Three Essays on the Economics of Education

Alexander Apt Smith Swarthmore, PA

B.A., Williams College, 2006 M.A., University of Virginia, 2011

A Dissertation presented to the Graduate Faculty of the University of Virginia in Candidacy for the Degree of Doctor of Philosophy

Department of Economics

University of Virginia May, 2015

ABSTRACT

In my first chapter, I investigate the impact of the minimum wage on the schooling decisions of teenagers. While the possible disemployment effect of the minimum wage on teenagers has been the subject of contentious debate, comparatively little attention has been paid to the impact of the minimum wage on teen educational outcomes. This is surprising given that education is more directly linked to the later-life success of teenagers than is teen employment. In this paper, I investigate the educational effects of changes in the minimum wage, looking specifically at high school dropout decisions. I identify the effect of the minimum wage using two sources of variation (within state over time and cross-border at one point in time) and three individual-level datasets (ACS, CPS, and SIPP). I consistently find that a 10% increase in the minimum wage lowers the likelihood of dropping out for low-SES teenagers by 0.5-0.9 percentage points, roughly 4-10% of the group's dropout rate, but has no effect on higher-SES teenagers. Additionally, I find that an increase in the minimum wage has a negative effect on hours worked that is concentrated at the upper tail of the hours distribution (not at the employment margin) for low-SES teens, but not for other young and similarly low-skilled groups. Taken together, these findings suggest that an increase in the minimum wage generates an income effect on low-SES teens, which leads them to shift their allocation of time to schoolrelated activities and away from paid work.

In my second chapter, I investigate the long-run impacts of universal pre-kindergarten (UPK) on criminal activity. There is little evidence to date regarding the long-run effects of statewide universal preschool programs, only studies of programs targeted at more at-risk populations (e.g. Head Start and Perry Preschool) that are often more resource-intensive. I estimate the impact of Oklahoma's universal prekindergarten program (UPK) on later criminal activity, an outcome that accounted for 40-65% of the large estimated long-run benefits of Perry Preschool. I assemble data on criminal charges in the state of Oklahoma and identify the effect of UPK availability using a regression discontinuity design that leverages the birthdate cutoff for UPK in the program's first year of implementation. I find significant negative impacts of UPK availability on the likelihood that black children are later charged with a crime at age 18 or 19 of 7 percentage points for misdemeanors and 5 percentage points for felonies. I find no impact on the likelihood of later charges for white children. The results suggest that universal Pre-K can, like more targeted programs, have dramatic effects on later criminal outcomes, but these effects are concentrated among more at-risk populations.

In my third chapter, I investigate the impacts of performance-based compensation systems with co-authors James Wyckoff and Thomas Dee. We examine systems of differential teacher pay based on student test performance and classroom observations implemented by districts across the country as part of the Teacher Incentive Fund (TIF), a large competitive grant program operated by the U.S. Department of Education. We use student-level data from six TIFparticipating districts in the state of Oregon to provide early evidence on how these high-profile reforms influenced measures of student performance. We estimate this impact of TIF by leveraging the discontinuous rule that defines a school's TIF-eligibility in a regressiondiscontinuity design. Our approach provides an effective proof of concept for TIF in two ways. First, we capture the treatment contrast between implementing the full bundle of TIF-reforms and the status quo. Second, we focus on districts with high fidelity of implementation compared to many of the participating districts. We find significant effects of TIF in improving student reading achievement but not math achievement.

CHAPTER 1

Dollars and Dropouts: The Minimum Wage and Schooling Decisions of Teenagers

1. INTRODUCTION

The minimum wage has generated hotly contested political debate in the United States for nearly a century. President Roosevelt passed the federal Fair Labor Standards Act of 1938 that established the first minimum wage (25 cents per hour) only after his landslide victory in 1936, protracted debate in Congress, and earlier Supreme Court rulings of unconstitutionality.¹ In his 2014 State of the Union address, President Obama called for an increase in the federal minimum wage to \$10.10 an hour (from \$7.25) as a way to reduce income inequality and improve economic opportunities for low-skilled workers. The American public broadly supports an increase in the minimum wage, by 71% in a 2014 Quinnipiac poll, but Congress remains deadlocked on the issue.^{2,3} In the absence of federal action, many state governments have voted in the past year to raise their minimum wage.⁴

The economic literature on the consequences of the minimum wage is both vast and contentious. Researchers concentrate mainly on the possible disemployment effects of the minimum wage, with a frequent focus on teenagers. Teenagers are the age group most subject to

¹ Jonathan Grossman, "Fair Labor Standards Act of 1938: Maximum Struggle for a Minimum Wage," http://www.dol.gov/oasam/programs/history/flsa1938.htm.

 $^{^{2}}$ The large increase advocated by the president enjoys a slimmer majority of support (51%).

³ "Obama Approval Plunge Levels Off, Quinnipiac University National Poll Finds; Hike Minimum Wage, Extend Jobless Benefits, Voters Say," http://www.quinnipiac.edu/news-and-events/quinnipiac-university-poll/national/release-detail?ReleaseID=1993, (January 8, 2014).

⁴ In 2014, 34 states considered increases to the minimum wage. 10 states enacted increases: CT, DE, HI, MD, MA, MI, MN, RI, VT, and WV.

the minimum wage: one quarter of employed teenagers earn the minimum wage, and many more earn only slightly above the minimum, making them likely to be swept up by any change.^{5,6} The recent literature on minimum wage and teen employment is divided between traditional stateand-year-fixed-effect studies, which find large negative effects, and newer studies which use local area controls, such as cross-border designs, and find small negative or null effects.⁷ Comparatively few U.S. studies have focused on the effect of the minimum wage on teen educational outcomes. The results from these studies are mixed and they suffer from data limitations such as substantial sample selection problems, short time windows, and coarse, noisy measures. Moreover, these earlier studies do not make the critical distinction between teens of varying socio-economic status (SES). This distinction is important because higher-SES teens, with more-educated parents and more family resources, are much less likely than lower-SES teens to be on the margin of leaving school.

The lack of attention to educational outcomes in the minimum wage literature is surprising given the primary importance generally accorded to teen educational attainment by policymakers and researchers, who typically view education as more closely linked to the laterlife success of teenagers than teen employment.⁸ While the effect of the minimum wage on educational outcomes is indirect, it is an intuitive byproduct of human capital theory. Teenagers

⁵ Only a tenth of 20-24 year old and a twentieth of 25-34 year old wage earners earn the minimum wage (calculations from March CPS).

⁶ Turner and Demiralp (2001) estimated that two thirds of employed teenagers were affected by the minimum wage increase in 1992.

⁷ Neumark et al (2013) argue that these differences arise because the cross-border designs throw out too much relevant variation, while Allegretto et al (2013) argue that the differences are due to bias in the state and year fixed effect estimates due to time-varying spatial heterogeneity.

⁸ Table A1 shows naïve regressions of various life outcomes at age 26 on education and employment outcomes at age 18 using data from the National Longitudinal Survey of Youth 1997 (NLSY97). When controlling for other covariates, dropping out of high school is associated with decreases in later-life outcomes (i.e. employment, income, family income, and arrest record at age 26) that are more than twice the magnitude of the positive associations with employment at age 18.

face a tradeoff between investing in human capital (i.e. time and effort spent on school) and immediate consumption from time spent in the labor market.⁹ To the extent that changes in the minimum wage alter the labor market opportunities faced by teenagers, these changes will also alter their investment-consumption tradeoff.

In this paper, I investigate the impact of changes in the minimum wage on high school dropout decisions using two distinct sources of variation and three individual-level datasets. I use two decades of the March Current Population Survey (March CPS), one decade of the American Community Survey (ACS), and the 1996, 2001, 2004, and 2008 4-year panels of the Survey of Income and Program Participation (SIPP). To identify the effect of minimum wage change, I leverage both variation in minimum wage rates within states over time and variation in minimum wage rates between neighboring localities on either side of a state border at a given point in time. Each of these sources of data and variation yield different advantages and disadvantages in measuring the educational effect of the minimum wage. I use them in concert to address many more threats to the internal validity of my estimates than would be possible if I used only one source of data and variation.

Consistently across data sources, sources of variation, and empirical specifications, I find that an increase in the minimum wage lowers the likelihood that low-SES teenagers will drop out of high school but has no effect on other teenagers.¹⁰ The effect on low-SES teens, who are at higher risk of dropping out *ex ante*, is substantial. I find that a 10% increase in the minimum

⁹ Teenagers can also consume using transfers from their family. These transfers could be affected by changes in household income caused by minimum wage changes.

¹⁰ The robustness across empirical approaches in the educational effects of the minimum wage, but not in the employment effects in prior studies, suggests that minimum wage may be endogenous with respect to employment outcomes but not education outcomes. This could occur if state minimum wage policy changes are made in response to the overall unemployment rate.

wage, equivalent to an increase of 73 cents per hour at the current federal minimum, lowers the likelihood of dropping out by 0.4-0.9 percentage points, or between four and ten percent of the average dropout rate for the low-SES group.

I investigate the mechanism by which the minimum wage lowers the dropout rate for low-SES teens and find evidence that the affected teens work fewer hours after a minimum wage increase. This impact is concentrated at the upper tail of the hours distribution rather than at the employment margin. A substantial literature finding that high work hours increases the chance of dropping out lends further support to this mechanism for the observed educational effects. Interestingly, there is no similar effect on hours of an increase in the minimum wage for other young low-skilled groups comparable to low-SES teens. This suggests that the effect on low-SES teens arises from an income effect of the minimum wage rather than from a change in labor demand, which would have to disproportionately affect low-SES teens and not other young lowskilled workers.

These findings suggest that the near-exclusive focus of prior research on the teen employment effects of the minimum wage may have missed the bigger picture. While increasing the minimum wage reduces hours worked by teenagers, especially low-SES teenagers, this may be due to an income effect which leads low-SES teens to shift toward human capital investment and away from work for immediate consumption. Such a shift will benefit teenagers and society in the long run if teenage educational outcomes are more important than teenage employment outcomes in determining later-life well-being.¹¹

¹¹The long-run benefits of increased human capital investment by teenagers are public as well as private. The public cost of high school dropouts is substantial. Compared with a high school graduate, a high school dropout yields less

The rest of this paper is organized as follows. In Section 2, I review the existing literature on minimum wage and educational outcomes. In Section 3, I discuss the possible effects of the minimum wage suggested by economic theory in order to frame my empirical approach. In Section 4, I describe the various data sources, sample constructions, and variable definitions that I employ. In Section 5, I present my various empirical strategies. In Section 6, I discuss my empirical results. In Section 7, I conclude.

2. THE MINIMUM WAGE, TEEN EMPLOYMENT, AND TEEN SCHOOLING

In this section, I briefly review three areas of the literature relevant to my analysis of the effect of minimum wage increases on high school dropout decisions. First, I give a brief overview of the controversies in the broader minimum wage literature on teen employment. Second, I discuss the literature that looks at the minimum wage's effects on teen educational outcomes. Third, I summarize the literature that examines the relationship between a teen's employment and her decision to drop out of school.

2.1 Minimum Wage and Teen Employment

The economic literature examining the effects of the minimum wage is both large and contentious. It has primarily concentrated on the possible disemployment effects of the minimum wage. These studies have generally focused on specific populations likely to be most affected by the minimum wage. Teenagers, in particular, have received much of the attention, as minimum wage earners make up a larger proportion of workers in this age group than any other.

tax revenue, requires more benefits, and is more likely to be arrested or incarcerated. This leads to a \$200,000 higher lifetime cost to the government for a high school dropout than for a graduate (Levin et al., 2007).

The most recent literature on the minimum wage's effect on teen employment has been divided into two camps. The first, led by Neumark and Wascher in several works (Neumark, 1992, 2006; Neumark and Wascher, 1995, 2007; Neumark, Salas, and Wascher, 2013a), use the traditional two-way (i.e. state and year) fixed effect approach and find substantial disemployment effects (elasticities between -0.1 and -0.3). The second, pioneered by Card and Krueger (1992) and exemplified by Allegretto et al (2013), criticize the traditional approach for failing to account for spatial heterogeneity in labor market conditions and advocate the use of local area controls. Primarily, they use cross border designs which compare neighboring localities that cross a state border. These studies tend to find small or null disemployment effects.

2.2 Minimum Wage and Teen Enrollment

Despite the spotlight on teenagers in the minimum wage literature, relatively little attention has been paid to the possible effect that the minimum wage may have on their educational outcomes, which have been shown to affect teens' later-life earnings, crime, and health outcomes (Card, 1999; Lochner, 2011). Recent studies in the U.S. and other developed countries have come to conflicting conclusions regarding the effects of the minimum wage on teenage schooling. A number of studies have found negative enrollment effects (Neumark and Wascher, 1995, 1995b, 2003; Turner and Demiralp, 2001; Chaplin et al. 2003) while others have found mixed or null enrollment effects (Warren and Hamrock 2010; Campioleti et al, 2005; Pacheco and Cruickshank, 2007) or positive enrollment effects (Matilla, 1978, 1982). I discuss the most recent studies in the U.S. below.

Neumark and Wascher (1995b, 2003) and Turner and Demiralp (2001) employ multinomial logit approaches that use joint enrollment-employment outcomes. They both find that increases in the minimum wage positively affect transitions from enrolled to not enrolledemployed, but Neumark and Wascher also find a positive effect on transitions to not enrollednot employed while Turner and Demiralp do not. Turner and Demiralp use data from 1991 and 1992 waves of the 1990 SIPP, spans only two years with only one change in minimum wage with which to identify the effect of the minimum wage. Neumark and Wascher (1995b, 2003) use matched observations in consecutive years of the May CPS for the teenagers (16-19 year old). However, in order to be observed in the second year, the household as a whole must not have moved and the teenager must not have moved out of the household. This is an especially problematic sample selection problem for this age group, as only 65% of teenagers remain in the matched sample and being in the matched sample is likely to be endogenous in that "moving out" is likely related to a teenager's employment-enrollment status. For example, 19 year olds that leave home to go to college will be excluded from the matched sample leading the enrollment estimates to be biased downward.

Chaplin et al. (2003) and Warren and Hamrock (2010) use state-year panel data from the Common Core of Data for 1982-2005 and 1979-1986, respectively. Both use state and year fixed effects specifications. Chaplin et al. find that increases in the minimum wage reduce teen school enrollment, specifically at the 9th to 10th grade transition. They measure changes in enrollment using total enrollments in grade-state-years. As Warren and Hamrock point out, this measurement approach does not account for interstate migration, grade retention, and changes in incoming cohorts over time. Not accounting for grade retention is likely to be particularly problematic for measuring the transition from ninth to tenth grade, since upper-level students are often held back in 9th grade (Heckman and LaFontaine, 2010).

Warren and Hamrock attempt to mitigate these measurement problems by constructing a cohort-specific state-level measure of the high school completion rate based on a comparison of public high school completers with the *estimated* first-time ninth grade enrollment three years earlier (accounting for interstate migration in the intervening years). They find null results on the effect of the minimum wage on high school completion, with large standard errors on the minimum wage coefficient estimate. This imprecision is not surprising given the coarseness of the data (no individual-level or local area controls), the noisiness of a completion measure which uses an estimate as the denominator, and the assumption that the minimum wage only affects dropout behavior in a teenager's senior year.

I investigate a slightly different outcome than these prior studies. Rather than looking at aggregate continuation or graduation ratios (9th-12th graders) or joint enrollment-employment outcomes (age 16-19), I focus on high school drop out and enrollment outcomes (age 16-18). Individual-level outcomes rather than aggregate ratios enable me to control for individual characteristics that might influence drop out decisions. Limiting my sample to 16-18 year olds narrows my focus to decisions regarding completing high school, whereas including 19 year olds, as in Neumark and Wascher(1995b, 2003), means also capturing decisions of whether to go to college. Focusing on education rather than education and employment jointly follows from the greater weight of teen education compared to teen employment outcomes in predicting later-life success. Table A1 shows this with naive regressions of outcomes at age 26 (i.e. employment, income, family income, arrest record) on employment and dropout at age 18 using data from the 1997 National Longitudinal Survey (NLSY97). The magnitude of the association of dropping out at age 18 with later-life outcomes is more than twice that of employment at age 18, and the interaction between the two is insignificant.

My work improves on the existing literature in a number of ways. First, I look at the effect of the minimum wage on high school dropout behavior separately by SES status. This is a critical distinction as my analysis concentrates on the teenagers who are at the highest risk of dropping out and who are likely to receive the largest income effect from minimum wage changes. Second, I use multiple individual-level data sources in concert, which improve on prior data used in U.S. studies. I have a direct measure of dropout and enrollment rather than estimates of aggregate rates, as in Chaplin et al. (2003) and Warren and Hamrock (2010). I have data spanning up to 20 years of minimum wage changes (compared to two years in Turner and Demiralp, 2001) and including more recent changes than elsewhere in the literature. In my analysis using the SIPP, I observe nearly all teens and can therefore rule out that my results are driven by sample selection, a real concern for Neumark and Wascher (1995b, 2003). Third, I am the first to use local cross-border variation in minimum wage in a given year while investigating educational effects of the minimum wage. Fourth, to my knowledge, I am the first to provide evidence that the minimum wage improves low-SES teen educational outcomes through an income effect that reduces the likelihood of these teens working long hours.

2.3 Teen Employment and Teen Schooling

A large literature finds a substantial negative relationship between high intensity work (i.e. greater than 20 hours) and academic achievement in high school. The central question in this literature is whether this relationship is causal. In other words, does paid work take time and effort away from other activities that improve achievement (e.g. studying and doing homework)? Or, are teens that work long hours different from those who don't in unobserved ways that are related to academic achievement (e.g. ability)? Controlling for the rich set of covariates available in the 1979 National Longitudinal Survey of Youth, Ruhm(1997) finds that high school seniors working 20 (40) hours in the interview week complete 0.21 (0.68) years less schooling than non-workers. Turner (1994) uses the High School and Beyond survey of high school sophomores and seniors in 1980. He accounts for the potential endogeneity of work decisions by employing an individual fixed effects strategy that instruments for work decisions with labor market conditions. He finds that working 30 or more hours significantly reduces standardized test scores and grade point averages, and significantly increases the likelihood of dropping out.

Rothstein (2007) employs a similar empirical strategy using the National Longitudinal Survey of Youth 1997 to measure the effect of hours worked on grade point average in grades 10, 11, and 12. She finds small but significant negative effects for the individual fixed effects specification that imply a reduction of less than -0.04 in GPA points from working 20 hours per week. Her instrumental variables results show large, but insignificant negative effects for men of 0.24 in GPA points from working 20 hours per week. These small and insignificant estimates are not surprising, given the assumption, contrary to descriptive evidence in the prior literature, that the relationship between work intensity and academic achievement is linear. Taken as a whole, the literature is suggestive of a non-linear effect of hours worked on academic achievement, concentrated at the high end of the hours distribution. However, the internal validity of the effect estimates remains unclear.¹²

3. ECONOMIC FRAMEWORK

In this section, I briefly discuss the various theoretical channels – including labor demand and labor supply (income and substitution effects) – through which the minimum wage could

¹² Local labor market conditions may be correlated with other factors affecting drop out, such as family income.

impact the educational outcomes of teenagers and discuss how these could differ by SES status. While the interrelation of work and schooling decisions is a fundamental notion of human capital theory and the connection between the minimum wage and work decisions is clear from economic theory more broadly, the direction of the impact of the minimum wage on schooling is theoretically ambiguous. As Neumark and Wascher (2006) conclude, "Theoretical models of how minimum wages might affect schooling decisions have quite a few layers of complexity and provide no clear predictions." (p. 210).

3.1 Teens' Work, School, and Leisure Time

As a framework for this discussion, I will assume that a teen allocates her time between work (*h*), school (*s*), and leisure (*l*) such that 1 = h + s + l. However, she may not be able to choose her hours of work if this is limited by the demand for her labor. Her utility is given by consumption and leisure in the first period (high school age) and discounted expected value of future lifetime consumption: $U = u(c_0, l) + \beta E(c_1)$. Consumption in the current period is defined by $c_0 = W^e + wh$, where W^e is a parental transfer for teens of SES status $e \in \{H, L\}$ and w is the minimum wage. I assume that teens cannot borrow, a realistic simplification given my focus on high school drop out outcomes rather than post-secondary educational outcomes.

Utility in the following period is given simply by the discounted expected future consumption, $\beta E(c_1) = \beta g^e(s)$, where β is the discount rate. Since my discussion focuses on the high school graduation margin, I assume that $g^e(s)$ captures the relative payoffs of graduating from high school (compared to dropping out), as well as the likelihood of graduating at a given level of time spent in school. I assume that $g^e(s)$ is strictly concave in time spent on school (decreasing marginal returns to *s*), and that the probability of high school graduation is more sensitive to time spent on schoolwork for low-SES teens, $\frac{\partial g^{H}(s)}{\partial s} < \frac{\partial g^{L}(s)}{\partial s}$ for all $s \in [0,1]$. The latter assumption follows from the widely-held understanding that higher-SES teens have a variety of advantages relative to low-SES teens that help to ensure their graduation (e.g. more parental involvement, parental modeling of academic achievement, and better school quality).

Given the assumptions above, a minimum wage change that affects time at work – through labor supply or labor demand – implies either an increase in time spent on school (decrease in dropout likelihood), or an increase in time spent on leisure (no change in dropout likelihood), or both. A teen will increase her time spent on school if the expected future return exceeds the current marginal utility of leisure, $\beta \frac{\partial g^e(s^*)}{\partial s} > u_2(c_0^*, l^*)$.¹³ An identical change in time at work for high and low-SES teens could produce differential effects on the likelihood of drop out. First, the dampened dropout sensitivity of higher-SES teens will make shifting from work to leisure (rather than school) relatively more appealing for higher-SES teens than lower-SES teens, all else equal. Second, differences between the *ex ante* time allocation decisions of high and low-SES teens could result in differential effects. For example, if low-SES teens allocate less time to school prior to the minimum wage change, then they would receive a higher return to additional school time, making them more likely to shift towards school.

3.2 Labor Demand

Neoclassical economic theory generates a relatively simple prediction for the effect of the minimum wage on total hours worked. Firms will respond to an increase in the minimum wage

¹³ The concavity of $g(\cdot)$ suggests that a teen will be more likely to shift her time toward school rather than leisure if she has a low level of time allocated to school (high risk of drop out). Additionally, the increased demand for higher-skilled workers may increase the future return to school (g'(s)) further incentivizing a shift towards schooling.

by shifting their inputs away from these workers and toward other (substitutable) production inputs, such as higher-skilled workers or capital.¹⁴ This yields an unambiguous prediction that demand for low-skilled labor will decrease. If firms see all young workers without a HS diploma as equally low-skilled, then high and low-SES teens (as well as slightly older dropouts) would see similar drops in hours worked. As discussed in subsection 3.2, this exogenous (from the workers perspective) reduction in time spent working implies either a null or negative effect on the dropout rate of teens. If firms do not view high-SES and low-SES teens as interchangeable, then those teens seen as higher-skilled (most likely the high-SES teens) would see an increase in labor demand, which could result in an increase work time and a decrease in school time (more likely to drop out) for these teens.

The simple neoclassical model does not generate predictions regarding the effect of the minimum wage on the distribution of hours worked for low-SES teens, only an effect on average hours worked. If firms face a fixed cost for each worker (e.g. the costs of hiring and training), then firms would likely reduce total hours by cutting at the low end of the hours distribution. This would mean a larger negative effect of the minimum wage at the employment margin compared to the effect at higher levels of hours worked. On the other hand, if firms face additional costs in the form of government-mandated benefits (e.g. health insurance coverage) for workers at higher hours levels, then we would see a larger negative effect at higher levels of hours worked than at the employment margin. In either case, the pattern of the minimum wage's effects should be similar for similarly low-skilled workers.

3.3 Labor Supply

¹⁴ This will occur only if the minimum wage exceeds the productivity of low-skill workers

The direction of the effect of a minimum wage increase on labor supply will depend on the relative magnitudes of the opposing income and substitution effects that the increase produces. The higher opportunity cost of non-work time will push a teen to shift her time allocation toward work (and away from leisure, school, or both). At the same time, the additional income gained from the wage increase will reduce the marginal utility of working relative to other uses of time, pushing the teen to shift her time allocation away from work (and toward leisure, school, or both). This income effect will be particularly large for two groups. Teens with high levels of time allocated to work *ex ante* will see the largest increase in available income, all else equal. Teens with low levels of parental transfers (i.e. low-SES teens) will see the largest relative increase in their available income, for a given allocation of time at work, since their own income accounts for much more of their first period consumption.¹⁵ These low-SES teens that receive little in parental transfers are also more likely to be constrained by their inability to borrow, thereby increasing the impact of additional income on school time. As in the canonical model of credit constraints in education (Lochner and Monge-Naranjo, 2012), individuals who are constrained by the inability to borrow will underinvest in school because they are unable to bring the additional future consumption into the present.

3.4 Summary

The theoretical framework presented here suggests that an observed negative effect of the minimum wage on the likelihood that low-SES teens drop out, but no observed effect on high-SES teens, could be generated through two primary mechanisms. First, the increase in minimum wage could cause a decrease in the demand for low-skilled labor. This would lead to a similar

¹⁵ It is also possible that the minimum wage increases the teen's household income (and thereby parental transfers) through high wages for her family members.

reduction in time spent working for both low and high-SES teens (assuming they are similarly low-skilled). However, a similar reduction in work may only translate into improved educational attainment for low-SES teens because their academic success is more sensitive to time inputs.¹⁶ Second, the increase in minimum wage could cause a large income effect (relative to the substitution effect) among low-SES teens because of their low parental transfers and inability to borrow. This would lead low-SES teens, particularly those already working long hours (the group receiving the largest income effect), to re-allocate their time away from work and toward activities that lower their chances of dropping out.

4. DATA

I match data on state-level minimum wage rates and local labor market characteristics to three individual-level datasets with information on teenagers' labor market and educational outcomes: the Current Population Survey (CPS), the 2000 Census and American Community Survey (ACS), and the Survey of Income and Program Participation (SIPP). Each of these datasets has different advantages and disadvantages in measuring the impact of minimum wage changes on teenage educational outcomes. Taken as a group, these datasets allow me to avoid the major drawbacks of the data used in prior research (e.g. imprecise measures and substantial endogenous sample selection), while also allowing me to differentiate effects on high and low-SES teenagers.

4.1 March Current Population Survey (CPS)

¹⁶ Alternatively, this could be a result of differing preferences for leisure or ex ante time allocations between low and high-SES teens.

I obtain data from the March CPS, a nationally representative annual cross-section of approximately 60,000 households, from the Integrated Public-Use Microdata Series (IPUMS). This survey allows for analysis over a relatively long time horizon since it has asked the same education question for two decades.¹⁷ However, it has a relatively small sample size, when considering narrowly defined age groups, and only contains coarse geographic information (i.e. state of residence) for all individuals.¹⁸ The data contains information on recent labor market outcomes (e.g. employment status, hours worked, and income) and educational outcomes (e.g. enrollment status and educational attainment), as well as demographic information. For the 86% of 16-18 year olds who are observed in a household with their parent/guardian, the CPS also provides information on the income and education of their parents and guardians. I use the CPS to construct an annual sample of 16-18 year olds for the years 1992-2012, which I will refer to as CPS-Teens. The sample consists of 499,610 observations. Table 2, Column 1 shows summary statistics for the CPS-Teens sample.

4.2 2000 Census and American Community Survey (ACS)

I obtain data from the 2000 Census and American Community Survey (ACS) from IPUMS. Like the March CPS, the 2000 Census and ACS are nationally representative, individual-level annual cross-sections. The 2000 Census and 2005-2011 ACS provide a large sample (1-in-20 and 1-in-100, respectively) and residence information at the Public-Use Microdata Area (PUMA) level for all individuals. PUMAs are geographic areas built on counties and census tracts that contain at least 100,000 residents and do not cross state lines.

¹⁷ Prior to 1992, the IPUMS-CPS does not differentiate between having attended 12 years of school and obtaining a high school diploma or equivalent.

¹⁸ The IPUMS-CPS contains county of residence and metropolitan area of residence for some individuals in large counties or metropolitan areas, but the sample size is not large enough for meaningful analysis at this level of granularity.

County of residence is also provided for counties with at least 100,000 residents. The 2001-2004 ACS provides a smaller sample (less than 1-in-230) and only state of residence information. The ACS contains similar educational outcomes and demographic information as the CPS, but the reference period for hours worked is the prior year. Consequently, I use only the educational outcomes from this data.¹⁹ Similar to the CPS, parental or guardian education and income is observed only for the 86% of 16-18 year olds who live in the same household as the parent(s) or guardian(s).

I use the 2000 Census and ACS data to create five distinct samples with increasing specificity about geographic location. With these samples, I account for educational and labor differences between local areas and leverage variation in minimum wage within a local labor market at given point in time. First, I construct a sample similar to the CPS-Teens sample, which I call ACS-Teens. This sample is restricted to16-18 year olds with observed demographic, education, and parental education information. It covers the period 2000-2011 and identifies residence at the state-level. Second, I restrict the ACS-Teens to 2005-2011, so that residence can now be identified at the PUMA-level (ACS-P-Teens). Third, I restrict the ACS-P-Teens sample to individuals with county of residence information (ACS-C-Teens).²⁰ Fourth, I conduct a probabilistic match between the Public Use Microdata Areas (PUMAs) in the ACS-P-Teens sample and their corresponding commuting zones.²¹ Commuting zones are clusters of counties,

¹⁹ In the ACS, I am not able to determine a teenagers' enrollment status at the time of their reported hours worked. Also, hours worked for 18 year olds would be observed at age 19, when parental education is much less likely to be observed.

²⁰ Public Use Microdata Areas (PUMAs) are statistical geographic areas built on counties and census tracts that contain at least 100,000 residents and do not cross state lines. In sparsely populated areas PUMAS can be made up of multiple counties (and/or portions of multiple counties), while in highly populated areas, a county may contain multiple PUMAs. PUMA of residence is provided for all individuals in the ACS for 2005-2011, while county of residence is provided only for individuals in counties with at least 100,000 residents.

²¹ I obtain the geographic crosswalk file matching PUMAs to 1990 Commuting Zones from Autor and Dorn (2013).

sometimes spanning state borders, which have strong within-cluster commuting ties and weak between-cluster commuting ties. Since PUMAs do not perfectly overlap county boundaries, some are mapped to multiple commuting zones. Rather than throwing out observations from PUMAs whose commuting zone is ambiguous, the probabilistic matching procedure multiplies the population weight by the fraction of the area of the teen's PUMA in the given commuting zone. I will refer to this sample as ACS-PCZ-Teens. The ACS-Teen, ACS-P-Teen, ACS-PCZ-Teen, and ACS-C-Teen samples consist of 1.7 million, 0.9 million, 0.9 million, and 0.5 million unique observations, respectively. Table 2, Columns 2-5 shows summary statistics for these samples.

4.3 Survey of Income Program Participants (SIPP)

The SIPP is a longitudinal survey that interviews individuals every four months for 3-4 years. Since its 1996 redesign, the U.S. Census Bureau has conducted four SIPP panels, beginning in 1996, 2001, 2004, and 2008. Similar to the CPS, these panels provide information on education outcomes, demographics, and recent labor market outcomes, with residence identified at the state-level. The sample sizes of 40,000-50,000 households are substantial given the level of detail and frequency of the data, and the oversampling of low-income households improves statistical power when examining low-SES teenagers. The SIPP has two advantages over the CPS and ACS data. First, unlike in the cross-sectional datasets, parental/guardian education is observed for nearly all 16-18 year olds in the SIPP (99%). The SIPP follows all individuals in the initial sample as they move to new residences, which means parental education can be observed if the teenager was initially (or at any point) observed in the same household as their parents/guardians. Therefore, the SIPP provides a way to check that results obtained from the CPS and ACS are not driven by endogenous sample selection of teenagers with observable

parental education. Second, the panel nature of the data allows for analyses of labor market outcomes that control for the unobserved time-invariant traits of individuals. I create a sample of 16-18 year olds with three observations per year for each individual by combining four SIPP panels (1996, 2001, 2004, and 2008). Table 2, Column 7 shows the summary statistics for this sample, which I will refer to as SIPP-Teens.

4.4 Defining Key Variables

Given my focus on high school dropout behavior, defining when a teenager has "dropped out" is of central concern. All three datasets provide two pieces of information on 16-18 year olds: attainment and enrollment. Attainment gives the highest level of education *completed* by an individual at the time they are surveyed. Enrollment shows whether an individual attended school in a given period at or prior to the time they are surveyed. The timeframe of the enrollment question varies slightly between the surveys. The CPS asks whether the respondent was enrolled in school in the previous week, the ACS asks for the past 3 months, and the SIPP asks for the past 4 months.²² Using these two sources of information I define a simple dropout indicator applicable to all three datasets as follows: teenager *i* has dropped out if she is not currently enrolled in school and she has not obtained a high school diploma (or greater level of education).^{23,24} This is a stock variable indicating those who are *currently dropouts* (regardless of how recently they dropped out), rather than a flow variable, which would indicate those who had newly transitioned from enrolled to dropout. Table 2 shows the proportion of teens that are defined as high-SES (80-84%) and the average dropout rates by SES (9-12% for low-SES and 3-

²² Those on holiday or season vacations were told to answer that they were enrolled.

²³ For comparability between datasets and over time within datasets, I count GED recipients as equivalent to HS diploma holders in the primary analysis. I also include students enrolled part-time as currently enrolled in my primary analysis. ²⁴ For the SIPP, I only count the individual as "dropped out" if they satisfy this definition for two waves in a row.

4% for high-SES) for select data samples. Figure 1 shows the modest downward trends in these dropout rates over time, particularly among low-SES teens. It is possible that some portion of this trend is driven by increased GED recipients, but my data sources do not treat GED recipients separately from high school graduates for sufficient time periods to analyze. Therefore, to ensure that this or some other aspect of my dependent variable construction is not driving my results, I repeat all analyses using a simple enrollment indicator as the dependent variable.²⁵

In my primary analysis, I define a teenager's socio-economic status using the educational attainment of her parents. Specifically, I define a teenager as "high SES" if *all* of her observed parents (or guardians) have graduated from high school. I define a teenager as "low SES" if *any* of her observed parents (or guardians) has not graduated from high school. In all data samples, roughly 80% of 16-18 year olds with observed SES are classified as high SES (see Table 2). I provide robustness checks with alternative definitions of high-SES as those with all observed parents attended "some college" or as those with household income (excluding their own income) above the p^{th} percentile of the yearly distribution, for $p \in \{20,30,40,50\}$. My primary formulation has two advantages over these potential alternatives. First, family income is relatively volatile from year-to-year and may be affected by the minimum wage. Second, using parental high school education to define low SES effectively identifies teenagers who are at high risk of dropping out. Using this definition, the dropout rate of low-SES teenagers is more than three-times that of high-SES teenagers (see Table 2).

The potential drawback of my primary SES definition, or any other that relies on observing a teenager's parents, is that the presence of parents in the household may be

²⁵ Additional analyses using the October CPS, which enables GEDs to be counted as dropouts, do not differ substantively from the March CPS results. These are available upon request.

endogenous. That is, whether teenagers remain in their parents' household may be affected by the minimum wage. Table 2 shows the proportion of each data sample that is missing the SES measure: 14% of ACS-Teen, ACS-P-Teen, ACS-PCZ-Teen, and ASC-C-Teen; 8% of CPS-Teen; and 1% of SIPP-Teen. I use the SIPP longitudinal data, where parents or guardians are observed for nearly all teenagers to corroborate the results from the ACS and CPS cross-sections.

4.5 Other Data

I merge state-level minimum wage and unemployment rate data with the individual-level information on teenagers in each sample. I obtain yearly state and county unemployment rates from the Bureau of Labor Statistics' Local Area Unemployment Statistics. I obtain state-by-month information on state minimum wage rates for 1992-2012 from the Tax Policy Center at the Urban Institute and Brookings Institution.²⁶ Figures 3A and 3B show the substantial variation in minimum wages (constructed as the maximum of federal and state minimum wage laws) over the last two decades. Figure 3A depicts the variation in the minimum sabove the federal minimum is depicted as the black line, while states with minimums above the federal minimum are in gray (the size of the bubble denotes the number of states in a given \$0.25 bin). Figure 3B maps the difference between state and federal minimum wages over time, in percentage terms.

5. EMPIRICAL STRATEGY

²⁶ This data is compiled by the Tax Policy Center from January issues of the Bureau of Labor Statistics' Monthly Labor Review, the 1968-1999 Book of the States published by the Council of State Governments (for 1990-1999), and U.S. Department of Labor data (for 2000-2012).

I utilize two different sources of variation to identify the effect of minimum wage increases on teen educational outcomes. First, I use a traditional quasi-experimental framework with two-way (state and year) fixed effects which leverages the variation within-states, over-time in the minimum wage. Second, I use a cross-border design which leverages variation in the minimum wage at a given point in time between nearby PUMAs in the same commuting zone on either side of a state border. I then investigate household income and hours worked as potential mechanisms through which changes in the minimum wage may affect educational outcomes. I employ an individual fixed effect framework using panel data to investigate these mechanisms. This approach leverages within-state variation in the minimum wage over time while controlling for time-invariant individual characteristics.

5.1 Traditional Two-way Fixed Effects

I first adopt the approach that has been used frequently in the minimum wage literature to investigate employment effects (Neumark and Wascher, 1992) and apply it to all three datasets (CPS, ACS, and SIPP). This approach includes state fixed effects to remove time-invariant differences between states that may be related to both differences in teen outcomes and minimum wage levels, such as the industrial structure of the state economy, the generosity of social welfare programs, and the quality of the state educational system. In my preferred specifications, statespecific polynomial time trends are included to account for these differences evolving smoothly overtime. Year fixed effects are included to remove differences between years, common to all states, that may be related to both outcomes and minimum wage levels, such as shocks to the national economy and the political situation at the federal level. The effect of the minimum wage is identified by variation over time in a state's effective minimum wage (the maximum of the state and federal minimum wages). I diverge from the traditional analysis of employment by including an interaction between the minimum wage term and an individual's socio-economic status, allowing me to examine the differential effects of minimum wage by family background. The basic specification is as follows:

$$y_{ist} = \beta_0 + \beta_1 l n(mw_{st}) + \beta_2 HSES_{ist} + \beta_3 l n(mw_{st}) \cdot HSES_{ist}$$
$$+ \gamma X_{ist} + \nu_t + \theta_s + \tau_s(t) + \epsilon_{ist}, \tag{1}$$

where y_{ist} is the outcome of interest. For the main results, the outcome of interest is an indicator for whether individual *i* in state *s* at time *t* is identified as a high school dropout. $ln(mw_{st})$ is the log of the minimum wage in state s at time t.²⁷ HSES_{ist} is an indicator equal to one if individual *i* is high SES, that is, if *all* of his observed parents/guardians have at least a high school diploma. X_{ist} are demographic characteristics of individual *i* (i.e. indicators for age, sex, race, and whether she is above the state's compulsory schooling age) and characteristics of the labor market in state *s* at time *t* (i.e. state unemployment rate). v_t and θ_s are year and state fixed effects. $\tau_s(t)$, included in some specifications, is a state-specific polynomial time trend to account for differential trends across states. The primary coefficients of interest are β_1 , which captures the impact of changes in the minimum wage on the likelihood that low-SES teenagers will drop out of high school, and $\beta_1 + \beta_3$, which captures the same effect for high-SES teenagers. I estimate this equation using OLS with standard errors clustered at the state-level.

The finer geographic granularity available in the ACS-P-Teens and ACS-C-Teens samples allows for a similar specification which uses within-state minimum wage variation over time for identification, but removes time-invariant differences at the PUMA or county-level,

²⁷ Using the log of the minimum wage, which measures minimum wage changes as percent rather than level changes, has become standard for the literature. It allows for simple comparisons over time and across states without concern for inflation.

rather than the state-level. While unobserved (time-invariant) spatial heterogeneity at this more local level is unlikely to create endogeneity problems since minimum wage policy is generally determined at the state-level, removing it should improve estimates by reducing noise from persistent differences across localities in industrial structure and school quality. This specification is as follows:

$$y_{igst} = \beta_0 + \beta_1 l n(mw_{st}) + \beta_2 HSES_{igst} + \beta_3 l n(mw_{st}) \cdot HSES_{igst}$$
$$+ \gamma X_{igst} + \nu_t + \eta_g + \tau_s(t) + \epsilon_{igt}, \qquad (2)$$

where most elements remain the same as equation (1), but η_g is a county or PUMA-level fixed effect. I estimate this equation using OLS with standard errors clustered at the county-level for the ACS-C-Teens sample and at the PUMA-level for the ACS-P-Teens sample.

To rule out the possibility that sample selection in the CPS and ACS is driving the results for equation (1) and (2) (see Section 4.4), I estimate a similar equation using the SIPP-Teens sample. The equation varies slightly for (1) due to differences in the construction of this sample. First, since the SIPP-Teens sample is an aggregation of four panels, I include panel-by-state fixed effects (σ_{sl}) in order to account for any systematic differences between panels (indexed by l) for a given state, such as the samples selected. Second, the time period in the SIPP-Teens sample is four months rather than a year. Therefore, I include trimester fixed effects ($\zeta_{TRIMESTER(t)}$) in addition to year fixed effects ($\nu_{YR(t)}$), and I use four month periods as the unit of time for state-specific time trends ($\tau_s(t)$). These variations result in the following specification:

$$y_{islt} = \beta_0 + \beta_1 l n(mw_{st}) + \beta_2 HSES_{ist} + \beta_3 l n(mw_{st}) \cdot HSES_{ist}$$

$$+\gamma X_{islt} + v_{YR(t)} + \zeta_{TRIMESTER(t)} + \sigma_{sl} + \tau_s(t) + \epsilon_{islt}.$$
 (3)

I estimate this equation using OLS with standard errors clustered at the individual-level.

5.2 Internal Validity of Traditional Two-way Fixed Effects

There are a number of potential concerns regarding the internal validity of equations (1), (2), and (3). In this subsection, I discuss how I address these concerns by taking advantage of the large sample size (particularly within state-year-SES cells) and geographic granularity of the ACS.

First, the estimates of the effect of the minimum wage may simply be capturing elements of states' labor or education environments that pre-date minimum wage changes and are not accounted for by state-specific polynomial time trends. Endogenous policy change would be one example of this possibility, where state politicians adjust the minimum wage in response to changes in the state that are correlated with state dropout rates. I address this concern by testing whether minimum wage changes "affect" outcomes prior to their implementations. Specifically, I use the ACS to estimate an equation similar to (1), but with 3 year lags and leads of *changes* in the log minimum wage. Using log minimum wage changes rather than log minimum wage is necessitated by the high correlation from year-to-year in log minimum wage, muddying attempts to separately identify effects from different years. The specification is as follows:

$$y_{ist} = \beta_0 + \sum_{\tau \in [-3,3]} (\alpha_\tau \Delta l \, n \, m w_{s,t-\tau} + \delta_\tau \Delta l \, n \, m w_{s,t-\tau} \cdot HSES_{ist}) + \eta HSES_{ist}$$
$$+ \gamma X_{ist} + \nu_t + \eta_s + \tau_s(t) + \epsilon_{ist}, \tag{4}$$

where $\Delta l n m w_{s,t-\tau}$ is the year-to-year change in the log minimum wage τ years prior to year t ($\tau < 0$ refers to changes after year t). If my effect estimates in equations (1)-(3) are capturing pre-existing conditions rather than effects of minimum wage changes, then estimates of α_{τ} or $\alpha_{\tau} + \delta_{\tau}$ (or both) would be significant for $\tau < 0$.

Second, the estimate of the differential effect of the minimum wage on high and low-SES teens (β_1 vs. $\beta_1 + \beta_3$) may be driven by differential trends in the dropout rate of high and low-SES teens that are not accounted for by the common state-specific trends (and common year and state fixed effects) included in equations (1)-(3). To address this concern, I use the ACS to estimate equation (1) separately for high and low-SES teen subsamples.

Third, there may be time-varying heterogeneity in local labor markets that bias the effect estimates (β_1 and $\beta_1 + \beta_3$). For example, Allegretto et al. show that states experiencing greater increases in minimum wages differ systematically from other states in terms of the severity of economic downturns, the reduction of routine task intensive jobs, and the growth in upper-half wage inequality. To the extent that these types of differential trends across states are not sufficiently smooth to be captured by state-specific polynomial time trends, they will bias the estimates of equations (1), (2), and (3). The alternative approach discussed in the following section is designed to account for this concern.

5.3 Cross-Border Design

I address the possibility of endogeneity due to time-varying heterogeneity in local labor markets by employing an approach used by Allegretto et al (2013) to look at the employment effects of the minimum wage. This framework leverages variation in minimum wage within a commuting zone, that spans a state border, in a given year.²⁸ I apply this approach to the ACS-PCZ-Teen and ACS-C-Teen sample, which have sample sizes large enough to allow for analysis at geographic levels finer than state of residence. The ACS-PCZ-Teen sample includes all teens in the ACS, while ACS-C-Teen is limited to teens residing in counties with populations greater than 100,000 (the only counties identified in the public-use ACS).

The specification is largely the same as equation (1) except that a commuting zone by year fixed effect (ρ_{zt}) and a county or PUMA fixed effect (ϕ_g) replace the state and year fixed effects and state-specific time trends in (1):

$$y_{igzt} = \beta_0 + \beta_1 l n(mw_{gzt}) + \beta_2 HSES_{igzt} + \beta_3 l n(mw_{gzt}) \cdot HSES_{igzt}$$
$$+ \gamma X_{igzt} + \rho_{zt} + \phi_g + \epsilon_{igzt}.$$
(5)

I estimate this equation using OLS with standard errors clustered at the county-level for the ACS-C-Teens sample and at the PUMA-level for the ACS-PCZ-Teens sample.

This approach accounts for time-varying local labor market heterogeneity as well as timeinvariant differences between PUMAs. The cost of this improved internal validity is a reduction in external validity. The estimates are identified by comparing teenagers in the same commuting zone on either side of a state border, where the difference in minimum wages on either side of the border changes during 2005-2011. If teenagers in these border-spanning commuting zones are more or less responsive to minimum wage than typical American teenagers, the estimates of β_1 and β_3 will not represent the average effect nationwide of a minimum wage increase. Figure 2A shows the commuting zones that have minimum wage variation within commuting zone-

²⁸ I obtain commuting zones from Autor and Dorn (2013).

years during the period 2005-2011. Figure A1 gives an example of the identifying variation in one such commuting zone, Jacksonville, FL, which includes 7 counties: 5 in Florida and 2 in Georgia. For half of the years from 2005-2011, residents on the Florida side of the border faced a higher minimum wage than residents on the Georgia side. Figure 2B shows the counties observed in the ACS that are in commuting zones with minimum wage variation within commuting zone-years during the period 2005-2011.

5.4 Individual Fixed Effects

To complement my investigation of the impact of the minimum wage on teenagers' educational outcomes, I present evidence that hours worked may act as a mechanism through which this effect is operating. Specifically, I estimate the effect of a minimum wage increase on hours worked for teenagers that are still in school.²⁹ To identify this effect, I take advantage of the individual panel nature of the SIPP to use changes in outcomes for a given teenager before and after a minimum wage change.³⁰ This specification is similar to equation (1), but includes individual fixed effects (ψ_i) which account for time-invariant differences between teenagers (e.g. differences in fixed ability and family background). The specification is as follows,

$$y_{ist} = \beta_0 + \beta_1 l n(mw_{st}) + \beta_2 HSES_{ist} + \beta_3 l n(mw_{st}) \cdot HSES_{ist}$$

$$+\gamma X_{ist} + \nu_{YR(t)} + \zeta_{TRIMESTER(t)} + \psi_i + \epsilon_{ist}, \tag{6}$$

²⁹ I also estimate this equation using household earnings (excluding the teens own earnings) as the dependent variable.

³⁰ This approach lends itself to an analysis of household earnings and hours worked and not educational outcomes, as these outcomes vary from period to period whereas most students only dropout once and then remain unenrolled.

where y_{ist} is the number of hours worked (or an indicator a threshold of hours worked) for individual *i* in state *s* in reference period *t* (the SIPP-teen sample has 3 observations per year). I estimate this equation using OLS with standard errors clustered at the individual-level.

6. RESULTS

Below, I present the coefficient estimates for the empirical strategies discussed in Section 5. Subsection 6.1 discusses OLS estimates of the effect of minimum wage on the high school dropout likelihood using two-way fixed effect and cross-border approaches for multiple samples. Subsection 6.2 discusses additional robustness checks for the main results. Subsection 6.3 investigates possible mechanisms for the dropout effects observed in subsection 6.1, namely hours worked and earnings by others in the household.

6.1 High School Dropout Results

Table 3 presents OLS estimates for β_1 (effect of minimum wage change on low-SES teens) and β_3 (differential effect of minimum wage change on high-SES teens compared to low-SES teens) in equations (1)-(3), which leverage with-in state variation, and equation (5), which leverages with-in commuting zone by year variation. For each regression, Table 3 shows my preferred specification.³¹ The results are broadly consistent across samples and empirical approaches (two-way fixed effects and cross-border design). Raising the minimum wage significantly reduces the likelihood of dropping out among low-SES teens (β_1) across all samples

³¹ For the CPS-Teen sample I select a cubic trend (following Neumark et al, 2013a) and for the samples with shorter timeframes (i.e. SIPP-Teen, ACS-Teen, ACS-P-Teen, and ACS-C-Teen) I select a linear trend. While the SIPP-Teen timeframe is four-fifths of CPS-Teen, the select SIPP-Teen specification also includes state by panel fixed effects, leaving much less variation remaining to accommodate a higher polynomial state-specific time trend.

and specifications, but it has a much smaller or null impact on the likelihood of dropping out among high-SES teens ($\beta_1 + \beta_3$). A 10% increase in the minimum wage produces a 0.5-0.9 percentage point decrease in the dropout likelihood of low-SES teens (approximately 4-10% of this group's dropout rate) and a near zero impact on high-SES teens.^{32,33} The point estimates are relatively smaller in magnitude for the CPS-Teens and SIPP-Teens samples than for all of the ACS samples/approaches. These results suggest that President Obama's proposed federal minimum wage increase to \$10.10 (39%) would reduce the dropout likelihood of low-SES teens by 2-4 percentage points (16-40% of their dropout rate).³⁴ However, caution should be used when predicting the effects of an increase that is larger than much of the variation used to identify the estimates.

The results for equation (5) using the ACS-C-Teen sample (Table 3, column 7) differ somewhat from the rest. They show a larger negative impact on dropout likelihood among low-SES teens of 1.8 percentage points and among high-SES teens of 0.9 percentage points (for a 10% minimum wage increase). These results may differ from the rest because they are estimated using a much narrower group: teens in large counties (100,000+) that are part of commuting zones that cross state borders and have differential changes in minimum wages on either side of that border.

The consistency of the results in Table 3 across data samples and empirical approaches alleviate two major internal validity concerns. First, the SIPP-Teen coefficient estimates are

³² At the current federal minimum wage rate of \$7.25 per hour, a 10% increase represents a 73 cent per hour raise (roughly \$124 per month for a full-time worker).

³³ I find similar results using Pooled-Probit and Logit regressions models.

³⁴This estimate would apply only to states currently at the federal minimum. In the highest minimum wage state, Washington, the estimated low SES dropout effect of a \$10.10 federal minimum wage would be only 0.4-0.7 percentage points.

similar to the CPS-Teen and assorted ACS estimates, suggesting that the ACS and CPS estimates are not substantially biased by the selection of a sample with observed parental education (parental education is observed for nearly all teens in the SIPP). Second, the similarity of the cross-border estimates (columns 6-7) and two-way fixed effects estimates (columns 1-5) suggest that time-varying spatial heterogeneity is not driving the latter estimates.³⁵

6.2 Robustness Checks for High School Dropout Results

Table 4 shows the results for the same specifications as Table 3, but varies the definitions of high SES. Column 2 defines a teen as high SES if all of her observed parents have gone to college, coded as "some college" in the various datasets. Columns 3-6 define high SES by whether household income (excluding the teen) is above various percentile thresholds of household income distribution for teens in that year. The results are consistent with the estimates in Table 3. The magnitudes of the estimates for low SES teens are smaller for the "some college" definition than for the "high school diploma" definition. Similarly, the effect is smaller as the household income threshold for High SES increases. This is consistent with the variance the using an enrollment indicator as the dependent variable.

Tables A2-A5 provide robustness checks for the preferred specifications shown in Table 3 using the same definition of high/low SES.³⁶ They show results using both enrollment and high

³⁵The robustness across empirical approaches in the educational effects of the minimum wage, but not the employment effects (as seen in the contentious literature), suggests that state minimum wage policy decisions may be endogenous with respect to employment outcomes but not with respect to educational outcomes (e.g. changes are made in response to the overall unemployment rate).

³⁶ Table A9 provides an additional robustness check by regressing minimum wage at Age 17 on high school graduation by age 18 or 19 using the SIPP18 sample restricted to those who are enrolled at age 17 (and similar samples that span ages 17-19 and 18-19). The estimates show a 10% increase in the minimum wage increases the graduation likelihood for low-SES teens (still enrolled at age 17 or 18) by 1-3 percentage points. The effect on low-

school dropout dependent variables and including different specifications and/or time periods. Table A2 shows the OLS estimates of equation (1) using the CPS-Teen sample with various state-specific trends and a time period that excludes the Great Recession (1992-2007). The high school dropout results are largely consistent, though the standard errors are larger for the 1992-2007 subsample, leading to a loss of significance for the estimate of β_1 when using the statespecific cubic time trend. The enrollment results are significant and consistent across specifications and time periods. They are similar in magnitude to the high school dropout results, though slightly larger (and with an opposite sign).

Table A3 shows the OLS estimates of equations (1) and (2) using the ACS-Teens, ACS-P-Teens, and ACS-C-Teens samples for the two dependent variables, dropout and enrolled, and various state-specific time trends. The results are significant and similar in magnitude across all specifications. Table A4 shows the OLS estimates of equation (3) using the SIPP-Teens sample for various state-specific time trends and state or state by panel fixed effects. The high school dropout results are consistent across specifications with the exception of the state fixed effect and state-specific quadratic time trend specification, which provides an insignificant estimate of β_1 . The enrollment results are similar in magnitude, but slightly smaller than the high school dropout results. The estimates are not significantly different than zero in some of the specifications, though the statistically significant difference between the effects on high and low-SES teens remains. Table A5 shows the OLS estimates of equation (4) using the ACS-PCZ-Teens and ACS-C-Teens samples for the two dependent variables and two time periods (2005-2011 and 2000, 2005-2011). The results are significant and similar across time periods and dependent

SES teens is significantly more positive than high-SES teens, but the effects are only significantly different from zero for the age 18-19 group.

variables, though the β_1 estimates using the ACS-C-Teens are consistently larger than those using the ACS-PCZ-Teens.

Table A6 shows OLS separate estimates for equation (1) on high and low SES subsamples of the ACS-Teens sample. Column 2 shows a statistically significant 0.45 percentage point decrease in dropout likelihood for low-SES teens (and no change for high-SES teens) from a 10% increase in the minimum wage. These estimates are in line with those in Table 3, though the estimated effect on low-SES teens is slightly below the range of estimates in Table 3. This suggests two possibilities. The relatively small size of the low-SES sample may result in a larger role for measurement error in dampening the effect estimate. Alternatively, the existence of differential trends by SES status may mean that using common trends and fixed effects moderately exaggerates the true magnitude of effect on low-SES teens.

Finally, Figure 4 depicts OLS estimates of the falsification test defined in equation (4) using the ACS-Teens sample. It shows the estimated effect of a roughly 10% minimum wage increase in year t on the dropout likelihood of low-SES and high-SES teens in year $t + \tau$ for $\tau \in [-3,3]$.³⁷ The figure shows significant negative effects of minimum wage changes on low-SES teens in year t through t + 2, but not in years prior to the change ($\tau < 0$). There are no significant effects on high-SES teens. These results provide evidence that the negative dropout effects for low-SES teens in Table 3 are not driven by pre-existing conditions in states where the minimum wage increased.

6.3 Mechanism: Hours Worked

³⁷High and low-SES effect estimates are $0.1 \cdot \alpha_{\tau}$ and $0.1 \cdot (\alpha_{\tau} + \delta_{\tau})$, respectively This assumes a minimum wage change in year t given by $\Delta l n m w_{s,t} = 0.1$, which implies a 10.5% minimum wage change.
In the previous subsection, I show that raising the minimum wage improves educational outcomes for low-SES teens. In this subsection, I provide evidence to address the question of *how* the minimum wage affects these outcomes. The theoretical framework in Section 3 suggests that a likely mechanism for this effect is a reduction in hours worked that leads low-SES teens to allocate more time for school-related activities. I investigate the effect of the minimum wage changes on hours worked to assess the evidence for this story. Then, I discuss whether the hours results are more consistent with the minimum wage producing effects on low-SES teens through labor demand or labor supply (income effect) channels.

Table 5A presents OLS estimates of β_1 and β_3 from equation (5), an individual fixed effects specification with various measures of hours worked as the dependent variable. These regressions use the SIPP-Teen sample restricted to individuals who are consistently enrolled in high school. I make this restriction so that the hours changes that I observe are not driven by teenagers who left school and entered the workforce full-time. This would obscure the role of hours worked as a *mechanism* for the dropout effect.

Table 5A, Column 2 shows the estimates of β_1 and β_3 from equation (5) where the dependent variable is a continuous measure of the 4-month average of the hours worked by the teen (including zeroes for non-workers).³⁸ The estimate for low-SES teenagers is statistically significant and suggests that a 10% increase in the minimum wage reduces work by one hour on

³⁸ Table 5A, Column 1 shows similar estimates where dependent variable is the log of household income (excluding the teenager's own income) averaged over the trimester. The estimate of the minimum wage effect on household income is positive for low-SES teens and roughly zero for high-SES teens. Both estimates are insignificant with large standard errors. These estimates do not provide strong support for household income as a mechanism for the minimum wage effect on low-SES teen dropout likelihood. Additional estimates available upon request, show the main results in Table 3 for the SIPP-teens sample, but excluding teens with a parent who earns within 30% of the minimum wage. The similarity of these results to the original Table 3 results confirm that household income is not a mechanism for the dropout effects of the minimum wage.

average (17% of the mean for this group). For, high-SES teens the effect estimate is significant, but smaller in magnitude, 0.35 hours on average (4% of the mean for this group). Columns 3–8 show similar estimates of β_1 and β_3 , but replacing the dependent variable with an indicator equal to one if the 4-month average of hours worked is above a threshold $h \in \{0,5,10,15,20,25,30\}$. Essentially, these results provide estimates of the effect of minimum wage changes on the complementary cumulative distribution function (CCDF) of hours worked at various levels of hours. The CCDF is simply P(Hours Worked > h), or 1 - F(h), where $F(h) = P(Hours Worked \le h)$ is the cumulative distribution function of hours worked.

These estimates suggest that a 10% increase in the minimum wage affects high and low-SES teens at different points in the hours distribution. For low-SES teens, there is no significant effect at the employment margin (column 3), but large and significant negative effects at higher hours levels, $h \in \{15,20,25,30\}$ (columns 5-8). These estimates can be interpreted as a 2-3 percentage point drops in the likelihood that a low-SES teen works more than 15, 20, 25, or 30 hours. For high-SES teens, there is a significant negative effect of 1 percentage point at the employment margin (column 3) and at higher hours levels $h \in \{20,25\}$ (columns 6-7).

Figure 5A depicts the estimates from Table 5A graphically. The overall hours worked CCDFs for high and low-SES enrolled teens are the solid lines and the estimated CCDFs following a 10% minimum wage change are the dashed lines.³⁹ This figure shows that the major effects of the minimum wage on the hours worked of enrolled teens are concentrated among low-SES teens, particularly those in the upper tail of the hours distribution. This evidence is

³⁹ P(Hours > h) is the mean of the *Hours* > h indicator by SES status. The estimated P(Hours > h) after a 10% minimum wage increase is constructed as the mean (by SES status) of the *Hours* > h indicator plus the estimate of the minimum wage impact on the *Hours* > h indicator, i.e. $0.1 \cdot \beta_1$ or $0.1 \cdot (\beta_1 + \beta_3)$.

consistent with a minimum wage increase affecting low-SES teens' educational outcomes by causing a reduction in the time at work, which they replace with additional time on school-related activities. Several prior studies find significant associations between high work-intensity and dropping out further supporting this story.⁴⁰

The hours worked effect on enrolled low-SES teens observed in Table 5A and the Figure 5A could, in theory, be a result of either labor demand or labor supply (income) effects of a change in the minimum wage. To determine which of these possible channels is more likely, I attempt a simple test. If the changes in enrolled low-SES teen hours worked are due to changes in labor demand, then the pattern of these changes should be similar for similarly low-skilled workers who are not enrolled low-SES teens. If the changes are only observed among enrolled low-SES teens, then they likely reflect different labor supply responses between these teens and other similarly low-skilled (but not low-SES) workers.

I identify three groups that are potentially comparable in skill to enrolled low-SES teens (age 16-18): enrolled high-SES teens (age 16-18), non-enrolled high school dropouts who are slightly older (age 19-21), and non-enrolled individuals with only a high school diploma who are slightly older (age 19-21).⁴¹ Wages provide the most direct way to determine how the skill-levels of these groups compare. Figure 6A and 6B show the distribution of wages as a percent of the minimum wage for workers in each of these groups using the SIPP-teens sample (expanded to include 19-21 year olds).⁴² The wage distributions of enrolled high and low-SES teens are nearly identical, suggesting that these workers are indeed viewed as equally skilled by

⁴⁰ Goldschmidt and Wang (1999), Monahan, Lee, and Steinberg (2011), Warren and Lee (2003), and Warren and Cataldi (2003).

⁴¹ Non-enrolled refers to individuals who are never observed enrolled from age 19-21. High school dropout and high school diploma only groups include only those that remain at that level of education from age 19-21.

 $^{^{42}}$ Wages below 100% of the minimum wage are recoded as 100%.

employers. The slightly older, unenrolled workers appear much less comparable to enrolled low-SES workers.

Taken together, Table 5 and Figure 5 provide evidence that the pattern of the hours effect of an increase in the minimum wage differs substantially between enrolled low-SES teens and the comparison groups. In particular, the difference between enrolled low-SES and high-SES teens, two similarly skilled groups, is consistent with an income effect on low-SES teens, rather than a labor demand effect.⁴³

7. CONCLUSION

In this paper, I investigate the impact of the minimum wage on high school dropout decisions using three individual-level datasets and two distinct sources of variation. I leverage minimum wage variation both within states over time and between nearby localities on either side of a state border at a given point in time. I find that an increase in the minimum wage substantially lowers the likelihood of dropping out for low-SES teenagers, but has no effect on other teenagers. My estimates suggest that an increase in the minimum wage from \$7.25 to \$10.10 (which is 39%), as proposed by President Obama, would lead to a 2-4 percentage point decrease in the likelihood that a low-SES teen will drop out, roughly 16-40% of the rate for this group (in states that currently have the federal minimum wage). However, caution should be used in applying the effect estimates in this paper to a minimum wage change substantially larger than those used to identify the estimates.

⁴³ Table A8 shows that the industrial distribution of low and high-SES enrolled teen workers are also very similar.

Investigating the mechanism by which the minimum wage lowers the dropout rate, I show that increases in the minimum wage reduce the work hours for teenagers while they are enrolled in school. The effects are concentrated at the upper tail of the hours distribution and are more pronounced for low-SES teens. This effect of the minimum wage on hours is not found for high-SES teens, a comparably-skilled group (with a nearly identical wage distribution). These results are consistent with an income effect of the minimum wage on low-SES teens that leads them to reduce their work hours and invest more time in school.

These findings suggest that the current minimum wage literature's focus on teen employment neglects important aspects of the policy's broader implications for the later-life outcomes of teens.

REFERENCES

- Allegretto, Sylvia, Arindrajit Dube and Michael Reich. 2009. "Spatial Heterogeneity and Minimum Wages: Employment Estimates for Teens Using Cross-State Commuting Zones."
 Working Paper 181-09 Institute for Research on Labor and Employment, UC Berkeley.
- Allegretto, Sylvia, Arindrajit Dube and Michael Reich. 2011. "Do Minimum Wages Really Reduce Teen Employment? Accounting for Heterogeneity and Selectivity in State Panel Data." Industrial Relations 50(2): 205-40.
- Angrist, Joshua and Jorn-Steffen Pischke 2009. *Mostly Harmless Econometrics*. Princeton, NJ: Princeton University Press.
- Autor, David and David Dorn. "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market." American Economic Review, 103(5), 1553-1597, 2013.
- Belfield, Clive R., and Henry M. Levin, eds. *The price we pay: Economic and social consequences of inadequate education*. Brookings Institution Press, 2007.
- Campolieti, M., T. Fang and M. Gunderson. 2005. "How Minimum Wages affect Schooling-Employment Outcomes in Canada, 1993-1999." Journal of Labor Research 26(2):533-45.
- Card, David and Alan Krueger. 1994. "Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania." American Economic Review 84(4): 772-98.
- Card, David and Alan Krueger. 2000. "Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania: Reply." American Economic Review 90(5): 1397-1420.
- Chaplin, Duncan D., Mark D. Turner and Andreas D. Pape. 2003. "Minimum Wages and School Enrollment of Teenagers: A Look at the 1990's." Economics of Education Review 22(1):11-21.
- Dube, Arindrajit, T. William Lester and Michael Reich. 2010. "Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties." Review of Economics and Statistics 92(4): 945-64.
- Dube, Arindrajit, T. William Lester and Michael Reich. 2013. "Minimum Wage Shocks, Employment Flows and Labor Market Frictions." Working Paper 149-13. Institute for Research on Labor and Employment. UC Berkeley.
- Dube, Arindrajit, Suresh Naidu and Michael Reich 2007. "The Economic Effects of a Citywide Minimum Wage." Industrial and Labor Relations Review 60(4): 522-43.
- James J. Heckman and Paul A. LaFontaine, 2010. "The American High School Graduation Rate: Trends and Levels," The Review of Economics and Statistics 92(2): 244-262.
- King, Miriam, Steven Ruggles, J. Trent Alexander, Sarah Flood, Katie Genadek, Matthew B. Schroeder, Brandon Trampe, and Rebecca Vick. Integrated Public Use Microdata Series,

Current Population Survey: Version 3.0. [Machine-readable database]. Minneapolis: University of Minnesota, 2010

- Lochner, Lance. 2011. "Non-Production Benefits of Education: Crime, Health, and Good Citizenship" in E. Hanushek, S. Machin and L. Woessmann (eds.), *Handbook of the Economics of Education*, Vol. 4, Chapter 2, Amsterdam: Elsevier Science.
- Lochner, Lance and Alexander Monge-Naranjo. 2012. "Credit Constraints in Education." Annual Review of Economics 4:225-256.
- Meer, Jonathan and Jeremy West 2013 "Effects of the Minimum Wage on Employment Dynamics." NBER Working Paper 19262.
- Matilla, J. Peter. 1978. "Youth Labor Markets, Enrollments, and Minimum Wages." Industrial Relations Research Association Proceedings, 134-140.
- Matilla, J. Peter. 1982. "Determinants of Male School Enrollments: A Time-Series Analysis." Review of Economics and Statistics 64(2): 242-251.
- Neumark, David, and William Wascher. 1992. "Employment Effects of Minimum and Subminimum Wages: Panel Data on State Minimum Wage Laws." American Economic Review 85(2):244-49.
- Neumark, David, and William Wascher. 1995. "Minimum-Wage Effects on School and Work Transitions of Teenagers." American Economic Review 85(2):244-49.
- Neumark, David, and William Wascher. 1995. "Minimum Wage Effects on Employment and School Enrollment." Journal of Business and Economic Statistics 13(2):199-206.
- Neumark, David, and William Wascher. 2003. "Minimum Wages and Skill Acquisition: Another Look at Schooling Effects." Economics of Education Review 22(1):1-10.
- Neumark, David, and William Wascher. 2007. "Minimum Wages and Employment: A Review of Evidence from the New Minimum Wage Research." NBER Working Paper 12663.
- Neumark, David, and William Wascher. 2008. Minimum Wages. Cambridge, MA: MIT Press.
- Neumark, David, J. M. Ian Salas and William Wascher 2013a. "Revisiting the Minimum Wage and Employment Debate: Throwing out the Baby with the Bathwater?" Industrial and Labor Relations Review, forthcoming.
- Neumark, David, J. M. Ian Salas and William Wascher 2013b. "Revisiting the Minimum Wage and Employment Debate: Throwing out the Baby with the Bathwater?" NBER Working Paper 18681.
- Pacheco, Gail A., and Amy A. Cruickshank. 2007. "Minimum Wage Effects on Educational Enrollments in New Zealand." Economics of Education Review 26(5):574-87.
- Rothstein, Donna. 2007. "High School Employment and Youths' Academic Achievement." The Journal of Human Resources 42(1): 194-213.

- Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota, 2010.
- Ruhm, Christopher. 1997. "Is High School Employment Consumption or Investment?" Journal of Labor Economics 15(4): 735-776.
- Turner, Mark D. 1994. "The Effects of Part-Time Work on High School Students' Academic Achievement." University of Maryland at College Park. Unpublished.
- Turner, Mark D., and Berna Demiralp. 2001. "Do higher minimum wages harm minority and inner-city teens?." The Review of Black Political Economy 28(4): 95-116.
- United States Department of Commerce. Bureau of the Census. Survey of Income and Program Participation (SIPP) 2004 Panel. ICPSR04517-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2009-03-30.
- Warren, John Robert. 2005. "State-Level High School Completion Rates: Concepts, Measures, and Trends." Education Policy Analysis Archives 13(51):1-35. Available at: http://epaa.asu.edu/epaa/v13n51/.
- Warren, John Robert and Andrew Halpern-Manners. 2007. "Is the Glass Emptying or Filling Up? Reconciling Divergent Trends in High School Completion and Dropout." Educational Researcher 36(6):335-43.
- Warren, John Robert and Caitlin Hamrock. 2010. "The Effect of Minimum Wage Rates on High School Completion." Social Forces 88(3): 1379-1392.
- Warren, John Robert, Krista N. Jenkins and Rachael B. Kulick. 2006. "High School Exit Examinations and State-Level Completion and GED Rates, 1975 Through 2002." Educational Evaluation and Policy Analysis 28 (2): 131-152.

Figure 1: HS Dropout Rate by SES (CPS 1992-2012, ACS 2000-2011)











Figure 3A: State and Federal Minimum Wages (1992-2012)



Note: State minimum wages that exceed the federal minimum wage are grouped into \$0.25 bins. The size of the of a circle denotes the number of states in the given minimum wage bin (dot is one state).







Note: States are shaded according to the percent by which their minimum wage exceeds the federal minimum wage in each year.

Figure 4: Effect of 10% Increase in MW in Year t on HS Dropout in Years t-3 to t+3



Figure 5: Estimated Effect of 10% MW Increase on CCDF of Hours Worked [1-F(H)]







Figure A1: Example of Within CZ-Year Variation - Jacksonville, FL



50

 Table 1 - Overview of Select Data Samples

Sample Name	Years	Frequency	Geography Used	Source/Construction	Analysis Sample
CPS-Teen	1992-2012	Annual Cross-Section	State	IPUMS - March CPS	16-18 year olds with parent education observed
ACS-Teen	2000-2011	Annual Cross-Section	State	IPUMS - Census 2000 & ACS 2001-2011	16-18 year olds with parent education observed
ACS-P-Teen	2005-2011	Annual Cross-Section	State, PUMA	IPUMS - ACS 2005-2011	16-18 year olds with parent education observed
ACS-PCZ-Teen	2005-2011	Annual Cross-Section	State, PUMA, Commuting Zone	Probabilistic match of IPUMS - ACS 2005-2011 PUMAs to CZs (observations weighted by proportion of PUMA in each CZ)	16-18 year olds with parent education observed
ACS-C-Teen	2005-2011	Annual Cross-Section	State, County	IPUMS - ACS 2005-2011	16-18 year olds with parent education observed and county of residence observed.
SIPP-Teen	1996-2012	Individual Panel (3 Observations per Annum)	State	SIPP 1996, 2001, 2004, and 2008 panels appended	16-18 year olds with parent education observed who were in the initial SIPP sample (restricted to one observation per wave)

	CPS-Teen	ACS-Teen	ACS-P-Teen	ACS-PCZ-Teen ¹	ACS-C-Teen	SIPP-Teen
	(1)	(2)	(3)	(4)	(5)	(6)
Samula Description.						
Sumple Description:	1002 2012	2000 2011	2005 2011	2005 2011	2005 2011	1006 2012
Time Period	1992-2012	2000-2011	2005-2011	2005-2011	2005-2011	1996-2012
Frequency	Annual	Annual	Annual	Annual	Annual	Three per Annum
Observations	182,202	1,691,209	892,740	1,323,146	499,610	147,322
Unique Individuals	182,202	1,691,209	892,740	892,362	499,610	33,500
PUMAs			2.066	2.066		
Commuting Zones			,	741	190	
Counties				, , , ,	373	
Countres					575	
Full Sample Statistics (shown as	percentages of observa	ations with non-missing	values, unless specified):			
High SES	79.4	79.9	80.8	80.8	79.3	82.1
Percentage of Sample	8.4	13.9	14.4	14.4	13.8	1.0
Missing SES Measure						
HS Dropout (by SES)						
Low SES	11.9	10.6	9.5	9.5	8.5	12.6
High SES	3.4	3.1	2.9	2.9	2.7	3.4
Enrolled Subsample Statistics (sl	hown as percentages of	f observations with non-	missing values, unless sp	ecified):		
Hours Worked (by SES) ²	r U					
Low SES	3.9	6.6	5.6	5.6	5.0	5.9
High SES	4.6	8.9	7.9	7.9	7.2	8.1
C						
Worked Greater Than 20 Hour	s (by SES)					
Low SES	5.9	11.8	10.1	10.1	8.9	11.5
High SES	5.5	13.3	11.6	11.6	10.2	14.0
High SES	5.5	13.3	11.6	11.6	10.2	14.0

Table 2 - Summary Statistics for Select Data Samples

Table displays variable means. Population weights are used in all samples.

¹ ACS-PCZ-Teen includes a probabilistic match between PUMAs and commuting zones (see text for details). This means that individuals will be assigned probability weight for any

commuting zone-PUMA pairs that they could reside in (if a PUMA overlaps multiple commuting zones, creating more observations than individuals.

² Hours worked refers to the prior week in the CPS, the prior year's weekly average in the ACS, and the past month's weekly average in the SIPP.

			/				
	CPS-Teen (1992-2012) (1)	SIPP-Teen (1996-2012) (2)	ACS-Teen (2000-2011) (3)	ACS-P-Teen (2005-2011) (4)	ACS-C-Teen (2005-2011) (5)	ACS-PCZ-Teen (2005-2011) (6)	ACS-C-Teen (2005-2011) (7)
Ln(State Min Wage)	-0.049 ** (0.022)	-0.063 ** (0.030)	-0.094 *** (0.023)	-0.084 *** (0.011)	-0.076 *** (0.015)	-0.093 *** (0.027)	-0.179 *** (0.044)
Ln(State Min Wage) X High SES	0.062 *** (0.015)	0.071 **** (0.021)	0.102 *** (0.023)	0.083 *** (0.008)	0.091 *** (0.013)	0.086 *** (0.008)	0.092 *** (0.013)
Observations	165,829	158,525	1,455,883	764,209	430,298	1,132,279	430,421
Mean Dropout Rate Low SES High SES	12% 3%	13% 4%	11% 3%	9% 3%	8% 3%	9% 3%	8% 3%
Specification Fixed Effects	State Year	State x Panel Year Calendar Month	State Year	PUMA Year	County Year	PUMA CZ x Year	County CZ x Year
State-Specific Time Trend	Cubic	Linear	Linear	Linear	Linear		

Table 3 - Effect of Minimum Wage on H.S. Dropout Likelihood (By SES Level)

See text for descriptions of data samples CPS-Teen, SIPP-Teen, ACS-Teen, ACS-P-Teen, ACS-PCZ-Teen, ACS-C-Teen, and ACS-CCZ-Teen.

Regressions include indicators for age, race, sex, and whether the individual is above the state compulsory schooling age, as well as state unemployment rate (and county unemployment rate in column 5 and 7). Dropout indicator = 1 if not currently enrolled and have no H.S. diploma or GED. High SES indicator = 1 if all of teenager's parent/guardians have high school diploma (or equivalent).

Standard errors clustered at the state level for column 1 and 3, at the individual-level for column 2, at the PUMA-level for columns 4 and 6, and at the county-level for column 5 and 7, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4 - Effect of Minimum Wage on H.S. Dropout Likelihood (Various SES Measures)

	Definition of High SES							
	Parents	Education		Household Incom	e (Excluding Teen)			
	All have HS	All have some						
	Diploma	college	>20 percentile	>30 percentile	>40 percentile	>50 percentile		
	(1)	(2)	(3)	(4)	(5)	(6)		
CPS-Teen (1992-2012)								
[State and Year FE. State-specific Cubic Time Trend]								
Ln(State Min Wage)	-0.049 **	-0.012	-0.088 ***	-0.074 ***	-0.055 ***	-0.041 **		
	(0.022)	(0.015)	(0.020)	(0.019)	(0.019)	(0.019)		
Ln(State Min Wage) X High SES	0.062 ***	0.036 ***	0.123 ***	0.112 ***	0.100 ***	0.089 ***		
	(0.015)	(0.007)	(0.010)	(0.010)	(0.011)	(0.011)		
Observations	165,829	165,829	180,881	180,881	180,881	180,881		
SIPP-Teen (1996-2012)								
[State x Panel, Year, and calendar month FE, State-specifi	ic LinearTime Trend]							
Ln(State Min Wage)	-0.063 **	-0.035	-0.069 ***	-0.068 ***	-0.046 *	-0.042 *		
	(0.030)	(0.024)	(0.027)	(0.025)	(0.024)	(0.024)		
	0.071.444	0.050 +++	0.050 ++++	0.005 ***	0.0.00 ++++	0.0.00 ++++		
Ln(State Min Wage) X High SES	0.0/1 ***	0.050 ***	0.078 ***	0.085 ***	0.062 ***	0.062 ***		
	(0.021)	(0.010)	(0.015)	(0.012)	(0.011)	(0.010)		
Observations	158,525	158,525	158,476	158,476	158,476	158,476		
ACS-Teen (2000-2011)								
[State and Year FE, State-specific Linear Time Trend]								
Ln(State Min Wage)	-0.096 ***	-0.033 ***	-0.069 ***	-0.056 ***	-0.047 ***	-0.042 ***		
	(0.023)	(0.008)	(0.012)	(0.011)	(0.011)	(0.010)		
I p(State Min Wage) Y High SES	0 101 ***	0.041 ***	0.060 ***	0.061 ***	0.055 ***	0.054 ***		
En(State Will Wage) X High SES	(0.024)	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)		
			. ,			. ,		
Observations	1,455,883	1,455,883	1,578,768	1,578,768	1,578,768	1,578,768		
ACS-P-Teen (2005-2011)								
[Puma and Year FE, State-specific Linear Time Trend]								
Ln(State Min Wage)	-0.084 ***	-0.031 ***	-0.048 ***	-0.041 ***	-0.037 ***	-0.033 ***		
	(0.011)	(0.008)	(0.010)	(0.009)	(0.008)	(0.008)		
In (State Min Wege) V High SES	0.092 ***	0.024 ***	0.041 ***	0.020 ***	0.027 ***	0.026 ***		
En(State Will Wage) A fligh SES	(0.008)	(0.004)	(0.007)	(0.006)	(0.005)	(0.004)		
	(0.000)	(0.004)	(0.007)	(0.000)	(0.005)	(0.004)		
Observations	764,209	764,209	763,599	763,599	763,599	763,599		
ACS-PCZ-Teen (2005-2011)								
[CZ x Year and PUMA FE]								
Ln(State Min Wage)	-0.085 ***	-0.032	-0.048 **	-0.042 *	-0.039 *	-0.034		
	(0.024)	(0.023)	(0.024)	(0.023)	(0.023)	(0.023)		
Ln(State Min Wage) X High SFS	0.084 ***	0.034 ***	0.043 ***	0.040 ***	0.038 ***	0.037 ***		
En(State Ivini Wage) A Tingi SLS	(0.010)	(0.005)	(0.007)	(0.006)	(0.005)	(0.005)		
	(0.010)	(0.000)	(0.007)	(0.000)	(0.000)	(0.000)		
Observations	1,132,279	1,132,279	1,131,429	1,131,429	1,131,429	1,131,429		

See text for descriptions of data samples CPS-Teen, SIPP-Teen, ACS-P-Teen, ACS-PCZ-Teen, ACS-C-Teen, and ACS-CCZ-Teen.

Regressions replicate those in Table 3 for various definitions of SES.

Regressions replicate index in 1 and 5 for various definitions of SES. Household income percentile uses the yearly (tri-mester for SIPP-Teen) distribution of household income (excluding teen's own income) for all 16-18 year olds. Dropout indicator = 1 if not currently enrolled and have no H.S. diploma or GED. High SES indicator = 1 if all of teenager's parent/guardians have high school diploma (or equivalent). Standard errors clustered at the state level for CPS-Teen and ACS-Teen, at the individual-level for SIPP-Teen, at the PUMA-level for ACS-P-Teen and ACS-PCZ-Teen are in parentheses.

Table SA - Effect of Millin	num wage on nou	is worked and I	Iousenoiu Lain	ings (Enroneu re	J-10 Teal Olus)				
	Log HH Earnings				He	ours Threshold Indicate	ors		
	(Excluding Teen)	Hours	>0 Hours	>5 Hours	>10 Hours	>15 Hours	>20 Hours	>25 Hours	>30 Hours
VARIABLES	(1)	(2)	(3)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(State Min Wage)	0.246	-10.063 ***	-0.047	-0.158	-0.151	-0.276 **	-0.259 ***	-0.268 ***	-0.217 ***
	(0.194)	(3.212)	(0.124)	(0.124)	(0.117)	(0.113)	(0.098)	(0.086)	(0.070)
Ln(State Min Wage) X High SES	-0.231 (0.203)	6.482 * (3.406)	-0.074 (0.130)	0.087 (0.130)	0.092 (0.122)	0.184 (0.118)	0.139 (0.103)	0.160 * (0.089)	0.163 ** (0.073)

109,914

30,045

25%

37%

109,914

30,045

21%

30%

109,914

30,045

17%

22%

109,914

30,045

11%

14%

109,914

30,045

8%

10%

Table 54 - Effect of Minimum Wage on Hours Worked and Household Farnings (Enrolled 16-18 Vear Olds)

7.94 Regression is on SIPP-Teen sample restricted to individuals aged 16-18 that are always observed enrolled. See text for descriptions of SIPP-Teen data sample.

5.63

109,914

30,045

109,914

30,045

27%

41%

Observations are at the individual by trimester level.

Regressions include individual, trimester, year, and age fixed effects, as well as log of state unemployment rate and an indicator for age>compulsory schooling age.

Hours refers to total hours worked per week in a month, averaged over the four months of a given trimester.

101,855

28,428

7.74

8.39

Standard errors clustered at the individual-level are in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Observations

Mean Low SES

Unique Individuals

High SES

109,914

30,045

5%

6%

Table 5B - Effect of Minimum Wage on Hours Worked and Household Earnings (Non-enrolled 19-21 Year Olds)

				He	ours Threshold Indicat	ors		
	Hours	>0 Hours	>5 Hours	>10 Hours	>15 Hours	>20 Hours	>25 Hours	>30 Hours
VARIABLES	(1)	(2)	(2)	(3)	(4)	(5)	(6)	(7)
Subsample: Non-enrolled HS Drop	oout (Age 19-21)							
Ln(State Min Wage)	-8.519	-0.114	-0.129	-0.030	0.010	-0.097	-0.113	-0.150
	(8.962)	(0.208)	(0.208)	(0.194)	(0.194)	(0.197)	(0.189)	(0.199)
Observations	9,291	9,291	9,291	9,291	9,291	9,291	9,291	9,291
Unique Individuals	3,191	3,191	3,191	3,191	3,191	3,191	3,191	3,191
Mean	23.11	61%	60%	57%	55%	51%	48%	43%
Subsample: Non-enrolled HS Diple	oma Only (Age 19-21)							
Ln(State Min Wage)	2.773	-0.011	0.019	0.040	0.063	0.091	0.054	0.055
	(6.485)	(0.128)	(0.130)	(0.133)	(0.136)	(0.143)	(0.144)	(0.141)
Observations	19,681	19,681	19,681	19,681	19,681	19,681	19,681	19,681
Unique Individuals	6,947	6,947	6,947	6,947	6,947	6,947	6,947	6,947
Mean	29.47	75%	74%	72%	70%	66%	62%	57%

Regression is uses 1996, 2001, 2004, and 2008 SIPP panels restricting to individuals aged 19-21 that are never observed enrolled. See text for descriptions of SIPP-Teen data sample.

Observations are at the individual by trimester level.

Regressions include individual, trimester, year, and age fixed effects, as well as log of state unemployment rate.

Hours refers to total hours worked per week in a month, averaged over the four months of a given trimester.

Standard errors clustered at the individual-level are in parentheses.

Table A1 - Naïve Regression of Age 26 Outcomes on Age 18 Outcomes

	Log Income	e (Age 26)	Log Family Inc	ome (Age 26)	Employed	(Age 26)	Arrested (A	ge 19-26)
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HS Dropout (Age 18)	-0.514 *** (0.085)	-0.439 *** (0.100)	-0.598 *** (0.072)	-0.576 *** (0.084)	-0.189 *** (0.025)	-0.163 *** (0.029)	0.223 *** (0.023)	0.207 *** (0.027)
Employed (Age 18)	0.218 ***	0.151 ***	0.255 ***	0.178 ***	0.094 ***	0.066 ***	-0.014	0.008
	(0.035)	(0.037)	(0.035)	(0.036)	(0.012)	(0.013)	(0.011)	(0.012)
HS Dropout (Age 18) x Employed (Age 18)	-0.065 (0.107)	0.072 (0.120)	0.030 (0.091)	0.169 (0.103)	0.026 (0.031)	0.028 (0.036)	0.002 (0.030)	-0.012 (0.035)
Adjusted R-Squared	0.09	0.11	0.11	0.14	0.07	0.08	0.10	0.10
Observations	4,934	4,101	5,788	4,748	6,798	5,545	7,677	6,219
Specification Cohort Fixed Effect Controlling for Cognitive Test Score	Х	X X	х	X X	х	X X	х	X X

Regressions using data from the National Longitudinal Survey of Youth 1997 (NLSY97). Each column refers to different regression, all of which include controls for race and sex. Robust standard errors are in parentheses. HS Dropout (Age 18) indicator = 1 if individual has no H.S. diploma and is not enrolled at age 18.

The cognitive test score measure is the Armed Services Vocational Aptitude Battery (ASVAB), standardized by the age at which it was taken.

Table A2 - Education Outcome Robustness Checks (CPS)

		Dependent Variable = HS Dropout							Dependent Variable = Enrolled							
		1992	-2012			1992	2-2007			1992	-2012			1992	-2007	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Ln(State Min Wage)	-0.066 *** (0.019)	-0.061 *** (0.021)	-0.063 *** (0.019)	-0.049 ** (0.022)	-0.053 ** (0.026)	-0.058 * (0.030)	-0.023 (0.032)	-0.040 (0.035)	0.103 *** (0.027)	0.103 *** (0.027)	0.084 *** (0.020)	0.074 *** (0.026)	0.094 *** (0.032)	0.095 *** (0.032)	0.047 (0.033)	0.090 ** (0.038)
Ln(State Min Wage) X High SES	0.062 *** (0.015)	0.062 *** (0.015)	0.062 *** (0.015)	0.062 *** (0.015)	0.057 ** (0.024)	0.055 ** (0.025)	0.055 ** (0.025)	0.055 ** (0.025)	-0.066 *** (0.016)	-0.066 *** (0.016)	-0.066 *** (0.016)	-0.066 *** (0.016)	-0.071 *** (0.020)	-0.070 *** (0.020)	-0.068 *** (0.020)	-0.069 *** (0.020)
Observations	165,829	165,829	165,829	165,829	118,064	118,064	118,064	118,064	165,785	165,785	165,785	165,785	118,033	118,033	118,033	118,033
Mean Dropout Rate																
Low SES	12%	12%	12%	12%	13%	13%	13%	13%	84%	84%	84%	84%	83%	83%	83%	83%
High SES	3%	3%	3%	3%	4%	4%	4%	4%	93%	93%	93%	93%	93%	93%	93%	93%
Specification:																
Year FE	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х
State FE	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х
State-Specific Time Trend		Linear	Quadratic	Cubic		Linear	Quadratic	Cubic		Linear	Quadratic	Cubic		Linear	Quadratic	Cubic

See text for descriptions of CPS-Teen data sample.

Regressions include indicators for age, race, sex, and whether the individual is above the state compulsory schooling age, as well as state unemployment rate.

Dropout indicator = 1 if not currently enrolled and have no H.S. diploma or GED. High SES indicator = 1 if all of teenager's parent/guardians have high school diploma (or equivalent).

Standard errors clustered at the state-level are in parentheses.

Table A3- Education Outcome Robustness Checks (ACS - Twoway Fixed Effects)

		Dependent Variable = HS Dropout							Dependent Variable = Enrolled					
		2000-2011			200	5-2011			2000-2011			2005	-2011	
VARIABLES	(1)	(2)	(3)	(3)	(4)	(5)	(6)	(7)	(1)	(8)	(9)	(10)	(11)	(12)
Ln(State Min Wage)	-0.094 *** (0.023)	-0.096 *** (0.023)	-0.101 *** (0.022)	-0.080 *** (0.010)	-0.084 *** (0.011)	-0.082 *** (0.014)	0.078 *** (0.021)	0.078 *** (0.021)	0.080 *** (0.022)	0.082 *** (0.021)	0.084 *** (0.011)	0.082 *** (0.013)	0.090 *** (0.016)	0.080 *** (0.017)
Ln(State Min Wage) X High SES	0.102 *** (0.024)	0.101 *** (0.024)	0.102 *** (0.024)	0.083 *** (0.008)	0.083 *** (0.008)	0.089 *** (0.013)	-0.088 *** (0.022)	-0.088 *** (0.022)	-0.088 *** (0.022)	-0.088 *** (0.022)	-0.086 *** (0.009)	-0.086 *** (0.009)	-0.086 *** (0.013)	-0.087 *** (0.013)
Observations PUMAs	1,455,883	1,455,883	1,455,883	764,209 2,066	764,209 2,066	430,298	430,298	1,455,883	1,455,883	1,455,883	764,209 2,066	764,209 2,066	430,298	430,298
Counties						374	374						374	374
Specification:														
Year FE	Х	Х	х	Х	х	х	Х	Х	х	Х	х	х	х	х
State FE	х	х	х					Х	х	х				
PUMA FE				Х	х						х	Х		
County FE						Х	х						х	х
State-Specific Time Trend		Linear	Quadratic		Linear		Linear		Linear	Quadratic		Linear		Linear

See text for descriptions of ACS-Teen, ACS-P-Teen,, and ACS-C-Teen data samples.

Regressions include indicators for age, race, sex, and whether the individual is above the state compulsory schooling age, as well as state unemployment rate.

Dropout indicator = 1 if not currently enrolled and have no H.S. diploma or GED. High SES indicator = 1 if all of teenager's parent/guardians have high school diploma (or equivalent).

Standard errors clustered at the state-level are in parentheses.

Table A4 - Education Outcome Robustness Checks (SIPP)

			Dependent Varia	ble = HS Dropout	:		Dependent Variable = Enrolled					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	0.052 *	0.0(2.**	0.062 **	0.057 **	0.050 **	0.020	0.020	0.050	0.025	0.051	0.054 *	0.028
Ln(State Min wage)	-0.053 *	-0.063 **	-0.062 **	-0.05/ **	-0.058 **	-0.039	0.039	0.050	0.035	0.051	0.054 *	0.028
	(0.030)	(0.030)	(0.030)	(0.028)	(0.028)	(0.030)	(0.036)	(0.036)	(0.037)	(0.031)	(0.032)	(0.036)
Ln(State Min Wage) X High SES	0.072 ***	0.071 ***	0.071 ***	0.077 ***	0.074 ***	0.073 ***	-0.069 ***	-0.069 ***	-0.068 ***	-0.075 ***	-0.072 ***	-0.069 ***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Observations	158,525	158,525	158,525	158,525	158,525	158,525	158,525	158,525	158,525	158,525	158,525	158,525
Mean Dropout Rate												
Low SES	13%	13%	13%	13%	13%	13%	81%	81%	81%	81%	81%	81%
High SES	4%	4%	4%	4%	4%	4%	91%	91%	91%	91%	91%	91%
Specification:												
Year FE	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х
State FE				Х	Х	Х				Х	Х	Х
State x Panel FE	х	х	Х				Х	Х	Х			
State-Specific Time Trend		Linear	Quadratic		Linear	Quadratic		Linear	Quadratic		Linear	Quadratic

See text for descriptions of SIPP-Teen data sample.

Regressions include indicators for trimester, age, race, sex, and whether the individual is above the state compulsory schooling age, as well as state unemployment rate.

Dropout indicator = 1 if not currently enrolled and have no H.S. diploma or GED. High SES indicator = 1 if all of teenager's parent/guardians have high school diploma (or equivalent).

Standard errors clustered at the state-level are in parentheses.

		Dependent Varia	ble = HS Dropout		Dependent Variable = Enrolled					
	2005-	2011	2000, 20	05-2011	2005-	2011	2000, 20	05-2011		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
Ln(State Min Wage)	-0.093 *** (0.027)	-0.179 *** (0.044)	-0.089 *** (0.020)	-0.182 *** (0.041)	0.117 *** (0.033)	0.159 *** (0.059)	0.088 *** (0.024)	0.139 *** (0.053)		
Ln(State Min Wage) X High SES	0.086 *** (0.008)	0.092 *** (0.013)	0.089 *** (0.006)	0.097 *** (0.010)	-0.092 *** (0.009)	-0.088 *** (0.013)	-0.080 *** (0.007)	-0.081 *** (0.009)		
Observations Commuting Zones PUMAs Counties	1,132,279 741 2,066	430,421 190 373	1,953,865 741 2,066	705,546 190 373	1,132,279 741 2,066	430,421 190 373	1,953,865 741 2,066	705,546 190 373		
Specification: Commuting Zone x Year PUMA FE County FE	X X	X X	X X	X X	X X	X X	X X	X X		

Table A5 - Education Outcome Robustness Checks (ACS - Cross Border)

See text for descriptions of ACS-PCZ-Teen and ACS-C-Teen data samples.

Regressions include indicators for age, race, sex, and whether the individual is above the state compulsory schooling age, as well as state unemployment rate.

Dropout indicator = 1 if not currently enrolled and have no H.S. diploma or GED. High SES indicator = 1 if all of teenager's parent/guardians have high school diploma (or equivalent). Standard errors clustered at the PUMA-level are in parentheses.

		Subsample	Low SES		Subsample: High SES					
	Dependent Varia	able = HS Dropout	Dependent V	ariable = Enrolled	Dependent Va	riable = HS Dropout	Dependent V	ariable = Enrolled		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
Ln(State Min Wage)	-0.033 *	-0.045 **	0.029	0.035	-0.006	-0.007	0.000	0.002		
_	(0.019)	(0.018)	(0.023)	(0.022)	(0.005)	(0.005)	(0.007)	(0.009)		
Observations	292,382	292,382	292,382	292,382	1,163,501	1,163,501	1,163,501	1,163,501		
Mean	10%	10%	85%	85%	3%	3%	94%	94%		
Specification:										
Year FE	Х	Х	Х	Х	Х	Х	Х	Х		
State FE	Х	Х	Х	Х	Х	Х	Х	Х		
State-Specific Time Trend		Linear		Linear		Linear		Linear		

Table A6- Split Sample Robustness Checks (ACS - Twoway Fixed Effects)

Data from ACS-Teen sample (2000-2011).

Regressions include indicators for age, race, sex, and whether the individual is above the state compulsory schooling age, as well as state unemployment rate.

Dropout indicator = 1 if not currently enrolled and have no H.S. diploma or GED. High SES indicator = 1 if all of teenager's parent/guardians have high school diploma (or equivalent).

Standard errors clustered at the state-level are in parentheses.

Table A7 - Robustness Checks for SES Measure - Enrolled

	Definition of High SES							
	Parents Education Household Income (Excluding Teen)							
	All have HS	All have some						
	Diploma	college	>20 percentile	>30 percentile	>40 percentile	>50 percentile		
	(1)	(2)	(3)	(4)	(5)	(6)		
CPS Toon (1992 2012)								
[State and Year FE_State specific Cubic Time Trend]								
Ln(State Min Wage)	0.074 ***	0.036 *	0.127 ***	0.114 ***	0.091 ***	0.078 ***		
	(0.026)	(0.021)	(0.020)	(0.021)	(0.020)	(0.021)		
Ln(State Min Wage) X High SES	-0.066 ***	-0.040 ***	-0.131 ***	-0.118 ***	-0.100 ***	-0.094 ***		
	(0.016)	(0.008)	(0.017)	(0.015)	(0.016)	(0.015)		
Observations	165,785	165,785	180,830	180,830	180,830	180,830		
SIPP-Teen (1996-2012)								
[State x Panel, Year, and calendar month FE, State-specifi	ic LinearTime Trend]							
Ln(State Min Wage)	0.050	0.029	0.046	0.052	0.035	0.032		
	(0.036)	(0.031)	(0.034)	(0.032)	(0.031)	(0.031)		
L n(State Min Wage) Y High SES	0.060 ***	0.060 ***	0.055 ***	0.070 ***	0.050 ***	0.047 ***		
En(State Will Wage) A Tigh SES	(0.023)	(0.013)	(0.018)	(0.015)	(0.013)	(0.012)		
	(0.025)	(0.015)	(0.010)	(0.015)	(0.015)	(0.012)		
Observations	158,525	158,525	158,476	158,476	158,476	158,476		
ACS-Teen (2000-2011)								
[State and Year FE, State-specific Linear Time Trend]								
Ln(State Min Wage)	0.080 ***	0.020 *	0.071 ***	0.053 ***	0.045 ***	0.038 ***		
	(0.022)	(0.011)	(0.017)	(0.016)	(0.015)	(0.014)		
In (State Min Wage) X High SES	-0.088 ***	-0.025 ***	-0.077 ***	-0.063 ***	-0.058 ***	-0.055 ***		
En(State Will Wage) X High SES	(0.022)	(0.007)	(0.012)	(0.010)	(0.011)	(0.009)		
				((,		
Observations	1,455,883	1,455,883	1,578,768	1,578,768	1,578,768	1,578,768		
ACS-P-Teen (2005-2011)								
[Puma and Year FE, State-specific Linear Time Trend]								
Ln(State Min Wage)	0.083 ***	0.029 ***	0.059 ***	0.043 ***	0.039 ***	0.033 ***		
	(0.013)	(0.010)	(0.013)	(0.011)	(0.011)	(0.010)		
Ln(State Min Wage) X High SES	-0.086 ***	-0.037 ***	-0.059 ***	-0.046 ***	-0.047 ***	-0.044 ***		
	(0.009)	(0.006)	(0.009)	(0.007)	(0.006)	(0.006)		
Observations	764,209	764,209	763,599	763,599	763,599	763,599		
ACS BC7 Toon (2005 2011)								
[CZ x Year and PUMA FE]								
Ln(State Min Wage)	0.104 ***	0.051 *	0.079 ***	0.064 **	0.062 **	0.055 **		
	(0.028)	(0.027)	(0.028)	(0.028)	(0.027)	(0.027)		
Ln(State Min Wage) X High SES	-0.087 ***	-0.038 ***	-0.060 ***	-0.047 ***	-0.048 ***	-0.045 ***		
	(0.010)	(0.006)	(0.009)	(0.008)	(0.007)	(0.006)		
Observations	1 122 270	1 122 270	1 121 420	1 121 420	1 121 420	1 121 420		
Observations	1,132,279	1,132,279	1,131,429	1,131,429	1,131,429	1,131,429		

See text for descriptions of data samples CPS-Teen, SIPP-Teen, ACS-P-Teen, ACS-PCZ-Teen, ACS-C-Teen, and ACS-CCZ-Teen.

Regressions replicate those in Table 3 for various definitions of SES.

regressions replicate tnose in 1 and 5 for various definitions of SES. Household income percentile uses the yearly (tri-mester of SIPP-Teen) distribution of household income (excluding teen's own income) for all 16-18 year olds. Dropout indicator = 1 if not currently enrolled and have no H.S. diploma or GED. High SES indicator = 1 if all of teenager's parent/guardians have high school diploma (or equivalent). Standard errors clustered at the state level for CPS-Teen and ACS-Teen, at the individual-level for SIPP-Teen, at the PUMA-level for ACS-P-Teen and ACS-PCZ-Teen are in parentheses. *** p<0.01, ** p<0.01, ** p<0.01

	Low SES	High SES
	(1)	(2)
Construction	2.55%	2.40%
Ed/Health/Social Services	10.08%	9.95%
Food Services, Accommodations, Recreation	42.33%	42.91%
Manufacturing	2.68%	2.21%
Mining & Agriculture	3.88%	3.58%
Trade (Retail & Wholesale)	24.76%	23.10%
Professional/Science/Management/Administrative/Waste Services	4.44%	4.92%
Other Services	4.05%	5.81%

Table A8 - Industrial Distribution of Working Teens by SES

Percent of working and enrolled teens (age 16-18) in each industry by SES

Data from ACS 2008-2011.

High SES indicator = 1 if all of teenager's parent/guardians have high school diploma (or equivalent).

	Dependent Variable = HS Graduation by Age 18				Dependent Variable = HS Graduation by Age 19			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample of Enrolled 17 Year Olds								
Ln(State Min Wage at Age 17)	0.114	0.101	0.111	0.068	0.100	0.123	0.126	0.154
	(0.126)	(0.127)	(0.130)	(0.138)	(0.108)	(0.110)	(0.112)	(0.122)
Ln(State Min Wage at Age 17) X High SES	-0.124 ** (0.058)	-0.118 ** (0.058)	-0.119 ** (0.058)	-0.119 ** (0.059)	-0.221 *** (0.063)	-0.224 *** (0.063)	-0.219 *** (0.064)	-0.219 *** (0.064)
Observations Graduation Rate:	15,241	15,241	15,241	15,241	10,026	10,026	10,026	10,026
Low SES	60%	60%	60%	60%	78%	78%	78%	78%
High SES	78%	78%	78%	78%	93%	93%	93%	93%
Sample of Enrolled 18 Year Olds								
Ln(State Min Wage at Age 18)					0.265 *** (0.087)	0.264 *** (0.096)	0.244 *** (0.095)	0.287 *** (0.105)
Ln(State Min Wage at Age 18) X High SES					-0.155 ** (0.061)	-0.156 ** (0.061)	-0.154 ** (0.061)	-0.152 ** (0.061)
Observations					9,365	9,365	9,365	9,365
Graduation Rate:								
Low SES					85%	85%	85%	85%
High SES					95%	95%	95%	95%
Specification:								
Birth Year FE	Х	Х	Х	Х	Х	Х	Х	Х
State x Panel FE	Х	Х	Х	Х			Х	Х
State-Specific Time Trend		Linear	Quadratic	Cubic		Linear	Quadratic	Cubic

Table A9 - Education Outcome Robustness Checks - Regressions of Minimum Wage at Age 17-18 on HS Graduation by Age 18-19 (SIPP)

SIPP Panels are collapsed from the individual-trimester level to the individual level. Observations have information from age 17, 18, and/or 19.

Regressions include indicators for race and sex, as well as state unemployment rate at age 17 (top panel) or age 18 (bottom panel).

HS Graduation indicator = 1 if observed having received a H.S. diploma or GED by the given age (18 or 19). High SES indicator = 1 if all of teenager's parent/guardians have high school diploma (or equivalent).

The maximum state minimum wage at the given age is used.

Standard errors clustered at the state by panel level are in parentheses.

CHAPTER 2

The Long-Run Effects of Universal Pre-K on Criminal Activity

1. INTRODUCTION

In the last decade, campaigns to provide high quality prekindergarten (Pre-K) to four year olds have achieved remarkable success. The fraction of four year olds attending Pre-K reached 28% in 2013, double what it was in 2002 (Barnett et al, 2013). The Wall Street Journal called this increase "one of the most significant expansions in public education in the 90 years since World War I" (Solomon, 2007). Supporters of increasing funding for early childhood education make the case that major economic and social problems, such as crime and teen pregnancy, can be traced to low cognitive and socio-emotional skill levels. Differences in these skill levels in advantaged and disadvantaged children appear early in childhood, but these differences can be alleviated by intervening early, leading to substantial reductions in negative later-life outcomes and therefore very high benefit-cost ratios and rates of return (Heckman et al, 2010a). These arguments have won the support of many policymakers at the state and federal level, including President Obama. In his 2013 State of the Union address, he announced the Preschool for All Initiative to expand high-quality preschool access to every child in America by allocating federal funds to finance states' provision of Pre-K. The plan's goal is to reduce the achievement gap by helping to "level the playing field" for children from low income families.¹

¹ "Fact Sheet President Obama's Plan for Early Education for All Americans," https://www.whitehouse.gov/the-press-office/2013/02/13/fact-sheet-president-obama-s-plan-early-education-all-americans (February, 12, 2013).

Seven in ten Americans support expanding preschool programs using federal funds, according to an August 2014 Gallup poll.² Despite this widespread public support, important disagreements arise around the scope of early childhood educational interventions: Should interventions be targeted only at the most at-risk children or should they be universal? The commonly cited evidence showing high long-run returns to early childhood education programs is largely limited to preschool programs targeted at especially at-risk children, and often to programs that are highly resource-intensive. This evidence includes quasi-experimental evaluations of Head Start (Deming, 2009; Ludwig and Miller, 2007; Garces et al, 2002) and experimental evaluations of HighScope Perry preschool (Heckman et al, 2010a) and the Abcedarian program (Campbell et al, 2012). Universal Pre-K differs dramatically from these programs because it is available to all children and typically involves lower levels of funding per child, especially when compared to Abcedarian and Perry preschool.

Evidence regarding the effectiveness of less resource-intensive but universally available preschool programs has thus far been necessarily restricted to short-run impacts. This is largely because the first of these programs were only implemented relatively recently, in the mid to late 1990s.³ These evaluations look at outcomes in middle school or earlier and tend to find mixed results. Some find evidence of substantial "fade out" in early test score effects, though this does not rule out large long-run effects. Similar fade out in cognitive effects have also been observed in a number of other early childhood interventions that nonetheless found large long-run effects (e.g. Heckman et al, 2013; Deming, 2009). The lack of evidence on the long-run effectiveness of universal Pre-K programs is remarkable given growing political momentum behind these

² Jones, Jefferey. "In U.S., 70% Favor Federal Funds to Expand Pre-K Education," (September 8, 2014). ³ See section 2 for a review of this literature.

programs. The result is a critical blind spot for state policymakers deciding how to best allocate early childhood education funding.

I estimate the impact of Oklahoma's universal Pre-K program (UPK), introduced in 1998, on an important later-life outcome: teenage criminal activity. I assemble data on criminal charges in the state of Oklahoma and use a regression discontinuity design which leverages the birthdate cutoff for UPK eligibility in the program's first year of implementation. This approach yields estimates of the effect of UPK availability (or the intent-to-treat effect of state Pre-K) compared to the prior mix of preschool services. I compare the effect of UPK availability differentially by race, as black children in Oklahoma are four times more likely to be charged at age 18 or 19 than are white children (31% vs. 7%). I find a significant negative impact of UPK availability on the likelihood that a black child is later charged with a misdemeanor or felony at age 18 or 19 (7 and 5 percentage points, respectively), but no impact on the likelihood of later charges for white children. This suggests that, like more targeted preschool programs, UPK has a large and important impact on a measure frequently associated with socio-emotional skills, but the impact on this measure is concentrated within a higher-risk population.

The rest of this paper is organized as follows. In Section 2, I review the existing evidence on the long-run effects of preschool. In Section 3, I discuss the details of the universal Pre-K program in Oklahoma. In Section 4, I describe my data sources. In Section 5, I present my empirical strategy. In Section 6, I discuss my empirical results. In Section 7, I use my results to approximate an alternative estimand. In Section 8, I conclude and discuss future work.

2. EXISTING EVIDENCE ON LONG-RUN PRESCHOOL EFFECTS

In this section I review the existing literature on the long-run impacts of early childhood education interventions. This literature can be broadly categorized as evaluating three types of preschool programs: small pilot programs targeted to the most disadvantaged children, larger scale targeted programs, and state-run universal Pre-K programs. Evidence of potentially large long-run effects of preschool programs comes primarily from studies of the targeted programs. Studies of universal Pre-K programs involve shorter timeframes and obtain more mixed results.

2.1 Targeted Preschool Pilot Programs

The most frequently cited evidence of the long-run impacts of early education interventions comes from the HighScope Perry Preschool Program. This was a program for three and four year olds conducted in Ypsilanti, Michigan during the 1960s, where children received 2.5 hours of preschool each school day and weekly home visits from teachers at a cost of \$20,854 per student in inflation-adjusted 2015 dollars (Barnett, 1996). Only children judged to be disadvantaged by family socioeconomic status and IQ scores were eligible to participate, and the eligible 123 children were randomly assigned to the program or the control group.⁴ Heckman et al (2010a) uses data that includes periodic follow-up interviews to age 40 to estimate that Perry Preschool produced an annual social rate of return of 7-10%. This large estimated return was a product of the program's beneficial effects on criminal, welfare, and earnings outcomes, the bulk of which were mediated by persistent changes in personality skills (i.e. reduced externalizing behavior) rather than changes in cognitive skills or academic motivation (Heckman et al, 2013).

⁴ Heckman et al (2010) adjusts their effect estimates for the potentially problematic reassignment of treatment and controls after the initial random assignment.
Perry Preschool's impact on crime played a central role in generating its large social returns, accounting for roughly 40-65% of the benefits of the program (Heckman et al, 2010a). Among men, the program caused an average reduction in the number of arrests by age 40 of 4.2, and a 13 percentage point reduction in the likelihood of arrest by age 40. However, it is uncertain whether similarly large impacts of the program could be expected in other contexts, given the extent to which the Perry sample was selected on disadvantage. For example, 37 of 39 (95%) of men assigned to the control group in the Perry study were arrested by age 40, and averaged 12.4 arrests.

The Abcedarian Project was similar to Perry Preschool in its scale (111 children), its focus on disadvantaged children, and its use of random assignment, but it differed somewhat in the intensity of its treatment. Children assigned to the Abcedarian treatment group attended an educational child care program from infancy to the start of Kindergarten (mean entry age was 4.4 months). The preschool component of the program cost roughly \$22,000 per child per year (2015 dollars).⁵ Campbell et al (2012) find a substantial effect on educational attainment at age 30 from Abcedarian, but no effect on the likelihood of criminal conviction by age 30. The differences in crime effects between Perry Preschool and Abcedarian are not directly comparable, since studies of the former observe arrests while studies of the latter observe convictions (likely a noisier measure).

2.2 Large-scale Targeted Preschool Programs

Evidence of the effectiveness of small-scale, high-intensity single-site interventions like Perry Preschool and Abcedarian naturally raises concerns about the replicability of these

⁵ Ramey et al state that the cost of the preschool component of Abcedarian was \$6,000 per child per year. Children entered the preschool component between 1972 and 1983.

programs and their results. The federal Head Start and the Chicago Child-Parent Centers (CPC) programs are also targeted to disadvantaged children, but at a much larger scale. This larger scale alleviates some of the external validity and scalability concerns with the smaller pilot programs. However, there is not yet experimental evidence on the impact of Head Start or CPC on adult outcomes, making it much more difficult to determine their effects.⁶

Head Start began as a federal summer program for low income children in 1966 and expanded to a full-year program by the early 1970s. The program currently costs between \$8,000 and \$10,000 per child (2015 dollars), much less than Perry Preschool or Abcedarian, and enrolls 900,000 children (Deming, 2009). While quality standards for Head Start are set at the federal level, the program is administered locally, leading to substantial heterogeneity in implementation quality across localities and over time.

Two studies, Garces et al (2002) and Deming (2009) identify the long-run effect of Head Start by comparing siblings who attended the program with those who did not, and assuming the siblings do not differ systematically. Garces et al use the 1964-1977 birth cohorts of the Panel Survey of Income Dynamics. The authors find that Head Start increases educational attainment, but only for white children. They also find dramatic reductions in the likelihood of facing criminal charges, but only for black children (12 percentage points). Deming looks at a later cohort of children, born in the early 1980s, using data from the National Longitudinal Mother-Child Supplement. He finds that Head Start participation increases a summary index of later adult outcomes by 0.23 standard deviations, despite the fadeout of short-run cognitive effects. He finds no impact on later crime, but measures crime differently than Garces et al., as an

⁶ There is a national experimental evaluation of Head Start, but thus far it only observes outcomes to third grade (Puma et al, 2012).

indicator equal to one if an individual is currently incarcerated or reports having been convicted, sentenced, or on probation.

Ludwig and Miller (2007) employ an alternative approach to examine the effects of Head Start on its early cohorts. Using a regression discontinuity design which compares counties around an eligibility threshold for grant writing assistance in 1965, they find that Head Start reduces childhood mortality rates and possibly increases educational attainment. They do not observe any criminal outcome measures.

The CPC program began in 1967 and was designed to provide educational and family support to children in high-poverty neighborhoods in Chicago that did not have access to Head Start. The program is operated by Chicago Public school system and provides a variety of services to children ages 3 to 9, including a preschool program. Preschool teachers are required to have a college degree and child-to-staff ratios are relatively low, 17:2 (Reynolds and Ou, 2011). In 1985, the average cost per child of the preschool component was \$9,636 in 2015 dollars (Reynolds et al, 2014).

Reynolds et al (2007) identifies the impact of CPC on adult outcomes by comparing CPC participants with a non-experimental comparison group of children in Chicago and controlling for covariates. At age 24, they find that CPC preschool participants were 4.6 percentage points (22%) less likely to have been arrested for a felony and 5.0 percentage points (20%) less likely to have been incarcerated. They also find that CPC participants were more likely to be employed, more likely to have a high school degree, and less likely to have depressive symptoms.

2.3 State-run Universal Pre-K

72

To date, there is no evidence on the impacts of state-run universal Pre-K programs on later adult outcomes. There have, however, been short-run impact studies in a number of states, including Georgia (Fitzpatrick, 2008), Tennessee (Lipsey et al, 2013), and Oklahoma (Gormley and Gayer, 2005; Gormley et al, 2011). Fitzpatrick uses a difference-in-difference approach with data from the National Assessment of Educational Progress and finds that the availability of Georgia Pre-K improved fourth grade test scores and on-grade percentage, but only for disadvantaged students in small towns and rural areas. Lipsey et al use random assignment of students applying to over-enrolled Pre-K programs in Tennessee to find that Pre-K improved achievement by the end of the Pre-K year. However, these gains were lost in the following year.

Gormley and Gayer (2005) investigate the achievement impact of Oklahoma's universal Pre-K program using data from Tulsa Public Schools in 2001. They use a birthdate regression discontinuity approach which leverages a strict cutoff in whether students were eligible to attend Pre-K in 2000 or 2001. However, since students on either side of this cutoff will both eventually receive the same access to Pre-K, they can only look at outcomes immediately after Pre-K, when a treatment contrast still exists. In essence, Gormley and Gayer compare the outcomes of Kindergarten students who have just finished Pre-K with students who are just starting Pre-K. They find a 0.39 standard deviation increase in cognitive test scores for those who have attended Pre-K. The impact is concentrated among blacks and Hispanics, with little effect on whites.

Looking again at universal Pre-K in Tulsa Public Schools, Gormley et al (2011) investigate the socio-emotional effects of the program using teacher ratings of kindergarteners in 2006. Using a teacher fixed effects approach with propensity score matching, they find that Pre-K participation was associated with higher attentiveness and lower timidity. The effects on these socio-emotional measures predict lower rates of future delinquency in adolescence (Moffitt et al, 1990).

3. OKLAHOMA UNIVERSAL PRE-K

Universal Pre-K (UPK) in Oklahoma differs from other early state universal programs (e.g. Georgia, New York) in both its scale and its quality. Unlike other states that operate Pre-K similar to a voucher system, Oklahoma treats Pre-K largely as another grade in the public education system, though an optional one. This distinction is evident in three important features of Oklahoma's pre-K provision. First, the majority of students attend school-based classrooms rather than independent centers. Second, four year olds are included in the formula for allocating state funds to districts rather than centers applying for funds directly from the state. Third, quality standards with regard to teachers and class sizes are enforced similarly for Pre-K and other grades. Pre-K teachers are required to be certified in early childhood education and to be paid at the same rate as other teachers. The adult-to-child ratio in classrooms cannot exceed 1:10 (Rose, 2011). These features of pre-K in Oklahoma make the quality of provision much more consistent than in other states that contracted their pre-K classrooms to outside providers. However, this consistent quality comes with a higher per student price tag than other universal pre-K programs, \$7,700 per student in 2013, with roughly half coming from the state and the remainder coming from federal and local sources (Barnett et al, 2013).⁷ This per student cost is similar to Head Start, but less than half that of Perry Preschool or Abcedarian.

⁷ Inflation adjusted to 2015 dollars.

In 1980, Oklahoma initiated a small-scale Pre-K pilot program targeted toward children from low-income families, but its funding was limited and it reached few children for the first two decades of its existence. By the 1997-98 school year, only 4% of four year olds who would eventually enter Oklahoma public schools were enrolled in state Pre-K.^{8,9} The rest of the preschool market was divided as follows: roughly one quarter of four year olds attended private preschool, 20% attended head start, and half did not attend any preschool.¹⁰ Starting in 1998-99, Oklahoma expanded the program to make all four year olds (as of September 1) eligible while simultaneously increasing the per-pupil funding for four year olds in the state funding formula. The result was a large jump in Pre-K access and enrollment, which I exploit to identify the impact of the program.

I measure the extent and distribution of the 1998-99 increase in Pre-K access and enrollment using data from the Common Core of Data's (CCD) Public Elementary/Secondary School Universe Survey, administered by the National Center for Education Statistics (NCES). Figure 1A shows that the total Pre-K enrollment in Oklahoma increased more than eight-fold (2,000 to 17,000) from 1997-98 to 1998-99. Figures 1B and 1C show the percent of future Oklahoma public school first graders who enrolled in Pre-K and who went to a school that provided Pre-K, respectively, by race.¹¹ They show that from 1997-98 to 1998-99 Pre-K enrollment (access) increased from 4% to 32% (10% to 45%) for white children and 6% to 41%

⁸ Pre-K enrollment rate is calculated as the number of students enrolled in Pre-K divided by the number of students in first grade two years later (Source: National Center for Education Statistics, Common Core of Data).

⁹ Given that 46% of students in the state were eligible for free or reduced price lunch (FRL) in 1998-99, only a very small fraction of children from low-income families were served by state Pre-K prior to 1998-99 (National Center for Education Statistics, Common Core of Data).

¹⁰ Private preschool and no preschool attendance rates are calculated from the October Current Population Survey. Head Start attendance rate is calculated as the total head start enrollment divided by the number of students in first grade two years later.
¹¹ Percent Pre-K enrollment for a given race is the calculated as the aggregate state Pre-K enrollment of that race in

¹¹ Percent Pre-K enrollment for a given race is the calculated as the aggregate state Pre-K enrollment of that race in year t divided by the aggregate state Grade 1 enrollment of that race in t+2. Pre-K enrollment by grade and race is not available prior to 1998-99, therefore it is imputed for all years by multiplying Pre-K enrollment in each school by the fraction of school enrollment of that race (it is then aggregated to the state-level).

(20% to 60%) for black children.¹² Consequently, the contrast in Pre-K enrollment (access) in the first year of UPK is 6 percentage points (5 percentage points) larger for black children than white children.

The increase in Pre-K enrollment and access differed substantially by socio-economic status. Figures 2A and 3A show the relationship between the percent of free and reduced luncheligible (FRL) students in a given zip code and the percent Pre-K enrollment and access in that zip code, respectively.¹³ The top panel depicts a locally weighted regression of percent Pre-K enrollment (or access) on percent FRL separately for 1997-98 and 1998-99, where the zip code is the unit of observation.¹⁴ The bottom panel shows the grade 1 enrollment distribution by zip code percent FRL. The top panel of both Figures 2A and 3A show a dramatically larger 1998-99 increase in Pre-K enrollment and access in higher percent FRL (lower socio-economic status) zip codes. In other words, children from low-income families, those at higher risk of later criminal charges, were more likely than children from higher-income families to have access to and enroll in Pre-K in the first year of UPK (relative to the prior year). Figures 2B and 3B show the same story using a zip code's family median income in 1999 as its measure socio-economic status, rather than percent FRL.¹⁵ Cascio and Schanzenbach (2013) present a similar picture of preschool enrollment responses to universal Pre-K implementation in Georgia and Oklahoma with data from the October Current Population Survey. Using a difference-in-differences strategy, they estimate a much larger increase in preschool enrollment among children with lesseducated mothers.

¹² Figure A5 shows similar increases in public preschool enrollment (Head Start and Pre-K) for the full sample of Oklahoma four year olds using data from the October Current Population Survey.

¹³ A student is considered to be part of a zip code if they attend a school located in that zip code.

¹⁴ Zip codes with grade 1 enrollment less than 20 are excluded.

¹⁵ Zip code median family income obtained from Table P077 of the 2000 Census.

In addition to the risk levels of children that enrolled, the effect of Pre-K will depend on the preschool services that those enrolled children would have experienced if they did not attend Pre-K. Figure 4 shows the mix of preschool services in Oklahoma in the two years before and after UPK implementation. The fractions in the figure are imprecise, as they are calculated using only the 115 observations of Oklahoma four year olds in the October Current Population Survey (1996-1999). Despite this limitation, the figure suggests that the UPK increase in Pre-K enrollment drew four year olds who would not have attended preschool or who would have attended private preschool. There is no evidence that UPK drew enrollment away from Head Start.

4. OKLAHOMA CRIME AND BIRTH DATA

I use two sources of data to measure the long-run impact of universal Pre-K availability on criminal outcomes in Oklahoma. First, I obtain data on criminal charges filed against 18 and 19-year-olds in Oklahoma for the cohort of individuals that turned five in the years surrounding UPK implementation. Second, I obtain data on births in the state of Oklahoma for the same cohort of individuals. I use these two data sources to construct a measure of the likelihood of criminal charges at age 18-19 for individual birthdate cohorts.

4.1 Criminal Court Data

Two organizations provide public access to criminal court data in Oklahoma: The Oklahoma State Courts Network (OSCN) and On Demand Court Records (ODCR). Together, they cover 71 of the 77 counties in Oklahoma. These counties account for 96% of the arrests in

Oklahoma.¹⁶ Each organization maintains a website that provides public access to detailed charge and defendant information, presented separately on individual webpages by court case.¹⁷ I systematically scrape relevant information from the html code on these webpages and compile it into a dataset with observations at the defendant by case level, which contains the names, birthdates, demographics, and charges of defendants.

The compiled data includes criminal charges filed against 18 and 19-year-olds in Oklahoma from January 1, 2010 to May 1, 2014. For those born between January 1, 1992 and May 1, 1994 (the cohorts surrounding the first exposed to UPK), the data cover the timeframe between their 18th and 20th birthdays (the full years they were age 18 and 19). For a larger range of birthdates, from January 1, 1992 to May 1, 1995, the data covers the timeframe between 18th and 19th birthdays (the full year they were age 18). The number of unique criminal court cases observed in my sample for 18-19 year olds (5,346 in 2013) is 54% of the number of arrests of 18-19 year olds reported in Oklahoma (9,892 in 2013).^{18,19} This relationship mirrors the 53% of arrests that result in charges for southern 18-19 year olds in the National Longitudinal Survey of Youth 1997 (NLSY97), suggesting that my sample of court records is broadly consistent with officially reported law enforcement data.²⁰

I collapse the criminal court data to the birthdate-level to obtain the number of unique individuals with a given birthdate who were charged with a crime at age 18 or 19. I repeat this

¹⁶ Source: Oklahoma State Bureau of Investigation, "State of Oklahoma Uniform Crime Report, Annual Report January-December 2013."

¹⁷ Figures A1 and A2 show example screenshots of these webpages.

¹⁸ Unique criminal cases are defined as unique combinations of court ID, defendant first and last name, defendant birthdate, and case filing date. The number of arrests in Oklahoma is obtained from the Oklahoma State Bureau of Investigation, "State of Oklahoma Uniform Crime Report, Annual Report January-December 2013." Arrest statistics exclude traffic offenses while the court records do not.

¹⁹ In 2012, the number of unique criminal cases (5,717) is 51% of the total number of arrests (11,112).

²⁰ This statistic reflects the ratio of the total arrests and total charges reported "since last interview" by 18-20 year olds in the south region (which includes Oklahoma) in the NLSY97. 20 year olds are including because arrests and charges that they report may have occurred while they were age 19.

process for various subsamples by race (i.e. black only and white only), age (i.e. 18 only and 19 only), and charge severity (i.e. felony only or misdemeanor only). Table 1 shows the number of these individuals in the year surrounding the birthdate eligibility cutoff for UPK availability in 1998-99. In all, 2,517 unique 18-19 year olds were charged with a misdemeanor and 2,005 were charged with a felony.

ODCR and OSCN are missing race information for 55% of the criminal defendants in my sample, complicating my subsample analyses.²¹ For these defendants I impute race using a race prediction index based on the likelihood that a person is of a particular race given their first and last name. Specifically, the index is the predicted probability from a probit regression using observations in my sample that observe race. The dependent variable of this regression is an indicator for a given race and the independent variables consist of cubics of the fraction of criminal defendants with the same first or last name that are of a given race. Appendix Figures A3 and A4 show the distribution of the race prediction index by observed race for the white name and black name index. These figures show that the prediction index fits the observed race of defendants quite well. For my main analysis, I select a relatively tight threshold for imputing a defendant's missing race in order to minimize type I error. Specifically, if a defendant's prediction index for a given race is greater than or equal to 0.9, then I include them in subsamples limited to that race. In supplementary analyses, I exclude individuals with missing race and also use looser imputation thresholds (i.e. 0.1, 0.25, 0.5, and 0.75). This has no qualitative effect on my results, though it does affect the magnitudes of the coefficient estimates

²¹ OSCN provides the defendant's race, but this information is missing or listed as "Unknown" for 15% of defendants. ODCR does not provide information on defendant's race.

as would be expected.²² After imputing race based on first and last names, I have race information for 97% of the defendants in my sample. Table 1 shows the number of unique individuals charged with a crime by race both including (column 2 and 3) and excluding charges with imputed race (column 4 and 5). Excluding charges with imputed race leads to subsample charge counts that are between one quarter and one half the size of the size of the imputed subsample charge counts.

4.2 Births Data

I obtain data on all births in the state of Oklahoma for the years 1992-1995 from the public-use Natality File maintained by the National Center for Health Statistics (NCHS) at the Centers for Disease Control and Prevention. This data is recorded at the individual-birth level and contains the year, month, and day of the week of the birth (i.e. Sunday, Monday, etc.), as well as the sex of the child and the state of residence and race of the mother. I restrict the sample to mothers residing in Oklahoma and collapse the data to obtain the number of births by year-month-day of the week. Table 1 shows the number of births in Oklahoma in the year surrounding the birthdate eligibility cutoff for UPK in 1998-99. For the full sample and separately by mother's race, I impute the number of births on a given date by dividing the number of births in a year-month-day of the week bin by the number of days in that bin.²³

4.3 Criminal Outcomes for Individual Birthdate Cohorts

²² Reducing the number of individuals whose charges are included in the analysis lowers the numerator of the outcome variable (count of unique individuals with a criminal charge) but not the denominator (count of births), therefore it lowers the baseline mean of the outcome variable. The expected result is a reduction in the magnitude of the coefficient estimate.

²³ I find no change in my results when I account for the reduced number of births on holidays by deflating the births on these days by the average difference between births on Sundays and weekdays (and adjusting the other days in the same bin).

I combine the birthdate-level criminal charge and birth data to construct my primary outcome measure: the likelihood of criminal charges at age 18-19 by date of birth. This is the count of unique individuals with a given birthdate that were charged with a crime in Oklahoma at age 18-19, divided by the count of individuals born in Oklahoma on that date. I repeat this construction for various subsamples by race (i.e. black only and white only), age when charged (i.e. 18 only and 18-19), and charge severity (i.e. felony only or misdemeanor only). For the outcome measure constructed separately by race, I use mother's race to determine the denominator (number of births) and the individual's reported or imputed race to determine the numerator (number of individuals charged).

5. EMPIRICAL STRATEGY

I estimate the impact of Oklahoma's universal pre-K program (UPK) on later criminal behavior using a regression discontinuity (RD) design and difference-in-regression-discontinuity (DRD) design. These strategies yield estimates of the effect of UPK availability, or equivalently the intent-to-treat effect of state Pre-K, compared to the prior mix of preschool services in Oklahoma. This prior mix consisted mainly of head start, private preschool, and no preschool (approximately 20%, 25%, and 50%). The identification of the effect of UPK eligibility in both approaches leverages the treatment contrast between children just below and just above the Kindergarten birthdate cutoff in the first year of the UPK implementation. In this section, I first discuss this treatment contrast in further detail, then I present the standard RD framework, and finally I present the DRD framework.

5.1 Treatment Contrasts

Students in Oklahoma must be 5 years old on September 1 to attend Kindergarten in the public school system. This creates a birthdate cutoff where children born on or before September 1 in a given year are assigned to a different school cohort than children born after September 1. Table 2 shows how the schooling experiences of these children differed in the years surrounding the implementation of universal Pre-K (UPK). In the table, PK denotes that a child is *eligible* to attend state Pre-K as part of UPK. Therefore, the contrast between PK and no PK reflects the intent-to-treat (ITT) effect of the state's Pre-K program relative to the pre-existing mix of preschool services, rather than the treatment effect of actually attending state Pre-K. While not capturing the effectiveness of local Pre-K implementation because 60% of students did not attend state Pre-K, the ITT effect is the parameter of primary interest for state policymakers contemplating a similar policy of voluntary universal Pre-K.

In the 1998-99 school year, the first year of UPK implementation, Child C in Table 2 (born September 1, 1993) just meets the birthdate cutoff and therefore attends Kindergarten in 1998-99. She never experiences UPK eligibility because it had not been available the prior year. On the other hand, Child D in Table 2 (born September 2, 1993) misses the birthdate cutoff and is therefore eligible for UPK in 1998-99. This means that Child C and Child D differ in their access to UPK. But this is not the only difference between them. They also differ in relative age to their classmates: Child C is the youngest in her cohort and Child D is the oldest in her cohort. This difference is problematic for my identification strategy to the extent that it affects later criminal outcomes. Dobkin and Ferreira (2010) find that students who are young relative to their peers attain slightly higher levels of education, but do not perform as well academically. They find that these differences do not impact later job market outcomes but they do not observe criminal outcomes. Except for the first year of UPK implementation, there is no policy contrast at the Kindergarten birthdate cutoff. In 1997-98, prior to UPK, Child A in Table 2 (born September 1, 1992) meets the September 1 cutoff and attends Kindergarten in 1997-98 as the youngest in her grade. Child B in Table 2 (born September 2, 1992) just misses the birthdate cutoff and so she does not attend Kindergarten in 1997-98, but attends in 1998-99 as the oldest in her grade. Critically, Child A and B differ in their relative ages, but *neither* are eligible for UPK because it is not implemented until they are too old. Similarly, Child E and F in Table 2 (born September 1 and 2, 1994) differ in their relative ages, but *both* are eligible for UPK. I will leverage the lack of policy contrast in these other years to test whether discontinuities in the likelihood of criminal charges at the Kindergarten birthdate cutoff in the first year of UPK are a result of UPK or relative age differences.

5.2 Regression Discontinuity Design

My basic regression discontinuity (RD) design looks at a one year window of birthdates around the Kindergarten birthdate cutoff in the first year of UPK implementation (March 2, 1993 to March 1, 1994) and estimates the discontinuity in the likelihood of later criminal charges that occurs at this eligibility threshold (September 1, 1993). Intuitively, this strategy compares children born just before and just after the eligibility threshold (Child C and Child D in Table 2). The regression specification is as follows,

$$Y_b = \alpha + \beta(z_b \ge 0) + f(z_b) + (z_b \ge 0) \cdot f(z_b) + \epsilon_b, \tag{1}$$

where Y_b is the proportion of children born on date b who face criminal charges at ages 18 or 19 (or age 18 only) and z_b is the difference between date of birth b and the eligibility threshold for kindergarten in the 1998-99 school year (September 1, 1993). In most specifications, $f(z_b)$ is a linear function of z_b . By interacting $f(z_b)$ and $(z_b \ge 0)$, the specification allows the function to vary on either side of the eligibility threshold. The coefficient of interest, β , can be interpreted as the ITT effect of state Pre-K on children who just missed the Kindergarten cutoff (i.e. the oldest students relative to their grade).

The key identifying assumption of this RD approach is that variables related to the outcome must vary smoothly, and not discontinuously, through the cutoff. In other words, except for their UPK eligibility, Child C and Child D (from Table 2) must be the same in ways that might affect the outcome variable. A potential problem is that there is a substantial difference between Child C and Child D: they differ in relative age to their classmates (C is the oldest in her cohort and D is the youngest). This difference could impact their likelihood of criminal charges later in life, and therefore bias my RD estimates of the impact of universal pre-K. In one approach to investigating this concern, I estimate RD specification (Equation 1) for one year windows around the Kindergarten birthdate cutoff in the prior (or subsequent) year, when children on either side of the cutoff are either both UPK ineligible (i.e. Child A and B) or both UPK eligible (i.e. Child E and F).

Another potential internal validity problem would be differential exit from the state (by age 18-19) at the UPK eligibility threshold. I only observe criminal charges for those that stay in the state, so differential exit would affect the numerator of my outcome variable at the cutoff. Figure A6 shows the percent of Oklahoma-born 18-19 year olds who stay in Oklahoma, by quarter of birth, using data from the American Community Survey 2005-2013.²⁴ Though the measure of birth timing is coarse, the figure does not show any evidence of differential exit around the UPK eligibility threshold (1993Q3 vs. 1993Q4).

²⁴ Individual birthdates are not available in the public-use American Community Survey.

5.3 Difference-in-Regression-Discontinuity Design

In another approach to account for the potentially confounding factor of relative differences at the UPK eligibility threshold, I use a difference-in-regression-discontinuity (DRD) design. This design explicitly incorporates any outcome discontinuity at the eligibility threshold in the prior (or subsequent) year, when there was no UPK policy contrast, by measuring the outcome discontinuity in the first UPK implementation year relative to the discontinuity in the baseline year. Therefore, this approach accounts for differences between the oldest and youngest students (relative to their grade) that are not related to UPK eligibility.

Under additional assumptions, the DRD design can be used to make inferences regarding the impact of UPK on children that are not the relative oldest in their grade, thereby extending the external validity of the RD strategy discussed in Section 5.2. The intuition behind this extension is that Child A and C in Table 2 are the equivalent in terms their UPK eligibility and relative ages, and Child D and F are similarly equivalent. Therefore, I will assume that:

(A1)
$$Y^A = Y^C$$
, and

(A2)
$$Y^D = Y^F$$

where Y^x is the outcome for child *x* in Table 2. Given these assumptions, I will infer the effect of UPK on the relatively youngest children $(Y^E - Y^C)$ by taking the difference of the outcome discontinuities at the Kindergarten birthdate cutoff in the first year of UPK $(Y^D - Y^C)$ and the second year of UPK $(Y^F - Y^E)$. This yields the identity $Y^E - Y^C = (Y^D - Y^C) - (Y^F - Y^E)$. Similarly, I will infer the effect of UPK on the relatively oldest children from the identity $Y^D - Y^B = (Y^D - Y^C) - (Y^B - Y^A)$. For the DRD design, I employ following empirical specification using various birthdate ranges:

$$Y_{yb} = \alpha + \beta_1 (UPK YEAR1_y) \times (z_b \ge 0) + \beta_2 (z_b \ge 0) + \beta_3 (UPK YEAR1_y)$$

+ $f(z_b) + (z_b \ge 0) \times f(z_b) + (UPK YEAR1_y) \times (z_b \ge 0) \times f(z_b) + \epsilon_{yb},$ (2)

where the variables are defined similarly to Equation 1, but *y* indexes the school year of the relevant Kindergarten eligibility threshold around which the one year birthdate window is constructed. *UPK YEAR*1_{*y*} is an indicator equal to one for the 1998-99 school year Kindergarten eligibility threshold, UPK's first implementation year. This specification estimates Equation 1 separately for the 1998-99 Kindergarten eligibility threshold and for the threshold in the baseline year within one regression (the polynomial in z_b is allowed to vary by school year *y* and on either side of the eligibility cutoff). β_1 is the difference between the discontinuity at the eligibility threshold in 1998-99 and in the baseline year.

If assumptions A1 and A2 hold, and the analysis sample includes windows around the 1998-99 and 1999-00 birthdate cutoffs (UPK years 1 and 2), β_1 can be interpreted as the effect of UPK availability on the youngest children relative to their cohort (i.e. $Y^E - Y^C$). This is because the only group *not eligible* for UPK in this timeframe was the relatively young children at the 1998-99 birthdate cutoff (Child C in Table 2). When the analysis sample includes windows around the 1998-99 and 1997-98 birthdate cutoffs (UPK Year 1 and prior year), β_1 can be interpreted as the effect of UPK availability on the oldest children relative to their grade (i.e. $Y^D - Y^B$). This is because the only group *eligible* for UPK in this timeframe was the relatively old children at the 1998-99 birthdate cutoff (Child D in Table 2).

6. RESULTS

Table 3 shows least squares estimates of β in Equation 1, weighted by the number of births on a given date.²⁵ The dependent variable is the likelihood of criminal charges at age 18 or 19. Estimates of β are presented by race and charge type (i.e. felony or misdemeanor) for one year windows (plus or minus six months) around Kindergarten birthdate eligibility thresholds in two years: the year prior to UPK implementation (1997-98) and the first year of UPK implementation (1998-99).²⁶ These are the latest cohorts for whom I observe criminal charges at age 18 and 19. In the first year of UPK implementation, I find large negative effects on later charge likelihood for black children who just missed the Kindergarten age cutoff (and were therefore eligible for UPK), but no effect for white children. The estimates can be interpreted as UPK eligibility causing a 4.5 percentage point (26% of the mean rate) reduction in the likelihood of a felony charge (Column 3) and a 6.8 percentage point (38% of the baseline rate) reduction in the likelihood of a misdemeanor charge (Column 4) at ages 18-19. I find no significant effects of missing Kindergarten birthdate cutoff in the year prior to UPK implementation, suggesting that these large estimated effects are not driven by differences on either side of the Kindergarten birthdate cutoff in students' ages relative to their classmates. I will test this explicitly using the DRD design.

Figures 5 and 6 depict the results in Table 3 graphically, showing the likelihood of a misdemeanor (Panel A) or felony (Panel B) charge at age 18-19 by birthdate (each dot shows the mean for a two week bin). Figure 6, Panel A shows no discontinuity in the likelihood of misdemeanors for blacks in the year prior to UPK implementation, but a large discontinuity in

 ²⁵ The weights account for the lower variance of the outcome variable on dates when more births occur.
 ²⁶ In some analyses, I also use a window around the birthdate cutoff in the second year of UPK.

the first year. Figure 6, Panel B shows a small (and insignificant according to Table 3) discontinuity in the prior year and a larger discontinuity in the first year of implementation, suggesting that the felony effect in 1998-99 may be capturing both the effect of UPK and differences in the later felony rates of children that are relatively old and relatively you compared to their classmates.

Table 4 shows the same estimates of β in Equation 1 for the first year of UPK as Table 3, but for various bandwidths (plus or minus 1 to 6 months) around the Kindergarten birthdate cutoff. The insignificant estimates for white felonies and misdemeanors are consistent across bandwidths. The black misdemeanor estimates remain similar and statistically significant (at the six percent level or less) for all bandwidths, though the magnitudes are larger for smaller bandwidths. The black felony estimates also remain similar for smaller bandwidths, though the three month bandwidth estimate drops slightly below the ten percent significance level and the one month bandwidth is not significant. Table A1 and A2 show that these results are robust to various imputation thresholds for the race name prediction index, as well as to excluding charges with missing race. Table A3 and A4 also show similar results when using the count of individuals charged as the dependent variable (without dividing by the number of births).

Table 5 shows results from local linear (kernel regression) rather than a linear least squares estimation of β in Equation 1. Estimates are presented for various kernel bandwidth choices, including the optimal bandwidth following Imbens and Kalyanaraman (2011). The results are somewhat unstable at small bandwidths, but converge at higher bandwidths. At these higher bandwidths the results are similar to Table 3, though felony estimates are not significant and misdemeanor estimates are somewhat larger in magnitude (and significant or nearly

significant). Figures A7 and A8 depict these results graphically, showing the local linear estimates and 95% confidence intervals by bandwidth choice.

Table A5 shows the same results as Table 3, but the dependent variable is the likelihood of criminal charges at age 18 only. Unlike the age 18-19 charge likelihood, this outcome can be observed for the window around the Kindergarten birthdate cutoff in the second year of UPK implementation (1999-00). Therefore, Table A5 presents results for two years where there should be no contrast in UPK availability, one year prior to UPK implementation (Child A and B in Table 2) and one year after (Child E and F in Table 2). Again, I find no significant effects in the years without treatment contrast and no significant effects on white children. I find a large negative impact of 3.5 percentage points on the likelihood of later misdemeanor charges for black children (40% of the mean rate), but no impact on the likelihood of later felony charges. Figure A9 and A10 present these results graphically (as in Figure 5 and 6). Table A6 and A7 show the same results for various bandwidths and race imputation thresholds. The black age 18 misdemeanor effect is less robust than for age 18-19, losing statistical significance at 1, 3, and 4 month bandwidths.

Table 6 shows least squares estimates for β_1 in Equation 2, the DRD design, for various comparison years. Panel A and C show the differential effect of missing the Kindergarten birthdate cutoff in the first year of UPK compared to the prior year (Child B vs. Child D in Table 2). The sample includes birthdates six months before the 1997-98 Kindergarten cutoff to six months after the 1998-99 Kindergarten cutoff. Panel B shows the differential effect of missing the Kindergarten cutoff in the first year of UPK compared to the prior year *and* the following year. The sample includes birthdates six months before the 1997-98 Kindergarten cutoff to six months after the 1999-00 Kindergarten cutoff. The Panel D shows the differential effect of

missing the Kindergarten cutoff in the first year of UPK compared to the following year (Child C vs. Child E in Table 2). The sample includes birthdates six months before the 1997-1998 Kindergarten birthdate cutoff to six months after the 1999-2000 Kindergarten birthdate cutoff.

As in the basic RD results, there appears to be a negative impact on black misdemeanors, though it is a similar magnitude and statistically significant for misdemeanors at age 18-19. The estimates show a first year differential effect of UPK on the likelihood of misdemeanor for blacks of 5.7 percentage points at age 18-19. There is some evidence of a smaller impact on black felonies at age 18-19 of 2.8 percentage points, but it is not significant. Taken together, Table 6 lends further support to the hypothesis that UPK availability rather than relative age differences generating the change in criminal outcomes for blacks at the K cutoff.

While not significant, the similarity in the effect estimates of Panels C and D of Table 6 suggest that the effect of UPK availability may be similar on the youngest and oldest children (relative to their peers). That is, $Y^E - Y^C \approx Y^D - Y^B$. This is suggestive evidence that the RD estimates may be generalizable beyond the oldest students (relative to their peers) on which these estimates are identified.

7. APPROXIMATE TREATMENT-ON-TREATED EFFECT

While the ITT effect estimates in Section 6 are likely to be the relevant parameters for state policymakers considering similar Pre-K policies, local administrators may be more interested in the impact of state Pre-K attendance rather than UPK eligibility. In Appendix 1, I calculate a back-of-the-envelope estimate of the treatment-on-treated (TOT) effect on Oklahoma residents implied by the ITT effect estimates in Section 6. The TOT effect is the impact of UPK on UPK *compliers*, those who attend Pre-K as a result of the new policy. I calculate the TOT effect on 18-19 black misdemeanors by dividing the ITT effect estimate in Table 6 by the change in the Pre-K enrollment rate of black four year olds in the first year of UPK. The resulting implied TOT effect is a roughly 19 percentage point reduction in the likelihood of a charge. However, this approximation will over-estimate the true TOT effect if other preschool services (e.g. Head Start) responded to UPK by improving their own quality. There is some anecdotal evidence that this may have been the case. For example, the Community Action Project of Tulsa County established partnerships between Head Start providers and local school districts, where Head Start centers adopted some of the quality components of UPK.²⁷

The interpretation of the magnitude of the TOT effect depends critically on the underlying baseline rate, which is likely to be higher than the general population if UPK compliers are typically from lower income families than the general population. In Appendix 1, I calculate the income distribution for UPK compliers based on the median family income of their school's zip code, and reweight the baseline criminal charge rate using this distribution. Figure 8 shows the zip code median family income distribution of UPK compliers vs. the overall grade 1 enrolled population. Figure 9 shows the TOT effect on Oklahoma residents as a percent of the calculated baseline criminal charge rate for various assumptions regarding UPK compliers' family income relative to their zip code median income (η) and criminal charge rates of new migrants to Oklahoma relative to those who have stayed in the state (α). For example, for $\eta = 0.45$ and $\alpha = 0.6$, the TOT effect on the likelihood of black misdemeanors at age 18-19 is a roughly 50% decrease relative to the baseline rate for black UPK compliers who reside in Oklahoma.

8. CONCLUSION

In this chapter, I leverage a contrast in the availability of Oklahoma universal Pre-K (UPK) in the first year of its implementation to estimate the effect of UPK availability on later criminal outcomes at age 18 and 19. I find a significant negative impact of UPK availability on the likelihood that black children are later charged with a crime at age 18 or 19, but no impact on the likelihood of later charges for white children. As the first estimates of the long-run impacts of universal pre-K availability, these results are an important first step towards bridging the disconnect between the policy debate around universal Pre-K and the evidence of its potential long-run effectiveness. The results suggest that universal Pre-K can, like more targeted programs, have dramatic effects on later criminal outcomes, but these effects are concentrated among more at-risk populations. As with Perry Preschool, these large crime reductions are likely to determine a major part of the benefits of universal Pre-K in future benefit-cost analyses. The program's benefits may outweigh its costs, but if effects on other outcomes (e.g. earnings) are also concentrated among at-risk populations, then it is liable to be a less efficient use of public funding than a more targeted program of equal quality.

REFERENCES

- Barnett, W.S. (1996). Lives in the Balance: Age 27 Benefit–Cost Analysis of the High Scope Perry Preschool Program. High/Scope Press, Ypsilanti, MI.
- Barnett, W.S., Carolan, M.E., Squires, J.H., Clarke Brown, K. (2013). The state of preschool 2013: State preschool yearbook. New Brunswick, NJ: National Institute for Early Education Research.
- Barnett, S., Jung, M., Youn, M., Frede, E. (2013). Abbott Preschool Program Longitudinal Effects Study: Fifth Grade Follow-Up. National Institute for Early Education Research.
- Bartik, T., Gormley, W., Adelstein, S. (2012). Earnings benefits of Tulsa's pre-K program for different income groups. Economics of Education Review, 31, 1143–1161
- Campbell, F. A., Pungello, E. P., Burchinal, M., Kainz, K., Pan, Y., Wasik, B. H., Barbarin, O. A., Sparling, J. J., & Ramey, C. T. (2012). Adult Outcomes as a Function of an Early Childhood Educational Program: An Abecedarian Project Follow-Up. Developmental Psychology.
- Cascio, E. U. and Schanzenbach, D. W. (2013). The Impacts of Expanding Access to High-Quality Preschool Education. Brookings Papers on Economic Activity, Fall 2013.
- Deming, D. (2009). Early Childhood Interventions and Life-Cycle Skill Development: Evidence from Head Start. American Economic Journal: Applied Economics, 1:3, 111–134.
- Fitzpatrick, M. (2008). Starting School at Four: The Effect of Universal Pre-Kindergarten on Children's Academic Achievement. The B.E. Journal of Economic Analysis & Policy: Vol. 8: Iss. 1 (Advances), Article 46.
- Garces, E., Thomas, D., and Currie, J. (2002). Longer-term Effects of Head Start. The American Economic Review, 92(4), 999-1012.
- Gibbs, C., Ludwig, J., and Miller, D. (2012). Does Head Start Do Any Lasting Good? Working Paper.
- Gormley, W. T., Jr. (2010). Small miracles in Tulsa: The effects of universal pre-k on cognitive development. In A. J. Reynolds, A. J. Rolnick, M. M. Englund, & J. A. Temple (Eds.), Childhood programs and practices in the first decade of life: A human capital integration (pp. 188–198). New York: Cambridge University Press.
- Gormley, W. T., Jr., & Gayer, T. (2005). Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-k program. Journal of Human Resources, 40(3), 533–558.
- Gormley, W. T., Jr., Gayer, T., Phillips, D. A., & Dawson, B. (2005). The effects of universal pre-k on cognitive development. Developmental Psychology, 41(6), 872–884.
- Gormley, W. T., Jr., Phillips, D. A., & Gayer, T. (2008). Preschool programs can boost school readiness. Science, 320(5884), 1723–1724.

- Gormley, W. T., Jr., Phillips, D. A., Adelstein, S., & Shaw, C. (2010). Head Start's comparative advantage: Myth or reality? Policy Studies Journal, 38(3), 397–418.
- Gormley, W., Phillips, D., Newmark, K., Welti, K., & Adelstein, S. (2011). Social-emotional effects of early childhood education programs in Tulsa. Child Development, 82, 2095–2109.
- Heckman, J. J., Malofeeva, L., Pinto, R., & Savelyev, P. A. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. American Economic Review 2013, 103(6), 2052–2086.
- Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010a). The rate of return to the high/scope Perry preschool program. Journal of Public Economics, 94(1–2), 114–128.
- Heckman, J. J., Moon, S., Pinto, R., Savelyev, P., Yavitz, A. (2010b). Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. Quant Econom, 1(1), 1–46.
- Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. Journal of Labor Economics, 24(3), 411–482.
- Ladd, H., Muschkin, C., Dodge, K. (2014). From Birth to School: Early Childhood Initiatives and Third-Grade Outcomes in North Carolina. Journal of Policy Analysis and Management, Vol. 33, No. 1, 162–187.
- Lipsey, M. W., Hofer, K. G., Dong, N., Farran, D. C., & Bilbrey, C. (2013). Evaluation of the Tennessee Voluntary Prekindergarten Program: Kindergarten and First Grade Follow-Up Results from the Randomized Control Design (Research Report). Nashville, TN: Vanderbilt University, Peabody Research Institute.
- Ludwig, J. and D. L. Miller (2007). Does Head Start Improve Children's Life Chances? Evidence from a Regression-Discontinuity Design. Quarterly Journal of Economics. 122(1), 159-208.
- Ludwig, J. and D. A. Phillips (2007). The benefits and costs of Head Start. Cambridge, MA: NBER Working Paper 12973.
- Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., Mashburn, A. and Downer, J. (2012). Third Grade Follow-up to the Head Start Impact Study Final Report, OPRE Report # 2012-45, Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Reynolds, A. J. Temple, J. A., Ou, S., Robertson, D. L., Mersky, J. P., Topitzes, J. W. & Niles, M. D. (2007). Effects of a School-Based, Early Childhood Intervention on Adult Health and Well Being: A 19-Year Follow Up of Low-Income Families. Archives of Pediatrics & Adolescent Medicine. 161(8), 730-739.

- Reynolds, A. and S. Ou, (2011). Paths of Effects from Preschool to Adult Well-Being: A Confirmatory Analysis of the Child-Parent Center Program. Child Development, 82(2), 555– 582.
- Reynolds, A. J., Richardson, B., Hayakawa, C. M., Lease, E. M. Warner-Richter, M. Englund, M. M., Ou, S., Sullivan, M. (2014). Association of a full-day vs part-day preschool intervention with school readiness, attendance, and parent involvement. Journal of American Medical Association, 312(20):2126-34. doi: 10.1001/jama.2014.15376
- Rose, Elizabeth (2011). Pre-Kindergarten in Oklahoma In E. Zigler, W. S. Gilliam, and W. S. Barnett, *The Pre-K Debates: Current Controversies and Issues*. Brookes Publishing Company. PO Box 10624, Baltimore, MD 21285.
- Solomon, D. (2007). As States Tackle Poverty, Preschool Gets High Marks. Retreived April 29, 2015, from http://www.wsj.com/article_email/SB118660878464892191lMyQjAxMDE3ODA2OTYwMDk4Wj.html.
- Temple, J. A. & Reynolds, A. J. (2007). Benefits and costs of investments in preschool education: Evidence from the Child-Parent Centers and related programs. *Economics of Education Review*, 26 (1), 126-144.
- Weiland, C. and Yoshikawa, H. (2013). Impacts of a Prekindergarten Program on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills. Child Development, 84: 2112–2130.
- White, J. L., Moffitt, T. E., Earls, F., Robins, L., and Silva, P. A. (1990). How early can we tell? Predictors of childhood conduct disorder and adolescent delinquency. Criminology, 28, 507– 533.

APPENDIX 1: Back-of-the-Envelope Calculation of Implied TOT Effect

In this section, I calculate an implied back-of-the-envelope estimate of the treatment-ontreated (TOT) effect of UPK. This is the effect on UPK compliers: those that attended state Pre-K as a result of the new policy. Because I only observe criminal charges of 18-19 year olds who remain in Oklahoma, I will focus on the impact on the UPK compliers who remain in Oklahoma.

Recall that the observed outcome measure is given by $Y_b = \frac{C_b^R}{N_b^B}$, where C_b^R is the number of Oklahoma residents born on date *b* who were charged with a crime at a given age. N_b^B is the number of births in Oklahoma on date *b*. We can think of β in Equation 1, the ITT effect of UPK on this measure, as representing $E[Y_{b^*} | eUPK = 1] - E[Y_{b^*} | eUPK = 0]$, where b^* is the birthdate eligibility cutoff for UPK and eUPK is an indicator for UPK eligibility. In order to adjust β to reflect the ITT effect of UPK on the likelihood of criminal charges for current Oklahoma residents I multiply it by the ratio of residents born on b^* ($N_{b^*}^R$) to Oklahoma births ($N_{b^*}^R$).

$$\beta^{R} = \frac{E[C_{b^{*}}^{R} \mid eUPK=1] - E[C_{b^{*}}^{R} \mid eUPK=0]}{N_{b^{*}}^{R}} = \frac{N_{b^{*}}^{B}}{N_{b^{*}}^{R}} \cdot \beta.$$
(3)

I approximate $\frac{N_{b^*}^B}{N_{b^*}^B}$ as the ratio of the number of children born in Oklahoma to the number of 19 year olds residing in Oklahoma 19 years later, using population data from the Surveillance, Epidemiology, and End Results Program (SEER). For the black subsample, this yields the approximation $\frac{N_{b^*}^B}{N_{b^*}^B} \approx 0.76$.

Under the assumption that UPK has no impact on those that did not receive the treatment, we can write the TOT effect of UPK on those who remained in Oklahoma, τ^R , as follows:

$$\beta^R = p_{b^*} \cdot r_{b^*} \cdot \tau^R \tag{4}$$

where p_{b^*} is the fraction of Oklahoma four-year-olds born on b^* that were UPK compliers in the first year of implementation. r_{b^*} is the fraction of Oklahoma residents born on b^* who resided in Oklahoma at age four. I cannot observe p_{b^*} or r_{b^*} and so I approximate them from the Common

Core of Data and the 2000 Census, respectively.^{28,29} For the black subsample, $p_{b^*} \approx 0.35$ and $r_{b^*} \approx 0.64$. Using the approximations of $\frac{N_{b^*}^B}{N_{b^*}^R}$, p_{b^*} , and r_{b^*} , discussed above, and the ITT effect estimate on black misdemeanors in Table 6 (-0.057), Equations 3 and 4 yield $\tau^R = -0.194$.

Interpreting the magnitude of this TOT effect depends critically on the baseline outcome variable (i.e. expected charge likelihood of the population of UPK compliers who remain in Oklahoma). Therefore, I reweight the sample mean of the relevant outcome variable, $\overline{Y^R} = \frac{c^R}{N^R}$, to reflect the expected value for the UPK compliers who remain in Oklahoma, \tilde{Y}_B^R .

First, I assume that Oklahoma residents who were in Oklahoma at age 4 differ in their criminal charge likelihood from those who were not by a factor α . Therefore,

$$\overline{Y^R} = \overline{Y^R_B} \left(r + \alpha (1 - r) \right) \tag{5}$$

where $\overline{Y_B^R}$ is the mean charge likelihood of residents who were residing in Oklahoma at age 4 and r is the fraction of residents who were in Oklahoma at age 4. This distinction is motivated by substantial observed differences in lifelong Oklahoma residents and new residents. Among black Oklahoma residents age 18-20, those that were born in Oklahoma are 8 percentage points (76%) more likely to be unemployed, 9 percentage points (38%) more likely to not have a high school diploma, and 13 percentage points (28%) less likely to have attended college.³⁰

Next, I assign each Pre-K aged child in 1998-99 the median family income of their first grade school's zip code and make the conservative assumption that children did not leave Oklahoma differentially by income.³¹ This means that I can rewrite the $\overline{Y_B^R}$ and \tilde{Y}_B^R as

$$\overline{Y_B^R} = \frac{\sum_j E_j \cdot c(x_j)}{\sum_j E_j}$$
 and $\widetilde{Y}_B^R = \frac{\sum_j E_j \cdot p_j \cdot c(\eta x_j)}{\sum_j E_j}$,

 $^{^{28}}$ I approximate *q* as the difference in OK Pre-K participation rates in the 1997-98 and 1998-99 school years, where participation rates are constructed as in Figure 1B.

²⁹ I approximate r_{b^*} from the 2000 Census using information on state of residence five years prior. I construct the approximation as follows, $r_{b^*} \approx p(0K4 \mid 0K9) \cdot (0K9 \mid 0K14) \cdot p(0K14 \mid 0K19)$, where OKx is an indicator for whether an individual resided in Oklahoma at age x.

³⁰ Means are calculated from the American Community Survey 2005-2013 using population weights.

³¹ Zip code median income for 1999 is obtained from the 2000 Census.

Where E_j is the 1998-99 grade 1 enrollment in zip code *j*, p_j is the increase in Pre-K enrollment in 1998-99 in zip code j, x_j is the median family income of zip code j, and η reflects the fraction of median family income of UPK compliers. c(x) is the probability an individual will be charged with a crime at age 18-19 given their family income, *x*. It is approximated as,

$$c(x) = c_0 \left(1 + \epsilon \left(\frac{x - x_0}{x_0} \right) \right),$$

Where ϵ is the elasticity of family income with respect to crime and x_0 is the lowest zip code median family income. I estimate ϵ using a probit regression of an indicator for any criminal charge at age 18-19 on family income (at or before age 17) using black southerners in the National Longitudinal Survey of Youth 1997. This yields $\hat{\epsilon} = -0.41$.

I calculate
$$\theta(\eta) = \frac{\tilde{Y}_B^R(\eta)}{Y_B^R}$$
 using the Common Core of Data and the 2000 Census. I use this ratio, an approximation of *r* from the 2000 Census and an approximation of $\overline{Y^R} = \frac{N^B}{N^R} * \overline{Y}$ to

rewrite Equation 5 as,

$$\tilde{Y}_B^R(\eta,\alpha) = \theta(\eta) \cdot \overline{Y_B^R}(\alpha) = \theta(\eta) \cdot \frac{N^B}{N^R} \cdot \frac{\bar{Y}}{(r+\alpha(1-r))},$$
(6)

Where \overline{Y} , $\frac{N^B}{N^R}$, and *r* are observed or approximated. Figure 9 shows the TOT effect as percentage of $\widetilde{Y}_B^R(\eta, \alpha)$ for various choices of η and α .

Figure 1A: Total Pre-K Enrollment in Oklahoma Public Schools



Notes: Oklahoma statewide Pre-K enrollment obtained by aggregating data from the School Universe Survey of the Common Core of Data produced by the National Center for Education Statistics.

Figure 1B: Percent Pre-K Enrollment in Oklahoma Public Schools, by Race



1998-99, therefore it is imputed for all years by multiplying Pre-K enrollment in each school by the fraction of school enrollment of a produced by the National Center for Education Statistics. % Pre-K enrollment is defined as the aggregate state Pre-K enrollment in Notes: School-level Pre-K enrollment for Oklahoma is obtained from the School Universe Survey of the Common Core of Data year t divided by the aggregate state Grade 1 enrollment in t+2. Pre-K enrollment by grade and race is not available prior to given race. These enrollment counts are then aggregated to the state level. Figure 1C: Percent Pre-K Access in Oklahoma Public Schools, by Race



Notes: School-level Pre-K enrollment for Oklahoma is obtained from the School Universe Survey of the Common Core of Data produced by the National Center for Education Statistics. % Pre-K access is defined as the aggregate state Pre-K enrollment in schools with Pre-K enrollment greater than zero in year t divided by the aggregate state Grade 1 enrollment in t+2.





Pre-K enrollment for Oklahoma is obtained from the School Universe Survey of the Common Core of Data produced by the National Center for Education Statistics. % Pre-K enrollment is defined as the aggregate zip code Pre-K enrollment in year t divided by the where the school building zip code is the unit of observation. The bottom panel shows the distribution zip code %FRL. School-level aggregate zip code Grade 1 enrollment in t+2. Zip code %FRL is defined as the total FRL students divided by the total enrollment Notes: The top panel depicts a locally weighted regression of %Pre-K enrollment on %FRL (seperately for 1997-98 and 1998-99), in zip code in 1998-99.





code median family income in 1999 from the 2000 Census. % Pre-K enrollment is defined as the aggregate zip code Pre-K enrollment code %FRL. School-level Pre-K enrollment for Oklahoma is obtained from the School Universe Survey of the Common Core of Data produced by the National Center for Education Statistics. Enrollment information, aggregated to the zip code-level is merged to zip Notes: The top panel depicts a locally weighted regression of %Pre-K enrollment on zip code median family income (seperately for 1997-98 and 1998-99), where the school building zip code is the unit of observation. The bottom panel shows the distribution zip in year t divided by the aggregate zip code Grade 1 enrollment in t+2.

School-level Pre-K enrollment for Oklahoma is obtained from the School Universe Survey of the Common Core of Data produced by 1998-99), where the school building zip code is the unit of observation. The bottom panel shows the distribution zip code %FRL. Notes: The top panel depicts a locally weighted regression of an indicator of Pre-K access on %FRL (seperately for 1997-98 and the National Center for Education Statistics. Zip code Pre-K access is defined as one if the Pre-K enrollment in the zip code is greater than zero. Zip code %FRL is defined as the total FRL students divided by the total enrollment in zip code in 1998-99.



Figure 3A: Percent Pre-K Access in Oklahoma Public Schools,





the National Center for Education Statistics. Enrollment information, aggregated to the zip code-level is merged to zip code median School-level Pre-K enrollment for Oklahoma is obtained from the School Universe Survey of the Common Core of Data produced by 1998-99), where the school building zip code is the unit of observation. The bottom panel shows the distribution zip code %FRL. Notes: The top panel depicts a locally weighted regression of an indicator of Pre-K access on %FRL (seperately for 1997-98 and family income in 1999 from the 2000 Census. Zip code Pre-K access is defined as one if the Pre-K enrollment in the zip code is greater than zero.
Figure 4: Fraction of Oklahoma Four Year Olds by Preschool Type (October CPS)



Notes: Data from October Current Population Survey. The survey is collected in October of each year, so 1998 falls in the 1998-99 school year (UPK Year 1). The fraction of four year olds attending public preschool is allocated to head start and Pre-K based on each program's share of the statewide total enrollment both programs in the relevant years.





Likelihood of Criminal Charge at Age 18-19 = Number of individuals charged with a crime at age 18-19 in OK with given birthday / Number of individuals born on that day in OK. Notes: Data from Oklahoma criminal court records and birth records. Each dot corresponds to the mean for a bin of 15 days.

Figure 6: Likelihood of Criminal Charge at Age 18-19, by Birthdate (Blacks)



Likelihood of Criminal Charge at Age 18-19 = Number of individuals charged with a crime at age 18-19 in OK with given birthday / Number of individuals born on that day in OK. Notes: Data from Oklahoma criminal court records and birth records. Each dot corresponds to the mean for a bin of 15 days.



Notes: Data from Oklahoma birth records. Each dot corresponds to the mean for a bin of 15 days.

Figure 8: Distribution of Black Grade 1 Population and Black UPK Compliers, by Zip Code Black Median Income







Screenshot
Records
Court
Demand
On
A1:
Figure

On Demé	and	Court Rec	ords			Pricing Log	n Sign Up
							New search
Case Info STATE Case Identi Type of Cas Date Filed Amount Ow	ormation E OF O tiffer Alfalf sse Crimi 03/26 wed \$0.00	httanta vs fa OK iinal Felony Proceedi 6/2009 0 (as of 11/12/2014 fm-	A Montor this case lings		Offense or Cause UNLAWFUL POSSESSION OF MARIHU INTENT TO DISTRIBUTE	JANA WITH	
			Name				
Parties Involv	/ed		Date of Birth				
Agency	QK	KLAHOMA HIGHWAY F	Address		cription	mount	
Officer	WA	ALLACE, R G			DRMATION FILED (1) 3	103.00	
Attorney	GU	INNINGHAM, RICK of ,	Phone		/ LIBRARY FEE	00	4
Judge	AN	IGLE, LOREN E.	Description		EST FEE TO SHERIFF	5.00	۵,
DA	Ę	HMANN, DANNY G. of			ET ASSESSMENT		RACTS
Defendant				Monitor this person	AUTO FINGER PRINTING INFO SYSTEM	SIGN SIGN	8
					CRIME VICTIMS COMPUTATION	45.00	
Calandar aver	unte				STATE TREASURERS FORENSIC FEE	5:00	Ę
Date	Time	Description			10% OF AFS5	150 FA	
03/27/2009		INITIAL APPEARANCE	E Completed : 03/27/2009 Code: X		10% of CLT9	06 C	100
04/08/2009	9:30am	FELONY DOCKET CAL	VLL Completed : 04/08/2009 Code: X		10% OF FORE	3.50 AU	
05/06/2009	9:30am	FELONY DOCKET CAL	VLL Completed : 05/06/2009 Code: X		10% OF VCA1	4.50 GETAF	RE REIAL V
					MEDICAL EXPENSE LIABILITY REV FUND FEE	10.00	and the second

\$10.00

MEDICAL EXPENSE LIABILITY REV FUND FEE

07/01/2009 9:30am FELONY DOCKET CALL Completed : 07/01/2009 Code: X

Figure A2: Oklahoma State Courts Network Screenshot

WWW PACE-ARE	Help	ess of this information. Verify all information with the official record keeper. The information contained in this report is provided in comp	IN THE DISTRICT COURT IN AND F
E COU	alendar	s or completen	
. ГАНОМА БТАТ	Legal Research 0	Do not rely on the correctnes	
	Court Dockets	pe is NOT an official record	
SC	Courts	on on this pag	
P	Home	te informati	

	IN THE DISTRICT COURT IN AND	FOR ELLIS COUNTY, OKLAHOMA	
		No. CF 2013-1 (Criminal Felony)	
State of Oklahoma v.		Eiket: 01/02/2013 Ctored: 05/07/2013	
		Judge: Jackson, Joe L.	
Parties			
. Defendant STATE OF OKLAHOMA , Plaintiff			
Attorneys			
Attorney		Represented Parties	
Events			
Event Thursday, January 10, 2013 at 9:00 AM Attornety DataFiFEA)	Part	٩	Docket
Thursday, February 7, 2013 at 10:00 AM Per LiMiukary HEARING CONFERENCE/PRELIMC) Wednesday, February 20, 2013 at 13:30 PM HEARING(HEA)		1	Joe L. Jackson
Counts			
Paties appear only under the counts with which they were charged. For complete sentence information, see	ee the court minute on the docket.		
Count # 1.	Count as Filed: LAS, Larceny of Automobile with Supplemental Informatic Date of Offense: 12/31/2012	on For Prior Felony Conviction , in violation of 21 O.S. 1720	
Darty Namo:	Dienosition Information:		

Count # 1.		Court as Filed: LMS, Larceny of Automobile with Supplemental Information For Prior Felony Conviction , in violation of <u>21 O.S. 1720</u> Date Of Offense: 12/31/2012
	Party Name:	Disposition. Information:
		Disposed: COMVCTION, 0X97/2013, Guilty Plea. Court as Descenduration of Automobie with Supplemental Information For Prior Felony Conviction (LAS) Voltation de 2005, TCS
Count # 2.		Count as Filed. FE1A, Euding/ATTEMPT TO ELUDE Police OFFICER , in violation of <u>21.0.5. 540A (B).</u> Date Of Offense: 1231/2012
	Party Name:	Disposition Information:
		Dispossed: CONVICTION, 03/07/2013. Guilty Plea. Court as DiscosedEludino/ATTEMPT TO ELUDE Police OFFICER (FE 1A)

SOLVER CALLING A STATE CURRENT OF A CONTRACT OF A CONTRACT. OF A CONTRACT OF A CONTRAC

Record:			Alias or Alternate Names:	
			None Found.	
ersonal Profile [°]				
decord Date	Marital Status	Birth Date		Birth City
March 26, 2007 at 2:52:26 PM				
hysical Profile [°]				
100				

BirthPlace

Jecord Date	Hair	ye	Sex	Race	kin	Weight	Height	Blood Type
ugust 4, 2009 at 1:48:17 PM			Male	White	•	150	6.7	
ddress Information								
secord Date	Status	ſ	Ape		Address			
anuary 2, 2013 at 11:44:29 AM	Current		Home Address					
anuary 2, 2013 at 11:44-29 AM	1		Home Address					
october 2, 2012 at 11:58:52 AM	'		Home Address					
eptember 18, 2012 at 10:1:46 AM	'		Home Address					
ugust 20, 2012 at 10:47:35 AM	'		Home Address					
uy 5, 2012 at 9:36:32 AM	•	-	Home Address					
elephone Information								
ecord Date	Type					Number		
October 2, 2012 at 11:58:52 AM	Mobil	e Phone N	umber					
billet months and have been actived monthing. Becaute off months and and and active distribution and months active first								

Figure A3: Distribution of White Name Prediction Index by Observed Race



Note: White Name Prediction Index is the predicted probability from a probit regression with an indicator for white as the dependent variable and cubics of the fraction of each race observed with the same first name and cubics of the fraction of each race observed with the same last name.



dependent variable and cubics of the fraction of each race observed with the same first name and cubics of the fraction of each race Note: Black Name Prediction Index is the predicted probability from a probit regression with an indicator for black as the observed with the same last name.





school year (UPK Year 1). The fraction of four year olds attending public preschool includes both Head Start and State Pre-K. The shaded area depicts the 95% confidence interval. Population weights are used. Notes: Data from October Current Population Survey. The survey is collected in October of each year, so 1998 falls in the 1998-99

Figure A6: Percent of Oklahoma-born 18 and 19-Year Olds Residing in Oklahoma American Community Survey 2005-2013



1994q3

1994q1

1993q3 Quarter of Birth

1993q1

1992q3

1992q1

0

for White Felony and Misdemeanor Charges, by Kernel Bandwidth (in Days) Figure A7: Local Linear Regression Discontinuity Estimates



Figure A8: Local Linear Regression Discontinuity Estimates for Black Felony and Misdemeanor Charges, by Kernel Bandwidth (in Days)







Likelihood of Criminal Charge at Age 18 = Number of individuals charged with a crime at age 18 in OK with given birthday / Number of individuals born on that day in OK. Notes: Data from Oklahoma criminal court records and birth records. Each dot corresponds to the mean for a bin of 15 days.

Figure A10: Likelihood of Criminal Charge at Age 18, by Birthdate (Blacks)



Likelihood of Criminal Charge at Age 18 = Number of individuals charged with a crime at age 18 in OK with given birthday / Number of individuals born on that day in OK. Notes: Data from Oklahoma criminal court records and birth records. Each dot corresponds to the mean for a bin of 15 days.

		Including Cha Imputed	arges with Race	Excluding Ch Imputed	arges with Race
	Full Sample (1)	White (2)	Black (3)	White (4)	Black (5)
Total Births	46,217	36,134	4,904	36,134	4,904
Percent of Total Births		78.2%	10.6%	78.2%	10.6%
Charged at Age 18-19 (Unique	Individuals):				
Misdemeanor Only					
Total	2,517	1,724	867	720	260
Percent of Births	5.4%	4.8%	17.7%	2.0%	5.3%
Felony Only					
Total	2,005	1,169	831	514	410
Percent of Births	4.3%	3.2%	16.9%	1.4%	8.4%
Misdemeanor or Felony					
Total	4,005	2,579	1,501	1,107	587
Percent of Births	8.7%	7.1%	30.6%	3.1%	12.0%
Charged at Age 18 (Unique Ind	lividuals)				
Misdemeanor Only					
Total	1,214	827	424	321	112
Percent of Births	2.6%	2.3%	8.6%	0.9%	2.3%
Felony Only					
Total	1,058	614	442	267	233
Percent of Births	2.3%	1.7%	9.0%	0.7%	4.8%
Misdemeanor or Felony					
Total	2,073	1,325	792	544	310
Percent of Births	4.5%	3.7%	16.2%	1.5%	6.3%

Mother's race is used to define race subsamples of births.

For charged individuals missing race, race is imputed using race name index thresholds of 0.9.

	1997-98		199	1998-99		2000
	Child A	Child B	Child C	Child D	Child E	Child F
Birthdate	9/1/1992	9/2/1992	9/1/1993	9/2/1993	9/1/1994	9/2/1994
Age Relative to Others in Grade	Youngest	Oldest	Youngest	Oldest	Youngest	Oldest
Grade at Age 4	-	-	-	-	РК	-
Grade at Age 5	Κ	-	К	PK	K	РК
Grade at Age 6	1	Κ	1	K	1	Κ
Criminal Charges Observed in Sample	<i>e</i> :					
Age 18	yes	yes	yes	yes	yes	yes
Age 19	yes	yes	yes	yes	no	no
				$\rightarrow OK Unive$	ersal Pre-K	

Table 2 - Treatment Contrasts at Kindergarten Birthdate Cutoffs in Various School Years

	WHITE		BL	ACK	
	Felony	Misdemeanor	Felony	Misdemeanor	
	(1)	(2)	(3)	(4)	
1997-1998 School Year (Birthdates: 3/2/1991-3/1/1993)					
Missed Kindergarten	0.005	-0.006	-0.016	-0.010	
-	(0.00)	(0.00)	(0.02)	(0.02)	
Obs	365	365	365	365	
1998-1999 School Year [FIRST YEAR OF U	UNIVERSAL PRE	-K]			
(Birthdates: 3/2/1993-3/1/1994)					
Missed Kindergarten	-0.001	0.003	-0.045 **	-0.068 ***	
-	(0.00)	(0.00)	(0.02)	(0.02)	
Obs	365	365	365	365	
Avg. Fraction Charged (by Birthdate)	0.03	0.06	0.17	0.18	
Avg. Daily Births (Denominator)	99.0	99.0	13.4	13.4	

Table 3 - RD Estimates of Missing Kindergarten Birthdate Cutoff by School Year

OUTCOME: Likelihood of Criminal Charge at Age 18-19

Each entry represents the estimate of being born after the birthdate cutoff (Sept. 1) in the given year from a different RD regression.

Regression: $y = a + b^*(z \ge 0) + c^*z^*(z \ge 0) + d^*z^*(z < 0)$, where z is birthdate minus 9/1/YYYY and b is the coefficient of interest.

y = # of individuals charged with a crim at age 18 in OK with given birthday / # individuals born on that day in OK. Robust standard errors in parentheses.

	WHITE		BLAC	CK
	Felony	Misdemeanor	Felony	Misdemeanor
	(1)	(2)	(3)	(4)
andwidth: +/- 6 Months (Birthdates: 3/2/1993-3/1/1994)				
Missed Kindergarten	-0.001 (0.00)	0.004 (0.00)	-0.045 ** (0.02)	-0.068 *** (0.02)
Obs	365	365	365	365
andwidth: +/- 5 Months (Birthdates: 4/1/1993-2/1/1994)				
Missed Kindergarten	-0.002 (0.00)	0.003 (0.00)	-0.054 ** (0.02)	-0.081 *** (0.03)
Obs	307	307	307	307
andwidth: +/- 4 Months (Birthdates: 5/1/1993-1/1/1994)				
Missed Kindergarten	-0.004 (0.00)	0.001 (0.01)	-0.067 *** (0.03)	-0.081 *** (0.03)
Obs	246	246	246	246
andwidth: +/- 3 Months (Birthdates: 6/1/1993-12/1/1993)				
Missed Kindergarten	-0.005 (0.00)	-0.003 (0.01)	-0.047 (0.03)	-0.087 *** (0.03)
Obs	184	184	184	184
andwidth: +/- 2 Months (Birthdates: 7/1/1993-11/1/1993)				
Missed Kindergarten	-0.003 (0.01)	-0.011 (0.01)	-0.057 * (0.03)	-0.156 *** (0.04)
Obs	124	124	124	124
andwidth: +/- 1 Months (Birthdates: 8/1/1993-10/1/1993)				
Missed Kindergarten	-0.002 (0.01)	0.001 (0.01)	-0.051 (0.05)	-0.109 * (0.06)
Obs	62	62	62	62

Table 4 - RD Estimates of Missing Kindergarten Birthdate Cutoff in UPK Year 1 (Various Bandwidths) OUTCOME: Likelihood of Criminal Charge at Age 18-19

Each entry represents the estimate of being born just after the birthdate cutoff (Sept. 1) in the first implementation year from a different RD regression. Regression: $y = a + b*(z \ge 0) + c*z*(z \ge 0) + d*z*(z < 0)$, where z is birthdate minus 9/1/1993 and b is the coefficient of interest.

y = # of individuals charged with a crime at age 18-19 in OK with given birthday / # individuals born on that day in OK.

	W	HITE	BL	АСК
	Felony	Misdemeanor	Felony	Misdemeanor
	(1)	(2)	(3)	(4)
Kernel Bandwidth (Days)				
4	-0.005	-0.001	0.167 **	-0.376 ***
	(0.01)	(0.01)	(0.08)	(0.05)
10	0.024 *	0.010	0.106	-0.164
	(0.01)	(0.01)	(0.06)	(0.10)
20	0.016	0.005	0.032	-0.117
	(0.01)	(0.01)	(0.05)	(0.08)
30	0.006	0.000	0.003	-0.128 **
	(0.01)	(0.01)	(0.05)	(0.06)
40	0.002	-0.001	-0.033	-0.110 **
	(0.01)	(0.01)	(0.05)	(0.05)
50	0.000	-0.003	-0.045	-0.109 **
	(0.01)	(0.01)	(0.04)	(0.05)
60	-0.001	-0.005	-0.043	-0.119 ***
	(0.01)	(0.01)	(0.04)	(0.04)
Ontineal Dandwidth.	1 10	129	5.24	5 47
Optimal Banawiain:	4.19	4.30	5.24	3.4/
	-0.005	0.003	0.144 **	-0.354 ***
	(0.01)	(0.01)	(0.07)	(0.06)

 Table 5 - Local Linear RD Estimates of Missing Kindergarten Birthdate Cutoff in UPK Year 1

 OUTCOME: Likelihood of Criminal Charge at Age 18-19

Each entry represents the estimate of being born just after the birthdate cutoff (Sept. 1) in the first implementation year from a different RD regression. Regression: $y = a + b^*(z \ge 0) + f(z)$, where z is birthdate minus 9/1/1993 and b is the coefficient of interest.

Estimation uses a local linear (kernel regression) approach (Nichols, 2011).

y = # of individuals charged with a crime at age 18-19 in OK with given birthday / # individuals born on that day in OK.

Optimal bandwidth is calculated following Imbens and Kalyanaraman (2011). Standard errors in parentheses.

	J J	WHITE	F	BLACK
	Felony	Misdemeanor	Felony	Misdemeanor
	(1)	(2)	(3)	(4)
A) UPK Year 1 and Prior Year [Charges at (Birthdates: 3/2/1992-3/1/1994)	Age 18-19]			
Missed Kindergarten x UPK Year 1	-0.006	0.009	-0.028	-0.057 *
	(0.01)	(0.01)	(0.03)	(0.03)
Avg. Daily Fraction Charged	0.03	0.05	0.17	0.18
Obs	730	730	730	730
B) UPK Years 1-2 and Prior Year [Charge a (Birthdates: 3/2/1992-3/1/1995)	it Age 18]			
Missed Kindergarten x UPK Year 1	0.001	0.005	0.010	-0.019
	(0.00)	(0.00)	(0.02)	(0.02)
Avg. Daily Fraction Charged	0.02	0.02	0.09	0.09
Obs	1,095	1,095	1,095	1,095
C) UPK Year 1 and Prior Year [Charge at A (Birthdates: 3/2/1992-3/1/1994)	Age 18]			
Missed Kindergarten x UPK Year 1	-0.001	0.004	0.017	-0.017
	(0.00)	(0.00)	(0.02)	(0.02)
Avg. Daily Fraction Charged	0.02	0.02	0.09	0.09
Obs	730	730	730	730
D) UPK Years 1-2 [Charge at Age 18] (Birthdates: 3/2/1993-3/1/1995)				
Missed Kindergarten x UPK Year 1	0.002	0.006	0.003	-0.022
	(0.00)	(0.00)	(0.02)	(0.02)
Avg. Daily Fraction Charged	0.02	0.02	0.09	0.09
Obs	730	730	730	730

Table 6 - RD Estimates of Missing Kindergarten Birthdate Cutoff in UPK Year 1 vs. Other Years Outcome Measure: Likelihood of Criminal Charge at Age 18 or Age 18-19

Each entry represents the estimate of being born just after the birthdate cutoff (Sept. 1) in the first implementation year

(relative to other years) from a different RD regression.

Regression: $y = a + b1^{*}(YEAR 1)^{*}(z \ge 0) + b2^{*}(z \ge 0) + c^{*}(YEAR 1)^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z \ge 0)$... (other interaction terms),

where z is birthdate minus 9/1/YYYY and b is the coefficient of interest.

y = # of individuals charged with a crime at age 18 in OK with given birthday / # individuals born on that day in OK.

			FEL	ONY		MISDEMEANOR						
	Not Imputed	$WNI \ge 0.1$	$WNI \ge 0.25$	$WNI \ge 0.5$	$WNI \ge 0.75$	$WNI \ge 0.9$	Not Imputed	$WNI \ge 0.1$	$WNI \ge 0.25$	$WNI \ge 0.5$	$WNI \ge 0.75$	$WNI \ge 0.9$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Bandwidth: +/- 6 Months	-0.002	-0.002	-0.002	-0.002	-0.001	-0.001	0.003	0.003	0.003	0.005	0.006	0.004
(Birthdates: 3/2/1993-3/1/1994)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Bandwidth: +/- 5 Months	-0.002	-0.004	-0.004	-0.004	-0.003	-0.002	0.004	0.004	0.003	0.005	0.006	0.003
(Birthdates: 4/1/1993-2/1/1994)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Bandwidth: +/- 4 Months	-0.003	-0.006	-0.006	-0.007	-0.006	-0.004	0.002	0.003	0.002	0.004	0.003	0.001
(Birthdates: 5/1/1993-1/1/1994)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Bandwidth: +/- 3 Months	-0.005	-0.007	-0.007	-0.008	-0.006	-0.005	0.003	-0.002	-0.002	-0.001	-0.001	-0.003
(Birthdates: 6/1/1993-12/1/1993)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Bandwidth: +/- 2 Months	-0.003	-0.003	-0.002	-0.003	-0.003	-0.003	0.003	-0.008	-0.009	-0.008	-0.009	-0.011
(Birthdates: 7/1/1993-11/1/1993)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Bandwidth: +/- 1 Months	-0.005	0.002	0.002	0.000	-0.001	-0.002	0.008	0.005	0.004	0.004	0.004	0.001
(Birthdates: 8/1/1993-10/1/1993)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)

See text for construction of White Name Index (WNI).

Regression: $y = a + b^*(z \ge 0) + c^*z^*(z \ge 0) + d^*z^*(z < 0)$, where z is birthdate minus 9/1/1993 and b is the coefficient of interest.

y = # of individuals charged with a crime at age 18-19 in OK with given birthday / # individuals born on that day in OK.

Table A2 - Black Sample Robustness Checks for RD Estimates in UPK Year 1 (Likelihood of Criminal Charge at Age 18-19)

	FELONY							MISDEMEANOR					
	Not Imputed $BNI \ge 0.1$ $BNI \ge 0.25$ $BNI \ge 0.5$ $BNI \ge 0.75$					$BNI \ge 0.9$	Not Imputed	$BNI \ge 0.1$	$BNI \ge 0.25$	$BNI \ge 0.5$	$BNI \ge 0.75$	$BNI \ge 0.9$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	
Bandwidth: +/- 6 Months	-0.035 **	-0.059 **	-0.045 *	-0.042 *	-0.041 *	-0.045 **	-0.035 ***	-0.076 ***	-0.094 ***	-0.081 ***	-0.067 ***	-0.068 ***	
(Birthdates: 3/2/1993-3/1/1994)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	
Bandwidth: +/- 5 Months	-0.043 **	-0.075 ***	-0.056 **	-0.049 **	-0.051 **	-0.054 **	-0.042 ***	-0.093 ***	-0.110 ***	-0.101 ***	-0.083 ***	-0.081 ***	
(Birthdates: 4/1/1993-2/1/1994)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	
Bandwidth: +/- 4 Months	-0.054 ***	-0.088 ***	-0.067 **	-0.063 **	-0.066 **	-0.067 ***	-0.042 ***	-0.086 ***	-0.103 ***	-0.095 ***	-0.080 ***	-0.081 ***	
(Birthdates: 5/1/1993-1/1/1994)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	
Bandwidth: +/- 3 Months	-0.037 *	-0.061 *	-0.045	-0.042	-0.046	-0.047	-0.035 *	-0.084 **	-0.105 ***	-0.096 ***	-0.085 **	-0.087 ***	
(Birthdates: 6/1/1993-12/1/1993)	(0.02)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	
Bandwidth: +/- 2 Months	-0.053 **	-0.057	-0.051	-0.049	-0.063 *	-0.057 *	-0.063 ***	-0.147 ***	-0.162 ***	-0.160 ***	-0.159 ***	-0.156 ***	
(Birthdates: 7/1/1993-11/1/1993)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.02)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	
Bandwidth: +/- 1 Months	-0.077 **	-0.037	-0.033	-0.027	-0.051	-0.051	-0.081 ***	-0.084	-0.103 *	-0.100 *	-0.105 *	-0.109 *	
(Birthdates: 8/1/1993-10/1/1993)	(0.04)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)	(0.03)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	

See text for construction of Black Name Index (BNI).

Regression: $y = a + b^*(z \ge 0) + c^*z^*(z \ge 0) + d^*z^*(z < 0)$, where z is birthdate minus 9/1/1993 and b is the coefficient of interest.

y = # of individuals charged with a crime at age 18-19 in OK with given birthday / # individuals born on that day in OK.

Table A3 - White Sample Robustness Checks for RD Estimates in UPK Year 1 (Number of Individuals Charged at Age 18-19)

			FEL	ONY		MISDEMEANOR						
	Not Imputed	$WNI \ge 0.1$	$WNI \ge 0.25$	$WNI \ge 0.5$	$WNI \ge 0.75$	$WNI \ge 0.9$	Not Imputed	$WNI \ge 0.1$	$WNI \ge 0.25$	$WNI \ge 0.5$	$WNI \ge 0.75$	$WNI \ge 0.9$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Bandwidth: +/- 6 Months	-0.299	-0.325	-0.356	-0.373	-0.280	-0.188	0.316	0.330	0.301	0.471	0.634	0.348
(Birthdates: 3/2/1993-3/1/1994)	(0.25)	(0.41)	(0.41)	(0.40)	(0.39)	(0.36)	(0.30)	(0.52)	(0.51)	(0.51)	(0.51)	(0.50)
Bandwidth: +/- 5 Months	-0.304	-0.605	-0.637	-0.628	-0.538	-0.358	0.421	0.323	0.291	0.507	0.598	0.279
(Birthdates: 4/1/1993-2/1/1994)	(0.28)	(0.44)	(0.44)	(0.43)	(0.42)	(0.39)	(0.32)	(0.57)	(0.56)	(0.55)	(0.55)	(0.55)
Bandwidth: +/- 4 Months	-0.406	-0.786	-0.781	-0.800	-0.727	-0.516	0.297	0.347	0.303	0.508	0.492	0.168
(Birthdates: 5/1/1993-1/1/1994)	(0.31)	(0.50)	(0.50)	(0.50)	(0.49)	(0.45)	(0.35)	(0.61)	(0.61)	(0.60)	(0.61)	(0.61)
Bandwidth: +/- 3 Months	-0.511	-0.778	-0.787	-0.816	-0.672	-0.542	0.479	0.108	0.057	0.213	0.173	-0.063
(Birthdates: 6/1/1993-12/1/1993)	(0.36)	(0.56)	(0.56)	(0.55)	(0.54)	(0.50)	(0.40)	(0.69)	(0.68)	(0.68)	(0.69)	(0.70)
Bandwidth: +/- 2 Months	-0.420	-0.461	-0.411	-0.489	-0.443	-0.436	0.497	-0.730	-0.791	-0.692	-0.829	-1.018
(Birthdates: 7/1/1993-11/1/1993)	(0.43)	(0.65)	(0.65)	(0.65)	(0.65)	(0.59)	(0.48)	(0.83)	(0.83)	(0.82)	(0.85)	(0.87)
Bandwidth: +/- 1 Months	-0.654	-0.229	-0.229	-0.350	-0.371	-0.541	0.912	0.426	0.281	0.403	0.372	0.043
(Birthdates: 8/1/1993-10/1/1993)	(0.60)	(0.89)	(0.89)	(0.88)	(0.90)	(0.81)	(0.64)	(1.13)	(1.12)	(1.11)	(1.14)	(1.18)

See text for construction of White Name Index (WNI).

Regression: $y = a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$, where z is birthdate minus 9/1/1993 and b is the coefficient of interest.

y = # of individuals charged with a crime at age 18-19 in OK with given birthday.

Table A4 - Black Sample Robustness Checks for RD Estimates in UPK Year 1 (Number of Individuals Charged at Age 18-19)

FELONY							MISDEMEANOR					
Not Imputed	$BNI \ge 0.1$	$BNI \ge 0.25$	$BNI \ge 0.5$	$BNI \ge 0.75$	$BNI \ge 0.9$	Not Imputed	$BNI \ge 0.1$	$BNI \ge 0.25$	$BNI \ge 0.5$	$BNI \ge 0.75$	$BNI \ge 0.9$	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	
-0.441 **	-0.651 *	-0.444	-0.441	-0.434	-0.484 *	-0.544 ***	-1.024 ***	-1.288 ***	-1.158 ***	-0.971 ***	-0.982 ***	
(0.22)	(0.34)	(0.32)	(0.30)	(0.29)	(0.29)	(0.18)	(0.38)	(0.36)	(0.35)	(0.35)	(0.35)	
-0.537 **	-0.854 **	-0.572 *	-0.521	-0.543 *	-0.586 *	-0.633 ***	-1.209 ***	-1.468 ***	-1.387 ***	-1.145 ***	-1.128 ***	
(0.23)	(0.37)	(0.34)	(0.32)	(0.31)	(0.31)	(0.19)	(0.41)	(0.40)	(0.38)	(0.38)	(0.38)	
-0.733 ***	-1.136 ***	-0.817 **	-0.792 **	-0.818 **	-0.834 **	-0.649 ***	-1.210 ***	-1.466 ***	-1.387 ***	-1.171 ***	-1.177 ***	
(0.26)	(0.42)	(0.39)	(0.37)	(0.36)	(0.36)	(0.21)	(0.45)	(0.43)	(0.42)	(0.42)	(0.43)	
-0.560 *	-0.919 *	-0.663	-0.639	-0.672 *	-0.677 *	-0.597 **	-1.383 ***	-1.697 ***	-1.590 ***	-1.420 ***	-1.439 ***	
(0.30)	(0.48)	(0.46)	(0.42)	(0.40)	(0.39)	(0.25)	(0.50)	(0.49)	(0.47)	(0.48)	(0.48)	
-0.742 **	-0.800	-0.666	-0.672	-0.829 *	-0.750	-0.975 ***	-2.223 ***	-2.444 ***	-2.448 ***	-2.408 ***	-2.372 ***	
(0.36)	(0.54)	(0.51)	(0.48)	(0.46)	(0.46)	(0.29)	(0.62)	(0.60)	(0.59)	(0.59)	(0.59)	
-1.071 **	-0.543	-0.489	-0.383	-0.663	-0.663	-1.225 ***	-1.350	-1.657 *	-1.619 *	-1.655 *	-1.706 *	
(0.52)	(0.82)	(0.79)	(0.72)	(0.70)	(0.70)	(0.43)	(0.92)	(0.90)	(0.87)	(0.89)	(0.90)	
	Not Imputed (1) -0.441 ** (0.22) -0.537 ** (0.23) -0.733 *** (0.26) -0.560 * (0.30) -0.742 ** (0.36) -1.071 ** (0.52)	$\begin{tabular}{ c c c c c } \hline Not Imputed & BNI \ge 0.1 \\ \hline (1) & (2) \\ \hline (2) \\ \hline (2) & -0.441 ** & -0.651 * \\ (0.22) & (0.34) \\ \hline (0.22) & (0.34) \\ \hline -0.537 ** & -0.854 ** \\ (0.23) & (0.37) \\ \hline -0.733 *** & -1.136 *** \\ (0.26) & (0.42) \\ \hline -0.560 * & -0.919 * \\ (0.30) & (0.48) \\ \hline -0.742 ** & -0.800 \\ (0.36) & (0.54) \\ \hline -1.071 ** & -0.543 \\ (0.52) & (0.82) \\ \hline \end{tabular}$	$\begin{tabular}{ c c c c c c c c c c c c c c c c c c c$	$\begin{tabular}{ c c c c c c c c c c c c c c c c c c c$	$\begin{tabular}{ c c c c c c c c c c c c c c c c c c c$	FELONY Not Imputed BNI ≥ 0.1 BNI ≥ 0.25 BNI ≥ 0.5 BNI ≥ 0.75 BNI ≥ 0.9 (1) (2) (3) (4) (5) (6) -0.441 ** -0.651 * -0.444 -0.441 -0.434 -0.484 * (0.22) (0.34) (0.32) (0.30) (0.29) (0.29) -0.557 ** -0.854 ** -0.572 * -0.521 -0.543 * -0.586 * (0.23) (0.37) (0.34) (0.32) (0.31) (0.31) -0.733 *** -1.136 *** -0.817 ** -0.792 ** -0.818 ** -0.834 *** (0.26) (0.42) (0.39) (0.37) (0.36) (0.36) -0.560 * -0.919 * -0.663 -0.639 -0.672 * -0.677 * (0.30) (0.48) (0.46) (0.42) (0.40) (0.39) -0.742 ** -0.800 -0.666 -0.672 -0.829 * -0.750 (0.36) (0.54) (0.51) (0.48) (0.46)	FELONY Not Imputed BNI ≥ 0.1 BNI ≥ 0.25 BNI ≥ 0.5 BNI ≥ 0.75 BNI ≥ 0.9 Not Imputed (1) (2) (3) (4) (5) (6) (7) -0.441 ** -0.651 * -0.444 -0.441 -0.434 -0.484 * -0.544 *** (0.22) (0.34) (0.32) (0.30) (0.29) (0.29) (0.18) -0.537 ** -0.854 ** -0.572 * -0.521 -0.543 * -0.586 * -0.633 *** (0.23) (0.37) (0.34) (0.32) (0.31) (0.31) (0.19) -0.733 *** -1.136 *** -0.817 ** -0.792 ** -0.818 ** -0.834 ** -0.649 *** (0.26) (0.42) (0.39) (0.37) (0.36) (0.21) -0.597 ** -0.560 * -0.919 * -0.663 -0.639 -0.672 * -0.677 * -0.597 ** (0.30) (0.48) (0.46) (0.42) (0.40) (0.39) (0.25) -0.742 **	FELONY Not Imputed BNI ≥ 0.1 BNI ≥ 0.25 BNI ≥ 0.5 BNI ≥ 0.75 BNI ≥ 0.9 Not Imputed BNI ≥ 0.1 (1) (2) (3) (4) (5) (6) (7) (8) -0.441 ** -0.651 * -0.444 -0.441 -0.434 -0.484 * -0.544 *** -1.024 *** (0.22) (0.34) (0.32) (0.30) (0.29) (0.29) (0.18) (0.38) -0.537 ** -0.854 ** -0.572 * -0.521 -0.543 * -0.586 * -0.633 *** -1.209 *** (0.23) (0.37) (0.34) (0.32) (0.31) (0.31) (0.19) (0.41) -0.733 *** -1.136 *** -0.817 ** -0.792 ** -0.818 ** -0.834 ** -0.649 *** -1.210 *** (0.26) (0.42) (0.39) (0.37) (0.36) (0.21) (0.45) -0.560 * -0.919 * -0.663 -0.639 -0.672 * -0.677 * -0.597 ** -1.383 ***	HELONY MISDEM Not Imputed BNI ≥ 0.1 BNI ≥ 0.25 BNI ≥ 0.5 BNI ≥ 0.75 BNI ≥ 0.9 Not Imputed BNI ≥ 0.1 BNI ≥ 0.25 (1) (2) (3) (4) (5) (6) (7) (8) (9) -0.441 ** -0.651 * -0.444 -0.441 -0.434 -0.484 * -1.024 *** -1.288 *** (0.22) (0.34) (0.32) (0.30) (0.29) (0.29) (0.18) (0.38) (0.36) -0.537 ** -0.854 ** -0.572 * -0.521 -0.543 * -0.586 * -0.633 *** -1.209 *** -1.468 *** (0.23) (0.37) (0.34) (0.32) (0.31) (0.31) (0.19) (0.41) (0.40) -0.733 *** -1.136 *** -0.817 ** -0.792 ** -0.818 ** -0.834 ** -0.649 *** -1.210 *** -1.466 *** (0.26) (0.42) (0.39) (0.37) (0.36) (0.36) (0.21) (0.45) (0.43)	HELONY MISDEMEANOR Not Inputed BNI ≥ 0.1 BNI ≥ 0.25 BNI ≥ 0.5 BNI ≥ 0.75 BNI ≥ 0.9 Not Imputed BNI ≥ 0.1 BNI ≥ 0.25	HELONY MISDEMEANOR Not Imputed BNI ≥ 0.1 BNI ≥ 0.25 BNI ≥ 0.5 BNI ≥ 0.75 BNI ≥ 0.9 Not Imputed BNI ≥ 0.1 BNI ≥ 0.5 BNI ≥ 0.75 DOS	

See text for construction of Black Name Index (BNI).

Regression: $y = a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$, where z is birthdate minus 9/1/1993 and b is the coefficient of interest.

y = # of individuals charged with a crime at age 18-19 in OK with given birthday.

	I I	WHITE	E	BLACK
	Felony	Misdemeanor	Felony	Misdemeanor
	(1)	(2)	(3)	(4)
1997-1998 School Year (Birthdates: 3/2/1991-3/1/1993)				
Missed Kindergarten	0.001	-0.001	-0.008	-0.018
	(0.00)	(0.00)	(0.02)	(0.02)
Obs	365	365	365	365
1998-1999 School Year [FIRST YEAR OF (Birthdates: 3/2/1993-3/1/1994)	F UNIVERSAL	LPRE-K]		
Missed Kindergarten	0.001	0.003	0.009	-0.035 **
	(0.00)	(0.00)	(0.02)	(0.02)
Obs	365	365	365	365
1999-2000 School Year (Birthdate Window: 3/2/1994-3/1/1995)				
Missed Kindergarten	-0.002	-0.003	0.006	-0.013
	(0.00)	(0.00)	(0.02)	(0.02)
Obs	365	365	365	365
Avg. Fraction Charged (by Birthdate)	0.02	0.02	0.09	0.09
Avg. Daily Births (Denominator)	99.0	99.0	13.4	13.4

Table A5 - RD Estimates of Missing Kindergarten Birthdate Cutoff by School Year

OUTCOME: Likelihood of Criminal Charge at Age 18

Each entry represents the estimate of being born after the birthdate cutoff (Sept. 1) in the given year from a different RD regression.

Regression: $y = a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$, where z is birthdate minus 9/1/YYYY and b is the coefficient of interest.

y = # of individuals charged with a crim at age 18 in OK with given birthday / # individuals born on that day in OK.

			FELO	ONY		MISDEMEANOR						
	Not Imputed	$WNI \ge 0.1$	$WNI \ge 0.25$	$WNI \ge 0.5$	$WNI \ge 0.75$	$WNI \ge 0.9$	Not Imputed	$WNI \ge 0.1$	$WNI \ge 0.25$	$WNI \ge 0.5$	$WNI \ge 0.75$	$WNI \ge 0.9$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Bandwidth: +/- 6 Months	-0.001	0.000	0.000	-0.001	0.000	0.001	0.004 **	0.003	0.002	0.003	0.004	0.003
(Birthdates: 3/2/1993-3/1/1994)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Bandwidth: +/- 5 Months	-0.002	-0.002	-0.002	-0.003	-0.002	-0.001	0.005 **	0.003	0.003	0.004	0.004	0.004
(Birthdates: 4/1/1993-2/1/1994)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Bandwidth: +/- 4 Months	-0.003	-0.004	-0.004	-0.004	-0.004	-0.002	0.003	0.005	0.005	0.005	0.005	0.005
(Birthdates: 5/1/1993-1/1/1994)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Bandwidth: +/- 3 Months	-0.002	-0.002	-0.002	-0.003	-0.001	0.000	0.005 *	0.004	0.004	0.005	0.005	0.005
(Birthdates: 6/1/1993-12/1/1993)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Bandwidth: +/- 2 Months	0.000	0.004	0.004	0.004	0.005	0.004	0.005	-0.004	-0.004	-0.004	-0.004	-0.004
(Birthdates: 7/1/1993-11/1/1993)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Bandwidth: +/- 1 Months	-0.001	0.009	0.009	0.009	0.009	0.007	0.003	-0.002	-0.002	-0.001	-0.002	-0.004
(Birthdates: 8/1/1993-10/1/1993)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)

See text for construction of White Name Index (WNI).

Regression: $y = a + b^*(z \ge 0) + c^*z^*(z \ge 0) + d^*z^*(z < 0)$, where z is birthdate minus 9/1/1993 and b is the coefficient of interest.

y = # of individuals charged with a crime at age 18-19 in OK with given birthday / # individuals born on that day in OK.

 Table A7 - Black Sample Robustness Checks for RD Estimates in UPK Year 1 (Likelihood of Criminal Charge at Age 18)

			FEL	ONY		MISDEMEANOR						
	Not Imputed	$BNI \ge 0.1$	$BNI \ge 0.25$	$BNI \ge 0.5$	$BNI \ge 0.75$	$BNI \ge 0.9$	Not Imputed	$BNI \ge 0.1$	$BNI \ge 0.25$	$BNI \ge 0.5$	$BNI \ge 0.75$	$BNI \ge 0.9$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Bandwidth: +/- 6 Months	-0.003	-0.002	0.010	0.010	0.011	0.009	-0.016 *	-0.042 **	-0.046 ***	-0.041 **	-0.033 **	-0.035 **
(Birthdates: 3/2/1993-3/1/1994)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Bandwidth: +/- 5 Months	-0.004	-0.003	0.010	0.011	0.013	0.012	-0.022 **	-0.045 **	-0.048 **	-0.044 **	-0.035 *	-0.037 **
(Birthdates: 4/1/1993-2/1/1994)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Bandwidth: +/- 4 Months	-0.009	-0.002	0.014	0.014	0.014	0.014	-0.021 *	-0.026	-0.027	-0.027	-0.020	-0.023
(Birthdates: 5/1/1993-1/1/1994)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Bandwidth: +/- 3 Months	0.010	0.023	0.033	0.036	0.038 *	0.037 *	-0.012	-0.026	-0.029	-0.025	-0.020	-0.022
(Birthdates: 6/1/1993-12/1/1993)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Bandwidth: +/- 2 Months	0.010	0.043	0.043	0.044	0.038	0.042 *	-0.019	-0.057 *	-0.060 **	-0.057 **	-0.061 **	-0.060 **
(Birthdates: 7/1/1993-11/1/1993)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.01)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Bandwidth: +/- 1 Months	-0.004	0.067	0.068 *	0.075 **	0.067 *	0.067 *	-0.026	-0.009	-0.015	-0.020	-0.022	-0.022
(Birthdates: 8/1/1993-10/1/1993)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.02)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)

See text for construction of Black Name Index (BNI).

Regression: $y = a + b^*(z \ge 0) + c^*z^*(z \ge 0) + d^*z^*(z < 0)$, where z is birthdate minus 9/1/1993 and b is the coefficient of interest.

y = # of individuals charged with a crime at age 18-19 in OK with given birthday / # individuals born on that day in OK.

	Mother	's Race
	White	Black
	(1)	(2)
Bandwidth: +/- 6 Months (Birthdates: 3/2/1993-3/1/1994)		
Missed Kindergarten	-1.904 (3.89)	0.734 (0.47)
Daily Mean Obs	99.0 365	13.4 365
Bandwidth: +/- 5 Months (Birthdates: 4/1/1993-2/1/1994)		
Missed Kindergarten	-3.180 (4.22)	0.870 * (0.51)
Daily Mean Obs	99.2 307	13.3 307
Bandwidth: +/- 4 Months (Birthdates: 5/1/1993-1/1/1994)		
Missed Kindergarten	-1.603 (4.77)	0.396 (0.60)
Daily Mean Obs	99.7 246	13.4 246
Bandwidth: +/- 3 Months (Birthdates: 6/1/1993-12/1/1993)		
Missed Kindergarten	1.822 (5.46)	-0.211 (0.69)
Daily Mean Obs	100.8 184	13.4 184
Bandwidth: +/- 2 Months (Birthdates: 7/1/1993-11/1/1993)		
Missed Kindergarten	0.508 (6.68)	0.107 (0.88)
Daily Mean Obs	101.4 124	13.5 124
Bandwidth: +/- 1 Months (Birthdates: 8/1/1993-10/1/1993)		
Missed Kindergarten	-3.027 (9.41)	-0.046 (1.33)
Daily Mean Obs	102.6 62	13.7 62

Table A8 -	RD Estimates	of Effect of Mi	ssing Kinderg	arten Birthdate	Cutoff (UPK	Year 1) on	Number of Births
------------	---------------------	-----------------	---------------	-----------------	-------------	------------	------------------

Each entry represents the estimate of being born just after the birthdate cutoff (Sept. 1) in the first implementation year from a different RD regression. Regression: $y = a + b^*(z \ge 0) + c^*z^*(z \ge 0) + d^*z^*(z < 0)$, where z is birthdate minus 9/1/1993 and b is the coefficient of interest.

y = # of individuals charged with a crime at age 18-19 in OK with given birthday / # individuals born on that day in OK. Robust standard errors in parentheses.

CHAPTER 3

The Effects of Holistic Performance-Based Compensation: Evidence from Oregon's Teacher Incentive Fund

Co-authored work with Thomas Dee and James Wyckoff

1. INTRODUCTION

There are substantial challenges to designing compensation systems to identify and retain workers; these challenges are only compounded for teachers. This is, in part, because effective teaching is multi-dimensional and, in some cases, only realized over students' longer-term development. A further complication is that the identification and reliable measurement of effective teaching practice is still developing and is likely to depend on the specific context.

In most current practice, assessments of public-school teachers tend to be informal, uninformative (e.g. binary ratings of satisfactory or unsatisfactory), and unconnected to both professional-development activities and long-term retention decisions (Weisberg et al, 2009). In fact, over 99% of teachers in districts with these sorts of evaluations were deemed satisfactory (Weisberg et al, 2009). Meanwhile, a number of studies have found substantial variation in teacher quality, suggesting that replacing a 25th percentile teacher with a 75th percentile teacher would increase math test scores by roughly 0.2 standard deviations (Hanushek and Rivkin, 2012). There is also a growing realization that this large variance in teacher quality has economically meaningful implications for students' long-run outcomes (Lazeaer, 2003; Hanushek and Zhang, 2009; Hanushek and Woessman, 2009; Chetty et al., 2011). This has motivated recent policy efforts to design and implement new systems that can reliably assess, support, and encourage teacher quality. The U.S. Department of Education recently sponsored one of the most high profile and large-scale of these initiatives under the aegis of the Teacher Incentive Fund (TIF).

The TIF program was initially authorized in 2006 to provide support to public schools introducing "performance-based compensation systems" (PBCS). In 2010, the U.S. Department of Education increased the scale of this program, using \$600 million in stimulus funds to make awards to over 60 multi-district grantees implementing reforms in more than one thousand schools. The revised federal guidance associated with these 2010 TIF awards (referred to as Cohort 3) required that districts implement integrated systems that blend both test-score data and observation-based assessments that guide teachers' professional-development activities as well as decisions about their retention. Additionally, grantees were required to limit the use of TIF funds to high-need schools, defined as those schools with 50 percent or more of their students eligible for free or reduced-price lunches.

We examine the effects of these TIF reforms by focusing on one Cohort 3 TIF awardee, a coalition of six Oregon districts with 131 schools. Using student-level data, we track outcomes through the TIF planning year and two years of full implementation. We identify the effects of the TIF reforms through a regression-discontinuity (RD) design that leverages the federally defined eligibility criteria that limited TIF to high-need schools.

Our study of this new initiative can be situated in an emerging literature on multi-faceted systems of teacher assessment, feedback, incentives, and support (e.g., Dee and Wyckoff 2013). These more holistic styles of compensation reform pair tools for teachers to improve their performance with incentives to do so. They stand in contrast to recent narrowly targeted

experimental pilots, which provided only cash awards to teachers for test-score improvements and that have generally not been found to increase achievement (Springer et al, 2010; Springer et al, 2012; Fryer, 2013; Glazerman and Seifullah, 2012).

The rest of this paper is organized as follows. In Section 2, we review the economic literature on PBCS. In Section 3, we provide some background detail on TIF. In Section 4, we describe our data. In Section 5 we discuss our research design. In Section 6, we present our findings. Finally, in Section 7, we provide a concluding discussion.

2. EXISTING EVIDENCE ON PERFORMANCE-BASED COMPENSATION SYSTEMS

There is a large, decades-old literature that explores variants of performance-based compensation systems (PBCS) for teachers. However, until recently, the methodological rigor of the available empirical evidence has generally been weak.¹

Most PBCS are designed to improve student achievement (1) by providing financial incentives for teachers to exert more effort or develop stronger teaching skills, and (2) by increasing incentives for effective teachers to enter and remain in schools. Several recent city-level studies provide evidence solely on the effort margin by examining the extent to which the productivity of existing teachers increases in "cash for test score" incentive schemes. First, the Project on Incentives in Teaching (POINT) was a 3-year study that provided randomly assigned middle-school mathematics teachers in Nashville bonuses of as much as \$15,000 if their students met ambitious performance thresholds (Springer et al. 2010). The availability of these incentives led to no detectable effects on measured student performance or on measures of teacher effort

¹ For an overview of this literature, see Springer (2009) or Johnson and Papay (2009).

and classroom practice. A second random-assignment study provided New York City teachers with rewards up to \$3,000 for meeting performance targets (Fryer, 2013). In this study, treatment schools had flexibility in designing their incentive schemes, and most chose group-based incentives. The impact estimates from this study suggest that the presence of incentives did not raise school performance and may have even lowered it. A third study was conducted in 34 Chicago schools that were randomly assigned to when (but not if) they implemented the Teacher Advancement Program (TAP). Under this program, which was somewhat broader than a simple cash for test score scheme, teachers were eligible to receive payouts of as much as \$6,400 for their contribution to the achievement-based value-added of their students (at the school and school-grade level) and their performance on a classroom observation rubric. Under TAP, teachers could also earn extra pay for undertaking the increased responsibilities associated with promotion to a mentoring or master status. The evidence from this study suggests that random assignment to TAP did not raise student achievement (Glazerman and Seifullah, 2012). However, the program implementation did not occur entirely as intended. Teacher payouts were smaller than the originally stated targets and there were no rewards based on value-added because the requisite linked data systems were inadequate.

In contrast to the null findings in studies of cash for test score incentive schemes, Dee and Wyckoff (2013) find substantial positive effects of incentives on teacher performance and retention in the context of IMPACT, the teacher-evaluation system introduced in the District of Columbia Public Schools. Unlike the incentive schemes previously studied, IMPACT based its high-powered incentives (i.e. dismissal threats and financial rewards) on multiple measures of teacher performance, including both test performance and structured observational measures. Dee and Wyckoff use a RD design that leverages discontinuities in rewards or threats at performance score thresholds to identify the effects of these incentives. They find that dismissal threats increased the voluntary attrition of low-performing teachers more than 50 percent and improved the performance of the low-performing teachers who remained by 0.27 of a teacher-level standard deviation. They also find evidence that financial incentives further improved the performance of high-performing teachers (effect size = 0.24).

Aside from the motivational impacts of incentives, PBCS advocates argue that these systems have the potential to attract and retain a different type of teacher, one who expects to realize the compensation benefits of meeting performance targets (Lazear, 2003). These advocates fault rigid compensation systems that reward only experience and education for the sharp decline in the aptitude of those selecting into the teaching profession over the last 40 years (Corcoran et al. 2004, Hoxby and Leigh 2004). The potential of PBCS to impact such patterns of self-selection into and out of the teaching profession may be a particularly important mechanism for improving overall teacher quality. However, we are aware of relatively little empirical research that explores this important aspect of PBCS directly. One example is a panel-based evaluation of a targeted teacher bonus in North Carolina (Clotfelter et al. 2008), which finds that a \$1,800 bonus reduced teacher turnover by 17 percent. The TAP evaluation described above also found limited evidence that incentives improved teacher retention (Glazerman and Seifullah 2012). Another example is the Dee and Wyckoff's study of IMPACT, which found dismissal threats increased the likelihood that teachers left the district. However, the field-experimental studies in New York and Nashville provide no evidence that compensation incentives influenced teacher retention (Springer et al. 2010, Fryer et al. 2011).

Our contributions speak to the evaluation of the resource-intensive TIF program and the more general effectiveness of PBCS that employ incentives based on multiple measures (i.e.

classroom observations and school value-added) in combination with high quality evaluative feedback and professional development. First, we analyze a program which differs markedly from most other PBCSs that have been studied because of its more holistic approach. Second, we provide direct, timely evidence on the potential effectiveness of a future expansion of the TIF program to less needy schools by estimating the effect of TIF *at* the high-need threshold. Third, our credible evaluation of the student performance effects of the TIF program is important simply because the program is an unusually large-scale and exceptionally expensive federal effort (i.e., \$1.2 billion over five years to 62 multiple-partner programs in 27 states).

3. TEACHER INCENTIVE FUND (TIF)

In this section, we provide institutional details on the Teacher Incentive Fund (TIF). First, we describe the evolution of the federal grant program. Then, we discuss the specifics of the implementation of TIF in Oregon, focusing in particular on the contrast between TIF and non-TIF schools.

3.1 Federal Competitive Grant Program

The U.S. Department of Education's (USDOE) TIF was initially authorized with \$99 million in federal appropriations in 2006.² Its goal was to support education agencies that develop and implement performance-based compensation systems (PBCS) in primary and secondary education settings. In 2007, 33 grants were given to state or local education agencies

² "Teacher Incentive Fund: Funding," http://www2.ed.gov/programs/teacherincentive/funding.html (October 8, 2014).
in 18 states (cohorts 1 and 2).³ In 2010, dramatically increased funding from the American Recovery and Reinvestment Act of 2009 (\$200 million) and the Consolidated Appropriations Act of 2010 (\$400 million) permitted a larger third cohort of 63 grants to agencies in 27 states.⁴ For Cohort 3, USDOE imposed a number of new, explicit requirements to grant applicants. First, applicants were required to develop and implement a PBCS that rewards both teachers and principals who demonstrate their effectiveness in improving student achievement. PBCSs had to give significant weight to student growth based on test score data (with a preference given to value-added models) and also include observation-based assessments. They also had to provide substantial incentive amounts to teachers and principals. Second, applicants were required to present an integrated strategy that incorporated test score data and observation-based assessments in professional development as well as retention and tenure decisions. Third, applicants were directed to demonstrate the fiscal sustainability of their PBCS and accept responsibility for funding an increasing share of the PBCS using non-TIF sources. Fourth, and particularly important for our methodological design, TIF funds could only be used to fund the incentive awards and professional development of teachers and principals in high-need schools. These eligible schools were defined as those where free and reduced price lunch eligible (FRL) students made up at least 50% of their enrollment in 2009-10.⁵

3.2 Oregon TIF

In 2010, the Chalkboard Project, an Oregon-based non-profit organization in partnership with six Oregon school districts, was awarded \$24.4 million over five years by the Teacher

³ "Teacher Incentive Fund: Awards," http://www2.ed.gov/programs/teacherincentive/2007-awards.html (June 11, 2012).

⁴ "Teacher Incentive Fund: Funding," http://www2.ed.gov/programs/teacherincentive/funding.html (October 8, 2014).

⁵ Under the Richard B. Russell National School Lunch Act, students from families below 185% of the federal poverty level are eligible for free or reduced price lunch.

Incentive Fund.⁶ After a planning year (2010-11), the TIF implementation began in the 2011-12 school year in the selected TIF schools in each district.⁷ Not all TIF-eligible schools (those with at least 50% FRL) were selected as TIF schools. Among their TIF-eligible schools, most districts gave preference to those schools that were underperforming in at least one subject compared to similar schools in the state. Of the 147 schools in the six participating districts (Bend-La Pine, Crook County, Greater Albany, Lebanon, Redmond, and Salem-Keizer), 41 schools received the TIF performance-based awards (TIF schools).⁸

Prior to TIF, these six districts had received CLASS (Creative Leadership Achieves Student Success) grants of up to \$30,000 (as well as up to \$30,000 of in-kind technical support) from the Chalkboard Project to offset the costs of designing and planning a framework to integrate performance evaluation, relevant professional development, and expanded career opportunities for teachers and principals.⁹ Due in part to CLASS, much of each district's policies around career paths, performance evaluation, and professional development were the same for TIF and non-TIF schools. However, important differences remained, which together constitute the TIF "treatment" in Oregon. First, teachers in TIF schools received incentive awards based on their performance, while teachers in non-TIF schools did not. The details of the award models differed by district, but all districts provided awards of up to approximately \$6,000 for teachers (\$12,000 for principals) based on school-level value-added measures and teacher-

⁶ Initial award of \$13.2 million was later increased to \$24.4 million.

⁷ The first performance awards were made during the 2012-13 school year using results from the teacher evaluations for 2011-12.

⁸ In two districts, Bend-Lapine and Greater Albany, 22 of the 25 schools selected to participate in TIF also participated in the national TIF evaluation. This meant that these schools were randomly assigned to either be TIF schools or control schools, where the control condition involved a stipend given to all teachers rather than performance-based awards. In all, 11 schools were in this control group.

⁹ "Request for Proposals, CLASS Project Pilot Program", The Chalkboard Project (February 2012).

level (or principal-level) evaluations.¹⁰ Second, only teachers and principals in TIF schools could receive TIF-funded professional development.¹¹ This restriction was important because CLASS did not include implementation funding to ensure the adoption of new policies, such as increased professional development, district-wide. This made professional development for teachers in TIF schools easier for districts to provide, but it also resulted in higher quality teacher evaluations (i.e. more observations and more feedback) by administrators in TIF schools, according to district interviews conducted by the Chalkboard Project.¹² These higher quality evaluations were a result of TIF school administrators receiving substantially more training than their non-TIF counterparts. For example, they were required to complete Teachscape, an evaluator training program which includes scoring videos of classrooms and two multiple hour tests; they also attended a 3-day summer institute which included training on how to give effective feedback.¹³

4. DATA

In this section, we discuss the data used in our empirical analysis (subsection 4.1), the construction of the assignment variable in our regression discontinuity design (subsection 4.2), and the selection of our analytical sample (subsection 4.3).

4.1 Student Achievement Data

¹⁰ "Grantee Profiles: Chalkboard Project," http://www.cecr.ed.gov/TIFgrantees/granteeProfiles/cohort3.cfm?id=9 (March 30, 2012).

¹¹ "Teacher Incentive Fund: Frequently Asked Questions for the 2010 Competition and Grant Awards," http://www2.ed.gov/programs/teacherincentive/2010-faqs.doc (June 28, 2010).

¹² B. Pratt (personal communication, June 4, 2014).

¹³ B. Pratt (personal communication, June 4, 2014).

We obtained student-level achievement data for all public schools in Oregon for school years 2007-08 through 2012-13 from the Oregon Department of Education through a restricted-use data agreement. Observations in this data include all of the students who took the math and reading Oregon Assessment of Knowledge and Skills (OAKS). The OAKS is taken by all students in Oregon public schools in grades 3-8 and then once in high school. It measures student performance against the Oregon Content Standards and assigns raw scores between 150 and 300. The data contain student and school identifiers, student grade level, raw test scores, and student demographic information (e.g. free and reduced-price lunch eligible, ethnicity, and sex). We normalize the raw test scores by subtracting the state-wide mean score for each grade, subject, and year combination and then dividing by the state-wide standard deviation for that test-subject-year.¹⁴ The resulting standard deviation scores allow for easy comparisons across grades, years, and subjects.

4.2 Assignment Variable Construction

Our RD design identifies the effect of TIF by comparing students in schools that were just eligible for TIF with those that were just ineligible. Consequently, the measure by which this eligibility is determined, the assignment variable, is critical to our empirical approach. It is intended to reflect the eligibility of a school at the time when the school's district made the decision to select or not select it as a TIF school. While we were unable to obtain the precise data used by the districts for their TIF selection, we recreated this data as accurately as possible given Oregon Department of Education and Chalkboard Project documentation and in consultation with representatives from those organizations.

¹⁴ For students with multiple test records on the same test in the same day likely as a result of a data entry error, we use the maximum score. For students that took the same test multiple times on different days, we use the score from their most recent test.

According to USDOE's eligibility criteria, a school could receive TIF funding if at least 50% of students were eligible for free and reduced price lunch (FRL).¹⁵ Middle and high schools could also be deemed TIF-eligible if their "feeder schools" exceeded the 50% FRL threshold.¹⁶ We construct an assignment variable z_s for a given school *s* that represents how close the school was to satisfying these eligibility criteria at the time that TIF schools were selected by the participating districts. We define this assignment variable as follows:

$$z_s = \max(FRL_s, FEEDFRL_s) - 50,$$

where school *s* was TIF-eligible if and only if $z_s \ge 0$. FRL_s is the percent of the enrollment of school s that was eligible for FRL at the time of selection. For middle and high schools, $FEEDFRL_s$ is the percent of the enrollment that was eligible for FRL in schools that fed into *s*. For elementary schools, $FEEDFRL_s$ is set to zero so that it has no effect on z_s .

To the extent possible, we use the FRL data that is recorded nearest to the time that final TIF selections were made in the Summer of 2010. For the Bend-Lapine School District, we were able to obtain FRL data from the district recorded in May 2010. For the remaining TIF-participating districts, we were only able to obtain FRL data from the Oregon Department of Education's state-wide database, recorded in Fall 2009.^{17,18}

¹⁵ This eligibility restriction applies only to funding for incentive awards and professional development. More general funding, such as for data system upgrades could benefit schools that were not "TIF-eligible".

¹⁶ A school's "feeder schools" are those schools that send their graduating students to the school. If 5^{th} grade students from elementary schools A, B, and C go to middle school D as 6^{th} graders, then A, B, and C are feeder schools for D.

¹⁷ If Fall 2009 state data is used for Bend-Lapine, not all TIF schools are

¹⁸ For the one school in TIF-participating disctricts that was opened in the 2010-11 school year, Rosland Elementary, we use FRL data for the 2010-11 school year (recorded in Fall 2009) from the Oregon Department of Education's state-wide database.

Following Oregon Department of Education guidelines, we define $FEEDFRL_s$ as the average percent FRL of students entering school *s* from all of its feeder schools.¹⁹ This is constructed by taking the sum of FRL students in all of a school's feeders and dividing it by the sum of enrollment in all those feeders, where each sum is weighted by the fraction of students coming from each feeder school. This can be written as follows:

$$FEEDFRL_{s} = 100 \cdot \frac{\sum_{f \in F_{s}} p_{sf} \cdot NFRL_{f}}{\sum_{f \in F_{s}} p_{sf} \cdot NENROLL_{f}}$$

where F_s is the set of feeder schools for school *s*. $NFRL_f$ is the count of FRL students in feeder school *f* and $NENROLL_f$ is the count of all students in feeder school f^{20} p_{sf} is the fraction of students at school *s* that came from feeder school f^{21} . Figure 1 shows the school-level distribution of z_s for the analytical sample. Few schools (and students) have z_s below -25 (25% FRL) or above 40 (90% FRL).

4.3 Analytical Sample

We make a number of restrictions to the universe of Oregon public schools to obtain our analytical sample. First, we restrict to schools that were in existence in 2010-11, the year before the implementation of TIF, in one of the six TIF-participating districts (145 schools). This ensures that our analysis only compares schools in districts that chose to apply for the TIF grant with Chalkboard. Second, we exclude non-traditional schools (8 charter schools and 4 K-2

²⁰ When calculating $FEEDFRL_s$ for a high school, if a feeder school f is a middle school then its own feeder FRL and enrollment is used, $NFRL_f = \sum_{g \in F_f} p_{fg} * NFRL_g$ AND $NENROLL_f = \sum_{g \in F_f} p_{fg} * NENROLL_g$. ²¹ We calculated p_{sf} using unduplicated counts of students by school attended 10/1/2011 and school attended 10/1/2010, obtained from Oregon Department of Education. We restricted this data to transitions representing students graduating from one school and entering another school within the same district (e.g. 5th grade to 6th grade transitions). p_{sf} was then defined as the fraction of students entering school s that had been enrolled in feeder school f in the prior year.

¹⁹ Janet Bubl. 2014. "Establishing Poverty Levels." Oregon Department of Education Memo.

schools), leaving a sample of 133 schools (80 elementary schools, 22 middle schools, 17 high schools, and 14 schools with other grade configurations). Under our definition of the assignment variable, the sample can be categorized into three groups: 41 schools were TIF-ineligible ($z_s < 0$) and *did not receive* the TIF treatment, 51 schools were TIF-eligible ($z_s \ge 0$) and *did not receive* the TIF treatment, 51 schools were TIF-eligible ($z_s \ge 0$) and *did not receive* the TIF treatment, and 41 schools were TIF-eligible and *received* the TIF treatment.²² Among the TIF-eligible schools ($z_s \ge 0$), the decision to assign schools to the TIF treatment was made by districts, often based on whether they had underperformed in the previous year (relative to a comparison group of schools in the state).²³ Table 1 shows summary statistics for schools in the analytical sample in 2009-10. Tests scores are observed for 44,624 students, their average in math and reading was slightly below the state average. Students in TIF-ineligible schools were more likely to be white, less likely to be economically disadvantaged, and averaged much high test scores (0.3-0.4 SD) than students in TIF-eligible schools. Students in TIF schools, but averaged lower test scores (0.09 SD).

5. EMPIRICAL STRATEGY

We evaluate the effects of the stimulus-funded TIF programs in a fuzzy regression discontinuity (fuzzy RD) design that leverages the TIF-eligibility requirement mandating that these reforms be targeted only to high-need schools. The basic intuition behind the strong causal warrant of this RD design is straightforward. Within the districts receiving TIF awards, schools that just barely meet the high-need standard are much more likely to implement TIF. In contrast,

²² The list of TIF schools was obtained from the Chalkboard Project. "TIF treatment" schools exclude TIF schools randomly assigned to the control group as part of the national evaluation of TIF.

²³ B. Pratt (personal communication, June 4, 2014).

schools that just fail to meet the high-need standard are ineligible for TIF. That is, whether a school is just above or below this threshold implies a strong contrast in the PBCS regime that is experienced. This localized variation in whether a school just meets or fails to meet this condition can be viewed as effectively random so long as it is not manipulated, an issue that we will examine below. Any differences in outcomes among students in schools just above and below the high-need threshold can therefore be credibly attributed to the effects of TIF eligibility and use. Consequently, the RD design alleviates concerns that treated and non-treated groups differ in unobservable ways.

5.1 Regression Discontinuity Specification

We examine this treatment contrast through first-stage specifications that take the following general form:

$$TIF_{is} = \alpha I(z_s \ge 0) + f(z_s) + \theta X_{is} + \epsilon_{is}, \tag{1}$$

where α identifies the discrete change in TIF-funded reforms experienced by students in schools meeting the eligibility threshold (i.e. $I(z_s \ge 0)$). This identification conditions on how the probability of undertaking TIF reforms relates to the assignment variable, $f(z_s)$ and to control variables reflecting the observed traits of a given student *i*, X_{is} . A similarly structured reducedform (or intent-to-treat) specification is applied to the outcome measures (i.e. student reading and math achievement), Y_{is} , for student *i*:

$$Y_{is} = \gamma I(z_s \ge 0) + \kappa(z_s) + \beta X_{is} + \nu_{is}.$$
(2)

That is, estimates of γ identify the discrete shift in outcomes associated with being a TIF-eligible school.

We examine the causal warrant of this method by performing a series of tests commonly used when implementing RD designs as well as tests made possible by the panel-nature of our data. First, we restrict the analytical sample to include only students in schools with values of z_s close to the threshold (e.g. plus or minus 20 percentage points). Second, we provide falsification tests that look at the estimated placebo effect of TIF prior to its implementation. Third, we check for covariate balance around the threshold by estimating Equation 2 with student covariates as the outcome variable. Fourth, following McCrary (2006), we test for manipulation of the assignment variable by investigating whether there is a discontinuity in the density of students at the eligibility cutoff. This test is done at the school-level rather than the student-level, as the school is the unit that could conceivably manipulate their FRL in order to be eligible for TIF. Manipulation is unlikely, however, since in most cases it would require schools to anticipate their district's TIF award more than half a year before the application was sent. Fifth, we alter our assignment of students to schools to address concerns that changes in student mobility (e.g. more talented students switching to TIF schools) may be driving estimated TIF effects. In this analysis, we restrict our analytical sample to students in grade 3-6 in 2009-10 (or alternatively 2010-11) and assign these students to a predicted school in the TIF treatment years based on the most common school in each year for students in a given school-grade in 2009-10 (or 2010-11). This approach also allows us to control for a student's baseline test score (in 2009-10 or 2010-11). We therefore remove individual school choice as a potential factor, while also controlling for a clean baseline measure of achievement at the individual student-level.

5.2 Difference-in-Regression Discontinuity Specification

We expand on our RD design using a Difference-in-Regression Discontinuity (DRD) approach, which compares the differences between just eligible and just ineligible students

across multiple years simultaneously.²⁴ The primary purpose of this approach is to explicitly test whether discontinuities in the TIF years (2010-11 to 2012-13) statistically differ from any preexisting discontinuity at the eligibility threshold in the pre-TIF years, which are grouped together (2007-08 to 2009-10). If the discontinuities in TIF years do not differ statistically from pre-TIF years, it would suggest that any effects estimated in the TIF years using Equation 2 are the product of pre-existing differences around the eligibility threshold. The reduced-form specification is as follows,

$$Y_{ist} = \gamma_0 I(z_s \ge 0) + \kappa_0(z_s) + \sum_{\tau=2011}^{2013} I(t=\tau) \cdot \left(\gamma_\tau I(z_s \ge 0) + \kappa_\tau(z_s)\right) + \beta X_{ist} + \nu_{ist}, \quad (3)$$

where t is the school year (2011 refers to 2010-11), $I(t = \tau)$ is an indicator equal to 1 for TIF school year $\tau \in \{2011, 2012, 2013\}$, and the other variable definitions follow from Equation 2. γ_0 provides the change in achievement at the eligibility threshold in the Pre-TIF years, while γ_{τ} provides the differential change in achievement in year $\tau \in \{2011, 2012, 2013\}$ compared to the pre-TIF period.

5.3 Interpretation of Regression Discontinuity Estimates

If schools were randomly selected into the TIF treatment from among the group of TIFeligible schools, then the ratio of the reduced-form estimate from Equation 2 and the first stage estimate from Equation 1 $\left(\frac{\gamma}{\alpha}\right)$ gives the local average treatment effect (LATE) of TIF, that is, the average treatment effect of PBCS for students in schools near the eligibility cutoff. However, this was not the case in Oregon, where TIF schools were selected by the district from TIFeligible schools, often because they had underperformed in the previous year. This means that

²⁴ A similar design is used by Fitzpatrick (2010), though she compares birthdate discontinuities across treated and untreated states rather than across treated and untreated years.

 $\frac{\gamma}{\alpha}$ can be better interpreted as the local average treatment on treated (LATOT) effect. If district administrators selected for TIF treatment the TIF-eligible schools that were most likely to benefit from the treatment, then the LATOT would exceed the LATE. In this case, estimates of $\frac{\gamma}{\alpha}$ can be considered an upper bound of the LATE, and expanding the TIF program to additional eligible schools that were not initially selected would be expected to yield more modest effects.

6. RESULTS

In this section, we first present regression discontinuity (RD) estimates of the impact of TIF on math and reading achievement. Second, we discuss a number of robustness and falsification tests of our main results. Third, we present an alternative RD approach which alleviates concerns that differential student sorting determines our main results. Fourth, we present difference-in-regression discontinuity results that explicitly test differences in the achievement discontinuities in different years. Finally, we summarize our findings.

6.1 Main Results for Regression Discontinuity Design

Table 2A and 2B show the OLS regression discontinuity estimates of α and γ from Equation 1 (first stage) and Equation 2 (reduced-form) for math and reading standardized scores in each TIF-treatment year (2011-12 and 2012-13). The tables include results for specifications with different sets of individual covariates, all with standard errors clustered at the school-level. The estimates for α (first row) show a statistically and economically significant first stage, which does not vary across specifications or years. However, in some specifications, the first stage Fstatistic falls slightly below 10, the commonly applied rule of thumb for avoiding a weak instrument problem, which could inflate effect estimates. The magnitude of the estimates suggests that TIF eligibility increases the probability of receiving the TIF treatment by roughly 40 percentage points. Figure 2 depicts the first stage graphically. For 4 percentage point bins of the assignment variable, z_s , the figure shows the proportion of test-takers in TIF-treatment schools, and superimposes the regression lines from Equation 1.²⁵

The second row of Table 2A shows the reduced form estimates (γ in Equation 2) for math standard deviation scores in the same years and using the same specifications. Figure 3A depicts these results graphically (equivalent to column 1). These results show no significant effects of TIF eligibility on math standard deviation scores, with the exception of a marginally significant (at the 10% significance level) and unexpectedly negative impact in 2012-13 (column 10) that is not robust across specifications. Table 2B shows positive and statistically significant reduced-form effects of TIF on reading standard deviation scores in 2011-12, but not in 2012-13. This can also be seen graphically in Figure 3B (equivalent to column 1). The estimates for 2011-12 can be interpreted as a 0.10-0.16 standard deviation increase in reading achievement from TIF eligibility. The resulting estimate of $\frac{\gamma}{\alpha}$ is 0.26-0.41 standard deviations that is statistically different from zero, but imprecisely estimated. As discussed in Section 5, this magnitude likely represents the upper bound of the local average treatment effect (LATE) rather than the LATE itself.

6.2 Tests of Internal Validity

Table 3A and 3B present robustness checks for the results in Table 2A and 2B by showing the same specifications estimated on a smaller window around the eligibility cutoff (i.e.

²⁵ Figure 2 uses the sample of reading test-takers, but the graph is identical for math test-takers.

 $z_s \in [-30, 30]$ and $z_s \in [-20, 20]$ or equivalently %*FRL* $\in [20, 80]$ and %*FRL* $\in [30, 70]$).²⁶ The reduced-form results remain qualitatively similar for smaller windows, with two minor exceptions. First, the 2011-12 reduced-form results for reading increase slightly in magnitude for smaller bandwidths. Second, the 2012-13 reduced-form results for reading become statistically significant for two of three specifications when using the $z_s \in [-30, 30]$ bandwidth (equivalent to %*FRL* $\in [20,80]$). Figures A1 and A2 depict the reduced-form results for %*FRL* $\in [30,70]$ graphically.

Table 4A and 4B provide a falsification test of whether our empirical strategy finds a TIF placebo effect prior to its implementation, in which case the validity of our approach would be doubtful. These tables show estimates of α and γ from Equations 1 and 2, as in Table 2A and 2B, but for the years prior to TIF implementation. The results show small and statistically insignificant reduced-form effects for either math or reading in the years prior to TIF (2007-2008 through 2009-2010) or in the TIF planning year (2010-2011). This evidence lends substantial support to interpreting the estimated achievement effects in Table 2A and 2B as products of TIF and not as a consequence of observed or unobserved differences between students on either side of the eligibility threshold.

Table 5 provides tests of whether test-taker covariates differ discontinuously on either side of the eligibility threshold. The existence of such a difference would present the possibility that achievement differences at the eligibility threshold are due to observable differences between schools rather than TIF eligibility. The table includes estimates of γ in Equation 2 for various bandwidths, where the dependent variable is a student characteristic (X_{is} is excluded). The results show no significant discontinuities in the likelihood a test-taker is

²⁶ The relatively few number of schools prevents estimating tighter bandwidths.

economically disadvantaged, but does show significant discontinuities in the likelihood a testtaker is white (11-12 percentage points) and special education students (1-2 percentage points, marginally significant in 2012-2013 only).

We assess the importance of this possible covariate imbalance in two ways. First, we note that no discontinuities in covariates remain significant for the smallest bandwidth, suggesting that these significant estimates may be driven by the influence of schools far from the threshold on the slopes of $\kappa(\cdot)$. The significance of the results in Table 3 were not similarly affected by bandwidth restrictions. Second, we construct indexes of the reading and math test scores predicted from a student's covariates and then check for a threshold discontinuity in this index. This is a more appropriate test of imbalance at the threshold, as covariate imbalance is only problematic to the extent that those covariates affect achievement. We find no significant discontinuities in predicted score indexes at the eligibility threshold for any bandwidth size.

Figure 4 shows the results of the McCrary density test at the school-level, which checks for a discontinuity in the distribution of the assignment variable at the eligibility threshold (McCrary, 2006). Such a discontinuity could be evidence of schools manipulating the assignment variable in order to make themselves eligible for TIF. We find no evidence of a discontinuity, which is not surprising since the assignment variable was measured more than half a year prior to the TIF application for most schools. We address the possibility that students are "manipulating" their assignment variable with their choice of schools in the following subsection.

6.3 Cohort-tracking Regression Discontinuity Results

Table 6A and 6B address the possibility that the estimated TIF effects in Table 2A and 2B are due to changes in student mobility to TIF and non-TIF schools. For example, if relatively high-achieving students are more likely to move to TIF schools, then increases in achievement in those schools could be due to the composition of their students rather than improvements to teacher/principal quality from TIF. These tables provide estimates of α and γ in Equations 1 and 2 on two cohort subsamples (grade 3-6 students in 2009-10 and in 2010-11), where students are assigned to the school in 2011-12 and 2012-13 attended by the majority of students that were in their school and grade in the baseline year.²⁷ In this way, we ensure that differential sorting by achievement is not driving our estimates. With this construction, we are also able to control for individual-level baseline test scores (in 2009-2010 or 2010-2011), rather than school by grade test score means.

Table 6A shows no statistically significant effects of TIF on math achievement in 2011-12 or 2012-13. In the main math achievement results in Table 2A, only the estimate in 2012-13 that controlled for school by grade test score means (in 2009-10) was marginally significant (column 6). In Table 6A, the individual baseline test score specification (column 6) produces a 2012-13 estimate that is not significant. Table 6B shows large and significant reduced-form estimates of improved reading achievement in 2011-12. In fact, our preferred individual baseline test score specification (column 3) finds results in line with Table 2B. These results imply estimates of $\frac{\gamma}{\alpha}$ of 0.21 to 0.28 standard deviations. There are also large and significant reducedform estimates for reading in 2012-13, but not for our preferred specification (column 6). Overall, the results in Table 6A and 6B are qualitatively similar to our main results in Table 2A

 $^{^{27}}$ Estimates for 2012-13 are only shown for the 2010-11 cohort because the 2009-10 cohort's youngest members are 6th graders in 2012-13 leading to the loss of elementary schools from the sample (half of all schools) and therefore the loss of the school-level variation necessary for the RD design to work.

and 2B, suggesting that those main results are not driven by student mobility or insufficiently precise baseline achievement controls.

6.4 Difference-in-Regression Discontinuity (DRD) Results

Table 7A and 7B present DRD estimates of γ_0 and γ_τ from Equation 3 for two school year samples. The first sample (columns 1 and 2) includes school years 2008-09 to 2012-13 and groups the pre-TIF school years together to test whether the achievement discontinuities at the eligibility threshold in the planning and TIF treatment years differ substantially from the prior years. The tables show insignificant estimates of pre-TIF achievement discontinuities (γ_0) in both math and reading. In Table 7A, there are no significant differences between math achievement discontinuities in TIF years and pre-TIF years. In Table 7B, there is a positive difference between the 2011-12 reading achievement discontinuity and the pre-TIF discontinuity that is significant at the 10 percent level when controlling for student covariates in column 2 (column 1, without controlling covariates, finds an estimate significant at the 16 percent level). The magnitude of these differential discontinuity estimates are in line with those in Table 2A, columns 1-3.

The second sample presented in Tables 7A and 7B, columns 3-4, contains only the TIF treatment years (2011-12 and 2012-13). These results provide a test of whether the difference in the achievement discontinuities between 2011-2012 and 2012-2013 are statistically significant. For both math and reading achievement, estimates of the "ELIGIBILITY X (SY2012-13)" coefficient (γ_{2013}) are not significant, suggesting that while the differences in magnitude between estimated effects in 2011-12 and 2012-13 may be meaningful from a policy stand-point, they are not statistically meaningful.

6.5 Summary of Results

Across multiple specifications, sample constructions, and bandwidth selections, we consistently find that TIF improved reading achievement significantly in its first year (2011-12) but not in its second year (2012-13). We also find that the estimated effect in 2011-12 differs statistically from the Pre-TIF placebo effect. The preferred reduced-form estimates of these reading improvements (Table 2B, columns 3 and 6) are 0.1 SD in 2011-12 and 0.05 SD in 2012-13, implying LATOT effects of 0.26 SD in 2011-12 and 0.11 SD in 2012-13. While the differences in these magnitudes are clearly important, they are not statistically different due to imprecision in the estimates (Table 7B, columns 3-5).

In general, we find no effects on Math across our various specifications, sample constructions, and bandwidth selections. The one exception is a negative estimate in 2012-13 that is statistically significant at the 9 percent level when including controls for student characteristics, grade fixed effects and baseline mean test score (Table 2A, column 6). This result is not robust to alternate specifications (Table 2A, column 4 and 5) or alternate sample constructions (Table 6A). The estimated math effect in 2012-13 is also not statistically different from the placebo effect in the Pre-TIF years (Table 7A).

7. CONCLUSION

We use student-level data from six TIF-participating districts in the state of Oregon to provide early evidence on how high-profile reforms, implemented as part of Teacher Incentive Fund (TIF), influenced measures of student performance. Leveraging an eligibility threshold in TIF-eligibility based on the percentage of FRL enrollment in a school, we use an RD design to find that Oregon's TIF-reforms *can* produce substantial improvements in reading achievement in selected schools, but we find no improvements in math achievement. While significantly different from zero, the imprecision of our estimates of TIF's impact on Reading in 2011-12 results in wide range of possible true effect sizes. Moreover, our estimates provide an upper bound for the impact of TIF if administrators positively selected schools into TIF from the pool of TIF-eligible schools.

Our RD design provides an important complement to the ongoing, federally sponsored randomized trial study in 12 TIF-participating districts. The treatment contrast fielded in the federal evaluation is narrower than that implied by the introduction of PBCS under the TIF program. Specifically, the staff in the experimental control schools receives individualized feedback on their effectiveness, professional-development opportunities, an automatic pay bonus, and additional pay for taking on new roles and responsibilities (e.g., peer mentoring, leading professional development activities). In the treatment schools, staff were eligible for these TIF elements and for differentiated awards based on their performance that were "a little larger on average, than in the control schools" (Max et al. 2014).

The contrast used to identify our RD estimates, leverages more components of the full treatment contrast created by the introduction of a PBCS under TIF because the comparison is between TIF-eligible and ineligible schools. Specifically, teachers in TIF schools experienced more highly trained administrators, higher-quality evaluation feedback, and increased availability of professional development compared to their counterparts in ineligible schools. Furthermore, the compensation contrast in our design compares incentive awards with the status quo, rather than with an automatic pay increase. Early evidence from the federal evaluation (Max et al. 2014) suggests that implementation fidelity among these awardees has been quite poor. Fewer than half of districts implemented all of the required TIF elements, incentive awards appear to have been insufficiently challenging (most teachers received one), and many teachers may have misunderstood the performance measures and bonuses. In light of these issues, our focus on a single awardee with comparatively high fidelity implementation is likely to provide a more powerful proof of concept than the federal evaluation. Under the relatively favorable conditions in Oregon, we find promising but uneven effects of TIF on achievement.

REFERENCES

- Aaronson, D., L Barrow, W Sander (2007). Teachers and Student Achievement in the Chicago Public Schools. Journal of Labor Economics 25(1) 95-135.
- Atteberry, A., S. Loeb, and J. Wyckoff. "Do First Impressions Matter? Improvement in EarlyCareer Effectiveness," CALDER Working Paper No. 90, February 2013.
- Ballou, D. (2001). "Pay for Performance in Public and Private Schools." Economics of Education Review 20(1): 51-61.
- Boyd, D., Lankford, H., Loeb, S., Ronfeldt, M., & Wyckoff, J. (2011). The role of teacher quality in retention and hiring: Using applications-to-transfer to uncover preferences of teachers and schools. Journal of Policy and Management, 30(1), 88–110.
- Camerer, C. and R.M. Hogarth. "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework," Journal of Risk and Uncertainty 19(1-3), December 1999, 7-42.
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood: National Bureau of Economic Research.
- Clotfelter, C., Ladd, H. F., & Vigdor, J. L. (2006). Teacher-student matching and the assessmentof teacher effectiveness. Journal of Human Resources, 41(4), 778.
- Clotfelter, C., E, Glennie, H. Ladd, and J. Vigdor (2008) "Would Higher Salaries Keep Teachers in High-Poverty Schools? Evidence from a Policy Intervention in North Carolina" Journal of Public Economics, 92: 1352-1370.
- Danielson, Charlotte & Thomas L. McGreal. (2000). Teacher evaluation to enhance professional practice. Alexandria, Va.: Association for Supervision and Curriculum Development.
- Dee, T.S. and B. Keys (2004). "Does Merit Pay Reward Good Teachers? Evidence from a Randomized Experiment," Journal of Policy Analysis and Management, 23(3), pages 471-488.
- Dee, T.S. and J. Wyckoff (2013). "Incentives, Selection, and Teacher Performance: Evidence from IMPACT," NBER Working Paper 19529.
- Fryer, R. (2013). "Teacher Incentives and Student Achievement: Evidence form New York City Public Schools," Journal of Labor Economics 31(2): 373-427.
- Fryer, R., S. Levitt, J. List, and S. Sadoff. "Enhancing the Efficacy of Teacher Incentives through Loss Aversion," NBER Working Paper 18237, 2012.
- Glazerman, S. and A. Seifullah (2012). An Evaluation of the Chicago Teacher Advancement Program (Chicago TAP) After Four Years. Princeton: Mathematica Policy Research.
- Glazerman, S., E. Isenberg, S. Dolfin, M. Bleeker, A. Johnson, M. Grider, and M. Jacobus. (2010). Impacts of Comprehensive Teacher Induction: Final Results From a Randomized

Goe, Laura. and Andrew Croft. (2009). Methods of evaluating teacher effectiveness. Washington, DC: National Comprehensive Center for Teacher Quality.

- Goldhaber, D., B. Gross and D. Player (2007) "Are Public Schools Really Losing Their "Best"? Assessing the Career Transitions of Teachers and Their Implications for the Quality of the Teacher Workforce," CALDER working paper.
- Hanushek, E. (2007) "The Single Salary Schedule and Other Issues of Teacher Pay" Peabody Journal of Education, 82(4), 574-586.
- Hanushek, E., J. Kain, D. O'Brien and S. Rivkin (2005) "The Market for Teacher Quality," NBER working paper.
- Hanushek, E. and S. Rivkin (2012). "The Distribution of Teacher Quality and Implications for Policy," Annual Review of Economics 4: 131-157.
- Imbens, G. and K. Kalyanaraman (2012). "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," Review of Economic Studies 79 (3): 933-959
- Johnson, S. M., and John Papay (2009) Redesigning Teacher Pay, Washington, D.C.: EPI.
- Joseph, C. "New Evaluation System for New York Teachers," New York Times, June 2, 2013, page A20.
- Lee, D.S. & Lemieux, T. (2009). Regression discontinuity designs in economics. Journal of Economic Literature, 48, 281-355.
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: a density test. Journal of Econometrics, 142(2), 698-714.
- McNeil, M. "Feds, States Dicker Over Evaluations," Education Week, Vol. 32, Issue 12, March 16, 2013, pages 1,24.
- Measures of Effective Teaching (MET) Project (2013). Ensuring Fair and Reliable Measures of Effective Teaching: Culminating Findings from the MET Project's Three-Year Study. The Bill and Melinda Gates Foundation, January 2013.
- Murnane, R. (1984) "Selection and Survival in the Teacher Labor Market," The Review of Economics and Statistics 66(3) 513-518.
- Murnane, R. J., & Cohen, D. K. (1986). Merit pay and the evaluation problem: Why most merit pay plans fail and a few survive. Harvard Educational Review, 56(1), 1–17.
- Pianta, R.C. and B.K. Hamre. "Conceptualization, Measurement, and Improvement of Classroom Processes: Standardized Observation Can Leverage Capacity," Educational Researcher, March 2009 vol. 38 no. 2, 109-119.
- Rivkin, S., E. Hanushek, and J. Kain (2005). "Teachers, Schools, and Academic Achievement," Econometrica, 73(2), 417-458.

- Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. American Economic Review, 94(2), 247-252.
- Rothstein, J. (2012). "Teacher Quality Policy When Supply Matters," NBER Working Paper No. 18419, September 2012.
- Sanders, W.L. and Rivers, J.C. (1996). "Research Project Report: Cumulative and Residual Effects of Teachers on Future Student Academic Achievement," University of Tennessee Value-Added Research and Assessment Center.
- Schochet, P. Z., Cook, T., Deke, J., Imbens, G., Lockwood, J.R., Porter, J., & Smith, J. (2010).
- Standards for Regression Discontinuity Designs. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_rd.pdf.
- Springer, M. (2009) "Rethinking Teacher Compensation Policies: Why Now, Why Again?" in Performance Incentives, M. Springer Ed., Washington DC: Brookings Institution.
- Springer, M., D. Ballou, L. Hamilton, V. Le, J.R. Lockwood, D. McCaffrey, M. Pepper, B. Stecher (2010) Teacher Pay for Performance, Experimental Evidence from the Project on Incentives in Teaching, Nashville, TN: National Center on Performance Incentives at Vanderbilt University
- Springer, M., J. Pane, V. Le, D. McCaffrey, S. Burns, L. Hamilton, and B. Stecher. "Team Pay for Performance: Experimental Evidence From the Round Rock Pilot Project on Team Incentives," Educational Evaluation and Policy Analysis December 2012, Vol. 34, No. 4, pp. 367–390.
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect. Brooklyn, NY: The New Teacher Project.
- Weiss, E. and D. Long (2013) Market-Oriented Education Reforms' Rhetoric Trumps Reality. Broader, Bolder Approach to Education. http://www.epi.org/files/2013/bba-rhetorictrumps-reality.pdf.
- Taylor, E.S. and Tyler, J.H. 2012. The effect of evaluation on teacher performance. American Economic Review, 102(7), pp. 3628-3651. 2012.
- Turque, B. "D.C. teachers likely to ratify contract," Washington Post, June 1, 2010, B01.
- Ujifusa, A. "Buffalo Battle Rages over Evaluations," Education Week, Vol. 32, Issue 30, May 8, 2013, Page 21.
- Yoon, Kwang Suk, Teresa Duncan, Silvia Wen-Yu Lee, Beth Scarloss, and Kathy L. Shapley. 2007. Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement: Issues & Answers Report, REL 2007 No. 033. Washington, DC: U.S. Department of Education, Institute of Education Sciences.
- Zubrzycki, Jacklyn. "Districts Abandon Grants Targeting Teacher Quality," Education Week, Vol. 32, Issue 01, August 22, 2012, pages 1, 20-21.

Figure 1: School-level Distribution of the Assignment Variable (%FRL)





Figure 2: First Stage Regression Discontinuity, by School Year



Note: Figure uses the sample of OAKS reading test-takers in TIF-participating districts (excluding charter schools). Dots depict mean TIF treatment in 4 percentage point %FRL bins. Lines show first stage RD regression (Equation 1). %FRL is defined as $\max(\%FRL,\%FEEDFRL)$ in 2009-10.

Figure 3A: Reduced-Form Regression Discontinuity, by School Year (Math)



Note: Figure uses the sample of OAKS math test-takers in TIF-participating districts (excluding charter schools). Dots depict mean SD test scores in 4 percentage point %FRL bins. Lines show reduced-form RD regression (Equation 2). %FRL is defined as max(%FRL,%FEEDFRL) in 2009-10.

Figure 3B: Reduced-Form Regression Discontinuity, by School Year (Reading)



Note: Figure uses the sample of OAKS reading test-takers in TIF-participating districts (excluding charter schools). Dots depict mean SD test scores in 4 percentage point %FRL bins. Lines show reduced-form RD regression (Equation 2). %FRL is defined as max(%FRL,%FEEDFRL) in 2009-10.

Figure 4: School-level McCrary Density Test, by School Year



Note: Figure applies the density test from McCrary (2006) to the sample of OAKS reading test-takers in TIF-participating districts (excluding charter schools), collapsed to the school-level. Binwidth of 2 is used. %FRL is defined as max(%FRL,%FEEDFRL) in 2009-10.

Figure A1: Reduced-Form RD for 20-70% FRL Window, by School Year (Math)



Note: Figure uses the sample of OAKS math test-takers in TIF-participating districts (excluding charter schools). Dots depict mean SD test scores in 4 percentage point %FRL bins. Lines show reduced-form RD regression (Equation 2). %FRL is defined as max(%FRL,%FEEDFRL) in 2009-10.

Figure A2: Reduced-Form RD for 20-70% FRL Window, by School Year (Reading)



Note: Figure uses the sample of OAKS reading test-takers in TIF-participating districts (excluding charter schools). Dots depict mean SD test scores in 4 percentage point %FRL bins. Lines show reduced-form RD regression (Equation 2). %FRL is defined as max(%FRL,%FEEDFRL) in 2009-10.

	All Schools	TIF-Ineligible	TIF-Eligible	TIE Schoole*	
	(1)	(2)	(3)	(4)	
Sample (Counts)					
Test-takers	44,624	13,935	17,146	13,660	
Schools	132	41	51	40	
Student Achievement					
Math SD Score	-0.03	0.19	-0.10	-0.19	
Reading SD Score	-0.03	0.21	-0.11	-0.19	
Student Covariates					
Economically Disadvantaged	0.59	0.36	0.69	0.68	
White	0.67	0.79	0.60	0.65	
Hispanic	0.25	0.13	0.32	0.28	
Black	0.01	0.01	0.01	0.01	
Special Education	0.17	0.16	0.17	0.17	
Grade 3	0.13	0.15	0.12	0.12	
Grade 4	0.14	0.16	0.13	0.12	
Grade 5	0.13	0.16	0.12	0.13	
Grade 6	0.13	0.13	0.14	0.13	
Grade 7	0.13	0.13	0.14	0.12	
Grade 8	0.13	0.12	0.14	0.12	
High School	0.21	0.15	0.21	0.27	

Table 1 - Summary Statistics for TIF-Participating Districts (2009-10)

Each entry is the mean for the given variable (unless otherwise noted).

Includes data on all OAKS test-takers for 2009-2010 in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer). Sample counts and student covariate means are for reading test-takers (similar for math test-takers).

	2011-12			2012-13			
	(1)	(2)	(3)	(4)	(5)	(6)	
First Stage	0.38 ***	0.39 ***	0.39 ***	0.43 ***	0.43 ***	0.44 ***	
C	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	
Reduced Form (Elgibility)	-0.03	-0.05	-0.04	-0.07	-0.08	-0.09 *	
	(0.09)	(0.08)	(0.05)	(0.08)	(0.06)	(0.05)	
First Stage F	9.04	9.10	9.12	10.98	11.05	11.04	
Obs	42,816	41,710	41,341	41,989	40,917	40,530	
Specifications							
Student Characteristics		Х	Х		Х	Х	
Grade Fixed Effects		Х	Х		Х	Х	
Test Score Mean 2009-10 (G	rade x School)		Х			Х	

Table 2A - RD Estimates of TIF Impact on Math Achievement (SD Scores)

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: TIF = $a + b^{(z)=0} + c^{z^{(z)}=0} + d^{z^{(z)}=0}$

Reduced Form regressions: $y = a + b^*(z \ge 0) + c^*z^*(z \ge 0) + d^*z^*(z < 0)$

	2011-12			2012-13			
	(1)	(2)	(3)	(4)	(5)	(6)	
First Stage	0.39 ***	0.39 ***	0.39 ***	0.40 ***	0.40 ***	0.41 ***	
C C	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	
Reduced Form (Elgibility)	0.16 **	0.12 **	0.10 **	0.08	0.06	0.05	
	(0.07)	(0.06)	(0.05)	(0.07)	(0.05)	(0.05)	
First Stage F	9.15	9.12	9.36	9.64	9.73	9.89	
Obs	42,730	41,567	41,199	41,974	40,853	40,466	
Specifications							
Student Characteristics		Х	Х		Х	Х	
Grade Fixed Effects		Х	Х		Х	Х	
Test Score Mean 2009-10 (G	rade x School)		Х			Х	

Table 2B - RD Estimates of TIF Impact on Reading Achievement (SD Scores)

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: $TIF = a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$

Reduced Form regressions: $y = a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$

		2011-12			2012-13	
	(1)	(2)	(3)	(4)	(5)	(6)
FULL SAMPLE						
First Stage	0.38 ***	0.39 ***	0.39 ***	0.43 ***	0.43 ***	0.44 ***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Reduced Form (Elgibility)	-0.03	-0.05	-0.04	-0.07	-0.08	-0.09 *
	(0.09)	(0.08)	(0.05)	(0.08)	(0.06)	(0.05)
Obs	42,816	41,710	41,341	41,989	40,917	40,530
20 < FRL < 80						
First Stage	0.37 ***	0.37 ***	0.38 ***	0.41 ***	0.41 ***	0.42 ***
	(0.13)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)
Reduced Form (Elgibility)	0.01	-0.01	-0.03	-0.04	-0.07	-0.10 *
	(0.10)	(0.09)	(0.06)	(0.09)	(0.07)	(0.05)
Obs	38,631	37,695	37,414	37,950	37,045	36,748
30 < FRL < 70						
First Stage	0.35 **	0.35 **	0.35 **	0.40 **	0.40 **	0.40 **
	(0.15)	(0.15)	(0.15)	(0.16)	(0.16)	(0.16)
Reduced Form (Elgibility)	-0.02	-0.02	-0.02	-0.09	-0.09	-0.11 **
	(0.12)	(0.10)	(0.06)	(0.10)	(0.07)	(0.05)
Obs	29,494	28,885	28,658	28,771	28,172	27,938
Specifications						
Student Characteristics		Х	Х		Х	Х
Grade Fixed Effects		Х	Х		Х	Х
Test Score Mean 2009-10 (Grad	e x School)		Х			Х

Table 3A - RD Estimates of TIF Impact on Math Achievement (SD Scores) for Various Bandwidths

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: TIF = $a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$

Reduced Form regressions: $y = a + b^*(z \ge 0) + c^*z^*(z \ge 0) + d^*z^*(z < 0)$

		2011-12			2012-13	
	(1)	(2)	(3)	(4)	(5)	(6)
FULL SAMPLE						
First Stage	0.39 ***	0.39 ***	0.39 ***	0.40 ***	0.40 ***	0.41 ***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Reduced Form (Elgibility)	0.16 **	0.12 **	0.10 **	0.08	0.06	0.05
	(0.07)	(0.06)	(0.05)	(0.07)	(0.05)	(0.05)
Obs	42,730	41,567	41,199	41,974	40,853	40,466
20 < FRL < 80						
First Stage	0.39 ***	0.38 ***	0.38 ***	0.39 ***	0.39 ***	0.39 ***
	(0.14)	(0.13)	(0.13)	(0.14)	(0.14)	(0.14)
Reduced Form (Elgibility)	0.24 ***	0.21 ***	0.18 ***	0.14 *	0.11 *	0.09
	(0.08)	(0.07)	(0.06)	(0.08)	(0.06)	(0.05)
Obs	38,462	37,500	37,220	37,758	36,832	36,536
30 < FRL < 70						
First Stage	0.38 **	0.37 **	0.36 **	0.39 **	0.39 **	0.39 **
	(0.16)	(0.15)	(0.15)	(0.16)	(0.15)	(0.15)
Reduced Form (Elgibility)	0.17 *	0.17 **	0.13 **	0.09	0.08	0.04
	(0.09)	(0.08)	(0.05)	(0.09)	(0.07)	(0.05)
Obs	29,389	28,752	28,526	28,626	28,009	27,774
Specifications						
Student Characteristics		Х	Х		Х	Х
Grade Fixed Effects		Х	Х		Х	Х
Test Score Mean 2009-10 (Grad	e x School)		Х			Х

Table 3B - RD Estimates of TIF Impact on Reading Achievement (SD Scores) for Various Bandwidths

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: TIF = $a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$

Reduced Form regressions: $y = a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$

176

								TIF Planning Year	
	2007	2007-08		8-09	2009	2009-10		2010-11	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
First Stage	0.36 ***	0.35 ***	0.38 ***	0.37 ***	0.36 ***	0.36 ***	0.39 ***	0.38 ***	0.39 ***
C	(0.12)	(0.12)	(0.12)	(0.13)	(0.12)	(0.12)	(0.13)	(0.13)	(0.13)
Reduced Form (Elgibility)	0.01	-0.02	0.09	0.05	0.00	-0.02	-0.01	-0.03	-0.04
	(0.08)	(0.07)	(0.07)	(0.06)	(0.08)	(0.07)	(0.07)	(0.07)	(0.05)
First Stage F	8.73	8.44	9.43	8.90	8.64	8.36	9.13	9.02	9.18
Obs	43,571	43,571	44,920	44,920	44,832	44,830	45,570	44,438	44,180
Specifications									
Student Characteristics		Х		Х		Х		Х	Х
Grade Fixed Effects		Х		Х		Х		Х	Х
Test Score Mean 2009-10 (C	Grade x School)								Х

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: $TIF = a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$

Reduced Form regressions: $y = a + b^*(z \ge 0) + c^*z^*(z \ge 0) + d^*z^*(z < 0)$

								TIF Planning Year	
	2007-08		2008	2008-09		9-10	2010-11		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
First Stage	0.40 ***	0.39 ***	0.42 ***	0.41 ***	0.39 ***	0.38 ***	0.38 ***	0.38 ***	0.38 ***
·	(0.12)	(0.12)	(0.13)	(0.13)	(0.12)	(0.12)	(0.13)	(0.13)	(0.13)
Reduced Form (Elgibility)	0.05	0.01	0.04	0.00	0.07	0.02	0.08	0.04	0.03
	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)	(0.07)	(0.06)	(0.05)
First Stage F	10.30	9.76	10.96	10.56	9.83	9.33	9.14	8.90	9.16
Obs	41,070	41,070	43,518	43,517	43,431	43,431	44,639	43,526	43,266
Specifications									
Student Characteristics		Х		Х		Х		Х	Х
Grade Fixed Effects		Х		Х		Х		Х	Х
Test Score Mean 2009-10 (Grade x School)								Х

Table 4B - RD Estimates of TIF Impact on Reading Achievement (SD Scores) in Pre-Treatment Period

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: TIF = $a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$

Reduced Form regressions: $y = a + b^*(z \ge 0) + c^*z^*(z \ge 0) + d^*z^*(z < 0)$

Standard errors clustered at the school level are in parentheses.
Table 5 - RD Estimates of TIF Impact on Test-Taker Covariates for Various Bandwidths

	2011-12				2012-13					
				Predicted	Predicted Score Index				Predicted Score Index	
	Econ Disadv	White	Special Ed	Math	Reading	Econ Disadv	White	Special Ed	Math	Reading
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FULL SAMPLE										
Reduced Form (Elgibility)	-0.01	0.12 ***	0.01	0.02	0.04	0.00	0.10 ***	0.02 *	0.02	0.02
	(0.03)	(0.04)	(0.01)	(0.03)	(0.02)	(0.03)	(0.03)	(0.01)	(0.03)	(0.02)
Obs	42,730	41,567	42,730	41,710	41,567	41,974	40,853	41,974	40,917	40,853
20 < FRL < 80										
Reduced Form (Elgibility)	-0.04	0.11 ***	0.02	0.03	0.04	-0.05	0.11 ***	0.02 *	0.04	0.04
	(0.03)	(0.04)	(0.01)	(0.03)	(0.03)	(0.03)	(0.04)	(0.01)	(0.03)	(0.03)
Obs	38,462	37,500	38,462	37,695	37,500	37,758	36,832	37,758	37,045	36,832
30 < FRL < 70										
Reduced Form (Elgibility)	-0.03	0.05	0.02	-0.01	0.01	-0.03	0.05	0.02	0.01	0.01
	(0.04)	(0.05)	(0.01)	(0.04)	(0.03)	(0.04)	(0.05)	(0.01)	(0.04)	(0.03)
Obs	29,389	28,752	29,389	28,885	28,752	28,626	28,009	28,626	28,172	28,009

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS reading test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

Columns 1, 2, 3 and 6,7,8 are estimated using OAKS reading test-takers (results are similar for OAKS math test-takers).

Reduced Form regressions: $y = a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$

	2011-12			2012-13			
	(1)	(2)	(3)	(4)	(5)	(6)	
Grade 3-5 Students in 2009-2010							
First Stage	0.46 ***	0.47 ***	0.47 ***				
	(0.15)	(0.15)	(0.15)				
Reduced Form (Elgibility)	0.04	0.01	-0.10				
	(0.08)	(0.07)	(0.06)				
First Stage F	9.95	10.05	10.38				
Obs	19,104	18,605	18,605				
Grade 3-6 Students in 2010-2011							
First Stage	0.46 ***	0.46 ***	0.46 ***	0.47 ***	0.47 ***	0.47 ***	
-	(0.13)	(0.13)	(0.13)	(0.15)	(0.15)	(0.15)	
Reduced Form (Elgibility)	0.07	0.02	0.02	0.00	-0.04	-0.07	
	(0.08)	(0.08)	(0.06)	(0.08)	(0.07)	(0.07)	
First Stage F	13.14	13.45	13.48	9.78	9.89	9.99	
Obs	20,155	19,651	19,651	18,476	18,019	18,019	
Specifications							
Student Characteristics		Х	Х		Х	Х	
Grade Fixed Effects		Х	Х		Х	Х	
Individual Student's Baseline Test Score			Х			Х	

Table 6A - RD Estimates of TIF Impact on Math Achievement (SD Scores) by Cohort

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: $TIF = a + b^*(z \ge 0) + c^*z^*(z \ge 0) + d^*z^*(z < 0)$

Reduced Form regressions: $y = a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$

	2011-12			2012-13			
	(1)	(2)	(3)	(4)	(5)	(6)	
Grade 3-6 Students in 2009-2010							
First Stage	0.46 ***	0.47 ***	0.47 ***				
	(0.15)	(0.15)	(0.15)				
Reduced Form (Elgibility)	0.25 ***	0.19 ***	0.13 **				
	(0.07)	(0.07)	(0.06)				
First Stage F	9.89	9.98	10.16				
Obs	18,980	18,487	18,487				
Grade 3-6 Students in 2010-2011							
First Stage	0.46 ***	0.46 ***	0.46 ***	0.46 ***	0.47 ***	0.47 ***	
	(0.13)	(0.13)	(0.13)	(0.15)	(0.15)	(0.15)	
Reduced Form (Elgibility)	0.27 ***	0.19 ***	0.10 ***	0.20 ***	0.12 *	0.00	
	(0.07)	(0.07)	(0.03)	(0.08)	(0.07)	(0.04)	
First Stage F	13.06	13.37	13.47	9.70	9.79	10.00	
Obs	20,003	19,504	19,504	18,347	17,892	17,892	
Specifications							
Student Characteristics		Х	Х		Х	Х	
Grade Fixed Effects		Х	Х		Х	Х	
Individual Student's Baseline Test Score			Х			Х	

Table 6B - RD Estimates of TIF Impact on Reading Achievement (SD Scores) by Cohort

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

First Stage regressions: $TIF = a + b^*(z \ge 0) + c^*z^*(z \ge 0) + d^*z^*(z < 0)$

Reduced Form regressions: $y = a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$

	2008-09	2008-09 to 2012-13		2011-12 and 2012-1	13
	(1)	(2)	(3)	(4)	(5)
Eligibility	0.04	0.01	-0.03	-0.05	-0.05
	(0.04)	(0.04)	(0.09)	(0.08)	(0.05)
Eligibility X (SY 2010-2011)	-0.04	-0.05			
	(0.09)	(0.08)			
Eligibility X (SY 2011-2012)	-0.07	-0.06			
	(0.10)	(0.08)			
Eligibility X (SY 2012-2013)	-0.10	-0.09	-0.03	-0.03	-0.04
	(0.09)	(0.07)	(0.12)	(0.10)	(0.07)
Obs	263,698	260,386	84,805	82,627	81,871
Specifications					
Student Characteristics		Х		Х	Х
Grade Fixed Effects		Х		Х	Х
Test Score Mean 2009-10 (Grade x School)					Х

Table 7A - Difference-in-RD Reduced-Form Estimates of TIF Impact on Math Achievement (SD Scores)

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers 2008-2009 to 2012-2013 in traditional public schools in TIF-participating districts

(Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

Pre-TIF years (2008-2009 to 2009-2010) are grouped together.

	2008-09 to 2012-13		2	2011-12 and 2012-13		
	(1)	(2)	(3)	(4)	(5)	
Eligibility	0.05	0.01	0.16 **	0.12 **	0.10 *	
	(0.03)	(0.03)	(0.07)	(0.06)	(0.05)	
Eligibility X (SY 2010-2011)	0.03	0.04				
	(0.07)	(0.07)				
Eligibility X (SY 2011-2012)	0.11	0.12 *				
	(0.08)	(0.07)				
Eligibility X (SY 2012-2013)	0.03	0.06	-0.08	-0.06	-0.06	
	(0.08)	(0.06)	(0.10)	(0.08)	(0.07)	
Obs	257,362	253,964	84,704	82,420	81,665	
Specifications						
Student Characteristics		Х		Х	Х	
Grade Fixed Effects		Х		Х	Х	
Test Score Mean 2009-10 (Grade x School)					Х	

Table 7B - Difference-in-RD Reduced-Form Estimates of TIF Impact on Reading Achievement (SD Scores)

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers 2008-2009 to 2012-2013 in traditional public schools in TIF-participating districts

(Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

Pre-TIF years (2008-2009 to 2009-2010) are grouped together.

	2011-2012			2012-2013			
	All TIF-		Excluding	All TIF-		Excluding	
	Participating	Excluding	Lebanon and	Participating	Excluding	Lebanon and	
	Districts	Lebanon	Redmond	Districts	Lebanon	Redmond	
	(1)	(2)	(3)	(4)	(5)	(6)	
First Stage	0.39 ***	0.36 ***	0.35 **	0.44 ***	0.42 ***	0.43 ***	
-	(0.13)	(0.13)	(0.14)	(0.13)	(0.14)	(0.15)	
Reduced Form (Elgibility)	-0.04	-0.04	-0.06	-0.09 *	-0.09 *	-0.09 *	
	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)	
First Stage F	9.12	7.56	6.54	11.04	9.42	8.68	
Obs	41,341	39,196	35,227	40,530	38,463	35,103	
Specifications							
Student Characteristics	Х	Х	Х	Х	Х	Х	
Grade Fixed Effects	Х	Х	Х	Х	Х	Х	
Test Score Mean 2009-10 (Grade x School)	Х	Х	Х	Х	Х	Х	

Table A1 - Math Achievement RD Results Excluding Districts with Possible TIF Treatment Contamination in 2012-2013

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

Lebanon school district ceased its participation in Spring 2013.

In 2012-2013, Lebanon and Redmond school districts participated in the SB 252 Pilot program connected to Oregon's NCLB Waiver application

First Stage regressions: $TIF = a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$

Reduced Form regressions: $y = a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$

	2011-2012			2012-2013			
	All TIF- Participating Districts	Excluding Lebanon	Excluding Lebanon and Redmond	All TIF- Participating Districts	Excluding Lebanon	Excluding Lebanon and Redmond	
	(1)	(2)	(3)	(4)	(5)	(6)	
First Stage	0.39 ***	0.37 ***	0.38 ***	0.41 ***	0.39 ***	0.40 ***	
	(0.13)	(0.13)	(0.14)	(0.13)	(0.13)	(0.14)	
Reduced Form (Elgibility)	0.10 **	0.11 **	0.10 *	0.05	0.05	0.06	
	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.05)	
First Stage F	9.36	7.96	7.19	9.89	8.50	7.95	
Obs	41,199	39,297	35,473	40,466	38,650	35,297	
Specifications							
Student Characteristics	Х	Х	Х	Х	Х	Х	
Grade Fixed Effects	Х	Х	Х	Х	Х	Х	
Test Score Mean 2009-10 (Grade x School)	Х	Х	Х	Х	Х	Х	

Table A2 - Reading Achievement RD Results Excluding Districts with Possible TIF Treatment Contamination in 2012-2013

Each entry represents the coefficient estimate of interest from a different RD regression.

Data includes all OAKS test-takers in traditional public schools in TIF-participating districts (Bend-Lapine, Crook, Greater Albany, Lebanon, and Salem-Keizer) that have baseline FRL data.

Lebanon school district ceased its participation in Spring 2013.

In 2012-2013, Lebanon and Redmond school districts participated in the SB 252 Pilot program connected to Oregon's NCLB Waiver application

First Stage regressions: $TIF = a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$

Reduced Form regressions: $y = a + b^{*}(z \ge 0) + c^{*}z^{*}(z \ge 0) + d^{*}z^{*}(z < 0)$