control groups are given the same questionnaire; and (4) treatment and control groups are in the same economic (or social) environment. Heckman et al. (1997) suggest that features (2–4) are much more important than has been previously suggested, and feature (1), selection bias, "is a *relatively* small part of bias as conventionally measured" (Heckman et al. 1997, p. 606).

Features (2) and (3) are met quite easily in this study: exact matching by definition ensures that treatment and control groups have precisely the same observed characteristics, and all studied individuals completed a common questionnaire. Feature (4) common environments, is also met, since all students under study are in a single state. Feature (1) should be minimized, but as always with observational studies, there can be no guarantee that treatments and controls do not have different distributions of unobserved characteristics.

Even given these limitations, this study does further the idea that increased academic intensity does causally impact transfer rates. Using an exact matching procedure that paired high course credit students with identical background and institutional characteristics with those who had low numbers of course credits, I find a positive and substantively meaningful increase in transfer rates.

Given this finding, more policy action on this issue seems warranted. Larger scale experimentation with policies that provide both support for students to take more class credits, and incentives to encourage students to do so, would provide a basis for better understanding and implementation of policies across all states that would benefit students.

Policies that support more course credit taking should include both academic and non-academic support. Academic support should involve communication of placement standards and general requirements once students are enrolled in community college. In addition specific targeting of marginal students and the creation of individualized study plans in concert with counselors or faculty members could provide additional momentum for many students who could succeed in postsecondary education with a very minimal additional institutional investment. Many of the types of reforms undertaken by institutions described by Barefoot et al. (2005) may aid in helping community college students maintain the higher course loads that can lead to increased transfer rates.

Non-academic support should also be in place to provide students with the types of help that they need in order to arrive on campus ready to take additional credits. This should include additional need-based financial aid to reduce the need to work, transportation assistance and child care, among others. The essence of the support structure should be a holistic learning plan that enables the students to take additional course credits and build momentum toward the postsecondary degree.

Last, incentives should also be a part of the package for encouraging course credits. Institutions should consider targeted incentives such as "buy three, get one free" plans that provide a three credit course for free, given the student pays for the first nine credits. Other incentive plans could include tuition remittances for 20 credits of course completion in the first year, or financial aid support targeted to help students overcome barriers that stand in the way of full-time attendance.

Acknowledgements The author gratefully acknowledges the staff at the Tennesse Higher Education Commission, who provided the data for this study. Linda Sax, John Chesslock, and Jennifer Delaney all provided detailed comments and suggestions that improved the paper. Last, the author would like to thank the editor and two anonymous reviewers for their thoughtful and constructive comments. The author bears sole responsibility for the content of this paper.

References

Adelman, C. (1999). Answers in the tool box: Academic intensity, attendance patterns, and bachelor's degree attainment. Technical report, United States Department of Education, Washington, DC.

Adelman, C. (2004). Principal indicators of student academic histories in postsecondary education. Technical report, United States Department of Education.

Adelman, C. (2006). The toolbox revisited: Paths to degree completion from high school through college. Technical report, United States Department of Education, Washington, DC.

Agodini, R., & Dynarski, M. (2004). Are experiments the only option? A look at dropout prevention programs. *The Review of Economics and Statistics, 86*(1), 180–194.

Anderson, G. M., Alfonso, M., & Sun, J. C. (2006). Rethinking cooling out at public community colleges: An examination of fiscal and demographic trends in higher education and the rise of statewide articulation agreements. *Teachers College Record, 108*(3), 422–451.

Angrist, J. D. (1998). Estimating the labor market impact of voluntary military service using social security data on military applicants. *Econometrica, 66*(2), 249–288.

Angrist, J. & Krueger, A. (1998). Empirical strategies in labor economics. Working Papers. Princeton Department of Economics-Industrial Relations Section.

Bailey, T., & Weininger, E. B. (2002). Performance, graduation, and transfer of immigrants and natives in City University of New York community colleges. *Educational Evaluation and Policy Analysis, 24*(4), 359–377.

Barefoot, B. O., Gardner, J. N., Cutright, M., Morris, L. V., Schroeder, C. C., Schwartz, S. W., et al. (2005). *Achieving and sustaining institutional excellence for the first year of college*. San Francisco: Jossey-Bass.

Bauer, P. F., & Bauer, K. W. (1994). The community-college as an academic bridge—academic and personal concerns of community-college students before and after transferring to a 4-year institution. *College and University, 69*(3), 116–122.

Bradburn, E. M., & Hurst, D. G. (2001). Community college transfer rates to 4-year institutions using alternative definitions of transfer. Technical report, United States Department of Education, National Center for Education Statistics, Washington, DC.

Braxton, J., Milem, J. F., & Sullivan, A. S. (2000). The influence of active learning on the college student departure process. *Journal of Higher Education, 71*(5), 569–590.

Card, D., & Sullivan, D. G. (1988). Measuring the effect of subsidized training programs on movements in and out of employment. *Econometrica, 56*(3), 497–530.

Cheslock, J. J. (2005). Differences between public and private institutions of higher education in the enrollment of transfer students. *Economics of Education Review, 24*(3), 263–274.

D'Agostino, R. B. (1998). Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine, 17*, 2265–2281.

Dehejia, R. H., & Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluation the evaluation of training programs. *Journal of the American Statistical Association, 94*(448), 1053–1062.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics, 84*(1), 151–161.

Dougherty, K. J., & Kienzl, G. S. (2006). It's not enough to get through the open door: Inequalities by social background in transfer from community colleges to 4-year colleges. *Teachers College Record, 108*(3), 452–487.

Fortin, B., Lacroix, G., & Drolet, S. (2004). Welfare benefits and the duration of welfare spells: Evidence from a natural experiment in Canada. *Journal of Public Economics, 88*(7–8), 1495–1520.

Grubb, N. W. (1991). The decline of community college transfer rates: Evidence from national longitudinal surveys. *The Journal of Higher Education, 62*(2), 194–222.

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica, 69*(1), 201–209.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica, 46*, 931–961.

Heckman, J., Ichimura, H., Smith, J., & Todd, P. (1998a). Characterizing selection bias using experimental data. *Econometrica, 66*(5), 1017–1098.

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training program. *The Review of Economic Studies, 64*(4), 605–654.

Heckman, J. J., Ichimura, H., & Todd, P. E. (1998b). Matching as an econometric evaluation estimator. *The Review of Economic Studies, 65*(2), 261–294.

Heckman, J., & Robb, R. (1985). *Longitudinal analysis of labor market data, chapter alternative methods for evaluating the impact of interventions* (pp. 156–245). New York: Cambridge University Press.

Hilmer, M. J. (1997). Does community college attendance provide a strategic path to a higher quality education? *Economics of Education Review, 16*(1), 59–68.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2005). *Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference*. Cambridge: Harvard University.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81*(396), 945–960.

Kane, T. (1999). Has financial aid policy succeeded in ensuring access to college? In *The price of admission*. Washington, DC: Brookings.

Kane, T. J., & Rouse, C. E. (1995). Labor-market returns to 2-year and 4-year college. *American Economic Review, 85*(3), 600–614.

Lee, V. E., & Frank, K. A. (1990). Students characteristics that facilitate the transfer from 2-year to 4-year colleges. *Sociology of Education, 63*(3), 178–193.

Lee, V. E., Mackielewis, C., & Marks, H. M. (1993). Persistence to the baccalaureate degree for students who transfer from community-college. *American Journal of Education, 102*(1), 80–114.

Leigh, D. E., & Gill, A. M. (2003). Do community colleges really divert students from earning bachelor's degrees? *Economics of Education Review, 22*(1), 23–30.

Leigh, D. E., & Gill, A. M. (2004). The effect of community colleges on changing students' educational aspirations. *Economics of Education Review, 23*(1), 95–102.

Lemieux, T., & Milligan, K. (2006). Incentive effects of social assistance: A regression discontinuity approach. Analytical Studies Branch Research Paper Series 2006280e, Statistics Canada, Analytical Studies Branch.

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

Morgan, S. L., & Harding, D. J. (2006). Matching estimators of causal effects: Prospects and pitfalls in theory and practice. *Sociological Methods and Research, 35*(1), 3–60.

Prager, C. (1993). Transfer and articulation within colleges and universities. *Journal of Higher Education, 64*(5), 539–554.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41–55.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician, 39*, 33–38.

Rouse, C. E. (1995). Democratization or diversion—the effect of community-colleges on educational-attainment. *Journal of Business & Economic Statistics, 13*(2), 217–224.

Rouse, C. E. (1998). Do 2-year colleges increase overall educational attainment? Evidence from the states. *Journal of Policy Analysis And Management, 17*(4), 595–620.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology, 66*, 688–701.

Rubin, D. B. (1976). Assignment of treatment group on the basis of a covariate. *Journal of Educational Statistics, 2*, 1–26.

Rubin, D., & Neal, T. (2000). Combining propensity score matching with additional adjustments for prognostic covariates. *Journal of the American Statistical Association, 95*(450), 573–585.

Shaw, K. M., & Coleman, A. B. (2000). Humble on sundays: Family, friends, and faculty in the upward mobility experiences of African American females. *Anthropology & Education Quarterly, 31*(4), 449–470.

Shaw, K. M., & London, H. B. (2001). Culture and ideology in keeping transfer commitment: Three community colleges. *Review of Higher Education, 25*(1), 91.

Shulock, N., & Moore, C. (2005). Diminished access to the baccalaureate for low-income and minority students in california: The impact of budget and capacity constraints on the transfer function. *Educational Policy, 19*(2), 418–442.

Smith, J. A., & Todd, P. E. (2001). Reconciling conflicting evidence on the performance of propensity-score matching methods. *The American Economic Review, 91*(2), 112–118.

Surette, B. J. (2001). Transfer from 2-year to 4-year college: an analysis of gender differences. *Economics of Education Review, 20*(2), 151–163.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 89–125.

Townsend, B. K. (1995). Community-college transfer students—a case-study of survival. *Review of Higher Education, 18*(2), 175–193.

United States Department of Education (2005). Beginning postsecondary students: Data analysis system. Technical report, National Center for Education Statistics.

U.S. Department of Education (2004). Digest of education statistics, 2002. Technical report, National Center for Education Statistics.

Vinha, K. (2002). A primer on propensity score matching estimators. Documents Cede 001928, Universidad de los Andes Cede.

Wassmer, R., Moore, C., & Shulock, N. (2004). Effect of racial/ethnic composition on transfer rates in community colleges: Implications for policy and practice. *Research in Higher Education, 45*(6), 651–672.