Federal Reserve Bank of New York
Staff Reports

Financial Aid, Debt Management, and Socioeconomic Outcomes: Post-College Effects of Merit-Based Aid

Judith Scott-Clayton
Basit Zafar

Staff Report No. 791
August 2016

Financial Aid, Debt Management, and Socioeconomic Outcomes: Post-College Effects of Merit-Based Aid

Judith Scott-Clayton and Basit Zafar
Federal Reserve Bank of New York Staff Reports, no. 791
August 2016
JEL classification: I22, I26, J24

Abstract

Prior research has demonstrated that financial aid can influence both college enrollments and completions, but less is known about its post-college consequences. Even for students whose attainment is unaffected, financial aid may affect post-college outcomes via reductions in both time to degree and debt at graduation. We utilize two complementary quasi-experimental strategies to identify causal effects of the WV PROMISE scholarship, a broad-based state merit aid program, up to ten years post-college-entry. This study is the first to link college transcripts and financial aid information to credit bureau data later in life, enabling us to examine important outcomes that have not previously been examined, including homeownership, neighborhood characteristics, and financial management (credit risk scores, defaults, and delinquencies). We find that even as graduation impacts fade out over time, impacts on other outcomes emerge: scholarship recipients are more likely to earn a graduate degree, more likely to own a home and live in higher-income neighborhoods, less likely to have adverse credit outcomes, and more likely to be in better financial health than similar students who did not receive scholarships.

Key words: merit aid, debt management, financial health

Scott-Clayton: Teachers College, Columbia University and NBER (e-mail: scott-clayton@ tc.columbia.edu). Zafar: Federal Reserve Bank of New York (e-mail: basit.zafar@ny.frb.org).
The authors of this paper are listed alphabetically and are equally responsible for the research presented herein, which was supported by the Spencer Foundation (Grant #201500101). They are especially grateful to Neal Holly, David Bennett, and Chancellor Paul Hill of the West Virginia Higher Education Policy Commission, and to Henry Korytkowski of Equifax, for facilitating data access, and Katherine Strair of the Federal Reserve Bank of New York, for essential programming and research assistance. Angela Bell provided essential early support for the project during her time at WVHEPC. Elizabeth Mason of the New York Fed and Sandra Spady of Teachers College provided essential support with negotiating the data agreements, and Anna Wen of Teachers College provided top-notch assistance cleaning the WVHEPC data. We thank Sarah Cohodes, Raji Chakrabarti, Josh Goodman, and participants at the Association for Education Finance and Policy 2016 spring meetings, Princeton University Education Seminar Series, and Federal Reserve Bank of New York brown bag seminars for comments.

1. Introduction

As college costs have risen, financial aid has become an increasingly integral feature of the U.S. postsecondary landscape, with 7 in 10 undergraduates now receiving some form of financial aid (Baum, Elliot, & Ma 2014). At the state level, the largest expansions in financial aid spending in the past two decades have come from the introduction and growth of broad-based merit aid programs (College Board, 2012). Since 1991, several states have instituted large-scale, merit-based grant programs to defray the costs of higher education among their residents who meet basic academic criteria, regardless of financial need. For example, the West Virginia PROMISE scholarship required a high school grade point average (GPA) of 3.0 and an ACT or SAT score above the median (21 or 1000, respectively) when it was implemented in 2002. Many of these programs fully cover tuition and fees (at least initially) at in-state public institutions, and require minimal paperwork to claim. The simplicity of their eligibility and application processes, as well as their broad constituency including many middle class families, has contributed to their popularity (Dynarski & Scott-Clayton, 2006).

Nonetheless, these state merit-based programs have been controversial. Advocates point to evidence that such programs have led to improvements in college readiness metrics; increases in college enrollment and performance; improved rates of degree attainment; and decreases in the "brain drain" of talented students to other states (Bruce & Carruthers, 2014; Castleman, 2014; Carruthers & Ozek, 2013; Cornwall, Mustard, & Sridhar, 2006; Dynarski, 2004, 2008; Pallais, 2009; Scott-Clayton, 2011; Zhang & Ness, 2010). Most of this research is quasi-experimental, utilizing variation in program access due to discontinuities in eligibility criteria or in the timing of implementation. But a recent randomized-control trial of a similarly-designed private scholarship—the Buffet Scholarship in Nebraska—also finds evidence of positive effects on college persistence (Angrist, Hudson & Pallais, 2014; Angrist, Autor, Hudson, & Pallais, 2015).

Critics, however, point to important caveats, conflicting findings, and unanswered questions within this growing body of research. A pair of recent studies using Census data to examine a broader set of merit-aid programs suggest that single-state, early estimates of the impact of merit aid may overstate the impacts experienced more generally (Fitzpatrick & Jones, 2012; Sjoquist & Winters, 2012). In Massachusetts—a state with high baseline levels of educational attainment and strong private institutions—the merit-based Adams Scholarship resulted in students switching to in-state public institutions away from higher quality alternatives, ultimately reducing students' likelihood of timely degree attainment (Cohodes &

Goodman, 2014). Even when programs have generated positive effects, they may exacerbate socioeconomic gaps in attainment (Dynarski, 2000), result in unproductive strategic behavior (Cornwell, Lee, & Mustard 2005), or simply subsidize too many students who would have gone to college anyway (Fitzpatrick & Jones, 2012).

A full assessment of these programs' value relative to their cost requires knowing what happens to students after college. Yet there is little extant evidence regarding the effects of merit aid—or any type of financial aid for that matter—on post-college outcomes such as further education, employment and earnings, mobility, homeownership, and other socioeconomic outcomes. It might be reasonable to suspect that if financial aid increases educational attainment, it should improve later life outcomes as well, and if it doesn't affect attainment, then one may question the utility of looking any further. Impacts on post-college outcomes, however, do not *have* to go through impacts on enrollment and completion. It could be that marginal enrollees and completers have particularly low returns to education (although related evidence suggests otherwise; see, e.g. Zimmerman 2014). On the other hand, even for so-called "inframarginal" students whose educational attainment is unaffected by the program, more generous financial aid may influence later life trajectories by promoting faster degree completion, increasing the quality or quantity of human capital acquired during college (e.g., by improving college GPAs), or by reducing the amount of debt that students hold at graduation.

We begin to fill this evidence gap by examining the long-term effects of the WV PROMISE scholarship, which had a documented positive impact on college GPAs, credit accumulation, and degree completion after five years (Scott-Clayton, 2011). To facilitate this analysis, we construct a dataset we believe is without precedent in the literature: we link transcript, financial aid, degree completion and employment data from a state higher education agency to data from one of the nation's largest credit reporting agencies, up to eleven years after college entry (when our sample is between 28 and 30 years old). This enables us to examine important outcomes that have not been previously studied, including homeownership, neighborhood characteristics, and financial management (credit risk scores, defaults, and delinquencies). To identify causal effects of receiving the scholarship, we utilize two complementary quasi-experimental approaches: a regression discontinuity (RD) that compares students just above and below the test score cutoff for initial eligibility, and a difference-in-difference (DD) that compares eligible students before and after program implementation to ineligible students over the same time period.

Our main threat to identification is that the scholarship program induced increases in enrollments of qualified students (either due to new students enrolling in college as a result of the increased aid, some students working harder or taking the test repeatedly to score above the cutoff, and/or students choosing to attend college in-state instead of out-of-state post PROMISE). As a result, our RD analysis fails standard tests for continuity of density, and continuity of some covariates. Such violations are not uncommon in real-world applications, and economists have worked to develop reasonable strategies for causal inference even in the context of imperfect identification (Manski, 1990; Dong, 2015; Gerard, Rokkanen, & Rothe, 2015). Indeed, restricting research only to cases with seemingly perfect identification may lead to other problems including limited generalizability, selective reporting, publication bias, and lack of replicability. We address identification concerns in a number of ways. First, we use two alternative identification strategies (RD and DD), each with distinct strengths and weaknesses, and give greatest credence to results for which the RD and DD generate broadly consistent estimates across specifications. Second, we show all results under a variety of specifications, including with and without rich covariates, to assure readers our results are robust. Finally, we carefully assess the potential role of selection bias using multiple bounding strategies in which we throw out top performers (in terms of outcomes as well as in terms of covariates, separately) from our treatment group.[1]

To preview our results, we find that PROMISE recipients continue to benefit from the program more than a decade after they entered college, with modest impacts dissipating and shifting across margins over time. For example, we find that even as bachelor's degree (BA) completion impacts fade out over time, impacts on graduate school attainment emerge: recipients are 3-4 percentage points more likely to have a graduate degree after ten years (a 15-30 percent increase from baseline rates, depending upon the sample and specification). Similarly, while the program significantly reduced undergraduate borrowing, it significantly increased graduate borrowing, such that overall student borrowing was no different at the end of the follow-up period. Point estimates for earnings, for those who remain in the state and are employed year-round, are consistently positive and of an economically-meaningful magnitude ($1,500-$2,700 annually), but these estimates are noisy and not consistently significant. Scholarship recipients live in higher-income neighborhoods, have slightly better credit scores, and are generally less

[1] Our setting is quite unusual in that we have knowledge of the distribution of the running variable for cohorts that enroll after PROMISE as well as for those enrolling before. Therefore, under plausible assumptions, we can categorize the different subgroups that lead to an increase in the enrollment, and come up with informative bounds by removing groups that are likely to be positively selected (and likely to bias estimates upwards).

likely to have adverse credit outcomes (such as delinquencies and accounts in collections) than similar non-recipients, and students who just barely qualified for the scholarship appear more likely to have purchased a home. Scholarship recipients also seem to be in better financial health, as measured by an index that combines three arguably unambiguously positive financial outcomes (residing in a high-income neighborhood; not having accounts in collection; not having delinquent debt).

The paper most similar to ours – in that it also looks at the impact merit aid on post-college outcomes – is a concurrent study by Bettinger, Gurantz, Kawano, and Sacerdote (2016) which examines graduate school attainment, mobility, and earnings using data 15 years after college entry. Using data from the National Student Clearinghouse and U.S. income tax records for students just above and below the income and GPA cutoffs for a merit-based scholarship in California (Cal Grant A), they find increases in both undergraduate and graduate attainment, as well as earnings gains of about 5 percent (0.047 log points) for students just above the GPA cutoff.[2]

The remainder of the paper proceeds as follows: in Section 2, we describe the policy background and related research on the WV PROMISE scholarship. In Section 3, we describe our data sources, sample, and outcomes. Section 4 describes our approach to causal identification, and key threats to validity. Section 5 presents our main findings, Section 6 describes robustness checks, and Section 7 provides a concluding discussion and interpretation.

2. Policy background and related research on PROMISE

In 2002, West Virginia began offering PROMISE (Providing Real Opportunities to Maximize In-state Student Excellence) scholarships to approximately one-quarter of their in-state recent high school graduates (or about 40 percent of their in-state first-time freshmen). The program had multiple motivations: to reduce the cost of college, to provide incentives for increased achievement in both high school and college, and to retain more of the "best and brightest" students in-state. For the first two cohorts of recipients, who are the focus of the present analysis, graduates had to have a 3.0 high school grade point average (GPA) both overall and within a set of "core courses", as well as at least a 21 overall on the ACT or 1000 on the

[2] The Bettinger et al. (2016) study also demonstrates the importance of post-college follow up: looking at the pattern of impacts across groups in their study, impacts on undergraduate attainment do not necessarily predict impacts on longer-term outcomes. Specifically: impacts on B.A. attainment were positive and significant for recipients around both the GPA and income cutoff (slightly larger for those around the income cutoff), but impacts on graduate school attainment and earnings are only found for students around the GPA cutoff.

SAT, and had to start college within two years of high school graduation. While academic requirements have become more stringent over time, eligibility has always been based entirely on a student's academic record, not financial need.

For early cohorts, the scholarship provided full tuition and required fees for up to four years for eligible first-time freshmen who enrolled full-time at a West Virginia public two- or four- year institution, or an equivalent amount to attend an eligible West Virginia private institution (for later cohorts, the scholarship was capped at a fixed amount). To renew the scholarship, undergraduates had to complete at least 30 credits per year and maintain a 3.0 cumulative GPA, although the first two cohorts were allowed a 2.75 GPA in their first year. The average value of the award in 2002-03 was $2,900 for the first year (over $3,700 in 2016 dollars). Those who initially qualified received about $10,000 ($12,300 in 2016 dollars) on average over four years.[3] Scott-Clayton (2011) analyzed the impact of the scholarship program on college outcomes for the first two cohorts of recipients, using two complementary quasi-experimental approaches to identify causal effects: a regression-discontinuity (RD) analysis based on the ACT score threshold for PROMISE eligibility, and an event-study analysis based on the discontinuous timing of program implementation. Focusing on the event-study results (which provide a broader estimate of program effects), she found that PROMISE receipt increased GPAs and credit completion rates particularly during the first through third years of college, culminating in a 6.7 percentage point increase in bachelor's degree completion rates after four years (from a baseline rate of just 26 percent). The BA completion results faded to a 3.7 percentage point impact after five years, though remained statistically significant. The RD analysis suggested even larger impacts for students near the initial eligibility cutoff. Additional analyses suggested the program may have reduced student employment and student debt as well.

Questions remain, however, about the longer term effects of the program. Policymakers have questioned whether recipients may be more likely to leave the state after graduation. In addition, the shrinking of the BA completion impact between years four and five begs the question of whether it might ultimately fade out completely, and if so, whether improving time to degree alone is enough to substantially affect post-college outcomes.

On the other hand, the scholarship could still impact post-college outcomes even if the BA completion impact eventually fades out. Even those whose graduation status was completely unaffected may graduate with less debt as a result of the program, which may give them an

[3] This average includes students who failed to renew the scholarship for all four years.

advantage financially. Graduating a year or more earlier than a student would have otherwise not only gives the student an experience advantage in the labor market (or a head start on graduate school), it may further reduce students' debt at graduation. In addition, if the GPA impacts of the program represent real human capital gains, this could show up as an advantage in either the labor market or graduate school admissions.

3. Data, Sample, and Outcomes

Data. Our data come from two primary sources: the West Virginia Higher Education Policy Commission (WVHEPC), a state agency that maintains a comprehensive database on the state's public college enrollees, and Equifax Inc., one of the three main consumer credit reporting bureaus in the United States. WVHEPC provided de-identified data on four cohorts (2000-01 through 2003-04) of new public college entrants under a restricted-use data agreement. The data include limited background information such as age, race, gender, overall high school GPA, and ACT and SAT scores as reported on the college application.[4] No direct measure of family income or wealth is available for the full sample, but we observe the student's county of residence at entry. The data include complete college transcripts and records of financial aid receipt for all cohorts for ten years after initial enrollment (note: financial aid application data from the FAFSA is only available for our post-cohorts, 2002 and later). The data also include administrative records of quarterly employment and earnings for students who worked in-state, acquired by WVHEPC from the state's Employment Security agency which uses it to administer unemployment insurance (UI), for up to 10 years after college entry. The data are the same as used in Scott-Clayton (2011), but updated to include more years of follow-up.

We use these data to examine several outcomes of interest: long-term degree completion, including both bachelor's and graduate degree attainment; long-term student loan accumulation, separately for undergraduate, graduate, and parent loans; and in-state employment outcomes including whether individuals were employed in-state at all 10 years post-entry, whether they were employed year-round, and annual earnings conditional on year-round employment. "Employment" is simply defined as non-zero earnings in any quarter of the year, while year-round employment requires non-zero earnings in every quarter of the year. Note that it is typically not possible to distinguish non-employment from out-of-state mobility in state UI databases. While in-state employment is of primary importance to state policymakers, our

[4] We have at most one ACT or SAT score. We presume students report their highest score at the time of the application.

merged data allow us to directly determine which students ultimately moved out of state (we find that approximately two-thirds of those who are not employed in WV are indeed living out of state).[5]

The Equifax data provide a longitudinal panel of the agency's individual consumer credit files from 2005 to 2014, with one observation per year. The data are de-identified and include no demographic information other than birth year and geographic location of the file holder's residence at the zip code level. By definition, the sample only consists of those with credit reports, and it includes information on debt accounts including their type, balance, and status. We use a number of consumer debt metrics as our outcome variables. We look at the Equifax risk score of the individual. This risk score is similar to the FICO score, in that both model 24 month default risk as a function of credit report measures. It varies between 280 and 840 and represents an assessment of the individual's credit-worthiness. Using the panel, we construct several measures of the individual's repayment behavior. These include an indicator for whether the individual has ever had a delinquent student, auto, or home loan account, where delinquency is defined as a debt payment that is reported as 30 or more days past due, and an indicator for ever having had an account in collections.

Exploiting the panel nature of the dataset, we also study whether the individual ever had any housing debt (indicative of home ownership). In a sample of consumers in their twenties, any history of home-secured debt is a reasonably complete proxy for past or present homeownership. Few homeowners this young own their homes outright. The panel is also exploited to study whether the individual has ever taken out a student loan (note that the Equifax loan measure is broader than the one constructed using financial aid data from WVHEPC, which includes only federal student loans taken out for enrollment at WVHEPC institutions). The maximum mortgage amount at origination, credit card balances (debt that is typically used to support consumption), and student loan balance are other outcomes that we also analyze.

Finally, using the Zip code information on the credit report, we construct a dummy variable for whether the individual resides outside West Virginia (WV) at a given point in time. Our analysis also uses Zip code-level income data for 2010, drawn from the IRS Individual Income Tax Statistics zip code data. The zip code level income information can be interpreted as a proxy for neighborhood quality or socioeconomic status, that does not rely on individuals being

[5] We do not make full use of this information. Our employment outcomes are currently estimated for the full WVHEPC sample, without conditioning on state of current residence as indicated in the Equifax data. Our earnings outcome is conditional on year-round employment in the state.

homeowners to measure. In addition, this outcome may be less sensitive to the turbulence in the housing market during the Great Recession.

Several of the credit report outcomes have ambiguous implications for the consumer's well-being. For example, homeownership or higher levels of consumer debt may be desirable in certain states of the world and undesirable in other. To study the impact on the consumer's financial well-being, we construct an index based on three outcomes that we believe are unambiguously positive: residing in a zip code in the top quartile of the income distribution, never having past due loans, and never having debts in collections. Each of the three outcomes are indicators that take the value 1 if that is the case, and zero otherwise. The index hence varies on a scale of 0 to 3, with higher values indicative of better financial health.

The financial outcomes are all measures of credit activity on the intensive margin (that is, conditional on having a credit report). Our treatment (receipt of the scholarship) may also impact the likelihood of the individual having any credit activity and, thus, a credit report. Being matched with Equifax data is, therefore, another outcome of interest that we analyze.

The individual credit report data have been analyzed in several papers (Mian and Sufi, 2011; Brown et al., 2015). Due to limited demographic information on credit reports, these studies have been constrained to exploiting (arguably exogenous) variation at the geographic level to identify causal impacts. Linking credit bureau data to other proprietary datasets is fairly rare- there are instances where credit bureau data have been matched with other financial data (such as Bhutta, Skiba, and Tobacman, 2012), but we are unaware of any prior efforts to match a postsecondary education dataset to these individual credit reports.

Matching. Executing the data match required six separate data agreements between the researchers, WVHEPC, the Federal Reserve Bank of New York (FRBNY), and Equifax. To facilitate the match without compromising data security, the research team facilitated a multi-step match. First, WVHEPC provided the research team with de-identified administrative data, similar to what was provided for Scott-Clayton (2011), containing only a random scrambled identifier. Then, Equifax received a crosswalk file containing only the scrambled identifier and actual identifiers needed to conduct the match, with no other variables included. After conducting the match, Equifax stripped the file of all identifiers except the scrambled identifier and transferred the file to a secure location at FRBNY. The research team then matched the de-identified Equifax and de-identified WVHEPC files using the dummy identifier.

Sample Description. Table 1 provides descriptive statistics on our sample. The first column, for comparison, provides statistics for first-time college entrants in the nationally-representative Beginning Postsecondary Students (BPS) 2003 sample. The second column provides summary statistics for all young (19 or under), WV-resident entrants in the WVHEPC data. The third column describes our RD sample: those entering in 2002 or 2003, with at least a 3.0 high school GPA and an ACT score (or SAT equivalent) between 16 and 25. The final column describes our DD sample: those with at least a 3.0 high school GPA, entering in the two years before and after PROMISE began (2000-2003 cohorts). The gender composition of our WV sample is comparable to national statistics, but unsurprisingly, the WV sample is exceedingly white relative to enrollees nationally (95 percent versus 62 percent in national figures). The WV sample, by construction, is younger than the typical pool of college entrants (with an average age of 19 versus 22). It is not an economically advantaged sample, however. The rate of Pell receipt in our RD sample is virtually identical to the national average (37 percent, compared to 36 percent among young entrants nationally), and only modestly lower in the higher-scoring DD sample (32 percent). About 80 percent of our sample overall (or 90 percent of PROMISE eligible students) initially enrolled at a four-year institution while the remainder started at a two-year institution (not shown).

Description of outcomes. Table 2 provides statistics on our various outcomes of interest. The top panel shows academic, loan, and employment outcomes based on the WVHEPC data. The national comparison group here is the 2003 BPS sample. The WV samples have notably higher than average BA completion rates, likely due to the younger average age of our samples (19 versus 22 among first-time entrants nationally). One notable finding from this table is just how much degree completion continues to increase even in the later years of follow up. Bachelor’s degree completion doubles between four and six years post-entry in our DD sample (from 29 to 60 percent), and more than doubles in the RD sample (from 21 to 50 percent). But it continues to increase through ten years post entry, to 66 percent in the DD sample and 57 percent in the RD sample (implying that at least 1 in 10 graduates takes more than 6 years to complete – the maximum follow up of the BPS survey). The WV samples have somewhat higher-than-average levels of undergraduate borrowing (53 percent had any federal undergraduate loans after 6 years in the BPS data, compared to 56-58 percent after 5 years in the WV data). Average amounts borrowed are also higher in the WV data. Again, this is likely a consequence of our sample being younger than average, with more students pursuing bachelor’s degrees than among all first-time entrants nationally. About 60% of the full WVHEPC sample had some in-state earnings in the 10^{th} year post-entry, but only 47% had earnings in all quarters of that year. For those working in-state year round, average earnings were $41,510. The lower panel of Table 2 shows Equifax outcomes. The national comparison column here reports the statistics for a national sample of individuals of similar age as our sample (28-32 in 2014). The reported statistics for the outcomes here are based on the 2014 data, except those which use the entire history of the individual (such as, "ever" past due). Roughly 92 percent of the WV sample is matched with the credit bureau data. This match rate compares favorably with the coverage of credit bureau data for 28-32 year olds nationally.

Nearly a quarter of WV enrollees had ever lived outside WV at some point during 2005-2014. Mean zip code incomes and the rate of residing in zip code in the lowest quartile of the (national) income distribution are similar for our sample and the national comparison group. Homeownership, as proxied by ever having any housing-related debt, in our sample varies between 35% and 40% depending on the subsample that one looks at, substantially higher than the national rate for this age group over this period. Maximum mortgage log balances, conditional on having a mortgage, are however on average similar for the WV and national samples. Average credit card balances for the WV sample are $2,120, somewhat higher than the national average.

In the full WV sample, 40.2% of the enrollees have had a delinquent (student, auto, or housing) debt at some point, and nearly half have had an account in collection. These statistics vary across the subsamples, and are higher than those of the national sample. Roughly 62% of our sample ever takes out a student loan, with an average maximum student loan balance of $16,400. Both statistics are higher than the national average, which should not be surprising since the national sample does not condition on pursuing post-secondary education.

The average credit score in our sample is 663, higher than the corresponding national average. Our index of financial health takes an average value of 1.2 (with a standard deviation of 0.996) in the full sample, quite similar to the average for the national sample.

In addition, for the female respondents, revisions to our initial matching procedure inadvertently resulted in a noisy proxy of "marriage." The initial matching of the WVHEPC files with the Equifax data was done by year based on last names, which revealed the surprising pattern that coverage rates for females decreased with age (when in fact the opposite is expected, since credit bureau data's coverage increases with age as the propensity to enter credit markets increases in early adulthood). An investigation of this puzzling pattern revealed that females who matched in one year were no longer considered as matches in a subsequent year if their names changed, even though Equifax could identify them as the same person based on other identifying information. This led Equifax to revise their algorithm. But we use this information to construct a proxy for marriage. Needless to say, that this is a noisy measure, since females need not change their last name upon marriage. Just over one-third (35.5%) of females in our sample are coded as ever being married under this definition, far lower than the actual marriage rate observed for similarly aged WV residents in national data sources.[6]

4. Approach to Causal Identification

Overview. Following Scott-Clayton (2011), we utilize two complementary quasi-experimental strategies to identify causal effects of PROMISE receipt: the first is a regression-discontinuity (RD) that estimates the effect of being just above rather than just below the test score threshold for initial eligibility; and the second approach is a difference-in-difference (DD) comparing eligible students before and after program implementation, with ineligible students as a comparison group.[7] For both approaches, we layer on an instrumental variables approach to address the issue that we do not observe all of the information needed to precisely determine PROMISE eligibility.[8] We thus use estimated eligibility, based on high school GPA and ACT/SAT scores, as an instrument for actual receipt (which is observed). Our first stage is very strong under both approaches: estimated eligibility increases the likelihood of actual receipt by 70 to 80 percentage points.[9] As such, the IV serves primarily to correct for measurement error in our treatment indicator rather than as an identification strategy per se. We describe each of our two main strategies in more detail below, followed by a discussion of key threats to validity.

Regression discontinuity specifications. For this analysis, we limit the sample to West Virginia residents entering in the first two years after PROMISE implementation who earned at least a 3.0 high school GPA. For these students PROMISE receipt is largely determined by ACT score (or SAT equivalent), though grades in high school "core courses" is another factor

---

[6] American Community Survey data for 2013 indicate ever-marriage rates of 50-60 percent for WV residents between age 28 and 30 with at least some college (authors' estimates using publicly available data).

[7] Note that Scott-Clayton's (2011) preferred second approach was a simple event study analysis comparing eligible students before and after implementation, though she also included a DD specification as a robustness check (the academic impacts she examined were actually larger in the DD than the event study estimates). We think an event study analysis is too simplistic to rigorously estimate the labor market, housing, and credit outcomes examined here, however, given that our follow-up period spans the Great Recession. It is much more plausible to assume no substantial cohort effects on GPAs, credits, or graduation rates during 2004-2007 than it is to assume no substantial cohort effects on employment, homeownership, or credit delinquencies during 2008-2013.

[8] Specifically, we have only an overall high school GPA as reported on the college application, though PROMISE eligibility also required a 3.0 in a set of "core courses." In addition, because GPA and test score data are self-reported from the college application, it is possible that they may change in between the time of college application and the time of PROMISE eligibility determination.

[9] Also note that the difference between estimated eligibility and actual receipt is unlikely to be explained in any large part by imperfect take-up for our sample of WVHEPC enrollees. The program was heavily advertised and the application itself was minimal; students who might not have been aware initially are likely to have learned of their eligibility during the college application/registration process.

which we do not observe. The vast majority of those who score a 20.50 (and thus are rounded to a score of 21) have access to the program while those who score only 20.49 do not. Except for PROMISE, students scoring just above 20.5 should not systematically differ from those scoring just below. If this assumption holds, then one can examine outcomes by ACT score and attribute any discontinuous jumps at the threshold to the effects of PROMISE.

Following Imbens and Lemieux (2008), our main specification utilizes a two-stage local linear regression specification in which we first predict PROMISE receipt using the test score discontinuity, and then estimate the effect of predicted receipt on a given outcome:

(1a) $P_i = \lambda + \psi(above_i) + \gamma(ACTdist_i * below_i) + \varphi(ACTdist_i * above_i) + X_i\phi + \varepsilon_i$

(1b) $y_i = \alpha + \beta(\hat{P}_i) + \varsigma(ACTdist_i * below_i) + \pi(ACTdist_i * above_i) + X_i\delta + \varepsilon_i$

where P_i represents actual PROMISE receipt, $\hat{P}_i$ represents predicted PROMISE receipt, $above_i$ is an indicator that the student is above the score threshold, $below_i$ is an indicator that the student is below the threshold, $ACTdist_i$ is the distance between the student's individual score and the underlying cutoff score (20.5), X_i is a vector of covariates including gender, race/ethnicity, age, high school GPA and high school GPA squared, high school type, and county of residence at entry fixed effects, and ε_i is an idiosyncratic error term.[10] The parameter β estimates the difference in outcome y_i at the threshold. In practical terms, this IV-RD specification provides essentially identical results to what we would get by running a simple RD specification for all outcomes (using equation 1a), and simply scaling the resulting estimates up by a factor of 1.43 (i.e., 1.00/0.70) to account for the fact that crossing the ACT threshold increases PROMISE receipt by 70 percentage points.

To ensure that our results are driven by the RD specification itself and not by differences in covariates around the cutoff, we test our main specification with and without covariates included. Our main specification focuses on a bandwidth +/- 5 score points (16<=ACT<=25), but we test sensitivity to narrower and wider bandwidths. We also test out a specification that uses the full range of test scores, with additional quadratic controls for ACT score distance that are allowed to vary above and below the score cutoff. Finally, as a falsification test, we run our preferred specification for the two cohorts that entered before PROMISE was available.

[10] Lee and Card (2008) suggest clustering standard errors by values of the forcing variable (ACT score, in this case) when the forcing variable is discrete rather than continuous. This procedure is not clearly an improvement here due to the unbalanced size and small number of clusters (10) and does not, in practice, necessarily result in larger standard errors.

Difference-in-difference specifications. For this analysis, we limit the sample to entrants with at least a 3.0 high school GPA, and compare changes over time among those above the test score cutoff (who were eligible for PROMISE after 2002) to those below the test score cutoff (who were never eligible). We again utilize a two-stage model in which estimated eligibility (being above the threshold, after implementation) is the instrument for actual PROMISE receipt. Specifically, we estimate the two-stage model:

$$\text{(2a)}\ P_{it} = \lambda + \gamma(ABOVE_i * AFTER_t) + \theta(ACTFE_i) + \vartheta(COHORTFE_t) + X_i\phi + u_{it}$$

$$\text{(2b)}\ y_{it} = \alpha + \beta(\hat{P}_{it}) + \xi(ACTFE_i) + \tau(COHORTFE_t) + X_i\delta + \varepsilon_{it}$$

where P_{it} represents actual PROMISE receipt, $\hat{P}_{it}$ represents predicted PROMISE receipt based on the parameter estimates from (2a), *ACTFE* represents fixed effect controls for ACT score (a more flexible way of controlling for the main effects of being above versus below the score cutoff), *COHORTFE* represents fixed effects controls for college entry cohort (a more flexible way of controlling for the main effects of entering before versus after implementation), and X is a vector of covariates as previously defined.

As with the RD, we test our main specification with and without covariates to ensure that our results do not rest upon their inclusion. Our main specification uses heteroskedasticity-robust standard errors. However, we also test alternative ways of computing standard errors, including running our analyses with data aggregated to the ACT-cohort level.

Finally, we also run a version of the DD that limits the sample to students just above and below the score cutoff (19<=ACT<=22). Note that the RD and DD not only rest upon different core assumptions, but also generate different estimands (the former is local to the lowest-scoring PROMISE recipients, while the latter estimates the average effect for all recipients). Thus, this specification provides a point of translation between the RD and DD findings: it applies the DD specification to estimate impacts for recipients near the threshold.

Threats to validity. The RD and DD approaches are stronger together than either would be alone. The advantage of the RD is that it tightly links any observed impacts to an arbitrary program rule, eliminating several alternative explanations for the findings. Neither institutional policies, labor market conditions, nor students' background characteristics should vary discontinuously around the ACT threshold. One concern, though, is that the cutoff was well-known and thus students could influence their eligibility by working harder and/or retesting in order to meet the scholarship criteria. An additional limitation is that the RD provides local average treatment effects (LATEs) that apply only to those near the eligibility threshold, who represent the lowest-scoring 20 percent of all PROMISE recipients and who may differ from other students in their response to the program. In contrast, the DD does not suffer from the same test manipulation concern, and also provides average treatment effects (ATEs) across all recipients, not just those near the threshold. However, the DD estimates may be more sensitive to dramatic changes in economic conditions during our follow-up period (roughly 2007-2013), if these changes affect eligible and ineligible students differentially.

In addition, as discussed at length in Scott-Clayton (2011), both approaches are potentially contaminated by some selection bias given that our sample includes only those students who enrolled in the WV public college system, when the enrollment decision itself could have been affected by scholarship eligibility. Bias, not always in the same direction, could arise from three sources: 1) eligible individuals who otherwise would have attended college out-of-state may now choose to enroll in-state (potentially inducing upward bias in both the RD and DD analysis), 2) eligible individuals who otherwise would not have enrolled in college may now choose to do so (likely to negatively bias both analyses), and 3) individuals who would have enrolled in college but failed to meet the eligibility criteria could work harder or retest in order to reach them (likely to negatively bias the DD analysis, but bias is ambiguous in the RD).

As in Scott-Clayton (2011), we do find evidence of significant heaping just above the ACT threshold for PROMISE eligibility (see Figure 1). Based on a comparison of frequency distributions before and after PROMISE, we clearly see excess observations above the threshold after PROMISE (compared to before). However, the majority of this additional density above the cutoff is due to new enrollments, rather than to students shifting from below to above the cutoff. And as we will show later, less than a fifth of these new enrollments can be explained by students shifting from out-of-state to in-state colleges, the most plausible source of upward bias. The net result of these enrollment shifts is a number of small, but statistically significant differences in covariates around the cutoff, predominantly in the IV-DD specification (most notably, treated students have 0.03 higher high school GPA in the RD, 0.12 points higher in the DD; see Appendix Table A.1 for McCrary test and complete covariate checks).[11] The simplest way to address our partial identification challenge is simply to control for any observable differences between the treated and control groups, but bounding strategies can also be used to assess concerns about selection.

[11] To test covariate balance, we run our basic IV-RD and IV-DD specifications (with no covariates on the right hand side) but with each covariate in turn treated as an outcome, on the left side of the equation. Although we detect a number of statistically significant differences in the IV-DD specification, the magnitude of the coefficients is small.

Scott-Clayton (2011) undertakes a number of sensitivity analyses, including a bounding analysis, to explore the severity of these selection concerns. She finds that selection bias cannot explain the positive academic impacts of the program even under extreme assumptions about marginal enrollees. In addition, she finds that effects on students' annual GPA and credit attainment are concentrated precisely around those margins most directly connected to the scholarship's renewal requirements, and are evident only in the first three years of college, when students face these renewal incentives[12] – a pattern we would not expect if the impacts were a spurious result of selection. Still, there is no guarantee that results for these new outcomes will be similarly robust.

Thus, after presenting our main findings, we examine these selection concerns in more detail and provide robustness checks including three bounding analyses. In the first, we assume all "excess enrollees" are top performers, and exclude them from our treated group. In the second, we trim only the fraction of excess enrollees that are most likely to generate positive bias. In the third, we exclude observations in a narrow region around both sides of the cut-off (a "donut hole" analysis), with the idea that the running variable is more likely (if at all) to be manipulated in that region. The preferred approach we take in our main analysis, however, is to address compositional bias in a straightforward fashion by adjusting for a rich set of covariate controls, including two of the best predictors of college success—high school GPA and ACT score—as well as gender, race/ethnicity, age at entry, high school type (public/private), and fixed effects for county of residence at entry. Differential sample selection remains a concern only to the extent it occurs on other unmeasured dimensions that are also correlated with post-college outcomes.

5. Main results

For the academic and student loan outcomes, we pick up where Scott-Clayton (2011) left off, examining students both at 5-6 years post-entry and again at the end of the follow-up period, 10 years after college entry. For the in-state employment outcomes as well as for the Equifax-based outcomes, we focus on realizations 9 to 11 years after enrollment. In the RD analysis, this means focusing on outcomes during 2011-2013 (2012-2014) for the 2002 (2003) entering cohort, implying that the average age of the respondents would be 28 to 30 years. Certain outcomes,

[12] Since students can only receive the scholarship for four years, students face no renewal incentives in the fourth year of receipt. If program effects were driven by spurious selection, there is no reason we would expect them to suddenly disappear in the fourth year of college. See Scott-Clayton (2011) for additional details.

including graduate degree completion, cumulative loan borrowing, ever past due or owning a home, take the respondent's entire history into account up to that point in time.

Regression discontinuity results. RD results for the WVHEPC-based outcomes are provided in Table 3, with our preferred specification in column (1). Note that all models shown are two-stage, RD-IV models that use estimated PROMISE eligibility as an instrument for actual receipt (with a first stage of 0.70).[13] The first line simply replicates one of the key findings from Scott-Clayton (2011) that PROMISE receipt increased on-time graduation for students near the score threshold dramatically: a nearly 10 percentage point increase off a base of just 21 percent. Over time, however, this effect fades out as non-recipients eventually catch up. After 10 years there is no longer a significant difference in BA completion.

As the BA completion results fade out, however, other impacts fade in. PROMISE recipients are significantly more likely to earn a graduate degree: a 2.6 percentage point impact after 6 years grows to 4.2 percentage points after 10 years (off base rates of just 7 and 13 percent, respectively). These results are quite robust across specifications, although they lose significance in some alternative models.

The student loan borrowing outcomes reveal a very interesting pattern: while PROMISE recipients have significantly lower undergraduate borrowing (both in terms of ever borrowed and cumulative amounts borrowed) and cumulative parent loan amounts (only relevant for undergraduates), they also have substantially higher graduate borrowing. After 10 years, this nets out to an overall 7.4 percentage point decline in the likelihood of having ever taken any federal student loan, but no difference in cumulative amounts. Like the graduate degree attainment findings, the student and parent loan impacts are highly robust across specifications. The effects on parent borrowing are particularly notable because parent borrowing generally is much smaller than student borrowing. The RD estimates in column (1) suggest parent borrowing was reduced by about two-thirds, by over $1000 off of an average of just over $2000.

We find no statistically significant impacts on in-state employment rates or earnings conditional on year-round employment in our main specification. In general, the point estimates are close to zero for employment rates but modestly positive for conditional earnings. The conditional earnings estimates are significant in some alternative specifications, and if taken

[13] OLS versions of these RD models provide the same pattern of findings but with magnitudes of approximately 70% of what is shown here.

seriously, would represent an increase of about 7% over the mean earnings of $40,000 in the RD sample.[14]

Reassuringly, the falsification results (which implement an RD based on the ACT score cutoff, but prior to the existence of PROMISE) reported in the last column of Table 3, find no significant effects on any of these outcomes, with point estimates generally close to zero.

RD results for the Equifax-based outcomes are presented in Table 4. The first row looks at the impact of predicted Promise receipt on the extensive margin of participating in credit markets. Focusing on our preferred specification in column (1), individuals above the cut-off are not any more likely to be matched than those below the cut-off. This lack of impact on the extensive margin makes it easier to interpret the intensive margin results that we next look at. We see no impact of the scholarship receipt on the propensity to ever live outside WV. We find a robust positive impact of predicted PROMISE receipt on mean zip code-level income and a negative impact on the propensity to live in a zip code that is in the bottom quartile of the adjusted gross income (AGI) distribution (a decline of 4.3 percentage points on a base of 15.5%). We find no impact on the marriage proxy.

Turning to financial outcomes on the intensive margin, predicted scholarship receipt leads to a (economically and statistically) significant 6 percentage point increase in the likelihood of home ownership (on a base of 36.1%). Conditional maximum mortgage log balances (that is, conditional on having housing debt) are on average lower by 1 log point, but imprecisely estimated. Average credit card debt – debt that is a good proxy for consumption – is unaffected. Predicted scholarship receipt leads to a significant decline of 5 percentage points in the likelihood of ever having student loans (on a base of 62.2%) 9-11 years after enrollment. Student loan balances, however, are higher by nearly $2,800 (though not precisely estimated), consistent with scholarship receipt leading individuals to either persist through college or pursue post-bachelors' degrees at higher rates. Conditional student loan balances are, however, much higher for recipients. The significant estimate of $7,700 is considerably larger than the impacts on amounts borrowed based on the WVHEPC data in Table 3. This could reflect differential repayment patterns (particularly if PROMISE recipients defer repayment during graduate enrollment), or it could reflect additional loans accrued at educational institutions outside the WV public education system or due to private loans, none of which would is captured in the WVHEPC borrowing data.

[14] This is quite consistent with Bettinger et al. (2016), who found a 4.7% impact on post-college earnings for students who just met the GPA threshold to qualify for a Cal Grant A scholarship.

In terms of credit performance, we observe overall improved outcomes relative to those who are predicted to not have received the scholarship. Specifically, we see a lower propensity to be past due on accounts or have any in collections. Predicted scholarship receipt leads to an average increase of 3.8 points in individuals' credit score. While none of these outcomes are significant at conventional levels, the signs of coefficients are indicative of improved credit standing and financial outcomes among scholarship recipients. This is also reflected in the positive impact on our index of financial well-being (though again the estimate is imprecise).[15]

Columns (2) through (5) of the table report various robustness checks. They yield similar qualitative conclusions. Finally, column (6) of the table reports estimates from a falsification exercise, where we run the RD specification on the 2000 and 2001 cohorts that enroll prior to the implementation of the PROMISE scholarship. None of the estimates that are significant in the baseline specification are found to be significant in this placebo test except on the propensity to live in a zip code that is in the bottom quartile of the income distribution (though in the placebo this result switches sign).

Difference-in-difference results. Difference-in-difference estimates for the WVHEPC-based outcomes are presented in Table 5. Again, we use a two-stage approach in which estimated PROMISE eligibility is used as an instrument for actual receipt (with a first-stage of 0.79). Our preferred model is shown in column (1). Note that the DD estimates impacts for all PROMISE recipients, not just those near the score cutoff. As a result the differences between the RD and DD findings may reflect true heterogeneity in impacts for students depending upon ability (as proxied by test scores). The model in column (3), which applies the DD specification to a narrow subset of students around the threshold, is our attempt to provide a crosswalk between the two empirical approaches. If these results mimic the RD it would suggest that discrepancies between the RD and DD are due primarily to true impact heterogeneity rather than to differences in identification assumptions. However, the results in column (3) are much noisier and not always easy to interpret: in some cases they do appear more similar to the RD findings, while in other cases, they appear more similar to the DD.

As in the RD, the DD suggests large increases in on-time BA completion, but for the DD sample the effects remain large and significant even 10 years after entry (an increase of 8 percentage point from a baseline rate of 66 percent). The DD results also indicate significant increases in graduate school borrowing and attainment, although the magnitudes are somewhat

[15] We also can examine the Equifax outcomes at earlier points in time. Doing so does not change the overall pattern of results.

smaller and more sensitive to the inclusion of controls than was the case in the RD. The DD suggests no substantial reduction in undergraduate borrowing (except when the sample is limited to students near the score threshold), and smaller reductions in parent borrowing than in the RD.

Finally, as in the RD, the DD suggests no significant impact on employment rates or conditional earnings, with point estimates that are generally smaller in magnitude than in the RD.

Estimates of the difference-in-difference specifications for the Equifax-based outcomes are presented in Table 6. Results for the credit performance measures match those from the RD specification and are stronger, well-summarized by an approximate average increase of 6.5 points on credit score in the baseline specification, alongside a 3% lower likelihood of having ever had a past due student, auto, or home loan, or an account in collections. Predicted PROMISE receipt leads to a 0.06 increase in the index of financial health. This is a sizable impact, equivalent to nearly 0.06 of the standard deviation in the index. The predicted impacts of the scholarship on average zip code level income and on residing in the bottom quartile are similar to those in the RD specification.

A few outcomes lose precision or even reverse signs relative to the RD specification. While the RD showed an increase in the likelihood of homeownership, here we see no impact on the ownership margin. In addition, we do not see any impact on the likelihood of taking out student loans (while the RD showed a decline). Column 3 of the table shows the estimates of the DD that limits the sample to students just above and below the score cutoff (19<=ACT<=22). Estimates from this specification should, in principle, be most comparable to those from the RD specification, though they rest on different assumptions. We do see that the estimates of the DD restricted to those individuals around the cut-off are qualitatively similar to those from the RD, except for the impact on home ownership.

Graphical checks. Graphical analyses of selected outcomes are presented in Figures 2 and 3. We focus on those outcomes for which there appear to be significant and reasonably robust impacts in the regression specifications. The graphs can help provide intuition regarding where the underlying variation comes from and thus help readers to evaluate the plausibility of the regression results. In general, these graphs are reassuring: for outcomes for which we consistently find impacts, across specifications, we do see visual evidence of impacts in the graphs. The graphs show estimates for the baseline RD specification as well as the falsification RD. Take, for example, the impact on cumulative graduate school loans by the end of year 10 in the lower panel of Figure 2. We can see a clear discontinuity for the baseline RD (denoted by "After Promise"). The falsification RD ("Before Promise") does not exhibit a discontinuity at the

ACT score threshold, as one would expect. Note that the DD implemented on students around the threshold is effectively the difference between the After Promise and Before Promise lines, above the cutoff and below the cutoff. Similarly, the bottom panel in Figure 3 shows a clear discontinuity in mean zipcode income for the baseline RD specification, and no jump for the falsification RD.

Subgroup analyses. We examined results for our main RD and DD specifications for three sets of subgroups: by Pell eligibility status, by gender, and by whether the student came from a public or private high school. These results are presented in full in Appendix Tables A.2-A.4. In many cases (particularly when cutting the sample by type of high school), sample sizes become quite small and estimates become noisy—we are thus underpowered in terms of detecting subgroup differences. This lack of power is reflected in the fact that it is difficult to discern consistent patterns across outcomes and across the RD vs. DD specifications, and makes us hesitant to put much weight on their interpretation. However, the results generally suggest that the effects may be stronger for students who are Pell recipients and those who attend public high schools. Since we might expect test re-taking and out-of-state migration to be most likely among higher socioeconomic groups, it is reassuring that our positive results do not appear to be limited to these groups.

6. Robustness checks

Alternative standard errors. One concern is that our standard errors may be underestimated because individual student outcomes are not fully independent. Since the PROMISE treatment in our sample varies only by ACT score and entry cohort, the most problematic type of intraclass correlation for us is correlation that occurs within ACT score bins or within cohorts. With only 10 ACT score bins in our preferred RD analysis and only four cohorts in our DD analysis, we have too few clusters to use typical clustered standard errors. Instead, in Table 7 we undertake a very conservative approach that assumes perfect intraclass correlation: we aggregate the data into 20 ACT score-by-cohort bins, and run our regression on the aggregated data (in addition to imposing perfect intraclass correlation, this approach also ensures that statistical significance is evaluated using the small sample t-distribution). In order to still incorporate covariate adjustments, we residualize all outcomes, as well as our PROMISE receipt and eligibility flags, prior to aggregation.[16] Table 7 shows that this alternate way of

[16] This results in much more comparable coefficients than if we attempt to control for aggregated covariates (i.e. means of covariates within cells) in the aggregated data.

computing standard errors has, perhaps surprisingly, no impact on our conclusions. Almost all estimates that were statistically significant in Tables 3-6 continue to be so. This suggests that there is not much correlation within ACT score bins or within cohorts.

Bounding strategies. Economists increasingly utilize bounding techniques to help assess the extent of potential bias in contexts of imperfect identification (Manski, 1990; Lee, 2009). Lee's (2009) critical insight was that if the researcher can identify the proportion of "excess observations" present in the treated (or control) group, bounds can be obtained by making different extreme assumptions about which observations are the excess ones (e.g., the excess observations may be the ones with the very best outcomes, or the very worst outcomes in the relevant treatment group), and trimming them from the sample.

Recent work has extended these approaches to the RD context (Dong, 2015; Gerard, Rokkanen, & Rothe, 2015). A key challenge in the RD context is that the proportion of excess observations must be estimated not just overall, but for different values of the running variable. Unlike the cases considered in Dong (2015) and Gerard, Rokkanen, & Rothe (2015), however, our data enable us to estimate the proportion of excess observations by ACT score directly, by comparing the ACT score distribution before and after PROMISE implementation (see Figure 1). Comparing the ACT score distribution after PROMISE to that before it, we estimate that about 35 percent of our treatment group are new/excess observations (that is, about 35 percent of our post-implementation cohorts with scores above the cutoff either would not have enrolled or would not have scored above the cutoff, in the absence of PROMISE).

In the first column of Table 8, we show what happens if we assume these students come from the top of the cumulative college GPA distribution. That is, at each ACT score above the cut-off, the excess observations are removed by removing students from the top of the cumulative college GPA distribution. Under this extreme assumption, many effects are diminished and some effects—particularly those linked closely to college academics, like graduate attainment—even switch sign. Interestingly, however, even with this extreme assumption the undergraduate, overall & parent loan impacts are virtually unchanged, and the earnings, homeownership, and neighborhood income outcomes remain consistent in sign. This assumption is particularly extreme because GPA is an outcome clearly affected by the program. If we trim based on high school GPA instead (see last two columns in Table 8), our results are barely distinguishable from baseline.

As in Scott-Clayton (2011), we can also draw upon additional data sources to estimate tighter bounds. As explained in Section 4, the excess mass above the cutoff comes from a combination of individuals who would not have enrolled in college at all in the absence of PROMISE, individuals who otherwise would have attended college out-of-state, and individuals who would have enrolled in college but failed to meet the eligibility criteria (that is moved from below the cutoff to above it). The first source of selection is likely to lead to downward bias, while the latter two could plausibly generate upward bias in our results.

Using data from the WV DOE and IPEDS on enrollment patterns of all WV high school graduates, we can calculate what fraction of excess PROMISE enrollees plausibly come from students switching from out-of-state to in-state enrollment – the source of differential selection most likely to generate a positive bias. Using aggregate data on out-of-state enrollment patterns (see Figure A.1), we estimate that only about 6% of treated students (that is, about 0.06/0.35=17% of the excess enrollees) could be students who would otherwise have attended college out-of-state. To come up with an estimate of the number of excess enrollees attributable to retesting or studying harder to score above the test score cutoff (what are referred to "manipulators" in RD settings, as in Gerard et al., 2015), we estimate the missing observations below the cut-off (see Figure 1).[17] These would constitute about 22% of the excess enrollees.

Together these two groups constitute 40% of the excess enrollees, or 14% (0.40 x 0.35) of the treated group overall. We estimate the worst-case bounds under the assumption that these 14% are the ones with the very best outcomes and trim them from the sample unconditional on ACT scores (à la Lee, 2009). Estimates of this bounding exercise are presented in Table 9. The magnitude of several RD estimates decreases and, unsurprisingly, some lose precision. However, with the exception of the outcomes related to location, all the other estimates are qualitatively similar to those in the baseline model. Turning to the DD estimates in the second column of Table 9, we see that they are remarkably robust to this exercise and very similar (economically and statistically) to the corresponding estimates in Tables 5 and 6.

Finally, we test the sensitivity of our results by excluding students with ACT scores around the threshold, in the spirit of leaving a "donut hole" in the regression analysis. The idea behind this robustness check is that selection bias due to the manipulators is likely to be worst near the cutoff (since arguably retakers are most likely to move from just below the cutoff to just

[17] The implicit assumption here is that individuals with scores below the ACT cutoff who would have enrolled prior to PROMISE would do so after PROMISE as well. Then any missing mass has to be due to individuals moving to above the test score cutoff. This assumption is quite reasonable since PROMISE implementation should generally not dissuade such students from attending. It would, however, be violated if there were capacity constraints such that low-performing students (in terms of ACT scores) who would have earlier gotten in are now not able to. Conversations with WVHEPC administrators in 2008 suggest that this was not the case. In addition, comparing observable characteristics of entry cohorts with scores below the cut-off before and after PROMISE reveals little difference in student observable characteristics (such as race, gender, age, and private school attendance).

above). Appendix Table A.5 presents the baseline RD and DD specifications excluding the set of students with ACT scores in the range of 20-21. We see that the results are quite similar to those in tables 3-6 (except that we lose power for some outcomes), suggestive of the bias due to manipulation not being a serious concern in our context.

Alternative source of identification: first year GPA. Students who initially qualify for PROMISE must maintain a 2.75 GPA in their first year, or they lose eligibility for subsequent years. We can thus perform another RD analysis with first year GPA as the forcing variable, using a local linear specification (+/- 0.75 bandwidth) with the same demographic controls specified in equation (1a). Although this strategy is not without its own limitations, it provides another way to triangulate the results. Table 10 provides our results, which are quite consistently positive and significant across all outcomes examined. While we suggest interpreting these results cautiously given that GPA is lumpy (something found in other contexts which use GPA as the running variable too; for example, see Bettinger et al., 2016) and we again find excess mass above the cutoff, it is reassuring that this entirely separate source of variation generates findings consistent with (indeed even stronger than) our main results.

7. Discussion

Although our alternative approaches generate broadly consistent patterns, results are not always identical across specifications. In this context, how can we assess which results are most convincing? While the DD results, if plausible, would be more useful for policy due to their broader generalizability (because they apply to all PROMISE recipients, not just those near the score cutoff), the potential tradeoff for more questionable internal validity may be too great in this case, at least for those outcomes we expect to be highly sensitive to economic conditions. Dramatic changes over time in the economy are hitting right in the middle of our follow up period. We thus have to heavily rely on PROMISE-ineligible students from the same cohorts to provide us a way to control for underlying trends over time. But if the recession affects high and low-scoring students differentially, it is a problem for the DD analysis.

It is worth noting that if we compare DD Model 3 (which limits to students around the score cutoff, and is thus is more comparable in theory to the RD strategy), results between the RD and DD are quite comparable for those outcomes that are arguably *least likely to be impacted by the great recession*: BA completion, undergraduate borrowing, and parent borrowing (all of which primarily occur in 2007 or earlier), as well as mobility, zipcode-level income (which is based upon fixed Census data), and our marriage proxy. Debt delinquencies and credit scores are also quite consistent between the two strategies, suggesting that our PROMISE-ineligible comparison group may adequately capture the effects of the recession on these outcomes for eligible students. Impact estimates for graduate school enrollment and home ownership, however, are inconsistent between the RD and limited DD specification, suggesting that the assumption of parallel trends may be violated for these outcomes. This is plausible if higher-ability individuals tend to have higher propensities to pursue graduate school and much higher home ownership in general. For example, home ownership among PROMISE-eligible students could have been affected more by the recession than were ineligible students, generating the observed negative point estimates in DD Model 1.

Despite their more limited generalizability, the RD results may nonetheless be of particular interest to both state and national policymakers. While the RD sample is not reflective of all PROMISE recipients, it is more reflective of the typical student at a public institution, since the ACT score cutoff was the median ACT score at the time. The RD results can thus inform policymakers regarding the potential effects of offering PROMISE-like scholarships to students of median ability (among the test taking population).

Focusing, then, on the RD results, the story they reveal is generally a positive, though not miraculous one. It is also a story of the dissipation of treatment effects across many margins. For the RD sample, receiving PROMISE had an average value of $8,400 over four years. Parents appear to "take their cut" of this benefit, reducing their own federal PLUS loan borrowing by about $1,000 while students reduce their undergraduate borrowing by nearly $1,400. Graduate school enrollment (and borrowing) increases substantially as a result of the scholarship, meaning that even a 10-year follow up window may be insufficient to fully assess the labor market outcomes of recipients. Nonetheless, point estimates are suggestive of positive earnings effects ($2,724 [s.e.=1,809], p=0.13). While a full cost-benefit analysis is beyond the scope of this paper, earnings effects of this magnitude would on their own more than compensate for the direct per-student costs of the scholarship, after just four years in the labor market. We also find improvements in socioeconomic status, as measured by income of the individual's residing zip code, for PROMISE recipients. Given the growing evidence on the importance of neighborhood quality for intergenerational mobility and outcomes of children (Chetty et al., 2014; Chetty, Hendren, and Katz, 2016), this is an important result. The credit outcomes also suggest an improvement in financial health of recipients, as indicated by a lower incidence of adverse outcomes and a modest improvement in credit scores and our index of financial well-being.

Our results also add new perspective to the ongoing conversation about the mechanisms underlying financial aid impacts, as well as the long-term consequences of student debt. Scott-Clayton (2011) provides strong evidence that the academic impacts of WV PROMISE are linked to its annual academic performance requirements: treated students' GPAs and credit accumulation are substantially higher than control students in years when performance requirements are binding, but disappear in the fourth year when requirements cease, even though students continue to receive the scholarship money. The WV PROMISE scholarship significantly reduced undergraduate debt for students with ACT scores near the cutoff, and it might be tempting to infer that this is the mechanism that leads to higher graduate attainment and improved financial and neighborhood outcomes later in life. But the acceleration of time to degree is a separate and equally plausible mechanism. Although there is no clean way to separately identify the precise channels for the longer-term impacts, exploratory analyses suggest that improving time to degree may be more important than the reductions in undergraduate loans. First, graduate attainment increased for both our RD and our DD samples, even though only the RD sample saw a reduction in undergraduate loans. Second, due to these increases in graduate enrollment, current loan balances are actually higher—not lower—for the treated group in both the RD and DD samples, yet they still have higher rates of homeownership and live in wealthier neighborhoods. Finally, controlling for undergraduate loan amounts has little effect on subsequent impact estimates, while controlling for time-to-degree reduces subsequent impacts noticeably (see Appendix Table A.6). While we cannot directly examine the effect of student loans on outcomes like homeownership, our results are consistent with the argument that later life outcomes depend more on college outcomes (completion, performance, time to degree) than on debt loads per se (Dynarski, 2016; Executive Office of the President, 2016).

A final important conclusion of our study is that bachelor's degree attainment alone is not sufficient to measure the impact of aid. Ten years after starting college, PROMISE recipients were no more likely to have a BA than their counterparts. But they earned these degrees significantly earlier, enabling them to get a head start on the next phase of their lives. This appears to have paid off in terms of higher graduate attainment, higher homeownership, better neighborhoods, and indications of improved credit and earnings. Our results show that aid improves important life outcomes long after students leave college, even when impacts on bachelor's degree completion fade out.

References

Angrist, J., Autor, D., Hudson, S., & Pallais, A. (2015). Evaluating econometric evaluations of post-secondary aid. *American Economic Review*, *105*(5), 502-07.

Angrist, J., Hudson, S., & Pallais, A. (2014). *Leveling Up: Early Results from a Randomized Evaluation of Post-Secondary Aid* (No. w20800). National Bureau of Economic Research.

Baum, S. R., Elliott, D. C. & Ma, J. (2014). *Trends in student aid, 2014*. New York: NY: The College Board.

Bettinger, E., Gurantz, O., Kawano, L, & Sacerdote, B. (2016). *The Long Run Impacts of Merit Aid: Evidence from California's Cal Grant.* NBER working paper no. 22347. Cambridge, MA: NBER.

Bhutta, Neil, Paige Marta Skiba and Jeremy Tobacman (2012). "Payday Loan Choices and Consequences." Vanderbilt Law and Economics Research Paper No. 12-30.

Brown, Meta, John Grigsby, Wilbert van der Klaauw, Jaya Wen, Basit Zafar (2015). "Financial Education and the Debt Behavior of the Young." *Review of Financial Studies*, Forthcoming.

Bruce, D. J. & Carruthers, C. K., (2014). Jackpot? The impact of lottery scholarships on enrollment in Tennessee. *Journal of Urban Economics* 81, 30-44.

Castleman, Benjamin (2014). *The Impact of Partial and Full Merit Scholarships on College Entry and Success: Evidence from the Florida Bright Futures Scholarship Program.* EdPolicyWorks working paper. Charlottesville, VA: University of Virginia.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. (2014). Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States, *Quarterly Journal of Economics*, 129(4): 1553-1623.

Chetty, Raj, Nathaniel Hendren, and Lawrence Katz. (2016). The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic Review*, 106(4): 855-902.

Cohodes, S. R., & Goodman, J. S. (2014). Merit aid, college quality, and college completion: Massachusetts' Adams scholarship as an in-kind subsidy. *American Economic Journal: Applied Economics*, *6*(4), 251-285.

College Board (2015). *Trends in College Pricing 2015*. Washington, DC: The College Board.

College Board (2012) College is Affordable: A pilot study. The College Board. Washington, DC Available online at http://www.advocacy.collegeboard.org/sites/default/files/college-is-affordable-summary-9-6-12-final.pdf.

Cornwell, Christopher M., Kyung Hee Lee, and David B. Mustard. 2005. "Student Responses to Merit Scholarship Retention Rules." Journal of Human Resources 40(4):895–917.

Cornwell, C., Mustard, D., & Sridhar, D. (2006). The enrollment effects of merit-based financial aid: Evidence from Georgia's HOPE Scholarship. *Journal of Labor Economics* 24: 761–86.

Dong, Y. (2015). *Regression Discontinuity Designs with Sample Selection*. Unpublished manuscript, University of California - Irvine.

Dynarski, S. M. (2000). Hope for whom? Financial aid for the middle class and its impact on college attendance. *National Tax Journal* 53 (3): 629–61.

Dynarski, S. M. (2004). The new merit aid. In Hoxby, C. M. (ed.) *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It.* University of Chicago Press and the National Bureau of Economic Research, pp. 63–100.

Dynarski, S. M. (2008). Building the stock of college-educated labor. *Journal of Human Resources 43(*3): 576–610.

Dynarski, S. M., & Scott-Clayton, J. (2006). The cost of complexity in federal student aid: Lessons from optimal tax theory and behavioral economics. *National Tax Journal* 59 (2): 319–56.

Dynarski, S.M. (2016). The dividing line between haves and have-nots in home ownership: Education, not student debt. Evidence Speaks Reports, Vol 1, #17. Washington, DC: The Brookings Institution.

Executive Office of the President (2016). *Investing in higher education: benefits, challenges, and the state of student debt.* Washington, DC: Executive Office of the President.

Fitzpatrick, M. D., & Jones, D. (2012). *Higher education, merit-based scholarships and post-baccalaureate migration.* Working Paper No. 18530. Cambridge, MA: National Bureau of Economic Research.

Gerard, F., Rokkanen, M. & Rothe, C. (2015). *Partial Identification in Regression Discontinuity Designs with Manipulated Running Variables*. Unpublished manuscript, Columbia University.

Imbens, Guido, and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142(2): 615-635.

Lee, D. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economic Studies*, 76(3), 1071-1102.

Manski, C. (1990). Nonparametric Bounds on Treatment Effects. *American Economic Review:Papers and Proceedings*, 80(2), 319-323.

Mettler, S. (2014). How U.S. higher education promotes inequality—and what can be done to broaden access and graduation. Cambridge, MA: Harvard University, Scholars Strategy Network.

Mian, Atif and Amir Sufi. 2011. "House Prices, Home Equity-Based Borrowing, and the US Household Leverage Crisis." *American Economic Review*, 101 (august 2011): 2132-2156.

Pallais, A. (2009). Taking a chance on college: Is the Tennessee Education Lottery Scholarship Program a winner? *Journal of Human Resources, 44*(1), 199-222.

Scott-Clayton, J. (2011). On money and motivation: A quasi-experimental analysis of financial incentives for college achievement. *Journal of Human Resources 46* (3): 614–46.

Sjoquist, D. L., & Winters, J. V. (2012). Building the stock of college-educated labor revisited. *Journal of Human Resources*, 47(1), 270-285.

Zhang, L., & Ness, E. C. (2010). Does state merit-based aid stem brain drain? *Educational Evaluation and Policy Analysis*, *32*(2), 143-165.

Zimmerman, Seth D.. 2014. "The Returns to College Admission for Academically Marginal Students". Journal of Labor Economics 32 (4).

Figure 1: Distribution of ACT Scores Pre- and Post- PROMISE

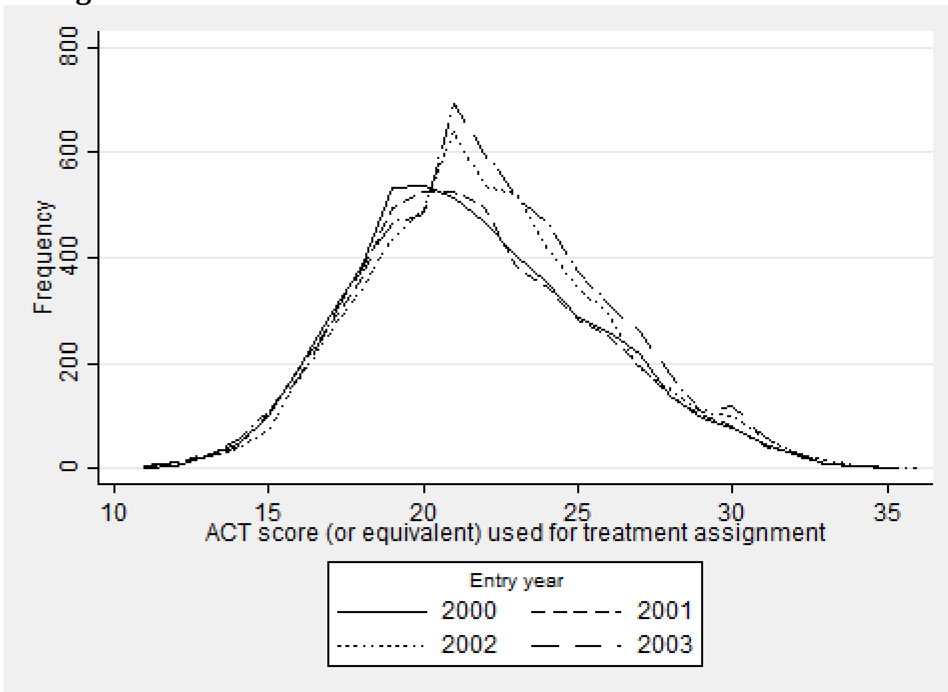


Figure 2: Graphical RD Checks for Selective Educational Outcomes

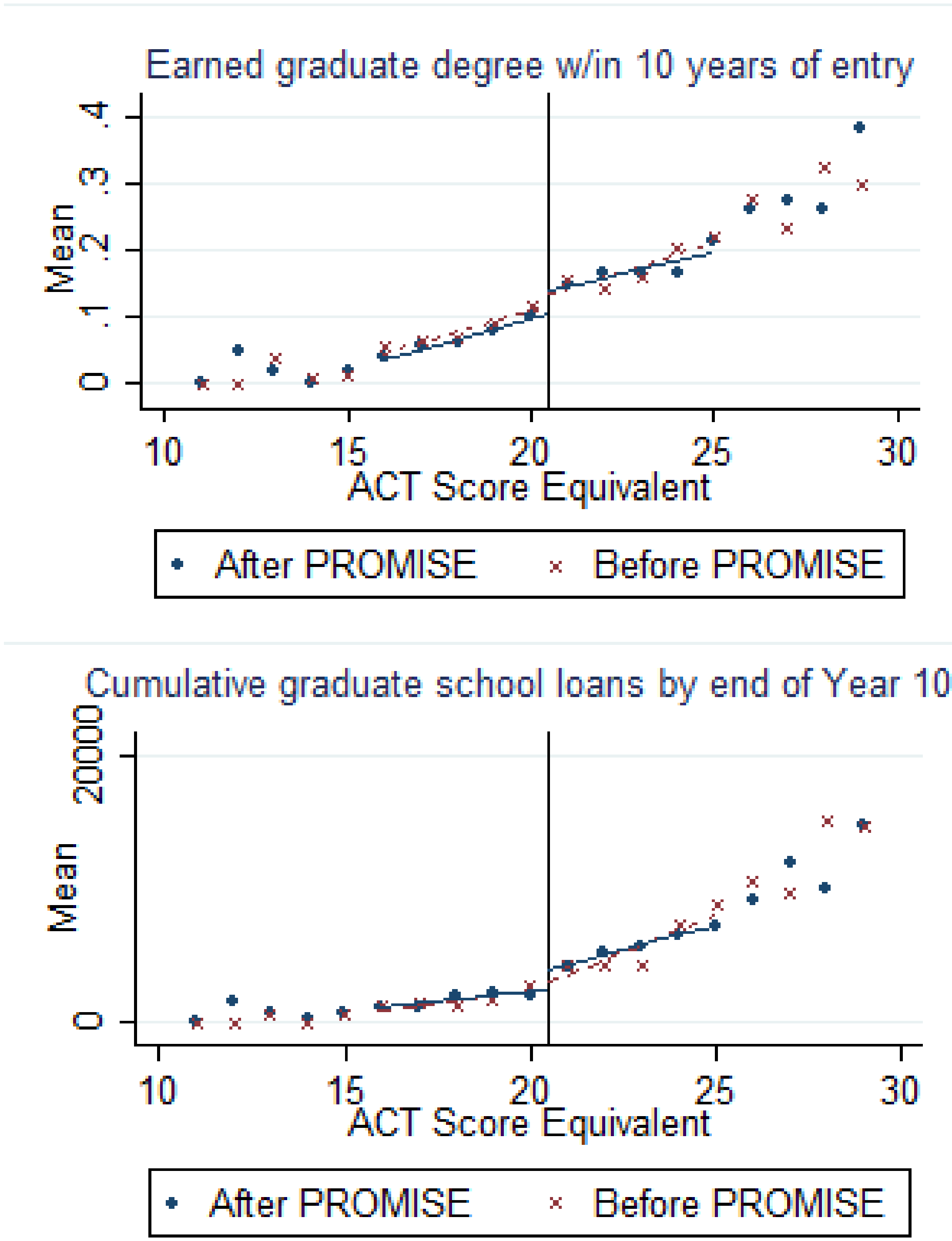


Figure 3: Graphical RD Checks for Selective Credit Outcomes

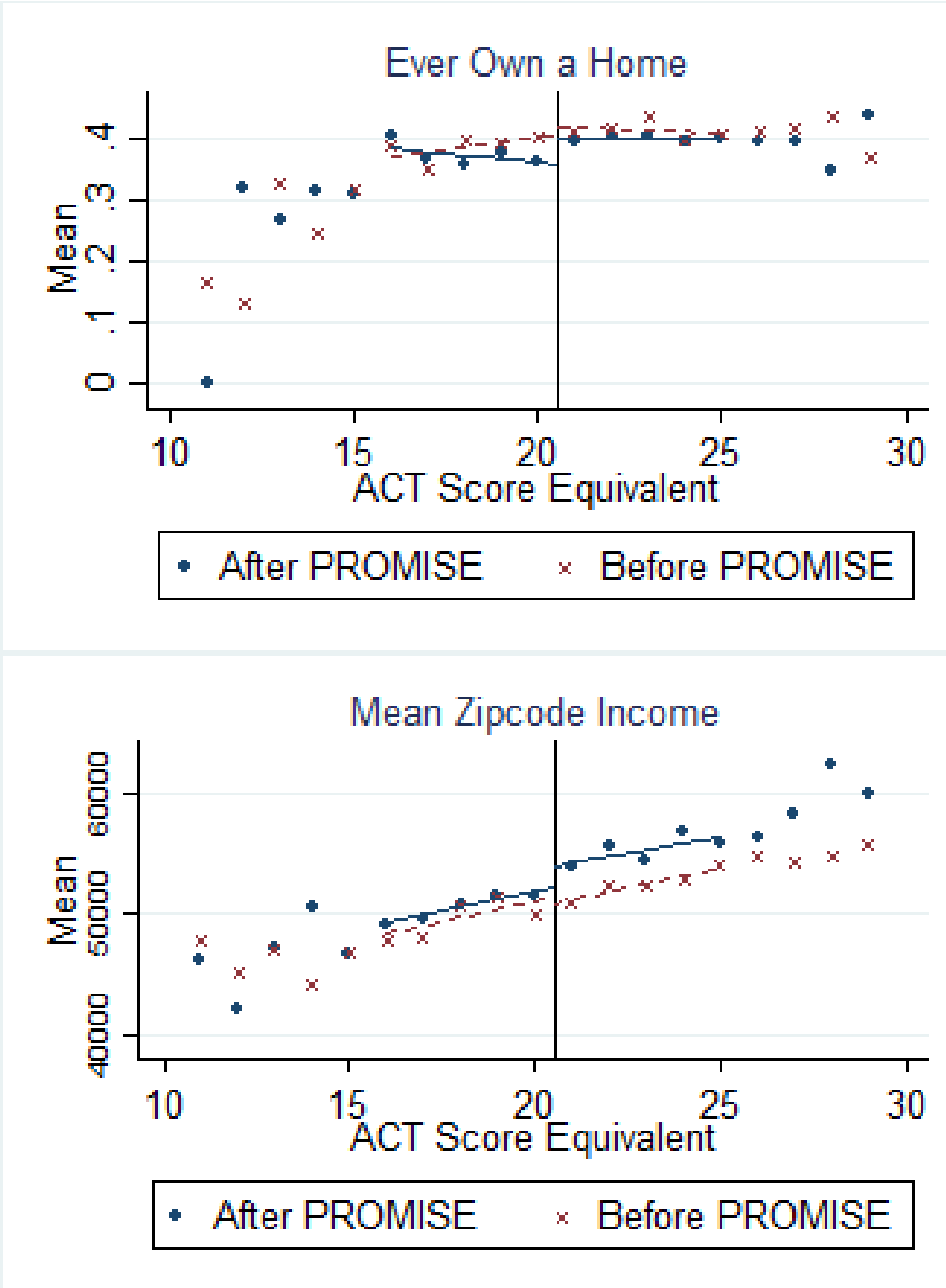


Table 1

Sample Descriptives - Background Characteristics

Outcome	BPS Sample, 2003	Full WV Analysis Sample	Main RD Sample	Main DD Sample
	(1)	(2)	(3)	(4)
Female	0.575	0.546	0.604	0.590
White	0.615	0.951	0.959	0.960
Black, non-hispanic	0.138	0.033	0.024	0.023
Hispanic	0.149	0.006	0.005	0.005
Other race/ethnicity	0.097	0.011	0.011	0.011
Age at entry	22.1	18.6	18.6	18.6
Graduated from private HS	0.082	0.031	0.026	0.030
HS GPA	2.8	3.2	3.5	3.5
Took SAT	0.640	0.105	0.106	0.101
Took ACT		0.926	0.942	0.924
ACT (or equivalent score)	21.0	20.8	21.0	21.9
Had 3.0+ HSGPA and 20+ ACT	n/a	0.429	0.595	0.619
Pell Recip. in 1st year	0.357	0.370	0.367	0.347
Sample size	16,500	30,107	8,578	20,866

SOURCE: Authors' calculations using WVHEPC administrative data on first-time degree-seeking freshmen aged 19 and younger who were West Virginia residents, had non-missing HS GPA and either ACT or SAT scores, and entered in the Fall terms of 2000 through 2003. BPS National from the US Department of Education, National Center for Education Statistics, 2003-04 Beginning Postsecondary Students Longitudinal Study.

NOTES: The first column provides comparison data from the Beginning Postsecondary Student survey, which follows a nationally representative sample of first-time college entrants for six years. The second column includes all first-time degree seeking entrants age 19 or under in the 2000 through 2003 fall entry cohorts, who had valid high school GPA and ACT/SAT data. The RD sample is limited to students entering in 2002 or 2003, with an ACT score between 16 and 25, and at least a 3.0 high school GPA.

DD sample is limited to students entering in 2000-2003, with a high school GPA of at least 3.0.

Table 2
Sample Descriptives - Selected Outcomes

Outcome	National Comparison Group*	Full WV Analysis Sample	Main RD Sample	Main DD Sample	DD Sample: Eligible But Ent.<2002
Earned BA within 4 Years		0.161	0.207	0.293	0.267
Earned BA within 6 Years	0.307	0.412	0.500	0.595	0.591
Earned BA within 10 Years	n/a	0.484	0.566	0.656	0.656
Any grad degree w/in 6 years		0.063	0.074	0.115	0.118
Any grad degree w/in 10 years	n/a	0.118	0.134	0.204	0.209
Any undergraduate loan w/in 5 yrs**	0.530	0.569	0.580	0.555	0.569
Cum. undergraduate loan over 5 yrs**	$6,333	$7,117	$7,626	$7,571	$8,100
Any graduate loan w/in 10 yrs	n/a	0.101	0.115	0.169	0.173
Cum. graduate loan over 10 yrs	n/a	$4,026	$4,147	$7,521	$7,692
Any student loan w/in 10 yrs	n/a	0.626	0.635	0.634	0.645
Cum. Student loan over 10 yrs	n/a	$13,280	$13,744	$16,839	$17,514
Any parent loan w/in 10 yrs	n/a	0.111	0.123	0.113	0.114
Cum parent loan over 10 yrs	n/a	$1,325	$1,536	$1,391	$1,532
Any employment, Yr10 post	n/a	0.594	0.608	0.568	0.564
Employed year-round, Yr 10	n/a	0.471	0.491	0.458	0.462
Earnings|year round emp, Yr 10	n/a	$41,510	$41,991	$46,724	$47,881
<u>*Equifax Outcomes*</u>					
Match to Equifax***	0.922**	0.916	0.931	0.895	0.884
Ever living outside WV	0.995	0.255	0.241	0.265	0.323
Zipcode: mean income	$55,957	$53,864	$53,383	$53,174	$55,758
Zipcode: bottom AGI quartile	0.193	0.194	0.259	0.249	0.249
Marriage proxy (females only)	n/a	0.355	0.366	0.375	0.389
Ever owned a home	0.179	0.34913	0.386	0.395	0.413
Log of mortgage value | Mortgage	12.04	12.26	12.27	12	12.6
Credit card balance	$1,775	$2,120	$2,609	$2,394	$2,398
Ever take on a student loan	0.418	0.617	0.723	0.715	0.623
Student loan balance (inc. 0s)	$10,788	$16,418	$15,531	$20,862	$20,227
Student loan balance (no. 0s)	$28,918	$29,329	$25,993	$35,861	$38,784
Ever past due: stud., auto, home loan	0.306	0.402	0.344	0.306	0.261
Ever had account in collections	0.419	0.501	0.445	0.421	0.387
Credit score	647.98	662.64	664.07	670.08	685.36
Index+	1.27	1.22	1.32	1.39	1.45

SOURCE: Authors' calculations using WVHEPC administrative data on first-time degree-seeking freshmen aged 19 and younger who were West Virginia residents, had non-missing HS GPA and either ACT or SAT scores, and entered in the Fall terms of 2000 through 2003. These data are matched to Equifax consumer credit data.
NOTES: The RD sample is limited to students entering in 2002 or 2003, with ACT score (or SAT equivalent) of 16-25 and high school GPA of 3.0 or higher.
* National comparisons use BPS for academic outcomes. For Equifax outcomes, comparison population is composed of 28-32 year olds in 2014 from the 0.1% sample of the credit bureau data. BPS sample size is 16,500. Equifax sample size is 19,534.
**BPS data examines loans over 6 years rather than 5.
*** The match rate for the comparison population is obtained by using the 2010 Census Population Estimates by age.
+ "Index" ranges from 0 to 3, the sum of three indicators of unambiguously positive credit outcomes: being in the top quartile AGI, never having past due loans, and never having debt in collections.

Table 3

RD Estimates of the Effect of the WV PROMISE Scholarship, Using Estimated Eligibility as Instrument for Actual Receipt (First Stage=0.70)

Outcome	Mean at ACT=20	(1) IV-RD ACT: 16-25		Robustness Checks: (2) No covariates		Alternate Bandwidths: (3) ACT: 18-23		Alternate Bandwidths: (4) ACT: 11-30		(5) Full sample, local quadratic		(6) Falsification: RD Before 2002
Earned BA within 4 Years	0.154	0.096 (0.022)	***	0.103 (0.021)	***	0.099 (0.029)	***	0.080 (0.017)	***	0.085 (0.025)	***	0.012 (0.015)
Earned BA within 6 Years	0.460	0.040 (0.028)		0.060 (0.028)	**	0.015 (0.036)		0.023 (0.023)		0.044 (0.033)		-0.002 (0.020)
Earned BA within 10 Years	0.543	0.019 (0.028)		0.038 (0.028)		-0.013 (0.037)		-0.006 (0.023)		0.018 (0.033)		-0.024 (0.020)
Any grad degree w/in 6 years	0.052	0.026 (0.014)	*	0.031 (0.014)	**	0.029 (0.018)		0.024 (0.011)	**	0.023 (0.016)		0.015 (0.011)
Any grad degree w/in 10 years	0.106	0.042 (0.019)	**	0.050 (0.018)	***	0.045 (0.024)	*	0.023 (0.015)		0.032 (0.022)		0.016 (0.014)
Any UG loan w/in 5 yrs	0.614	-0.088 (0.029)	***	-0.082 (0.028)	***	-0.076 (0.037)	**	-0.065 (0.023)	***	-0.066 (0.033)	**	0.003 (0.020)
Cum. UG loan over 5 yrs	$8,294	-$1,371 (553)	**	-$1,168 (548)	**	-$1,783 (711)	**	-$1,513 (443)	***	-$1,005 (640)		-$347 (400)
Any graduate loan w/in 10 yrs	0.092	0.038 (0.018)	**	0.045 (0.017)	***	0.045 (0.023)	*	0.026 (0.014)	*	0.042 (0.021)	**	0.015 (0.013)
Cum. graduate loan over 10 yrs	$2,071	$1,834 (841)	**	$2,245 (827)	***	$2,220 (1062)	**	$1,018 (738)		$2,359 (996)	**	$373 (650)
Any student loan w/in 10 yrs	0.660	-0.074 (0.028)	***	-0.066 (0.028)	**	-0.063 (0.036)	*	-0.049 (0.023)	**	-0.051 (0.033)		0.004 (0.020)
Cum. Student loan over 10 yrs	$12,580	$689 (1144)		$1,163 (1128)		$546 (1459)		-$598 (965)		$1,724 (1345)		-$514 (858)
Any parent loan w/in 10 yrs	0.140	-0.028 (0.019)		-0.027 (0.019)		-0.043 (0.025)	*	-0.042 (0.015)	***	-0.023 (0.023)		-0.015 (0.013)
Cum parent loan over 10 yrs	$2,088	-$1,013 (364)	***	-$958 (354)	***	-$1,163 (478)	**	-$957 (295)	***	-$1,010 (433)	**	-$155 (211)
Any employment, Yr10 post	0.624	0.003 (0.028)		0.004 (0.028)		-0.004 (0.037)		0.019 (0.023)		-0.004 (0.033)		0.015 (0.020)
Employed year-round, Yr 10	0.511	-0.014 (0.029)		-0.012 (0.029)		-0.026 (0.037)		0.017 (0.024)		-0.029 (0.034)		0.014 (0.021)
Earnings|year round emp, Yr 10	$40,086	$2,724 (1809)		$3,561 (1902)	*	$4,438 (2270)	*	$1,769 (1518)		$2,838 (2169)		$1,297 (1364)
Sample size	974	8,578		8,800		6,096		10,732		10,953		7,830

SOURCE: Authors' calculations using WVHEPC administrative data on first-time degree-seeking freshmen aged 19 and younger who were West Virginia residents, entered in 2002-03 or 2003-04, and met the high school GPA requirement for PROMISE (3.0+).
NOTES: Estimated eligibility for PROMISE (scoring above the ACT cutoff) is used as an instrument for PROMISE receipt (first stage=0.71 [0.013]). Robust standard errors are in parentheses. All regressions include indicator controls for gender, race/ethnicity, high school type, county of residence at entry, and age, as well as a quadratic function of high school GPA. Except where otherwise noted, regressions use a fuzzy RD, local linear regression for students with ACT scores of 16 to 25. *, **, *** indicate the significance of individual findings at the p<0.10, p<0.05, or p<0.01 level.

Table 4

RD Estimates of the Effect of the WV PROMISE Scholarship - Equifax Outcomes

Outcome	Mean at ACT=20	(1) IV-RD ACT: 16-25	Robustness Checks: (2) No covariates	Alternate Bandwidths: (3) ACT: 18-23	Alternate Bandwidths: (4) ACT: 11-30	(5) Full sample, local quadratic	(6) Falsification: RD Before 2002
Match to Equifax data	0.923	-0.002 (0.015)	0.000 (0.016)	-0.001 (0.019	0.004 (0.012)	-0.000 (0.017)	0.013 (0.013)
Ever living outside WV	0.228	-0.003 (0.025)	-0.002 (0.025)	0.023 (0.032)	-0.002 (0.025)	0.009 (0.029)	0.012 (0.019)
Zipcode: mean income	$51,815	$2,409 (1193) **	$2,576 (1208) **	$3,714 (1540) **	$1,796 (1011) *	$2,739 (1346) **	-$39.72 (895)
Zipcode: bottom AGI quartile	0.155	-0.043 (0.023) *	-0.048 (0.025) *	-0.050 (0.031	-0.016 (0.020)	-0.058 (0.029) **	0.031 (0.017) *
Changed name (marriage proxy)	0.348	0.013 (0.034)	-0.000 (0.034)	0.051 (0.043	0 (0.029)	0.043 (0.040)	0.009 (0.026)
Ever owned a home	0.361	0.060 (.029) **	0.054 (0.029) *	0.046 (0.037)	0.054 (0.029) *	0.068 (0.034) **	0.011 (0.022)
Log of mortgage value | Mortgage	11.968	-0.923 (4.038)	-2.113 (4.147)	-3.387 (5.262)	-1.228 (3.372)	-2.294 (4.903)	-2.910 (3.357)
Credit card balance	$2,430	$75 (196)	$76 (196)	$275 (252)	-$48 (157)	$224 (226)	$121 (180)
Ever take out a student loan	0.622	-0.052 (.027) **	-0.049 (.027) *	-0.029 (.034)	-0.034 (.022)	-0.036 (.0316)	-0.018 (0.020)
Student loan balance (incl. 0s)	$15,931	$2,805 (1739)	$3,321 (1734) *	$1,946 (2201)	$1,440 (1445)	$3,066 (2032)	-$1,038 (1178)
Student loan balance (no 0s)	$29,556	$7,723 (2584) ***	$8,160 (2577) ***	$5,144 (3204)	$4,203 (2204) *	$8,488 (3048) ***	-$1491 (1852)
Ever past due: stud/auto/home loan	0.367	-0.026 (0.029)	-0.029 (0.029)	-0.029 (0.037	-0.010 (0.024)	-0.029 (0.034)	-0.005 (0.021)
Ever had account in collections	0.479	-0.047 (0.029)	-0.051 (0.030) *	-0.073 (0.038 *	-0.003 (0.024)	-0.066 (0.034) *	-0.006 (0.022)
Credit score	665.3	3.863 (5.838)	4.998 (5.954)	7.486 (7.554	-1.739 (4.851)	8.131 (6.899)	3.393 (4.283)
Index	1.218	0.062 (0.089)	0.103 0.059) *	0.117 (0.074)	0.008 (0.048)	0.118 (0.068) *	0.043 (0.043)
Sample size		8,578	8,578	6,096	10,732	10,953	7,830

Source: Authors' calculations using WVHEPC administrative data on first-time degree-seeking freshmen aged 19 and younger, entering in the fall semester of school years 2000-01 through 2003-04 matched to Equifax/Federal Reserve Bank of New York Consumer Credit Panel files. The sample is restricted to West Virginia residents who met the high school GPA (3.0+) requirement for PROMISE eligibility.

Notes: Estimated eligibility for PROMISE (scoring above the ACT cutoff) is used as an instrument for PROMISE receipt (first stage=0.71 [0.013]). Robust standard errors are in parentheses and number of observations is in brackets. All regressions include indicator controls for gender, race/ethnicity, high school type, county of residence at entry, and age, as well as a quadratic function of high school GPA. Except where otherwise noted, regressions use a fuzzy RD, local linear regression for students with ACT scores of 16 to 25. *, **, *** indicate the significance of individual findings at the $p<0.10$, $p<0.05$, or $p<0.01$ level. The marriage proxy is estimated for the subsample of females only.

"Index" ranges from 0 to 3, the sum of three indicators of unambiguously positive credit outcomes: being in the top quartile adjusted gross income (AGI) zip code, never having past due loans, and never having debt in collections.

Table 5

IV Difference-in-Difference Estimates of Academic, Student Loan, and Employment Outcomes

Outcome	Eligible stud. Pre-Mean	(1) Baseline IV-Diff-in-Diff		(2) No Controls		(3) Limited to ACT 19-22	
Earned BA within 4 Years	0.267	0.085 (0.012)	***	0.127 (0.012)	***	0.081 (0.021)	***
Earned BA within 6 Years	0.591	0.090 (0.016)	***	0.152 (0.016)	***	0.044 (0.028)	
Earned BA within 10 Years	0.656	0.081 (0.016)	***	0.138 (0.016)	***	0.043 (0.028)	
Any grad degree w/in 6 years	0.118	0.013 (0.008)		0.033 (0.008)	***	0.012 (0.014)	
Any grad degree w/in 10 years	0.209	0.028 (0.011)	***	0.061 (0.011)	***	0.019 (0.019)	
Any UG loan w/in 5 yrs	0.569	-0.019 (0.017)		-0.039 (0.016)	**	-0.056 (0.028)	*
Cum. UG loan over 5 yrs	$8,100	-$505 (314)		-$624 (310)	**	-$859 (555)	
Any graduate loan w/in 10 yrs	0.173	0.022 (0.010)	**	0.041 (0.010)	***	0.019 (0.018)	
Cum. graduate loan over 10 yrs	$7,692	$589 (571)		$1,701 (560)	***	$203 (869)	
Any student loan w/in 10 yrs	0.645	-0.003 (0.016)		-0.017 (0.016)		-0.042 (0.028)	
Cum. Student loan over 10 yrs	$17,514	$51 (722)		$804 (709)		-$155 (1164)	
Any parent loan w/in 10 yrs	0.114	-0.021 (0.011)	**	-0.023 (0.010)	**	-0.013 (0.019)	
Cum parent loan over 10 yrs	$1,532	-$736 (186)	***	-$741 (179)	***	-$799 (341)	**
Any in-state employment, Yr10 po	0.564	-0.008 (0.016)		-0.006 (0.016)		0.019 (0.028)	
Employed year-round, Yr 10	0.462	-0.024 (0.017)		-0.014 (0.017)		0.001 (0.029)	
Earnings|year round emp, Yr 10	$47,881	$1,540 (1155)		$3,688 (1197)	***	$1,640 (1793)	
Sample size	5,779	20,866		20,866		8,440	

Source: Authors' calculations using WVHEPC administrative data on first-time degree-seeking freshmen aged 19 and younger, entering in the fall semester of school years 2000-01 through 2003-04. The sample is restricted to West Virginia residents who met the high school GPA (3.0+) requirement for PROMISE eligibility.
Notes: Robust standard errors are in parentheses. First stage impact on PROMISE receipt is 0.79 percentage points (S.E.=0.006).
*, **, *** indicate the significance of individual findings at the p<0.10, p<0.05, or p<0.01 level.
Unless otherwise noted, regressions include fixed effects for entry cohorts and ACT scores, as well as controls for age, high school GPA and GPA squared, gender, and race/ethnicity, high school type (public/private) and county of residence at entry indicators.

Table 6

IV Difference-in-Difference Estimates of Equifax Outcomes

Outcome	Eligible stud. Pre-Mean	(1) Baseline IV-Diff-in-Diff			(2) No Controls			(3) Limited to ACT 19-22		
Match to Equifax data	0.884	0.001	(0.009)		0.009	(0.010)		-0.024	(0.020)	
Ever living outside WV	0.323	0.024	(0.015)	*	0.026	(0.015)	*	-0.012	(0.025)	
Zipcode: mean income	$55,758	$2,838	(736.9)	***	$3,955	(745.2)	***	$1,901	(1257)	
Zipcode: bottom AGI quartile	0.249	-0.034	(0.014)	**	-0.061	(0.014)	***	-0.046	(0.023)	**
Changed name (marriage proxy)	0.389	-0.036	(0.021)		-0.026	(0.021)		0.03	(0.046)	
Ever owned a home	0.413	-0.020	(0.017)		-0.009	(0.0171)		0.018	(0.029)	
Log of mortgage value | Mortgage	12.604	-1.323	(2.367)		-3.305	(2.459)		2.709	(4.906)	
Credit card balance	$2,398	$71	(139)		$99	(137)		$16	(235)	
Ever take out a student loan	0.632	0.000	(0.016)		-0.008	(0.016)		-0.016	(0.027)	
Student loan balance (inc. 0s)	$20,227	$1,050	(1018)		$993.9	(1018)		$2,768	(2032)	***
Student loan balance (no 0s)	$38,784	$988.1	(1573)		$1,103	(1566)		$8,260	(3053)	***
Ever past due: stud/auto/home loan	0.261	-0.034	(0.017)	**	-0.053	(0.017)	***	-0.026	(0.029)	
Ever had account in collections	0.387	-0.031	(0.017)	*	-0.055	(0.018)	***	-0.034	(0.029)	
Credit score	685.36	6.454	(3.446)	*	11.765	(3.532)	***	2.312	(5.833)	
Index	1.450	0.061	(0.034)	**	0.128	(0.035)	***	0.030	(0.072)	
Sample size	5,779	20,866			20,866			8,440		

Source: Authors' calculations using WVHEPC administrative data on first-time degree-seeking freshmen aged 19 and younger, entering in the fall semester of school years 2000-01 through 2003-04 matched to Equifax/Federal Reserve Bank of New York Consumer Credit Panel files. The sample is restricted to West Virginia residents who met the high school GPA (3.0+) requirement for PROMISE eligibility.

Notes: Robust standard errors are in parentheses. First stage impact on PROMISE receipt is 0.79 percentage points (S.E.=0.006). *, **, *** indicate the significance of individual findings at the p<0.10, p<0.05, or p<0.01 level.

Unless otherwise noted, regressions include fixed effects for entry cohorts and ACT scores, as well as controls for age, high school GPA and GPA squared, gender, and race/ethnicity, high school type (public/private) and county of residence at entry indicators. The marriage proxy is estimated for the subsample of females only.

"Index" ranges from 0 to 3, the sum of three indicators of unambiguously positive credit outcomes: being in the top quartile AGI, never having past due loans, and never having debt in collections.

Table 7

Estimates with Alternative Standard Errors,
Assuming Perfect Intraclass Correlation

Outcome	(1) IV Baseline RD		(2) Baseline IV Diff-in-Diff	
Earned BA within 4 Years	0.095 (0.022)	***	0.079 (0.014)	***
Earned BA within 6 Years	0.039 (0.029)		0.090 (0.018)	***
Earned BA within 10 Years	0.017 (0.032)		0.079 (0.020)	***
Any grad degree w/in 6 years	0.026 (0.015)	*	0.009 (0.010)	
Any grad degree w/in 10 years	0.041 (0.019)	**	0.022 (0.015)	
Any undergraduate loan w/in 5 yrs	-0.090 (0.027)	***	-0.024 (0.018)	
Cum. undergraduate loan over 5 yrs	-$1,397 (446)	***	-$536 (365)	
Any graduate loan w/in 10 yrs	0.037 (0.012)	***	0.017 (0.016)	
Cum. graduate loan over 10 yrs	$1,815 (991)	*	$250 (944)	
Any student loan w/in 10 yrs	-0.076 (0.024)	***	-0.011 (0.017)	
Cum. Student loan over 10 yrs	$645 (1029)		-$318 (1070)	
Any parent loan w/in 10 yrs	-0.028 (0.023)		-0.024 (0.011)	**
Cum parent loan over 10 yrs	-$1,026 (329)	***	-$760 (175)	***
Any employment, Yr9 post	-0.011 (0.029)		-0.009 (0.018)	
Employed year-round, Yr 9	0.002 (0.028)		-0.011 (0.018)	
Earnings|year round emp, Yr 9	$1,087 (2141)		-$33 (1203)	
Any employment, Yr10 post	0.004 (0.033)		-0.006 (0.019)	
Employed year-round, Yr 10	-0.013 (0.035)		-0.020 (0.018)	
Earnings|year round emp, Yr 10	$2,648 (2264)		$1,563 (1391)	
<u>*Equifax Outcomes*</u>				
Match to Equifax data	-0.002 (0.018)		-0.001 (0.008)	
Ever living outside WV	-0.004 (0.022)		0.019 (0.017)	
Zipcode: mean income	$2,389 (1046)	**	$2,748 (790)	***
Zipcode: bottom AGI quartile	-0.042 (0.018)	**	-0.035 (0.014)	**
Changed name (marriage proxy)	0.008 (0.018)		-0.014 (0.012)	
Ever owned a home	0.061 (0.019)	***	-0.009 (0.015)	
Log of mortgage value | Mortgage	-0.316 (4.311)		-1.413 (2.472)	
Credit card balance	$77 (152)		$68 (111)	
Ever take out a student loan	-0.071 (0.025)	***	-0.011 (0.019)	
Student loan balance (inc. 0s)	$2,754 (1782)		$603 (1270)	
Student loan balance (no 0s)	$7,829 (2642)	***	$842 (2006)	
Ever past due: stud/auto/home loan	-0.025 (0.023)		-0.040 (0.013)	***
Ever had account in collections	-0.048 (0.027)	*	-0.035 (0.014)	**
Credit score	3.970 (4.377)		7.538 (2.648)	***
Index	0.083 (0.046)	*	0.078 (0.028)	***
Sample Size	20		101	

Notes: Standard errors are in parentheses. First stage impact on PROMISE receipt is 0.79 percentage points (S.E.=0.006). *, **, *** indicate the significance of individual findings at the p<0.10, p<0.05, or p<0.01 level.

Table 8
Bounding Exercise: Baseline Specifications
Trimming All "Excess Observations" from Treated Group at Each ACT Score Above the Cutoff

	Trimming by Cumulative College GPA				Trimming by High School GPA			
Outcome	(1) IV-RD ACT: 16-25		(2) Baseline IV-Diff-in-Diff		(1) IV-RD ACT: 16-25		(2) Baseline IV-Diff-in-Diff	
Earned BA within 4 Years	0.035 (0.02)		0.053 (0.02)	***	0.091 (0.02)	***	0.092 (0.02)	***
Earned BA within 6 Years	-0.047 (0.03)		0.054 (0.02)	***	0.054 (0.03)	*	0.103 (0.02)	***
Earned BA within 10 Years	-0.059 (0.03)	*	0.048 (0.02)	**	0.041 (0.03)		0.094 (0.02)	***
Any grad degree w/in 6 years	-0.026 (0.01)	**	-0.021 (0.01)	**	0.032 (0.02)	**	0.015 (0.01)	
Any grad degree w/in 10 years	-0.064 (0.02)	***	-0.033 (0.01)	***	0.039 (0.02)	*	0.029 (0.01)	**
Any UG loan w/in 5 yrs	-0.076 (0.03)	**	0.010 (0.02)		-0.096 (0.03)	***	-0.008 (0.02)	
Cum. UG loan over 5 yrs	-$1,302 (591)	**	-$385 (398)		-$1,458 (603)	**	-$602 (402)	
Any graduate loan w/in 10 yrs	-0.041 (0.02)	**	-0.017 (0.01)		0.031 (0.02)	*	0.021 (0.01)	*
Cum. graduate loan over 10 yrs	$570 (933)		-$154 (654)		$1,130 (848)		$1,554 (632)	
Any student loan w/in 10 yrs	-0.077 (0.03)	***	0.014 (0.02)		-0.083 (0.03)	***	0.007 (0.02)	
Cum. Student loan over 10 yrs	-$132 (1248)		-347 (844)		8.69 (1193)		-577 (830)	
Any parent loan w/in 10 yrs	-0.026 (0.02)		-0.034 (0.02)	**	-0.033 (0.02)		-0.040 (0.02)	***
Cum parent loan over 10 yrs	-$1,002 (382)	***	-$,1199 (466)	**	-$1,078 (397)	***	-$1,300 (472)	***
Any employment, Yr10 post	-0.001 (0.03)		-0.007 (0.02)		0.028 (0.03)		-0.005 (0.02)	
Employed year-round, Yr 10	-0.022 (0.03)		-0.027 (0.02)		0.014 (0.03)		-0.018 (0.02)	
Earnings|year round emp, Yr 10	$1,148 (1957)		$1,689 (1263)		$2,598 (1936)		$2,296 (1266)	*
<u>*Equifax Outcomes*</u>								
Match to Equifax data	-0.005 (0.02)		0.002 (0.01)		0.003 (0.02)		0.003 (0.01)	
Ever living outside WV	-0.008 (0.03)		0.024 (0.02)		-0.015 (0.03)		0.020 (0.02)	
Zipcode: mean income	$1,945 (1324)		$3,368 (1050)	***	$2,539 (1315)	*	$3,190 (1055)	***
Zipcode: bottom AGI quartile	-0.024 (0.03)		-0.031 (0.02)	*	-0.038 (0.03)		-0.032 (0.02)	**
Changed name (marriage proxy)	0.004 (0.022)		-0.013 (0.015)		0.005 (0.022)		-0.014 (0.015)	
Ever owned a home	0.030 (0.03)		-0.015 (0.02)		0.089 (0.03)	***	-0.002 (0.02)	
Log of mortgage value | Mortgage	2.113 (5.12)		-2.411 (3.34)		-4.940 (5.08)		-3.473 (3.38)	
Credit card balance	-$20 (199)		$64 (142)		$43 (199)		$75 (141)	
Ever take out a student loan	-0.053 (0.03)	*	0.013 (0.02)		-0.085 (0.03)	***	-0.002 (0.02)	
Student loan balance (inc. 0s)	$3,425 (1911)	*	$1,326 (1339)		$2,266 (1871)		$456 (1329)	
Student loan balance (no 0s)	$8,628 (2822)	***	$384 (2055)		$7,578 (2777)	***	-$399 (2052)	
Ever past due: stud/auto/home loan	0.028 (0.03)		-0.022 (0.02)		-0.016 (0.03)		-0.047 (0.02)	**
Ever had account in collections	0.013 (0.03)		-0.026 (0.02)		-0.047 (0.03)		-0.058 (0.02)	***
Credit score	-10.716 (6.30)	*	3.007 (4.22)		3.239 (6.42)		9.542 (4.25)	**
Index	-0.044 (0.06)		0.064 (0.04)	*	0.079 (0.06)		0.120 (0.04)	***
Sample size	7,548		19,623		7,548		19,623	

Notes: Robust standard errors are in parentheses.
*, **, *** indicate the significance of individual findings at the p<0.10, p<0.05, or p<0.01 level.
Unless otherwise noted, regressions include fixed effects for entry cohorts and ACT scores, as well as controls for age, high school GPA and GPA squared, gender, and race/ethnicity, high school type (public/private) and county of residence at entry indicators. The marriage proxy is estimated for the subsample of females only.
"Index" ranges from 0 to 3, the sum of three indicators of unambiguously positive credit outcomes: being in the top quartile AGI, never having past due loans, and never having debt in collections.

Table 9

Bounding Exercise: Baseline Specifications

Trimming Out-of-State Enrollees & Test Retakers

Outcome	(1) IV-RD ACT: 16-25		(2) Baseline IV-Diff-in-Diff	
Earned BA within 4 Years	0.068 (0.022)	***	0.086 (0.014)	***
Earned BA within 6 Years	0.022 (0.028)		0.101 (0.019)	***
Earned BA within 10 Years	0.006 (0.028)		0.091 (0.019)	***
Any grad degree w/in 6 years	0.005 (0.013)		0.018 (0.009)	**
Any grad degree w/in 10 years	0.021 (0.018)		0.031 (0.012)	***
Any undergraduate loan w/in 5 yrs	-0.066 (0.029)	**	-0.011 (0.019)	
Cum. undergraduate loan over 5 yrs	-$1,052 (557)	*	-$669 (383)	*
Any graduate loan w/in 10 yrs	0.044 (0.018)	**	0.022 (0.011)	**
Cum. graduate loan over 10 yrs	$1,952 (856)	**	$464 (622)	
Any student loan w/in 10 yrs	-0.045 (0.028)		0.005 (0.019)	
Cum. Student loan over 10 yrs	$1,237 (1156)		-$222 (808)	
Any parent loan w/in 10 yrs	-0.023 (0.020)		-0.040 (0.014)	***
Cum parent loan over 10 yrs	-$948 (368)	***	-$1,283 (444)	***
Any employment, Yr10 post	-0.015 (0.029)		-0.011 (0.018)	
Employed year-round, Yr 10	-0.038 (0.029)		-0.028 (0.018)	
Earnings|year round emp, Yr 10	-$32 (1627)		$1,142 (1054)	
<u>*Equifax Outcomes*</u>				
Match to Equifax data	-0.005 (0.015)		0.003 (0.011)	
Ever living outside WV	-0.027 (0.025)		0.025 (0.016)	
Zipcode: mean income	-$1,214 (963)		$2,544 (944)	***
Zipcode: top AGI quartile	0.007 (0.020)		0.025 (0.016)	
Changed name (marriage proxy)	0.008 (0.022)		-0.013 (0.015)	
Ever owned a home	0.033 (0.030)		-0.004 (0.020)	
Log of mortgage value | Mortgage	-0.568 (0.448)		-0.141 (0.236)	
Credit card balance	-$29 (190)		$70 (136)	
Ever take out a student loan	-0.037 (0.030)		0.001 (0.021)	
Student loan balance (inc. 0s)	$3,661 (1760)	**	$662 (1275)	
Student loan balance (no 0s)	$10,594 (2806)	***	$646 (2149)	
Ever past due: stud/auto/home loan	-0.014 (0.028)		-0.050 (0.018)	***
Ever had account in collections	-0.030 (0.030)		-0.056 (0.020)	***
Credit score	0.075 (5.857)		10.45 (4.048)	***
Index	0.032 (0.053)		0.137 (0.037)	***
Sample size	8,175		20,463	

Notes: Robust standard errors are in parentheses.
*, **, *** indicate the significance of individual findings at the $p<0.10$, $p<0.05$, or $p<0.01$ level.
Unless otherwise noted, regressions include fixed effects for entry cohorts and ACT scores, as well as controls for age, high school GPA and GPA squared, gender, and race/ethnicity, high school type (public/private) and county of residence at entry indicators. The marriage proxy is estimated for the subsample of females only.

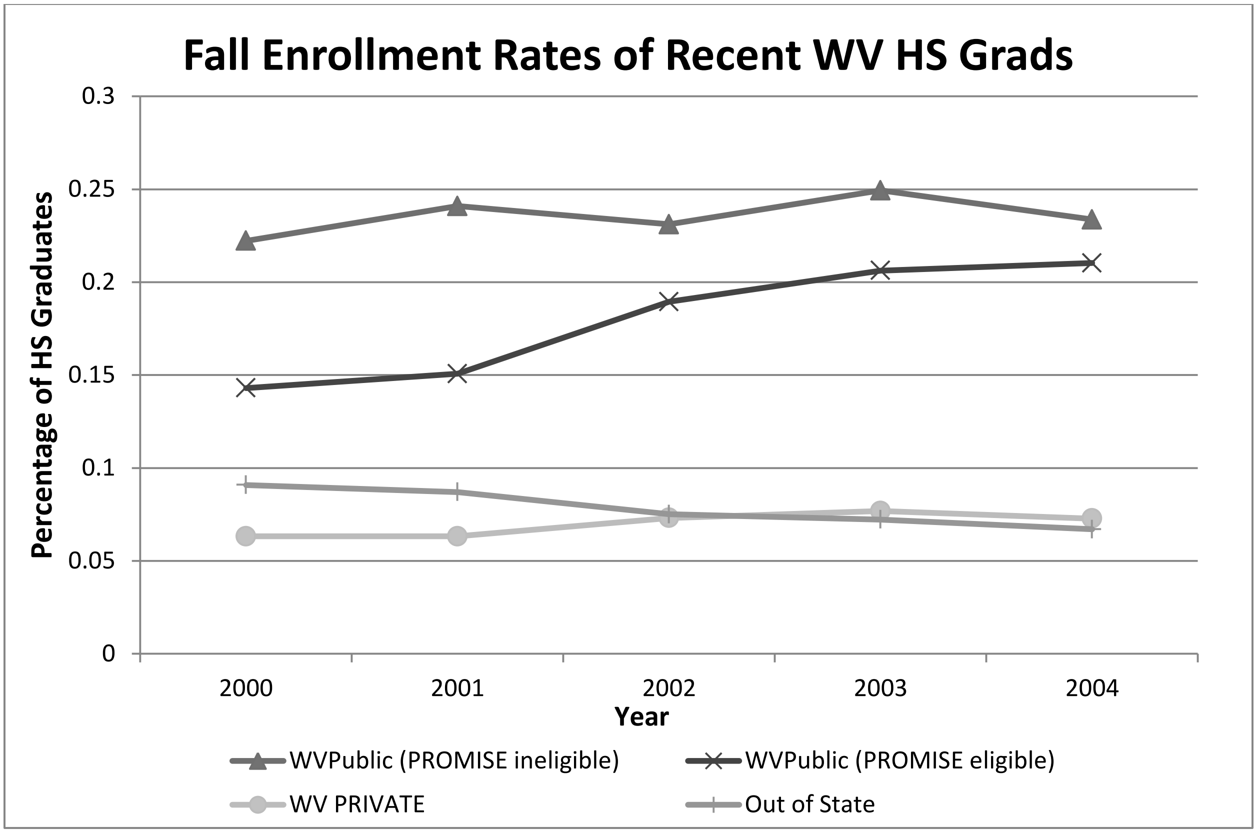
Table 10

Meeting GPA Requirments After Year 1: RD Estimates

Outcome	Estimate	Std Error	
Receive PROMISE Year 2	0.832	(0.013)	***
Cumulative PROMISE Amt	$6,838	(135)	***
Earned BA within 4 Years	0.208	(0.016)	***
Earned BA within 6 Years	0.262	(0.024)	***
Earned BA within 10 Years	0.234	(0.023)	***
Any grad degree w/in 6 years	0.069	(0.009)	***
Any grad degree w/in 10 years	0.103	(0.014)	***
Any undergraduate loan w/in 5 yrs	-0.106	(0.022)	***
Cum. undergraduate loan over 5 yrs	-$1,027	(450)	**
Any graduate loan w/in 10 yrs	0.083	(0.014)	***
Cum. graduate loan over 10 yrs	3869	(635)	***
Any student loan w/in 10 yrs	-0.081	(0.021)	***
Cum. Student loan over 10 yrs	$1,890	(894)	**
Any parent loan w/in 10 yrs	-0.050	(0.017)	***
Cum parent loan over 10 yrs	-$591	(291)	**
Any employment, Yr10 post	-0.022	(0.023)	
Employed year-round, Yr 10	0.005	(0.023)	
Earnings|year round emp, Yr 10	$3,471	(1569)	**
Equifax Outcomes			
Match to Equifax data	0.014	(0.012)	
Ever living outside WV	0.062	(0.020)	***
Zipcode: mean income	$4,057	(974)	***
Zipcode: bottom AGI quartile	-0.029	(0.019)	
Changed name (marriage proxy)	0.007	(0.013)	
Ever owned a home	0.090	(0.023)	***
Log of mortgage value | Mortgage	-10.43	(4.45)	**
Credit card balance	$341	(175)	*
Ever take out a student loan	-0.089	(0.023)	***
Student loan balance (inc. 0s)	$2,452	(1574)	
Student loan balance (no 0s)	6499	(2234)	***
Ever past due: stud/auto/home loan	-0.156	(0.023)	***
Ever past due on any account	-0.156	(0.023)	***
Ever had account in collections	-0.179	(0.024)	***
Credit score	35.69	(4.54)	***
Index	0.389	(0.046)	***
Sample Size		2,791	

Notes: Robust standard errors are in parentheses.
*, **, *** indicate the significance of individual findings at the p<0.10, p<0.05, or p<0.01 level.
First-year GPA is the running variable, with a +/- 0.75 bandwidth around 2.75. Unless otherwise noted, regressions include fixed effects for entry cohorts and ACT scores, as well as controls for age, high school GPA and GPA squared, gender, and race/ethnicity, high school type (public/private) and county of residence at entry indicators. The marriage proxy is estimated for the subsample of females only.

Figure A.1: Fall Enrollment Rates of WV High School Graduates (Source: IPEDS)



Appendix Table A.1. McCrary Test and Covariate Balance

McCrary Test	Coef. (S.E.)		N
Frequency of observations at the eligibility cutoff			
IV-RD	202.9 (30.0)	***	20
IV-DD	128.8 (28.8)	***	101

IV-RD specification	Coef. (S.E.)		N
High School GPA	0.03 (0.02)	**	8578
Race/ethnicity			
White, non-hispanic	-0.02 (0.01)		8578
Black, non-hispanic	0.01 (0.01)		8578
Hispanic (any race)	0.00 (0.00)		8578
Other race	0.01 (0.01)	*	8578
Age at entry	-0.02 (0.02)		8578
Attended private high school	-0.01 (0.01)		8578
Pell amount, Year 1	-62.66 (88.05)		8578
IV-DD specification	Coef. (S.E.)		N
High School GPA	0.12 (0.01)	***	20866
Race/ethnicity			
White, non-hispanic	0.03 (0.01)	***	20866
Black, non-hispanic	-0.03 (0.01)	***	20866
Hispanic (any race)	-0.01 (0.00)	***	20866
Other race	0.01 (0.00)	*	20866
Age at entry	-0.04 (0.01)	***	20866
Attended private high school	0.01 (0.01)	*	20866
Pell amount, Year 1	-198.47 (47.27)	***	20866

SOURCE: Authors' calculations using WVHEPC administrative data on first-time degree-seeking freshmen aged 19 and younger who were West Virginia residents, had non-missing HS GPA and either ACT or SAT scores, and entered in the Fall terms of 2000 through 2003. BPS National from the US Department of Education, National Center for Education Statistics, 2003-04 Beginning Postsecondary Students Longitudinal Study.
NOTES: Coefficients for covariate checks represent results of running our basic IV-RD and IV-DD specifications (with no covariates) separately treating each covariate as an outcome on the left side of the regression equation.

Appendix Table A.2

Subgroup Analysis: Pell Recipients and Non-recipients

	RD-IV (16<=ACT<=25)		DD-IV (Main specification)	
Outcome	Pell Recipient	Non-recipient	Pell Recipient	Non-recipient
Earned BA within 4 Years	0.092 (0.031) ***	0.092 (0.030) ***	0.088 (0.018) ***	0.081 (0.016) ***
Earned BA within 6 Years	0.046 (0.044)	0.021 (0.037)	0.085 (0.026) ***	0.081 (0.020) ***
Earned BA within 10 Years	0.032 (0.045)	0.001 (0.036)	0.078 (0.026) ***	0.070 (0.020) ***
Any grad degree w/in 6 years	0.013 (0.020)	0.031 (0.019)	-0.004 (0.012)	0.019 (0.011) *
Any grad degree w/in 10 years	0.044 (0.028)	0.036 (0.025)	0.014 (0.017)	0.033 (0.014) **
Any UG loan w/in 5 yrs	-0.139 (0.042) ***	-0.063 (0.038) *	-0.046 (0.025) *	-0.006 (0.022)
Cum. UG loan over 5 yrs	-$1,893 (857) **	-$1,162 (719)	-$1,375 (519) ***	-$212 (394)
Any graduate loan w/in 10 yrs	0.043 (0.028)	0.027 (0.023)	0.015 (0.017)	0.018 (0.013)
Cum. graduate loan over 10 yrs	$1,329 (1256)	$1,988 (1141) *	$321 (853)	$597 (748)
Any student loan w/in 10 yrs	-0.103 (0.040) **	-0.061 (0.037)	-0.035 (0.024)	0.009 (0.021)
Cum. Student loan over 10 yrs	$222 (1788)	$837 (1496)	-$1,351 (1137)	$484 (924)
Any parent loan w/in 10 yrs	0.002 (0.026)	-0.046 (0.027) *	-0.025 (0.014) *	-0.026 (0.015) *
Cum parent loan over 10 yrs	$0 (284)	-$1,740 (569) ***	-$225 (148)	-$1,174 (287) ***
Any employment, Yr10 post	0.050 (0.045)	-0.030 (0.037)	-0.005 (0.027)	-0.015 (0.021)
Employed year-round, Yr 10	0.020 (0.046)	-0.035 (0.038)	-0.024 (0.028)	-0.029 (0.021)
Earnings|year round emp, Yr 10	$3,179 (2945)	$2,392 (2293)	$1,714 (1782)	$991 (1489)
Equifax Outcomes				
Match to Equifax data	-0.017 (0.038)	0.021 (0.029)	-0.057 (0.023) **	0.009 (0.017)
Ever living outside WV	-0.041 (0.039)	0.017 (0.032)	0.019 (0.024)	0.030 (0.019)
Zipcode: mean income	-$1,947 (1836)	$4,903 (1605) ***	$2,603 (1042) **	$2,914 (1010) ***
Zipcode: bottom AGI quartile	-0.021 (0.039)	-0.051 (0.030) *	-0.025 (0.023)	-0.037 (0.017) **
Changed name (marriage proxy)	-0.010 (0.036)	0.017 (0.028)	-0.06 (0.022) **	0.008 (0.016)
Ever owned a home	0.090 (0.045) **	0.044 (0.040)	-0.014 (0.028)	-0.027 (0.022)
Log of mortgage value | Mortgage	-5.906 (8.615)	1.079 (5.354)	-0.889 (4.978)	-1.180 (3.220)
Credit card balance	-$97 (291)	$104 (263)	-$74 (192)	$67 (152)
Ever take out a student loan	-0.028 (0.043)	-0.093 (0.040) **	-0.031 (0.027)	0.015 (0.023)
Student loan balance (inc. 0s)	$3,139 (2771)	$2,462 (2234) **	-$1,070 (1597)	$2,160 (1314)
Student loan balance (no 0s)	$7,495 (3903) *	$8,020 (3476) **	-$998 (2242)	$2,331 (2181)
Ever past due: stud/auto/home loan	-0.034 (0.047)	-0.019 (0.037)	-0.066 (0.029) **	-0.005 (0.021)
Ever had account in collections	-0.059 (0.046)	-0.046 (0.038)	-0.048 (0.028) *	-0.014 (0.022)
Credit score	7.06 (9.55)	0.915 (7.28)	12.70 (5.94) **	0.249 (4.196)
Index	0.048 (0.082)	0.106 (0.068)	0.125 (0.050) **	0.017 (0.039)
Sample size	3152	5426	7233	13633

Notes: Robust standard errors are in parentheses.

*, **, *** indicate the significance of individual findings at the p<0.10, p<0.05, or p<0.01 level.

Appendix Table A.3

Subgroup Analysis: Females and Males

	RD-IV (16<=ACT<=25)				DD-IV (Main specification)			
Outcome	Males		Females		Males		Females	
Earned BA within 4 Years	0.119 (0.035)	***	0.085 (0.027)	***	0.076 (0.018)	***	0.093 (0.016)	***
Earned BA within 6 Years	0.095 (0.050)	*	0.005 (0.034)		0.085 (0.026)	***	0.092 (0.020)	***
Earned BA within 10 Years	0.062 (0.050)		-0.010 (0.034)		0.076 (0.027)	***	0.081 (0.020)	***
Any grad degree w/in 6 years	0.025 (0.022)		0.022 (0.018)		0.001 (0.011)		0.020 (0.011)	*
Any grad degree w/in 10 years	0.039 (0.030)		0.039 (0.024)		-0.001 (0.016)		0.048 (0.014)	***
Any UG loan w/in 5 yrs	-0.125 (0.051)	**	-0.064 (0.034)	*	-0.003 (0.028)		-0.027 (0.021)	
Cum. UG loan over 5 yrs	-$1,232 (956)		-$1,469 (676)	**	$217 (513)		-$958 (399)	**
Any graduate loan w/in 10 yrs	0.026 (0.028)		0.042 (0.023)	*	0.006 (0.015)		0.030 (0.014)	**
Cum. graduate loan over 10 yrs	$983 (1669)		$2,083 (935)	**	-722.2 (1031)		$1,583 (661.5)	**
Any student loan w/in 10 yrs	-0.105 (0.050)	**	-0.052 (0.033)		0.012 (0.027)		-0.011 (0.020)	
Cum. Student loan over 10 yrs	-$275.3 (2145)		$1,048 (1333)		-168.0 (1256)		$359.2 -(867)	
Any parent loan w/in 10 yrs	-0.040 (0.035)		-0.021 (0.023)		0.020 (0.018)		-0.046 (0.013)	***
Cum parent loan over 10 yrs	-$919 (690)		-$1,057 (415)	**	-$347 (348)		-$978 (213)	***
Any employment, Yr10 post	0.017 (0.049)		-0.001 (0.034)		-0.026 (0.027)		-0.000 (0.021)	
Employed year-round, Yr 10	0.038 (0.051)		-0.043 (0.035)		-0.025 (0.028)		-0.028 (0.021)	
Earnings|year round emp, Yr 10	$3,608 (3965)		$2,103 (1732)		-$1,830 (2264)		$3,853 (1209)	***
<u>*Equifax Outcomes*</u>								
Match to Equifax	-0.002 (0.016)		-0.003 (0.022)		-0.011 (0.009)		0.008 (0.014)	
Ever living outside WV	-0.012 (0.039)		-0.003 (0.030)		0.039 (0.023)	*	0.014 (0.019)	
Zipcode: mean income	-$299.3 (1836)		$3,936 (1562)	**	$3,488 (1284)	***	$2,425 (886)	***
Zipcode: bottom AGI quartile	-0.087 (0.039)		-0.022 (0.029)		-0.028 (0.021)		-0.039 (0.018)	**
Changed name (marriage proxy)	n/a		n/a		n/a		n/a	
Ever owned a home	0.138 (0.045)	**	0.013 (0.035)		-0.054 (0.028)	*	0.006 (0.022)	
Log of mortgage value | Mortgage	-0.412 (8.615)		-1.343 (6.588)		-1.294 (2.945)		-3.206 (3.964)	
Credit card balance	-$58 (339)		$137 (239)		-$34 (189)		$79 (152)	
Ever take out a student loan	-0.093 (0.043)		-0.052 (0.035)		0.006 (0.028)		-0.008 (0.022)	
Student loan balance (inc. 0s)	-$166 (2771)		$3,968 (1987)	**	-$515 (1780)		$2,449 (1225)	**
Student loan balance (no 0s)	$5,478 (3903)	*	$8,324 (2833)	***	-$1,622 (2845)		$3,032 (1860)	
Ever past due: stud/auto/home loan	-0.026 (0.047)		-0.023 (0.034)		-0.026 (0.026)		-0.032 (0.020)	
Ever had account in collections	-0.055 (0.046)		-0.041 (0.036)		-0.041 (0.027)		-0.022 (0.022)	
Credit score	9.099 (9.55)		-0.177 (7.275)		2.101 (5.305)		8.579 (4.535)	*
Index	0.048 (0.082)		0.214 (0.146)		0.186 (0.112)	*	0.075 (0.090)	
Sample size	3394		5183		8554		12312	

Notes: Robust standard errors are in parentheses.
*, **, *** indicate the significance of individual findings at the p<0.10, p<0.05, or p<0.01 level.

Appendix Table A.4

Subgroup Analysis: Public and Private HS Attendance

	RD-IV (16<=ACT<=25)		DD-IV (Main specification)	
Outcome	Public	Private	Public	Private
Earned BA within 4 Years	0.095 (0.022) ***	0.061 (0.146)	0.082 (0.012) ***	0.054 (0.089)
Earned BA within 6 Years	0.043 (0.029)	0.099 (0.170)	0.089 (0.016) ***	0.102 (0.101)
Earned BA within 10 Years	0.020 (0.029)	0.118 (0.170)	0.081 (0.016) ***	0.089 (0.104)
Any grad degree w/in 6 years	0.023 (0.014) *	0.101 (0.070)	0.011 (0.008)	0.041 (0.051)
Any grad degree w/in 10 years	0.038 (0.019) **	0.143 (0.094)	0.027 (0.011) **	0.047 (0.067)
Any UG loan w/in 5 yrs	-0.095 (0.029) ***	0.258 (0.182)	-0.022 (0.017)	-0.015 (0.110)
Cum. UG loan over 5 yrs	-$1,345 (558) **	-$453 (3834)	-$544 (319) *	-$1,759 (2096)
Any graduate loan w/in 10 yrs	0.037 (0.018) **	0.060 (0.080)	0.02 (0.010) **	-0.015 (0.060)
Cum. graduate loan over 10 yrs	$1,852 (847) **	$1,468 (7585)	$639 (578)	-$1,612 (4260)
Any student loan w/in 10 yrs	-0.079 (0.028) ***	0.129 (0.187)	-0.005 (0.017)	-0.068 (0.109)
Cum. Student loan over 10 yrs	$812 (1155)	-$2,400 (8946)	$81 (731)	-$4,635 (5098)
Any parent loan w/in 10 yrs	-0.028 (0.020)	0.038 (0.117)	-0.021 (0.011) **	-0.007 (0.062)
Cum parent loan over 10 yrs	-$993 (368) ***	-$600 (2156)	-$720 (190) ***	-$925 (990)
Any employment, Yr10 post	0.006 (0.029)	-0.058 (0.169)	-0.004 (0.017)	-0.195 (0.103) *
Employed year-round, Yr 10	-0.013 (0.029)	-0.033 (0.176)	-0.023 (0.017)	-0.059 (0.108)
Earnings|year round emp, Yr 10	$3,634 (1912) *	-$8,792 (15748)	$2,097 (1219) *	$5,121 (8754)
<u>*Equifax Outcomes*</u>				
Match to Equifax	-0.001 (0.015)	0.001 (0.083)	0.007 (0.010)	-0.004 (0.063)
Ever living outside WV	-0.003 (0.025)	0.009 (0.183)	0.022 (0.015)	0.125 (0.101)
Zipcode: mean income	$2,501 (1204) **	-$58 (8069)	$2,689 (744) ***	$5,806 (5054)
Zipcode: bottom AGI quartile	0.029 (0.021)	0.026 (0.149)	0.019 (0.013)	-0.049 (0.088)
Changed name (marriage proxy)	0.004 (0.022)	0.174 (0.130)	-0.021 (0.013)	0.082 (0.082)
Ever owned a home	0.059 (0.030) **	0.174 (0.200)	-0.014 (0.018)	-0.078 (0.117)
Log of mortgage value | Mortgage	-2.24 -4.84	-20.08 -22.31	-2.37 -2.85	6.95 -20.41
Credit card balance	$76 (199)	-$205 -1185	$12 (199)	244 (862)
Ever take out a student loan	-0.072 (0.030) **	0.065 (0.210)	-0.008 (0.018)	-0.100 (0.113)
Student loan balance (inc. 0s)	$2,846 (1768)	$721 -9610	$1,080 (1032)	-$1,170 (6272)
Student loan balance (no 0s)	$7,803 (2626) ***	$415 (14749)	$1,144 (1593)	-$5,278 (8988)
Ever past due: stud/auto/home loan	-0.028 (0.028)	0.019 (0.174)	-0.036 (0.016) **	0.015 (0.096)
Ever had account in collections	-0.049 (0.030) *	-0.118 (0.191)	-0.038 (0.018) **	0.063 (0.111)
Credit score	5.35 -5.93	-16.07 -39.06	7.59 -3.51 **	-19.80 (18.76)
Index	0.194 (0.121)	0.048 (0.795)	0.154 (0.072) **	-0.166 (0.438)
Sample size	8358	220	20257	609

Notes: Robust standard errors are in parentheses.
*, **, *** indicate the significance of individual findings at the p<0.10, p<0.05, or p<0.01 level.

Appendix Table A.5

Bounding Exercise: Donut Hole Specifications

Outcome	(1) RD-IV Baseline Excluding ACT 20-21			(2) DD-IV Baseline Excluding ACT 20-21		
Earned BA within 4 Years	0.115	(0.033)	***	0.097	(0.015)	***
Earned BA within 6 Years	0.077	(0.045)	*	0.132	(0.021)	***
Earned BA within 10 Years	0.054	(0.046)		0.116	(0.021)	***
Any grad degree w/in 6 years	0.042	(0.023)	*	0.013	(0.010)	
Any grad degree w/in 10 years	0.049	(0.030)	*	0.032	(0.013)	**
Any undergraduate loan w/in 5 yrs	-0.089	(0.047)	*	0.011	(0.021)	
Cum. undergraduate loan over 5 yrs	-$1,011	(889)		-$313	(415)	
Any graduate loan w/in 10 yrs	0.043	(0.028)		0.025	(0.012)	**
Cum. graduate loan over 10 yrs	$1,136	(1620)		$288	(688)	
Any student loan w/in 10 yrs	-0.085	(0.046)	*	0.029	(0.021)	
Cum. Student loan over 10 yrs	$379	(2025)		-$228	(887)	
Any parent loan w/in 10 yrs	-0.004	(0.031)		-0.014	(0.015)	
Cum parent loan over 10 yrs	-$628	(550)		-$134	(407)	
Any in-state employment, Yr10 post	0.038	(0.046)		-0.012	(0.020)	
Employed year-round, Yr 10	0.018	(0.047)		-0.023	(0.020)	
Earnings|year round emp, Yr 10	-$3,631	(2938)		$1,747	(1338)	
Equifax Outcomes						
Match to Equifax data	0.021	(0.025)		0.013	(0.012)	
Marrige proxy	-0.100	(0.071)		-0.031	(0.017)	*
Ever living outside WV	-0.030	(0.040)		0.037	(0.018)	**
Zipcode: mean income	$2,345	(1985)		$4,494	(1102)	***
Zipcode: bottom AGI quartile	-0.036	(0.038)		-0.041	(0.018)	**
Ever owned a home	0.056	(0.048)		-0.018	(0.023)	
Log of mortgage value | Mortgage	-0.182	(0.089)	**	-0.026	(0.205)	
Credit card balance	-$120	(713)		$312	(294)	
Ever take out a student loan	-0.089	(0.049)	*	0.009	(0.023)	
Student loan balance (inc. 0s)	$3,940	(3047)		$450	(1337)	
Student loan balance (no 0s)	$13,857	(5027)	***	-$2,121	(2242)	
Ever past due: stud/auto/home loan	-0.020	(0.045)		-0.051	(0.020)	**
Ever had account in collections	-0.048	(0.049)		-0.053	(0.022)	**
Credit score	1.125	(9.457)		11.64	(4.45)	***
Index	0.120	(0.088)		0.148	(0.041)	***
Sample size	6,274			20,899		

Notes: Robust standard errors are in parentheses.
*, **, *** indicate the significance of individual findings at the p<0.10, p<0.05, or p<0.01 level.

Appendix Table A.6
Exploring Possible Mechanism for Credit Impacts

Outcome	CONTROLLING FOR TIME TO DEGREE (1) RD Baseline	(2) DD Baseline	CONTROLLING FOR CUM. LOAN, 5 YRS (3) RD Baseline	(4) DD Baseline
Match to Equifax data	-0.005 (.015)	.0002 (0.010)	-0.000 (.015)	0.002 (.009)
Ever living outside WV	-0.013 (0.025) *	0.016 (0.014) *	-0.000 (0.025)	0.025 (0.015) *
Zipcode: mean income	$1,873 (1183)	$2,437 (735) ***	$2,429 (1193) **	$2,849 (737) ***
Zipcode: bottom AGI quartile	-0.037 (0.023)	-0.029 (0.014) **	-0.044 (0.023) *	-0.034 (0.014) **
Changed name (marriage proxy)	0.004 (0.022)	-0.023 (0.012) *	0.008 (0.022)	-0.019 (0.013)
Ever owned a home	0.042 (0.029)	-0.036 (0.017) *	0.053 -0.029 *	-0.025 (0.017)
Log of mortgage value | Mortgage	-0.247 (4.617)	-0.813 (2.715)	-1.702 (4.641)	-1.695 (2.703)
Credit card balance	-$5 (197)	-$40 (118)	$106 (196)	$36 (118)
Ever take out a student loan	-0.061 (0.029) **	0.003 (0.017)	-0.029 (0.026)	0.011 (0.015)
Student loan balance (inc. 0s)	$2,324 (1748)	$523 (1018)	$4,491 (1606) ***	$1,646 (951) *
Student loan balance (no 0s)	$7,012 (2586) ***	-$302 (1018)	$8,472 (2526) ***	$1,589 (1549)
Ever past due: stud/auto/home loan	-0.006 (0.027)	-0.02 (0.016)	0.001 (0.027)	-0.031 (0.016) **
Ever had account in collections	-0.024 (0.029)	-0.012 (0.017)	-0.016 (0.029)	-0.028 (0.017) *
Credit score	-1.48 (5.75)	2.232 (3.41)	-3.06 (5.71)	5.8 (3.43) *
Index	0.056 (0.117)	0.0311 (0.069)	0.027 (0.117)	0.117 (0.070) *
Sample size	8358	20257	8358	20257

Notes: Robust standard errors are in parentheses.
*, **, *** indicate the significance of individual findings at the p<0.10, p<0.05, or p<0.01 level.
Unless otherwise noted, regressions include fixed effects for entry cohorts and ACT scores, as well as controls for age, high school GPA and GPA squared, gender, and race/ethnicity, high school type (public/private) and county of residence at entry indicators. The marriage proxy is estimated for the subsample of females only.
"Index" ranges from 0 to 3, the sum of three indicators of unambiguously positive credit outcomes: being in the top quartile AGI, never having past due loans, and never having debt in collections.