

NBER WORKING PAPER SERIES

CASH-ON-HAND & COLLEGE ENROLLMENT:
EVIDENCE FROM POPULATION TAX DATA AND POLICY NONLINEARITIES

Dayanand S. Manoli
Nicholas Turner

Working Paper 19836
<http://www.nber.org/papers/w19836>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2014

This research represents our private research efforts and does not reflect the views or opinions of the U.S. Treasury or the National Bureau of Economic Research. We thank Peter Bergman, Sandy Black, David Card, Raj Chetty, Hilary Hoynes, Larry Katz, Paco Martorell, Judy Scott-Clayton, Steve Trejo, Emmanuel Saez, Eric Zwick and numerous seminar participants for valuable comments and suggestions. Day Manoli is grateful for funding from the Laura and John Arnold Foundation. We also thank Randeep Kaur for excellent research assistance. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2014 by Dayanand S. Manoli and Nicholas Turner. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Cash-on-Hand & College Enrollment: Evidence from Population Tax Data and Policy Nonlinearities
Dayanand S. Manoli and Nicholas Turner
NBER Working Paper No. 19836
January 2014, Revised April 2016
JEL No. H24,I23,I28

ABSTRACT

We estimate causal effects of cash-on-hand on college enrollment decisions of students from low-income families. Using population-level, administrative data from United States income tax returns, we exploit variation in tax refunds received in the spring of the high school senior year. The variation in tax refunds results from the kink point between the phase-in and maximum credit portions of the Earned Income Tax Credit schedule. The results suggest tax refunds received in the spring of the high school senior year have meaningful effects on college enrollment.

Dayanand S. Manoli
Department of Economics
University of Texas at Austin
2225 Speedway Stop C3100
Bernard and Audre Rapoport Building Room 1.116
Austin, TX 78712
and NBER
dsmanoli@austin.utexas.edu

Nicholas Turner
Office of Tax Analysis
U.S. Department of the Treasury
1500 Pennsylvania Avenue, NW
Washington, D.C. 20220
Nicholas.Turner@treasury.gov

1. Introduction

How does cash-on-hand affect college enrollment decisions of students from low-income families? Motivated by the private and social benefits of college education and the differences in college enrollment rates between students from higher and lower income families, a significant body of prior research has examined the impacts of family income on college enrollment.¹ While this research has demonstrated that additional family income can have meaningful impacts on college enrollment decisions, it is still unclear if additional cash transfers at the time of college enrollment decisions can also have meaningful impacts on college enrollment decisions. On the one hand, students may face up-front, out-of-pocket college enrollment costs, and these costs may represent a barrier to entry for students from low-income families. In this case, additional cash-on-hand at the time of college enrollment decisions may have positive impacts on college enrollment rates for students from low-income families. On the other hand, college preparedness and preferences for college enrollment may be entirely determined prior to the time of college enrollment decisions. In this case, additional cash-on-hand at the time of college enrollment decisions may have no impacts on these decisions. Quantifying the causal effects of cash-on-hand is challenging because students from households with more cash-on-hand may also have higher income, higher ability and longer planning horizons, or higher college preparedness, and each of these factors could obscure the causal impacts of additional cash-on-hand.

In this paper, we implement a novel research design to estimate the effects of cash-on-hand on college enrollment decisions of students from low-income families. To study the effects of cash-on-hand at the time of college enrollment decisions, we exploit quasi-experimental variation in tax refunds received during the spring of the high school senior year. The quasi-experimental variation in tax refunds arises from policy nonlinearities in the tax code, or more specifically, from nonlinearities in the benefit schedules for the Earned Income Tax Credit (EITC) and the Child Tax Credit (CTC). These nonlinearities cause observationally similar households to receive different amounts of tax refunds at a time when high school senior dependents are making college

¹ For a sample of references on the returns to schooling, see Card (1999), Currie and Moretti, (2003), Moretti (2004), Lochner (2004) and Milligan, Moretti and Oreopoulos (2004) and Goldin and Katz (2009). We discuss the literature on family income and college enrollment below.

enrollment decisions. More formally, we use population-level administrative tax data, and we implement a Regression Kink Design (RKD) to relate changes in tax refunds at kink points in the EITC and CTC benefit schedules to changes in enrollment rates around these kink points.

The results indicate that an additional \$1000 from tax refunds that arrive in the spring of the high school senior year increases college enrollment by roughly 0.50 percentage points. This result is based on our analysis sample around the first kink point in the EITC benefit schedule. In this sample, average pre-tax income is about \$11000, the average tax refund is about \$5000 and enrollment is about 21%. Further analysis indicates some additional results. First, we find effects of cash-on-hand from tax refunds received in the spring of the high school senior year, but not from tax refunds received in spring of the high school junior year. This indicates that the timing of the additional income may be important. Second, we find evidence of enrollment effects that are persistent up through 4 years after the high school senior year. This indicates that additional cash-on-hand in the spring of the high school senior year leads to new enrollments as opposed to just earlier enrollments. These results also indicate that some students from low-income families may use the additional cash-on-hand to enroll in 4-year college programs. Finally, we do not find evidence of similar cash-on-hand effects for students from higher income families where EITC benefits phase-out completely.

The research design and results offer multiple contributions to the prior literature. First, in the literature examining family income and college enrollment,² it is difficult to find credible quasi-experimental variation in income for low-income families, and even more difficult to find this variation in income at a time when students are making college enrollment decisions. Our research design contributes to this literature by overcoming these difficulties with quasi-experimental variation in tax refunds received during the spring of the high school senior year. Moreover, this research design allows us to examine issues related to timing of additional income, persistence in effects of cash-on-

² For evidence on the effects of long-run and short-run variation in family income on college enrollment, see Ellwood and Kane (2000), Shea (2000), Acemoglu and Pischke (2001), Carneiro and Heckman (2002), Keane and Wolpin (2001), Cameron and Taber (2004), Plug and Vijverberg (2005), Belley and Lochner (2007), Lovenheim (2011), Bailey and Dynarski (2011), Hilger (2013), Michelmore (2013), and Lovenheim and Reynolds (2013).

hand from the senior year and heterogeneity in cash-on-hand effects at different points in the income distribution.

Second, the research design and results contribute estimates of income effects to the literature examining how tax policies and student aid policies affect college enrollment.³ Previous studies generally focus on education tax credits and traditional student aid such as grants. Such aid policies typically operate through price effects, as they lower the relative price of college enrollment if a student enrolls in college. Building on this literature, our findings highlight income effects from the tax refunds. While tax refunds need not be spent on higher education, the lump-sum payments may help households cover any out-of-pocket college costs. Our findings support the idea that additional cash-on-hand increases college enrollment for families that benefit from the EITC, but we do not find similar effects at higher income levels, so similar cash-on-hand effects may not extend to other areas of the income distribution.⁴ Beyond the student aid literature, this study also builds upon earlier studies documenting how tax policies affect child outcomes over the life-cycle.⁵

Third, this study contributes to a growing literature that develops and applies RKD as an empirical strategy to estimate causal effects based on policy nonlinearities.⁶ The analysis is based on a large sample size: nearly all high school seniors in the United States between 2001 and 2011. This population-level administrative data allows us to

³ For evidence on credit constraints, student aid and college enrollment decisions, see van der Klaauw (2002), Dynarski (2003), Stinebrickner and Stinebrickner (2008), Nielsen, Sorensen and Taber (2010), Gurgand, Lorenceau and Melonio (2011), Lochner and Monge-Naranjo (2011), and Solis (2012). These papers primarily exploit quasi-experimental variation in enrollment-contingent student aid. Dynarski and Scott-Clayton (2013) summarize the student aid literature and indicate that additional \$1000 of student aid increases college enrollment by roughly 2 to 4 percentage points. For evidence on the effects of education tax credits on college enrollment, see Long (2004), Turner (2011a,b), LaLumia (2012) and Bulman and Hoxby (2015).

⁴ Recent work by Bulman and Hoxby (2015a,b) examines the impact of tax-based federal student aid over a broader segment of the income distribution and finds no enrollment effects. Bulman and Hoxby (2015a,b) study the American Opportunity Tax Credit and the Tuition and Fees Deduction. Unlike the EITC from the high-school senior year, this credit provides students and their families with a benefit after they incur college costs. This timing difference may also account for the different patterns of enrollment.

⁵ Hoynes, Miller and Simon (2015) present estimates of the impacts of EITC benefits on birth weight; Dahl and Lochner (2012) present estimates of the impacts of EITC benefits on early age test scores; Michelmore (2013) studies the impacts of state and federal EITC benefits on college enrollment, though the identification strategies and treatment populations differ from those used in this study.

⁶ See Calonico, Cattaneo, and Titiunik (2014), Landais (2015), Card Lee, Pei and Weber (2015), Ganong and Jaeger (2014), Marx and Turner (2015), Card, Johnston, Leung, Mas and Pei (2015), Hansen, Nguyen and Waddell (2015), and Gelber, Moore and Strand (2016).

implement a research design based on slope changes around tax kink points. Since it is unlikely that other factors change exactly at the tax kink points, this research design offers highly credible estimates of causal effects. As RKD relies on identifying kinks in the enrollment-income profile, it is important to distinguish between kinks and nonlinearities in the enrollment-income function. We consider multiple strategies to address this methodological concern. Specifically, in addition to looking at further analyses of timing, persistence and other points in the income distribution, we also consider placebo analyses, controlling for income polynomials and other strategies.

This paper is organized as follows. Section 2 describes the data, institutional background and cross-sectional analysis. Section 3 presents the main empirical analysis to estimate causal effects of tax refunds on college enrollment. Section 4 discusses evidence on liquidity and informational constraints to provide some context for the estimates presented in Section 3. Section 5 discusses the conclusions from the analysis.

2. Background & Data

2.1 The Earned Income Tax Credit and the Child Tax Credit

The EITC is a refundable tax credit that provides benefits to low-income working families. As the credit is refundable, taxpayers may benefit from the EITC even when they have no tax liability. The EITC amounts are primarily determined based on tax filing status, the number of qualifying children, and income.⁷ Qualifying children for the EITC are relatives who are under age 19 or permanently disabled and who resided with the tax filers for at least half of the year.⁸

The EITC benefit schedule is characterized by three regions: the phase-in region where the credit is increasing in income, the maximum credit region where the credit value is constant in income, and the phase-out region where the credit is decreasing in income. The value of the credit is a function of earned income and adjusted gross income (AGI). Earned income includes wages, salaries and tips and net earnings from self-employment. Benefits from unemployment insurance, workers compensation, food stamps, Medicaid, TANF, SSI, social security, disability and child support do not count

⁷ Eligibility for the EITC also includes a ceiling on investment income, from such sources as dividends, rental properties etc. In 2014, the limit on investment income was \$3350.

⁸ Children between ages 19 and 24 can also count as qualifying children if they were full-time students for any five months of the calendar year.

as earned income. Adjusted gross income includes total gross income for the taxpayer minus adjustments for certain expenses.⁹ As the name suggests, taxpayers must have positive earned income to claim the EITC. Starting from no earned income, the credit is phased in. As earned income continues to increase, taxpayers reach the maximum credit region of the EITC schedule.

The income variable for determining the value of EITC benefits changes at the beginning of the phase-out portion of the schedule. The rules for determining benefits are as follows.¹⁰ If earned income is equal to AGI, EITC benefits are computed using earned income. If earned income is not equal to AGI and AGI is below a phase-out earnings threshold, then EITC benefits are computed using earned income. If earned income is not equal to AGI and AGI is greater than or equal to the phase-out threshold, then EITC benefits are computed using the maximum of AGI and earned income. Appendix Table 1 lists the exact phase-out AGI thresholds.

Following these rules, we define EITC income as the income that is used to compute EITC benefits. Figure 1A illustrates a stylized EITC benefit schedule in the case where AGI equals earned income for a family with two dependents. The figure highlights that the EITC benefit schedule has three kink points: EITC Kink 1 where the phase-in portion meets the maximum benefit portion, Kink 2 where the maximum benefit begins to phase out and Kink 3 where the credit becomes fully phased out.¹¹ Appendix Table 1 lists the exact values for EITC Kink 1, and Appendix Table 2 lists the exact values for EITC Kinks 2 and 3.

In addition to the EITC, the Child Tax Credit (CTC) and Additional Child Tax Credit (ACTC) offer taxpayers benefits of up to \$1000 per qualifying child. These two tax credits are effectively a single tax credit with the CTC being the non-refundable portion and the ACTC being the refundable portion.¹² We refer to the combined CTC and ACTC as the CTC. Like the EITC, the CTC phases in with earnings and is phased out at higher levels of income. From 2001 through 2008, the CTC begins to phase in at EITC

⁹ The exact adjustments for computing AGI are listed on IRS Form 1040 (<https://www.irs.gov/pub/irs-pdf/f1040.pdf>)

¹⁰ IRS Publication 596 provides the official documentation of the rules and eligibility criteria for this credit.

¹¹ We recognize that there is also a kink point where the credit begins to phase in. However, empirically there are very few families with high-school seniors at this low level of income.

¹² IRS Publication 972 provides the official documentation of the rules and eligibility criteria for this credit.

Kink 1 (either near the 2 qualifying children Kink 1 value in 2001-2007 or near the 1 qualifying child Kink 1 value in 2008; Appendix Table 1 lists the income thresholds at which the CTC phases in). Unlike the EITC, qualifying children for the CTC must be younger than age 17. As the analysis sample consist of taxpayers with high school senior dependents who are age 17 or 18, the CTC applies only for those taxpayers in the analysis sample who have a younger child in addition to the high school senior dependent. Figure 1A presents an example of the CTC benefit schedule for a married filing jointly family with two dependents in 2005.

2.2 EITC Kink 1

Within a given year, the first EITC kink point varies only by the number of qualifying children. Appendix Table 1 lists the specific earnings thresholds for each tax year for the EITC as well as the earnings thresholds for the CTC. In the EITC phase-in portion, EITC benefits increase by \$0.34, \$0.40 and \$0.45 per dollar of earned income based on one, two, or three or more qualifying children respectively. In some years, the CTC begins to phase in at the first EITC kink point so that tax refunds continue to increase as earnings increase in the maximum credit region of the EITC benefit schedule.

To provide intuition for our RKD specification at the first EITC kink point, Figure 1B plots a simulated example of the change in tax refunds at EITC Kink 1. The figure illustrates that tax refunds increase at a faster rate for earnings levels below the kink point because EITC benefits increase in the phase-in region. To the right of the kink point, EITC benefits no longer continue to phase-in since individuals are in the maximum credit region. Intuitively, the slope of tax refunds decreases as tax refunds no longer increase at the higher rate. In some cases, tax refunds still continue to increase to the right of the kink because the CTC begins to phase in at a rate of \$0.10 per dollar of income at the first EITC kink point, though this credit does not apply to all families in the analysis sample. For families without a CTC qualifying dependent (a dependent child under age 17), the slope change in tax refunds at the first EITC kink point is roughly \$0.34 because the slope changes from \$0.34 (the EITC phase-in rate) to 0. For taxpayers in the analysis sample that have one high school senior dependent and at least one other younger child, the CTC does apply and the slope change in tax refunds at the first EITC kink point is roughly \$0.30 because the slope changes from about \$0.40 (the weighted average over

the EITC phase-in rates based on the fractions of taxpayers with two, or three or more qualifying children) to roughly \$0.10 (the Child Tax Credit phase-in rate).

2.3 Data

To analyze the effect of cash on hand on college enrollment we use information from the population of U.S. tax returns and from the Social Security Administration (SSA). To focus on high school seniors, we create our sample by first pulling all Social Security Numbers (SSNs) from the SSA data for individuals who are 17 or 18 during the years 2001 to 2011. For these observations, we assign high school cohorts based on the month and year of birth. In each year, the cohort of high school seniors is defined as individuals who were 18 as of December 31 and who were born in September through December and individuals who were 17 as of December 31 and who were born in January through August. In aggregate, this approach matches well to the number of high school seniors reported by the Department of Education. For example, for 2007 the U.S. Department of Education reports a total of 4.21 million high school seniors, whereas we find 4.09 million in the tax data.¹³ Next, we look for tax returns that claim these individuals as dependents during the sample period, retaining information on family structure (married, number of dependents) and income from the 1040 tax form. Given our focus on tax returns claiming high school seniors, we restrict the sample to returns that file as either head of household or married filing jointly.¹⁴

To measure college enrollment, we use the 1098-T tax form. To remain eligible for Title IV federal student aid, schools are required to send a 1098-T form to nearly all students, and to the IRS.¹⁵ Chetty et al (2011) and Chetty, Friedman and Rockoff

¹³ This approach may misclassify some individuals. In particular, since we do not have data directly from schools on their senior students, their graduation and subsequent enrollment, our ability to specifically identify high school seniors, on-time high school graduation, and college enrollment may be limited in the tax data. However, such misclassification is not likely to be problematic for our specifications. In the RKD case, as long as this misclassification does not vary across the EITC kink point we examine, this will not have an effect on our estimates. Intuitively, measurement error in defining the senior year may impact the average enrollment rate but should not have a differential effect on the slope of the enrollment profile at the tax kink. Further, we find evidence that the age profile is smooth through the tax kinks. For the difference in difference analysis, as long as the misclassification is constant across the different family structures affected by the policy change, then the measurement error will be differenced out.

¹⁴ There are relatively few married filing separate or qualifying widower returns claiming a high school senior.

¹⁵ This form is used to verify educational expenses for certain tax-based aid programs. Exceptions to the 1098-T filing rule include: courses for which no credit is earned; nonresident alien students; and students whose qualified tuition is covered by a formal billing arrangement between the institution and the student's

(2013a,b) also use the 1098-T to measure college enrollment. Chetty et al (2011) find that enrollment from as measured by the 1098-T form is comparable to enrollment reported in other data including the Current Population Survey and the U.S. Department of Education.

2.4 Analysis Sample

To construct the analysis sample we include individuals that are -\$6000 to +\$3000 (in earned/EITC income) around the first EITC kink point. We restrict to at most +\$3000 to avoid interactions with the second kink point. Additionally, we exclude high school seniors who died at any time between 2001 and 2012, late-filed tax returns (so we can ensure that the tax refunds are received in the spring of the high school senior year and not later), taxpayers with any self-employment income, and taxpayers with more than a \$1000 difference between W2 earnings and total earned income. We impose the non-W2 income restrictions because many non-W2 income sources cannot be third party verified. Previous studies (see Saez 2010 and Chetty, Friedman and Saez 2012) highlight evidence of individuals with these income sources sorting along the tax schedule, which violates the identifying assumptions behind the RKD approach Card, Lee, Pei and Weber (2015). We also remove a small number of returns that are likely to have errors in key income measures.¹⁶

Table 1 shows summary statistics for the analysis sample. The EITC Kink 1 sample consists of roughly 1.4 million high school seniors between 2001 and 2011. Mean enrollment for this sample is roughly 21 percent, suggesting that as many as four out of five children from these families do not attend college in the fall after their senior year. Roughly one in five families are married, with the remainder coming from single-parent households, of which nearly all are female headed families. On average, there are 1.7 children per family, so that many families in the sample benefit from both the EITC and the CTC. These tax refunds help boost income to \$16,200 on average after taxes, compared to just over \$11,300 of pre-tax income.

employer. As a check we also verify that our enrollment results do not change when we use data from the universe of Pell grant recipients, including those with Pell grants that cover all of their education costs who may not receive a F1098-T form.

¹⁶ We also exclude taxpayers with large differences between taxpayer entered values and computer verified values for total income, AGI, tax refunds and EITC benefits. Specifically, taxpayers with differences of more than \$100 in total income, AGI or tax refunds, and differences of more than \$10 in EITC benefits, are excluded.

3. Empirical Analysis

3.1 Regression Kink Design

To identify the impact of cash-on-hand on college enrollment near the first EITC kink point we use a fuzzy regression kink design.¹⁷ This approach relates the change in the slope of the enrollment function to the change in slope of tax refunds at the tax kink point. To implement the fuzzy RKD, we estimate both the change in the enrollment-income profile and the tax refund-income profile at EITC Kink 1. The estimate of the impact of cash-on-hand on college enrollment is the ratio of these slope changes at the tax kink point. We compute earned income relative to that kink point, denoted by *kinkdist*. This measure allows us to pool the data across groups to estimate changes in the slopes of enrollment and after-tax income at each kink point. We exploit differences in the location of these kink points across groups in the placebo tests.

Following Card, Lee, Pei and Weber (2015) and Nielsen et al (2010), we consider the following constant-effect, additive model to examine the effects of refunds on college enrollment,

$$enroll_i = \beta refund_i + g(kinkdist_i) + \varepsilon_i.$$

The subscript *i* refers to the individual who is a high school senior. The variable *enroll_i* is an indicator equal to one if individual *i* enrolls in college in the year after his or her high school senior year. The variables *refund_i* and *kinkdist_i* are based on tax returns filed in the spring of individual *i*'s senior year on which individual *i* is claimed as a dependent. The *refund_i* variable measures the tax refund and *kinkdist_i* measures the distance (\$2011) relative to the specified kink point. The function *g(.)* is a continuous function. The tax refund function, *refund_i=refund(kinkdist_i)*, is assumed to be a continuous and deterministic function of earnings (equivalently of earnings relative to the kink point) with a slope change at the kink point (i.e. at *kinkdist=0*). If *g(.)* and *E(ε|kinkdist=k)* have

¹⁷ Even though the tax refund function is deterministic we use a fuzzy RKD rather than a sharp RKD. The fuzzy approach allows us to empirically estimate the change in slope of tax refunds and show that it matches the statutory slope change. The trade-off of using the fuzzy RKD in place of the sharp RKD is a potential loss of precision. As a result, implementing the fuzzy RKD should result in relatively more conservative inferences about the impact of tax refunds on enrollment, compared to using a sharp RKD specification.

derivatives that are continuous in *kinkdist* at *kinkdist* = 0, then the fuzzy RKD estimator is given by

$$\beta = \frac{\lim_{k \rightarrow 0^+} \frac{\partial E[\text{enroll} | \text{kinkdist} = k]}{\partial k} \Big|_{k=0} - \lim_{k \rightarrow 0^-} \frac{\partial E[\text{enroll} | \text{kinkdist} = k]}{\partial k} \Big|_{k=0}}{\lim_{k \rightarrow 0^+} \frac{\partial E[\text{refund} | \text{kinkdist} = k]}{\partial k} \Big|_{k=0} - \lim_{k \rightarrow 0^-} \frac{\partial E[\text{refund} | \text{kinkdist} = k]}{\partial k} \Big|_{k=0}}.$$

The numerator of this expression captures the change in the slope of the conditional expectation of enrollment with respect to income at the kink point. The denominator reflects the change in the slope of tax refunds at the kink point.

We estimate the changes in enrollment and after-tax income, for the above numerator and denominator respectively, using regressions of the following form

$$\begin{aligned} \text{enroll}_i &= \alpha \text{kinkdist}_i + \delta^{\text{enroll}} D_i * \text{kinkdist}_i + \alpha_2 X_i + u_i \\ \text{refund}_i &= \gamma \text{kinkdist}_i + \delta^{\text{refund}} D_i * \text{kinkdist}_i + \gamma_2 X_i + v_i. \end{aligned}$$

where D_i is an indicator variable equal to one if earnings fall below the kink point, i.e. $D_i = 1(\text{kinkdist}_i < 0)$. The variable X denotes a vector of covariates included in the regressions. The fuzzy RKD estimator is then given by

$$\hat{\beta} = \frac{\widehat{\delta^{\text{enroll}}}}{\widehat{\delta^{\text{refund}}}}.$$

The vector of covariates includes dummies for year, filing status, and number of kids. Intuitively, the coefficient $\hat{\beta}$ reflects the impacts on enrollment of additional cash-on-hand coming from increases in tax refunds, or equivalently from increases in after-tax income, in the spring of the high school senior year.

When estimating these enrollment and refund regressions for Kink 1, we choose a baseline bandwidth of (-\$6000,+\$3000) around the kink point, which is the largest possible bandwidth that does not overlap with any other tax kink points.¹⁸ The large bandwidth also allows us to account for nonlinearity in the enrollment-income profile that is separate from discontinuous changes at the tax kink point. While it is possible to use different bandwidths for the numerator and denominator, we present results based on using the same bandwidths for the enrollment and after-tax income regressions. In this

¹⁸ The choice of -\$6000 is motivated by choosing a lower bound that does not get too close to the zero earned income threshold; since the lowest value of EITC Kink 1 is \$7140 in 2001, a distance of -\$6000 relative to this kink point is just slightly above the zero earnings threshold. The choice of the +\$3000 is motivated by choosing the highest bound that does not reach the beginning of the phase-out region (i.e. the lowest value of EITC Kink 2, see Appendix Tables 2 and 3 for key EITC points).

case we estimate the fuzzy RKD using an instrumental variables approach based on estimating the following regression

$$enroll_i = \beta refund_i + kinkdist_i + \varepsilon_i$$

in which we instrument for tax refunds using the interaction $D_i * kinkdist_i$.

3.2 Regression Kink Estimates

Figure 2 presents the main graphical evidence for the regression kink analysis. Figure 2A plots tax refunds against earnings relative to the kink point, and Figure 2B plots enrollment against earnings relative to the kink point. We construct these plots by computing average tax refunds and enrollment rates within \$100 bins of income relative to the respective kink points. We generate the fitted values by using the individual-level data and regressing the enrollment indicator on $kinkdist_i$ and $kinkdist_i * D_i$. We then plot the average of the fitted values in each \$100 bin of income relative to the kink. Note that while we bin the data in \$100 increments for the figures, the estimated slopes in the figures correspond to regression output from underlying individual-level data. Figure 2A highlights a kink in tax refunds at the first EITC kink point and Figure 2B shows a slope change in enrollment rates at the same kink point.

Table 2 presents the quantitative results corresponding to the graphical evidence in Figure 2. The estimated slope change in enrollment is -0.15 and the first stage change in the slope of after-tax income (tax refunds) is -0.34. Using the IV specification to estimate the ratio of these two coefficients, the RKD estimates indicate that a \$1000 increase in after-tax income (tax refunds) causes roughly a 0.45 percentage point increase in college enrollment. We get similar results when we consider enrollment that is at least half-time, with an IV result that is 0.42 percentage points per \$1,000.¹⁹ We find similar results for 1 and 2 child households separately, though in each case the standard errors are sufficiently large that we cannot rule out zero effects. Note that the income point corresponding to Kink 1 varies across 1 child and 2 child cases so that the separate results suggest that there is an enrollment kink for each group at the same location as the tax kink point but that the enrollment kink point occurs at different income levels. We return to this point below in the robustness checks, where we show that there is no

¹⁹ We are not able to perfectly observe enrollment intensity, as the F1098-T form indicates only whether enrollment is at least-half time or not.

corresponding kink for the alternate case (i.e. there is no enrollment kink among 1 child families at the income level corresponding to the kink for 2 children families).

3.3 Evidence on the Identifying Assumptions

Identification with the RKD methodology requires that (1) taxpayers do not sort along the tax schedule and that (2) other covariates do not change in the tax kink points. This section presents evidence that both of these key assumptions hold at the first EITC kink point.

To study sorting along the tax schedule, we examine frequencies of taxpayers around the kink point. Figure 2C presents plots of the frequencies of taxpayers around EITC Kink 1. Prior to our sample restrictions (i.e. when we include all tax returns around the first EITC kink point), we find significant evidence of bunching around the kink point. This is consistent with previous evidence in the income tax literature (see Saez 2010 and Chetty et al 2013). After excluding individuals with self-employment earnings or other non-third party verified income, as well as individuals with more than a \$1000 difference between earned income and AGI, we find no evidence of sorting along the tax schedule in the sample.²⁰

To examine if any covariates change at the tax kink point, we regress enrollment on a set of covariates, obtain predicted enrollment values, and then test for a kink in predicted enrollment using the above RKD regression specifications. Intuitively, if the aggregate effect of the covariates in our specification has a kink at the tax kink point, then the predicted enrollment values will also have a kink.²¹ However, as none of the covariates are expected to change at the tax kink point, we can verify that there is no evidence for a kink in the covariate predicted enrollment rates. In addition to dummies for year, filing status and number of kids, we also include linear controls for senior year

²⁰ We follow Card, Lee, Pei and Weber (2015) and formally test for a kink in the frequencies by estimating a series of polynomial models using the binned frequencies. The polynomial models allow the first and higher order derivatives to change at the kink point, and we test for a kink in the frequencies based on whether or not there is a statistically significant change in the linear piece of the polynomial at the kink point. Overall, the frequencies appear smooth.

²¹ Note that it is possible for the aggregate effect of the controls to be smooth through the kink while the individual controls have offsetting kinks. We do not find any evidence for this possibility. Instead, we find no evidence that each control has a kink. We use the predicted enrollment measure as a convenient way to summarize this overall finding.

income and junior year income (i.e. income in the calendar years of the start of the senior and junior years respectively) when computing covariate predicted enrollment values.

Figure 2D presents the graphical analysis of covariate predicted enrollment. The plot shows that there are no detectable changes in covariate predicted enrollment when predicting enrollment using a rich set of covariates from the tax data. Additionally, while the tax data do not contain data on federal student aid eligibility, we have verified that there are no specific changes in federal student aid eligibility that correspond to the income levels of the tax kink points. Individuals in the first EITC kink sample generally qualify for zero Expected Family Contribution and maximum Pell grants.

3.4. Robustness

We examine the robustness of the RKD estimates in a variety of ways. First, a natural concern may be that the RKD estimates result from a nonlinear relationship between enrollment and earnings. However, we emphasize that the RKD estimates are based on variation in earnings relative to EITC Kink 1 and not just variation in earnings levels. The variation in the EITC Kink 1 values across households with different numbers of qualifying children allows us to control for nonlinear enrollment-earnings relationships that are not co-linear with the RKD running variable (earnings relative to the kink). We can control for earnings polynomials that account for nonlinear enrollment-earnings relationships that are common across households. However, our identification strategy does not allow us to control for separate nonlinear enrollment-earnings relationships across households with different numbers of qualifying children because the RKD estimates are based exactly on this variation.

We start to address the concerns about nonlinear enrollment-earnings relationships by first plotting the empirical relationship between enrollment and earned income for the EITC Kink 1 earnings range. Using this data, we create \$100 bins of CPI-adjusted earned income and compute mean enrollment in each bin. Figure 3 presents a plot of mean enrollment (y-axis) for each earned income bin (x-axis). The graphical evidence suggests perhaps a relatively linear relationship between enrollment and earned income for this income range, or at least there does not appear to be evidence of a downward kink similar to the reduced form evidence plotted in Figure 2B. This suggests

that the reduced form evidence for the RKD estimates emerges only when earned income is re-centered relative to EITC Kink 1 for each group.

We further address concerns about nonlinearity in the enrollment-income relationship by explicitly accounting for this relationship when estimating the RKD results. We find that the baseline results are robust to a number of different income controls, including adding a 5th order polynomials in earned income and AGI, and separately 5th order polynomials in earned income from the senior and junior years.²² Moreover, in the controls with junior year income, we also include controls for junior year earnings relative to EITC Kink 1 as well as an interaction between this earnings relative to the kink and an indicator for earnings being above the junior year kink (i.e. we include the analogous variables to $kinkdist_i$ and $kinkdist_i * D_i$ based on junior year earned income). These results are in Appendix Table 3. Overall, these findings are consistent with the interpretation that the RKD estimates at EITC Kink 1 are not driven by nonlinearity in the enrollment-income relationship.

A second natural concern about the robustness of the RKD estimates at EITC Kink 1 relates to whether the estimated slope changes in tax refunds and enrollment are actually occurring at the statutory kink points, or if the estimated changes result from nonlinearities and/or kink points close to EITC Kink 1. In the spirit of the permutation test suggested by Ganong and Jaeger (2014), we implement a placebo test to address this concern. For this analysis, we vary a placebo kink around the true kink point and verify that the largest estimated kink in tax refunds and enrollment occurs at the true EITC kink point. We choose a distance from the true kink point $p = -4000, -3900, \dots, +2000$ and define a placebo kink point based on this distance, $pkink = kink1 + p$. Using this placebo kink point, we define earnings relative to the placebo kink point, $pkinkdist_i$ and an indicator $D_i^p = 1(pkinkdist_i < 0)$. We then estimate the slope changes in tax refunds and enrollment at the placebo kink using the following regressions,

$$\begin{aligned} enroll_i &= \alpha_1 pkinkdist_i + \alpha_2 D_i^p + \delta^{enroll,p} [D_i^p * pkinkdist_i] + \alpha_2 X_i + u_i \\ refund_i &= \gamma_1 pkinkdist_i + \gamma_2 D_i^p + \delta^{refund,p} [D_i^p * pkinkdist_i] + \gamma_2 X_i + v_i \end{aligned}$$

²² We define junior year earned income as earned income for the primary taxpayer (parent) in the year prior to the dependent's high school senior year. Because data for some information returns is only consistently available in 2001 onwards, we cannot track parents of the 2001 high school senior cohort to prior years. Thus, when including junior year income controls, the estimates are only based on the 2002-2011 cohorts.

We then plot the estimated slope changes, $\delta^{enroll,p}$ and $\delta^{refund,p}$, for each of the placebo kink points $p = -4000, -3900, \dots, +2000$ and verify that the highest estimated slope changes occurs at the true kink point.

Figure 4 presents the estimated slope changes when varying the placebo kink points in \$100 increments around the true EITC kink points. Figure 4A presents the estimated placebo slope changes for tax refunds, and Figure 4B presents the estimated slope changes for enrollment. The plots highlight that for tax refunds and enrollment, the largest estimated kink points seem to occur at Kink 1. This is consistent with the assumption that the RK estimates result from the kink in tax refunds at the statutory kink point, and not from a spurious nonlinear relationship between enrollment and earnings relative to the group-specific kink point.

We also examine the robustness of the RKD estimates at EITC Kink 1 in relationship to (1) additional control variables and alternative sample restrictions, (2) alternative bandwidths, (3) and alternative clustering of the standard errors. The results are presented in Appendix Tables 4, 5, and 6 respectively. In general, the baseline estimate appear robust to these changes. Finally, Appendix Table 7 presents estimates based on including quadratic and cubic function of the running variable. Specifically, we estimate regressions of the following form for the first stage and reduced form (the IV specification is analogous):

$$\begin{aligned} enroll_i &= \alpha g(kinkdist_i) + \delta^{enroll} D_i * g(kinkdist_i) + \alpha_2 X_i + u_i \\ refund_i &= \gamma g(kinkdist_i) + \delta^{refund} D_i * g(kinkdist_i) + \gamma_2 X_i + v_i. \end{aligned}$$

where $g(\cdot)$ denotes a linear function (baseline), a quadratic function, or a cubic function. In general, the standard errors for all of the estimates increase as we move from the linear function of the running variable to the quadratic and cubic functions. The first stage results are essentially unchanged. With the quadratic function, the reduced form results continue to indicate a statistically significant downward kink in enrollment at EITC Kink 1. Consistent with the reduced form, the IV estimates in the quadratic case indicate a statistically significant impact of cash-on-hand on enrollment, but the standard errors are large enough so that the baseline estimate from the linear case (0.45) is included within the 95% confidence interval. In the cubic case, we do not detect any statistically impacts

of cash-on-hand on enrollment as the standard errors are more than double the standard errors in the quadratic case.

3.5 Timing of Income and Enrollment

Our baseline results relate tax refunds in the high school senior year to college enrollment the next year. To determine the relative importance of the timing of tax refund payments, we examine whether tax refunds in the spring of the high school junior year also affect college enrollment. For this analysis, we replicate the RKD based on income in the high school junior year. In particular, we draw the sample of returns that are around the first EITC kink point in the high school junior year.²³ Thus, with this junior year sample, the first stage is based on the slope change in tax refunds at EITC Kink 1 in the junior year.

Table 3 presents regression kink estimates for both the junior and senior EITC Kink 1 samples.²⁴ In both samples, we test for kinks in tax refunds in the junior year and senior year. Thus, we verify that in the junior year K1 sample, we find a kink in tax refunds in the junior year and not in the senior year, and similarly for the senior year K1 sample, we find a kink in refunds in the senior year and not the junior year. The reduced form column highlights the main results for these samples: there does not appear to be any evidence of a kink in enrollment based on junior year earnings relative to EITC Kink 1. Thus, the estimates indicate that changes in tax refunds in the high school senior year affect college enrollment, but changes in tax refunds in the high school junior year do not.

We next examine the persistence of the enrollment effects from tax refunds received in the spring of the high school senior year by estimating the on enrollment in later years. On the one hand, if enrollment decisions just after high school are persistent then we expect that there would be a kink in college enrollment rates in later years. On the other hand, if enrollment decisions are not persistent so that some individuals who

²³ Similar to the analysis based on tax returns from the high school senior year, we examine evidence on the identifying assumptions for the regression kink analysis in the high school junior year, and we verify (1) that taxpayers do not sort along the tax schedule in the junior year and (2) that there is no kink in covariate predicated enrollment.

²⁴ The senior year K1 sample used in Table 3 is not identical to the baseline EITC K1 sample because it only includes cohorts 2002-2011, whereas the baseline sample includes the 2001 cohort. Because some data for parent taxpayers is only available from 2001 onward, we cannot compute junior year earned income for the 2001 cohort.

were enrolled drop out and others who were not enrolled become enrolled, then we would not expect to find evidence of a kink in enrollment two years after the high school senior year.

We estimate the impacts of tax refunds in the spring of the senior year on these additional enrollment outcomes by estimating separate regressions for each year after the high school senior year.²⁵ Figure 5A presents the IV estimates of the impacts of tax refunds in the spring of the high school senior year on enrollment in each year after the high school senior year. The baseline RKD estimate from Table 2 corresponds to the estimate shown at 1 year after the senior year. While the standard errors are sufficiently large that many of the estimates may not be statistically different from 0 or from each other, we focus on the pattern of the point estimates. The point estimates suggest a positive effect in each of the first four years, consistent with the interpretation of increased enrollment and persistence at four-year colleges. To put the estimates from Figure 5A into context, Figure 5B plots mean enrollment for each year after high school.²⁶

3.6 Heterogeneity by Geography and Prior Income

Motivated by recent evidence that shows meaningful differences in intergenerational mobility across geographic locations in the United States (Chetty, Hendren, Kline and Saez 2014), we examine geographic heterogeneity in enrollment.²⁷ Figure 6 illustrates a heat map of county-level enrollment rates for high school seniors in the 2001 cohort.²⁸ This map is based on the full population of high school seniors in the

²⁵ We note that, because data is available only for cohorts 2001-2012, this analysis drops one cohort for each additional year after the high school senior year. For example, we can use cohorts 2001-2011 when looking at enrollment 1 year after high school, cohorts 2001-2010 when looking at enrollment 2 years after high school, cohorts 2001-2009 when looking at enrollment 3 years after high school, etc.

²⁶ Figure 5B indicates that, 10 years after high school (i.e. at age 27 and 28), enrollment is at roughly 13%. This estimate is consistent with population estimates from the National Center for Education Statistics (NCES) and Census:

http://nces.ed.gov/programs/coe/indicator_cea.asp

http://nces.ed.gov/programs/digest/d14/tables/dt14_103.20.asp

<http://www.census.gov/hhes/school/data/cps/2014/tables.html> [Table 2]

²⁷ We have also examined heterogeneity across (1) urban versus rural counties and (2) high versus low EITC bunching counties (Chetty, Friedman and Saez 2014). There did not appear to be any systematic differences across these dimensions. These results are available upon request.

²⁸ For this map, we restrict the sample of counties based on the sample size with the county. Specifically, we compute the number of high school seniors in each county in each year, and we restrict the mapped counties to those that are in the top 75% of the sample size distribution in 2001 (i.e. those counties that

respective cohort, not just the seniors in the EITC Kink 1 sample, and enrollment is measured in the year after the high school senior year. The heat map illustrates significant geographic heterogeneity in college enrollment. Lower enrollment places have enrollment rates below 30%, and high enrollment places have enrollment rates about 55%. The high enrollment places are concentrated in the Midwest, Northeast and California coast. Appendix Figure 1 illustrates similar heat maps for the 2006 and 2011 high school senior cohorts, and overall, the geographic heterogeneity in enrollment appears consistent across years.

Using the county-level enrollment rates in each year, we calculate county-enrollment quintiles. Specifically, in each year, we categorize counties into the top 20% of enrollment rates, the next 20%, etc. down to the lowest 20% of enrollment rates. The first row of Table 4 presents the average county-level enrollment rates for counties in each quintile. Similar to the heat maps, the average county enrollment rates range from below 30% for counties in the lowest quintile to above 60% in counties in the highest quintile. Table 4 also presents average enrollment for seniors in the EITC Kink 1 sample who reside in counties in each quintile. Consistent with prior literature highlighting the importance of geographic locations in determining economic outcomes (Chetty et al 2014 and Chetty et al 2016), we note that, even conditional on being in the EITC Kink 1 sample, enrollment rates vary across the county-enrollment quintiles, ranging from below 15% in the lowest county quintile to roughly 30% in the highest county quintile.

Below the enrollment rates, Table 4 presents the regression kink estimates for each county-enrollment quintile. The first stage slope changes in tax refunds are essentially identical across county-quintiles as expected. The reduced form slope changes in enrollment vary slightly across the quintiles. Given that the IV point estimates are all positive and of similar magnitudes, and given that the standard errors are sufficiently large that we cannot rule out that the estimates are equal across quintiles, we conclude that cash-on-hand effects appear to be similar for students from low-income families in higher and lower enrollment counties.

have at least 144 high school seniors in 2001, and 144 is the 25th percentile of the sample size distribution across counties in 2001).

Next, we examine heterogeneity based on prior income to understand if cash-on-hand effects are similar across more permanently low-income families and more transitorily low-income families. Intuitively, cash-on-hand effects may be smaller for transitorily low-income families since these families may be less likely to be liquidity constrained and students from families with higher permanent income may have higher college preparedness. For this analysis, we restrict the sample to high school senior cohorts in 2005 through 2011 and then calculate average AGI over the 4 years prior to the high school senior year. We also restrict the sample to taxpayers who do not have any self-employment income in any of the four years prior to the high school senior year. Because the income levels for EITC Kink 1 differ based on the number of EITC qualifying children and year, we create quintiles of average AGI within groups based on the high school senior year and the number of qualifying children. We then estimate the RKD regression specifications within the samples defined by the average AGI quintiles.²⁹

Table 5 presents the results based on average AGI in the 4 years prior to the high school senior year. First, the full sample results verify that for this sample of high school seniors in the 2005-2011 cohorts, we find similar cash-on-hand effects as in the baseline EITC Kink 1 sample based on the 2001-2011 cohorts. In particular, the results in Table 5 serve as a robustness check for the baseline results in Table 2 since the sample is further restricted to not have any self-employment income in any of the four years prior to the high school senior year, whereas the baseline EITC Kink 1 sample only imposed a sample restriction based on not having any self-employment income in the high school senior year only.

Next to the full sample results, Table 5 presents the regression kink estimates for the lowest four prior AGI quintiles and the highest prior AGI quintile. While the lower quintiles are roughly similar, the highest quintile is noticeably different. Specifically, in the highest quintile, the mean of the average prior AGI is roughly \$51000, which is more than four times as large as the mean of the average prior AGI for taxpayers in the lowest four quintiles. Furthermore, enrollment is noticeably highest for high school seniors in the highest quintile at just under 30%, compared to roughly 20% for high school seniors

²⁹ Some high school seniors are missing average AGI if they were not claimed as a dependent on a filed tax return in any of the 4 years prior to the high school senior year.

in the lower prior AGI quintiles. This suggests that the highest quintile consists of some high school seniors from transitorily low-income families. The regression kink estimates highlight that, even though the first stage kink in tax refunds is similar across these groups, we only estimate a statistically significant kink in enrollment for the more permanently low-income taxpayers in the lowest four prior AGI quintiles.

Figures 7A and B present graphical evidence on the reduced forms for these two groups. While a noticeable kink can be observed in the lowest four prior AGI quintiles, no kink is noticeable in the highest quintile. Moreover, for the higher permanent income sample, enrollment is higher at every level of earnings relative to the kink, and enrollment does not appear to vary much by earnings relative to the kink. This is consistent with enrollment being relatively uncorrelated with transitory income in the senior year for taxpayers with higher permanent incomes. Overall, we conclude that cash-on-hand effects appear to be significant for high school seniors from more permanently low-income families, and these effects may not be as significant for seniors from families with higher permanent incomes.

3.7 Effects at Higher Income Levels: EITC Kink 2 and Kink 3

We also investigate RKDs at EITC Kinks 2 and 3. Appendix Table 2 lists the income values for EITC Kink 2 and Kink 3. As mentioned in section 2.1, the income measure used to determine EITC income benefits changes around EITC Kink 2. As a result, the definition of the running variable changes around EITC Kink 2, and this rules out the RKD at this kink point.

Appendix Figures 2A and 2B illustrate the main issue at EITC Kink 2. For these plots, we define EITC income as the income measure used to determine EITC benefits. Specifically, we follow the rules discussed in Section 2.1 (and in IRS Publication 596 or EIC Worksheet B) so that EITC income is equal to earned income when AGI is less than the phase-out threshold, and EITC income is equal to the maximum of earned income and AGI if AGI is above the phase-out threshold. As illustrated in Appendix Figure 2A, we are able to empirically reproduce the EITC benefit schedule. However, even though the EITC benefit schedule can be reproduced with this running variable, Appendix Figure 2B demonstrates that the frequencies change discontinuously at EITC Kink 2. Intuitively, taxpayers can only be on the maximum credit region close to EITC Kink 2 if their AGI is

less than the phase-out threshold and their earned income is less than EITC Kink 2. On the other hand, taxpayers can be on the phase-out portion of the benefit schedule just above EITC Kink 2 if their AGI is above the phase-out earnings threshold and either their AGI or earned income puts them in the phase-out income range. Thus, there is a discontinuous increase in the number of individuals who are in the phase-out income range as opposed to the maximum credit income range. The discontinuous change in the frequencies at EITC Kink 2 violates the RKD identifying assumptions and hence invalidates the RKD at EITC Kink 2.³⁰

We also estimate an RKD at Kink 3 (the point at which benefits are completely phased out), but do not find evidence of an enrollment kink. Appendix Figure 3 illustrates the graphical evidence on the (A) first stage, (B) reduced form, (C) frequencies and (D) covariate prediction, and Table 6 presents the first stage, reduced form and IV regression results. The first stage results show evidence of a slope change in tax refunds at EITC Kink 3: below the kink point, tax refund decline relatively quickly since EITC benefits are being phased out as AGI increases, and above the kink point, tax refunds decline relatively slowly since EITC benefits have declined to 0 and are no longer phased out. The regression results indicate a slope change in refunds of \$0.14 per dollar of income. Turning to the reduced form evidence, there does not appear to be evidence of a kink in enrollment at EITC Kink 3. However, in examining the magnitude of the first stage slope change and the standard error of the reduced form slope change, we note that there is not sufficient statistical power in the EITC Kink 3 setting to rule out effects similar to the EITC Kink 1 setting. We acknowledge that, given the magnitude of the first stage slope change in the EITC Kink 3 setting, the analysis is statistically under-powered to rule out that there are positive cash-on-hand effects that are smaller than the effects estimated at EITC Kink 1. As shown in Appendix Table 2, EITC Kink 3 occurs at higher income levels than EITC Kink 1, (for example, for head-of-household taxpayers with 1 qualifying child, EITC Kink 1 is at roughly \$9,100 in 2011 and EITC Kink 3 is \$36,052;

³⁰ We have also tried implementing an RKD at EITC Kink 2 using the restricted sample of individuals with earned income equal to AGI. However, even within this restricted sample, the frequencies are not smooth at EITC Kink 2. Some taxpayers who would have had AGI equal to earned income and been just above the phase-out earnings threshold report adjustments so that they have AGI below the phase-out earnings threshold and hence they qualify for maximum EITC benefits. Thus, within the restricted sample, the frequencies illustrate missing mass just above EITC Kink 2.

for head-of-household taxpayers with 2 qualifying children, EITC Kink 1 is at \$12,780 and EITC Kink 3 is at \$40,964, so it may be reasonable to expect that cash-on-hand effects may be smaller for households at higher income levels, and the EITC Kink 3 results are consistent with this intuition.

4. Discussion

There are multiple reasons why tax refunds in the spring of the high school senior year may have a meaningful effect on college enrollment. First, tax refunds for many families in the analysis samples are large enough to have a substantive effect on college enrollment, and they arrive in a lump sum at a time when many youths make their enrollment decisions. Based on data from the National Postsecondary Student Aid Study (NPSAS), Table 7 shows the distribution of annual college tuition and fees net of grant aid by institution type and year. Based on this distribution of costs and the summary statistics on tax refunds in Table 1, tax refunds could cover a significant portion of overall college costs or out-of-pocket college costs, particularly at 2-year and 4-year public colleges where students may not have room-and-board costs if they are living at home.

Another reason that tax refunds in the spring of the high school senior year could be economically significant is because of credit constraints or informational constraints.³¹ Even though financial aid may be widely available, we present evidence of information asymmetry and incomplete take-up of financial aid which suggests that existing programs may not effectively alleviate credit and information constraints for all students. Table 8 presents percentages of enrolled students who do not apply for federal student aid (e.g. Pell grants) using NPSAS data. The evidence highlights that, at the lowest income levels (i.e. below \$40000 of income), roughly 50% to 70% of enrolled students applied for aid. Below \$40000, virtually all of the enrolled students are likely to be eligible for aid. Yet as shown in Table 9, about 60% of low income students who do not apply for aid thought they were ineligible. This percentage does not vary much by income levels even though eligibility for income levels above \$50000 decreases sharply with income.

³¹ For evidence on credit and information constraints in education tax credits, financial aid and college enrollment, see Bettinger, Long, Oreopoulos and Sanbonmatsu (2012), Hoxby and Avery (2012), Turner (2011a,b), Hoxby and Turner (2013a,b), Dynarski and Scott Clayton (2013) Dynarski, Scott-Clayton and Wiederspan (2013) and GAO (2012).

5. Conclusions

We examine the impacts of cash-on-hand on college enrollment decisions. We find evidence of meaningful effects of tax refunds received in the spring of the high school senior year on college enrollment decisions of students from low-income families. Regression results indicate that an additional \$1000 of after-tax income in the spring of the high school senior year increases college enrollment by roughly 0.50 percentage points.

Consistent with the interpretation that tax refunds in the high school senior year relax credit constraints, we find evidence that tax refunds received in the junior year have little to no effects on college enrollment. While many tax benefits and financial aid programs offer aid with complicated forms or aid that arrives after individuals have financed their college costs, our findings are based on income that arrives prior to incurring any enrollment costs when many youths are finalizing their enrollment plans. Providing additional family resources for college through the tax code ensures that take-up is complete among tax filers, which may increase the effectiveness of the transfers. In the context of student aid, it is likely that most marginal youths are from tax filing families,³² so that targeted tax benefits can relax binding credit constraints. Studying the impacts of tax refunds and information on tax benefits on higher education outcomes remains an interesting avenue for future work.

References

- Acemoglu, Daron, and J-S. Pischke. 2001. "Changes in the wage structure, family income, and children's education." *European Economic Review* 45.4: 890-904.
- Bailey, M. J., & Dynarski, S. M. (2011). Gains and gaps: Changing inequality in US college entry and completion (No. w17633). National Bureau of Economic Research.
- Belley, P., & Lochner, L. (2007). The changing role of family income and ability in determining educational achievement (No. w13527). National Bureau of Economic Research.

³² Chetty et al. (2014) show since the mid-1980s roughly 97% of a given birth cohort appears in the tax data during their teenage years.

Bettinger, Eric B., Bridget T. Long, Philip Oreopoulos, and Lisa Sanbonmatsu. 2012. "The role of application assistance and information in college decisions: Results from the HR Block FAFSA experiment." *Quarterly Journal of Economics* 127 (3): 1205–1242.

Bulman, G. B., & Hoxby, C. M. (2015a). The Returns to the Federal Tax Credits for Higher Education (No. w20833). National Bureau of Economic Research.

Bulman, G. B., & Hoxby, C. M. (2015b). The Effects of the Tax Deduction for Postsecondary Tuition: Implications for Structuring Tax-Based Aid (No. w21554). National Bureau of Economic Research.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, 82(6), 2295-2326.

Cameron, Stephen V., and Christopher Taber. 2004. "Estimation of educational borrowing constraints using returns to schooling." *Journal of Political Economy* 112.1: 132-182.

Card, D. (1999). The causal effect of education on earnings. *Handbook of labor economics*, 3, 1801-1863.

Card, David, David Lee, Zhuan Pei, and Andrea Weber 2015. Inference on Causal Effects in a Generalized Regression Kink Design. Working paper, University of Mannheim.

Card, D., Johnston, A., Leung, P., Mas, A., & Pei, Z. (2015). The Effect of Unemployment Benefits on the Duration of Unemployment Insurance Receipt: New Evidence from a Regression Kink Design in Missouri, 2003-2013 (No. w20869). National Bureau of Economic Research.

Carneiro, P., & Heckman, J. J. (2002). The Evidence on Credit Constraints in Post-Secondary Schooling. *The Economic Journal*, 112(482), 705-734.

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics*, 129(4), 1553-1623.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *The Quarterly Journal of Economics*, 126(4), 1593-1660.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2013a. *Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates*. No. w19423. National Bureau of Economic Research.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2013b. *Measuring the Impact of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood*. No. w19424. National Bureau of Economic Research.

Chetty, Raj, John N. Friedman, and Emmanuel Saez. 2012. "Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings." NBER working paper 18232.

Chetty, Raj, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Michael Stepner, Nicholas Turner, Augustin Bergeron, David Cutler (2016). "The Relationship between Life Expectancy and Income in the United States, 2001-2014." forthcoming, *Journal of the American Medical Association*.

Currie, J., & Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics*, 118(4), 1495-1532.

Dahl, Gordon B. and Lance Lochner. 2012. "The impact of family income on child achievement: Evidence from the Earned Income Tax Credit." *American Economic Review* 102 (5):1927–1956.

Dynarski, S. M. (2003). Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion. *The American Economic Review*, 93(1), 279-288.

Dynarski, Susan M. and Judith Scott-Clayton 2013. *Financial Aid Policy: Lessons from Research*, *The Future of Children*, Vol. 23, No. 1.

Dynarski, Susan, Judith Scott-Clayton, and Mark Wiederspan. "Simplifying Tax Incentives and Aid for College: Progress and Prospects." *Tax Policy and the Economy, Volume 27*. University of Chicago Press, 2013.

Ellwood, D., & Kane, T. J. (2000). Who is getting a college education? Family background and the growing gaps in enrollment. *Securing the future: Investing in children from birth to college*, 283-324.

Ganong, P., & Jäger, S. (2014). A permutation test and estimation alternatives for the regression kink design.

Gelber, Alexander, Timothy Moore, and Alexander Strand (2016). The Effect of Disability Insurance Payments on Beneficiaries' Earnings. No. w21851. National Bureau of Economic Research.

Government Accountability Office, 2012. "Improved Tax Information Could Help Families Pay for College." *HIGHER EDUCATION*, GAO-12-560, Report to Committee on Finance, United States. Senate.

- Goldin, C. D., & Katz, L. F. (2009). *The race between education and technology*. Harvard University Press.
- Gurgand, M., Lorenceau, A., & Mélonio, T. (2011). Student loans: Liquidity constraint and higher education in South Africa. Agence Française de Développement Working Paper, (117).
- Hansen, B., T. Nguyen and G. Waddell (2015). Benefit Generosity and Injury Duration: Quasi-Experimental Evidence from Regression Kinks, University of Oregon working paper.
- Hilger, Nathaniel G. 2013. "How Does Family Income Affect College Enrollment? Evidence from Timing of Parental Layoffs." Working paper.
- Hoxby, C. M., & Avery, C. (2012). The Missing "One-Offs": The Hidden Supply of High-Achieving, Low Income Students (No. w18586). National Bureau of Economic Research.
- Hoxby, C., & Turner, S. (2013a). Expanding College Opportunities for High-Achieving, Low Income Students (No. 12-014).
- Hoxby, C. M., & Turner, S. (2013b). Informing Students about Their College Options: A Proposal for Broadening the Expanding College Opportunities Project. Discussion Paper 2013-03, Hamilton Project, Brookings Institution, Washington, DC.
- Hoynes, Hilary, Doug Miller, David Simon (2015). Income, the Earned Income Tax Credit, and Infant Health, *American Economic Journal: Economic Policy* 7(1): 172–211.
- Keane, M. P., & Wolpin, K. I. (2001). The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review*, 1051-1103.
- LaLumia, S. (2012). Tax preferences for higher education and adult college enrollment. *National Tax Journal*, 65(1), 59-90.
- Landais, C. (2015). Assessing the welfare effects of unemployment benefits using the regression kink design. *American Economic Journal: Economic Policy*, 7(4), 243-278.
- Lochner, L. (2004). EDUCATION, WORK, AND CRIME: A HUMAN CAPITAL APPROACH. *International Economic Review*, 45(3), 811-843.
- Lochner, L. J., & Monge-Naranjo, A. (2011). The Nature of Credit Constraints and Human Capital. *The American Economic Review*, 101(6), 2487-2529.
- Long, Bridget. 2004. College Choices: The Economics of Where to Go, When to Go, and How to Pay For It, chap. The Impact of Federal Tax Credits for Higher Education Expenses. University of Chicago Press, 101–168.

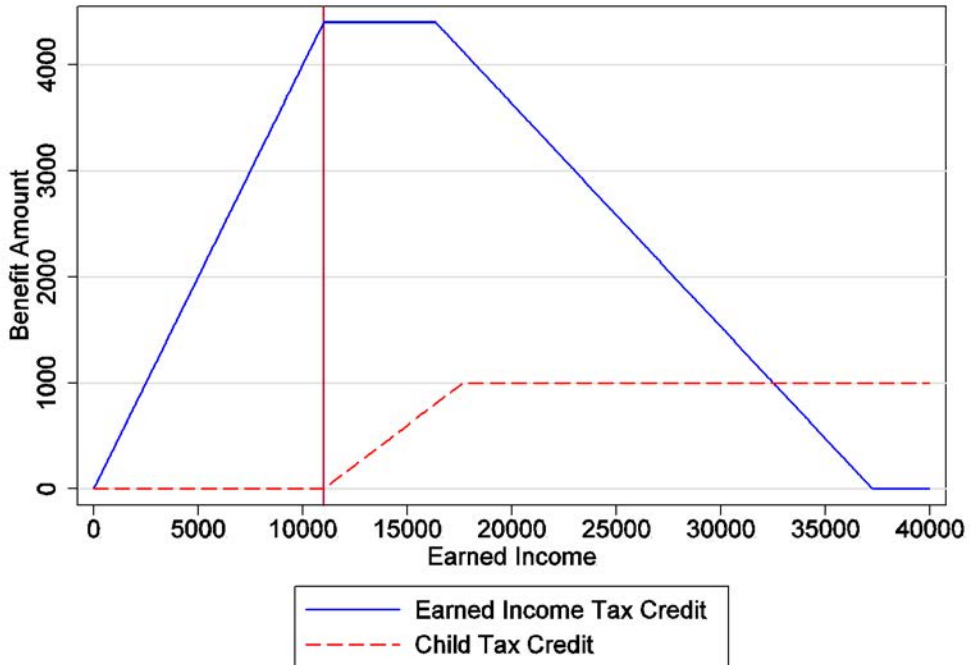
- Lovenheim, Michael F. 2011. "The Effect of Liquid Housing Wealth on College Enrollment." *Journal of Labor Economics* 29 (4):741–771.
- Lovenheim, M. F., & Reynolds, C. L. (2013). The effect of housing wealth on college choice: Evidence from the housing boom. *Journal of Human Resources*, 48(1), 1-35.
- Marx, B. M., & Turner, L. J. (2015). Borrowing Trouble? Student Loans, the Cost of Borrowing, and Implications for the Effectiveness of Need-Based Grant Aid (No. w20850). National Bureau of Economic Research.
- Milligan, K., Moretti, E., & Oreopoulos, P. (2004). Does education improve citizenship? Evidence from the United States and the United Kingdom. *Journal of Public Economics*, 88(9), 1667-1695.
- Micheltore, K. (2013). The effect of income on educational attainment: Evidence from state earned income tax credit expansions. Available at SSRN 2356444.
- Moretti, E. (2004). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, 121(1), 175-212.
- National Center for Education Statistics 2013. *Digest of Education Statistics*. http://nces.ed.gov/programs/digest/2013menu_tables.asp, accessed 12/3/2013.
- Nielsen, Helena Skyt, Torben Sørensen, and Christopher R. Taber 2010. "Estimating the Effect of Student Aid on College Enrollment: Evidence from a Government Grant Policy Reform," *American Economic Journal: Economic Policy*, 2 (2), 185–215.
- Plug, E., & Vijverberg, W. (2005). Does Family Income Matter for Schooling Outcomes? Using Adoptees as a Natural Experiment. *The Economic Journal*, 115(506), 879-906.
- Saez, Emmanuel 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy*, 2 (3), 180–212.
- Shea, J. (2000). Does parents' money matter?. *Journal of public Economics*, 77(2), 155-184.
- Solis, Alex. (2012). "Credit access and college enrollment" Working paper.
- Stinebrickner, R., & Stinebrickner, T. (2008). The Effect of Credit Constraints on the College Drop-Out Decision: A Direct Approach Using a New Panel Study. *The American Economic Review*, 98(5), 2163-84.
- Turner, Nicholas 2011a. "The effect of tax-based federal student aid on college enrollment." *National Tax Journal*, 64(3), 839-862.

Turner, Nicholas 2011b. "Why Don't Taxpayers Maximize their Tax-Based Student Aid? Salience and Inertia in Program Selection." *The BE Journal of Economic Analysis & Policy* 11.1.

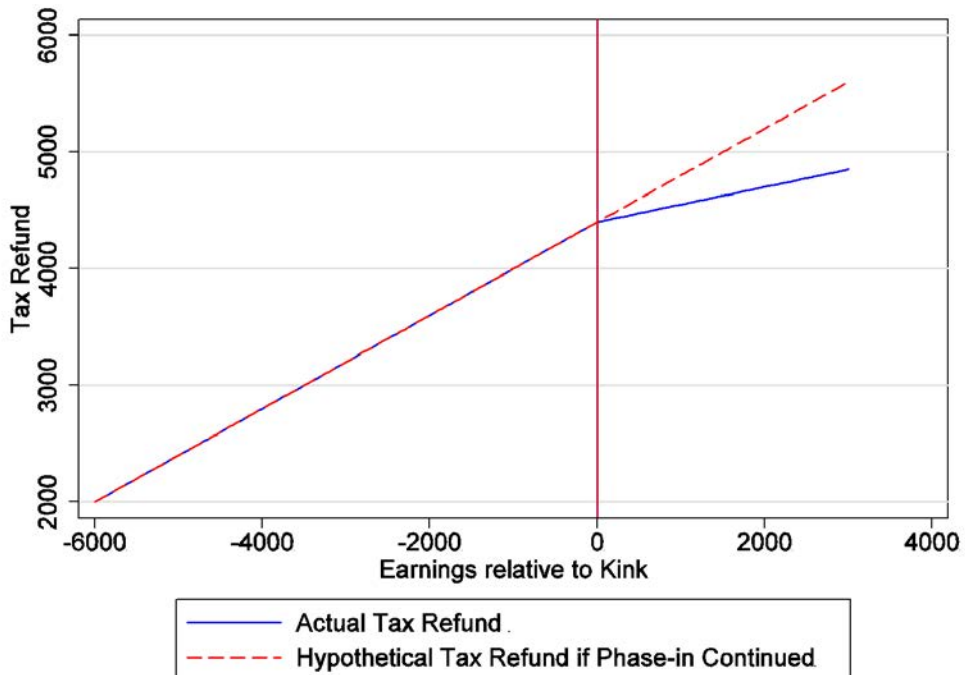
van der Klaauw, Wilbert. "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression–Discontinuity Approach*." *International Economic Review* 43.4 (2002): 1249-1287.

Figure 1. Institutional Background

A. EITC and CTC Benefit Schedule



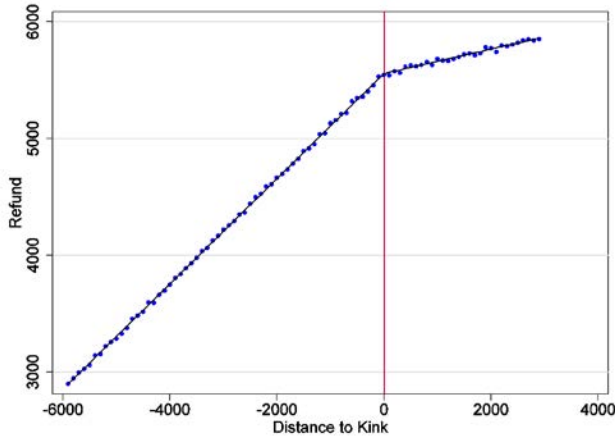
B. Tax Refunds at EITC Kink 1



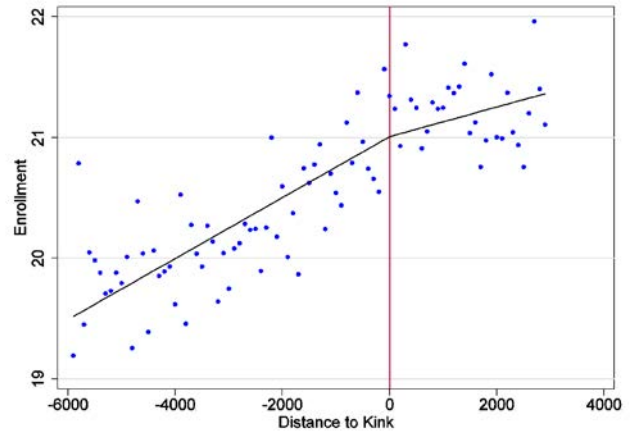
Notes: These figures plot simulated federal tax amounts that are calculated using the NBER TAXSIM calculator. Panel A shows the EITC and CTC schedules for the case of a taxpayer with two children (one kid of age 14 and another of 17) with married filing jointly status for the year 2005. The solid line represent the EITC benefit schedule and dashed line represents the CTC benefit schedule. Using the same case as Panel A, Panel B plots tax refunds as a function of earnings relative to the EITC kink point. For the tax calculations, earned income is from W2 wages only, and earned income is equal to AGI.

Figure 2. Regression Kink at EITC Kink 1

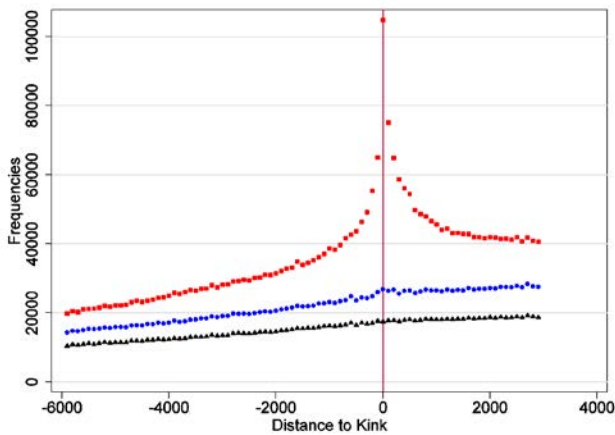
A. First Stage



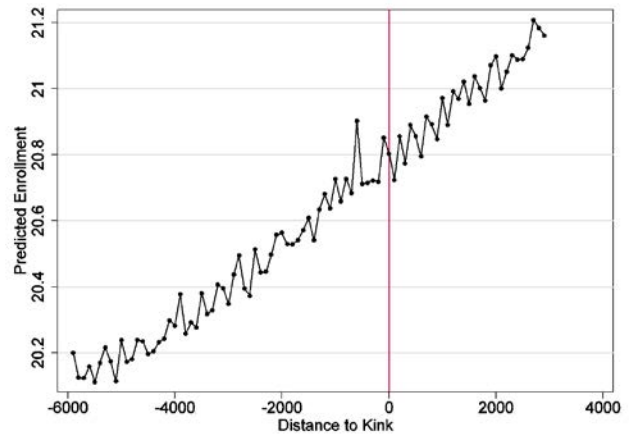
B. Reduced Form



C. Frequencies

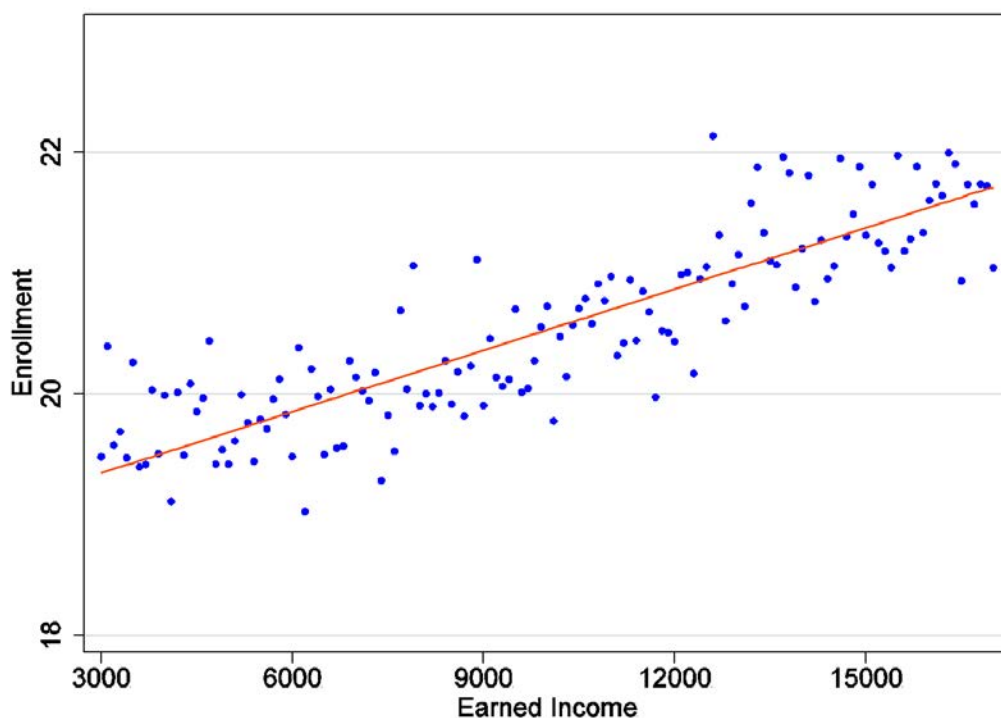


D. Covariate Prediction



Notes: For Panels A and B, the circles show the mean tax refund and enrollment rate within each \$100 bin of earnings relative to the tax kink points, respectively. The solid lines show fitted values within each \$100 bin of earnings relative to the kink points. Fitted values are obtained from regressions using the individual-level data in which tax refund or an enrollment indicator is regressed on a linear control for earnings relative to the kink point and a dummy for earnings less than the kink point interacted with the linear control. \$100 bins are assigned based on rounding earnings relative to the kink point to the nearest \$100 amount. Panel C plots the number of tax returns within \$100 bins around each tax kink point. The red squares are frequencies including the self-employed; the blue triangles are frequencies when excluding the self-employed and the black circles are frequencies when excluding individuals with a difference between W2 wages and wages reported on the 1040 form of more than \$1000. This difference is attributable to non-third party verified wages. For Panel D, the circles show mean predicted enrollment rates within each \$100 bin of earnings relative to the tax kink points. Predicted enrollment is computed by obtaining fitted values from a regression of an enrollment indicator on dummies for calendar year, filing status, and number of dependents and a linear control for earned income in the senior year.

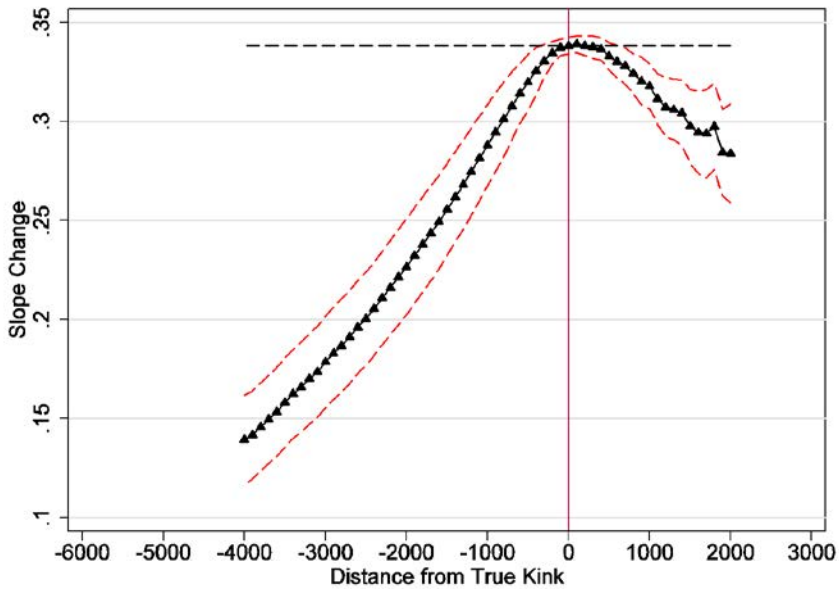
Figure 3. Enrollment by Earned Income



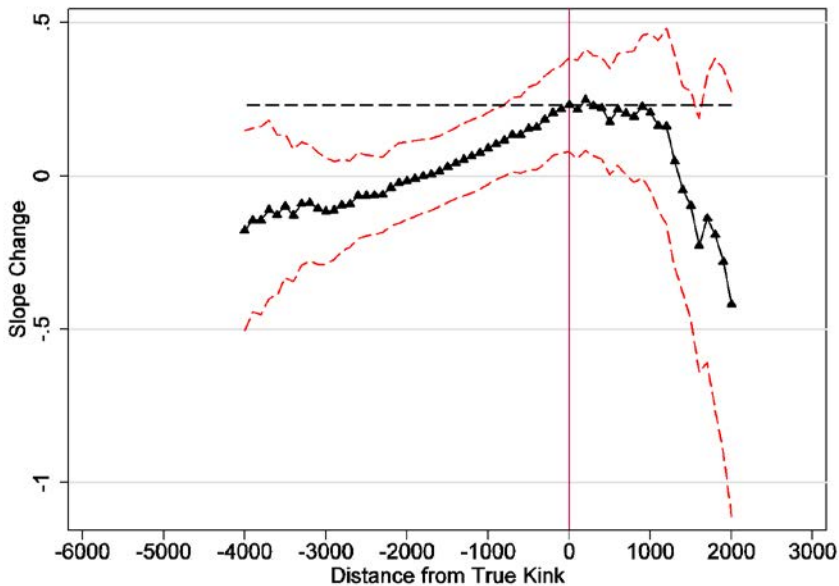
Notes: This plot is based on the full sample of high school seniors from 2001 through 2011 with CPI-adjusted income in the range shown, not just the EITC Kink 1 sample. Enrollment is measured in the year after the high school senior year, and income is based on earned income in the high school senior year. Similar to the EITC Kink 1 sample, individuals with any self-employment income are excluded from the sample. The circles represent mean enrollment within each \$100 bins of CPI-adjusted earned income. The solid line plots mean fitted values within each bin. The fitted values are obtained from regressing an enrollment indicator on a linear control for earned income. The estimated intercept and slope with respect to earned income are 18.835 and 0.169, respectively, where enrollment is scaled to be either 0 or 100% and the slope with respect to earned income is scaled to be in \$1000 of earned income (standard errors are in parentheses).

Figure 4. Placebo Analysis

A. Change in Tax Refund



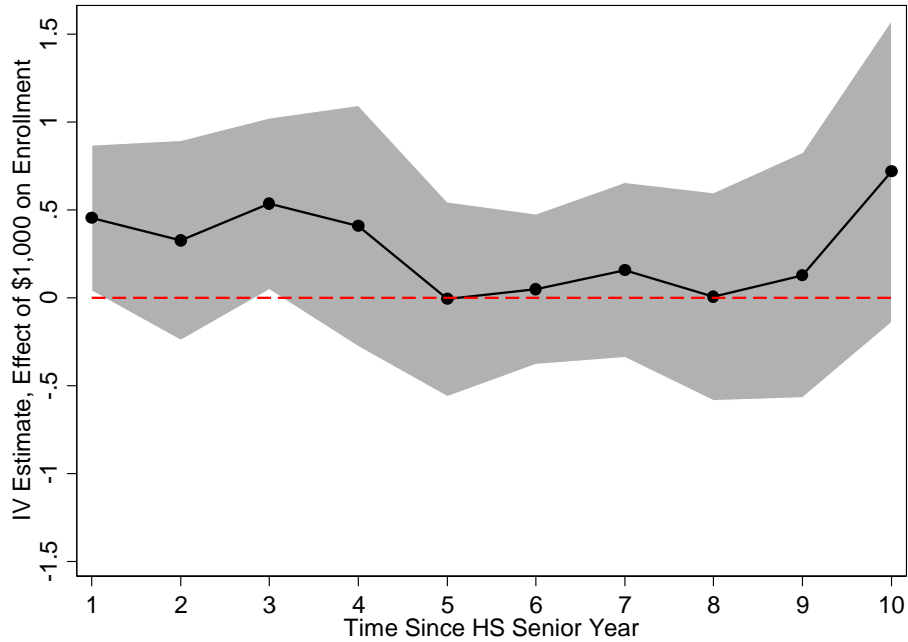
B. Change in Enrollment



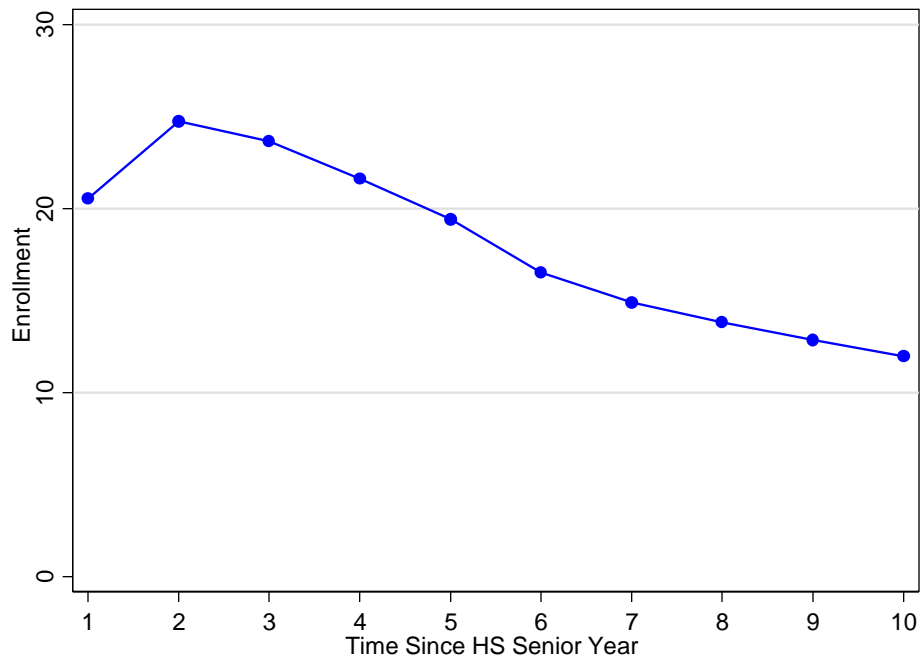
Notes: These figures plot the estimated slope changes when varying the placebo kink points in \$100 increments around the true EITC kink points. In particular, placebo kink = kink1+p, where $p \in \{-4000, -3900, \dots, +2000\}$. For each figure, the triangles show the slope change in tax refund and enrollment rate relative to the placebo kink points. The horizontal dotted lines show the estimated slope changes at the true kink point. Fitted slope changes are obtained from regressions using the individual-level data in which tax refund or an enrollment indicator is regressed on a linear control for earnings relative to the placebo kink points, a dummy for earnings less than the kink point and an interaction between the dummy variable and the linear control. Each figure also shows dotted lines for the 95% confidence interval bands around the estimates. Panel A shows the placebo results for the first stage and panel B shows the placebo results for the reduced form estimation.

Figure 5. Persistence of Cash-on-Hand Effects

A. IV Estimates by Time Since HS Senior Year

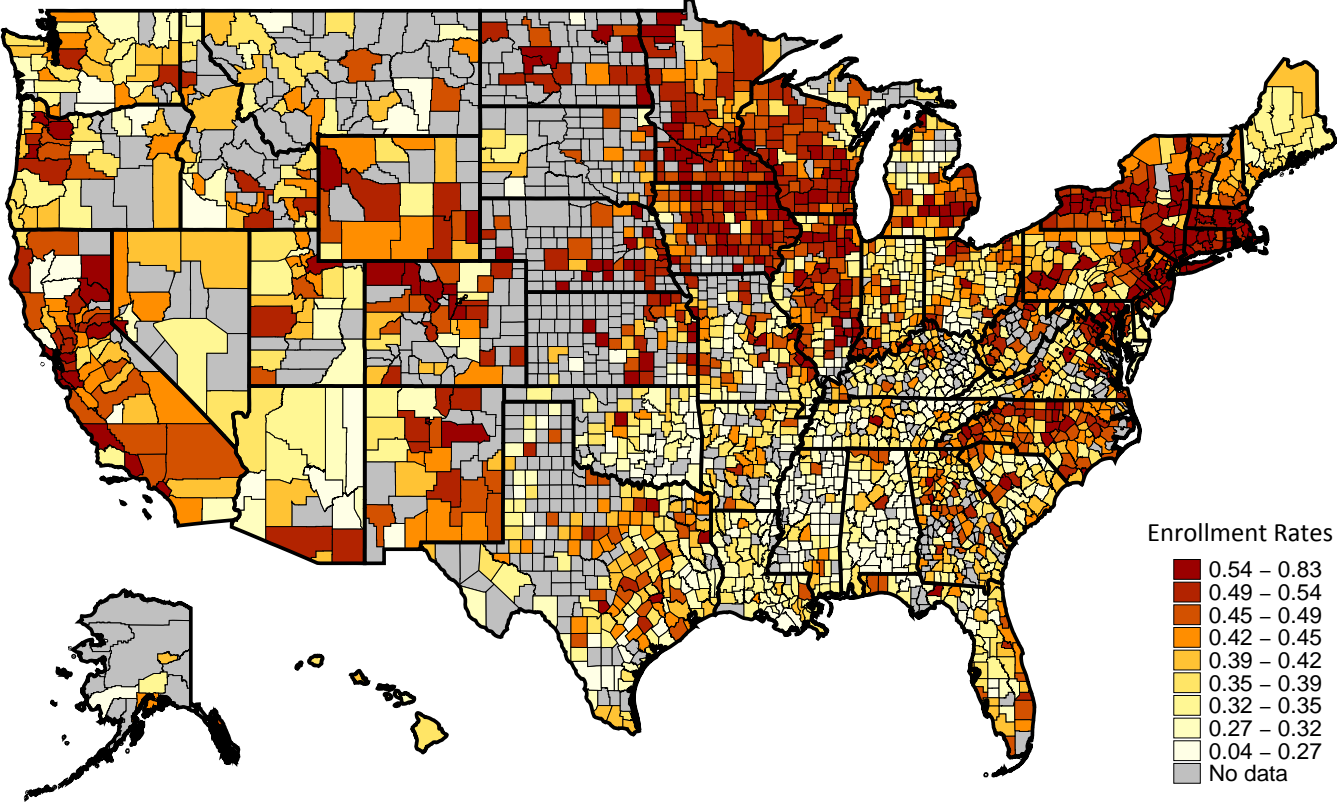


B. Mean Enrollment by Time Since HS Senior Year



Notes: Panel A plots the IV estimates of the impacts of tax refunds in the spring of the high school senior year on enrollment in each year after the high school senior year. Each circle represents a regression coefficient from a separate regression in which the dependent variable is an indicator for enrollment in the specified year after the high school senior year. The grey area plots the 95 percent confidence interval of the estimate. Panel B plots mean enrollment for each year after high school.

Figure 6. County-Level Enrollment



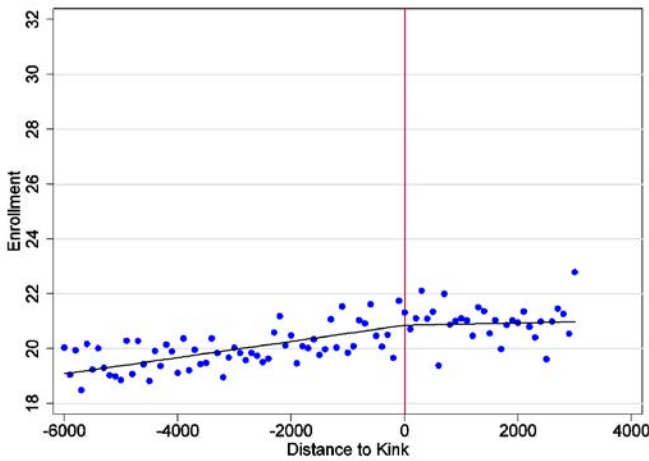
Notes: These figures plot heat maps of county-level enrollment rates for high school seniors in the 2001, 2006 and 2011 cohorts. Enrollment rates are listed as the fraction of high school seniors in the cohort who enroll in the year after the high school senior year. The sample of counties is restricted based on the sample size within the county. Specifically, we compute the number of high school seniors in each county in each year, and we restrict the mapped counties to those that are in the top 75% of the sample size distribution in 2001 (i.e. those counties that have at least 144 high school seniors in 2001, where 144 is the 25th percentile of the sample size distribution across counties in 2001).

Figure 7. Heterogeneity by Average AGI

Average AGI Computed over 4 Years Prior to HS Senior Year

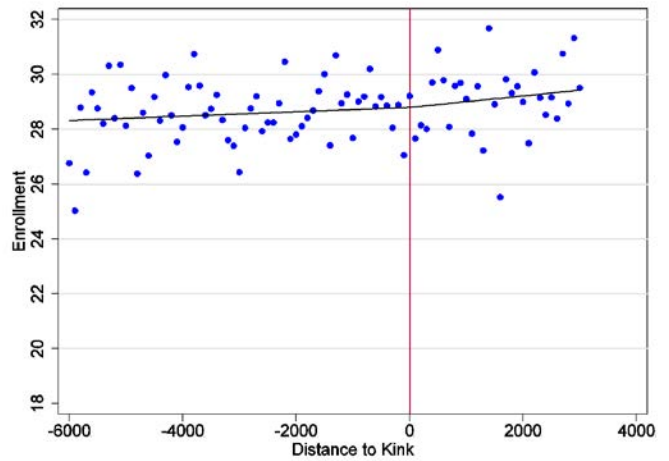
A. Lowest 4 Quintiles

Average Prior AGI=\$16,137



B. Highest Quintile

Average Prior AGI=\$51,200



Notes: The circles represent mean enrollment within each \$100 bins of earnings relative to EITC Kink 1. The solid lines show mean fitted values within each \$100 bin of earnings relative to the kink points. Fitted values are obtained from regressions using the individual-level data in which tax refund or an enrollment indicator is regressed on a linear control for earnings relative to the kink point and a dummy for earnings less than the kink point interacted with the linear control. \$100 bins are assigned based on rounding earnings relative to the kink point to the nearest \$100 amount. The sample is restricted to high school senior cohorts in 2005 through 2011 who do not have any self-employment income in any of the 4 years prior to the high school senior year. Quintiles of average AGI are computed within groups based on the high school senior year and the number of qualifying children. High school seniors who were not claimed as dependents in a year prior to the high school senior year have missing prior AGI, and these observations are included in the lowest 4 quintiles sample.

Table 1: Summary Statistics

	EITC Kink 1 Sample	
	Mean	SD
N	1,375,626	
Enroll (1 year after HS)	20.67	40.49
Married Filing Jointly	0.18	0.39
Head of Household	0.82	0.39
Child Dependents	1.69	0.88
After-Tax Income	16,194.59	40,634.06
Balance Due	-4,878.22	1,828.19
Pre-Tax Income	11,316.37	40,457.78
Has Refund	1.00	0.01

Notes: Dollar values are CPI adjusted to 2015 dollars. Pre-tax income is total income, which is measured on line 22 of Form 1040. This total income measure is the sum of all income listed on Form 1040. Balance due captures the net amount due; it is negative if a net refund to the taxpayer is due. After-tax income is the sum of pre-tax income and the balance due.

Table 2: EITC Kink 1 RKD Estimates

	Full Sample			Full Sample		
	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment	First Stage Dep Var = Refund	Reduced Form Dep Var = Full Time Enrollment	IV Dep Var = Full Time Enrollment
Earnings Relative to Kink (kinkdist)	0.442 [0.000722]	0.234 [0.0246]	0.0345 [0.0737]	0.442 [0.000722]	0.216 [0.0250]	0.0297 [0.0682]
Slope Change at Kink (D*kinkdist)	-0.338 [0.00199]	-0.153 [0.0705]		-0.338 [0.00199]	-0.142 [0.0670]	
Effect of \$1000 on Enrollment (IV)			0.451 [0.208]			0.420 [0.197]
N	1375626	1375626	1375626	1375626	1375626	1375626
	1 Child			2+ Children		
	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment
Earnings Relative to Kink (kinkdist)	0.406 [0.00115]	0.195 [0.0423]	0.0509 [0.0853]	0.461 [0.000775]	0.248 [0.0305]	0.0611 [0.121]
Slope Change at Kink (D*kinkdist)	-0.352 [0.00305]	-0.125 [0.104]		-0.33 [0.00219]	-0.134 [0.102]	
Effect of \$1000 on Enrollment (IV)			0.356 [0.295]			0.405 [0.309]
N	428587	428587	428587	820891	820891	820891

Notes: Each regression includes dummy variables for senior year, number of children and filing status. Standard errors are clustered based on \$100 bins of earnings relative to the kink point.

Table 3: EITC Kink 1, RKD Estimates for Junior and Senior Samples

	Junior Year K1 Sample			
	First Stage Dep Var = Junior Refund	Dep Var = Senior Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment
Earnings Relative to Kink (kinkdist)	0.451 [0.000760]	0.147 [0.00251]	0.192 [0.0300]	0.207 [0.0924]
Slope Change at Kink (D*kinkdist)	-0.320 [0.00217]	-0.0285 [0.00508]	0.0110 [0.0833]	
Effect of \$1000 on Enrollment (IV)				-0.0344 [0.259]
N	1180719	1180719	1180719	1180719

	Senior Year K1 Sample			
	Dep Var = Junior Refund	First Stage Dep Var = Senior Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment
Earnings Relative to Kink (kinkdist)	0.127 [0.00201]	0.443 [0.000834]	0.213 [0.0274]	0.0191 [0.0797]
Slope Change at Kink (D*kinkdist)	-0.00656 [0.00502]	-0.341 [0.00228]	-0.149 [0.0775]	
Effect of \$1000 on Enrollment (IV)				0.438 [0.227]
N	1189802	1189802	1189802	1189802

Notes: Each regression includes dummy variables for senior year, number of children and filing status. Standard errors are clustered based on \$100 bins of earnings relative to the kink point.

Table 4: Geographic Heterogeneity
 Dependent Variable = Enroll

	County-Enrollment Quintiles				
	(1)	(2)	(3)	(4)	(5)
Average County Enrollment Rate	27.93	37.70	44.04	50.42	60.39
<u>EITC K1 Sample, Regression Kink Estimates</u>					
Mean Enrollment (Percentage)	12.71	16.82	20.94	23.53	29.49
Slope Change in Tax at Kink (First Stage)	-0.336 [0.00426]	-0.344 [0.00357]	-0.344 [0.00352]	-0.336 [0.00397]	-0.33 [0.00534]
Slope Change in Enroll at Kink (Reduced Form)	-0.216 [0.113]	-0.0907 [0.136]	-0.236 [0.108]	-0.0449 [0.152]	-0.327 [0.190]
Effect of \$1000 on Enroll (IV)	0.642 [0.336]	0.264 [0.392]	0.686 [0.314]	0.134 [0.451]	0.991 [0.575]
N	221339	278094	337802	284937	222725

Notes: Within each year, county enrollment quintiles are computed by categorizing counties into the top 20% of enrollment rates, the next 20%, etc. down to the lowest 20% of enrollment rates). The first row presents the average county-level enrollment rates for counties in each quintile. The regression kink slope changes are then estimated using the EITC Kink 1 high school seniors who reside in these quintiles.

Table 5: Heterogeneity by Average AGI
Sample: HS Senior Year Cohorts 2005-2011, No Self-Employment Income in 4 Years Prior to HS Senior Year

	Full Sample	Avg AGI Quintile	
		Lowest 4 Quintiles	Highest Quintile
Average AGI over 4 Prior Years		16137.45	51200.19
Average AGI in Senior Year		11458.19	11549.13
<hr/>			
EITC K1 Sample, Regression Kink Estimates			
Mean Enroll (Percent)	21.78	20.40	28.79
Slope Change in Tax at Kink (First Stage)	-0.353 [0.00274]	-0.353 [0.00287]	-0.356 [0.00711]
Slope Change in Enroll at Kink (Reduced Form)	-0.223 [0.104]	-0.288 [0.123]	0.0532 [0.214]
Effect of \$1000 on Enroll (IV)	0.630 [0.295]	0.814 [0.348]	-0.149 [0.596]
N	723131	604187	118944

Notes: This table presents results based on average AGI in the 4 years prior to the high school senior year. The sample is restricted to high school senior cohorts in 2005 through 2011 who do not have any self-employment income in any of the 4 years prior to the high school senior year. Quintiles of average AGI are computed within groups based on the high school senior year and the number of qualifying children. High school seniors who were not claimed as dependents in a year prior to the high school senior year have missing prior AGI, and these observations are included in the lowest 4 quintiles sample.

Table 6: EITC Kink 3 RKD Estimates

	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment
Earnings Relative to Kink (kinkdist)	-0.134 [0.001]	0.589 [0.021]	0.563 [0.023]
Slope Change at Kink (D*kinkdist)	0.139 [0.002]	-0.026 [0.040]	
Effect of \$1000 on Enrollment (IV)			-0.188 [0.287]
N	1965835	1965835	1965835

Notes: Each regression includes dummy variables for senior year, number of children and filing status. Standard errors are clustered based on \$100 bins of earnings relative to the kink point.

Table 7: Tuition and Fees minus All Grants, 2000-2012

Type of Institution	Year	\$0	\$1 - 500	\$501 - 1,500	\$1,501 - 3,000	\$3,001 - 5,000	\$5,001 or more
Public 4-year	1999-00	24.4	4.5	8.6	22.8	25.3	14.4
	2003-04	26.0	3.6	6.3	13.6	24.5	26.0
	2007-08	28.3	2.8	5.5	7.6	16.0	39.8
	2011-12	30.5	2.5	4.8	7.0	9.1	46.1
Private not-for-profit 4-year	1999-00	12.0	2.1	3.8	7.2	10.9	64.0
	2003-04	9.8	1.6	3.7	5.9	8.2	70.9
	2007-08	9.9	1.6	2.0	3.1	6.0	77.3
	2011-12	14.5	1.0	2.4	2.6	6.0	73.4
Public 2-year	1999-00	29.9	16.8	25.3	22.8	3.5	1.746 !
	2003-04	31.4	8.6	22.5	25.8	9.2	2.5
	2007-08	36.1	5.5	18.2	24.8	13.6	1.8
	2011-12	48.7	4.9	11.9	15.3	15.4	3.7
Private for-profit	1999-00	3.977 !	3.950 !	2.819 !	10.6	14.1	64.5
	2003-04	5.3	0.635 !!	2.932 !	3.6	13.1	74.4
	2007-08	3.402 !	1.646 !	2.8	4.5	7.7	79.9
	2011-12	2.4	0.403 !!	1.3	2.630 !	3.4	89.8

Notes: The above table is created using NPSAS data and the PowerStats tool by NCES. The NPSAS data sample size was 95,000 for 2011-12; 113,500 for 2007-08; 79,900 for 2003-04; and 50,000 for 1999-2000. The subsample used in this table includes individuals between ages 18-20, who are full-time/full-year students. The rows show different type of institutions and the columns show different cost intervals for tuition minus all grants.

! Interpret data with caution. Estimate is unstable because the standard error represents more than 30 percent of the estimate.

!! Interpret data with caution. Estimate is unstable because the standard error represents more than 50 percent of the estimate.

Source: U.S. Department of Education, National Center for Education Statistics, National Postsecondary Student Aid Study (NPSAS). PowerStats Tool available at <http://nces.ed.gov/datalab/powerstats/default.aspx>

Table 8: Applications for Federal Aid by Income

Percentage of enrolled students who applied for federal aid			
Total Income	2000	2004	2008
All Income Groups	49	58	58
0 to 10,000	72	71	73
10,001 to 20,000	66	73	74
20,001 to 30,000	53	71	70
30,001 to 40,000	44	60	62
40,001 to 50,000	40	56	58
50,001 to 75,000	41	48	49
75,001 to 100,000	36	46	49
100,001 to 125,000	35	41	43
125,001 to 150,000	32	38	40
Greater than 150,001	30	38	42

Notes: Data is from U.S. Department of Education, National Center for Education Statistics, 1999-2000, 2003-04, 2007-2008 National Postsecondary Student Aid Study.

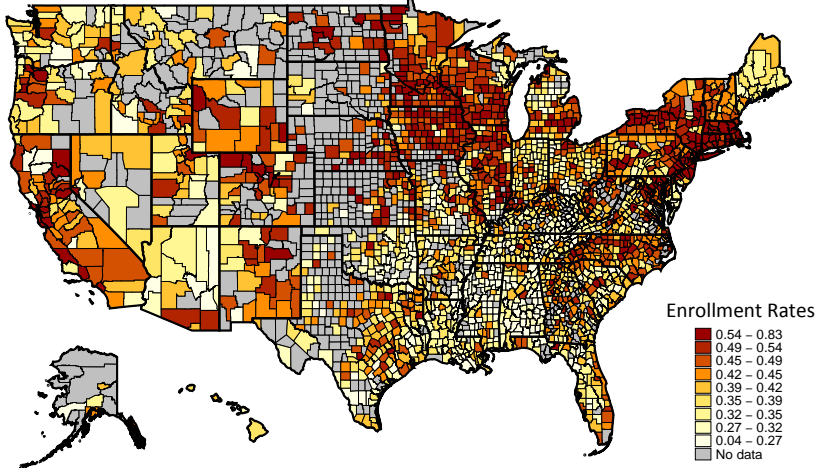
Table 9: Non-Application by Income

Total Income	Reason for not applying for federal student aid				
	Did not want to take on debt	Forms were too much work	No information on how to apply	No Need	Thought Ineligible
All Income Groups	40	19	23	51	61
0 to 10,000	41	22	28	41	58
10,001 to 20,000	42	21	29	46	59
20,001 to 30,000	45	19	27	37	62
30,001 to 40,000	43	21	24	41	60
40,001 to 50,000	43	21	28	41	60
50,001 to 75,000	42	20	24	48	62
75,001 to 100,000	39	18	23	55	61
100,001 to 125,000	37	17	16	64	60
125,001 to 150,000	35	12	12	68	63
Greater than 150,001	33	15	15	70	63

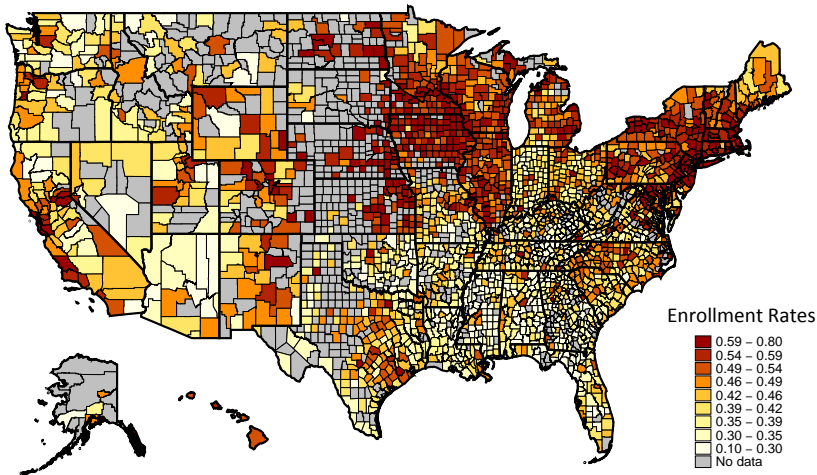
Source: Data is from U.S. Department of Education, National Center for Education Statistics, 2007-2008 National Postsecondary Student Aid Study.

Appendix Figure 1. County-Level Enrollment

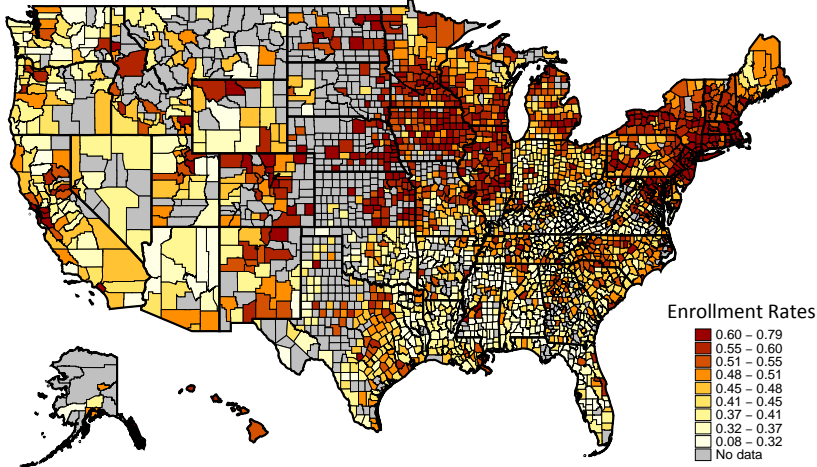
A. 2001 Senior Cohort



B. 2006 Senior Cohort



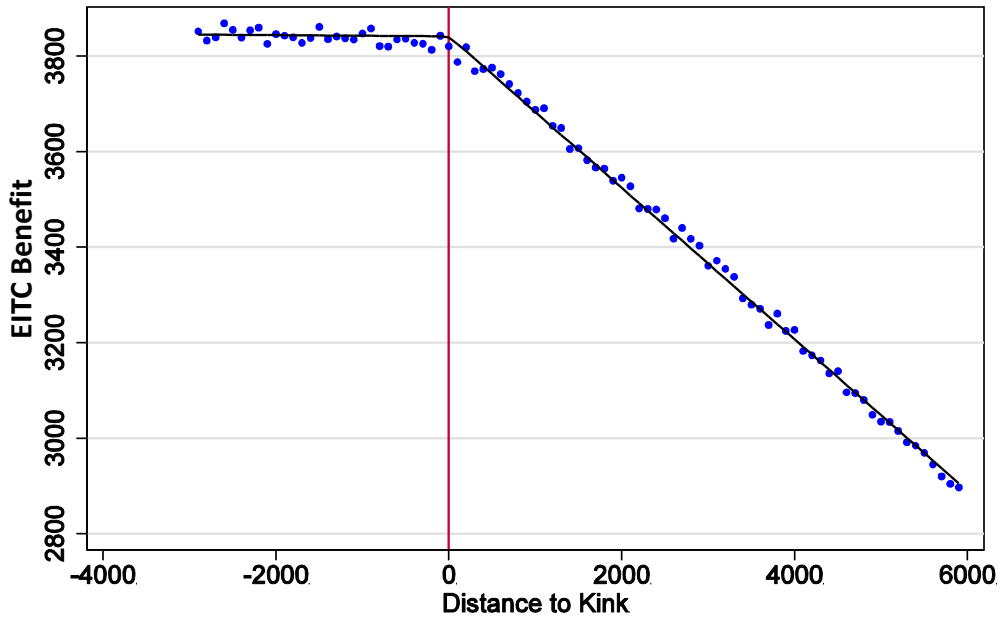
C. 2011 Senior Cohort



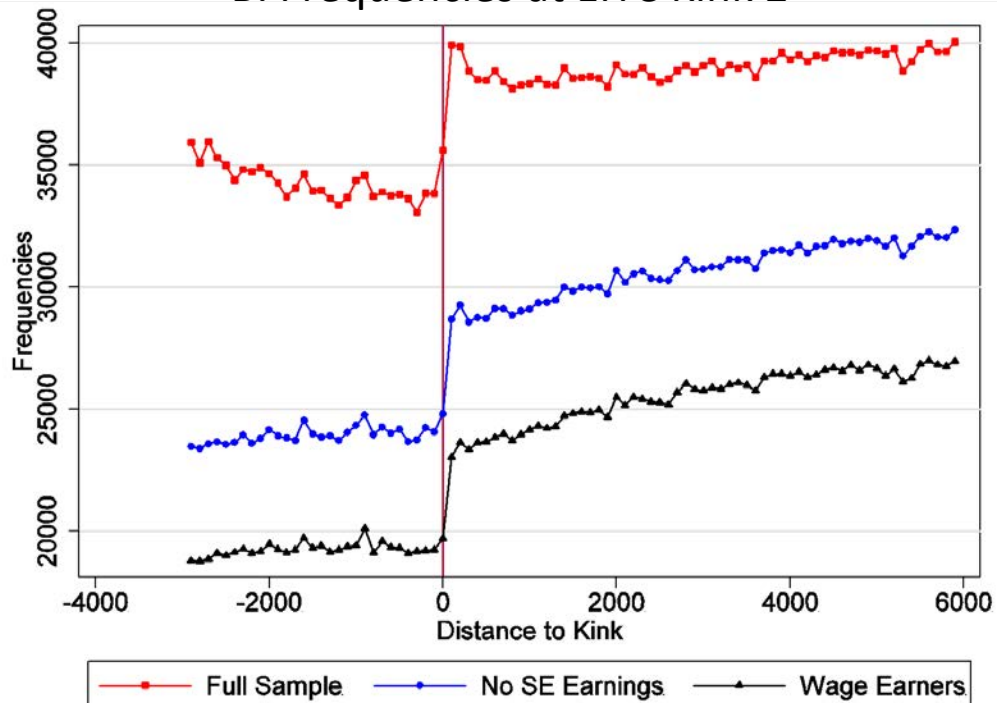
Notes: These figures plot heat maps of county-level enrollment rates for high school seniors in the 2001, 2006 and 2011 cohorts. Enrollment rates are computed as the fraction of high school seniors in the cohort who enroll in the year after the high school senior year. The sample of counties is restricted based on the sample size within the county. Specifically, we compute the number of high school seniors in each county in each year, and we restrict the mapped counties to those that are in the top 75% of the sample size distribution in 2001 (i.e. those counties that have at least 144 high school seniors in 2001, where 144 is the 25th percentile of the sample size distribution across counties in 2001).

Appendix Figure 2. Regression Kink at EITC Kink 2

A. First Stage, Slope Change in EITC Benefits at EITC Kink 2



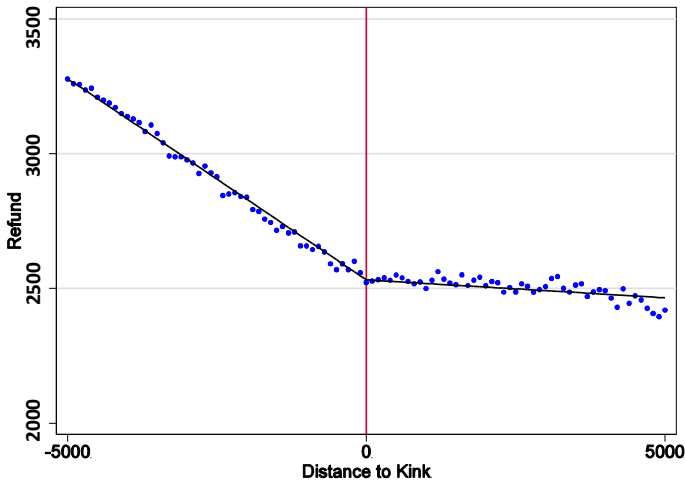
B. Frequencies at EITC Kink 2



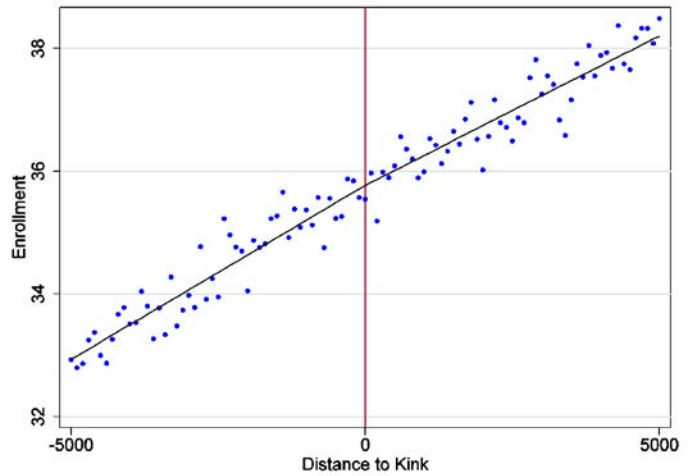
Notes: In Panels A and B, Distance to Kink measures EITC Income relative to EITC Kink 2. EITC income is defined as the income measure used to determine EITC benefits. Following official rules described in IRS Publication 596, EITC Income changes around EITC Kink 2. Specifically, EITC income is equal to earned income when AGI is less than the phase-out threshold, and EITC income is equal to the maximum of earned income and AGI if AGI is above the phase-out threshold. For Panel A, the circles show the mean tax refund within each \$100 bin of earnings relative to EITC Kink 2. The solid line shows fitted values within each \$100 bin of earnings relative to the kink 2. Panel B plots the number of tax returns within \$100 bins around each tax kink 2 point. The red squares are frequencies including the self-employed; the blue circles are frequencies when excluding the self-employed and the black triangles are frequencies for wage earners.

Appendix Figure 3. Regression Kink at EITC Kink 3

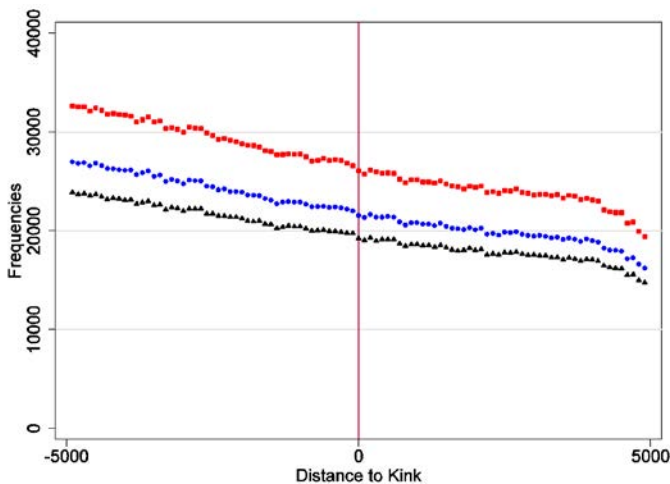
A. First Stage



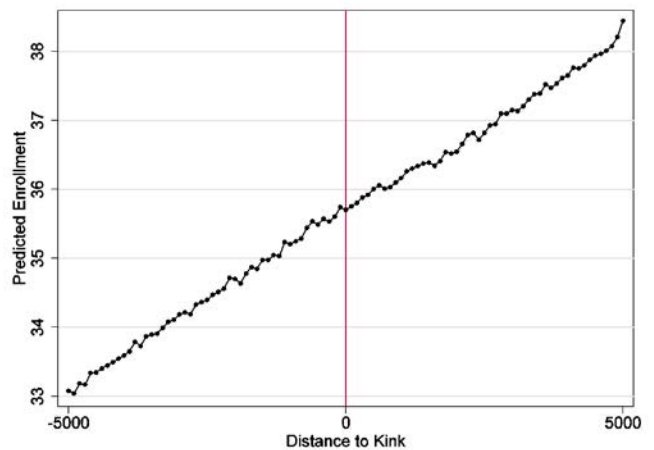
B. Reduced Form



C. Frequencies



D. Covariate Prediction



Notes: All panels plot values by \$100 AGI bins relative to the tax kink, where \$100 bins are assigned based on rounding AGI relative to the kink point to the nearest \$100 amount. For Panels A and B, the circles show mean tax refunds and enrollment rates. The solid lines show fitted values obtained from regressions using the individual-level data in which tax refund or an enrollment indicator is regressed on a linear control for AGI relative to the kink point, a dummy for earnings less than the kink point and an interaction between the dummy variable and the linear control. Panel C plots the number of tax returns in the following ways: the red squares are frequencies including the self-employed; the blue triangles are frequencies excluding the self-employed and the black circles are frequencies when excluding individuals with a difference between W2 wages and wages reported on the 1040 form of more than \$1000. This difference is attributable to non-third party verified wages. For Panel D, the circles show mean predicted enrollment rates. The solid lines show fitted values of predicted enrollment rates computed by regressing an enrollment indicator on dummies for calendar year, filing status, and number of dependents and linear controls for senior year income.

Appendix Table 1: Tax Kink Points

Tax Year	EITC Kink 1		EITC Phase-Out AGI Threshold		Child Tax Credit		
	1 Child	2+ Children	Married Filing Jointly	Head of Household	kink	Rate	max
2001	\$7,140	\$10,020	\$13,100	\$13,100	\$10,000	10	\$600
2002	\$7,370	\$10,350	\$14,550	\$13,550	\$10,350	10	\$600
2003	\$7,490	\$10,510	\$14,750	\$13,750	\$10,500	10	\$1,000
2004	\$7,660	\$10,750	\$15,050	\$14,050	\$10,750	15	\$1,000
2005	\$7,830	\$11,000	\$16,400	\$14,400	\$11,000	15	\$1,000
2006	\$8,080	\$11,340	\$16,850	\$14,850	\$11,300	15	\$1,000
2007	\$8,390	\$11,790	\$17,400	\$15,400	\$11,750	15	\$1,000
2008	\$8,580	\$12,060	\$18,750	\$15,750	\$8,500	15	\$1,000
2009	\$8,950	\$12,570	\$21,450	\$16,450	\$3,000	15	\$1,000
2010	\$8,970	\$12,590	\$21,500	\$16,500	\$3,000	15	\$1,000
2011	\$9,100	\$12,780	\$21,800	\$16,700	\$3,000	15	\$1,000

Notes: All dollar values are in nominal dollars.

Appendix Table 2: EITC Phase-Out Expansion

Panel A: 1 Child				
Year	Head of Household		Married Filing Jointly	
	Kink 2	Kink 3	Kink 2	Kink 3
	Beginning of Phase-Out	Ending of Phase-Out	Beginning of Phase-Out	Ending of Phase-Out
2001	13,090	28,281	13,090	28,281
2002	13,520	29,201	14,520	30,201
2003	13,730	29,666	14,730	30,666
2004	14,040	30,338	15,040	31,338
2005	14,370	31,030	16,370	33,030
2006	14,810	32,001	16,810	34,001
2007	15,390	33,241	17,390	35,241
2008	15,740	33,995	18,740	36,995
2009	16,420	35,463	21,420	40,463
2010	16,450	35,535	21,460	40,545
2011	16,690	36,052	21,770	41,132

Panel B: 2 Children				
Year	Head of Household		Married Filing Jointly	
	Kink 2	Kink 3	Kink 2	Kink 3
	Beginning of Phase-Out	Ending of Phase-Out	Beginning of Phase-Out	Ending of Phase-Out
2001	13,090	32,121	13,090	32,121
2002	13,520	33,178	14,520	34,178
2003	13,730	33,692	14,730	34,692
2004	14,040	34,458	15,040	35,458
2005	14,370	35,263	16,370	37,263
2006	14,810	36,348	16,810	38,348
2007	15,390	37,783	17,390	39,783
2008	15,740	38,646	18,740	41,646
2009	16,420	40,295	21,420	45,295
2010	16,450	40,363	21,460	45,373
2011	16,690	40,964	21,770	46,044

Panel C: ≥ 3 Children				
Year	Head of Household		Married Filing Jointly	
	Kink 2	Kink 3	Kink 2	Kink 3
	Beginning of Phase-Out	Ending of Phase-Out	Beginning of Phase-Out	Ending of Phase-Out
2001	13,090	32,121	13,090	32,121
2002	13,520	33,178	14,520	34,178
2003	13,730	33,692	14,730	34,692
2004	14,040	34,458	15,040	35,458
2005	14,370	35,263	16,370	37,263
2006	14,810	36,348	16,810	38,348
2007	15,390	37,783	17,390	39,783
2008	15,740	38,646	18,740	41,646
2009	16,420	43,279	21,420	48,279
2010	16,450	43,352	21,460	48,362
2011	16,690	43,998	21,770	49,078

Notes: Technical documentation on EITC program eligibility and benefit rules are available in IRS Publication 596. Nominal dollar values are listed. Historical EITC parameters are available at www.taxpolicycenter.org/. Beginning in 2002, the phase-out points for married filing jointly taxpayers were higher than the phase-out points for head-of-household filers. Conditional on the number of qualifying children, the differences between the phase-out points for head-of-household and married filing jointly taxpayers is 0 in 2001, \$1000 in 2002-2004, \$2000 in 2005-2007, \$3000 in 2008, \$5000 in 2009, \$5010 in 2010, \$5080 in 2011.

Appendix Table 3: EITC K1, Accounting for Nonlinearit in Enrollment-Income Relationship

	Polynomials in Earned Income & AGI		
	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment
Earnings Relative to Kink (kinkdist)	-0.189 [0.00259]	0.645 [0.0336]	0.742 [0.0527]
Slope Change at Kink (D*kinkdist)	-0.444 [0.00587]	-0.227 [0.0658]	
Effect of \$1000 on Enrollment (IV)			0.512 [0.149]
N	1375626	1375626	1375626

	Polynomials in Senior & Junior Income		
	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment
Earnings Relative to Kink (kinkdist)	-0.192 [0.00276]	0.662 [0.0406]	0.778 [0.0663]
Slope Change at Kink (D*kinkdist)	-0.448 [0.00625]	-0.271 [0.0800]	
Effect of \$1000 on Enrollment (IV)			0.606 [0.179]
N	1189802	1189802	1189802

Notes: Each coefficient is estimated from a separate regression. Each regression includes dummy variables for senior year, number of children and filing status. Standard errors are clustered based on \$100 bins of earnings relative to the kink point.

Appendix Table 4: EITC Kink 1, Additional Control Variables and Alternative Sample Restrictions

	Controlling for Year*# Kids*Filing Status Fixed Effects			earned income - wages < 5		
	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment
Earnings Relative to Kink (kinkdist)	0.442 [0.000687]	0.231 [0.0245]	0.0409 [0.0729]	0.442 [0.000760]	0.182 [0.0260]	-0.011 [0.0770]
Slope Change at Kink (D*kinkdist)	-0.339 [0.00193]	-0.146 [0.0700]		-0.339 [0.00204]	-0.148 [0.0741]	
Effect of \$1000 on Enrollment (IV)			0.430 [0.206]			0.437 [0.217]
N	1375626	1375626	1375626	1199807	1199807	1199807

	Earned Income = Wages			Total Income = AGI		
	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment
Earnings Relative to Kink (kinkdist)	0.445 [0.00110]	0.195 [0.0358]	-0.154 [0.116]	0.443 [0.000719]	0.228 [0.0239]	0.0325 [0.0705]
Slope Change at Kink (D*kinkdist)	-0.341 [0.00311]	-0.267 [0.109]		-0.339 [0.00203]	-0.149 [0.0679]	
Effect of \$1000 on Enrollment (IV)			0.784 [0.318]			0.441 [0.199]
N	542994	542994	542994	1331586	1331586	1331586

Notes: Each coefficient is estimated from a separate regression. Each regression includes dummy variables for senior year, number of children and filing status. Standard errors are clustered based on \$100 bins of earnings relative to the kink point.

Appendix Table 5: EITC Kink 1, Alternative Bandwidths

	[K1-\$6000,K1+\$3000]			[K1-\$5000,K1+\$2000]		
	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment
Earnings Relative to Kink (kinkdist)	0.442 [0.000722]	0.234 [0.0246]	0.0345 [0.0737]	0.442 [0.000924]	0.275 [0.0296]	-0.00643 [0.108]
Slope Change at Kink (D*kinkdist)	-0.338 [0.00199]	-0.153 [0.0705]		-0.337 [0.00384]	-0.215 [0.101]	
Effect of \$1000 on Enrollment (IV)			0.451 [0.208]			0.638 [0.294]
N	1375626	1375626	1375626	1077216	1077216	1077216
	[K1-\$4000,K1+\$1000]			[K1-\$3000,K1+\$3000], max symmetric BW		
	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment
Earnings Relative to Kink (kinkdist)	0.44 [0.00124]	0.293 [0.0409]	0.146 [0.195]	0.44 [0.00170]	0.349 [0.0462]	-0.0791 [0.0850]
Slope Change at Kink (D*kinkdist)	-0.328 [0.00867]	-0.11 [0.170]		-0.335 [0.00298]	-0.326 [0.0937]	
Effect of \$1000 on Enrollment (IV)			0.334 [0.512]			0.974 [0.278]
N	774322	774322	774322	1015643	1015643	1015643

Notes: Each coefficient is estimated from a separate regression. Each regression includes dummy variables for senior year, number of children and filing status. Standard errors are clustered based on \$100 bins of earnings relative to the kink point.

Appendix Table 6: EITC Kink 1, Alternative Clustering for Standard Errors

	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment
Earnings Relative to Kink [kinkdist]	0.442	0.234	0.0345
Std Errors clustered based in \$100 bins of earnings relative to kink [baseline]	[0.000722]	[0.0246]	[0.0737]
No clustering	[0.000625]	[0.0221]	[0.0634]
Std Errors clustered based on year and \$100 bins of earned income	[0.00703]	[0.0238]	[0.0713]
Std Errors clustered based on year and ZIP-3	[0.000848]	[0.0229]	[0.0653]
Slope Change at Kink [D*kinkdist]	-0.338	-0.153	
Std Errors clustered based in \$100 bins of earnings relative to kink [baseline]	[0.00199]	[0.0705]	
No clustering	[0.00172]	[0.0608]	
Std Errors clustered based on year and \$100 bins of earned income	[0.0242]	[0.0661]	
Std Errors clustered based on year and ZIP-3	[0.00207]	[0.0624]	
Effect of \$1000 on Enrollment [IV]			0.451
Std Errors clustered based in \$100 bins of earnings relative to kink [baseline]			[0.208]
No clustering			[0.180]
Std Errors clustered based on year and \$100 bins of earned income			[0.202]
Std Errors clustered based on year and ZIP-3			[0.184]
N	1375626	1375626	1375626

Notes: Each column is estimated from a separate regression. Each regression includes dummy variables for senior year, number of children and filing status. Standard errors are shown in brackets.

Appendix Table 7: EITC Kink 1, Nonlinear Functions of Running Variable

A. Linear Polynomial in Earnings Relative to Kink (Baseline)			
	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment
Earnings Relative to Kink (kinkdist)	0.442 [0.000722]	0.234 [0.0246]	0.0345 [0.0737]
Slope Change at Kink (D*kinkdist)	-0.338 [0.00199]	-0.153 [0.0705]	
Effect of \$1000 on Enrollment (IV)			0.451 [0.208]
N	1375626	1375626	1375626
B. Quadratic Polynomial in Earnings Relative to Kink			
	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment
Earnings Relative to Kink (kinkdist)	0.438 [0.00258]	0.497 [0.0890]	-0.344 [0.264]
Slope Change at Kink (D*kinkdist)	-0.327 [0.00808]	-0.627 [0.249]	
Effect of \$1000 on Enrollment (IV)			1.921 [0.756]
N	1375626	1375626	1375626
C. Cubic Polynomial in Earnings Relative to Kink			
	First Stage Dep Var = Refund	Reduced Form Dep Var = Enrollment	IV Dep Var = Enrollment
Earnings Relative to Kink (kinkdist)	0.43 [0.00582]	0.465 [0.225]	0.765 [0.644]
Slope Change at Kink (D*kinkdist)	-0.305 [0.0189]	0.213 [0.585]	
Effect of \$1000 on Enrollment (IV)			-0.699 [1.907]
N	1375626	1375626	1375626

Notes: Each coefficient is estimated from a separate regression. Each regression includes dummy variables for senior year, number of children and filing status. Standard errors are clustered based on \$100 bins of earnings relative to the kink point.