NBER WORKING PAPER SERIES

DESEGREGATION AND BLACK DROPOUT RATES

Jonathan Guryan

Working Paper 8345
http://www.nber.org/papers/w8345

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2001

The author thanks Daron Acemoglu, Joshua Angrist, David Autor, Adam Ashcraft, David Card, Aimee Chin, Sue Dynarski, Campe Goodman, Sean May, John Johnson, Steve Pischke, Jim Poterba, Canice Prendergast, Melissa Schettini, Jon Zinman, and participants at the MIT Public Finance/ Labor Seminar and the MIT Labor Lunch for their suggestions and guidance. Financial support was provided by the National Science Foundation through a Graduate Research Fellowship. The views expressed herein are those of the author and not necessarily those of the National Bureau of Economic Research.

Desegregation and Black Dropout Rates
Jonathan Guryan
NBER Working Paper No. 8345
June 2001
JEL No. H4, I2, J1, J7, N3

ABSTRACT

In 1954 the Supreme Court of the United States ruled that separate schools for black and white children were "inherently unequal." This paper studies whether the desegregation plans of the next 30 years in fact benefited the black students for whom the plans were designed. Analysis of data from the 1970 and 1980 censuses suggests that desegregation plans of the 1970's reduced the high school dropout rates of blacks by one to three percentage points during this decade. Desegregation plans can account for about half of the decline in dropout rates of blacks between 1970 and 1980. A similar analysis suggests that desegregation plans had no effect on the dropout rates of whites. The results are robust to controls for time-varying region and family income effects, as well as to tests for selective migration, though mean reversion may account for some portion of the larger estimated effects. Further investigation of conditions in segregated schools in 1970 suggests that peer effects explain at least some of the decline in the dropout rates of blacks due to desegregation plans.

Jonathan Guryan
University of Chicago GSB
1101 E. 58th Street
Chicago, IL 60637
and NBER
jonathan.guryan@gsb.uchicago.edu

1. Introduction

From *Plessy v. Ferguson*[1] in 1896 until *Brown v. Board of Education*[2] in 1954, Southern and Border States legally segregated their school systems by race. Black schools received fewer resources and black children were taught almost exclusively by black teachers. Outside the South, migration, housing patterns, and actions by state and local leaders contributed to similar racial isolation in the schools. With the *Brown* decision, the Supreme Court deemed segregated schools "inherently unequal" and therefore unconstitutional.

Over the next 30 years, federal courts ordered the implementation of desegregation plans for many of the largest school districts in the United States. It was the intent of these court orders to provide equal educational resources to blacks by eradicating segregation on the basis of race. Indeed, the desegregation of the public schools was the most significant innovation in the educational system of the post-World War II U.S. Nevertheless, few economists have studied the effect of desegregation on integration's intended beneficiaries, black students.[3] The fact that black high school dropout rates fell from the late 1960's through the early 1980's is documented in Figure I. The contribution of court-ordered integration to this decline is an important open question.

In this paper variation in the timing of desegregation plans in large school districts, the result of judicial enforcement of the *Brown* decision, is used to estimate integration's effect on black high school dropout rates. Using the 1970 and 1980

[1] 163 U.S. 537 (1896).
[2] 347 U.S. 483 (1954).
[3] Boozer, Krueger, and Wolkon (1992) and Card and Krueger (1992) are notable exceptions. St. John (1975) and Armor (1995) survey the education literature on the effect of desegregation on black achievement. They conclude that the results of studies, as well as the methods employed therein, vary significantly.

censuses, high-school-aged blacks in districts that desegregated between 1970 and 1980 are compared to those in districts that desegregated both before and after. Dropout rates among blacks declined between 1970 and 1980 by two to three percentage points in districts that desegregated in the interim relative to districts that desegregated both earlier and later. The results are robust to controls for district and time-varying region effects, controls for family income, and controls for potential confounds from the selective migration of blacks. Models that condition on lagged dropout rates produce estimates of a one percentage point decline in dropout rates as a result of desegregation. The smaller magnitude of these estimates suggests that mean reversion may account for some of the decline in dropout rates attributed to the effect of desegregation plans. As I describe in an appendix, however, under reasonable assumptions, estimates from the difference-in-differences and lagged dependent variable specifications provide upper and lower bounds on the effect of desegregation plans on dropout rates.[4] Therefore, on balance, the results reported here are consistent with a one to three percentage point decline in dropout rates due to desegregation.

Estimated effects are quite substantial. A one to three percentage point decline in dropout rates can account for about half of the decline in black dropout rates from 1970 to 1980. These estimates are also relevant to current policy issues. In September of 1999, a federal district judge in North Carolina ended the nearly thirty-year-old busing order in Charlotte.[5] In June of 1999, on the twenty-fifth anniversary of the busing order in Boston, white families filed suit claiming that racially based assignment plans in the Boston school system were unconstitutional.[6] The Boston School Committee then voted to end the use of race as a criterion for the assignment of students to schools. These and other decisions are made without the benefit of a comprehensive definitive study of the effect the original desegregation efforts had on black achievement.

The effects of desegregation are not only of interest to economists because of policy concerns. The forced integration of the public schools also provides an opportunity to evaluate how peers and school inputs affect students' educational outcomes. As a result of desegregation plans, large numbers of students lost control over

[4] Imbens, Liebman, and Eissa (1997) make a similar point.
[5] 57 F. Supp. 2d. 228 (1999).
[6] 379 F. Supp. 410 (1974), 1999 U.S. Dist. Lexis 12941 (1999).

the choice of both peers in school and the quality of the schools they attended. While the net effect of desegregation on student educational outcomes is interesting because of the inherent historical importance of integration, this reduced-form estimate is also of interest because it provides a well-identified empirical estimate of two economically important structural parameters.

There are large and growing literatures on both peer and school quality effects on economic outcomes. Theoretical work by Benabou (1993) on residential segregation, by Epple and Romano (1998) on private school tuition vouchers, and by Becker and Murphy (1992) on learning on the job, assumes that economic agents affect each other's productivity. Empirical evidence on peer effects is mixed, however. Case and Katz (1991) present evidence that a teenager's neighborhood peers affect various outcomes. The authors note, however, that this association need not be causal. Unobserved factors that determine a child's residential location also determine economic outcomes. Evans, Oates, and Schwab (1992) show that estimates of neighborhood peer effects that account for endogenous neighborhood choice are close to zero, while similar estimates that presume exogenous residential location yield large peer effect estimates. Cutler and Glaeser (1997) use an instrumental variables strategy and find evidence of significant neighborhood peer effects.

Empirical evidence on the effect of school quality on scholastic achievement and economic success is also mixed. Hanushek (1986) argues that there is little evidence that increased spending or improvements in crude measures of school quality have any effect on student performance. Krueger (1999) and Angrist and Lavy (1999) show direct evidence that class size impacts student test scores. Card and Krueger (1992) present direct evidence of a positive relationship between school quality and earnings, although Heckman, Layne-Farrar and Todd (1996) argue that the estimated relationship is driven by false assumptions. These studies notwithstanding, endogenous choice of residential location and educational expenditures generally makes the task of identifying the causal effect of school quality on economic outcomes difficult.

Desegregation provides a source of plausibly exogenous variation in both the peers and the resources at the schools students attend. Data limitations make it difficult to identify each effect separately, though a number of authors have noted that much of the

gap between black and white school inputs was closed before the desegregation plans of the 1970's were implemented.[7] Moreover, my examination of the characteristics of predominantly black and white schools in 1970 suggests that peer effects explain some portion of the overall decline in dropout rates attributed to desegregation. Together these findings imply that the benefits of integration could not be obtained simply by increasing observed school quality.

The paper is organized as follows. In the next section, I present a brief history of school desegregation and discuss the literature on the effectiveness of desegregation plans. Section 3 lays out the identification strategy. Section 4 describes the data. In Section 5, I present the results and examine potential threats to validity. In Section 6, I examine peer and school quality effects separately. Section 7 concludes.

2. Background

There was no nationally organized campaign that desegregated the public schools in the U.S. Rather, a series of court cases brought chiefly by private civil rights groups led to the court orders that were the most effective stimulus to desegregation. Political forces dictated that the enforcement of the *Brown* decision occurred mainly in the courts. What is germane here is that the courts' role in the process necessitated law enforcement on a case-by-case basis. As a result, the timing of integration varied at the school district level.

Although desegregation began first in the South, tabulations of the timing of major desegregation plans in the largest school districts in the country reveal significant inter- and intra-regional variation. Table I lists the sample of large school districts considered in this study.[8] School districts are listed chronologically by the year in which they instituted a major desegregation plan. There is a perhaps surprising amount of variation in the timing of desegregation plans not explained by regional differences. Of the 22 districts that implemented desegregation plans in the 1960's, seven are located

[7] See, for example, Card and Krueger (1992) or Margo (1990).
[8] I discuss the sampling procedure and data more specifically in Section IV. The sample of districts used is the same as in Welch and Light (1987).

outside of the South. Similarly, of the 77 districts that implemented plans in the 1970's, 35 are located outside of the South.

2.1. Legal History

A 1955 Supreme Court ruling, often called *Brown II*[9], gave the federal district courts responsibility for determining whether districts were segregated and for evaluating plans to remedy segregation. Litigation followed, mostly with the intent of allowing individual black students access to white schools, but very little integration occurred before the mid-1960's. The Civil Rights Act of 1964 gave the Department of Health, Education and Welfare (HEW) power to cut federal funding to school districts that discriminated on the basis of race. The law also authorized the Department of Justice to join school integration suits it deemed in the national interest. Many rural Southern school districts desegregated soon after the passage of the law.[10]

Some larger districts, where desegregation was more complex, allowed students the option of transferring schools within the district. In 1968, with its *Green v. New Kent County, Virginia*[11] decision, the Supreme Court outlawed plans that did not effectively integrate the schools, ordering the end to the use of so-called "freedom of choice" plans. The decision stimulated new litigation throughout the South. After *Green*, desegregation plans were more likely to include the pairing of nearby schools, or the redrawing of attendance zone boundaries. It was not until 1971, with the *Swann v. Charlotte-Mecklenburg County*[12] decision, that the Supreme Court approved busing of students outside of their neighborhood for the purpose of racial integration. The use of busing made desegregation feasible in many of the larger school districts of the North and West, where residential segregation was more severe.

Proving that school or state officials intentionally segregated the schools was necessary to win a legal battle, but finding such evidence was significantly more difficult outside of the South. In the absence of laws requiring separate school systems by race, plaintiffs had to sift through transcripts of school board meetings and interview school

[9] 349 U.S. 294 (1955).
[10] Orfield (1978) p. 279.
[11] 391 U.S. 430 (1968).

officials to produce evidence of intent to segregate. In 1973, the Court made it easier to warrant a desegregation order outside of the South when it ruled in *Keyes v. School District No. 1, Denver, Colorado*[13] that the plaintiff only needed to show segregative action in one school or neighborhood in the school district.

In 1974, the Court restricted remedies to include only school districts where plaintiffs could prove intentional segregation. The *Milliken v. Bradley*[14] ruling outlawed the inclusion of suburbs of Detroit in the city's busing plan.

By 1986, some school districts had been under court order for twenty years. In that year, a Fourth Circuit Court of Appeals declared the schools of Norfolk, Virginia integrated, ending the court order issued in 1971.[15] Decisions by the Supreme Court in 1991 and 1992 clarified what school districts could do once court orders were lifted, allowing districts to return to neighborhood schools.[16]

2.2. The Effectiveness of Desegregation Plans

Much research has focused on the effect of desegregation plans on the racial composition of school districts. Coleman (1975) suggested that court-ordered desegregation plans increased the speed of white migration out of cities. Subsequent research confirmed Coleman's claim, but also found that induced white migration was not extensive enough to offset fully the effect of desegregation plans on the integration of schools. In particular, Welch and Light (1987) show that desegregation plans of the 1960's, 1970's and 1980's decreased the index of dissimilarity in school districts by about 20 percentage points.[17] Rossell and Armor (1996) show that, net of effects on white enrollment, desegregation plans led to a 10 to 20 percentage point increase in the fraction of white students at the typical black student's school. There is reasonably strong evidence that desegregation plans led to a decrease in the segregation of public school districts. Through the reassignment of students to different schools, desegregation

---

[12] 402 U.S. 1 (1971).
[13] 413 U.S. 189 (1973).
[14] 418 U.S. 717 (1974).
[15] 784 F. 2d. 521 (4th Circuit 1986).
[16] 498 U.S. 237 (1991), 503 U.S. 467 (1992).

plans may have also improved the quality of educational resources available to black students.

2.3 The Role of Legal Precedent

Causes of the timing of desegregation are crucial to the identification strategy that will be explained in the following section. Desegregation plans were principally the ultimate result of court proceedings that were initiated by private civil rights groups. A concern that will arise is that these groups focused their energy and resources on school districts where desegregation would provide the greatest local benefit. If this were the case, then the timing of desegregation was at least in part a function of time-varying determinants of student achievement.

This argument overlooks the role of precedent in the U.S. legal process. Legal precedence played a major role in the determination of where to focus the resources of national civil rights groups, who where the primary force behind most legal challenges. Any agent working to benefit students nationally through desegregation must consider the long-standing effects of a legal failure early in the process. Thus, such national civil rights organizations chose to bring suit early in school districts where they felt there was a higher probability of victory, rather than in school districts where desegregation would benefit the largest number of students.

As is shown more formally in Appendix 2, when precedent has a strong effect on subsequent probability of success—as was the case with desegregation—an agent acting to desegregate the nation's schools should choose to weight the probability of success almost to the exclusion of any local benefits of desegregation when choosing where to bring legal challenges. The role of precedence suggests that any strategy to identify the effect of desegregation on dropout rates should focus on controlling for the idiosyncratic district characteristics that made legal victory more likely and thus led some districts to be desegregated earlier, while others were desegregated later.

[17] The index of dissimilarity can be thought of as the number of students that would have to move from their present school to fully integrate the district, relative to the number of students that would have to change schools to go from a fully segregated district to a fully integrated district.

3. Identification

Desegregation plans affect black dropout rates through three main channels. First, the reassignment of students within the school district affects the set of peers with which students attend school. New student assignment plans may cause parents to withdraw their children from the public schools, or to move out of the district altogether. Net of effects on the total enrollment and the racial composition of the district, desegregation plans alter the set of peers with which black children attend school.

Second, desegregation plans may move black students to better schools. If whites attended better schools than blacks did before integration, then on average desegregation should improve the quality of schools that blacks attend. Though total support for schools may decline as a result of desegregation-induced migration, integration may still lead to a change in the average quality of schools to which black students are assigned.

Third, there may be other effects of desegregation plans on black educational outcomes. Parents may become more involved in their children's education as a result of increased information, or in order to reap the benefits of the fight they have recently won. The legal victory that usually accompanies a desegregation plan may also make black children feel enfranchised. Any analysis of the effect of desegregation plans will estimate the net effect of these three changes.

The estimation of the net effect of desegregation plans on black educational outcomes is not completely straightforward, however. A comparison of integrated and segregated school systems at any point in time confounds the effect of the desegregation plans with the effect of factors that led to the imposition of the plan in the first place. The analysis to follow focuses on a sample of large school districts, 86 percent of which implemented desegregation plans between 1961 and 1982. Variation in the timing of the imposition of these plans is used to identify the effect of desegregation plans on black educational outcomes.

In its simplest form the identification strategy is to use a difference-in-differences estimator. An example should help to clarify. Consider two school districts: Birmingham, Alabama, which desegregated in 1970, and St. Louis, Missouri, which

desegregated in 1980. High-school-aged blacks are compared in the 1970 and 1980 censuses. In 1970, 17-year-olds in both cities had attended segregated schools all their lives. In 1980, 17-year olds in Birmingham had attended integrated schools since 2nd grade, while those in St. Louis had attended segregated schools throughout their education. The experiences in St. Louis are used to represent what would have happened in Birmingham in the absence of desegregation. A comparison of the change in the dropout rate in Birmingham relative to the change in the dropout rate in St. Louis is an estimate of the effect of desegregation.

In the full analysis, the treatment group comprises districts that implemented desegregation plans between 1970 and 1979. In these districts, high school students in April of 1980 attended at least one year of school after the implementation of a desegregation plan, while high school students in April of 1970 attended segregated schools throughout their education.[18] The change in dropout rates in these districts is compared to that in districts that desegregated before 1970 and after 1979. In the control districts, no desegregation plan was implemented between the time the 1970 and 1980 censuses were conducted. The analysis focuses on districts where there was ever a desegregation plan because these districts should be more comparable, but expanding the control group to encompass all districts that did not desegregate in the 1970's produces similar results.

The change in dropout rates is observed in districts that desegregated between 1970 and 1979, but it is unclear whether this change is due to desegregation or due to other factors that vary over time. Assuming that these other potential determinants of dropout status affect all school districts similarly, the change in dropout rates in the districts that desegregated both before 1970 and after 1980 is an estimate of what would have happened to dropout rates in districts that desegregated in the 1970's in the absence of desegregation. Suppose that the dropout rates of high-school-aged blacks can be written[19]

$$(1) \qquad E[D_i|t,g] = \beta_t + \gamma_g + \delta T_i$$

[18] Census data refer to April 1st of the census year.
[19] The notation that follows borrows heavily from Angrist and Krueger (1998).

where D_i is an indicator of dropout status of individual *i*, β_t is an effect for year *t* common to all school districts, γ_g is a time-constant effect that is allowed to vary by the decade of desegregation, indexed by *g*, T_i is an indicator for living in 1980 in a district that desegregated between 1970 and 1980, and δ is the effect of desegregation on dropout rates. The dropout status of high-school-aged blacks can now be written

$$D_i = \beta_t + \gamma_g + \delta T_i + \varepsilon_i \tag{2}$$

where ε_i is an error term such that $E[\varepsilon_i|t,g] = 0$. The simplest difference-in-differences comparison is

$$\{E[D_i|\ g=\text{treatment}, t=1980] - E[D_i|\ g=\text{treatment}, t=1970]\} - \{E[D_i|\ g=\text{control}, t=1980] - E[D_i|\ g=\text{control}, t=1970]\} = \delta. \tag{3}$$

Since T_i is equal to the product of a dummy that equals 1 for observations in 1980 and an indicator for districts that desegregated between 1970 and 1979, δ can be estimated with a simple regression of the model in (2). Suppose, however, that the timing of desegregation is correlated with time-varying determinants of dropout status. The regression framework allows convenient control for a vector of individual characteristics, X_i, by estimation of the equation,

$$D_i = X_i'\beta_0 + \beta_t + \gamma_g + \delta T_i + \varepsilon_i \tag{4}$$

where β_0 is a vector of coefficients that includes a constant. Validity of the estimate of δ now only requires that, conditional on X_i, inclusion in the treatment group is uncorrelated with unobserved time-varying determinants of dropout status.

Notice also that because there are many observations per school district, district effects can be included instead of the treatment main effect, γ_g. Since the treatment group is defined in terms of districts, the district indicators completely characterize the time-

constant effect specific to treatment districts and are thus more general. The model can now be written as,

$$(5) \qquad D_i = X_i'\beta_0 + \beta_t + \sum_{s=1}^{S-1} d_{is}\alpha_s + \delta T_i + \varepsilon_i$$

where d_{is} indicates that *i* resides in school district *s* and *S* denotes the total number of school districts in the sample.

The difference-in-differences estimator eliminates bias from any association between treatment and time-invariant characteristics of districts. But if desegregation is correlated with low past completion rates, i.e. a lagged dependent variable, then difference-in-differences estimates will exaggerate the effect of treatment.[20] A more general model would allow for fixed district effects and control for lagged district-level dropout rates. The model can be written as follows:

$$(6) \qquad D_{it} = X_{it}'\beta_0 + \beta_t + \sum_{s=1}^{S-1} d_{is}\alpha_s + \delta T_{it} + \gamma E[D_{it-1}|s] + \varepsilon_{it}.$$

Equation (6) is unidentified as written. Taking first differences and aggregating to the school district level yields

$$(7) \qquad y_{st} - y_{st-1} = (X_{st} - X_{st-1})'\beta_0 + \beta_t + \delta T_{st} + \gamma(y_{st-1} - y_{st-2}) + \varepsilon_{st} - \varepsilon_{st-1}$$

where y_{st} is the school district level dropout rate in year *t*. The model in equation (7) can be estimated using y_{st-2} as an instrument for $(y_{st-1} - y_{st-2})$.

In practice school districts cannot be identified in 1960, and thus the model with fixed effects that controls for lagged dropout rates cannot be estimated.[21] Therefore, estimates are reported that control for lagged dependent outcomes directly, without differencing or district effects. The model estimated in this case becomes,

$$(8) \qquad D_{it} = X_{it}'\beta_0 + \delta_{LD} g_i + \gamma E[D_{it-1}|s] + \varepsilon_{it}$$

[20] See, for example, Ashenfelter and Card (1985), and Appendix I.
[21] The public use 1960 census does not identify geographic areas smaller than states.

where g_i is an indicator for treatment districts. The estimate of δ_{LD} measures the conditional mean difference in dropout rates between treatment and control districts in 1980, controlling for dropout rates in 1970.

If the timing of desegregation plans is a function of either time-invariant district characteristics or lagged district-level dropout rates, then under reasonably plausible assumptions the differencing and lagged dependent variables estimates will provide an upper and lower bound of the true effect of desegregation on black dropout rates. Specifically, suppose selection into treatment is an increasing function of fixed determinants of dropout rates such that,

$$T_{st} = \begin{cases} 1 & \text{if } \alpha_s > \bar{y} \\ 0 & \text{otherwise} \end{cases}$$

where $\bar{y}$ is a constant. Call $\hat{\delta}_{LD}$ the estimator that controls for lagged dependent variables and estimates the treatment effect. Then, as is shown formally in Appendix 1,

$$\operatorname{plim} \hat{\delta}_{LD} = \delta + \frac{\operatorname{Cov}(\alpha_s, T_{st})}{\operatorname{Var}(\tilde{T}_{st})}\left(1 - \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\varepsilon^2}\right) \geq \delta$$

where $\tilde{T}_{st}$ is the residual from a regression of T_{st} on lagged dropout rates. Alternatively, suppose treatment is positively selected on lagged dropout rates, y_{st-1}, such that,

$$T_{st} = \begin{cases} 1 & \text{if } y_{st-1} > \bar{y} \\ 0 & \text{otherwise} \end{cases}$$

Call $\hat{\delta}_{DD}$ the difference-in-difference estimator of the treatment effect.

$$\operatorname{plim} \hat{\delta}_{DD} = \delta - \mathrm{E}\left[\varepsilon_{st-1} \middle| y_{st-1} > \bar{y}\right] \leq \delta$$

If treatment is *positively* selected either on lagged outcomes or on fixed determinants of lagged outcomes then, $\text{plim}\,\hat{\delta}_{DD} \leq \delta \leq \text{plim}\,\hat{\delta}_{LD}$. Thus, the two estimators should bracket the causal effect of interest.[22]

4. Data

In a report commissioned by the U.S. Commission on Civil Rights, Welch and Light (1987) evaluated the effect of desegregation plans on integration. They sampled 125 school districts for this purpose. This sample represents less than one percent of U.S. school districts, but about 20 percent of total enrollment and about half of minority enrollment in 1968.[23] The analysis to follow focuses on this sample of very large school districts. Table I lists the districts included in the study and when they implemented a desegregation plan. Column 3 shows whether the district is assigned to the treatment or control group.

I match 15-, 16- and 17-year-olds from the 1970 and 1980 censuses to the 125 districts based on individuals' county group or SMSA of residence. County group is the smallest geographic identifier available on public use census data files in 1970 and 1980.[24] Assigning individuals to districts is not straightforward, however, because the physical area described by county groups changed from 1970 to 1980. Maps provided by the census bureau are used to determine the smallest geographic area identifiable in both years.

Three different census samples cover the 125 geographic areas: the 1970 one-percent Metro sample, the 1980 one-percent Metro sample and the 1980 five-percent State sample. All three samples identify county groups, although the county groups in the State sample are slightly different because they do not cross state lines. The 1970 five-percent State sample does not identify any geographic area smaller than a state. The 1980 five-percent State sample identifies county groups and only SMSA's that do not cross state lines. Where possible, I use the 1980 State sample because of the larger

[22] The claim is derived more rigorously in Appendix I.
[23] Welch and Light (1987), p. 34.

sample size, but in many cases it is impossible to match the area identified in 1970 with the county groups or SMSA's identified in the 1980 State sample. Since the Metro samples are one-fifth the size of the State sample, observations are weighted to reflect the portion of the population they represent.

5. Results

As can be seen in the means presented in Table II, the high school dropout rate of blacks in districts that did not desegregate in the 1970's remained unchanged over that time period. In contrast, districts that desegregated between 1970 and 1980 experienced a decline in black dropout rates of 3.6 percentage points. The simplest difference-in-differences estimate as specified in equation (3) compares the change in dropout rates from 1970 to 1980 in districts that desegregated between 1970 and 1979 (-3.6 percentage points) to the change in dropout rates in districts that desegregated both before the 1970 census and after the 1980 census (0.2 percentage points). The relative change (-3.8 percentage points) is an estimate of the effect of desegregation plans on black dropout rates.

An alternative method of obtaining the simple difference-in-differences estimate is by estimation of equation (2). This method is useful because it allows convenient control for permanent characteristics that could potentially explain differential trends in dropout rates. Estimates in Table III show that neither changes in demographic characteristics nor changes in the effect of these characteristics on dropout rates can explain the decline in black dropout rates between 1970 and 1980 among districts that implemented a desegregation plan in the interim.

A nice feature of the 'experiment' is that there did not seem to be an appreciable shift in black dropout rates between 1970 and 1980 among districts that did not desegregate in the interim. As measured by the coefficient on the indicator for being an observation from 1980, the change in dropout rates from 1970 to 1980 among districts that *did not* desegregate during the decade is never significantly different from zero. The

[24] These groupings are very similar to PUMA's in the 1990 census.

results of a natural experiment are always more convincing when they are not measured amidst unexplained trends in the dependent variable.

The specification of the simple difference-in-differences model allows controls for fixed characteristics only at the treatment group level. There is significant variation in dropout rates across districts. Because the data include many observations per district, the model can be estimated controlling for district-level fixed differences in dropout rates. These fixed effects fully characterize the differences in dropout rates that existed in 1970. Estimates that allow pre-existing dropout rates to vary by school district, instead of only at the decade of desegregation level, produce nearly identical point estimates. These estimates, shown in Columns 4–6 of Table III, produce estimates that range from 3 – 3.4 percentage points, but are significantly more precise than those from the simple difference-in-differences model.

5.1. Regional Variation in Dropout Rates and in the Timing of Desegregation

The legal process that led to court-ordered desegregation plans proceeded differently in and out of the South. Additionally, there is significant regional variation in dropout rates. If not properly controlled for, region-specific trends in dropout rates may be attributed to the effect of desegregation plans. Such worries turn out to be unfounded as shown by the results from various specifications. Controls for permanent and time-varying region effects leave the results virtually unaffected (3.0 – 3.5 percentage point declines). Estimates that allow state-specific trends in dropout rates suggest a smaller decline (2 – 2.8 percentage points) in black dropout rates associated with desegregation. Alternatively, regressions run separately for each region, presented in Table IV, reveal that the estimated effect of desegregation plans on black dropout rates is remarkably consistent across regions. Estimates range from two to three percentage points in the Northeast to four to five percentage points in the West. The estimated effects in and out of the South are virtually the same, a 3 percentage point decline.

5.2. Two Control Groups

Another useful feature of the research design is that there are two obvious comparison groups that can be used to check the robustness of the estimates. If there are secular trends in the timing of desegregation plans that are correlated with determinants of dropout rates, then the estimates presented in Tables III and IV are biased. For instance, schools that desegregated later may have been less in need of compensatory intervention for black students. Natural convergence would then be spuriously attributed to the effect of desegregation. However, if such determinants of the timing of desegregation plans operated monotonically, then regressions that compare districts that desegregated in the 1970's to districts that desegregated in the 1960's and 1980's separately should yield more convincing estimates.

Another concern is that desegregation plans may not take effect immediately. Districts that desegregated in the late 1960's may not have felt the full effect of desegregation by 1970. These districts would have experienced some effect of desegregation between 1970 and 1980, leading to an underestimate of the effect of desegregation plans on the districts that desegregated in the 1970's.

Estimates that compare the districts that desegregated in the 1970's separately to districts that desegregated earlier and later, presented in Table V, suggest that the estimates from the standard specifications are reasonable. Estimated effects of desegregation plans on black dropout rates are slightly larger when comparing to districts that desegregated in the 1980's (3.2 – 3.9 percentage point declines compared to 2.3 – 3.0 percentage point declines), which lends weak support to each of the above concerns. However, the estimates do not differ significantly either from each other or from the estimates of the standard specifications in Table III.

5.3. Changing Demographics

Demographic characteristics may have changed differently in the districts that desegregated in the 1970's. For instance, if family income among high-school-aged

blacks increased more in these districts, then one would expect to see larger declines in dropout rates even if desegregation had no effect. Changing gender and age compositions could similarly explain the relative decline in dropout rates in the districts that desegregated in the 1970's. Additionally, the economic characteristics of black families may have remained relatively constant in all school districts, but the effect of family income on dropout rates may have changed over time. If, for instance, family income among blacks is permanently higher in districts that desegregated between 1970 and 1979 and the relationship between family income and dropout rates became stronger over time, imperfect controls for this shift would lead to spuriously negative estimates of desegregation on dropout rates.

The regression results presented in Columns 5 and 6 of Table III address these worries by controlling for the age, gender and family income of high-school-aged blacks and by allowing the effects of these characteristics on dropout rates to vary over time. Again, the estimated decline in black high school dropout rates due to desegregation plans remains consistently near three percentage points.

5.4. Selective Migration

Large-scale integration may have induced migration of blacks and whites. Assignment of observations to districts in the analysis thus far has been based on residence on April 1st of the current year. Validity of the estimates presumes that migration into and out of districts is not affected by desegregation plans. At the very least the analysis assumes that desegregation-induced migrants do not have different dropout propensities from the rest of the population in the absence of desegregation. While much has been written on desegregation-induced white migration—often termed "white flight"—little has been discussed about the parallel phenomenon for black families.

One might still worry that families of potential dropouts avoided desegregating areas when they moved, or that families of good students sought desegregating districts.

The former concern is likely to be more relevant as movers tended to be from lower-income families and had higher dropout rates.[25]

The data allow three empirical checks of whether selective migration did in fact bias the previous estimates. In both 1970 and 1980 there are questions on the public use census files asking whether individuals have moved during the previous five years. In Table VI, estimates are presented from models similar to those in Table III but with a discrete dependent variable that indicates whether the individual has moved from the county group where he lived five years ago. These results do not suggest that desegregation had any effect on migration. Columns 1–3 in Table VII present estimates from models that again use dropout status as the dependent variable, but control for whether the individual have moved from his county group of five years ago. Estimates of the effect of desegregation are unchanged from those in Table III.

While the analogous question is not asked in the 1970 census, there is a question asked of half of the 1980 sample that indicates individuals' county group of residence in 1975. Columns 4–6 of Table VII report estimates of models from Table III using county group in 1975 to assign districts to the 1980 data. Estimates of the effect of desegregation on black high school dropout rates are larger but more imprecise due to smaller sample sizes. Differential migration patterns into desegregating districts in the period from 1975 to 1980 do not seem to explain the estimated effects from the standard specifications in Table III.

The questions available on the public use census files address migration into districts, as opposed to migration out of districts. Welch and Light (1987) and Rossell and Armor (1996) show that the implementation of desegregation plans led to an increase in the speed of migration of whites out of urban school districts. If there is a similar phenomenon among blacks, the previous estimates are compromised. In fact, the problem is not as serious as it may seem because the districts referred to in this analysis are larger in physical area than actual school districts. In practice, districts as measured by county groups include many of the suburban areas to which migrants may have fled. Moreover, the data show no indication that desegregation plans led to a decrease in the

[25] The dropout rate in the sample among those who had moved from their county group of residence from five years ago is 18.3 percent, while that of non-movers is 11.7 percent. The average family income of movers is $11,608, while that of non-movers is $13,298.

population of high-school-aged blacks in districts. Estimates in Table VIII, together with those presented in Tables VI and VII, suggest that desegregation had little effect on migration in or out of districts.

5.5. Mean Reversion

Perhaps the most troubling concern with estimates in Table III is that mean reversion may explain the decline in dropout rates attributed to desegregation. The difference-in-differences strategy presumes that the timing of desegregation is dependent on permanent characteristics of districts. If that is the case, differencing over time eliminates the resulting permanent differences in dropout rates between districts that desegregated in the 1960's, 1970's and 1980's. On the other hand, the fact that dropout rates in 1970 were higher in districts that would desegregate in the 1970's than in other districts may lead to biased estimates. If the timing of desegregation was not determined by the fixed characteristics that led to permanently higher dropout rates, but rather on unusually high dropout rates in 1970, then the ensuing decline in dropout rates among desegregating districts may have been due to reversion to long-term means and not due to desegregation.

If the timing of desegregation was determined largely by recent lagged outcomes and not by permanent differences in dropout rates, then a more appropriate specification controls for 1970 dropout rates and compares dropout rates in 1980 of districts that desegregated in the ensuing decade to the dropout rates of those that would not. Estimates that control for lagged district-level dropout rates are presented in Table IX. The basic specification that controls for lagged outcomes associates a 2.2 percentage point decline in dropout rates with desegregation. Models that also control for age, region, gender and family income yield estimates of a smaller decline (0.9 – 1.4 percentage point). Estimated effects of desegregation on dropout rates are negative, but smaller in magnitude than those estimated by difference-in-differences models. The smaller magnitudes suggest that mean reversion may account for a large portion of the treatment effect estimated by the difference-in-differences estimator.

Controlling for lagged dependent variables assumes a linear relationship between dropout rates in 1970 and 1980 in the absence of changing desegregation status. An alternative specification matches districts based on 1970 dropout rates non-parametrically. For instance, the estimates in Columns 5–7 of Table IX are from a model in which the sample is separated based on deciles of the 1970 district-level dropout rate distribution computed separately by age. Dropout rates from 1980 of districts that desegregated in the 1970's are compared with dropout rates in 1980 of all other districts within each of these thirty strata, conditional on various characteristics. Estimates of full-sample coefficients are averages of the thirty coefficients weighted by the number of treatment observations in the stratum.[26] Matching estimates yield slightly smaller negative estimates of desegregation plans on black dropout rates.

Given that specifications that control for lagged dropout rates produce less negative estimates and that 1970 dropout rates are higher in districts that desegregated in the 1970's, it is possible that mean reversion can account for some of the difference-in-differences treatment effects. However, it is not clear that the estimators that control for 1970 dropout rates are more appropriate than the difference-in-differences estimator. If the timing of desegregation is selected on fixed characteristics, estimates that control for lagged outcomes will be biased in the direction of the permanent difference in dropout rates. Since dropout rates of districts that would desegregate in the 1970's are higher in 1970 than those of other school districts, estimates that control for dropout rates in 1970 will be biased positively. The claim is explained more rigorously in Appendix 1, but the implication is that under certain assumptions the lagged outcome and difference-in-differences estimators provide an upper and lower bound of the causal effect of desegregation plans on dropout rates. In particular, it is shown that if treatment is selected either on fixed (permanent) characteristics or on lagged outcomes then these two estimates will bracket the true value of the coefficient of interest—in this case the effect of desegregation on dropout rates.

An alternative view of the fixed effect model is that it is a version of the lagged dependent variable model with the coefficient on the lagged dropout rate fixed at unity. The argument above claims that if either model is misspecified, the estimate of the effect

[26] See Angrist and Lavy (1998) or Acemoglu and Angrist (1998) for an explanation of this estimator.

of lagged dropout rates on current dropout rates will be inconsistent. Another way to state the argument is that if the fixed effects model is estimated when the lagged dependent variable model is the correct one, setting the coefficient on the lagged dropout rate at unity is incorrect. And, if the lagged dependent variable model is specified when the fixed effect model is correct, the coefficient on the lagged dropout rate should be unity, but it is not estimated to be unity.

The coefficient on the lagged dropout rate is inconsistently estimated because the estimation compares the overlapping tails of two distributions. The district drawn from the distribution with the higher mean will on average revert to a higher outcome in the second period. Thus, unconditional on the distribution from which the district dropout rate was drawn—whether the district desegregated in the 1970's—the 1970 dropout rate does not predict the 1980 dropout rate with a coefficient of unity. In fact, some of the fixed difference between the dropout rates of the districts that desegregated in the 1970's and all other districts will be attributed to desegregation.

In reality, the true model for the determination of the timing of desegregation is probably somewhere between the lagged dropout rate model and the fixed effect model. As the discussion from Section 2 of the role of legal precedence illustrates, the timing of desegregation was the result of a complex process played out principally in the courts. Private civil rights groups played a major role in initiating legal challenges, which eventually led to the court orders that would effectuate desegregation. Concerns of the importance of legal precedence ensured that any agent of social change did not concentrate solely on the private benefits to the students in a school district when choosing a location for a legal challenge. The likelihood of legal victory—based on idiosyncratic characteristics of school districts—surely played a significant role in the determination of where to bring suit first. As a result, the timing of desegregation was based at least in part on idiosyncratic characteristics of school districts—most notably whether evidence of the intent to segregate was readily available. Thus, the true model of the timing of desegregation should consider both recent dropout rates and permanent school district characteristics.

The econometric argument above suggests that the two estimates bracket the true effect of desegregation. Thus, specifications that assume some mix of the two models

should yield estimates between those upper and lower bounds. One way to test this prediction is to estimate a specification that is implied by the discussion of the role of legal precedence—a partially differenced equation that constrains the effect of the lagged dropout rate to be various values ranging from 0 to 1.

Consider the econometric specification for the lagged dependent model that is expressed in equation (8). Table X presents estimates of this equation that constrain γ to be various values ranging from 0 to 1. When γ is held at unity, the model is essentially the same as the fixed-effect model. The fixed-effect specification corresponds to a model in which the timing of desegregation is determined solely by permanent school district characteristics. This model is realistic in a world where precedence is extremely important. When γ is held closer to zero, the model estimated is closer to the lagged dependent variable model. The lagged dropout model specification corresponds to a model in which the timing of desegregation is determined linearly by recent dropout rates. This model is realistic in a world where precedence is unimportant.

The estimates in Table X follow the pattern predicted by the bracketing argument. Estimates change monotonically from about one percentage point to about three percentage points as γ moves from zero to unity. If the timing of desegregation is dependent on a mix of recent dropout rates and permanent school district characteristics, then it seems that the causal effect of desegregation on black high school dropout rates falls somewhere between a one percentage point decline and a three percentage point decline.

5.6. The Effect of Desegregation Plans' Characteristics

While the ultimate goal of every desegregation plan was at some level the same, the methods each employed varied. Some plans redrew attendance zone boundaries in the district. Others assigned all children to the school closest to their home. Still others created schools, called magnets, that specialized in certain disciplines to attract students. More can be learned about the way desegregation plans affected black educational outcomes by examining differential effects of different plans.

An important distinguishing characteristic of a desegregation plan is whether it allows students any choice about which school to attend. Table XI presents difference-in-differences estimates from models that allow the effect of a desegregation plan to vary by whether the plan allowed choice. The results suggest that plans that assigned students to schools were associated with a larger decline in the dropout rates of blacks than plans that allowed choice.

5.7. Estimated Effects for Whites

There is reasonably consistent evidence that desegregation plans led to significant declines in high school dropout rates of blacks. The analysis thus far presumes that changes in unobservable characteristics of districts were not different in the districts that desegregated in the 1970's. One way to check this assumption is to perform the same analysis on whites. Additionally, the effect of desegregation plans on white dropout rates is interesting in its own right. Just as black students are placed in schools with white students, white students are placed in schools with black students. Just as black students are assigned to better schools, white students are assigned to worse schools.

Estimates of the standard specifications for high-school-aged whites are reported in Table XII. Whereas blacks experienced a decline in dropout rates as a result of desegregation plans, there did not seem to be any similar effect on whites. Estimated effects of desegregation plans on white dropout rates are all positive but insignificantly different from zero. Standard errors of the estimated effects are fairly small, but point estimates are too small to distinguish from zero.

As mentioned earlier, one might worry that desegregation-induced migration out of districts would bias the estimates. In fact, if desegregation caused white families with good students to leave the district then dropout rates among the remaining high-school-aged whites would be higher. Remember, however, that districts as defined in the analysis are larger in physical area than school districts. They include the suburban areas to which whites likely moved to avoid desegregation. So, white migration is not necessarily a large problem for the analysis. Indeed, examination of the effect of

desegregation plans on the population of high-school-aged whites in the district shows no evidence of desegregation-induced migration from districts as defined in this analysis.

The results for whites have two important implications. First, they lend credence to the results for blacks. If the results for whites had suggested that desegregation plans led to a decline in dropout rates of whites, one might think that district-specific trends in unobservables were driving the decline in black dropout rates. Second, it seems that desegregation plans did not have large effects on white dropout rates. Since black students are in the minority even in large school districts, whites should have, on average, experienced smaller changes in their set of peers and in the quality of schools they attended. Thus, it is not surprising that the effects of desegregation were smaller in magnitude on whites than on blacks. However, using the median point estimates from the standard specifications for blacks and whites, the analysis suggests that overall dropout rates declined as a result of desegregation plans.

6. School Quality or Peer Effects?

Given the apparent negative impact of desegregation on black dropout rates, a natural question is whether the improvement in black educational outcomes was a result of improved school quality or of interactions between blacks and whites. As mentioned in the introduction, the net effect of desegregation plans is interesting on its own, but desegregation also provides an opportunity to study the existence of human capital spillovers in schools and the effectiveness of school quality. In light of data limitations, however, it is not a simple task to identify each effect separately.

If school-level data on peers and resources were available, finer measures of the timing of desegregation plans could be used to identify peer and school quality effects on educational outcomes. Information that is available provides evidence that at least some of the effect of desegregation plans worked through their effect on the composition of students' peers in school. Data collected by the U.S. Department of Health, Education, and Welfare (HEW) in 1970 provide school-level full-time teacher and student enrollment counts by race for the largest school districts in the nation. With this data the

pupil-teacher ratio and the proportion of black students and teachers in each school can be computed.

One might suspect that in segregated school districts black students attended schools with higher pupil-teacher ratios. Coleman (1966) shows that, perhaps surprisingly, there were few large observable differences between the schools that black and white children attended in the mid-1960's. His findings are consistent with those of Card and Krueger (1992) who show that the convergence in observable school quality began well before the *Brown* decision in 1954, and that by the mid-1960's the schools that black and white students attended were observationally almost indistinguishable.

Indeed, the HEW data from 1970 suggest that, in districts that would desegregate in the ensuing decade, blacks did not attend schools with higher pupil-teacher ratios. If anything, schools that had proportionally more black students had lower pupil-teacher ratios. As seen in Table XIII, the average pupil-teacher ratio in schools with more than 75 percent black students was 26.7 while the average pupil-teacher ratio in schools with less than 25 percent black students was 27.9. The data do not seem to be consistent with the view that predominantly black schools had larger pupil-teacher ratios.

One observable dimension along which predominantly black and white schools did differ in 1970 was the characteristics of teachers. In school districts that were to desegregate in the next decade, predominantly black schools were much more likely to have black teachers. In schools with a more than 75-percent-black student body 56.5 percent of teachers were black. In schools with a less than 25-percent-black student body 7.9 percent of teachers were black. This relationship is striking. A regression of the fraction of black teachers in a school on the fraction of black students in a school yields a coefficient of 0.53 with a standard error of less than 0.01. This slope indicates that a 10 percentage point increase in the fraction of black students in a school was associated with a 5.3 percentage point increase in the fraction of black teachers in the school. What is also striking is that the relationship holds within school districts. Estimation of the same model including fixed school district effects shows that a 10 percentage point increase in the fraction of black students in a school was associated with a 4.5 percentage point increase in the fraction of black teachers in the school.

That black students in segregated school districts were more likely to be taught by black teachers is of interest for two reasons. First, black teachers may have had more or less education and experience than white teachers. Second, conditional on the skill level of the teacher his race may have an effect on his students. What evidence exists suggests that the latter effect is likely to be that a black teacher has a positive effect on black students.[27] Since desegregation likely caused black students to be less likely to be taught by black teachers, this effect cannot explain the decline in black dropout rates attributed to desegregation plans. [28] As for the former point, a comparison of black and white teachers in the March 1970 Current Population Survey (CPS), presented in Table XIV, shows that black teachers on average had 2.5 fewer years of potential experience and 1.5 fewer years of education.[29] What does this imply about the effect of desegregation on the average experience and education level of the typical black student's teacher?

Some simple notation will make the relevant calculation more clear. Let E_B and E_W be the population average level of education or experience of black and white teachers, respectively. Also, let N_{Bs} denote the enrollment of black students at school s and let τ_{Bs} denote the fraction of black teachers at school s. The fraction of black teachers at the typical black student's school is

$$\frac{\sum_s N_{Bs}\tau_{Bs}}{\sum_s N_{Bs}} \equiv \tau_B .$$

In segregated school districts the average education (or experience) level of the typical black student's teacher is

$$Q_S = \tau_B E_B + (1-\tau_B)E_W$$

assuming all teachers are either black or white. If school districts are fully integrated then the typical black student's teacher would have

[27] See Ehrenberg and Brewer (1994). Coleman (1966) finds little effect of the proportion of white teachers on the test scores of students.

[28] Freeman (1977) claims that desegregation led to a decline in demand for black teachers, but that a coincidental increase in black voting power offset this decline such that the relative employment of blacks in teaching remained stable.

$$Q_I = \tau_P E_B + (1 - \tau_P) E_W$$

years of education (or experience), where τ_P denotes the fraction of teachers in the population who are black. Using data from the 1970 census and from the 1970 HEW enrollment survey, I can calculate $Q_I - Q_S$ as an estimate of the effect of desegregation on the education (or experience) level of the typical black student's teacher.

As measured in the census, the average education levels of white and black teachers in 1970 were 15.9 and 15.6 years respectively. Similarly, white and black teachers in 1970 respectively had 16.4 and 15.5 years of potential experience. In school districts that would desegregate in the 1970's, the typical black student went to a school with 80 percent black teachers, while 29 percent of teachers in the population were black. These estimates imply that in these segregated school districts the typical black student was taught by teachers with 15.7 years of education and 15.7 years of experience. The estimates also imply that if these school districts were fully integrated the typical black student would be taught by teachers with 15.8 years of education and 16.1 years of experience.

Desegregation could have at most increased the educational attainment by 0.1 years and the experience level by 0.4 years of the typical black student's teacher. If these changes in the quality of schools that black students attended were to explain the decline in black dropout rates due to desegregation the combined elasticity of dropout rates with respect to teachers' education and experience levels would have to be near –9.

It seems reasonable to assume that changes in the characteristics of black students' teachers cannot explain the decline in black dropout rates in school districts that desegregated in the 1970's. The evidence suggests that pupil-teacher ratios were not larger in predominantly black schools. Similarly, previous work has shown that there were no marked differences in the term length and teacher salaries at schools that blacks and whites attended by the mid-1960's. Desegregation may have led to an improvement along other dimensions of the quality of schools that blacks attended. One might consider a student's peers one of those dimensions.

---

[29] The same relationship is true amongst teachers who lived in the SMSA's in which the school districts that desegregated in the 1970's are located.

Did the effect of desegregation plans on black dropout rates work through the plans' effect on the peer composition or the quality of the schools blacks attended? It seems that peer effects played at least some role. It may seem surprising that sitting in class next to a white student would induce a black student to finish high school. Race is a powerful proxy for socioeconomic status, however. It seems less surprising that a child whose parents are high school dropouts is more likely to stay in school because he attends school with children whose parents are high school graduates. This argument suggests that socioeconomic integration regardless of race may be at least as important as racial integration.

7. Conclusions

Despite desegregation's prominent role in post-World War II education policy, few economists have studied its impact on the educational outcomes of the affected students. Comparisons of the educational attainment of black students from segregated and integrated school systems confound the effect of desegregation with the determinants of desegregation itself. The analysis in this paper has exploited variation in the timing of desegregation plans to identify the effect of these plans on the high school dropout rates of blacks. Specifically, the change in black dropout rates from 1970 to 1980 in districts that desegregated in the interim is compared to the change in black dropout rates in districts that desegregated in the 1960's and the early 1980's. Using data from the 1970 and 1980 censuses, estimates control for time-varying region effects and for changes in family income across districts.

The results suggest that desegregation plans led to a one to three percentage point decline in the dropout rates of blacks, and that desegregation had little or no effect on the dropout rates of whites. Estimates from models that control for lagged dropout rates indicate that mean reversion may account for some of the apparent decline in dropout rates attributed to desegregation. It is shown, however, that under certain reasonable assumptions, the difference-in-differences and lagged dropout rate specifications should provide estimates that bracket the causal effect of desegregation on dropout rates.

The choice of econometric specification is largely dependent on the mechanism that determined the timing of desegregation. Most desegregation plans came as a result of court orders in legal proceedings brought by private civil rights groups. A simple model of legal precedence points out that these national organizations did not just consider the local benefits that would accrue to students when choosing where to bring a legal challenge. In addition, these civil rights organizations paid close attention to the likelihood that a legal challenge would result in victory. They tried hard to avoid cases that were likely to result in legal failure early on, even if the potential local benefit was large. Thus, the timing of desegregation was a function both of recent black dropout rates—a signal of the potential local benefit of desegregation—and of permanent district characteristics—signals of the likelihood of a legal victory.

Specifications that assume that the timing of desegregation was determined by a mix of lagged dropout rates and permanent district characteristics imply that desegregation led to a decline in black dropout rates between 1 and 3 percentage points in magnitude. A similar analysis shows not evidence that desegregation had any effect on white dropout rates.

Desegregation also provides an opportunity to study the effect of peers and school quality on the educational outcomes of students. The production function of a school has long been of interest to economists. Under most circumstances, the peers and quality of resources at a child's school are subject to his or his parents' choice. Subsequently, the estimation of peer and school quality effects is difficult in practice. Desegregation plans took away the ability to choose a child's peers and school resources. Accordingly, desegregation allows for estimation of the net effect of these two characteristics of schools. Data limitations make separate estimation of peer and school quality effects difficult, although examination of the conditions in segregated schools in 1970 suggests that peer effects had some role in the decline in dropout rates attributed to desegregation plans.

Further investigation of the mechanisms by which desegregation plans affected the educational outcomes of black students is clearly warranted. Other natural avenues for future work include an examination of the effect of desegregation plans on wages later in life, and an analysis of the termination of desegregation plans.

References

Acemoglu, Daron and Joshua Angrist, "Consequences of Employment Protection? The Case of the Americans with Disabilities Act," *NBER Working Paper No. 6670* (July 1998).

Angrist, Joshua D. and Victor Lavy, "Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools," *NBER Working Paper No. 6781* (November 1998).

Angrist, Joshua D. and Victor Lavy, "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics* (May 1999) pp. 533-575.

Angrist, Joshua D. and Alan B. Krueger, "Empirical Strategies in Labor Economics," *MIT Department of Economics Working Paper* (October 1998).

Armor, David J., *Forced Justice: School Desegregation and the Law*, Oxford University Press (New York: 1995).

Ashenfelter, Orley and David Card, "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs," *Review of Economics and Statistics* (November 1985) pp. 648-660.

Becker, Gary S. and Kevin M. Murphy, "The Division of Labor, Coordination Costs, and Knowledge," *Quarterly Journal of Economics* (November 1992) pp. 1137-1160.

Benabou, Roland, "Workings of a City: Location, Education, and Production," *Quarterly Journal of Economics* (August 1993) pp. 619-652.

Boozer, Michael A., Alan B. Krueger, and Shari Wolkon, "Race and School Quality Since *Brown v. Board of Education*," *Brookings Papers on Economic Activity: Microeconomics*, Martin N. Bailey and Clifford Winston, Eds. (1992) pp. 269-326.

Card, David and Alan B. Krueger, "School Quality and Black-White Relative Earnings: A Direct Assessment," *Quarterly Journal of Economics* (February 1992) pp. 151-200.

Card, David and Alan B. Krueger, "School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina," *Journal of Economic Perspectives* (Fall 1996) pp. 31-50.

Case, Anne C. and Lawrence F. Katz, "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths," *NBER Working Paper No. 3705* (May 1991).

Coleman, James S., *Equality of Educational Opportunity*, U. S. Department of Health, Education, and Welfare (Washington D. C.: 1966).

Coleman, James S., "Trends in School Segregation, 1968-1973," The Urban Institute (Washington, D.C.: August 1975).

Clotfelter, Charles T., "Are Whites Still Fleeing? Racial Patterns and Enrollment Shifts in Urban Public Schools, 1987-1996," *NBER Working Paper No. 7290* (August 1999).

Clotfelter, Charles T., "Public School Segregation in Metropolitan Areas," *NBER Working Paper No. 6779* (November 1998).

Crain, Robert and Rita Mahard, "The Effect of Research Methodology of Desegregation-Achievement Studies: A Meta-Analysis," *American Journal of Sociology* (Vol. 88 No. 5, 1983) pp. 839-854.

Cutler, David M. and Edward L. Glaeser, "Are Ghettos Good or Bad?," *Quarterly Journal of Economics* (August 1997) pp. 827-872.

Ehrenberg, Ronald G. and Dominic J. Brewer, "Do School and Teacher Characteristics Matter? Evidence from *High School and Beyond*," *Economics of Education Review* (Vol. 13 No. 1, 1994) pp. 1-17.

Epple, Dennis and Richard E. Romano, "Competition Between Private and Public Schools, Vouchers, and Peer-Group Effects," *American Economic Review* (March 1998) pp. 33-62.

Evans, William N., Wallace E. Oates and Robert M. Schwab, "Measuring Peer Group Effects: A Study of Teenage Behavior," *Journal of Political Economy* (October 1992) pp. 966-991.

Fancher, Betsy, "Voices From the South: Black Students Talk About Their Experiences in Desegregated Schools," Southern Regional Council (August 1970).

Finch, Minnie, *The NAACP: Its Fight for Justice,* The Scarecrow Press, Inc. (Metuchen, N.J.: 1981).

Hanushek, Eric A., "The Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economic Literature* (September 1986) pp. 1141-1177.

Heckman, James J., Robert J. LaLonde, and Jeffrey A. Smith, "The Economics and Econometrics of Active Labor Market Programs," *Handbook of Labor Economics, Vol. III*, Orley Ashenfelter and David Card Eds. (Amsterdam: 1999).

Heckman James, Anne Layne-Farrar and Petra Todd, "Does Measured School Quality Really Matter? An Examination of the Earnings-Quality Relationship," *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, Gary Burtless, Ed., Brookings Institution Press (Washington, D.C.: 1996) pp. 192-289.

Hill, Roscoe and Malcom Feeley, Eds., *Affirmative School Integration*, Sage Publications (Beverly Hills, CA: 1968).

Hoxby, Caroline, "The Effects of Class Size and Composition on Student Achievement: New Evidence from Natural Population Variation," mimeo (July 1996).

Imbens, Guido, Jeffrey B. Liebman, and Nada Eissa, "The Econometrics of Difference in Differences," mimeo (January 1997).

Jones, Leon, *From Brown to Boston: Desegregation in Education, 1954-1974: Vols. I & II*, The Scarecrow Press (Metuchen, N. J.: 1979).

Keynes, Edward and Randall K. Miller, *The Court vs. Congress: Prayer, Busing, and Abortion*, Duke University Press (Durham, NC: 1989).

Krueger, Alan B., "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics* (May 1999) pp. 497-532.

Manski, Charles F. "Academic Ability, Earnings, and the Decision to Become a Teacher: Evidence from the National Longitudinal Study of the High School Class of 1972," *Public Sector Payrolls*, David A. Wise Ed. (1987).

Margo, Robert A., *Race and Schooling in the South, 1880-1950*, University of Chicago Press (Chicago: 1990).

Miller, LaMar P. Ed., *Brown Plus Thirty: Perspectives on Desegregation*, Metropolitan Center for Educational Research, Development and Training, (New York: 1986).

Orfield, Gary, *Must We Bus? Segregated Schools and National Policy*, The Brookings Institution (Washington, D. C.: 1978).

Orfield, Gary, *Public School Desegregation in the United States, 1968-1980*, Joint Center for Political Studies (Washington, D. C.: 1983).

Orfield, Gary and Susan Eaton, *Dismantling Desegregation: The Quiet Reversal of* Brown v. Board of Education, The New Press (New York: 1996).

Peltzman, Sam, "The Political Economy of the Decline of American Public Education," *Journal of Law and Economics* (April 1993) pp. 331-370.

Peltzman, Sam, "Political Economy of Public Education: Non-College-Bound Students," *Journal of Law and Economics* (April 1996) pp. 73-120.

Rodgers, Harrell R., Jr. and Charles S. Bullock, III, *Law and Social Change: Civil Rights Laws and Their Consequences*, McGraw-Hill Book Company (New York: 1972).

Rossell, Christine and David Armor, "The Effectiveness of School Desegregation Plans, 1968-1991," *American Politics Quarterly* (July 1996) pp. 267-302.

U. S. Bureau of the Census, *Public Use Samples of Basic Records From the 1970 Census: Description and Technical Documentation* (Washington, D. C.: 1972).

U. S. Bureau of the Census, *Census of Population and Housing, 1980 (United States): Public Use Microdata Samples Technical Documentation* (Washington, D. C.: 1983).

U. S. Commission on Civil Rights, "Racial Isolation in the Public Schools," (1967).

U. S. Commission on Civil Rights, "Southern School Desegregation, 1966-67," (July 1967).

U. S. Commission on Civil Rights, "School Desegregation: The Courts and Suburban Migration," (December 1975).

U. S. Commission on Civil Rights, "Fulfilling the Letter and Spirit of the Law: Desegregation of the Nation's Public Schools," (August 1976).

U. S. Commission on Civil Rights, "Desegregation of the Nation's Public Schools: A Status Report," (February 1979).

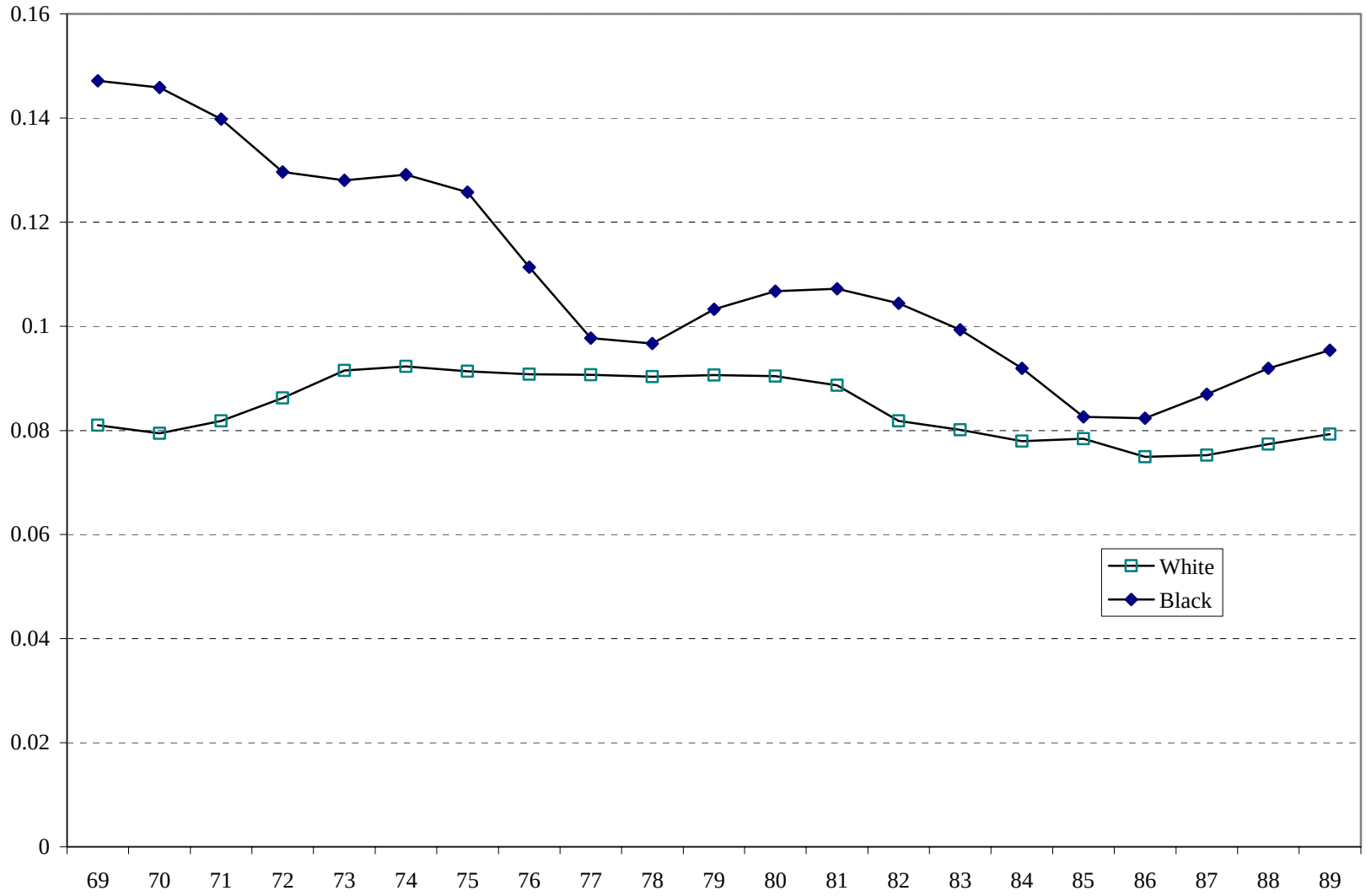
U. S. Department of Health, Education and Welfare, Office for Civil Rights, *Directory of Public Elementary and Secondary Schools in Selected Districts; Enrollment and Staff by Racial/Ethnic Group* (Fall 1970).

Welch, Finis and Audrey Light, "New Evidence on School Desegregation," *U. S. Commission on Civil Rights Clearinghouse Publication 92* (June 1987).

Wilder, Margaret G., "Black Assimilation in the Urban Environment; The Impact of Migration and Mobility," R & E Research Associates, Inc. (Palo Alto, CA: 1979).

Willis, Robert J. and Sherwin Rosen, "Education and Self-Selection," *Journal of Political Economy* (October 1979) pp. S7-S36.

FIGURE I: FRACTION OF 15-18-YEAR OLDS NOT ENROLLED IN SCHOOL, BY RACE



Note: Figure shows the fraction of 15-17 year olds not enrolled in school and 18-year-olds not enrolled in school and not high school graduates. Data are from the October Current Population Survey. Each data point represents the midpoint of a three-year moving average.

TABLE I: LIST OF SCHOOL DISTRICTS IN THE SAMPLE

School District		Year of Desegregation	Grade of Desegregation 1970	Grade of Desegregation 1980	Treatment or Control
NEW ORLEANS PARISH	LA	1961	Elem	Before	Control
NEWARK	NJ	1961	Elem	Before	Control
HARFORD COUNTY	MD	1965	JHS	Before	Control
OAKLAND	CA	1966	HS	Before	Control
HARTFORD	CT	1966	HS	Before	Control
GRAND RAPIDS	MI	1968	HS	Elem	Control
TACOMA	WA	1968	HS	Elem	Control
RICHMOND	CA	1969	HS	Elem	Control
BREVARD COUNTY	FL	1969	HS	Elem	Control
LEE COUNTY	FL	1969	HS	Elem	Control
PINELLAS COUNTY	FL	1969	HS	Elem	Control
POLK COUNTY	FL	1969	HS	Elem	Control
VOLUSIA COUNTY	FL	1969	HS	Elem	Control
CADDO PARISH	LA	1969	HS	Elem	Control
CALCASIEU PARISH	LA	1969	HS	Elem	Control
RAPIDES PARISH	LA	1969	HS	Elem	Control
TERREBONNE PARISH	LA	1969	HS	Elem	Control
CUMBERLAND COUNTY	NC	1969	HS	Elem	Control
NEW HANOVER COUNTY	NC	1969	HS	Elem	Control
SAN ANTONIO	TX	1969	HS	Elem	Control
PITTSYLVANIA COUNTY	VA	1969	HS	Elem	Control
BIRMINGHAM	AL	1970	After	Elem	Treatment
PASADENA	CA	1970	After	Elem	Treatment
STAMFORD	CT	1970	After	Elem	Treatment
BROWARD COUNTY	FL	1970	After	Elem	Treatment
DADE COUNTY	FL	1970	After	Elem	Treatment
PALM BEACH COUNTY	FL	1970	After	Elem	Treatment
EAST BATON ROUGE PARISH	LA	1970	After	Elem	Treatment
ROCHESTER	NY	1970	After	Elem	Treatment
GASTON COUNTY	NC	1970	After	Elem	Treatment
MECKLENBURG COUNTY	NC	1970	After	Elem	Treatment
CHARLESTON COUNTY	SC	1970	After	Elem	Treatment
GREENVILLE COUNTY	SC	1970	After	Elem	Treatment
RICHLAND COUNTY	SC	1970	After	Elem	Treatment
HOUSTON	TX	1970	After	Elem	Treatment
NORFOLK	VA	1970	After	Elem	Treatment
ROANOKE	VA	1970	After	Elem	Treatment
JEFFERSON COUNTY	AL	1971	After	Elem	Treatment
MOBILE	AL	1971	After	Elem	Treatment
LITTLE ROCK	AK	1971	After	Elem	Treatment
SAN FRANCISCO	CA	1971	After	Elem	Treatment

Note: Table lists school districts in the Welch and Light (1987) study, which are also the districts used in this study. Districts are chosen based on the following criteria, as described in Welch and Light (1987). Every district with 50,000 or more students in 1968 and 20 to 90 percent minority representation are included. Districts with 15,000 or more students in 1968 and ten to 90 percent minority representation were chosen with sampling probabilities proportional to their size and regional representation. The remaining districts—those with fewer than 15,000 students in 1968, less than ten percent minority representation—were excluded from the sample.

Grade of Desegregation columns identify the grade a 17-year-old in 1970 (1980) was in when the school district desegregated.

Year of Desegregation column reports the year the district's major desegregation plan was implemented according to Welch and Light (1987).

TABLE I (CONT.)

School District		Year of Desegregation	Grade of Desegregation 1970	Grade of Desegregation 1980	Treatment or Control
DUVAL COUNTY	FL	1971	After	Elem	Treatment
HILLSBOROUGH COUNTY	FL	1971	After	Elem	Treatment
MUSCOGEE COUNTY	GA	1971	After	Elem	Treatment
FORT WAYNE	IN	1971	After	Elem	Treatment
WICHITA	KS	1971	After	Elem	Treatment
JEFFERSON PARISH	LA	1971	After	Elem	Treatment
TULSA	OK	1971	After	Elem	Treatment
NASHVILLE	TN	1971	After	Elem	Treatment
DALLAS	TX	1971	After	Elem	Treatment
ARLINGTON COUNTY	VA	1971	After	Elem	Treatment
ORANGE COUNTY	FL	1972	After	Elem	Treatment
FAYETTE COUNTY	KY	1972	After	Elem	Treatment
LANSING	MI	1972	After	Elem	Treatment
CLARK COUNTY	NV	1972	After	Elem	Treatment
OKLAHOMA CITY	OK	1972	After	Elem	Treatment
AMARILLO	TX	1972	After	Elem	Treatment
ATLANTA	GA	1973	After	JHS	Treatment
ROCKFORD	IL	1973	After	JHS	Treatment
INDIANAPOLIS	IN	1973	After	JHS	Treatment
PRINCE GEORGE'S COUNTY	MD	1973	After	JHS	Treatment
CINCINNATI	OH	1973	After	JHS	Treatment
LAWTON	OK	1973	After	JHS	Treatment
MEMPHIS	TN	1973	After	JHS	Treatment
FORT WORTH	TX	1973	After	JHS	Treatment
WACO	TX	1973	After	JHS	Treatment
RALEIGH COUNTY	WV	1973	After	JHS	Treatment
DENVER	CO	1974	After	JHS	Treatment
BALTIMORE	MD	1974	After	JHS	Treatment
BOSTON	MA	1974	After	JHS	Treatment
SPRINGFIELD	MA	1974	After	JHS	Treatment
MINNEAPOLIS	MN	1974	After	JHS	Treatment
PORTLAND	OR	1974	After	JHS	Treatment
JEFFERSON COUNTY	KY	1975	After	JHS	Treatment
DETROIT	MI	1975	After	JHS	Treatment
SACRAMENTO	CA	1976	After	HS	Treatment
NEW BEDFORD	MA	1976	After	HS	Treatment
OMAHA	NB	1976	After	HS	Treatment
JERSEY CITY	NJ	1976	After	HS	Treatment
DAYTON	OH	1976	After	HS	Treatment
MILWAUKEE	WI	1976	After	HS	Treatment

Note: Table lists school districts in the Welch and Light (1987) study, which are also the districts used in this study. Districts are chosen based on the following criteria, as described in Welch and Light (1987). Every district with 50,000 or more students in 1968 and 20 to 90 percent minority representation are included. Districts with 15,000 or more students in 1968 and ten to 90 percent minority representation were chosen with sampling probabilities proportional to their size and regional representation. The remaining districts—those with fewer than 15,000 students in 1968, less than ten percent minority representation—were excluded from the sample.

Grade of Desegregation columns identify the grade a 17-year-old in 1970 (1980) was in when the school district desegregated.

Year of Desegregation column reports the year the district's major desegregation plan was implemented according to Welch and Light (1987).

TABLE I (CONT.)

School District		Year of Desegregation	Grade of Desegregation 1970	1980	Treatment or Control
SAN DIEGO	CA	1977	After	HS	Treatment
KANSAS CITY	KS	1977	After	HS	Treatment
KANSAS CITY	MO	1977	After	HS	Treatment
AKRON	OH	1977	After	HS	Treatment
FRESNO	CA	1978	After	HS	Treatment
LOS ANGELES	CA	1978	After	HS	Treatment
SAN BERNARDINO	CA	1978	After	HS	Treatment
NEW CASTLE COUNTY	DE	1978	After	HS	Treatment
PHILADELPHIA	PA	1978	After	HS	Treatment
EL PASO	TX	1978	After	HS	Treatment
LUBBOCK	TX	1978	After	HS	Treatment
SEATTLE	WA	1978	After	HS	Treatment
TUCSON	AZ	1979	After	HS	Treatment
CLEVELAND	OH	1979	After	HS	Treatment
COLUMBUS	OH	1979	After	HS	Treatment
LONG BEACH	CA	1980	After	After	Control
DOUGHERTY COUNTY	GA	1980	After	After	Control
ST. LOUIS	MO	1980	After	After	Control
BUFFALO	NY	1980	After	After	Control
TOLEDO	OH	1980	After	After	Control
PITTSBURGH	PA	1980	After	After	Control
AUSTIN	TX	1980	After	After	Control
SAN JOSE	CA	1981	After	After	Control
SOUTH BEND	IN	1981	After	After	Control
CHICAGO	IL	1982	After	After	Control
ECTOR COUNTY	TX	1982	After	After	Control
MESA	AZ	None			
MODESTO	CA	None			
VALLEJO	CA	None			
PUEBLO	CO	None			
GARY	IN	None			
SAGINAW	MI	None			
ALBUQUERQUE	NM	None			
LAS CRUCES	NM	None			
NEW YORK	NY	None			
LORAIN	OH	None			

Note: Table lists school districts in the Welch and Light (1987) study, which are also the districts used in this study. Districts are chosen based on the following criteria, as described in Welch and Light (1987). Every district with 50,000 or more students in 1968 and 20 to 90 percent minority representation are included. Districts with 15,000 or more students in 1968 and ten to 90 percent minority representation were chosen with sampling probabilities proportional to their size and regional representation. The remaining districts—those with fewer than 15,000 students in 1968, less than ten percent minority representation—were excluded from the sample.

Grade of Desegregation columns identify the grade a 17-year-old in 1970 (1980) was in when the school district desegregated.

Year of Desegregation column reports the year the district's major desegregation plan was implemented according to Welch and Light (1987).

TABLE II: MEANS OF SELECTED VARIABLES BY TREATMENT-CONTROL STATUS

Variables	Full Sample	1970			1980		
		Full	Desegregated 1970-1979	Control	Full	Desegregated 1970-1979	Control
Dropout	.120	.135	.143	.116	.110	.107	.118
	(.325)	(.342)	(.351)	(.321)	(.313)	(.309)	(.323)
Female	.50	.50	.50	.50	.50	.50	.51
	(.50)	(.50)	(.50)	(.50)	(.50)	(.50)	(.50)
Age 16	.33	.32	.33	.31	.33	.33	.33
	(.47)	(.47)	(.47)	(.46)	(.47)	(.47)	(.47)
Age 17	.32	.32	.32	.24	.33	.33	.34
	(.47)	(.46)	(.47)	(.43)	(.47)	(.47)	(.47)
Northeast	.11	.12	.10	.17	.11	.09	.16
	(.32)	(.33)	(.30)	(.37)	(.31)	(.28)	(.36)
Midwest	.29	.31	.25	.44	.28	.22	.42
	(.45)	(.46)	(.43)	(.50)	(.45)	(.41)	(.49)
South	.48	.44	.50	.32	.50	.55	.36
	(.50)	(.50)	(.50)	(.47)	(.50)	(.50)	(.48)
West	.11	.11	.12	.08	.11	.13	.07
	(.31)	(.31)	(.33)	(.26)	(.32)	(.34)	(.26)
Family Income	13,218	7,815	7,870	7,695	16,924	17,117	16,439
	(11,598)	(5,621)	(5,703)	(5,435)	(13,090)	(13,087)	(13,087)
Poverty	171	157	157	158	181	184	175
	(128)	(115)	(115)	(115)	(135)	(135)	(134)
Year of Desegregation	1973.8	1973.8	1973.8	1973.8	1973.7	1973.6	1974.0
	(5.1)	(5.2)	(3.1)	(8.2)	(5.0)	(3.1)	(8.0)
No. Obs. (unweighted count)	53,331	7,256	5,019	2,237	46,075	31,827	14,248

Note: Data are weighted using population weights. Desegregated 1970-1979 group is 15-, 16-, and 17-year-old blacks who live in districts that desegregated between 1970 and 1979. Control group is all other 15-, 16, and 17-year-old blacks in districts that desegregated between 1961 and 1982. Dropout is an indicator for not being enrolled in school. Poverty measures what percentage of the poverty line the individual's family income is.

TABLE III: DIFFERENCE-IN-DIFFERENCES ESTIMATES OF THE EFFECT OF DESEGREGATION ON DROPOUT RATES OF BLACKS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Desegregated 1970-1979 *1980	-.038 (.011)	-.035 (.011)	-.035 (.009)	-.034 (.005)	-.030 (.005)	-.030 (.005)	-.028 (.010)	-.021 (.007)
Desegregated 1970-1979	.027 (.008)	.026 (.008)	.023 (.007)				.044 (.008)	
1980	.001 (.008)	-.003 (.008)	-.004 (.007)	-.004 (.003)	.014 (.039)	.023 (.038)	.008 (.015)	.001 (.010)
Age 16		.048 (.004)	.048 (.004)	.048 (.004)	.048 (.004)	.048 (.004)	.048 (.004)	.048 (.004)
Age 17		.166 (.006)	.166 (.006)	.166 (.006)	.162 (.006)	.162 (.006)	.162 (.006)	.162 (.006)
Female					.011 (.004)	.022 (.008)	.011 (.004)	.011 (.004)
Female *1980						-.018 (.009)		
Poverty/10^3					-.198 (.031)	-.025 (.048)	-.196 (.030)	-.194 (.030)
Poverty/10^3 *1980						-.239 (.058)		
Family Income/10^6					-.028 (.028)	-.038 (.013)	-.031 (.028)	-.033 (.028)
Family Income/10^6 *1980						.042 (.013)		
Northeast			-.060 (.030)					
Northeast *1980					-.019 (.039)	-.017 (.039)		
Midwest			-.059 (.030)					
Midwest *1980					-.014 (.039)	-.014 (.039)		
South			-.046 (.030)					
South *1980					-.008 (.038)	-.008 (.038)		
West			-.084 (.030)					
West *1980					-.007 (.040)	-.004 (.039)		
District Effects	No	No	No	Yes	Yes	Yes	No	Yes
State Effects	No	No	No	No	No	No	Yes	Yes
State *1980 Effects	No	No	No	No	No	No	Yes	Yes
R^2	.002	.048	.049	.056	.061	.062	.060	.064
No. Obs.	53,331	53,331	53,331	53,331	52,416	52,416	52,416	52,416

Note: Data are weighted using population weights. Desegregated 1970-1979 group is 15-, 16-, and 17-year-old blacks who live in districts that desegregated between 1970 and 1979. Control group is all other 15-, 16, and 17-year-old blacks in districts that desegregated between 1961 and 1982. Standard error estimates are corrected for district*year correlation in the error term.

TABLE IV: MODELS RUN SEPARATELY BY REGION: DIFFERENCE-IN-DIFFERENCES ESTIMATES OF THE EFFECT OF DESEGREGATION ON DROPOUT RATES OF BLACKS

	Northeast		Midwest		West		South		Exclude South	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Desegregated 1970-1979 *1980	-.031 (.026)	-.021 (.007)	-.025 (.012)	-.027 (.009)	-.048 (.024)	-.042 (.010)	-.028 (.015)	-.029 (.010)	-.033 (.015)	-.029 (.005)
Desegregated 1970-1979	.020 (.018)		.014 (.008)		.052 (.015)		.024 (.011)		.019 (.011)	
1980	-.013 (.026)	.008 (.020)	-.002 (.006)	.019 (.013)	.020 (.015)	.003 (.013)	.004 (.011)	.012 (.011)	-.001 (.011)	.009 (.010)
Age 16	.063 (.010)	.064 (.010)	.050 (.007)	.050 (.007)	.025 (.006)	.026 (.006)	.045 (.006)	.045 (.006)	.049 (.005)	.050 (.005)
Age 17	.182 (.016)	.183 (.016)	.172 (.009)	.172 (.010)	.122 (.012)	.123 (.012)	.158 (.008)	.159 (.008)	.164 (.008)	.165 (.008)
Female	.007 (.007)	.022 (.012)	.019 (.009)	.028 (.017)	.010 (.009)	.004 (.020)	.009 (.006)	.027 (.012)	.013 (.006)	.017 (.012)
Female *1980		-.026 (.015)		-.015 (.019)		.011 (.020)		-.030 (.013)		-.007 (.012)
Poverty/10^3	-.227 (.063)	-.248 (.117)	-.148 (.006)	.015 (.046)	-.130 (.037)	-.083 (.097)	-.268 (.049)	.024 (.116)	-.170 (.031)	-.056 (.047)
Poverty/10^3 *1980		-.038 (.136)		-.290 (.059)		-.046 (.100)		-.323 (.133)		-.179 (.058)
Family Income/10^5	.047 (.060)	.193 (.318)	-.035 (.046)	-.284 (.159)	-.031 (.035)	-.218 (.253)	-.010 (.047)	-.743 (.255)	-.018 (.031)	-.190 (.138)
Family Income/10^6 *1980		-.104 (.326)		.353 (.161)		.190 (.256)		.777 (.261)		.232 (.140)
District Effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
R^2	.063	.069	.059	.063	.046	.058	.052	.061	.057	.064
No. Obs.	6,100	6,100	15,691	15,691	6,567	6,567	23, 885	23,885	28,531	28,531

Note: Data are weighted using population weights. Desegregated 1970-1979 group is 15-, 16-, and 17-year-old blacks who live in districts that desegregated between 1970 and 1979. Control group is all other 15-, 16, and 17-year-old blacks in districts that desegregated between 1961 and 1982. Standard error estimates are corrected for district*year correlation in the error term.

TABLE V: SPECIFICATIONS WITH ALTERNATIVE CONTROL GROUPS: DIFFERENCE-IN-DIFFERENCES ESTIMATES OF THE EFFECT OF DESEGREGATION ON DROPOUT RATES OF BLACKS

	Control Desegregated After 1980		Control Desegregated Before 1970	
	(1)	(2)	(3)	(4)
Desegregated 1970-1979 *1980	-.039	-.032	-.030	-.023
	(.007)	(.007)	(.005)	(.005)
1980	.001	.006	-.008	-.003
	(.006)	(.006)	(.002)	(.004)
Age 16	.046	.045	.048	.047
	(.005)	(.004)	(.005)	(.004)
Age 17	.165	.160	.168	.165
	(.007)	(.007)	(.006)	(.006)
Female		.008		.013
		(.004)		(.004)
Poverty/10^3		-.198		-.200
		(.034)		(.033)
Family Income/10^6		-.026		-.028
		(.031)		(.030)
District Effects	Yes	Yes	Yes	Yes
R^2	.057	.065	.057	.063
No. Obs.	44,694	43,888	45,843	44,684

Note: Data are weighted using population weights. Desegregated 1970-1979 group is 15-, 16-, and 17-year-old blacks who live in districts that desegregated between 1970 and 1979. Control group is all other 15-, 16, and 17-year-old blacks in districts that desegregated between 1961 and 1982. In Columns 1 and 2, the Control group is 15-, 16, and 17-year-old blacks in districts that desegregated *after* the Treatment districts (after 1979). In Columns 3 and 4, the Control group is 15-, 16, and 17-year-old blacks in districts that desegregated *before* the Treatment districts (before 1970). Standard error estimates are corrected for district*year correlation in the error term.

TABLE VI: DIFFERENCE-IN-DIFFERENCES ESTIMATES OF THE EFFECT OF DESEGREGATION ON MIGRATION

	(1)	(2)	(3)	(4)
Desegregated 1970-1979 *1980	-.003 (.016)	-.004 (.012)	-.005 (.006)	-.005 (.005)
Desegregated 1970-1979	.012 (.013)	.006 (.010)		
1980	-.029 (.014)	-.030 (.010)	-.022 (.005)	-.017 (.070)
Age 16		.008 (.003)	.007 (.003)	.007 (.003)
Age 17		.004 (.003)	-.001 (.003)	-.001 (.003)
Female			.003 (.003)	.003 (.003)
Poverty/10^3			.020 (.029)	.017 (.029)
Family Income/10^6			-.055 (.030)	-.051 (.030)
Northeast		-.000 (.020)		
Northeast *1980				-.007 (.070)
Midwest		-.004 (.019)		
Midwest *1980				-.005 (.069)
South		.014 (.019)		
South *1980				.002 (.069)
West		.045 (.022)		
West *1980				-.030 (.070)
District Effects	No	No	Yes	Yes
R^2	.005	.010	.028	.028
No. Obs.	53,331	53,331	52,416	52,416

Note: Data are weighted using population weights. Desegregated 1970-1979 group is 15-, 16-, and 17-year-old blacks who live in districts that desegregated between 1970 and 1979. Control group is all other 15-, 16, and 17-year-old blacks in districts that desegregated between 1961 and 1982. Standard error estimates are corrected for district*year correlation in the error term.

TABLE VII: DIFFERENCE –IN-DIFFERENCES ESTIMATES OF THE EFFECT OF DESEGREGATION ON DROPOUT RATES CONTROLLING FOR MIGRATION

	Control For Moving			Use Residence in 1975 for 1980		
	(1)	(2)	(3)	(4)	(5)	(6)
Desegregated 1970-1979 *1980	-.035 (.011)	-.034 (.005)	-.027 (.005)	-.045 (.015)	-.048 (.010)	-.039 (.009)
Desegregated 1970-1979	.025 (.008)			.025 (.008)		
1980	-.001 (.008)	-.002 (.003)	-.002 (.004)	.018 (.012)	.032 (.008)	.039 (.007)
Age 16	.048 (.004)	.048 (.004)	.048 (.004)	.050 (.006)	.052 (.006)	.051 (.006)
Age 17	.165 (.006)	.165 (.006)	.162 (.006)	.174 (.008)	.175 (.008)	.170 (.008)
Moved	.061 (.012)	.061 (.012)	.038 (.011)			
Female			.011 (.004)			.016 (.006)
Poverty/10^3			-.197 (.031)			-.140 (.040)
Family Income/10^6			-.269 (.280)			-.129 (.044)
District Effects	No	Yes	Yes	No	Yes	Yes
R^2	.050	.058	.062	.049	.060	.064
No. Obs.	53,331	53,331	52,416	21,669	21,669	21,222

Note: Data are weighted using population weights. Desegregated 1970-1979 group is 15-, 16-, and 17-year-old blacks who live in districts that desegregated between 1970 and 1979. Control group is all other 15-, 16, and 17-year-old blacks in districts that desegregated between 1961 and 1982. Columns 4-6 present regressions where the district of residence for 1980 observations is defined as the district of residence in 1975. Standard error estimates are corrected for district*year correlation in the error term.

TABLE VIII: ESTIMATED EFFECTS OF DESEGREGATION PLANS ON DISTRICT ENROLLMENT BY RACE

	1970	1980	1980-1970	Difference-in-Differences
Black Enrollment				
Desegregated in 70's	20,484	18,789	-1,695	
				410
Desegregated in 60's, 80's	21,369	19,264	-2,105	(8,872)
White Enrollment				
Desegregated in 70's	78,083	71,741	-6,342	
				2,935
Desegregated in 60's, 80's	78,920	69,643	-9,277	(25,534)

Note: Estimated enrollment counts are based on weighted counts of 15-, 16-, and 17-year-olds from the 1970 and 1980 censuses. Standard errors of the difference-in-differences estimated effect of desegregation plan on enrollment are reported in parentheses.

TABLE IX: SPECIFICATIONS THAT CONTROL FOR LAGGED DROPOUT RATES: ESTIMATES OF THE EFFECT OF DESEGREGATION ON DROPOUT RATES

	Lagged Dropout Rate				Matching		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Desegregated 1970-1979	-.022 (.006)	-.012 (.007)	-.014 (.007)	-.009 (.006)	-.009 (.004)	-.011 (.007)	-.007 (.007)
Lagged Dropout Rate	.548 (.058)	.108 (.048)	.086 (.048)	.055 (.045)			
Age 16		.043 (.005)	.045 (.005)	.045 (.005)			
Age 17		.148 (.012)	.152 (.011)	.154 (.010)			
Northeast			.045 (.026)	.066 (.016)		-.004 (.008)	-.000 (.008)
Midwest			.045 (.025)	.069 (.016)		.003 (.009)	.013 (.009)
South			.058 (.026)	.074 (.014)		.020 (.009)	.021 (.009)
West			.021 (.027)	.044 (.015)		.002 (.005)	.007 (.006)
Female				.004 (.003)			.003 (.003)
Poverty/10^3				-.276 (.035)			-.268 (.027)
Family Income/10^6				.054 (.031)			.047 (.028)
R^2	.029	.050	.051	.061			
No. Obs.	44,590	44,590	44,590	43,912	44,590	44,590	43,912

Note: Regressions are weighted using population weights. Note: Data are weighted to reflect the census population. Desegregated 1970-1979 group is 15-, 16-, and 17-year-old blacks who live in districts that desegregated between 1970 and 1979. Control group is all other 15-, 16, and 17-year-old blacks in districts that desegregated between 1961 and 1982. Columns 1 and 2 present estimates comparing the conditional mean difference in dropout rates between Treatment and Control districts in 1980, controlling for 1970 district-level dropout rates. Columns 3 and 4 present estimates of a matching model described in the paper. Standard error estimates are corrected for district*year correlation in the error term.

TABLE X: ALTERNATIVE SPECIFICATIONS TO CONTROL FOR LAGGED DROPOUT RATES

	Coefficient on Past Dropout Rate			
	$\gamma = .25$	$\gamma = .50$	$\gamma = .75$	$\gamma = 1.00$
Desegregated	-.013	-.018	-.024	-.029
1970-1979	(.006)	(.006)	(.007)	(.009)
Age 16	.035	.022	.009	-.004
	(.004)	(.005)	(.007)	(.009)
Age 17	.123	.083	.044	.005
	(.007)	(.007)	(.008)	(.010)
Northeast	.088	.016	.144	.172
	(.027)	(.041)	(.056)	(.071)
Midwest	.091	.119	.147	.175
	(.028)	(.043)	(.058)	(.073)
South	.094	.120	.146	.172
	(.025)	(.040)	(.054)	(.069)
West	.071	.105	.139	.174
	(.027)	(.042)	(.057)	(.072)
Female	.004	.004	.004	.004
	(.003)	(.004)	(.004)	(.004)
Poverty/10^3	-.273	-.269	.265	-.261
	(.035)	(.036)	(.037)	(.377)
Family	.534	.522	.511	.500
Income/10^6	(.305)	(.305)	(.307)	(.377)
R^2	.040	.025	.016	.012
No. Obs.	43,912	43,912	43,912	43,912

Note: The table shows estimates of the effect of desegregation on black high school dropout rates. Each column shows estimates based on a different specification. The specifications vary only in that they restrict the effect of the lagged dropout rate in the district to be different values between zero and unity. When γ is set at unity, the specification closely resembles the difference-in-difference specification. When γ is closer to zero, the specification more closely resembles the unrestricted lagged dropout rate specification. Standard errors, corrected for district-by-year correlation, are reported in parentheses.

TABLE XI: DIFFERENTIAL EFFECTS OF VOLUNTARY DESEGREGATION PLANS ON DROPOUT RATES OF BLACKS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Desegregated 1970-1979 *1980	-.041 (.011)	-.043 (.011)	-.040 (.011)	-.039 (.010)	-.032 (.010)	-.034 (.009)	-.035 (.009)
Desegregated 1970-1979 *1980 *Voluntary Plan	.018 (.008)	.028 (.011)	.029 (.011)	.024 (.010)	.016 (.010)	.020 (.009)	.021 (.009)
Desegregated 1970-1979	.027 (.008)	.021 (.008)	.020 (.009)	.023 (.009)	.024 (.008)	.023 (.007)	.024 (.007)
Voluntary Plan		-.011 (.007)	-.010 (.007)	.002 (.008)	.006 (.007)	.005 (.007)	.004 (.007)
Post	.002 (.008)	.001 (.007)	-.004 (.008)	-.004 (.007)	.001 (.007)	-.046 (.047)	-.034 (.047)
Age 16			.048 (.004)	.048 (.004)	.047 (.004)	.047 (.004)	.047 (.004)
Age 17			.166 (.006)	.166 (.006)	.161 (.006)	.161 (.006)	.161 (.006)
Female					.011 (.004)	.011 (.004)	.022 (.008)
Female *1980							-.018 (.009)
Poverty/10^3					-.198 (.031)	-.200 (.031)	-.027 (.048)
Poverty/10^3 *1980							-.243 (.058)
Family Income/10^6					-.025 (.029)	-.024 (.029)	-.372 (.125)
Family Income/10^6 *1980							.417 (.128)
Northeast				-.067 (.030)	-.049 (.030)	-.057 (.043)	-.057 (.043)
Northeast *1980						.032 (.048)	.033 (.048)
Midwest				-.061 (.030)	-.042 (.030)	-.058 (.043)	-.057 (.043)
Midwest *1980						.046 (.047)	.046 (.047)
South				-.047 (.030)	-.036 (.030)	-.056 (.043)	-.055 (.043)
South *1980						.053 (.047)	.052 (.047)
West				-.086 (.030)	-.069 (.030)	-.090 (.043)	-.090 (.044)
West *1980						.054 (.048)	.056 (.048)
R^2	.002	.003	.048	.050	.056	.056	.056
No. Obs.	53,331	53,331	53,331	53,331	52,416	52,416	52,416

Note: Regressions are weighted using population weights. Desegregated 1970-1979 group is 17-year-old blacks who live in districts that desegregated between 1970 and 1979, 16-year-old blacks who live in districts that desegregated between 1971 and 1980, and 15-year-old blacks who live in districts that desegregated between 1972 and 1981. Control group is all other 15-, 16, and 17-year-old blacks in districts that desegregated between 1961 and 1982. Standard error estimates are corrected for district*year correlation in the error term.

TABLE XII: DIFFERENCE-IN-DIFFERENCES ESTIMATES OF THE EFFECT OF DESEGREGATION ON DROPOUT RATES OF WHITES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Desegregated 1970-1979 *1980	.005 (.011)	.006 (.011)	.006 (.009)	.007 (.004)	.007 (.004)	.002 (.003)	.002 (.003)
Desegregated 1970-1979	.013 (.007)	.012 (.007)	.004 (.006)				
1980	.024 (.009)	.021 (.009)	.019 (.007)	.017 (.003)	.039 (.004)	.026 (.018)	.059 (.019)
Age 16		.049 (.003)	.049 (.003)	.049 (.003)	.050 (.003)	.050 (.003)	.050 (.003)
Age 17		.138 (.005)	.138 (.005)	.137 (.005)	.136 (.005)	.135 (.005)	.136 (.005)
Female					.010 (.002)	.010 (.002)	.014 (.003)
Female *1980							-.009 (.004)
Poverty/10^3					-.302 (.019)	-.302 (.018)	-.250 (.022)
Poverty/10^3 *1980							-.135 (.029)
Family Income/10^6					-.118 (.013)	-.118 (.013)	-.122 (.028)
Family Income/10^6 *1980							.049 (.029)
Northeast			-.045 (.022)				
Northeast *1980						-.005 (.018)	-.006 (.018)
Midwest			-.042 (.021)				
Midwest *1980						.014 (.018)	.016 (.018)
South			-.002 (.021)				
South *1980						.014 (.018)	.013 (.018)
West			-.026 (.022)				
West *1980						.035 (.018)	.035 (.018)
District Effects	No	No	No	Yes	Yes	Yes	Yes
R^2	.003	.040	.043	.049	.083	.083	.084
No. Obs.	203,063	203,063	203,063	203,063	200,379	200,379	200,379

Note: Data are weighted using population weights. Desegregated 1970-1979 group is 15-, 16-, and 17-year-old whites who live in districts that desegregated between 1970 and 1979. Control group is all other 15-, 16, and 17-year-old whites in districts that desegregated between 1961 and 1982. Standard error estimates are corrected for district*year correlation in the error term.

TABLE XIII: CHARACTERISTICS OF BLACK AND WHITE SCHOOLS IN 1970

	< 25 Percent Black Enrolment	25-75 Percent Black Enrollment	> 75 Percent Black Enrollment
Pupil-Teacher Ratio	27.9	23.9	26.7
	(12.9)	(10.6)	(6.5)
Fraction Black Teachers	.079	.241	.565
	(.113)	(.172)	(.243)
Number of Schools	3381	667	1079

Note: The unit of observation is a school. Data are from the U.S. Department of Health, Education and Welfare's Directory of Public Elementary and Secondary Schools in Selected Districts. Sample statistics only include districts that desegregated between 1970 and 1979. Standard deviations are reported in parentheses.

TABLE XIV: CHARACTERISTICS OF BLACK AND WHITE TEACHERS IN 1970

	Black Teachers	White Teachers
Years of Completed Education	15.6	15.9
	(2.1)	(1.8)
Potential Experience	15.5	16.4
	(11.7)	(14.5)
No. Obs.	1,404	13,558

Note: Data are from the 1970 Census. Standard deviations are reported in parentheses.

APPENDIX TABLE I: PROBIT ESTIMATES OF THE EFFECT OF DESEGREGATION ON DROPOUT RATES OF BLACKS

	(1)	(2)	(3)	(4)	(5)	(6)
Desegregated 1970-1979 *1980	-.028 (.010)	-.026 (.009)	-.026 (.008)	-.021 (.008)	-.025 (.008)	-.025 (.007)
Desegregated 1970-1979	.024 (.007)	.023 (.007)	.019 (.006)	.017 (.006)	.019 (.005)	.020 (.005)
1980	-.006 (.007)	-.011 (.007)	-.011 (.007)	-.005 (.006)	-.043 (.036)	-.026 (.035)
Age 16		.069 (.006)	.068 (.006)	.067 (.006)	.067 (.006)	.067 (.005)
Age 17		.185 (.007)	.185 (.007)	.181 (.007)	.181 (.007)	.180 (.007)
Female				.011 (.004)	.011 (.004)	.022 (.007)
Female *1980						-.019 (.007)
Poverty/10^3				-.176 (.033)	-.179 (.032)	-.033 (.042)
Poverty/10^3 *1980						-.241 (.054)
Family Income/10^6				-.055 (.032)	-.053 (.032)	-.031 (.011)
Family Income/10^6 *1980						.034 (.012)
Northeast			-.042 (.018)	-.030 (.019)	-.034 (.023)	-.035 (.023)
Northeast *1980					.032 (.040)	.034 (.040)
Midwest			-.044 (.018)	-.029 (.019)	-.039 (.024)	-.039 (.024)
Midwest *1980					.044 (.041)	.045 (.041)
South			-.032 (.020)	-.025 (.021)	-.039 (.026)	-.037 (.026)
South *1980					.048 (.039)	.047 (.039)
West			-.057 (.015)	-.047 (.016)	-.055 (.019)	-.056 (.019)
West *1980					.049 (.046)	.051 (.047)

Note: Marginal effects at the mean of explanatory variables are reported. Data are weighted using population weights. Desegregated 1970-1979 group is 15-, 16-, and 17-year-old blacks who live in districts that desegregated between 1970 and 1979. Control group is all other 15-, 16-, and 17-year-old blacks in districts that desegregated between 1961 and 1982. Standard error estimates are corrected for district*year correlation in the error term.

Appendix 1: Differences v. Lagged Dependent Variables Models

This section shows that if selection into treatment is based either on lagged outcomes or on fixed characteristics, the difference-in-differences estimator and an estimator that controls for a lagged dependent variable provide estimates that bracket the causal effect of interest.

Assume there is no secular trend in dropout rates so that we can write the dropout rate in school district *s* at time *t* as,

$$y_{st} = \alpha_s + \delta T_{st} + \varepsilon_{st}$$

where T_{st} is an indicator for being treated (living in 1980 in a district that desegregated in the 1970's).

Treatment is selected either on fixed characteristics such that,

$$T_{st} = \begin{cases} 1 & \text{if } \alpha_s > \bar{y} \\ 0 & \text{otherwise} \end{cases}$$

or on lagged values of *y* such that,

$$T_{st} = \begin{cases} 1 & \text{if } y_{st-1} > \bar{y} \\ 0 & \text{otherwise} \end{cases}$$

where $\bar{y}$ is a constant.

Define $\hat{\delta}_{LD}$ as the estimator that controls linearly for a lagged dependent variable. Define $\hat{\delta}_{DD}$ as the difference-in-differences estimator.

A1.1. The Lagged Dependent Variable Estimator, $\hat{\delta}_{LD}$:

$$\text{plim}\,\hat{\delta}_{LD} = \frac{\text{Cov}(y_{st}, \tilde{T}_{st})}{\text{Var}(\tilde{T}_{st})}$$

$$= \frac{\text{Cov}(\alpha_s + \delta\tilde{T}_{st} + \delta\hat{T}_{st} + \varepsilon_{st}, \tilde{T}_{st})}{\text{Var}(\tilde{T}_{st})}$$

where $\tilde{T}_{st}$ is the residual from a regression of T_{st} on y_{st-1}, and $\hat{T}_{st}$ is the predicted value from a regression of T_{st} on y_{st-1}. Therefore,

$$\text{plim}\,\hat{\delta}_{LD} = \delta + \frac{\text{Cov}(\alpha_s + \varepsilon_{st}, \tilde{T}_{st})}{\text{Var}(\tilde{T}_{st})}.$$

If treatment is selected on fixed characteristics,

$$\text{plim}\,\hat{\delta}_{LD} = \delta + \frac{\text{Cov}(\alpha_s + \varepsilon_{st}, T_{st} - k - \phi(\alpha_s + \varepsilon_{st-1}))}{\text{Var}(\tilde{T}_{st})}$$

where *k* is a constant and $\phi = \frac{\text{Cov}(T_{st}, y_{st-1})}{\text{Var}(y_{st-1})}$. Expanding the equation,

$$\text{plim}\,\hat{\delta}_{LD} = \delta + \frac{\text{Cov}(\alpha_s, T_{st})}{\text{Var}(\tilde{T}_{st})} - \frac{\text{Cov}(T_{st}, y_{st-1})}{\text{Var}(y_{st-1})}\left(\frac{\sigma_\alpha^2}{\text{Var}(\tilde{T}_{st})}\right)$$

assuming no serial correlation in ε. In the analysis in this paper, *t* and *t-1* are ten years apart. Thus, the assumption of no serial correlation in ε seems reasonable. Simplifying,

$$\text{plim}\,\hat{\delta}_{LD} = \delta + \frac{\text{Cov}(\alpha_s, T_{st})}{\text{Var}(\tilde{T}_{st})}\left(1 - \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\varepsilon^2}\right)$$

which implies that if treatment is *positively* selected on fixed characteristics, $\text{plim}\,\hat{\delta}_{LD} \geq \delta$. In other words, if treatment is *positively (negatively)* selected on fixed characteristics, the estimator that controls for lagged outcomes produces *positively (negatively)* biased estimates of the treatment effect.

A1.2. The Difference-in-Differences Estimator, $\hat{\delta}_{DD}$:

$$\text{plim}\,\hat{\delta}_{DD} = \frac{\text{Cov}(y_{st} - y_{st-1}, T_{st})}{\text{Var}(T_{st})}$$

$$= \frac{\text{Cov}(\delta T_{st} + \varepsilon_{st} - \varepsilon_{st-1}, T_{st})}{\text{Var}(T_{st})}$$

$$= \delta + \frac{\text{Cov}(\varepsilon_{st} - \varepsilon_{st-1}, T_{st})}{\text{Var}(T_{st})}$$

$$= \delta - \frac{\text{Cov}(\varepsilon_{st-1}, T_{st})}{\text{Var}(T_{st})}$$

assuming no serial correlation in ε.

If treatment is *positively* selected on lagged outcomes, $\text{plim}\,\hat{\delta}_{DD} \leq \delta$. In other words, if treatment is *positively (negatively)* selected on lagged outcomes, the difference-in-differences estimator produces *negatively (positively)* biased estimates of the treatment effect.

Therefore, if treatment is selected *positively* either on fixed characteristics or on a lagged dependent variable,

$$\text{plim}\,\hat{\delta}_{DD} \leq \delta \leq \text{plim}\,\hat{\delta}_{LD}.$$

If treatment is selected *negatively* either on fixed characteristics or on a lagged dependent variable,

$$\text{plim}\,\hat{\delta}_{LD} \leq \delta \leq \text{plim}\,\hat{\delta}_{DD}\,.$$

Appendix 2: A Simple Model of Legal Precedent

The following model is meant to illustrate the role of legal precedent in the decision of an agent seeking to promote social change through the courts. The agent's objective is to effect some change in a set of distinct local areas. To achieve this end, the agent brings legal challenges in each of these areas, one at a time. The model points out that the agent will pay close attention to precedent when choosing the timing of legal challenges. In the absence of precedent, it is clear that the agent will choose to bring suit first in the area with the largest potential benefit from success. He will then bring suit in the area with the second largest benefit, and so on until he is finished.

Precedent creates a spillover, where the national benefits and costs of bringing a suit are no longer the same as the local benefits and costs. The agent will internalize the spillovers created by the setting of precedent and, as a result, will weigh the probability of success more heavily than the benefit from success when choosing where to bring a legal challenge early on. When precedent is important in the legal system, and when the number of cases remaining to be brought is large, the agent can virtually ignore the potential benefits from success, and base his decision solely on the probability of success.

In the model, the agent chooses to bring a legal challenge in one location in each time period. The benefits of success vary by location, and the agent seeks to maximize the expected discounted sum of benefits. The legal challenge either succeeds or fails with some probability that depends both on the facts of the case specific to the locality and on the history of legal successes and failures up to the point of the current legal challenge. This second factor represents the role of precedent in the judicial system.

To illustrate the point more formally, consider an agent seeking to bring suit in two distinct locations, denoted *A* and *B*. The agent can bring forward only one legal challenge per time period. The agent wins the legal challenge with some probability, which depends on the characteristics of the location and the history of legal successes and failures. Specifically, let each location have an inherent probability of legal success, denoted P_i, $i \in (A, B)$. Since no precedent has been set when the agent brings the first legal challenge $P_i^{t=1} = P_i$.

At $t = 2$, precedent has been set by the outcome of the case filed in the first period. Thus, the probability of success in the second case brought depends on the outcome of the first case. For ease of exposition, assume that the agent must decide the order of cases at $t = 0$. Thus, the likelihood of success in the second period depends directly on the *ex-ante* likelihood of success of the case brought in the first period. Consequently, we can write the probability of success of the second case

$$P_i^{t=2} \equiv P_i^2 = P_i + \phi(P_j),\ j \neq i$$

where $\phi(.)$ is an increasing, non-negative function that represents the effect of precedent.

A2.1. When will the agent choose to file suit in location A first?

Each district has a potential benefit from legal success. For the purposes of this model, let each location be of a different size, denoted N_i. The larger the location, the more students will benefit from a legal victory. The agent seeks to maximize the expected number of student-years under court order. Thus, the agent will choose to file suit in location *A* first if

$$2P_A N_A + [P_B + \phi(P_A)]N_B \geq 2P_B N_B + [P_A + \phi(P_B)]N_A$$
$$\Leftrightarrow$$
$$P_A N_A + \phi(P_A)N_B \geq P_B N_B + \phi(P_B)N_A$$

By making this an equality, we can take the total derivative and find an indifference relationship between the local benefit (N_i), and the likelihood of success (P_i). The total derivative becomes

$$[P_A - \phi(P_B)]dN_A + [N_A + N_B\phi'(P_A)]dP_A + [P_B - \phi(P_A)]dN_B - [N_B + N_A\phi'(P_B)]dP_B = 0$$

Holding, N_B, P_B constant we get

$$-\frac{dN_A}{dP_A} = \frac{N_A + N_B\phi'(P_A)}{P_A - \phi(P_B)}$$

We can compare how the agent trades off N_A and P_A in a world without precedent. Here, the agent chooses location *A* first if

$$P_A N_A \geq P_B N_B$$

Thus, the total derivative yields

$$-\frac{dN_A}{dP_A} = \frac{N_A}{P_A}$$

Let us compare the agent's tradeoff in a world with precedent to the agent's tradeoff in a world without precedent. The difference in the marginal rate of substitution between local benefit and the probability of success is

$$\frac{N_A + N_B\phi'(P_A)}{P_A - \phi(P_B)} - \frac{N_A}{P_A} = \frac{P_A N_B\phi'(P_A) + N_A\phi(P_B)}{P_A{}^2 - P_A\phi(P_B)}$$

$$= \frac{N_B\phi'(P_A) + (N_A / P_A)\phi(P_B)}{P_A - \phi(P_B)}$$

There is some location for which $\phi(P_i) = 0$. Otherwise, the existence of precedent in the legal system would increase the probability of success in every case after the first period. In this case, let that be location *B*. The ratio can now be written

$$\left.\frac{dN_A}{dP_A}\right|_{no\ precedent} - \left.\frac{dN_A}{dP_A}\right|_{precedent} = \frac{N_B \phi'(P_A)}{P_A}$$

The magnitude of the above expression measures the amount that precedent adds to the agent's valuation of the probability of success, relative to the local benefit. The amount by which the marginal rate of substitution increases is a measure of how much precedent adds to the agent's valuation of P_i relative to N_i.

A measure of the importance of precedent in the legal system is the size of $\phi'(.)$. When $\phi'(.)$ is large, cases that are more likely to be successful increase the probability of success in subsequent trials by large amounts relative to cases that are unlikely to be successful. This effect is weighted by N_B; the larger the potential benefit to legal success in location *B*, the larger the incentive to increase the probability of success there, by winning the case in location *A* first.

A2.2. Additional Locations, Additional Time Periods

Adding more locations to the model requires the addition of structure to the model, but the result should be intuitively clear. The effect of precedent on subsequent legal challenges magnifies as the number of cases to be tried increases. When there are two locations, precedent creates a distinction between the global and local marginal benefits of choosing a higher P_i. This distinction becomes greater as the effect of precedent is allowed to compound over the course of many cases. When there are three locations, choosing a high P_i in the first period increases the probability of success in the second period. The increased likelihood of victory in the second period is a benefit in and of itself. Additionally, the increased likelihood of victory in the second period increases the *ex-ante* likelihood of victory in the third period.

Thus, in a legal system where precedent is important and in a situation where the agent wants to eventually bring suit in a large number of locations, the value of a high P_i will greatly outweigh the value of a high N_i. In other words, the agent seeking to pursue social change through the courts will virtually ignore the local benefit of success in the early stages of the process, and will choose to file suit in a location with a high probability of success.