Original Manuscript

Educational Evaluation and Policy Analysis
Month 202X, Vol. XX, No. X, pp. 1–26
DOI: 10.3102/01623737231210243
Article reuse guidelines: sagepub.com/journals-permissions

Is Working in College Worth It? How Hours on the Job Affect Postsecondary Outcomes

Avery M. D. Davis
Johns Hopkins University

Many students work during college to offset rising costs, but significant time on the job affects postsecondary outcomes. Analyzing the High School Longitudinal Study (N = 4,418), this article estimates the effects of hours worked on grades, credits earned, persistence, stopping out (i.e., unenrolling for 5 months before reenrolling), and dropping out. The polynomial regression analysis shows that after adjusting for background characteristics, prior academic achievement, institution types, and family obligations, "traditional" undergraduate students begin seeing deleterious effects at 20 hours, which becomes even more severe for those working 28+ hours (and the worst for Pell Grant recipients working long hours). While some work was good for students, on average, financial and family circumstances help explain the curvilinear relationships.

Keywords: *work–study hours, postsecondary performance, higher education, college persistence*

Working through college is common, but empirical studies have not adequately nailed down the thresholds for when jobs help and/or hurt postsecondary outcomes. In the United States, the jobs students take on vary dramatically, from being directly related to one's major, unrelated to one's course of study, explicitly part of a formal work–study program, part-time, full-time, internal (i.e., within a student's university), external (e.g., professional internships), and/or self-employed (e.g., an unreported side hustle). Since work can affect long-term outcomes (Darolia, 2014; DeSimone, 2008; Ecton et al., 2023; Ehrenberg & Sherman, 1987; Scott-Clayton & Minaya, 2016; Stinebrickner & Stinebrickner, 2003), students must know the tradeoffs of spending their time on the job (to help pay for college) so that they can make informed decisions.

How much does work affect academic achievement in the existing literature? Darolia (2014) argues that hours on the job do not affect grades but are significantly and negatively related to credit accumulation. Scott-Clayton and Minaya (2016) suggest that work is good for students, while Ecton et al. (2023) show consistently negative effects across outcomes. The one agreement in the literature is that these relationships and magnitudes differ for subgroups.

The bulk of prior research uses data from before the Great Recession, and the conventional wisdom based on these studies influences policy and practice. These analyses can no longer serve higher education stakeholders because the data are from a time remarkably different from what today's students are facing. In this article, I use the High School Longitudinal Study (HSLS) and third-degree polynomial regression analysis to estimate the effects of hours worked on postsecondary performance and persistence. I find that working up to 19 hours has little to no effect on postsecondary outcomes, but there is a stark and significant turn for the worse for students who work over 20 hours (and this is especially true for those above the 27-hour mark).

Background

Questioning the equity of student employment has been topical for at least a century, as noted in *The New York Times* ("Students Who Fight Their Way Through College," 1907):

> In every college in the land there are men today—average men—by the hundred who are earning their education by the labor of their own hands and brain. Without money to pay their way, they are forced to eke out, by all kinds of expedients, enough to defray the cost of their scholarships and their living expenses. (p. 39)

Even before Harry S. Truman became the only president without a postsecondary degree in 1945, he dropped out of college because he needed full-time work to support his family (Shermer, 2021). These anecdotes notwithstanding, the working college student (in part, as a way to pay for rising tuition rates) has become increasingly visible in the decades leading into the Great Recession: Full-time, "traditional" students' weekly hours worked increased from 6 hours per week in 1970 to 10 hours in 2008 (Scott-Clayton, 2012; see also Baum, 2010).

The Great Recession brought forth significant state divestments/tuition increases, and people turned to loans/private sources to shoulder the burden of these costs that have since outpaced inflation (Barr & Turner, 2013; Newfield, 2016; Webber, 2017). Since the Recession, the availability and character of jobs have changed, and recent generations of students are also more skeptical of working their way through college (Goldrick-Rab, 2016). Figure 1 provides trends for reported hours worked (and the share of students working 20+ hours) from 1990 through 2022 (the last available year of data from the Current Population Survey; Flood et al., 2022). In March 2008, nearly half (47.87%) of full-time college students (aged 16–24) were working before a sharp decline leading to 41.27% in 2013 (the time of this study). Although different from the Recession, the share of students working in 2013 is similar to that of today; in 2022, after another slight drop during the pandemic, 40.70% of students reported working in any job.

It has been argued that if jobs negatively affect postsecondary outcomes, students would be better off taking out loans so that they could finish faster with higher levels of achievement (Scott-Clayton, 2012; Soliz & Long, 2016). Still, credit-constrained students (from low-income backgrounds) often do not have sufficient financial aid, and they may have little "choice" but to work (Baum, 2010; Castleman & Meyer, 2019; Ecton et al., 2023; Goldrick-Rab et al., 2016; Scott-Clayton, 2012). Credit constraints prevent students from borrowing the minimum amount needed to pay tuition and nondiscretionary living expenses, but they can also prevent students from borrowing enough to optimally smooth consumption (i.e., balance between savings and spending, as to not live paycheck-to-paycheck; Lochner & Monge-Naranjo, 2011; Scott-Clayton, 2012). In addition to these factors, personal decisions to work are reflected through students' risk aversion to debt and parents' lack of willingness to pay for college (Avery & Turner, 2012; Boatman et al., 2017; Flaster, 2018; Scott-Clayton, 2012).

Once students "decide" whether to work, they are seemingly unrestricted with the number of hours at off-campus and unofficial jobs, but the most popular cutoff for students' employment intensity (in terms of considering negative effects) in any job is 20 hours per week (Bozick, 2007; Choi, 2018; Dundes & Marx, 2006). This cutoff for intense work–study hours is born out of the literature that examined this topic from the high school stage (D'Amico, 1984; Lee & Staff, 2007; Staff et al., 2010; Steinberg et al., 1982). In addition to scholars using the 20-hour cutoff, universities apply this in practice (e.g., Gallagher, 2022; Miller & Schmidt, 2022). While grounded in empirical evidence as an important threshold, an unwritten position can be argued that capping work at 20 hours guards against paying benefits since many states require employers to do such at 30 hours. Although many colleges limit on-campus work to 20 hours, this is far from universal and difficult to enforce with off-campus jobs. Returning to trends via the Current Population Survey, in 2008, 30.20% of students reported working at least 20+ hours, but this dropped to 25.01% in 2013 with a similar 25.94% in 2022 (see also National Center for Education Statistics [NCES], 2021).

The Federal Work–Study (FWS) program also has no formal limits to the number of hours students can work in combination with their awards (and non-FWS jobs)—the government only encourages people not to repeatedly surpass 40 hours in a single week (Federal Student Aid, 2020). Nonetheless, if FWS provides convenient jobs and money, the program may help alleviate

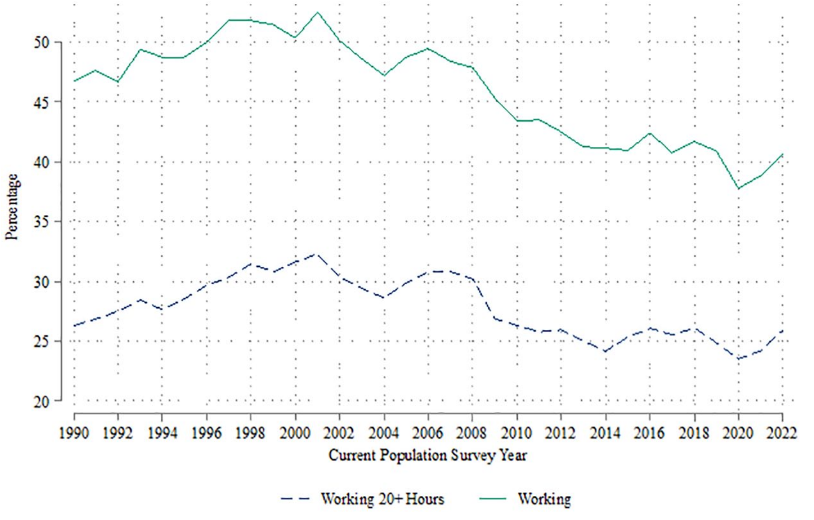

FIGURE 1. *The percentage of students working and working 20+ hours 1990–2022.*
Note. Source = Current Population Survey (Flood et al., 2022). *N* = 151,722. Data are weighted. Survey from March of each year, reporting the total number of hours the respondent (i.e., full-time college students aged 16–24) was at work during the previous week. For employers and the self-employed, this includes all hours spent attending to their operation(s) or enterprise(s). For employees, it is the number of hours they spent at work. For unpaid family workers, it is the number of hours spent doing work directly related to the family business or farm (not including housework).

credit constraints and help students succeed (Baum, 2010; Scott-Clayton, 2011). The overarching allocation of FWS funds is based on a government formula with much beholden to prior campus-based distributions and (less so) institutional need (Congressional Research Service, 2007; Higher Education Act of 1965, as amended in 2021). High-allocation colleges tend to be more established and attract high-achieving students, who are already more economically advantaged and who would perform better anyhow (Scott-Clayton, 2011; see also Soliz & Long, 2016).

Working may be a short-term solution to financing college, but jobs can positively and/or negatively affect postsecondary outcomes. Instead of forgoing studies, some students work long hours in college to support themselves (Perna, 2010; Scott-Clayton, 2011). These already economically disadvantaged and lower-performing students are likely to have a harder time succeeding in college from the get-go (Bozick, 2007; Goldrick-Rab et al., 2016; Keane & Wolpin, 2001). Several studies have documented the barriers inherent in college access and success when (low-income) students are faced with financial constraints that interact with academic constraints (Andrews et al., 2010; Long & Riley, 2007; Page & Scott-Clayton, 2016).

Empirical evidence on the specific effect work has on academics is inconclusive (Canabal, 1998; Dundes & Marx, 2006; Ecton et al., 2023; Ehrenberg & Sherman, 1987; Hawkins et al., 2005; Pike et al., 2008; Scott-Clayton & Minaya, 2016), even when plausibly accounting for exogenous variation (Darolia, 2014; DeSimone, 2008; Kalenkoski & Pabilonia, 2010; Soliz & Long, 2016; Stinebrickner & Stinebrickner, 2003). This confusion stems from the differences in data sources, job types, sample inclusion criteria, and the various constructions of such measures (Riggert et al., 2006). Essentially, there are three claims that surface in the literature:

1. Working does not affect postsecondary performance and persistence.
2. Working is negatively associated with academic achievement.
3. Some work improves student success, but too much work is bad.

Indeed, scholars recognize that these conclusions change when analyzing subsamples (e.g., Darolia, 2014; Ehrenberg & Sherman, 1987; Kalenkoski & Pabilonia, 2010; Scott-Clayton, 2011; Scott-Clayton & Minaya, 2016; Soliz & Long, 2016).

Students working 20+ hours typically have lower grade point averages (GPAs) than their peers (Dundes & Marx, 2006; Hawkins et al., 2005; see also Kalenkoski & Pabilonia, 2010). Pike et al. (2008) adjust for background characteristics, and these students still have significantly lower grades than others, yet nonworking students were not very different from those working between 1 and 20 hours. When using hours worked as a continuous variable, there are mixed results with both negative (Brint & Cantwell, 2010; DeSimone, 2008; Stinebrickner & Stinebrickner, 2003) and null (Canabal, 1998; Darolia, 2014; Ehrenberg & Sherman, 1987) effects on GPA (often the only outcome analyzed). Gleason (1993) shows that students who worked 1 to 10 hours or 11 to 20 hours had higher GPAs than those who did not work or worked more than 21 hours (see also NCES, 1994). Darolia (2014) also found a small positive association of up to 5 hours but consistently negative effects for hours worked on credits (and no significant association for grades; see also Ecton et al., 2023; Roksa & Kinsley, 2019). However, Soliz and Long (2016) show small, negative relationships for grades but positive increases in credits earned by the end of the first year (for their Ohio-based sample).

In addition to postsecondary performance, attention has been given to work's influence on persistence. Ehrenberg and Sherman's (1987) results differ from their null findings for grades—working 20 hours per week increased the probability of dropping out after the first year by 6.6% for 2-year students and 3.2% for 4-year students. Bozick (2007) reiterates this phenomenon, as working 20+ hours during a student's first year significantly decreased their ability to persist to the second year, but no effect was present for those working 1 to 20 hours (see also NCES, 1994). The matching results of Choi (2018) also agree: There were significant and deleterious effects on first-year retention for those working 20+ hours (versus other working students and/or nonworking students). Yet again, Soliz and Long (2016) dissent, showing no significant association between FWS participation and persistence.

The Present Study

What we know about the working college student is mostly based on data from before the Great Recession (e.g., Bozick, 2007; Ehrenberg & Sherman, 1987; Gleason, 1993) such as the 1997 National Longitudinal Survey of Youth (Choi, 2018; Darolia, 2014; Kalenkoski & Pabilonia, 2010), and the latest samples are of students who started college in the early 2000s (Pike et al., 2008; Scott-Clayton & Minaya, 2016). Furthermore, some studies have been limited to one state (Brint & Cantwell, 2010; Canabal, 1998; Ecton et al., 2023; Roksa & Kinsley, 2019; Scott-Clayton, 2011; Soliz & Long, 2016) or even a single case study (Dundes & Marx, 2006; Stinebrickner & Stinebrickner, 2003). Examining work's effect on postsecondary outcomes is crucial for students, practitioners, policymakers, and taxpayers, but the literature suffers from dated data and/or limited sampling. In this article, I use a recent, nationally representative sample to address these issues. Explicitly, I investigate the following research questions:

- What are the thresholds, if any, for when weekly hours worked significantly affect postsecondary outcomes?
- Do the turn points differ for Pell Grant recipients, federal student loan borrowers, and first-generation college students?
- Are students who work long hours also more likely to indicate that finances are the reason for their early exit (i.e., leaving college before completing one's intended degree)?

Although I will add to the mixed literature, the post-Recession landscape requires an update. Since the Recession's shock to higher education and the workforce, the way students pay for college has become an increasingly popular piece of political debate. Prospective students and their families are questioning whether a degree is really worth the time, money (work), and effort in this sky-high tuition era.

Data and Methods

Sample

The HSLS is a nationally representative sample of high school students, who started their secondary studies in 2009. I use the Base Year (2009) through Second Follow-up (2016), including Postsecondary Education Transcripts and Student Financial Aid Records (2017–2018). While the full sample includes 25,206 students, I study a subsample (N = 4,418) of full-time "traditional" undergraduates with valid information for all key variables.

This subsample of students is limited to those who studied primarily full-time in their first year. In addition to using 12 credits attempted (per semester) for "full-time," I include responses from postsecondary institutions (and respondents) as to whether the students were declared (or self-declared) as such. I justify this broad criterion, as institutions have various cutoffs, corresponding financial aid eligibility, and work–study budgets/policies in place. There were no notable differences for using university responses alone or just 12 credits attempted (as noted on the transcripts), so I keep all cases.

Measures

Weekly Hours Worked. The key variable at the center of this study is students' weekly hours worked (for pay). These self-reported hours come from the student questionnaire during the second follow-up. The respondents were first asked: "Did you ever work for pay during weeks you were also attending [Name of only college/trade school attended/college or trade school] in the following time period(s)?" Students were asked to not count occupations held when not attending school (e.g., a job held only during a summer break). They were also instructed to consider only paid experiences, including the following: part-time work, temporary/odd jobs, paid work experience programs (such as internships, apprenticeships, and co-ops), formal work–study jobs, self-employment, and military service.

If students said they did not work in response to the aforementioned question, they were logically coded zero and placed into the reference group. Students with a "yes" response to the above question were then asked: "How many hours per week did you usually work while attending [Name of only college/trade school attended/college or trade school] in the following time period(s)?" Self-reported hours between 0 and 40 were left unchanged, but for the 38 students who reported working over 40 hours, I rescaled their hours down to 40 as the maximum. Since this question was asked for multiple academic years, I use students' postsecondary start dates to connect the correct first year. For instance, consider the graduation date of June 2013: If a student took a full gap year before college, I would use hours worked during the 2014–2015 year while immediate enrollees would be coded from 2013–2014.

Postsecondary Outcomes. The outcomes for this study include the following: first-year GPA, first-year credits earned, first-year semester-to-semester persistence, year-to-year persistence (i.e., year 1 to year 2), ever stopout by June 2016, and ever dropout by June 2016. First-year GPA and first-year credits earned refer to the first 12 months of postsecondary enrollment, not necessarily the traditional fall-to-spring academic year. These performance-based outcomes are on "normalized" scales by NCES. Credit hours or units earned were placed on a common scale, so that they could be compared across colleges. Duplicate course records (created by transferring) were counted only once, and marks from Advanced Placement, auditing, or withdrawn courses were excluded from the calculation. Similar to hours worked, I connect these outcomes with postsecondary start dates to use all available data.

Next, semester-to-semester persistence and year-to-year persistence are coded via students' month-by-month enrollment status from their transcripts. For semester-to-semester persistence, the student persists if they enrolled by the following June (if coming from the previous summer/fall) or the following December (if coming from the previous spring/summer). Then, if a student had two consecutive semesters counted as semester-to-semester persistence, they were coded as having year-to-year persistence. In the online supplement, I also use more conservative cutoffs for these outcomes, such as September for the fall start and March for the spring (instead of

December and June). In the main text, I opt for giving students the best possible outcome given the variety of university census dates and documentation processes.

For the remaining outcomes concerning early exits (i.e., leaving college before a student completes their intended degree), I compare students' initial enrollment status with their standing as of June 2016. Those categorized as ever stopping out were students who were unenrolled for at least five consecutive months (Horn, 1998; Radford et al., 2016) but re-enrolled before June 2016. Since information concerning dropouts also comes from the student questionnaire during the second follow-up (when students were asked why they left college), I combine information from their responses and postsecondary transcripts. First, I take all available information for why one might be considered a dropout (or college graduate). Then, I recode the variable by assigning students with stopout episodes before spring 2016 and students who earned a certificate or degree to the reference group as nondropouts.

Finally, I extend these exit-related outcomes to student responses (in the second follow-up) for why they left their institution, specifically for financial reasons and family (personal) reasons. For these extended outcomes, the reference groups of nonstopouts and nondropouts remain unchanged, but those coded one are students who noted specific reasons for leaving early on the survey. Since the reasons are not mutually exclusive, there is some overlap between the outcomes, and I apply the reasons to both stopping out and dropping out—the questionnaire asked, "Generally, which of the following reasons describe why you left . . ."

Covariates. In addition to the abovementioned key variables, I use 20 adjustment variables for demographic characteristics, prior academic achievement, college characteristics, and (first-year) family obligations. Covariates for demographic characteristics include females (as identified in the 11th grade, compared to males in the reference group), Catholic private/other private schools (compared to public schools), Blacks/Hispanics/Asians/any other race (compared to Whites), family income (natural logarithm) as of 11th grade, and parent's/guardian's (henceforth parent's) education in years from the most highly educated parent. For each of these, I use information from the first follow-up (when students were in the spring term of their junior year); if missing, I pull information from the Base Year forward. Variables for prior academic achievement include high school GPA (unweighted and cumulative) and Scholastic Aptitude Test scores (or ACT equivalent as on the concordance scale; Dorans, 1999). College characteristics include net tuition price (in thousands of dollars) and enrollment status for students who were immediate highly selective 4-year enrollees/immediate moderately selective 4-year/immediate inclusive 4-year/immediate another (unclassified) 4-year/delayed any 4-year/delayed any 2-year (all compared to immediate 2-year enrollees).[1] Finally, I adjust for whether the student was married/divorced/widowed/living in a marriage-like situation during the first year of college (compared to being single) and whether the student was a parent during the first year (compared to having no children).[2]

Missing Data

Fifteen percent of the analytic sample had incomplete cases. To handle missing data, via Stata, I employ multiple imputation by chained equations, which uses separate conditional distribution for imputing values (van Buuren et al., 1999; White et al., 2011). Within this sequencing, I specify distributions through logistic regressions for dummy variables and predictive mean matching (five-nearest-neighbor) for continuous variables. To prevent perfect predictors, I use augmented regressions (White et al., 2010). The aforementioned variables (including the outcomes and weekly hours worked) are used as predictors of missing values. I ultimately yield 10 multiply imputed datasets with the following covariates holding imputed values (with corresponding percentages): high school GPA (4.35%), Scholastic Aptitude Test (7.24%), net tuition price (3.78%), marital status (1.31%), and parental status (0.38%).

Analysis

I will estimate the effects of weekly hours worked on postsecondary outcomes via third-degree polynomial regression analysis. Scholars who have previously examined hours often use categorical variables, such as the following: non-working versus 1 to 20 hours versus 21+ hours (Bozick, 2007; see also Choi, 2018; Dundes & Marx, 2006; Hawkins et al., 2005; Kalenkoski & Pabilonia, 2010; Pike et al., 2008; Roksa & Kinsley, 2019). For those who do not use categorical cutoffs, employing hours directly as a continuous predictor is common (Brint & Cantwell, 2010; Canabal, 1998; Stinebrickner & Stinebrickner, 2003), with some specifically noting that their data do not suggest a nonlinear relationship (Darolia, 2014; Ehrenberg & Sherman, 1987). Although DeSimone (2008) accounts for nonlinearity, he only analyzes GPA. Without assessing the full range of hours worked and several outcomes (with a rich set of adjustment variables), the question of "how much work is too much" cannot be effectively answered for today's students.

In this article, I employ ordinary least squares and linear probability models to address my research questions. Model 1 will be unadjusted to obtain reference points, expressed as:

$$\begin{aligned} Y_{ij} &= \beta_0 + \beta_1 \textit{Weekly Hours Worked}_{ij} \\ &+ \beta_2 \textit{Weekly Hours Worked}^2{}_{ij} \\ &+ \beta_3 \textit{Weekly Hours Worked}^3{}_{ij} + u_{ij}, \end{aligned} \tag{1}$$

where Y is the postsecondary outcome (i.e., first-year GPA, first-year credits earned, semester-to-semester persistence, year-to-year persistence, ever stopout, or ever dropout) for student i clustered in college j; β_0 is the intercept; β_1 is weekly hours worked for pay during the students' first year with β_2 and β_3, respectively reflecting the naturally squared and naturally cubed polynomial terms; and u is the error term. Model 2 then extends to the following:

$$\begin{aligned} Y_{ij} &= \beta_0 + \beta_1 \textit{Weekly Hours Worked}_{ij} \\ &+ \beta_2 \textit{Weekly Hours Worked}^2{}_{ij} \\ &+ \beta_3 \textit{Weekly Hours Worked}^3{}_{ij} \\ &+ \beta_4 \mathbf{X}_{ij} + u_{ij}, \end{aligned} \tag{2}$$

where all is the same as in equation (1) with additional adjustments for β_4, that is, a vector of covariates $\mathbf{X}$ for demographic characteristics, prior academic achievement, college characteristics, and students' first-year family obligations. Both models will use cluster-robust standard errors for students' postsecondary institutions (N = 1,055).

Via Stata, I will then use predictive margins to understand the effects at each hour, which will visualize and estimate the thresholds and turn points of statistical significance. These effects will then be estimated for Pell Grant recipients, federal student loan borrowers, and first-generation students. I will conclude my analysis with a discussion on causal inference. In the online supplement, I conduct several robustness checks and examine additional subsamples.

Results

Table 1 presents the means and standard deviations (*SD*; by any work-study participation) for all variables in the analysis. More than half of all students (53.76%) reported working during their first year.[3] Of those working, the average time spent per week was 21.26 (*SD* = 10.23) hours; for all students, though, the mean was 11.43 (*SD* = 12.99) hours. In terms of postsecondary outcomes, nonworking students had higher academic achievement in college than those who worked. For instance, nonworking students earned 26.28 (*SD* = 9.55) credits in their first year, and only 10.93% had a stopout episode by June 2016. Meanwhile, working students earned an average of 24.16 (*SD* = 10.15) credits, and 16.08% had a stopout episode.

For the covariates, from a descriptive standpoint, those who identified as women were more likely to work than men. Working students also came from slightly less socioeconomically advantaged backgrounds and were less likely to be Asian. High school GPAs were nearly identical! For college characteristics, working students' tuition costs were lower after merit-based scholarships and (especially) need-based grants; nonworking students had an average net tuition price of $7,943 versus working students' tuition bill of $5,090. Relatedly, 27.22% of nonworking students were in highly selective institutions, while only 14.36% of working students were in such colleges; expectedly, regardless of timing (immediate or delayed), 36.71% of working students enrolled 2-year colleges versus 21.03% of nonworking students.

TABLE 1

Descriptive Statistics of First-Time Full-Time "Traditional" Undergraduate Students by Employment Status

	All		Nonworking		Working	
Variable names	*M*	*SD*	*M*	*SD*	*M*	*SD*
Postsecondary outcomes						
First-year grade point average	2.75	0.92	2.80	0.90	2.70	0.94
First-year credits earned	25.14	9.93	26.28	9.55	24.16	10.15
Semester-to-semester persistence	0.93		0.95		0.92	
Year-to-year persistence	0.82		0.85		0.79	
Ever stopout by June 2016	0.14		0.11		0.16	
Ever dropout by June 2016	0.27		0.24		0.30	
Key variables						
Any work-study participation	0.54		0.00		1.00	
Weekly hours worked (for pay)	11.43	12.99	0.00	0.00	21.26	10.23
Covariates						
Female (11th grade)	0.53		0.46		0.58	
Black	0.11		0.11		0.12	
Hispanic	0.18		0.17		0.19	
Asian	0.05		0.07		0.03	
Another race (non-White)	0.08		0.09		0.07	
Family income (natural logarithm)	11.07	0.77	11.14	0.79	11.01	0.74
Parent highest education (years)	14.89	2.58	15.17	2.63	14.64	2.51
Catholic private high school	0.05		0.06		0.04	
Another private high school	0.05		0.06		0.04	
High school grade point average	3.07	0.64	3.09	0.65	3.05	0.63
Scholastic Aptitude Test Score	1,008.43	193.79	1,030.94	203.16	989.07	183.19
Net tuition price ($1,000)	6.41	8.83	7.94	9.97	5.09	7.47
Highly selective 4-year	0.20		0.27		0.14	
Moderately selective 4-year	0.34		0.37		0.33	
Inclusive 4-year	0.08		0.08		0.09	
Another 4-year	0.05		0.05		0.05	
Delayed enrollment at any 4-year	0.02		0.03		0.02	
Delayed enrollment at any 2-year	0.03		0.02		0.04	
First-year student is married	0.02		0.00		0.02	
First-year student is a parent	0.01		0.01		0.02	

Note. Source = High School Longitudinal Study. Data are weighted by W5W1W2W3W4PSRECORDS, modified with a ratio adjustment for having values across weekly hours worked and all outcomes. Missing values for covariates are multiply imputed by chained equations (10 datasets). N = 4,418. Delayed enrollment was after December 2014—all other timings were before as immediate entry. M = mean; SD = standard deviation.

Thresholds and Turn Points

The predicted marginal effects for the outcomes are presented in Figure 2 (performance), Figure 3 (persistence), and Figure 4 (early exits). These estimates provide the unadjusted values based on (1) visualized through the thin, green-cyan lines. The thick, dark blue lines denote the covariate-adjusted values based on (2) with the 95% confidence intervals (dashed lines). All underlying polynomial regression models are available and discussed in the online supplement (see Supplementary Tables A1 and A2 in the online version of the journal).

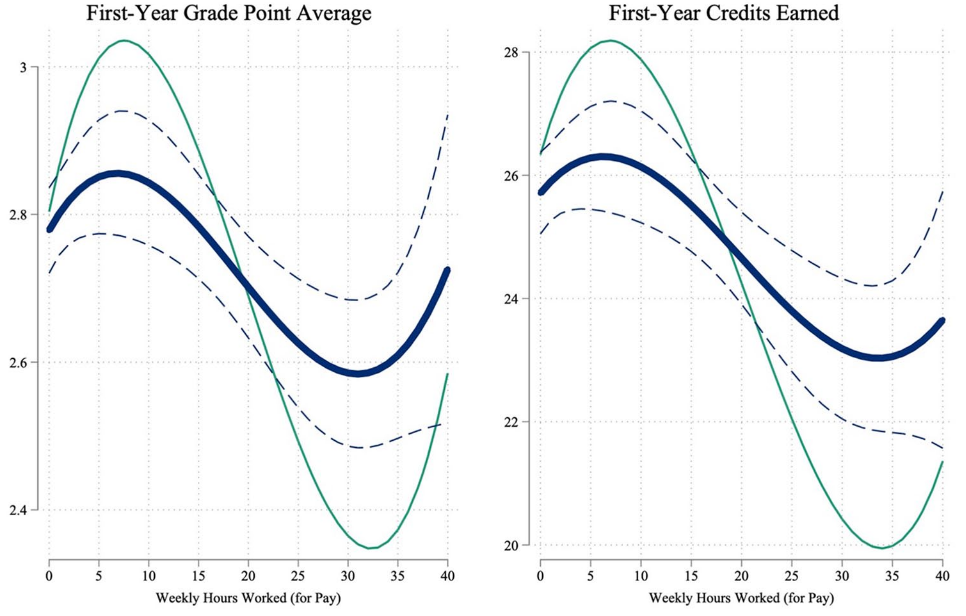

FIGURE 2. *Predictive margins of first-year postsecondary performance over weekly hours worked, adjusted for demographic characteristics, prior academic achievement, college characteristics, and family obligations.*
Note. Source = High School Longitudinal Study. Data are weighted by W5W1W2W3W4PSRECORDS, modified with a ratio adjustment for having values across weekly hours worked and all outcomes. Missing values for underlying covariates are multiply imputed by chained equations (10 datasets). Unadjusted estimates are shown on the thin green-cyan line. The thick blue line shows the adjusted estimates with corresponding confidence intervals (dashed blue lines). $N = 4{,}418$.

The graphs mostly reveal a curvilinear relationship, positing that some work is good for students but too much is bad. For GPA and credits earned (Figure 2), there was a brief bump for students who worked a few hours and then a downward trend starting around 12 hours that turned again/tapered off around 35 hours (with a starker positive return for GPA). A similar pattern was found for year-to-year persistence (Figure 3), though the downward trend turned later (i.e., around 15 hours versus 12); meanwhile, semester-to-semester persistence had a subtle but steady decline with a clearer turn around 35 hours. For stopping out and dropping out the patterns were flipped. So, students who worked a few hours were slightly less likely to have a stopout episode or (more so) dropout altogether, and this likelihood increased around 13 hours for stopping out (and 18 hours for dropping out). Again, there was slight tapering at the end for stopping out and virtually none for dropping out.

The first research question is particularly concerned with the threshold for when work negatively and significantly affect postsecondary outcomes. The point estimates (Table 2 and see Supplementary Table A3 in the online version of the journal) show the significant turn points. First, the good news: Working up to 19 hours had little to no effect on postsecondary outcomes (and some work was suggestively good). However, there was a stark and significant turn for the worse for students who worked over 20 hours, and this was robust across the board for those between 34 and 36 hours. The positive peak, visually, was between 6 and 9 hours; GPA was the only outcome suggestively affected by modest work, as working 1 to 4 hours was significantly better for students. Peaking at 4 hours, on average, students' GPA was 0.07 points ($p = .047$) higher compared to nonworking students.

In terms of the negative effects, I find identical results to the literature that suggests that 20 hours is the ostensible turn point (also a common cutoff

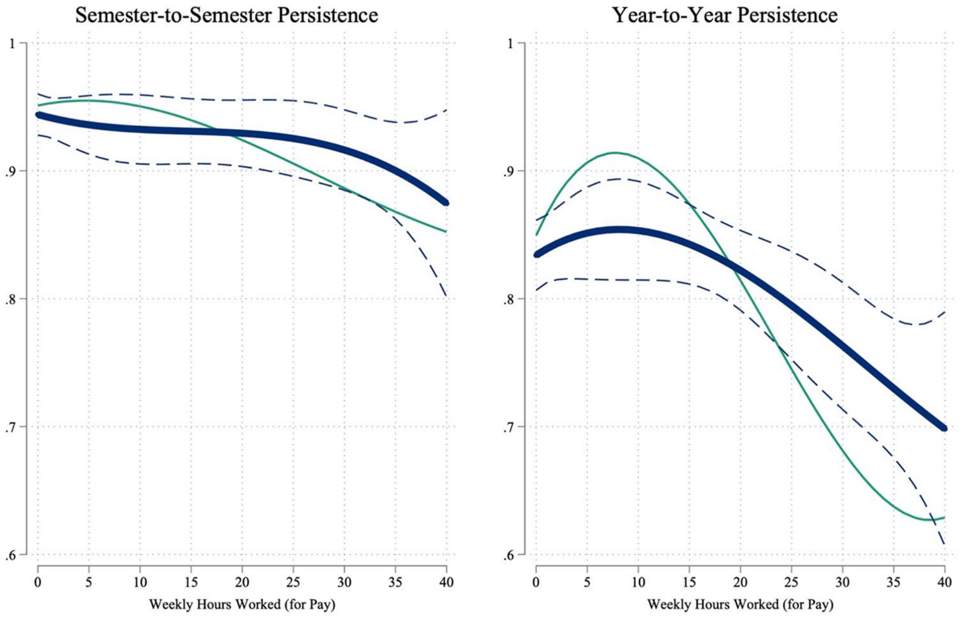

FIGURE 3. *Predictive margins of postsecondary persistence over weekly hours worked, adjusted for demographic characteristics, prior academic achievement, college characteristics, and family obligations.*
Note. Source = High School Longitudinal Study. Data are weighted by W5W1W2W3W4PSRECORDS, modified with a ratio adjustment for having values across weekly hours worked and all outcomes. Missing values for underlying covariates are multiply imputed by chained equations (10 datasets). Unadjusted estimates are shown on the thin green-cyan line. The thick blue line shows the adjusted estimates with corresponding confidence intervals (dashed blue lines). N = 4,418.

in practice). At 20 hours, the effects for credits earned and stopping out became significant (compared to nonworking students). In other words, students earned 1.06 (p = .032) credit fewer when working 20 hours versus their non-working counterparts. In addition, the probability that these working students would temporarily interrupt their studies increased by 4.08% (p = .049). For students who worked 21 hours, their GPA decreased by 0.09 (p = .045). Later on, at 28 hours, students were 5.80% (p = .045) less likely to have year-to-year persistence. At 31 hours, students were 6.63% (p = .044) more likely to dropout. The last turn point happened at 34 hours, where students were 3.99% (p = .047) less likely to return for their second semester. In short, as students worked more hours above the 20-hour turn point, their outcomes increasingly suffered.

The deleterious effects were strong between 28 and 36 hours (i.e., significant effects on most outcomes before tapering of effects). For instance, those who worked 32 hours lost 0.19 (p = .002) points on their GPA compared to nonworking students; at this level, students also lost 2.65 (p < .001) credits (or failed a class). For persistence, students working 32 hours were 8.42% (p = .006) less likely to make it to their second year and 7.86% (p = .013) more likely to stopout. These students were also 7.19% (p = .031) more likely to dropout.

There was some tapering and less precision at the top of the range. The last effect for GPA was at 36 hours, where students earned .15 (p = .031) fewer points. At 37 hours, the last threshold for semester-to-semester persistence was present with the probability decreasing by 5.28% (p = .043). At 38 hours, students earned 2.40 (p = .006) fewer credits compared to nonworking students. Finally, those who worked 40+ hours were 13.59% (p = .007) less likely to make it to their second year and 11.07% (p = .031) more likely to dropout.

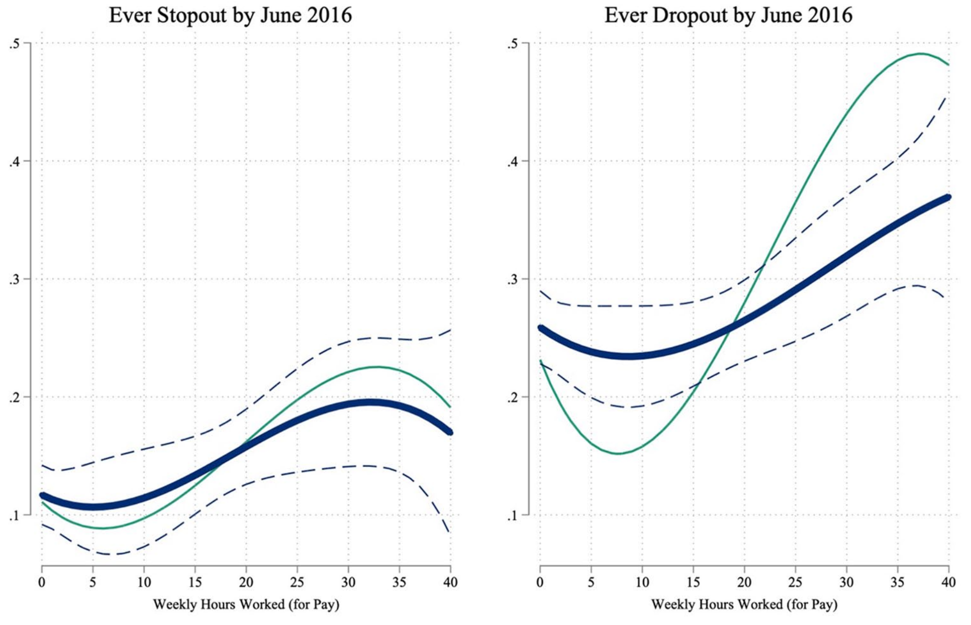

FIGURE 4. *Predictive margins of exiting early over weekly hours worked, adjusted for demographic characteristics, prior academic achievement, college characteristics, and family obligations.*
Note. Source = High School Longitudinal Study. Data are weighted by W5W1W2W3W4PSRECORDS, modified with a ratio adjustment for having values across weekly hours worked and all outcomes. Missing values for underlying covariates are multiply imputed by chained equations (10 datasets). Unadjusted estimates are shown on the thin green-cyan line. The thick blue line shows the adjusted estimates with corresponding confidence intervals (dashed blue lines). $N = 4{,}418$.

Financial Aid Subgroups

Given the public interest in unmet need during the sky-high tuition era, I examine the effects for two financial aid subgroups: Pell Grant recipients and student loan borrowers (including Parent PLUS Loans). Here, I limit the discussion to the effects for the longer-term outcomes (i.e., year-to-year persistence, stopping out, and dropping out), but the results for the shorter-term outcomes (constrained to the first year) are available in the online supplement (see Supplementary Figures A1 and A2 in the online version of the journal). As presented in Figure 5, Pell Grant recipients who worked 1 to 14 hours were slightly more likely to persist to the second year (compared to nonworking recipients); at the peak (i.e., 7 hours), the predicted probability that a Pell student would return for the second year was 80.26%, which was similar to that of nonrecipients working 30 hours (80.11%). However, starting at 34 hours, Pell students were significantly less likely to persist to the second year with the worst effects happening at 40 hours when they were 13.85% ($p = .046$) less likely to persist. In addition, working students without Pell Grants were less likely to persist starting at 32 hours with the largest effect also at 40 hours, where they were 14.36% ($p = .046$) less likely to return for their second year versus nonworking students without Pell Grants.

Pell Grant recipients also stopped out and dropped out (on average across the board) at higher rates than nonrecipients. For Pell students, the respective turn points were 34 hours for stopping out and 30 hours for dropping out, where recipients were 9.36% ($p = .046$) more likely to temporarily unenroll and 11.43% ($p = .043$) more likely to dropout all together (compared to nonworking recipients). The thresholds for stopping out and dropping out were nonexistent for nonrecipients, though the patterns suggest a similar association.

At the top of the range, a drastic difference occurred for dropping out. The predicted probability

TABLE 2

Difference in Predicted Point Estimates for Weekly Hours Worked Compared to Nonworking Students (Even Increments), Adjusted for Demographic Characteristics, Prior Academic Achievement, College Characteristics, and Family Obligations

Weekly Hours Worked	First-year grade point average	First-year credits earned	Semester-to-semester persistence	Year-to-year persistence	Ever stopout by June 2016	Ever dropout by June 2016
2 hours	0.041*	0.331	−0.004	0.009	−0.007	−0.011
	(0.019)	(0.210)	(0.006)	(0.009)	(0.009)	(0.009)
4 hours	0.065*	0.524	−0.007	0.015	−0.010	−0.018
	(0.033)	(0.356)	(0.010)	(0.015)	(0.016)	(0.016)
6 hours	0.077	0.594	−0.009	0.019	−0.010	−0.023
	(0.042)	(0.449)	(0.013)	(0.019)	(0.020)	(0.020)
8 hours	0.076	0.557	−0.011	0.020	−0.008	−0.025
	(0.046)	(0.498)	(0.014)	(0.021)	(0.022)	(0.023)
10 hours	0.065	0.428	−0.012	0.019	−0.003	−0.024
	(0.047)	(0.513)	(0.015)	(0.021)	(0.022)	(0.023)
12 hours	0.046	0.223	−0.012	0.016	0.004	−0.022
	(0.047)	(0.505)	(0.015)	(0.021)	(0.022)	(0.023)
14 hours	0.020	−0.043	−0.013	0.012	0.012	−0.017
	(0.045)	(0.488)	(0.015)	(0.020)	(0.020)	(0.023)
16 hours	−0.010	−0.356	−0.013	0.005	0.021	−0.011
	(0.044)	(0.473)	(0.015)	(0.020)	(0.020)	(0.022)
18 hours	−0.043	−0.698	−0.014	−0.003	0.031	−0.003
	(0.044)	(0.472)	(0.015)	(0.020)	(0.020)	(0.023)
20 hours	−0.077	−1.055*	−0.015	−0.012	0.041*	0.006
	(0.045)	(0.491)	(0.015)	(0.021)	(0.021)	(0.024)
22 hours	−0.110*	−1.412**	−0.016	−0.022	0.050*	0.016
	(0.048)	(0.525)	(0.016)	(0.023)	(0.023)	(0.025)
24 hours	−0.139**	−1.753**	−0.018	−0.033	0.059*	0.026
	(0.052)	(0.568)	(0.017)	(0.025)	(0.025)	(0.027)
26 hours	−0.164**	−2.062**	−0.020	−0.045	0.067*	0.038
	(0.055)	(0.607)	(0.017)	(0.027)	(0.028)	(0.029)
28 hours	−0.183**	−2.325***	−0.023	−0.058*	0.073*	0.049
	(0.058)	(0.637)	(0.018)	(0.029)	(0.030)	(0.031)
30 hours	−0.193**	−2.525***	−0.028	−0.071*	0.077*	0.061
	(0.060)	(0.654)	(0.018)	(0.030)	(0.031)	(0.032)
32 hours	−0.193**	−2.648***	−0.033	−0.084**	0.079*	0.072*
	(0.061)	(0.662)	(0.019)	(0.031)	(0.032)	(0.033)
34 hours	−0.180**	−2.677***	−0.040*	−0.097**	0.077*	0.083*
	(0.064)	(0.676)	(0.020)	(0.031)	(0.032)	(0.034)
36 hours	−0.154*	−2.599***	−0.048*	−0.111**	0.073*	0.093*
	(0.071)	(0.730)	(0.023)	(0.034)	(0.034)	(0.037)
38 hours	−0.112	−2.396**	−0.058	−0.124**	0.065	0.102*
	(0.087)	(0.865)	(0.030)	(0.040)	(0.038)	(0.042)
40+ hours	−0.052	−2.054	−0.070	−0.136**	0.052	0.111*
	(0.113)	(1.109)	(0.039)	(0.050)	(0.047)	(0.051)

Note. Source = High School Longitudinal Study. Data are weighted by W5W1W2W3W4PSRECORDS, modified with a ratio adjustment for having values across weekly hours worked and all outcomes. Missing values for underlying covariates are multiply imputed by chained equations (10 datasets). $N = 4{,}418$. Predicted values with delta-method standard errors in parentheses, adjusted for all covariates as in Equation 2.
*$p < .05$. **$p < .01$. ***$p < .001$.

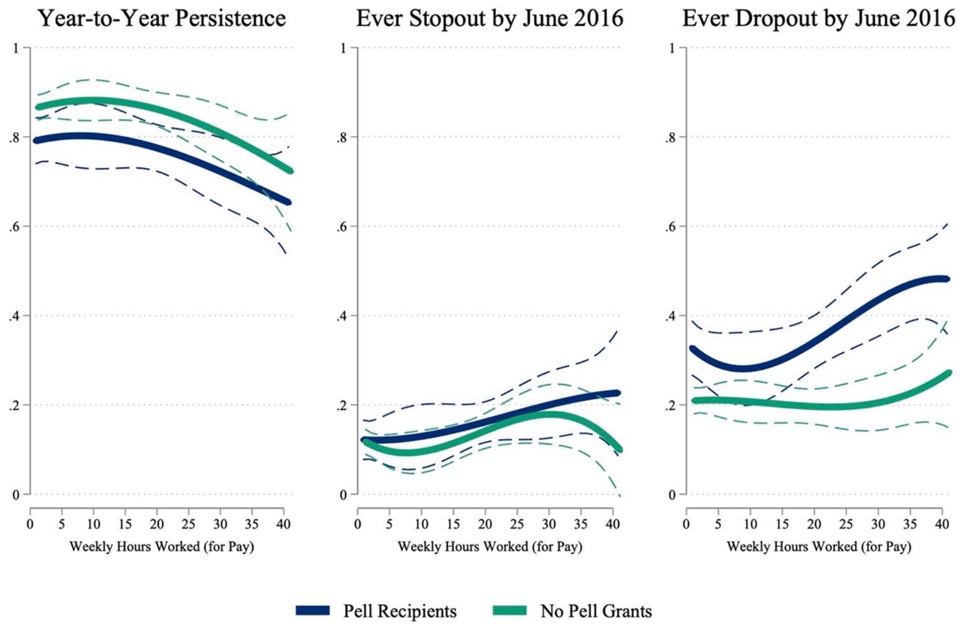

FIGURE 5. *Predictive margins of longer-term outcomes over weekly hours worked by Pell Grant status, adjusted for demographic characteristics, prior academic achievement, college characteristics, and family obligations.*
Note. Source = High School Longitudinal Study. Data are weighted by W5W1W2W3W4PSRECORDS, modified with a ratio adjustment for having values across weekly hours worked and all outcomes. Missing values for underlying covariates are multiply imputed by chained equations (10 datasets). Polynomial plots of predicted values over weekly hours worked. Pell Grant recipients N = 1,478. Nonrecipients N = 2,921.

of dropping out for Pell recipients who worked 40+ hours per week was 48.18% while this estimate was only 27.34% for nonrecipients at this level. In other words, these low-income (doubly full-time) students were nearly one-half *SD* as likely to dropout within the first 3 years compared to nonrecipients.

Both student loan borrowers and nonborrowers had nearly identical starting points for zero hours worked (for these longer-term outcomes), but employment intensity again produced heterogeneous effects (Figure 6). For year-to-year persistence, borrowers who worked 1 to 24 hours were slightly more likely to return for the second year compared to nonworking borrowers. Borrowers working 5+ hours were also more likely to return for the second year than nonborrowers—work also affected year-to-year persistence for nonborrowers at 28 hours, when they were 8.79% (p = .048) less likely to persist. Student loan borrowers with 22 weekly hours worked were 5.79% (p = .047) more likely to stopout, but this effect was null for nonborrowers. For dropping out, the adverse association was at 29 hours for nonborrowers, where they were 9.57% (p = .042) more likely to exit.

The turn points for nonborrowers suggests that taking out loans alleviates some of the challenges that interrupt studies, and those solely working their way through college are more sensitive to work. If one hypothesizes that money is a barrier to getting through school, then it makes sense that credit-constrained students are more likely to exit, and that those with loans are not as immediately burdened by rising costs (although student debt is a topic beyond the scope of this study). Interestingly, the predicted probability of dropping out (29.04%) for nonborrowers who worked 20 hours was identical to that of Pell recipients who worked 13 hours (29.48%).

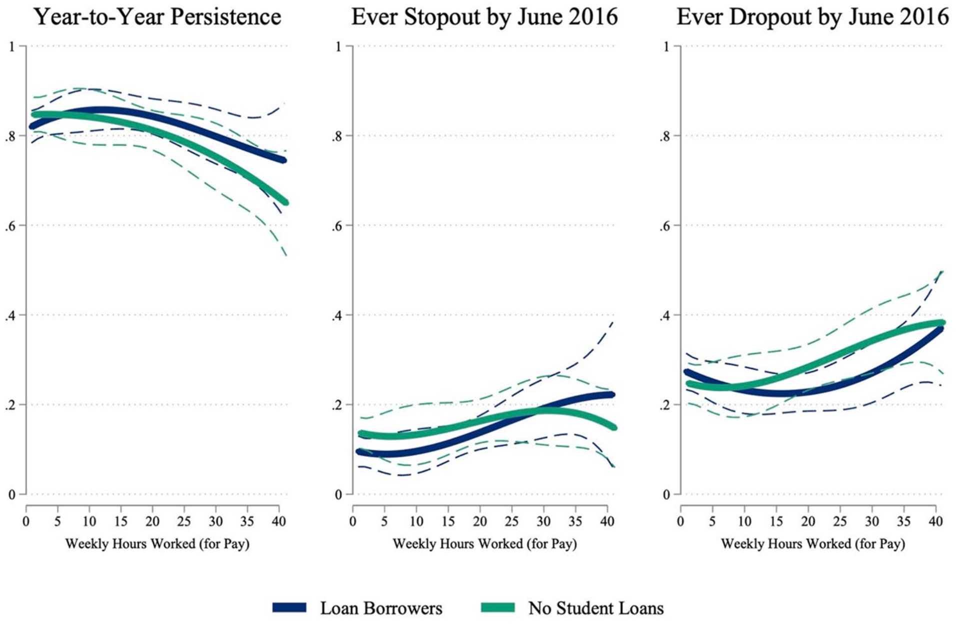

FIGURE 6. *Predictive margins of longer-term outcomes over weekly hours worked by student loan status, adjusted for demographic characteristics, prior academic achievement, college characteristics, and family obligations.*
Note. Source = High School Longitudinal Study. Data are weighted by W5W1W2W3W4PSRECORDS, modified with a ratio adjustment for having values across weekly hours worked and all outcomes. Missing values for underlying covariates are multiply imputed by chained equations (10 datasets). Polynomial plots of predicted values over weekly hours worked. Data are weighted. Federal student loan borrowers (including Parent PLUS) $N = 2{,}018$. Nonborrowers $N = 2{,}381$.

First-Generation Students

Although I adjust for parental education, there may still be variance by first-generation student status. Indeed, there is no shortage of definitions for what constitutes "first-generation" among higher education institutions, scholars, and policymakers (e.g., Ives & Castillo-Montoya, 2020; Toutkoushian et al., 2021). Here, I use three definitions: (1) parent(s) without at least a postsecondary certificate/degree, (2) parent(s) without at least an associate degree, and (3) parent(s) without at least a bachelor's degree. These are plotted against non-first-generation students at all levels (i.e., families with at least a postsecondary certificate).

Expectedly, first-generation students often performed worse than their non-first-generation counterparts (Figure 7; see also Supplementary Figure A3 in the online version of the journal). However, working a modest number of hours was still positively associated with higher outcomes, though the overall relationships were mostly identical between first-generation definitions. Consider year-to-year persistence. The predicted probability for those with parent(s) who did not earn at least a postsecondary certificate was 77.64% at 20 hours; those coming from families without at least an associate degree (76.60%) or without at least a bachelor's degree (77.50%) were also similar at 20 hours.

Since the economic returns are best for bachelor's degree recipients, and the results are nearly identical, I will limit the remaining discussion to the group of students from families without at least a bachelor's degree. At 30 hours, students from families without at least a bachelor's degree were 9.15% ($p = .043$) less likely to persist to the second year (below the predicted probability of 78.96% for nonworking first-generation students); there was a similar turn point for non-first-generation students at 28 hours, where they were 7.21% ($p = .036$) less likely to persist, but

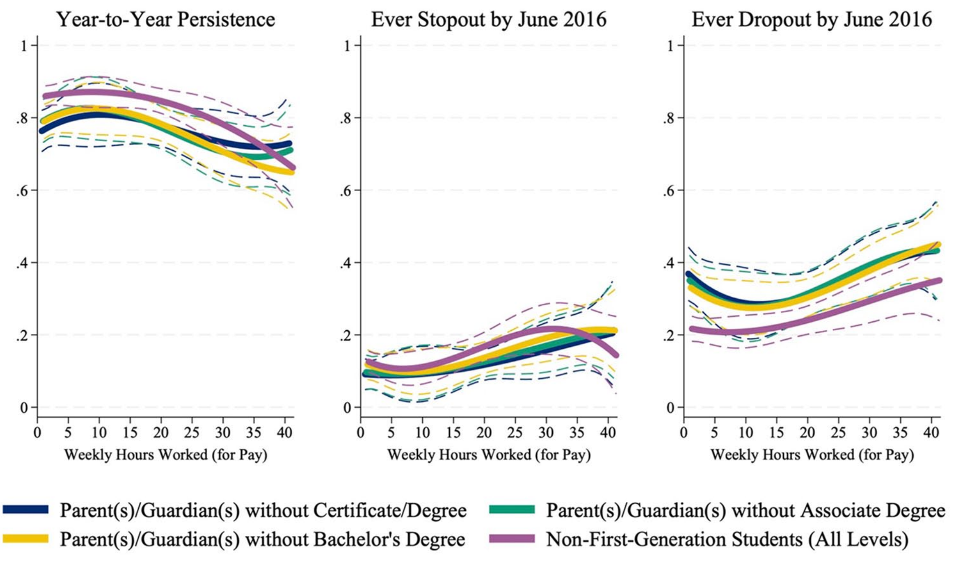

FIGURE 7. *Predictive margins of longer-term outcomes over weekly hours worked by first-generation student status, adjusted for demographic characteristics, prior academic achievement, college characteristics, and family obligations.*

Note. Source = High School Longitudinal Study. Data are weighted by W5W1W2W3W4PSRECORDS, modified with a ratio adjustment for having values across weekly hours worked and all outcomes. Missing values for underlying covariates are multiply imputed by chained equations (10 datasets). Polynomial plots of predicted values over weekly hours worked. All plots adjust for covariates. Parent(s) without at least a certificate/degree, N = 1,017. Parent(s) without at least an associate degree, N = 1,140. Parent(s) without at least a bachelor's degree, N = 1,749. Non-first-generation students, N = 3,401.

this was below a qualitatively higher starting point of 85.98% at zero hours. That said, at the top of the range of 40 hours, predicted probabilities were quite comparable for first-generation (66.24%) students and non-first-generation students (64.92%).

For stopping out, these first-generation students were less sensitive to work than their counterparts. At 31 hours, those from first-generation families were 8.26% (p = .046) more likely to stopout, but non-first-generation students were affected earlier at 21 hours where they were 5.73% (p = .042) more likely to stopout. Conversely, dropping out returned the expected pattern in terms of difference in rates, but the sensitivity to work was earlier for non-first-generation students. Across all hours, first-generation students were more likely to dropout compared to their non-first-generation peers. The turn point for first-generation students was at 36 hours, where they were 9.89% (p = .044) more likely to dropout above the predicted probability of 33.05%; meanwhile, non-first-generation students were 7.85% (p = .040) more likely to dropout above the predicted probability of 21.69%. In summary, working long hours is still not helpful in terms of students' longer-term outcomes. The differences in first-generation student status provide additional insight into the heterogeneity in work–study arrangements among an important population in the higher-education equity conversation.

Examining Reasons via Extended Outcomes

The last research question is concerned with the specific reasons students left. If students are working to help pay for college, then exiting early for financial and/or personal (family) reasons can shed light on the barriers inherent in the sky-high tuition era and whether the tradeoff of time is worth it. However, if students are leaving to support their families (or start one), then what is known about work–study hours can be expanded upon. The thresholds illuminated in the

TABLE 3

Descriptive Statistics by Weekly Hours Worked With Extended Outcomes and Selected Covariates

	Nonworking		1–19 hours		20–27 hours		28–36 hours		37–40+ hours	
Variable name	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Extended outcomes										
Ever stopout by June 2016	0.11		0.12		0.14		0.14		0.19	
Ever stopout (financial reasons)	0.24		0.16		0.34		0.34		0.49	
Ever stopout (family reasons)	0.01		0.01		0.02		0.02		0.07	
Ever dropout by June 2016	0.02		0.02		0.03		0.03		0.09	
Ever dropout (financial reasons)	0.05		0.02		0.09		0.09		0.23	
Ever dropout (family reasons)	0.08		0.05		0.10		0.10		0.23	
Selected covariates										
Family income ($1000)	90.85	63.23	87.24	54.57	75.76	53.69	71.80	53.24	58.18	37.13
Net tuition price ($1000)	7.94	9.97	7.00	8.67	4.21	6.49	3.55	6.10	3.61	6.26
Pell Grant amount ($1000)	1.59	2.28	1.59	2.26	2.04	2.40	2.32	2.44	2.35	2.42
Student loan amount ($1000)	3.94	6.81	4.68	6.73	3.48	5.11	3.10	6.15	2.83	5.20
Highly selective 4-year	0.27		0.26		0.10		0.04		0.04	
Moderately selective 4-year	0.37		0.39		0.34		0.25		0.18	
Inclusive 4-year	0.08		0.08		0.11		0.09		0.07	
Another 4-year	0.05		0.05		0.04		0.08		0.07	
Delayed enrollment at any 4-year	0.03		0.03		0.01		0.02		0.03	
Delayed enrollment at any 2-year	0.02		0.03		0.03		0.05		0.05	
First-year student is married	0.00		0.00		0.02		0.08		0.03	
First-year student is a parent	0.01		0.01		0.02		0.02		0.01	
Ever married by June 2016	0.05		0.06		0.10		0.20		0.19	
Ever a parent by June 2016	0.02		0.03		0.04		0.11		0.11	

Note. Source = High School Longitudinal Study. Data are weighted by W5W1W2W3W4PSRECORDS, modified with a ratio adjustment for having values across weekly hours worked and all outcomes. Missing values for covariates are multiply imputed by chained equations (10 datasets). *N* = 4,418. All covariates and *N*s for the extended outcomes (see Supplementary Table A4 in the online version of the journal). *M* = mean; *SD* = standard deviation.

previous section are now coded into five corresponding categories to examine the (extended) outcomes and selected covariates. The categories are as follows (based on the general patterns): nonworking students (reference group), 1 to 19 hours (no significant negative effects), 20 to 27 hours (start of significant negative effects), 28 to 36 hours (strong significant negative effects), and 37 to 40+ hours (tapering of effects).

Descriptive statistics for selected variables are available in Table 3 (and see Supplementary Table A4 in the online version of the journal for all variables). The results show that exiting early is mostly aligned with students coming from socioeconomically disadvantaged backgrounds with lower academic achievement in high school. Dropping out for financial reasons increased as students worked more. While 23.71% of nonworking students dropped out, only 5.20% indicated that financial reasons were the cause. Meanwhile, 33.84% of students working 20 to 27 hours dropped out, and 9.41% indicated that they did so for financial reasons. At the top of the range where students worked 37 to 40+ hours, 49.45% dropped out by June 2016, and 22.79% indicated that they left college specifically for financial reasons. Dropping out for family/personal reasons, as well as the patterns for stopping out were similar (see also Supplementary Figure A4 in the online version of the journal).

Students who work increasingly more hours had correspondingly lower net tuition prices, on average. The students at the top of the range, for example, had a tuition bill of $3,606 compared to $7,943 for nonworking students. The (mean) family income as of 11th grade for students working 37 to 40+ hours was also $58,179 versus $90,850 for nonworking students. Relatedly, the federal loan amounts (borrowed during the first year) were lowest among the students who worked 37 to 40+ hours. These students borrowed $2,832 (including Parent PLUS Loans), while nonworking students borrowed $3,939. Even more, the difference between Pell Grants was only $768 between these two groups. In other words, having less than half the tuition bill of nonworking students is still challenging to meet for those who work the most, as they come from families with nearly half the income, and Pell Grants did not fill the gap. They were also likely credit constrained as suggested with their student loan rates.

Looking at students' family circumstances also adds context to their decisions to leave and expounds upon the nonlinear effects. Only 0.41% of nonworking students were married, and 2.48% had children during their first year. Interestingly, for the highest achievement group (i.e., those who worked 1–19 hours), even fewer were married (0.23%) and had children (0.63%). Meanwhile, 7.88% of students working 28 to 36 hours were married and 2.16% had children. At the top of the range, 3.03% of first-year students were married and 0.98% had children.

Finally, I show students' marital and parental statuses at the end of the sample period. This further helps explain reasons for exiting early. Unsurprisingly, the magnitudes increased. At this point, 19.36% of those at the top of the work intensity range were married, and 10.72% had children. In addition, 19.88% of students who worked 28 to 36 hours were married, and 11.35% of this group had at least one child by June 2016. The rates for exiting early (in general and for specific reasons) were highest among the 37 to 40+ group, but the family circumstances of the 28 to 36 group were most pronounced. Neither the plots nor the model fit indices (not shown) change with the alternate indicators for marital/parental status, but the descriptive results help tell the story: Students who worked more hours were more likely to be married/have children and tended to exit early (especially for financial/family reasons), but these students also started off less academically and economically advantaged.

Propensity Score Weighting

One could ask how an explicitly causal approach to these questions could be pursued. While this is difficult to achieve with selection-on-observables (i.e., unconfoundedness), relaxing this assumption, one could estimate an average effect for a specific group as a binary "treatment." The above results (and the literature) suggest that the people who work 20 hours (i.e., intense work-study hours; Bozick, 2007; Choi, 2018; D'Amico, 1984; Lee & Staff, 2007; Staff et al., 2010; Steinberg et al., 1982) or more would be of particular interest. In this section, I provide propensity score-weighted regression estimates of the average treatment effect in the treated (ATT; see Austin & Stuart, 2015; Morgan & Winship, 2015). This method estimates potential outcomes based on the matched observations, balancing the observed covariates across the treated and untreated (i.e., attempting to make the students who worked 20+ hours identical to those who worked 0–19 hours).

This method starts with a logit model to generate students' estimated propensity scores, $\widehat{p}_i$, defined as

$$\widehat{p}_i = P(D = 1|\mathbf{X}), \tag{3}$$

which is the probability P that a student worked 20+ hours $D = 1$, conditional on covariates **X**. Using these estimated propensity scores, the treatment group is the target population of analysis:

$$\text{For } D_i = 1: w_i ATT = 1,$$

$$\text{For } D_i = 0: w_i ATT = \frac{\widehat{p}_i}{1 - \widehat{p}_i}, \tag{4}$$

where these weights leave the sampled treatment group unaltered, that is, $w_i ATT = 1$ when students participate in intense work-study (i.e., $D_i = 1$), and where the weight attempts to turn the control group into a representative sample of population-level treatment group because $w_i ATT = \frac{\widehat{p}_i}{1 - \widehat{p}_i}$ for the control group. These generated weights are then multiplied by the survey weights to account for design features of the data and used in the outcome models (Balancing diagnostics are available in the online supplement [see Supplementary Table A5 in the online version of the journal].)

TABLE 4

Average Treatment Effects in the Treated for Intense Work–Study (20+ Hours) With Some Sensitivity Analysis

	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Estimate	First-year grade point average		First-year credits earned		Semester-to-semester persistence	
ATT	−0.301***	−0.103*	−4.372***	−1.856***	−0.070***	−0.044*
	(0.052)	(0.052)	(0.592)	(0.524)	(0.017)	(0.019)
Y^1	2.846	2.638	26.586	23.998	0.955	0.927
	(0.030)	(0.034)	(0.324)	(0.355)	(0.006)	(0.011)
Y^0	2.545	2.535	22.215	22.142	0.885	0.883
	(0.045)	(0.038)	(0.546)	(0.429)	(0.016)	(0.015)
e value	2.039	1.437	2.377	1.632	1.897	1.556
Altonji ratio	—	0.522	—	0.738	—	1.716

Note. Source = High School Longitudinal Study. Data are weighted by W5W1W2W3W4PSRECORDS, modified with a ratio adjustment for having values across weekly hours worked and all outcomes. Missing values for underlying covariates are multiply imputed by chained equations (10 datasets). N = 4,418. Regression coefficients (with cluster-robust standard errors) via propensity score weighting and doubly robust adjustments for demographic characteristics, prior academic achievement, college characteristics, and family obligations. ATT = average treatment effect for the treated. Y^1 = potential outcome for the treated. Y^0 = potential outcome for the untreated. The "treatment" is intense work–study hours (i.e., 20+ hours per week). Model 1 is unadjusted and Model 2 is adjusted for all covariates. Altonji ratios are based on Altonji et al. (2005), where taking the difference between the empty model and the doubly robust model produces an implied ratio of standardized selection of unobservables to observables under the hypothesis that there is no intense work–study effect.
*$p < .05$. **$p < .01$. ***$p < .001$.

Students who worked 20+ hours had a predicted GPA 0.10 ($p = .048$) points lower than the potential outcome of 2.64 if these same students had worked 0 to 19 hours (Table 4). Even more, these students earned 1.86 ($p < .001$) fewer credits than their counterfactual state. Similar effects were present for the persistence-based measures with intense work–study students being 4.40% ($p = .021$) less likely to persist to the second semester than their potential outcome of 92.69% if these same students had worked less than 20 hours.

Two of the longer-term outcomes followed suit (Table 5). The ATT was −7.67% ($p = .002$) for year-to-year persistence. These students were also 5.79% ($p = .049$) more likely to dropout above the potential outcome of 33.24% if these same students were in the counterfactual state. Finally, with suggestive at best evidence, those with intense work-study hours were 4.11% ($p = .097$) more likely to stopout by June 2016 (above the potential outcome of 14.79%). In words, students who had intense work-study hours experienced deleterious effects (compared to if they had instead worked less), but these relationships were weaker for the early exit measures.

Sensitivity Analysis

The weighted estimates allow for a greater discussion of the challenges of causal inference with observational data. Selection-on-observables is a large assumption, but some sensitivity analysis can be done to see how much power an unobserved variable would need to have to explain away the relationship. Therefore, I estimate *e* values (Linden et al., 2020; VanderWeele & Ding, 2017) and Altonji ratios (Altonji et al., 2005; see also Nunn & Wantchekon, 2011), which provide the projected strength of association that an unmeasured confounder would need to have to explain away the "treatment"-outcome relationship.

The naïve (unadjusted) point estimate of the ATT could be explained away by an unmeasured confounder that was associated with both working 20+ hours and year-to-year persistence by a risk ratio of 2.16-fold each, above and beyond the measured confounders (but weaker confounding could not explain away this effect; VanderWeele & Ding, 2017). After introducing the covariates in the weighting scheme and doubly robust regression adjustment, this *e* value is slightly smaller at 1.64; the Altonji ratio was

TABLE 5

Average Treatment Effects in the Treated for Intense Work–Study (20+ Hours) With Some Sensitivity Analysis

Estimate	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
	Year-to-year persistence		Ever stopout by June 2016		Ever dropout by June 2016	
ATT	−0.142***	−0.077**	0.075**	0.041	0.177***	0.058*
	(0.025)	(0.025)	(0.022)	(0.025)	(0.026)	(0.029)
Y^1	0.865	0.798	0.112	0.148	0.212	0.332
	(0.011)	(0.018)	(0.011)	(0.017)	(0.014)	(0.023)
Y^0	0.723	0.721	0.187	0.189	0.389	0.390
	(0.023)	(0.018)	(0.019)	(0.018)	(0.023)	(0.018)
e value	2.156	1.636	1.741	1.446	2.248	1.476
Altonji ratio	—	1.178	—	1.207	—	0.486

Note. Source = High School Longitudinal Study. Data are weighted by W5W1W2W3W4PSRECORDS, modified with a ratio adjustment for having values across weekly hours worked and all outcomes. Missing values for underlying covariates are multiply imputed by chained equations (10 datasets). N = 4,418. Regression coefficients (with cluster-robust standard errors) via propensity score weighting and doubly robust adjustments for demographic characteristics, prior academic achievement, college characteristics, and family obligations. ATT = average treatment effect for the treated. Y^1 = potential outcome for the treated. Y^0 = potential outcome for the untreated. The "treatment" is intense work–study hours (i.e., 20+ hours per week). Model 1 is unadjusted and Model 2 is adjusted for all covariates. Altonji ratios are based on Altonji et al. (2005), where taking the difference between the empty model and the doubly robust model produces an implied ratio of standardized selection of unobservables to observables under the hypothesis that there is no intense work–study effect.
*$p < .05$. **$p < .01$. ***$p < .001$.

1.18. These values varied and were often weak across outcomes, suggesting that some external factor could be omitted. For instance, Nunn and Wantchekon (2011) consider Altonji ratios greater than 3.0 to be a solid threshold.

Although covariates in the literature are consistent with my models (i.e., demographics, prior achievement, financial aid, and family obligations; Bozick, 2007; Choi, 2018; Scott-Clayton & Minaya, 2016; Soliz & Long, 2016), sensitivity analyses suggests some unobserved factor of importance. I am unable to adjust for psychological stressors (such as depression, poor health, and tiredness), which are highly related to performance (Brint & Cantwell, 2010). There is also a line of scholarly work that examines the limited nature of racial categories in observational data alongside the long-standing notion of racism in society (e.g., Brantlinger et al., 2023; Castillo & Gillborn, 2022), which may impact minority student workloads and academic preparation. Others have shown strong selection into work by urbanicity (Scott-Clayton & Minaya, 2016) and smaller associations by region (Choi, 2018); adding these covariates (including whether students stayed in-state or in-region) did not change my results, so I opt for the more parsimonious models. It is possible that the locale of colleges could affect the character of jobs and the ability to commute to work more for non"traditional" students in the samples, hence the variance. To that end, without knowing the proportion of on-campus versus off-campus employment-to-living categorizations (e.g., Bozick, 2007; Ehrenberg & Sherman, 1987; Pike et al., 2008), as well as the distance/time taken to get between work and school, one could be muting relationships.

Additional Specifications and Alternative Samples (Online Supplement)

The main results show that working students performed worse overall. While a modest number of hours were good for students (with 6–9 hours being the sweet spot), working 20+ hours (especially above 26) significantly hurt postsecondary outcomes. Pell Grant recipients and first-generation students who worked more than 29 hours also had noticeably worse outcomes. In general, students at the top of the range were also more likely to exit early (for financial and/or family reasons), and this decision is influenced by the difficulty in paying tuition and shifting priorities with family circumstances.

One obvious concern is how these estimates would have fared under other conditions. Do the results differ when samples are restricted to complete cases, add in part-time students, or use alternate coding strategies? In the online supplement, I discuss a series of robustness checks and report on additional subsamples (see Supplementary Table A5 through Figure A25 in the online version of the journal). The robustness checks consider alternate coding schemes (different cutoffs for month-by-month enrollment status), various polynomial terms (e.g., orthogonal or fractional), and samples limited to listwise deletion of cases. For the subsamples, I explore the relationships by institution type, gender, race, hours worked outside of the academic year, postsecondary expectations, majors, job types, and initial performance. One could also model work as a continuous treatment with stabilized inverse-probability weights or with a dose-response function under generalized propensity scores (i.e., treating hours as "dosage" levels). These methods are also considered in the online supplement. In short, the magnitudes vary, but the results do not change appreciably, suggesting that the above interpretations are robust.

Conclusions

This study set out to answer the following question: What are the thresholds, if any, for when weekly hours worked significantly affect postsecondary outcomes? The main results in the previous section show that working up to 19 hours has little to no effect on postsecondary outcomes (and a few hours can be suggestively good). However, there are significant and deleterious effects for students working 20+ hours (and this is especially true for those working 28–36 hours). Indeed, not all outcomes react to work identically. For instance, credits earned and stopping out are the first outcomes significantly hurt by work (at 20 hours). First-year GPA (21 hours), year-to-year persistence (28 hours), dropping out (31 hours), and semester-to-semester persistence (34 hours), respectively, follow with later turn points for the worse. In other words, there is a limit for when the outcomes suffer.

The second purpose of this study was to examine how turn points differ for financial aid subgroups and first-generation students. Pell Grant recipients struggled with work more than nonrecipients. Notably, Pell students with long hours were much more likely to dropout within the first 3 years. For student loans, those who did not borrow were more likely to exit early, though there were some inconsistent patterns and cross-points. Dropout rates for nonborrowers who worked 20 hours were identical to those of Pell recipients who worked 13 hours. Borrowers working 5+ hours were more likely to return for the second year. Since loans help pay for college and optimally smooth consumption, students with the ability to borrow are set-up to succeed and return for the second year when working a modest number of hours. Finally, for first-generation college students, there was little difference between categorizations, as the sample consists only of "traditional" undergraduate students. That said, students from families with little to no postsecondary education were more likely to struggle.

The last question of this study asks if students who work more hours are also more likely to indicate that finances are the reason for their early exit. Indeed, students who stopout/dropout and work more hours are more likely to imply that finances are the reasons for their exit. For these groups (with many who are likely to be credit-constrained and skeptical of student loans post-Recession), the tradeoff is not working. Even though students who work 37 to 40+ hours have lower net tuition prices, their tuition bills are harder to pay (intensified by borrowing less). These nearly doubly full-time students might have half the tuition bill of nonworking students, on average, but they also come from a family with an income that is over half a standard deviation lower. These "decisions" to leave are also explained by their family circumstances, such as being married and having children, which is likely where the already-limited resources are being spent.

The focus on family and caution with income are indicative of post-Recession priorities. Analyzing data after the Great Recession elucidates an important update for families navigating various ways to pay for college. State funding, financial aid packages, and economic disparities have shifted since the Recession, and the public is more skeptical of higher education, especially taking out loans to pay tuition. In addition, prior research has been limited to categorical measures of hours worked and limited observed outcomes (e.g., GPA). In addition to looking at persistence-related measures, the continuous predictor of weekly hours worked illuminates specific turn points and thresholds for when work helps and/or hurts students.

Discussion

The negative turn points (showing up first) for performance make sense because credits earned and GPA are time-stamped in concert with hours worked. If students need to work long hours, academic performance signals as an initial warning to the deleterious effects. The covariates show that these students were already performing worse (and coming from less economically advantaged backgrounds), making college harder to succeed in from the get-go, hence the similar threshold for stopping out (e.g., unenrolling temporarily before changing programs/intensity). On the other end of the spectrum, semester-to-semester persistence and dropping out having the last thresholds is also understandable. Most "traditional"-aged students start in the fall semester, and there is not as much time during Christmas break versus summer break to rethink enrollment. The majority of these students also get financial aid packages set up for the full year, so there is an ostensibly lower financial risk than tuition changes year-over-year. For dropping out, much life happens between the first year on the job and that final status (as of June 2016), so this was also an outcome later-affected by work.

Analyzing work's effect on postsecondary outcomes is important to inform policy and practice. Student perceptions, higher education, and the economy of work are changing (Goldrick-Rab, 2016). Although the research in this area is extensive, the results are mixed (e.g., Bozick, 2007; Choi, 2018; Darolia, 2014; Ehrenberg & Sherman, 1987; Kalenkoski & Pabilonia, 2010; Pike et al., 2008; Scott-Clayton & Minaya, 2016). Furthermore, these studies use samples up to the early 2000s, and much of the conventional wisdom about work is based on a time different from what today's students are facing.

In this article, I explore how post-Recession students' hours worked affect their performance, persistence, and early exits. I center this study on the first year of college because this transition period requires support from high schools and higher-education institutions and sets in motion the trajectory of one's postsecondary success and adulthood. Families must also know the turn points to make an informed decision when determining the best way to pay for college. Although college is unique in its own right, first-year "traditional" undergraduates are most like high school students, where working above 20 hours has been shown to produce worse outcomes (D'Amico, 1984; Lee & Staff, 2007; Staff et al., 2010; Steinberg et al., 1982).

The results presented in the previous section complement the research of several scholars, especially those who have pointed out 20 hours as a key cutoff in college (Bozick, 2007; Choi, 2018; Dundes & Marx, 2006; Ehrenberg & Sherman, 1987). The overall trend shows that work has a curvilinear relationship with postsecondary outcomes (i.e., some work is good but too much work is bad; Dundes & Marx, 2006; Gleason, 1993; Pike et al., 2008). The sweet spot is 6 to 9 hours, when many of the outcomes were better (though insignificantly) and 1 to 4 hours was decidedly better for GPA. Nonetheless, when working 20+ hours, time on the job takes away from the time needed to navigate coursework (which is intensified by tougher starting points).

The notion of time on the job taking away from studies is one that has been previously identified (Brint & Cantwell, 2010; Dundes & Marx, 2006; Grave, 2011). Even more, scholars have long documented the link between time investment in one's studies and postsecondary performance (e.g., Andrietti & Velasco, 2015; Arulampalam et al., 2012; Ersoy, 2021; Metcalfe et al., 2019; Schwerter et al., 2022). If students take time away from studying to work (i.e., unable to put in the effort needed to academically succeed), then deleterious effects are to be expected. It is not known how much studying/time investment is picked up via high school GPA in this study, but it is clear that there are only so many hours in a day, and finding the right blend of work–study is crucial. Simply, students cannot do two things at once.

Intuitively, it makes sense that some work would be good for achievement as it provides structure, money, and flexibility. College employers want their students to academically succeed, so there is often flexibility during finals week and opportunities to study at desk jobs. In general, employers also want to hire higher-achieving students as this may reflect work ethic. Family background circumstances also come with social capital, that is, networks that connect them to suitable jobs and the transfer of information needed to get into coveted roles in a timely fashion (e.g., sending introductory inquiry emails to apply for jobs before the semester starts). Although my data show that starting points are largely associated with work–study intensity, some may argue that these students also have some stake in their education if they are working a modest number of hours (especially if they are using the earnings to pay for nondiscretionary living expenses). In other words, this modest mix provides them with enough funds to help pay for the costs of college, and they may correspondingly take their classes more seriously.

There is an obvious concern about the extent to which working on-campus versus off-campus shifts outcomes (e.g., commuting). Relatedly, students' job types—including whether they are aligned with their majors and long-term occupational trajectories—are important to consider. Future studies should detail the composition of work–study arrangements. Furthermore, my sensitivity analysis suggests that a rather small association could explain away the results, but the covariates in this study are consistent with the broader literature. In addition to exploring omitted variables (e.g., well-being) and exogenous variation, research should examine institutional engagement and individual motivation for work. Are people prioritizing work (as employees) or studies (as students) first? Do funds go to tuition or living expenses? Understanding these underlying situations could inform how colleges package financial aid. These limitations, notwithstanding, several implications can be gleaned from this study.

What Do We Do?

I suggest careful monitoring of hours for students across the board. Universities without 20-hour caps should strongly consider instituting this cutoff (at least as an average for full-time students). The federal government should also develop a plan that goes beyond its current language which counsels against repeatedly surpassing 40 hours (Federal Student Aid, 2020) to more focused monitoring of FWS. Another (all-too-familiar) recommendation is that the government should increase its support of Pell recipients, so that credit-constrained students are not forced to work long hours to pay their way through college. Still, if (low-income) students must work to serve their families because they (as Biggie said) got a "baby on the way, mad bills to pay . . . living every day like a hustle . . . another day, another struggle" (Wallace et al., 1994, 0:50), then honest and transparent conversations should be held—before working—about the tradeoffs. Students with tougher starting points might not want to quit their programs to work full-time (even if President Truman was a successful outlier), but the above evidence suggests that working long hours introduces deleterious effects on postsecondary outcomes.

Acknowledgments

The author would like to thank Stephen L. Morgan, Christopher C. Morphew, Robert Balfanz, Elizabeth A. Stuart, and Eric Rice for their early comments on this paper and conversations about this topic. In addition, this paper was substantially improved thanks to feedback from three anonymous reviewers and editor Geoffrey D. Borman.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: this study was funded by The National Academy of Education/Spencer Dissertation Fellowship.

ORCID iD

Avery M. D. Davis https://orcid.org/0000-0003-2511-1586

Supplemental Material

Supplemental material for this article is available online.

Notes

1. Net tuition price is the cost of tuition and fees minus all federal/state/institutional grants and scholarships; selectivity is based on the 2010 Carnegie Classification, which assigns SAT (or ACT equivalent) scores in the 25th-percentile greater than 1,020 to highly selective institutions (e.g., Princeton University, Johns Hopkins University, and the University of California Berkely), scores in the 25th-percentile between 850 and 1,020 to moderately selective institutions (e.g., DePaul University, Townson University, and Arizona State University), and scores in the 25th-percentile less than 850 to inclusive institutions (e.g., Liberty University, Spelman College, and Utah State University; S. Gast, personal communication, December 22, 2022).

2. Both marital and parental statuses are coded as of the students' first year, via responses on the Free Application for Federal Student Aid and the Second Follow-up; however, additional variables were coded through the end of the survey period for descriptive statistics discussed in the results section to help explain stopout/dropout reasons.

3. As discussed in Scott-Clayton (2012), surveys that ask about work during the school year (like HSLS in my study) elicit higher estimates than the Current Population Survey referenced in the literature review (background), which suggested that 41.27% of 2013 students worked. The observed 33.16% of students working 20+ hours in HSLS is also slightly higher than the 25.01% noted in March 2013 from the Current Population Survey (Flood et al., 2022).

References

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy*, *113*(1), 151–184. https://doi.org/10.1086/426036

Andrews, R. J., Ranchhod, V., & Sathy, V. (2010). Estimating the responsiveness of college applications to the likelihood of acceptance and financial assistance: Evidence from Texas. *Economics of Education Review*, *29*(1), 104–115. https://doi.org/10.1016/j.econedurev.2009.04.004

Andrietti, V., & Velasco, C. (2015). Lecture attendance, study time, and academic performance: A panel data study. *The Journal of Economic Education*, *46*(3), 239–259. https://doi.org/10.1080/00220485.2015.1040182

Arulampalam, W., Naylor, R. A., & Smith, J. (2012). Am I missing something? The effects of absence from class on student performance. *Economics of Education Review*, *31*(4), 363–375. https://doi.org/10.1016/j.econedurev.2011.12.002

Austin, P. C., & Stuart, E. A. (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine*, *34*(28), 3661–3679. https://doi.org/10.1002/sim.6607

Avery, C., & Turner, S. (2012). Student loans: Do college students borrow too much or not enough? *Journal of Economic Perspectives*, *26*(1), 165–192. https://doi.org/10.1257/jep.26.1.165

Barr, A., & Turner, S. E. (2013). Expanding enrollments and contracting state budgets: The effect of the Great Recession on higher education. *The ANNALS of the American Academy of Political and Social Science*, *650*(1), 168–193. https://doi.org/10.1177/0002716213500035

Baum, S. (2010). Student work and the financial aid system. In L. W. Perna (Ed.), *Understanding the working college student: New research and its implications for policy and practice* (pp. 3–22). Stylus Publishing.

Boatman, A., Evans, B. J., & Soliz, A. (2017). Understanding loan aversion in education: Evidence from high school seniors, community college students, and adults. *AERA Open*, *3*(1), 1–16. https://doi.org/10.1177/2332858416683649

Bozick, R. (2007). Making it through the first year of college: The role of students' economic resources, employment, and living arrangements. *Sociology of Education*, *80*(3), 261–285. https://doi.org/10.1177/003804070708000304

Brantlinger, A., Turner, B. O. N., & Valenzuela, A. (2023). Ties that bind? The teaching and post-teaching trajectories of Black and Latino/a community insiders and elite college graduates. *Educational Evaluation and Policy Analysis*. Advance online publication. https://doi.org/10.3102/01623737231162583

Brint, S., & Cantwell, A. M. (2010). Undergraduate time use and academic outcomes: Results from the University of California Undergraduate Experience Survey 2006. *Teachers College Record*, *112*(9), 2441–2470. https://doi.org/10.1177/016146811011200908

Canabal, M. E. (1998). College student degree of participation in the labor force: Determinants and relationship to school performance. *College Student Journal*, *32*(4), 597–605.

Castillo, W., & Gillborn, D. (2022). *How to "QuantCrit": Practices and questions for education data researchers and users* (Ed WorkingPaper No. 22-546). Annenberg Institute at Brown University. https://doi.org/10.26300/v5kh-dd65

Castleman, B., & Meyer, K. (2019). Financial constraints & collegiate student learning: A behavioral economics perspective. *Daedalus*, *148*(4), 195–216. https://doi.org/10.1162/daed_a_01767

Choi, Y. (2018). Student employment and persistence: Evidence of effect heterogeneity of student employment on college dropout. *Research in Higher Education*, *59*(1), 88–107. https://doi.org/10.1007/s11162-017-9458-y

Congressional Research Service. (2007, January 25). *The campus-based financial aid programs: A review and analysis of the allocation of funds to institutions and the distribution of aid to students*. https://crsreports.congress.gov/product/pdf/RL/RL32775/9

D'Amico, R. (1984). Does employment during high school impair academic progress? *Sociology of Education*, *57*(3), 152–164. https://doi.org/10.2307/2112599

Darolia, R. (2014). Working (and studying) day and night: Heterogeneous effects of working on the academic performance of full-time and part-time students. *Economics of Education Review*, *38*, 38–50. https://doi.org/10.1016/j.econedurev.2013.10.004

DeSimone, J. S. (2008). *The impact of employment during school on college student academic performance* (NBER Working Paper No. 14006). National Bureau of Economic Research. http://doi.org/10.3386/w14006

Dorans, N. J. (1999). Correspondences between ACT™ and SAT® I scores. *ETS Research Report Series*, *1999*(1), i–18. https://doi.org/10.1002/j.2333-8504.1999.tb01800.x

Dundes, L., & Marx, J. (2006). Balancing work and academics in college: Why do students working 10 to 19 hours per week excel? *Journal of College Student Retention*, *8*(1), 107–120. https://doi.org/10.2190/7UCU-8F9M-94QG-5WWQ

Ecton, W. G., Heinrich, C. J., & Carruthers, C. K. (2023). Earning to learn: Working while enrolled in Tennessee colleges and universities. *AERA Open*, *9*(1), 1–16. https://doi.org/10.1177/23328584221140410

Ehrenberg, R. G., & Sherman, D. R. (1987). Employment while in college, academic achievement, and postcollege outcomes: A summary of results. *The Journal of Human Resources*, *22*(1), 1–23. https://doi.org/10.2307/145864

Ersoy, F. (2021). Returns to effort: Experimental evidence from an online language platform. *Experimental Economics*, *24*(3), 1047–1073. https://doi.org/10.1007/s10683-020-09689-1

Federal Student Aid. (2020). *Federal student aid handbook*. United States Department of Education. https://fsapartners.ed.gov/sites/default/files/2021-02/Complete%202019-2020%20Federal%20Student%20Aid%20Handbook%20with%20Active%20Index.pdf

Flaster, A. (2018). Kids, college, and capital: Parental financial support and college choice. *Research in Higher Education*, *59*(8), 979–1020. https://doi.org/10.1007/s11162-018-9496-0

Flood, S., King, M., Rodgers, R., Ruggles, S., Warren, R. J., & Westberry, M. (2022). *Integrated public use microdata series, current population survey: Version 10.0* [dataset]. https://doi.org/10.18128/D030.V10.0

Gallagher, M. (2022, December 2). What is work-study? Everything you need to know. *Forbes*. https://www.forbes.com/advisor/education/is-work-study-worth-it/

Gleason, P. M. (1993). College student employment, academic progress, and postcollege labor market success. *Journal of Student Financial Aid*, *23*(2), 5–14. https://ir.library.louisville.edu/jsfa/vol23/iss2/1

Goldrick-Rab, S. (2016). *Paying the price: College costs, financial aid, and the betrayal of the American Dream*. University of Chicago Press.

Goldrick-Rab, S., Kelchen, R., Harris, D. N., & Benson, J. (2016). Reducing income inequality in educational attainment: Experimental evidence on the impact of financial aid on college completion. *American Journal of Sociology*, *121*(6), 1762–1817. https://doi.org/10.1086/685442

Grave, B. S. (2011). The effect of student time allocation on academic achievement. *Education Economics*, *19*(3), 291–310. https://doi.org/10.1080/09645292.2011.585794

Hawkins, C. A., Smith, M. L., Hawkins, R. C., II, & Grant, D. (2005). The relationships among hours employed, perceived work interference, and grades as reported by undergraduate social work students. *Journal of Social Work Education*, *41*(1), 13–27. https://doi.org/10.5175/JSWE.2005.200202122

Higher Education Act of 1965 [As Amended Through P.L. 117–81, Enacted December 27, 2021]. (2021). https://www.govinfo.gov/content/pkg/COMPS-765/pdf/COMPS-765.pdf

Horn, L. J. (1998). *Stopouts or stayouts? Undergraduates who leave college in their first year* (ED425683). Office of Educational Research and Improvement, United States Department of Education. https://files.eric.ed.gov/fulltext/ED425683.pdf

Ives, J., & Castillo-Montoya, M. (2020). First-generation college students as academic learners: A systematic review. *Review of Educational Research*, *90*(2), 139–178. https://doi.org/10.3102/0034654319899707

Kalenkoski, C. M., & Pabilonia, S. W. (2010). Parental transfers, student achievement, and the labor supply of college students. *Journal of Population Economics*, *23*(2), 469–496. https://doi.org/10.1007/s00148-008-0221-8

Keane, M. P., & Wolpin, K. I. (2001). The effect of parental transfers and borrowing constraints on educational attainment. *International Economic*

Review, *42*(4), 1051–1103. https://www.jstor.org/stable/826985

Lee, J. C., & Staff, J. (2007). When work matters: The varying impact of work intensity on high school dropout. *Sociology of Education*, *80*(2), 158–178. https://doi.org/10.1177/003804070708000204

Linden, A., Mathur, M. B., & VanderWeele, T. J. (2020). Conducting sensitivity analysis for unmeasured confounding in observational studies using E-values: The evalue package. *The Stata Journal*, *20*(1), 162–175. https://doi.org/0.1177/1536867X20909696

Lochner, L. J., & Monge-Naranjo, A. (2011). The nature of credit constraints and human capital. *American Economic Review*, *101*(6), 2487–2529. https://doi.org/10.1257/aer.101.6.2487

Long, B. T., & Riley, E. (2007). Financial aid: A broken bridge to college access? *Harvard Educational Review*, *77*(1), 39–63. https://doi.org/10.17763/haer.77.1.765h8777686r7357

Metcalfe, R., Burgess, S., & Proud, S. (2019). Students' effort and educational achievement: Using the timing of the World Cup to vary the value of leisure. *Journal of Public Economics*, *172*, 111–126. https://doi.org/10.1016/j.jpubeco.2018.12.006

Miller, G., & Schmidt, R. (2022, May 7). College raises student weekly work limit to 20 hours, TA positions to become hourly. *The Williams Record*. https://williamsrecord.com/460952/news/college-raises-student-weekly-work-limit-to-20-hours-ta-positions-to-become-hourly/

Morgan, S. L., & Winship, C. (2015). *Counterfactuals and causal inference: Methods and principles for social research* (2nd ed.). Cambridge University Press.

National Center for Education Statistics. (1994). *Undergraduates who work while enrolled in postsecondary education: 1989-1990*. Office of Educational Research and Improvement, United States Department of Education. https://nces.ed.gov/pubs94/94311.pdf

National Center for Education Statistics. (2021). *The condition of education 2020: College student employment*. Institute of Education Sciences, United States Department of Education. https://nces.ed.gov/programs/coe/pdf/2022/ssa_508.pdf

Newfield, C. (2016). *The great mistake: How we wrecked public universities and how we can fix them*. Johns Hopkins University Press.

Nunn, N., & Wantchekon, L. (2011). The slave trade and the origins of mistrust in Africa. *American Economic Review*, *101*(7), 3221–3252. https://doi.org/10.1257/aer.101.7.3221

Page, L. C., & Scott-Clayton, J. (2016). Improving college access in the United States: Barriers and policy responses. *Economics of Education Review*, *51*, 4–22. https://doi.org/10.1016/j.econedurev.2016.02.009

Perna, L. W. (Ed.). (2010). *Understanding the working college student: New research and its implications for policy and practice*. Stylus Publishing.

Pike, G. R., Kuh, G. D., & Massa-McKinley, R. C. (2008). First-year students' employment, engagement, and academic achievement: Untangling the relationship between work and grades. *NASPA Journal*, *45*(4), 560–582. https://doi.org/10.2202/1949-6605.2011

Radford, A. W., Velez, E. D., Bentz, A., Lew, T., & Ifill, N. (2016). *First-time postsecondary students in 2011–12: A profile*. National Center for Education Statistics, United States Department of Education. https://nces.ed.gov/pubs2016/2016136.pdf

Riggert, S. C., Boyle, M., Petrosko, J. M., Ash, D., & Rude-Parkins, C. (2006). Student employment and higher education: Empiricism and contradiction. *Review of Educational Research*, *76*(1), 63–92. https://doi.org/10.3102/00346543076001

Roksa, J., & Kinsley, P. (2019). The role of family support in facilitating academic success of low-income students. *Research in Higher Education*, *60*(4), 415–436. https://doi.org/10.1007/s11162-018-9517-z

Schwerter, J., Dimpfl, T., Bleher, J., & Murayama, K. (2022). Benefits of additional online practice opportunities in higher education. *The Internet and Higher Education*, *53*, 100834. https://doi.org/10.1016/j.iheduc.2021.100834

Scott-Clayton, J. (2011). The causal effect of federal work-study participation: Quasi-experimental evidence from West Virginia. *Educational Evaluation and Policy Analysis*, *33*(4), 506–527. https://doi.org/10.3102/0162373711421211

Scott-Clayton, J. (2012). What explains trends in labor supply among US undergraduates? *National Tax Journal*, *65*(1), 181–210. https://doi.org/10.17310/ntj.2012.1.07

Scott-Clayton, J., & Minaya, V. (2016). Should student employment be subsidized? Conditional counterfactuals and the outcomes of work-study participation. *Economics of Education Review*, *52*, 1–18. https://doi.org/10.1016/j.econedurev.2015.06.006

Shermer, E. T. (2021). *Indentured students: How government-guaranteed loans left generations drowning in college debt*. Harvard University Press.

Soliz, A., & Long, B. T. (2016). *Does working help or hurt college students? The effects of Federal Work-Study participation on student outcomes* (ED570706). Center for Analysis of Postsecondary Education and Employment. https://files.eric.ed.gov/fulltext/ED570706.pdf

Staff, J., Schulenberg, J. E., & Bachman, J. G. (2010). Adolescent work intensity, school performance, and academic engagement. *Sociology of Education*, *83*(3), 183–200. https://doi.org/10.1177/0038040710374585

Steinberg, L. D., Greenberger, E., Garduque, L., & McAuliffe, S. (1982). High school students in the labor force: Some costs and benefits to schooling and learning. *Educational Evaluation and Policy Analysis*, *4*(3), 363–372. https://doi.org/10.2307/1164066

Stinebrickner, R., & Stinebrickner, T. R. (2003). Working during school and academic performance. *Journal of Labor Economics*, *21*(2), 473–491. https://doi.org/10.1086/345565

Students who fight their way through college. (1907, February 24). *The New York Times*, p. 39. SM7. https://nyti.ms/3Q5WlPH

Toutkoushian, R. K., May-Trifiletti, J. A., & Clayton, A. B. (2021). From "first in family" to "first to finish": Does college graduation vary by how first-generation college status is defined? *Educational Policy*, *35*(3), 481–521. https://doi.org/10.1177/0895904818823753

van Buuren, S., Boshuizen, H. C., & Knook, D. L. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine*, *18*(6), 681–694. https://doi.org/10.1002/(SICI)1097-0258(19990330)18:6<681::AID-SIM71>3.0.CO;2-R

VanderWeele, T. J., & Ding, P. (2017). Sensitivity analysis in observational research: Introducing the E-value. *Annals of Internal Medicine*, *167*(4), 268–274. https://doi.org/10.7326/M16-2607

Wallace, C. G. L., Grusin, D., Mason, H., Glover, N., & Ellis, R. (1994). *Everyday struggle* [Song recorded by The Notorious B.I.G.]. Ready to Die and Bad Boy Records.

Webber, D. A. (2017). State divestment and tuition at public institutions. *Economics of Education Review*, *60*, 1–4. https://doi.org/10.1016/j.econedurev.2017.07.007

White, I. R., Daniel, R., & Royston, P. (2010). Avoiding bias due to perfect prediction in multiple imputation of incomplete categorical variables. *Computational Statistics & Data Analysis*, *54*(10), 2267–2275. https://doi.org/10.1016/j.csda.2010.04.005

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, *30*(4), 377–399. https://doi.org/10.1002/sim.4067

Author

AVERY M. D. DAVIS, MA, is a PhD candidate at the Johns Hopkins University School of Education. His research uses quantitative methods to investigate how ethnic and economic minorities navigate schools and society.

Manuscript Received February 11, 2023
Revised: May 30, 2023; August 29, 2023
Accepted October 3, 2023