EXHIBIT D

Expert Report of Jay P. Greene, Ph.D.

May 29, 2020

This report evaluates and responds to a number of the conclusions and opinions set forth in the Expert Report of Kirabo Jackson. Jackson's report is primarily a meta-analysis, or systematic review, of what he claims are the methodologically appropriate studies regarding the relationship between additional school funding and student outcomes. I examined the studies listed in Jackson's Expert Report, the studies listed in Jackson's 2018 systematic review (which he says is the basis for his report), as well as his characterizations of the individual and collective results of those studies.

Based on my reading of Jackson's report, the underlying research on the subject, as well as my experience, I have the following opinions, to a reasonable degree of professional certainty:

- **Key Claim 1 – The description of the scientific evidence on the relationship between additional spending and student outcomes in Jackson's Report is so riddled with errors, inconsistencies, and ambiguities that it is not a reliable summary of that evidence.**

- **Key Claim 2 – Jackson focuses on a set of studies that he claims demonstrate "credible causal relationships between school spending and outcomes" (p. 6) based on "natural experiments." (p. 44) While the research designs Jackson prioritizes can approximate the causal estimates derived from actual experiments if certain assumptions are strictly met, in the set of studies Jackson considers, those assumptions are routinely violated and their results should not be considered causal.**

- **Key Claim 3 – Jackson's method of counting study results with positive or negative results, regardless of statistical significance, and then calculating the odds of having that many positive results if there were truly no effect, is an inappropriate and misleading method of determining statistical confidence in the general findings of a research literature.**

- **Key Claim 4 – Jackson's list of studies is not a complete and unbiased summary of research on the relationship between additional spending and student outcomes because it is likely to suffer from "file drawer" bias, in which studies are missing because they are never reported or otherwise difficult to find.**

- **Key Claim 5 – In his Report, Jackson provides specific claims about the extent to which student outcomes would improve if school spending were increased. These claims are highly implausible given the experience with past changes in school spending and student outcomes.**

- **Key Claim 6 – Even if the studies in Jackson's list could truly be considered causal, and even if his characterization of that literature were complete, accurate, and unbiased, it is highly unlikely that the findings from this literature are applicable to the current circumstances in Delaware.**

- **Key Claim 7 – There are important inconsistencies between claims made by plaintiffs' experts, Kirabo Jackson, Jesse Rothstein, Hunter Gehlbach, and Clive Belfield.**

The documents and data upon which my opinions are based are cited throughout this report.

Qualifications

I am the 21st Century Distinguished Professor of Education Policy and Chair of the Department of Education Reform at the University of Arkansas. Since completing my doctorate in political science at Harvard University in 1995, I have been a professor at the University of Houston and University of Texas at Austin as well as a senior fellow at the Manhattan Institute prior to holding my current position at the University of Arkansas. During the last two decades I have published four books and more than two dozen peer-reviewed journal articles on a variety of education policy issues.

In particular, I have considerable knowledge about the issue of causal research in education policy, having conducted and studied nine different experimental interventions. These experiments have examined the effects of a variety of common educational practices and policies.

I have also conducted quasi-experimental research, like the kind listed in Jackson's Report, that attempts to approximate the causal nature of true experiments. That work has focused on the effects of test-based promotion policies as well as the effects of high-stakes testing on performance in low-stakes subjects. In addition, I conducted an influential meta-analysis of the effects of bilingual education that has been cited more than 800 times, according to Google Scholar.

Thus, I have extensive experience with issues raised in Jackson's Expert Report -- the challenges and opportunities in conducting causal research in education as well as synthesizing that research in a meta-analysis.

My curriculum vitae containing additional information about my qualifications and record of scholarship can be found at the bottom of this report.

Previous Expert Testimony and Rate of Compensation

In the past five years I testified as an expert witness in one case: Citizens for Strong Schools, et al vs. Florida State Board of Education, et al, Circuit Court of Leon County, Florida, Case No. 09-

CA-4534 (2016). I am being paid at a rate of $325 per hour in this case and have worked a total of 110 hours as of May 29, 2020.

Jay P. Greene

Key Claim 1 – The description of the scientific evidence on the relationship between additional spending and student outcomes in Jackson's Report is so riddled with errors, inconsistencies, and ambiguities that it is not a reliable summary of that evidence.

A. **Jackson's characterization of the evidence is inconsistent with his own prior description as well as the descriptions of the researchers on whose evidence he relies for his claim.**

In the Report, Jackson says, "On the whole, there is overwhelming evidence that school spending improves student outcomes." (p. 15) But Jackson, Jesse Rothstein (who is also an expert for the plaintiffs), and other researchers on whom Jackson relies for his conclusion, describe the research literature very differently in their scholarly work. In the studies included in Jackson's meta-analysis, these scholars describe the relationship between additional spending and student outcomes as "mixed," "contradictory," and "inconclusive." Listed below and in Appendix 1 are quotations from studies listed in Jackson's Report. Here is what these researchers, including Jackson and Rothstein, have written:

- Jackson, Johnson, Persico (2016)[1]: "Overall, the evidence on the effects of SFRs [school finance reforms] on academic outcomes is mixed, and the effects on long-run economic outcomes is unknown." (p. 160)

- Lafortune, Rothstein, and Schanzenbach (2018): "SFRs are arguably the most substantial national policy effort aimed at promoting equality of educational opportunity since the turn away from school desegregation in the 1980s. But there is little evidence about their effects on student achievement…. The literature regarding whether 'money matters' in education (Hanushek 1986, 2003, 2006; Card and Krueger 1992a; Burtless 1996) is contentious and does not offer clear guidance." (pp. 2-3)

- Cellini, Ferreira, and Rothstein (2010): "Despite the importance of capital spending, little is known about the overall impact of public infrastructure investment on economic output, and even less is known about the effects of school facilities investments…. Also closely related is the long literature on the effects of school spending more generally. Hanushek (1996) reviews more than ninety studies and concludes that '[s]imple resource policies hold little hope for

[1] Jackson sometimes refers to this study as being in 2015, see for example his Table 1 reproduced in my Exhibit 1. That is the date of a working paper version of his study, but the published version of this article is correctly listed as being from 2016 in his bibliography, which is the year I will use to reference this work. Similarly, Jackson sometimes describes Jackson, Wigger, & Xiang as being in 2020, but his bibliography lists it as being in 2018, which is the year I will also use.

improving student outcomes,' but Card and Krueger (1996) dispute Hanushek's interpretation of the literature…. Angrist and Lavy (2002) and Goolsbee and Guryan (2006) exploit credibly exogenous variation in school technology investments. Neither study finds shortrun effects on student achievement." (p. 216)

- Lafortune and Schonholzer (2018): "The empirical literature on capital expenditures offers little guidance with regard to these questions. Several studies find no or imprecise effects of capital expenditures on student achievement (see Cellini et al., 2010; Bowers and Urick, 2011; Goncalves, 2015; Martorell et al, 2016), while others find some evidence of positive impacts on student achievement, often only in reading and English-language arts (Welsh et al., 2012; Neilson and Zimmerman, 2014; Hong and Zimmer, 2016; Conlin and Thompson, 2017; Hashim et al., 2018). Other studies have looked at longer-run impacts of school construction programs in other countries that expand access to education (e.g. Duflo, 2001, 2004), measuring the effects of more general increases in human capital accumulation. Despite inconclusive evidence in the literature and general skepticism among economists, resource-based capital expenditure programs continue to be used by policymakers at the state and local level as tools to improve schools and reduce achievement gaps." (pp. 1-2)

- Lee and Polachek (2018): "Current analyses find contradictory evidence of the effect of school expenditures on dropout and graduation rates." (p. 131)

- Rauscher (2019): "Debates about the efficiency of education funding for student achievement have continued at least since the 1966 Coleman Report (e.g., Hanushek 1989, 1996; Burtless 1996; Greenwald et al. 1996; see Biddle and Berliner 2002 and Baker 2016 for reviews), including contemporary evidence of no relationship between funding and achievement (Morgan and Jung 2016)…. Existing research provides contradictory evidence about the effects of education funding on student achievement (e.g., Jackson, Johnson and Persico 2016; Morgan and Jung 2016)." (pp. 1, 27)

- Johnson (2015): "Despite its fiscal importance, evidence on the effectiveness of Title I is mixed (Matsudaira, Hosek, and Walsh 2012; Cascio, Gordon, and Reber 2013; Van der Klaauw 2008)." (p. 50)

- Papke (2008): "Yinger (2004) discusses education finance litigation and resulting reforms to state finance systems. He concludes that, while some of the evidence

indicates that state aid can boost student performance, none of the findings is definitive and some are quite ambiguous." (p. 466)

- Hyman (2017): "However, it is less clear whether the changes in spending affected student achievement, with some studies finding positive effects and others finding no effects."

- Conlin and Thompson (2017): "Recent literature has focused on using quasi-experimental designs to identify the causal effect of capital investment on student outcomes and housing prices. A set of quasi-experimental papers (Cellini et al., 2010; Hong & Zimmer, 2016; Kogan, Lavertu, & Peskowitz, 2017; Martorell, Stange, & McFarlin, 2016) estimate regression discontinuity designs using the majority rule cutoff in school bond referendum elections to compare outcomes (test scores and/or housing prices) for districts that just pass a bond referendum to fund additional capital expenditures to those that just fail to pass a bond referendum and generally find mixed evidence on the role of capital investments on student achievement." (p. 14)

Jackson's claim that "there is overwhelming evidence that school spending improves student outcomes" is at odds with his own prior description of the research literature, the description of another expert witness for the plaintiffs, and the bulk of the researchers in the studies on which Jackson relies for his claim. The fact that these contradictory descriptions are typically found in peer-reviewed articles further supports the view that Jackson's bold claim about "overwhelming evidence" is inconsistent with what the research community believes about the pattern of findings on this issue.

B. There are numerous inconsistencies between how Jackson characterizes the results of studies listed in his Expert Report and in his 2018 review of the literature on which his Report is based.

In the Report, Jackson compiles a list of what he claims are methodologically appropriate studies and classifies each as to whether the results are positive or not and whether those results are statistically significant or not. (See Exhibit 1) He says that he built this list "from Jackson (2018),"(p. 12) which contains a similar list of what he deems to be methodologically appropriate research. (See Exhibit 2) The credibility of how Jackson classifies studies as positive or not and statistically significant or not is seriously undermined by the fact that he classifies many studies differently in these two lists he has compiled. It is unclear why Jackson would describe the same studies fundamentally differently in these two reviews.

Of the 28 studies Jackson has in both the Report list and his 2018 list, he classifies the statistical significance of the results differently nine times. That is, roughly a third of the time, Jackson describes the results of studies differently in 2018 than he does now. In eight cases, Jackson describes studies as having positive and statistically significant results in 2018 that he describes as not having statistically significant results in the Expert Report: Brunner, Hyman, & Ju (2019); Candelaria & Shores (2019); Cascio, Gordon, & Reber (2013); Lafortune, Rothstein, & Schanzenbach (2018); Conlin and Thompson (2017); Holden (2016); Hong & Zimmer (2016); and Neilson & Zimmerman (2014).

In one case, the inconsistency goes in the other direction. In his 2018 review, Jackson described the results of Cellini, Ferreira, & Rothstein (2010), saying "Similar null impacts are found for capital spending (using a very similar design) in California (Cellini et al, 2010), and Ohio (Goncalves, 2015)." (p. 11) But in his Report, Jackson lists the results for Cellini, Ferreira, & Rothstein (2010) as positive and statistically significant.

Even within the Expert Report, Jackson does not code the statistical significance of studies consistently. In Table 1 in Jackson's Report (see Exhibit 1) he describes Guryan (2001) as having positive and statistically significant results, but in Figure 2 on p. 21 and Figure 6 on p. 29 of that same report, the bar representing the 95% confidence interval for that study encompasses a large area of possible negative as well as positive outcomes, indicating that the result is not statistically significant. (See Exhibit 3)

If Jackson cannot consistently classify the direction and statistical significance of the same studies in two different reviews he conducted recently, it is unlikely that we can rely on his claims about this research literature as a whole.

Exhibit 1: Jackson's Classification of Studies in Expert Report (2020)

Table 1: *List of Studies Meeting the Inclusion Criteria*

Study	pos	pos. & sig.	Outcome	Strategy	$ type	Low-I
Multi-State Studies						
Abott Kogan Lavertu Peskowitz, 2019	Yes	Yes	Test Scores, Grad. Rates	Regression Discontinuity	Operational	No
Biasi, 2019	Yes	Yes	Enroll Postsecondary, Income Mobility	Instrumental Variables	Any	Yes
Brunner Hyman Ju, 2019	Yes	No	Test Scores	Instrumental Variables	Any	n/a
Candelaria Shores, 2019	Yes	No	Grad. Rates	Event-Study DiD	Any	Yes
Card Payne, 2002	Yes	Yes	Test Scores	CO-SFR	Any	Yes
Cascio Gordon Reber, 2013	Yes	No	Dropout Rates	Instrumental Variables	Title I	n/a
Jackson Johnson Persico 2015, Johnson Jackson, 2019	Yes	Yes	Grad. Rates, Years of Ed., Wages	Event-Study DiD, Instrumental Variables	Any	Yes
Jackson Wigger Xiong, 2020	Yes	Yes	Test Scores, Enroll Postsecondary	Instrumental Variables	Any	Yes
Johnson, 2015	Yes	Yes	Grad. Rates, Other Ed. Outcomes, Wages	Event-Study DiD	Title I	Yes
Lafortune Rothstein Schanzenbach, 2018	Yes	No	Test Scores	Event-Study DiD	Any	Yes
Miller, 2018	Yes	Yes	Test Scores, Grad. Rates	Instrumental Variables	Any	n/a
Non-Multi-State Studies						
Baron, 2019	Yes	Yes	Test Scores	Regression Discontinuity	Operational	n/a
Carlson Lavertu, 2018	Yes	No	Test Scores	Regression Discontinuity	Any	n/a
Cellini Ferreira Rothstein, 2010	Yes	Yes	Test Scores	Regression Discontinuity	Capital	n/a
Clark, 2003	Yes	No	Test Scores	Event-Study DiD	Any	n/a
Conlin Thompson, 2017	Yes	No	Test Scores	Instrumental Variables	Capital	n/a
Gigliotti Sorensen, 2018	Yes	Yes	Test Scores	Instrumental Variables	Any	n/a
Goncalves, 2015	Yes	No	Test Scores	Event-Study DiD	Capital	Yes
Guryan, 2001	Yes	Yes	Test Scores	Instrumental Variables	Any	n/a
Holden, 2016	Yes	No	Test Scores	Regression Discontinuity	Operational	n/a
Hong Zimmer, 2016	Yes	No	Test Scores	Regression Discontinuity	Capital	n/a
Hyman, 2017	Yes	Yes	Enroll Postsecondary	Instrumental Variables	Any	No
Kogan Lavertu Peskowitz, 2017	Yes	Yes	Test Scores	Regression Discontinuity	Any	n/a
Kreisman Steinberg, 2019	Yes	Yes	Test Scores, Grad. Rates, Enroll Postsecondary	Instrumental Variables	Any	Yes
Lafortune Schonholzer, 2018	Yes	Yes	Test Scores	Event-Study DiD	capital	n/a
Lee Polachek, 2018	Yes	Yes	Dropout Rates	Regression Discontinuity	Any	n/a
Martorell Stange McFarlin, 2016	Yes	No	Test Scores	Regression Discontinuity	Capital	n/a
Matsudaira Hosek Walsh, 2012	Yes	No	Test Scores	Regression Discontinuity	Title I	No
Neilson Zimmerman, 2014	Yes	No	Test Scores	Event-Study DiD	Capital	n/a
Papke, 2008	Yes	Yes	Test Scores	Instrumental Variables	Any	n/a
Rauscher, 2019	Yes	No	Test Scores	Regression Discontinuity	Capital	Yes
Roy, 2011	Yes	Yes	Test Scores	Instrumental Variables	Any	Yes
Weinstein Stiefel Schwartz Chalico, 2009	Yes	No	Grad. Rates	Regression Discontinuity	Title I	n/a

Exhibit 2: Jackson's Classification of Studies in 2018 Review

Study	Pos and Sig	Neg and Sig	Not Sig	Outcomes	Variation	State	Type of Spending
Multi-State Studies							
Jackson Johnson and Persico (2015)	Y			Education, Wages	CO-SFR	ALL	Any
Johnson and Jackson (2018)	Y			Education, Wages, other	CO-SFR	ALL	Any
Lafortune, Rothstien, Shanzenbach (2018)	Y			Test Scores	SFR	ALL	Any
Candelaria and Shores (2018)	Y			Graduation Rates	CO-SFR	ALL	Any
Brunner, Hyman, and Ju (2018)	Y			Test Scores	SFR	ALL	Any
Biasi (2018)	Y			Income mobility	SFR	ALL	Any
Card and Payne (2002)	Y			SAT score inequality	CO-SFR	ALL	Any
Hoxby (2001)			Y	Dropout rates	SFR	ALL	Any
Downes and Figlio (1997)	Y			Test Scores	DiD-Tax Limit	ALL	Any
Johnson (2015)	Y			Graduation Rates	DiD	ALL	Title I
Jackson, Wigger and Xiong (2018)	Y			Test Scores	IV-Recession	ALL	Any
Miller (2018)	Y			Test Scores	IV-House Values	ALL	Any
Cascio, Gordon, and Reber (2013)	Y			Dropout	Event Study	SOUTH	Title I
Single-State Studies							
Hyman (2017)	Y			College-going	Rules-Based IV	MI	Any
Gigliotti (2018)	Y			Test Scores	Rules-Based IV	NY	Any
Papke (2008)	Y			Test Scores (pass rates)	IV-SFR	MI	Any
Roy (2011)	Y			Test Scores	Policy (SFR)	MI	Any
Guryan (2001)	Y			Test Scores	Rules-Based IV	MA	Any
Clark (2003)			Y	Test Scores	Rules-Based IV	KY	Any
Lee and Polachek (2018)	Y			Graduation Rates	RD-Referenda	NY	Any
Holden (2016)	Y			Test scores	RD	CA	Textbooks
Husted and Kenny (2000)	Y			SAT Scores	DiD	CA	Any
Cellini, Ferreira, and Rothstein (2010)			Y	Test Scores	RD-Bonds	CA	Capital
Lafortune and Schonholzer (2018)	Y			Test Scores	Event-Study	CA	Capital
Martorell, Stang, McFarlin (2016)			Y	Test Scores	RD-Bonds	TX	Capital
Conlin and Thompson (2017)	Y			Test Scores	IV	OH	Construction
Goncalves (2015)			Y	Test Scores	Event-Study	OH	Construction
Hong and Zimmer (2016)	Y			Test Scores	RD-Bonds	MI	Capital
Kogan, Lavertu, and Peskowitz (2017)	Y			Test Scores	RD-Referenda	OH	Any
Zimmerman and Neilson (2014)	Y			Test Scores	Event-study	New Haven, CT	Construction
Van Der Klaue (2008)			Y	Test Scores	RD Title I	NY City	Title-I
Matsudaira, Hosek, Walsh (2012)			Y	Test Scores	RD-Title I	NYC*	Title-I
Weinstein, M. G., Stiefel, et al (2009)			Y	Test Scores	RD-Title I	NYC*	Title-I

Exhibit 3: Figures in Jackson's Expert Report of Results by Study with 95% Confidence Intervals

Figure 2: Estimated School Spending Effect on Test Scores (Operational Spending)

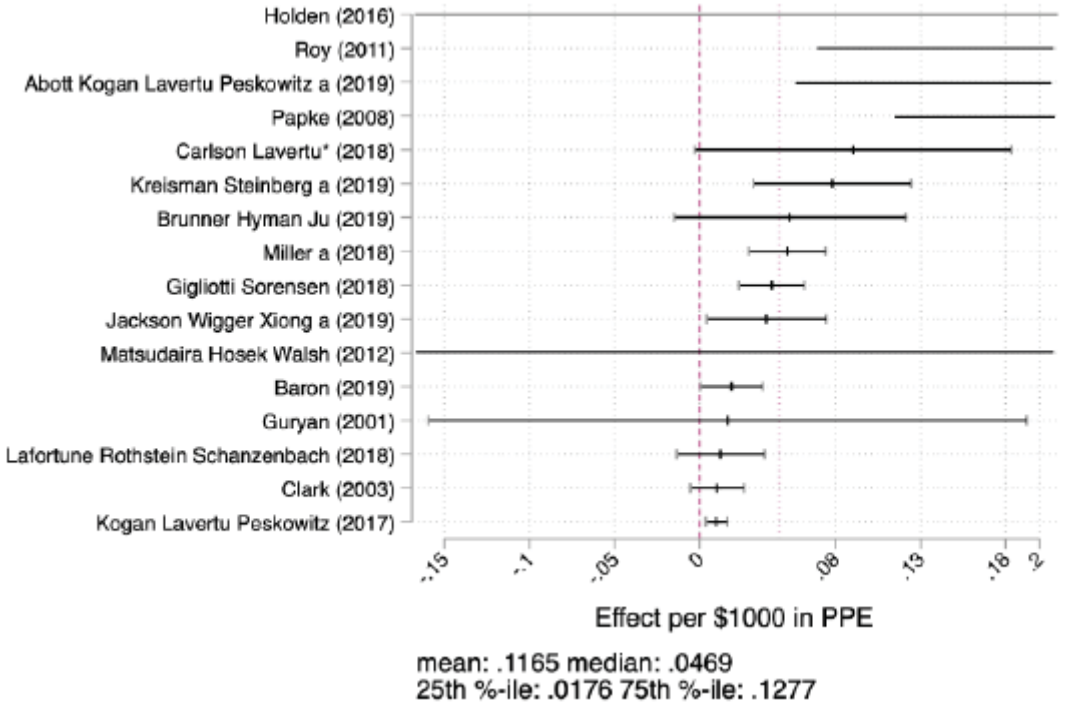


Figure 6: *Estimated School Spending Effects on Test Scores (capital and operational spending)*

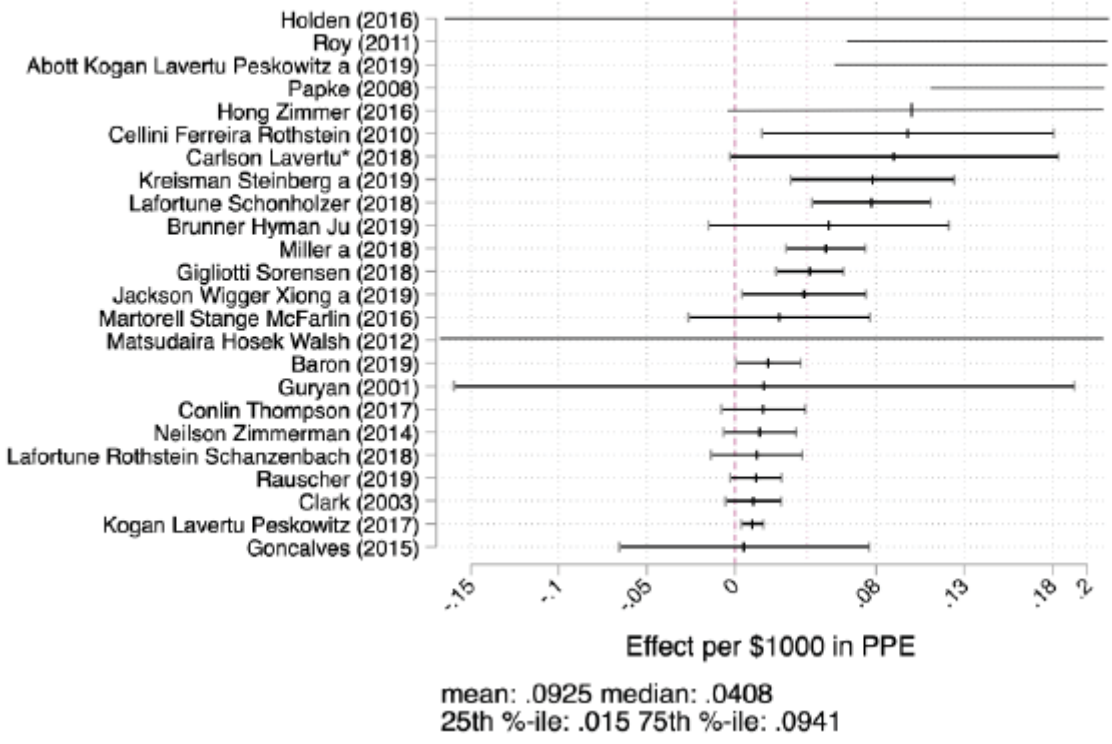


Figure 7: *Estimated School Spending Effects on Educational Attainment*

Overall estimates, non-test score outcomes:

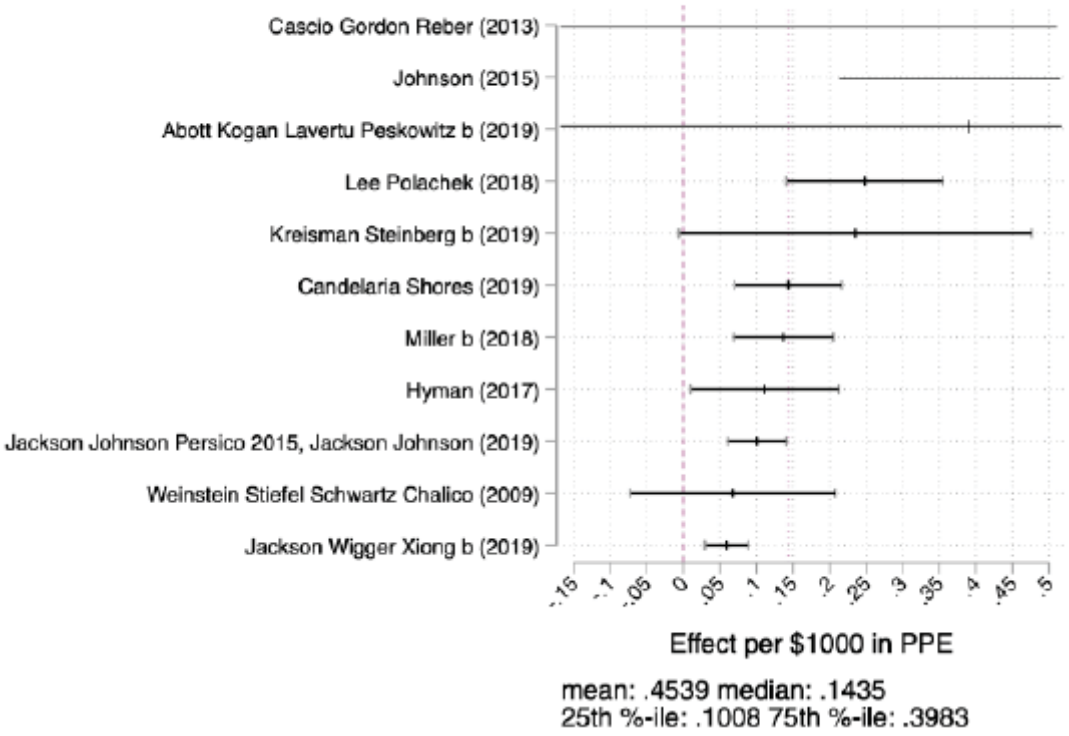
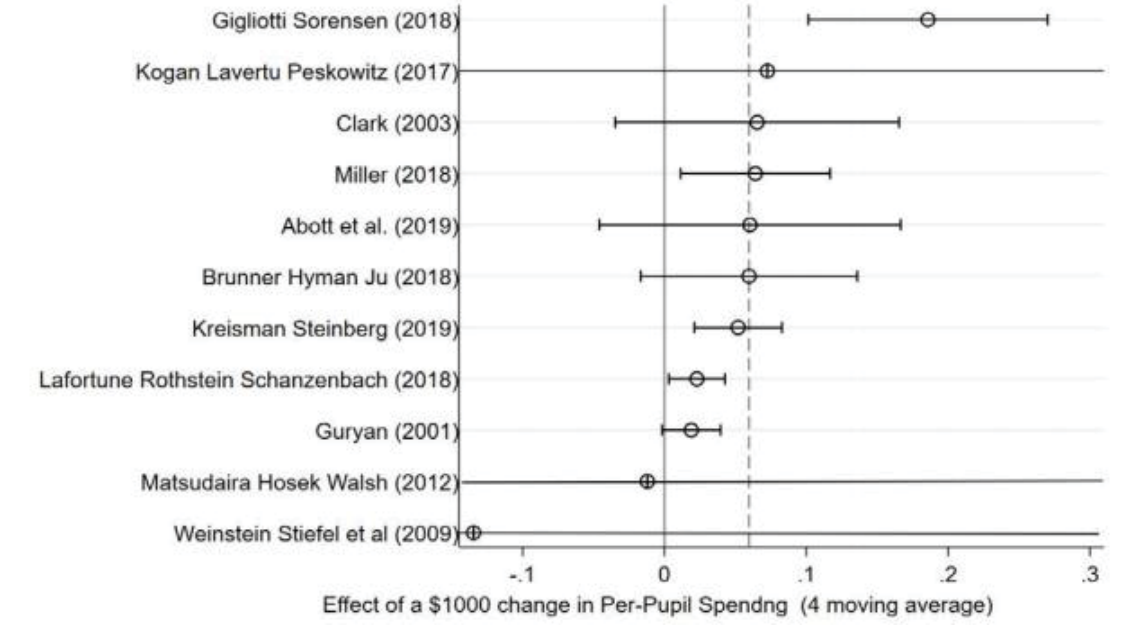


Exhibit 4: Figure in Jackson, Wigger, and Xiong (2018) of Results by Study with 95% Confidence Intervals

Figure 5. *Forrest Plot of Existing Studies*



C. There are numerous inconsistencies between how Jackson characterizes the results of studies listed in his Expert Report and how those results are reported in Jackson, Wigger, and Xiong (2018).

In a working paper that Jackson co-authored with Wigger and Xiong (2018), they provide a figure that summarizes the findings of other studies. (See Exhibit 4) Jackson, Wigger, & Xiong (2018) as well as the 11 studies whose results are represented in that "Forrest [sic] Plot" are all among the 33 studies in Jackson's Expert Report. But the results of those 11 studies are generally described very differently in the Jackson, Wigger, and Xiong (2018) working paper than how they are characterized in his Report.

Importantly, two of the studies, Weinstein, et al (2009) and Matsudaira, et al (2012) are described in the working paper as having negative but statistically insignificant results. In his Report, Jackson claims that these and all other methodologically appropriate studies show positive results.

In addition, the magnitude and statistical significance of results from the studies listed in Jackson, Wigger, & Xiong (2018) (Exhibit 4) are very different from what is listed in the figures in his Report (Exhibit 3). In Exhibit 4, three studies are shown as having statistically insignificant results because the bar representing the 95% confidence interval is shown crossing 0, but the results from those studies are described as being statistically significant in Jackson's Report.

Of the 11 studies whose results are characterized in Jackson, Wigger, and Xiong (2018), six have the direction or statistical significance of their findings reported differently in Jackson's Report. That is, a majority of these studies are described differently by Jackson in the Report than in his recent working paper.

D. In identifying the set of studies he deems to be methodologically appropriate, Jackson is not consistent in what studies meet his criteria between his Expert Report and 2018 review.

There are four studies in Jackson's 2018 review of the research that he does not include in the Expert Report list: Downes & Figio (1997); Hoxby (2001); Husted & Kenny (2000); and Van der Klaauw (2008).[2] (See Exhibits 1 and 2) Jackson deemed all four of these studies to be methodologically appropriate in his 2018 review, but he explicitly chooses to exclude them in the Report.

[2] Jackson incorrectly lists Van Der Klaauw (2008) as "Van Der Klaue (2008)" in his 2018 review. See Exhibit 2.

It is important to note that I would classify one of these studies as having negative and statistically significant results and three as having null findings. (See discussion below in my re-analysis in Key Claim 1, G).) Only by excluding this negative result is Jackson able to claim that every study finds positive effects of additional spending.

Jackson provides reasons in footnotes on p. 12 of his Report for why he excludes these four studies, but it is unclear why the arguments for ignoring these four studies were not persuasive to him two years earlier when he compiled his original list of methodologically appropriate studies. In addition, none of the explanations he offers is compelling, nor is their application to exclude these four studies consistent with the inclusion of other studies that did remain in the Expert Report list.

As another example of general imprecision, Jackson describes one of these negative studies, Husted & Kenny (2000), in his 2018 review as being a single-state study that examines California. (See Exhibit 2) This is factually incorrect. As Husted & Kenny describe their own study, "Our core data are taken from state reports on students taking the SAT compiled by the Educational Testing Service (ETS). The data begin in 1987, the first year in which data on parents' education were reported, and end in 1992, the last year for which we have data on school resource inequality. The ETS data cover 37 states for 6 years." (p. 288)

In addition, two of the studies Jackson lists in his Report, Biasi (2019) and Card & Payne (2002), do not appear in any of his Figures showing results by study. (See Exhibit 3) It is unclear why these two studies are missing or whether their results are included in his calculations of the mean and median effects of the combined studies.

E. It is not possible to replicate Jackson's classification of the direction and statistical significance of the studies he includes in his reviews, nor is there a plausible and consistent decision-rule that could account for how he classifies studies.

Almost every study in Jackson's Report list provides many estimates of effects, including multiple results across time, across types of outcomes, across grade-levels, and across different model specifications. For each study, Jackson must somehow combine these multiple estimates into a single estimate of how much student outcomes improve per $1,000 in additional spending. But Jackson never describes the exact method by which he combines these multiple results within each study other than to say in a footnote, "Most studies report overall effects. In a small number of cases, studies will report effects for different populations (e.g., 4th grade and 8th grade). In such cases, we take the average of the two effects as the overall effect." (p. 12)

But the reality of the multiplicity of reported results in each study is much more complicated than Jackson suggests. Simply taking an average or even a "precision-weighted average" (p. 17) of all reported results within each study would not replicate the way in which Jackson classifies each study. In fact, there is no consistent procedure that Jackson could have used to yield the pattern of how Jackson classifies the studies in his list. Let me illustrate the inconsistency with which he must code studies with some examples.

Jackson classifies Kogan, Lavertu, & Peskowitz (2017) as having positive and statistically significant results. The main results for Kogan, et al's examination of the effects of passage of tax referenda on student test scores are contained in Table 6 (p. 394) in that article. (See Exhibit 5) Note that Kogan, et al estimate the effect of the *defeat* of a school tax election on student achievement, so Jackson would treat a negative finding as an indication of a positive relationship between additional spending and student outcomes.

Exhibit 5: Results from Kogan, et al (2017)

Table 6. Impact of Tax Levy Failure on Student Achievement

	State "Value Added" Estimate (District SDs)				State Performance Index (District SDs)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2 years prior	−0.102	−0.00960	−0.144	−0.00754	−0.00593	−0.0127	−0.0118	−0.00703
	(0.129)	(0.0909)	(0.167)	(0.0577)	(0.0217)	(0.0174)	(0.0230)	(0.0121)
1 year prior	—	—	—	—	—	—	—	—
Election year	−0.0926	−0.0242	−0.155	−0.0429	0.0295	−0.00163	0.0271	−0.0141
	(0.113)	(0.0831)	(0.140)	(0.0552)	(0.0197)	(0.0150)	(0.0212)	(0.00998)
1 year after	−0.0894	−0.0342	−0.0666	−0.0486	−0.0187	−0.0351*	−0.0299	−0.0356**
	(0.117)	(0.0873)	(0.147)	(0.0548)	(0.0220)	(0.0167)	(0.0240)	(0.0123)
2 years after	−0.199^	−0.124	−0.178	−0.0913^	−0.0198	−0.0419*	−0.0251	−0.0312*
	(0.103)	(0.0782)	(0.146)	(0.0524)	(0.0245)	(0.0190)	(0.0265)	(0.0138)
3 years after	−0.179^	−0.126	−0.168	−0.00385	−0.0274	−0.0630**	−0.0317	−0.0281^
	(0.108)	(0.0826)	(0.148)	(0.0551)	(0.0273)	(0.0204)	(0.0290)	(0.0145)
4 years after	−0.0195	0.0722	0.0678	0.0115	0.0123	−0.0281	−0.0160	−0.0181
	(0.108)	(0.0868)	(0.154)	(0.0585)	(0.0296)	(0.0220)	(0.0309)	(0.0156)
5 years after	−0.120	−0.0701	−0.0344	−0.00224	−0.00610	−0.0394	−0.0364	−0.0148
	(0.117)	(0.0955)	(0.156)	(0.0614)	(0.0326)	(0.0246)	(0.0341)	(0.0173)
6 years after	−0.150	0.00258	−0.0456	−0.0312	−0.00776	−0.0213	−0.0339	−0.0126
	(0.123)	(0.0975)	(0.156)	(0.0667)	(0.0354)	(0.0270)	(0.0363)	(0.0194)
N	24,796	24,796	10,936	24,796	33,199	33,199	21,660	33,199
District count	571	571	509	571	571	571	541	571
Levy count	4,324	4,324	1,916	4,324	4,324	4,324	2,812	4,324
Mean dependent variable	0.03	0.03	0.03	0.03	0.09	0.09	0.06	0.09
Model	RD	RD	RD	Differences-in-Differences	RD	RD	RD	Differences-in-Differences
Specification	Quad.	Linear	Linear	N/A	Quad.	Linear	Linear	N/A
Restricted bandwidth	No	No	Yes	N/A	No	No	Yes	N/A
Levy type	Op. & Cap.	Op. & Cap.	Op. & Cap.	Op. & Cap.	Op. & Cap.	Op. & Cap.	Op. & Cap.	Op. & Cap.

Note: The results above are from models estimating the impact of levy failure (as opposed to passage) on district performance measures standardized by year. SEs clustered by district are presented in parentheses below the estimated coefficients. *p*-values were calculated using a two-tailed test.

^$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

Table 6 (Exhibit 5) presents four different models for two different types of test score results with outcomes for each of six years following the election for a total of 48 estimated effects. The first set of four models provides results for value added test scores and the second set of four models provides results for state performance index test scores. The authors describe the difference between the two types of test score measures: "Unlike district performance measures based on achievement levels (e.g., the performance index), which are confounded by student socioeconomic status and other differences in academic achievement unrelated to school and district quality, the value-added scores account for up to 5 years of students' previous test scores and, thus, account for student-level factors that may affect their performance." (pp. 387-8)

If Jackson were to focus on the more rigorous value added set of results, with four different model specifications across six years following the election, he would find no

statistically significant outcomes at the conventional level of $p < .05$. (See Exhibit 5) Three of those 24 results are significant at $p < .1$, but that is about what you might expect to see by chance. Perhaps Jackson's method for combining results within a study is to consider all results, including those that the authors describe as less rigorous. If Jackson considered the set of state performance index results, he would find that 5 of those 24 results are statistically significant. But those significant effects are only found in two of the four different model specifications and are only observed one to three years after the election. That is, there are no statistically significant benefits to passing a tax referenda four or more years after the election. And those benefits cannot be observed if one focuses on value added in test scores nor can they be observed if the model uses a quadratic or restricts the bandwidth of the regression discontinuity.

What decision-rule would have allowed Jackson to classify this study as showing positive and statistically significant results? He could not have simply taken an average, as he suggests in the Report, because only 5 of the 48 reported results are statistically significant. He could not have chosen to focus on the most rigorous measure of the outcome since the value added models had no statistically significant results. He could not have chosen to focus on more rigorous model specifications, since the limited bandwidth models also showed no statistically significant results. He could not have chosen to focus on results over the long-term since there were not statistically significant effects in years 4 through 6. Other than making a holistic and subjective judgment, the only plausible decision-rules that Jackson could have used to classify this study as having positive and statistically significant results would be 1) if there is any statistically significant result within a study, then that is how he would classify the study as a whole, or 2) the direction and significance of initial results, even if the direction or significance change over time, determine how he classifies the study as a whole, or 3) how the authors of the study describe their results determines how he classifies the study as a whole.

None of these decision-rules for how Jackson classifies the direction and significance of studies is consistent with how he classifies other studies. For example, if ever producing a statistically significant result determines how he would describe a study as a whole, then he would have classified Weinstein, Stiefel, Schwartz, & Chalico (2009) as showing negative and statistically significant results. As those authors report in Table 7 of their study, receiving Title I funds has a negative and statistically significant effect on elementary and middle school test scores in three of the four model specifications they present. (See Exhibit 6) But Jackson classifies this study as having positive but statistically insignificant results.

Exhibit 6: Results from Weinstein, et al (2009)

Table 7: Academic Outcomes, RD Estimates, Cubic Polynomial, Elementary and Middle Schools 1997-2003

	Math		Reading	
	(1)	(2)	(3)	(4)
Title I	-0.035*	-0.011	-0.037*	-0.031*
	(0.020)	(0.017)	(0.019)	(0.016)
%Poverty	-0.005**	-0.002	-0.002	-0.000
	(0.002)	(0.003)	(1.979)	(0.002)
%PovertySq	-0.000	-0.000	0.000*	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
%Poverty Cub	-1.37	-4.85	8.86***	-4.72
	(9.74)	(1.27)	(9.31)	(1.23)
TitleI*%Poverty	-0.004	0.005	-0.012***	0.000
	(0.004)	(0.004)	(0.004)	(0.003)
TitleI*%PovertySq	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
TitleI*%PovertyCub	1.56	5.76*	-3.65	3.26
	(3.31)	(3.07)	(3.12)	(2.89)
Total Registration	0.000***	-4.29	0.000	-7.35
	(7.48)	(0.000)	(7.27)	(0.000)
%ELL	-0.008***	-0.007***	-0.007***	-0.004***
	(0.000)	(0.001)	(0.000)	(0.001)
%Recent Immigrant	-0.005***	-0.000	0.007***	-0.002
	(0.000)	(0.001)	(0.001)	(0.001)
%Special Ed	-0.001***	0.000	-0.001**	0.001
	(0.001)	(0.001)	(0.000)	(0.001)
% black	-0.008***	-0.010***	-0.006***	-0.006***
	(0.000)	(0.001)	(0.000)	(0.001)
% Hispanic	-0.006***	-0.010***	-0.004***	-0.008***
	(0.000)	(0.001)	(0.000)	(0.001)
% Asian	0.004***	-0.001	0.001***	0.002***
	(0.000)	(0.001)	(0.000)	(0.001)
% female	0.018***	0.003***	0.019***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Constant	-0.237***	0.639***	-0.434***	0.270**
	(0.094)	(0.109)	(0.089)	(0.115)
Borough Effects	Yes	No	Yes	No
School Effects	No	Yes	No	Yes
N Schools	776	776	776	776
Observations	5432	5432	5432	5432
R-squared	0.66	0.83	0.63	0.81

(i) Robust standard errors in parentheses
(ii) * significant at 5%; ** significant at 1%
(iii)Year, middle school, and grade dummies are included but not shown.

If the initial direction and significance of results determined how Jackson would describe a study as a whole, then he would have classified Goncalves (2015) as negative and statistically significant. As that study shows in Table 4, districts that barely pass a school bond for construction experience a statistically significant decline in test scores for the first five years during construction. (See Exhibit 7) As Gonclaves describes these results, "Figure 5 and Table 4 show the effect of the program on test scores for all districts,

pooled across all grades. I find strong evidence of a negative effect during the construction period. Construction leads to drops of 2.2% of students proficient in math and 1.6% proficient in reading after 4 years of exposure, off a base of 80% proficiency rates. I do not find any significant evidence for positive effects from project completion." (pp. 13-14) Yet Jackson classifies this study as having positive but statistically insignificant results.

Exhibit 7: Results for Gonclaves (2015)

	(1) Math	(2) Reading	(3) Math	(4) Reading	(5) Math	(6) Reading
1 yr. Construction Exposure	-0.527	-0.651*	-0.525	-0.878	-0.698	-1.426**
	(0.376)	(0.322)	(0.541)	(0.489)	(0.620)	(0.518)
2 yr. Construction Exposure	-1.977***	-1.308**	-2.620***	-2.115*	-3.460***	-2.767***
	(0.442)	(0.412)	(0.737)	(0.835)	(0.727)	(0.697)
3 yr. Construction Exposure	-2.501***	-1.680**	-3.723***	-3.364**	-3.438***	-3.346***
	(0.596)	(0.566)	(0.979)	(1.150)	(0.862)	(0.872)
4 yr. Construction Exposure	-2.553**	-1.720*	-4.273**	-3.143*	-4.610**	-4.089***
	(0.945)	(0.765)	(1.437)	(1.454)	(1.383)	(1.070)
5 yr. Construction Exposure	-2.399*	-0.994	-4.001	-2.043	-4.280**	-3.630**
	(1.010)	(0.854)	(2.104)	(1.832)	(1.352)	(1.182)
6+ yr. Construction Exposure	-2.269	-1.076	-7.408	-4.123	-4.260	-4.682*
	(2.465)	(1.975)	(5.011)	(5.054)	(2.586)	(2.283)
1 yr. Completion Exposure	0.00874	0.129	0.689	0.120	0.635	0.522
	(0.587)	(0.254)	(0.589)	(0.484)	(0.687)	(0.324)
2 yr. Completion Exposure	-0.377	0.0665	-0.0853	-0.273	0.834	0.511
	(0.630)	(0.441)	(1.071)	(0.804)	(0.884)	(0.676)
3 yr. Completion Exposure	-0.135	-0.310	0.508	-0.772	1.386	0.0463
	(0.811)	(0.704)	(1.289)	(1.118)	(1.261)	(1.044)
4 yr. Completion Exposure	0.188	-0.186	0.652	-0.827	2.024	0.225
	(1.013)	(0.856)	(1.532)	(1.300)	(1.537)	(1.245)
5 yr. Completion Exposure	0.386	-0.123	0.948	-0.568	2.688	-0.169
	(1.144)	(1.010)	(1.739)	(1.584)	(1.754)	(1.511)
6+ yr. Completion Exposure	1.266	-1.442	2.216	-1.931	4.573*	-1.107
	(1.347)	(1.072)	(2.004)	(1.669)	(2.066)	(1.558)
N	14569	14570	7090	7091	5720	5720
R2	0.805	0.793	0.799	0.782	0.813	0.802
Avg. Proficiency	80.122	80.122	80.006	80.006	74.913	74.913
Districts	All	All	Poorest 25%	Poorest 25%	Lunch > 25%	Lunch > 25%

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

If how the authors describe their own results is the criterion Jackson uses to describe a study as a whole, then he would classify Martorell, Stange, & McFarlin (2016) as having a null result. As the authors describe their results in the abstract, "Event-study analysis

focused on the students actually affected by large campus renovations also generates very precise zero estimates of achievement effects." And in the text they conclude, "The results provide no indication of meaningful effects on test scores, as all of the post-intervention estimates are close to zero and precisely estimated." (p. 27) But Jackson classifies these null results as positive but not statistically significant.

Given the information Jackson provides in his Report, I am unable to successfully replicate the way in which he classifies the direction and statistical significance of the studies in his list, nor can I construct any consistent decision-rule that would allow those results to be replicated.

F. A large number of studies in Jackson's "list of all studies on the effect of school spending on student outcomes" do not meet his own criteria for inclusion in his meta-analysis.

In the Expert Report, Jackson identifies "a list of all studies on the effect of school spending on student outcomes" (p. 11) that meet three criteria[3] and constitute what he describes as "the causal literature" (p. 10) on the topic. Leaving aside whether these studies actually identify causal effects, which will be discussed in Key Claim 2, I believe that 12 of the 33 studies do not belong in his list according to his own criteria.

1) Eight studies do not provide any direct estimates of the effect of additional spending on student outcomes.

Eight of these studies should be excluded from his list because they do not provide any estimates of "the effect of school spending on student outcomes." Instead, these are studies of the effects of policy interventions on student outcomes. Those interventions may involve increased spending, but they do not exclusively, or, in some cases, primarily involve additional money. If one looked in these eight studies for the results showing how much student outcomes changed because of extra funding, you would not be able to find them. Rather than estimating the effect of spending, these studies focus on the effect of time since the policy intervention.

Some of the authors are very clear that they are studying policy interventions that involve several components in addition to increased spending. When a policy intervention provides schools with more money but also changes school operations in

[3] Those three criteria are that studies must have "relied on quasi-experimental or policy variation," "demonstrate that their analysis is based on policies (or policy-induced variation) that had a large enough effect on school spending to facilitate exploring the effect of school spending on student outcomes," and "demonstrated that the variation in school spending examined is unrelated to other determinants of student outcomes such as other policies or demographics." (p. 11)

other ways, we have no way of knowing whether any changes in student outcomes following the intervention were caused by the extra money or by the other changes that schools made. For all we know, the other ways in which schools were changed could have caused all of the improvement and the additional spending made no difference or even hindered their effort.

a) One of these 8 studies examines the effects of the School Improvement Grant (SIG) Program, not extra spending

For example, one of the studies Jackson includes in his list, Carlson & Lavertu (2018), conducts a study of the School Improvement Grant (SIG) Program, which was a multi-prong effort to improve struggling schools. As Carlson and Lavertu put it, "SIG was designed to achieve such improvements by providing schools with additional financial resources and by requiring significant changes to many aspects of schools' educational delivery, particularly their leadership and staffing, as well as their use of data to drive instructional and managerial decision making." (p. 287) They emphasize that SIG was highly prescriptive about the operational changes schools needed to make independent of receiving additional money: "Earlier we noted that schools receiving a SIG award were required to implement one of four federally approved turnaround models: Closure, Restart, Transformation, or Turnaround. Table 2 demonstrates that nearly every school implemented either the Turnaround or Transformation model" (p. 299) Carlson & Lavertu (2018) do show that schools participating in the SIG Program received extra money, but they do not and cannot isolate the effect of that extra money on student achievement from the effects of other components of that intervention. Their study only tells us about the effects of the multi-prong SIG Program as a whole and says nothing about the independent effect of additional spending. In fact, it does not contain a single estimate of the effect of spending on student outcomes. It is therefore inappropriate for Jackson to include this and other studies like this in his "list of all studies on the effect of school spending on student outcomes."

b) Two of these 8 studies examine the effects of school finance reforms that often include other significant policy changes, not only extra spending.

Two studies in Jackson's list report the effect of time since implementation of school finance reforms (SFRs) on student outcomes: Brunner, Hyman, & Ju (2019) and Lafortune, Rothstein, & Schanzenbach (2018). Despite being labeled as "finance" reforms, almost all of the major SFRs were accompanied by other significant reforms to how schools were governed and operated. Simply measuring the effect of time since adoption of this bundle of policy interventions does not provide us with any information about the effect of additional spending since we cannot distinguish extra funding from the effect of other contemporaneous policy changes.

As one of the studies in Jackson's list, Clark (2003), describes a typical SFR, "The Kentucky Education Reform Act (KERA), implemented in 1990, is one of the most ambitious and influential education reform policies ever attempted by any state. KERA's main components include a new funding system to correct large financial disparities between school districts, curriculum revision and standardization, and increased school and district accountability." (p. 1)

Another study in Jackson's list, Candelaria & Shores (2019), explicitly acknowledges the problem of SFRs being coincident with other policy changes: "Although our results suggest that school resources affect graduation rates, the primary threat to validity is whether changes in graduation rates following court-ordered finance reform can be wholly attributable to changes in school spending. For the exclusion restriction to hold, we must assume that the court reforms affect graduation rates only through their effect on spending. This assumption is violated in cases where court-ordered finance affects other unobserved policy changes that also affect graduation rates. For example, in addition to increasing spending in high-poverty districts, the state may also adopt an incentive policy to bring higher-quality teachers to more impoverished districts. In such a scenario, graduation rates resulting from court order are not separable from the change in spending and the unobserved programmatic change." (pp. 52-3)

These studies do not provide any direct estimates of the effect of spending and instead report the effect of time since policy adoption. It is therefore inappropriate for Jackson to include these studies in his list since they only speak to the effects of the adoption of reform packages, not to the independent effect of additional spending.

c) Five of these 8 studies examine the effects of school tax elections, not extra spending.

Five studies in Jackson's list estimate the effect of time since an election to raise school taxes. Since the main purpose of a school tax election is to spend more on schools, it may seem like studying the effects of those elections is the same as studying the effects of extra spending, but that is not correct. The outcomes of school tax elections affect schools in ways other than through the additional funds they may receive. For example, since whether and when to hold school tax elections is largely under the control of school district administration, districts that lose those elections often lose confidence in their leadership and can see altered responsibilities or even dismissal of senior administration officials. As former newspaper publisher and school board member, George Scott, observed, "Superintendents tend to lose their jobs if they lose two bond issues." (Gray, 2014) So, winning a tax election can affect student outcomes through

the maintenance of stable school district leadership and confidence in their decisions by school boards and staff.

In addition, the results of school tax elections can affect student outcomes by changing housing prices, which in turn change student composition in the schools. Most of the studies in Jackson's list that examine school tax elections find that passage significantly increases housing prices. (See for example Cellini, Ferreira, & Rothstein (2010)) Raising school taxes signals that a community is committed to its schools. That tends to attract people who share that willingness to sacrifice by paying more for their children's education and to drive away people unwilling or unable to pay the higher taxes for schools in the community. Student outcomes, therefore, could change because there are different students in the schools and not because those schools received additional funding. Raising taxes to build a new school theater complex may not improve math and reading test scores, but it may attract higher achieving families to a district who are more likely to be interested in this kind of attractive amenity.

There are five studies in Jackson's list that estimate the effects of time since passage of a school tax increase on student outcomes but provide no actual estimate of the effect of additional spending: Baron (2019); Cellini, Ferreira, & Rothstein (2010); Kogan, Lavertu, & Peskowitz (2017); Martorell, Stange, & McFarlin (2016); and Rauscher (2019). These studies do not belong in his list because they do not provide an estimate on "the effect of school spending on student outcomes."

2) There are logical and statistical errors with imputing the effect of extra spending from these eight studies that examine the effects of policy interventions or elections.

Jackson's list includes these eight studies showing the relationship between policy changes, such as the SIG Program, SFRs, or tax elections, and student outcomes but not directly between spending and outcomes. If these studies do not have analyses showing the change in student outcomes per change in units of funding, how does Jackson derive an estimated effect of spending on outcomes in these cases? It appears that what he is doing is taking the estimated effect of the policy on student outcomes and then dividing that by the estimated effect of the policy on spending.

This approach is mistaken both logically and statistically. First, as I have been describing, it is wrong to assume that because a policy affects spending and a policy also affects outcomes, then it must be that spending affects outcomes. There are pathways other than through spending by which policies may affect outcomes. The logical flaw of this approach can be shown formally by observing that X may cause Y and X may also cause Z, but that does not mean Y causes Z.

Second, Jackson's inclusion of these eight studies and dividing the effect of policies on outcomes by the effect of policies on spending is statistically flawed. Both estimated effects have error in them. Even if both the estimated effect of a policy on outcomes and its estimated effect on spending were statistically significant, the error is compounded when we divide one estimate by another such that the ratio may well not be statistically significant. If we take two blurry photos that still allow us to see the images in them, we may not be able to make out the image if we superimpose the two blurry photos on each other. Jackson is committing a statistical error if he describes a result as statistically significant simply because it shows significant relationships between a policy and student outcomes as well as between a policy and spending.

3) One study does not meet Jackson's criteria because it examines the effect of spending equalization, not the effect of changing the level of school spending.

Another study in Jackson's list, Biasi (2019), does not even examine how a policy-induced increase in spending affects student outcomes. Instead, it is interested in how policy changes altered the equalization of school spending and whether that equalization affected student outcomes. As the author puts it, "A few studies have used school finance reforms as a quasi-experimental source of variation in school expenditure to study both short-term outcomes, such as student achievement and educational attainment (Hoxby, 2001; Card and Payne, 2002; Lafortune et al., 2018), and long-term outcomes, such as earnings (Jackson et al., 2015). The focus of these studies, however, has been to estimate the effects of increases in the levels of revenues and expenditure, as opposed to changes in their distribution across states, which are at the center of this study." (p. 6) This study should not be included in Jackson's list because it never estimates how policies affect the level of spending nor how that change in the level of spending may affect student outcomes.

4) Two studies do not meet Jackson's criteria because they do not examine "policy-induced variation."(p. 11)

Two other studies in Jackson's list, Miller (2018) and Gigliotti & Sorensen (2018), do not examine the effects of policy-induced changes in spending on student outcomes. Unlike studies based on school finance reforms or school tax elections, where funding is altered by new policies, these studies examine situations where policies remain the same and funding is altered by changes in housing prices or enrollment. Jackson's criteria required that studies examine "policy-induced variation" and therefore these studies ought to be excluded.

5) One study acknowledges that it does not meet Jackson's criteria that "the variation in school spending examined is unrelated to other determinants of student outcomes such as other policies or demographics." (p. 11)

An additional study in Jackson's list, Kreisman & Steinberg (2019), acknowledges that changes in school funding formulae, like changes in school taxes, may lead to changes in student composition in school districts. Any analysis of the variation in school funding as a result of that change in funding formula would therefore be endogenous and not causal. As the authors put it, "It is important to note that we do not observe student outcomes before and after the implementation of this rule, beginning in 1975 and in 1984, implying that the impacts we estimate are the long-run effects of increased district funding on district level outcomes. That is, we are in no position to rule out endogenous responses on the part of families living in, or migrating to, districts according to the way funding is allotted as a result of this specific element of the formula." (p. 9) This problem of drawing causal conclusions from SFR or other changes in funding policies would apply to many other studies in Jackson's list, which is a point I will discuss under Key Claim 2 below. But since these authors explicitly acknowledge that their analysis is endogenous and therefore is not causal, Jackson should have excluded it from his list as it violates his requirement that it "demonstrated that the variation in school spending examined is unrelated to other determinants of student outcomes such as other policies or demographics…" (Jackson Report, p. 11)

In total, 12 of the 33 studies in Jackson's list of methodologically appropriate studies of "the effect of school spending on student outcomes" do not appear to meet his own criteria for inclusion.

G. My re-analysis of Jackson's list of studies

Given the high rate of errors and inconsistencies in Jackson's review of the literature, I have attempted to re-analyze the classification of the direction and statistical significance of results for studies he identifies in his 2018 or Expert Report reviews.[4] I am not attempting to replicate the process by which he found studies for potential inclusion in his lists because he simply does not provide enough information to permit replication. He only says that he conducted a Google search, examined bibliographies, and consulted other researchers, but he does not say what terms he used for his Google

[4] I added to my re-analysis the 2014 working paper version of Lee & Polachek's 2018 study. The 2014 version presented test score results showing a negative but statistically insignificant effect that was omitted from the later version. The omission of the test score analysis and its negative result from the published article raises questions about whether the literature more readily available for Jackson to review suffers from "file drawer bias," in which politically undesirable results are less likely to be reported and published. See Key Claim 4 below.

searches or name the other researchers he consulted, so it is not possible to replicate the process by which he identifies studies.

But it is possible to examine the studies he lists in 2018 or in the Report and see how I would classify their eligibility for inclusion given his stated criteria and how I would characterize the direction and statistical significance of results in those studies. I also do not offer this re-analysis to claim that it demonstrates with confidence what this literature really says, given all of the other problems with Jackson's meta-analysis. Instead, I only offer it to show how imprecise and ambiguous this enterprise is since different people may apply their judgment differently in classifying these studies. The results of my re-analysis can be found in Exhibit 8. As one would expect from the preceding discussion, how I code these studies differs considerably from how Jackson does so. In total, there are 28 discrepancies between how Jackson and I would characterize these studies.

In addition, Exhibit 8 documents the extent to which Jackson classifies studies in the Expert Report differently from how he classified them in his 2018 review, which is supposed to be the basis for the current one. In total, there are 13 discrepancies between how Jackson characterized the same studies between his two reviews of them. And of the 11 studies described in Jackson, Wigger, and Xiong (2018), six of them are described differently in Jackson's Report.

The high of rate of disagreement between my analysis and Jackson's, as well as between Jackson's previous assessments and his current one, clearly emphasizes the extent to which summarizing a research literature in this fashion is not a precise scientific enterprise. Instead, it involves judgment and differing assessments of what constitutes persuasive evidence. Jackson's review is not only flawed by the high rate of errors and inconsistencies, but also by the false confidence he conveys in his conclusions. The evidence is simply too ambiguous and conflicting to support claims that it is "overwhelming" or for the calculation of his being mistaken as being "one in 8,589,934,592." (p. 12)

Exhibit 8 -- Discrepancies And Differences in Classification of Studies

	Jackson (2018)				Jackson, Wigger, & Xiong (2018)			Jackson Report			Greene Re-Analysis			
Study	Pos. & Sig.	Not Sig.	Neg. & Sig.	Meets Criteria for Inclusion	Pos. & Sig.	Pos. & Not Sig.	Neg & Not Sig.	Pos. & Sig.	Pos. & Not Sig.	Meets Criteria for Inclusion	Pos. & Sig.	Not Sig.	Neg. & Sig.	Meets Criteria for Inclusion
Abott, Korgan, Lavertu, & Peskowitz (2019)				NA		Yes		Yes		Yes	Yes			Yes
Biasi (2019)	Yes			Yes				Yes		Yes	Yes			No
Brunner, Hyman, & Ju (2019)	Yes			Yes		Yes			Yes	Yes	Yes			No
Candelaria & Shores (2019)	Yes			Yes					Yes	Yes	Yes			Yes
Card & Payne (2002)	Yes			Yes				Yes		Yes	Yes			Yes
Cascio, Gordon, & Reber (2013)	Yes			Yes					Yes	Yes		Yes		Yes
Downes & Figio (1997)	Yes			Yes						No		Yes		Yes
Hoxby (2001)		Yes		Yes						No		Yes		Yes
Jackson, Johnson, & Persico (2016) Johnson & Jackson (2019)	Yes			Yes				Yes		Yes	Yes			Yes
Jackson, Wigger, & Xiang (2018)	Yes			Yes				Yes		Yes	Yes			Yes
Johnson (2015)	Yes			Yes				Yes		Yes	Yes			Yes
LaFortune, Rothstein, & Schanzenbach (2018)	Yes			Yes	Yes				Yes	Yes		Yes		No
Miller (2018)	Yes			Yes	Yes			Yes		Yes	Yes			No
Baron (2019)				NA				Yes		Yes	Yes			No
Carlson & Lavertu (2018)				NA					Yes	Yes	Yes			No
Cellini, Ferreira, & Rothstein (2010)		Yes		Yes				Yes		Yes		Yes		No
Clark (2003)		Yes		Yes		Yes			Yes	Yes		Yes		Yes
Conlin & Thompson (2017)	Yes			Yes					Yes	Yes		Yes		Yes
Gigliotti & Sorensen (2018)	Yes			Yes	Yes			Yes		Yes	Yes			No
Goncalves (2015)		Yes		Yes					Yes	Yes			Yes	Yes
Guryan (2001)	Yes			Yes		Yes		Yes		Yes	Yes			Yes
Holden (2016)	Yes			Yes					Yes	Yes		Yes		Yes
Hong & Zimmer (2016)	Yes			Yes					Yes	Yes		Yes		Yes
Husted & Kenny (2000)	Yes			Yes						No		Yes		No
Hyman (2017)	Yes			Yes				Yes		Yes	Yes			Yes
Kogan, Lavertu, & Peskowitz (2017)	Yes			Yes		Yes		Yes		Yes		Yes		No
Kreisman & Steinberg (2019)				NA	Yes			Yes		Yes	Yes			No
LaFortune & Schonholzer (2018)	Yes			Yes				Yes		Yes	Yes			Yes
Lee & Polachek (2018)	Yes			Yes				Yes		Yes	Yes			Yes
Lee & Polachek (2014)				No						No		Yes		NA
Martorell, Stange, & McFarlin (2016)		Yes		Yes					Yes	Yes		Yes		No
Matsudaira, Hosek, & Walsh (2012)		Yes		Yes			Yes		Yes	Yes		Yes		Yes
Neilson & Zimmerman (2014)	Yes			Yes					Yes	Yes	Yes			Yes
Papke (2008)	Yes			Yes				Yes		Yes	Yes			Yes
Rauscher (2019)				NA					Yes	Yes		Yes		No
Roy (2011)	Yes			Yes				Yes		Yes	Yes			Yes
Van der Klaauw (2008)		Yes		Yes						No			Yes	Yes
Weinstein, Stiefel, Schwartz, & Chalico (2009)		Yes		Yes			Yes		Yes	Yes			Yes	Yes
								4	10	4	4	6	3	15
								Total	18		Total	28		
						Discrepancies from Jackson (2018)					Differences from Jackson Report			
						Discrepancies from Jackson, Wigger, & Xiong (2018)								

Key Claim 2 – Jackson focuses on a set of studies that he claims demonstrate "credible causal relationships between school spending and outcomes" (p. 6) based on "natural experiments." (p. 44) While the research designs Jackson prioritizes can approximate the causal estimates derived from actual experiments if certain assumptions are strictly met, in the set of studies Jackson considers, those assumptions are routinely violated, and their results should not be considered causal.

A) Studies of School Finance Reforms (SFRs) using Difference in Difference Event Study (Diff in Diff) or Instrumental Variable (IV) Research Designs

A large portion of the studies Jackson claims to be causal examine the effects of school finance reforms (SFRs), five of which examine effects of SFRs across multiple states, three of which examine the effects of a SFR in Michigan, and one of which examines the effects of a SFR in Kentucky. These studies employ one of two research designs that can produce results that approximate being causal: Difference in Difference Event Study (Diff in Diff) or Instrumental Variable (IV).

Essentially, the Diff in Diff studies compare the trend in an outcome before and after adoption of a SFR relative to the trend before and after that date in places where that event did not occur. As Jackson rightly notes, to isolate something approximating the causal effect of additional spending on outcomes, this estimation strategy requires that "there were no other coincident policies or changes that occurred at the time of the event." (p. 34)

Those that employ the IV approach use the event of the SFR as an "instrument" to predict the resulting change in spending. They then examine the relationship between this predicted level of spending and student outcomes. This technique can produce results that approximate being causal as long as the SFR has no effect on student outcomes other than through how it alters school spending. Like with the Difference in Difference approach, this requires assuming that the SFR was not coincident with policy changes that could affect student outcomes in ways other than spending.

However, school finance reforms typically occur as part of reform packages that include multiple changes in how schools are governed and operated in addition to changing their funding. The contemporaneous changes in other policies affecting student outcomes violate the assumptions required for both Diff in Diff and IV analyses to be considered causal estimates of the relationship between spending and student outcomes.

As already discussed in Key Claim 1, F), 1), b) above, Clark (2003) and Candelaria & Shores (2019) describe in detail how SFRs are often coincident with policy changes that

extend beyond additional spending. And Candelaria & Shores (2019) explicitly acknowledge that this undermines the ability to draw causal conclusions: "the primary threat to validity is whether changes in graduation rates following court-ordered finance reform can be wholly attributable to changes in school spending. For the exclusion restriction to hold, we must assume that the court reforms affect graduation rates only through their effect on spending. This assumption is violated in cases where court-ordered finance affects other unobserved policy changes that also affect graduation rates." (pp. 52-53)

Lafortune, Rothstein, & Schanzenbach (2018) also acknowledge this threat to drawing causal conclusions, but they dismiss it, "Some of the reforms were accompanied by governance, curriculum, or accountability changes, though our assessment is that these additional changes were typically not very important or impactful." (p. 6) These authors may think that these other policy changes were unimportant, but the inability to rule out that student outcomes were caused by other policy changes means that their analysis is not causal and is no different from the observational studies that Jackson declares are uninformative. In both cases, researchers assert that other unexamined factors probably do not matter, but they cannot conclude that with scientific confidence.

Similarly, Jackson, Johnson, & Persico (2016) consider the possibility that "recent policy reforms that started in the late 1980s (such as charter schools and test-based accountability)" (p. 205) confound their results. They then divide their sample to see if more recent cohorts of students, who may have also been exposed to these recent reforms, show different effects from SFRs than previous ones. The problem with this approach is that it assumes that coincident changes in school spending and other policy changes have not been an issue in both time periods. The multi-state studies of SFRs simply cannot be considered causal because the effects of school spending they are studying are inseparable from other school reforms adopted around the same time.

The three studies focused on the effect of a SFR in Michigan, called Proposal A, (Hyman (2017), Papke (2008), and Roy (2011)) face the same intractable difficulty with drawing causal conclusions. While Proposal A did change spending for certain school districts, it was contemporaneous with the adoption of charter schools and accountability measures that may have also affected those same districts. As researchers Julie Berry Cullen and Susanna Loeb put it, "In addition to the changes in raising and delivering funds to school districts, Proposal A also included new school choice measures and led to a new accountability system. It is likely inevitable that such dramatic increases in the state role in education are accompanied by increased state oversight and involvement." (2004, p. 13) The inability of these studies to separate the effect of funding changes from the effect of these other manifestations of state oversight and involvement means

that they are unable to generate causal estimates on the relationship between spending and student outcomes.

The benefits of accountability and other reforms that have occurred at the same time as SFRs are not merely hypothetical. There is a large body of evidence showing improvements in student outcomes from these reforms that are as large, or larger, than those produced by SFRs. For example, Dee & Jacob (2011) examine the effects of national accountability reforms adopted in No Child Left Behind (NCLB) and find: "Our results indicate that NCLB generated statistically significant increases in the average math performance of fourth graders (effect size = 0.23 by 2007) as well as improvements at the lower and top percentiles. There is also evidence of improvements in eighth-grade math achievement, particularly among traditionally low-achieving groups and at the lower percentiles." (p. 418) Carnoy & Loeb (2002) conduct a study of the effects of state accountability reforms and find that "students in high-accountability states averaged significantly greater gains on the NAEP 8th-grade math test than students in states with little or no state measures to improve student performance." (p. 305)

Research by Sass, Zimmer, Gill, & Booker (2016) examine charter schools in Florida and conclude that "students attending charter high schools are more likely to persist in college, and that in their mid-20s they experience higher earnings." (p. 683) Using a randomized experiment, Angrist, Cohodes, Dynarski, Pathak, & Walters (2016) find that charter schools in Boston improve student performance on state, AP, and SAT tests, although the attainment benefits in college are less clear.

The fact that SFRs are often accompanied by other reforms and those other reforms have been demonstrated to significantly improve student outcomes means that it is quite plausible that some or all of the benefits Jackson attributes to additional spending from SFRs could actually be attributable to other contemporaneous policy changes.

B) Studies of School Tax Elections Using Regression Discontinuity (RD) Research Designs

Another common type of study that Jackson considers causal involves the use of Regression Discontinuity (RD) in school tax elections. (Abott, Korgan, Lavertu, & Peskowitz (2019); Baron (2019); Cellini, Ferreira, & Rothstein (2010); Hong and Zimmer (2016); Kogan, Lavertu, & Peskowitz (2017); Lee & Polachek (2018); Lee & Polachek (2014); Martorell, Stange, & McFarlin (2016); and Rauscher (2019)) These studies compare student outcomes for districts where voters barely pass measures to increase school spending to districts where those measures barely fail. The argument for why this approximates causal effects is that whether a measure barely passes or barely fails can be thought of as essentially random. If the districts that barely pass measures later

have better student outcomes, these researchers believe, it would have to be because they won the election and not any other pre-existing differences.

In Key Claim 1, F) 1) c) I already discussed how passage of a school tax can affect student outcomes in ways other than by increasing school spending. In particular, I noted that whether school tax measures pass or not could affect student outcomes by altering district leadership or by changing the mix of students in the district, as demonstrated by changing housing prices. If school tax elections alter outcomes in ways other than by changing spending, then these studies are unable to separate the effect of school spending from the effects of these other consequences of election outcomes.

In addition to these difficulties with drawing causal conclusions about the effects of increased spending on student outcomes, it is worth noting that these analyses routinely violate another assumption required to think of RD as approximating causal results. To believe that school districts with election outcomes barely above a passing threshold and those barely below are effectively randomly assigned, we would have to believe that school districts are unaware of how close they likely are to the cut-off and be unable to do anything to alter that outcome. But districts typically monitor the progress of their election campaigns through polling or by their informal sense of the community. And if they detect that measures are struggling, they can exert more effort and devote additional resources toward passage.

Losing a school tax election, even by a modest margin, is therefore likely to be associated with administrative incompetence, which is also very likely to be negatively associated with future student outcomes. As a Chamber of Commerce official commented following the defeat of a school tax increase in California, "I think the school district is mismanaging how they spend their money and mismanaging how they create a quality education for all their kids. [Before asking for more money] the district needs to get its house in order both fiscally and academically." (Blume, 2019) The bias introduced by the fact that districts that lose elections tend to be less administratively competent is exacerbated by the fact that many of these RD studies do not restrict their samples only to elections that are very close to the threshold for winning.

C) Changing school funding is likely to change student composition in schools, which could also alter outcomes independently of the additional spending.

The student composition of school districts is not only likely to change as a result of school tax increases, but also in response to any significant change in school spending. Whether schools spend more because of an election, a school finance reform, or programs like Title I, the additional funding is likely to attract families that value higher educational spending. This means that in almost all of the studies in Jackson's list, it is

impossible to fully distinguish changes in student outcomes caused by additional spending from the changes that would occur from different student composition. It is true that many of these studies run analyses to see if schools saw changes in the racial/ethnic make-up and free lunch status of their students following the influx of additional spending. But these analyses cannot observe or control for all dimensions on which student composition might change and therefore cannot rule out the confounding influencing of changing student composition. This also means that these "causal" studies ultimately rely on observational correlations of a handful of student characteristics to draw their conclusions, just like the observational research literature that Jackson dismisses as uninformative.

D) In general, changes in school spending are almost never exogenous, making it very difficult to draw any causal conclusions with confidence.

SFR studies that use Diff in Diff or IV research designs and school tax election studies that use RD constitute the bulk of the studies in Jackson's list. For the reasons described above, we should be dubious that these studies actually generate causal estimates of the relationship between school spending and student outcomes. The remaining studies in Jackson's list are no more likely to be considered causal. The general problem is that in the real world we rarely have "natural experiments" in which school spending varies for reasons that are effectively random. That is, changes in school spending are almost never exogenous.

Jackson talks about studying SFRs as if they approximate experiments in which "the timing and location of the money drop is random." (p. 46) The term "money drop" is just rhetorical flourish, not a metaphor for any actual school spending process. In reality, the timing and location of money allocated to schools almost never approximate randomness. Even SFRs are political events that unfold over many years, and are shaped by the characteristics and academic trajectory of the affected schools. These processes are so slow-moving and complicated that even the researchers who study SFRs cannot agree on where or when SFRs have occurred. (See Key Claim 7, C)) There is nothing magical about additional money generated by SFRs or other policy changes that makes studying those dollars causal while studying all other dollars allocated to schools uninformative.

Key Claim 3 – Jackson's method of counting study results with positive or negative results, regardless of statistical significance, and then calculating the odds of having that many positive results if there were truly no effect, is an inappropriate and misleading method of determining statistical confidence in the general findings of a research literature.

A) Jackson's "coin flip" method of calculating confidence in his conclusions is not an accepted technique in social science, in general, or meta-analyses, in particular.

Jackson describes his method of calculating the confidence we should have in his summary of the research literature: "If we treat each study as an independent datapoint [sic], the likelihood of this many positive studies or more occurring by random chance (i.e., if there were no effect) is the same as flipping a coin 33 times and getting all heads. The likelihood of this is one in 8,589,934,592." (p. 12) This "coin flip" method is virtually unheard of and very rarely if ever used by scholars. It is akin to the vote counting method, but cruder in that it does not consider the statistical significance of the results. The nearly universal adoption of hypothesis testing in science inclines researchers to treat statistically significant results differently and to consider statistically insignificant results as indistinguishable from zero. Jackson's coin flip method is inconsistent with this scientific norm.

B) Jackson's listing of appropriate studies and classification of results contains too many errors, inconsistencies, and ambiguities to allow him to use the "coin flip" method and draw such confident conclusions.

Given the numerous errors, inconsistencies, and ambiguities in how Jackson classifies the direction and statistical significance of findings, calculating odds based on the number of positive versus negative results is as flawed as the method of classifying those studies. As we have already seen, several studies in Jackson's Expert Report and 2018 reviews should be classified as having negative results. And on at least two occasions, Jackson himself characterizes studies as having negative results in a prior review that he claims as being positive in his Report. (See Key Claim 1, C)) Jackson's over-confidence of putting the odds of his being mistaken at "one in 8,589,934,592" (p. 12) is completely inappropriate given the general imprecision of his review and the inherent ambiguity of this type of meta-analysis.

C) Jackson's "coin flip" method assumes that each study is an independent data point, which is not true of the list he has compiled.

Jackson correctly notes that an assumption required to do his coin flip method is that each study has to be an "independent datapoint [sic]." (p. 12) That assumption is clearly violated in his list of 33 studies. A small number of researchers or their co-authors produce a large portion of these 33 studies. Jackson, Lavertu, Rothstein, and Hyman or their co-authors account for 11 of the studies. Studies with overlapping authors are not independent of each other. In addition, a large block of studies examine some of the same school finance reforms, another block examine some of the same school tax elections, and another set examine Title I. Studies that examine the same or

overlapping policy interventions are not independent of each other. And almost all of the studies employ one of three research designs (Diff in Diff, IV, or RD). Studies using the same research design share the same methodological assumptions and are not independent of each other. Jackson has a lot fewer than 33 independent coin flips in his research review.

D) Jackson's "coin flip" method assumes that he has the complete and unbiased set of studies to consider, which we know to be incorrect.

Another assumption required for Jackson's calculation to be correct is that his 33 studies would have to constitute the complete and unbiased set of studies that should be considered. This assumption is also clearly violated. Jackson explicitly excludes four studies that were contained in his 2018 review, one of which has negative results. In addition, as I noted in Key Claim 1, G) we know that Jackson's review only includes the 2018 published version of Lee & Polachek that omitted the negative test score effects that were contained in an earlier version of that study released in 2014. The omission of negative results in the published version raises serious concerns about "file drawer bias," whereby undesired results may be much less likely to be reported or published. This will be discussed more fully in Key Claim 4.

We are confident that Jackson's claim to have 33 positive versus 0 negative results is incorrect given the exclusion of these known negative findings, but the existence of file drawer bias means that there are likely to be numerous other negative results that neither Jackson nor anyone else can easily find because they are never reported or published.

In addition, we have no reason to believe that Jackson's list of 33 studies constitutes the complete and correct set of studies that should be considered because he excludes numerous other studies, many of which have negative or null results, for failing to meet his criteria for methodological rigor. But as we have already seen, many of the studies in his list also fail to meet those criteria. In addition, because Jackson's set of 33 studies routinely violate the assumptions required for their findings of the relationship between additional spending and student outcomes to be considered causal, it is unclear why Jackson should not consider all studies on this issue. If he considered all studies, Jackson's "coin flip" technique would yield dramatically different results.

Key Claim 4 – Jackson's list of studies is not a complete and unbiased summary of research on the relationship between additional spending and student outcomes because it is likely to suffer from "file drawer" bias, in which studies are missing because they are never reported or otherwise difficult to find.

A) **We have good reason to doubt that the set of research findings Jackson considers in his Expert Report is complete and unbiased, which could significantly distort his conclusions.**

Accurately summarizing a research literature, especially with great confidence, requires consideration of a complete and unbiased set of research findings, which is extremely difficult to do. First, some findings may be less likely to be published in peer-reviewed journals if the editors or reviewers of those journals are less inclined to believe or favor certain results. Second, researchers themselves may find certain results less desirable, and may continue to re-analyze data until more favorable results can be found. If more favorable results cannot be produced, researchers may fail to complete that work, post results in working papers, or submit that work for consideration at journals. These undesired findings are metaphorically filed away in the researchers' drawers and may never appear in print, which is why this problem is known as "file drawer" bias. For these reasons, the set of published as well as unpublished findings available for meta-analysis can be seriously incomplete and distorted.

Failing to identify the complete set of findings can greatly alter the conclusions one would draw. For example, a study in the *New England Journal of Medicine* found that if one limited a meta-analysis of the effects of 12 anti-depressant drugs to published studies, 94% of the trials showed positive results. (Turner, et al, 2008) If, however, the set of studies was expanded to include all unpublished drug trials registered with the FDA, only 51% were positive. Aware of the danger of conducting meta-analyses or systematic reviews of incomplete sets of research findings, researchers typically go to great lengths to find all published and unpublished findings and carefully document the details of their efforts in those reviews.

We have no evidence that Jackson engaged in an exhaustive search for both published and unpublished findings on the relationship between school spending and student outcomes. He provides almost no detail about how his search was conducted other than to say that he engaged in Google searches, examined bibliographies, and consulted other researchers. We do not know what terms were used for his searches. He did not search other databases of education research that might include more specialized and unpublished work that Google may miss. And we do not know how broadly he consulted other researchers because they are not named. We know that Jackson missed the negative results in Lee & Polachek (2014) only because it happened to be an earlier working paper version of a published study in his list.

Given the lack of evidence that Jackson conducted an exhaustive search and given that we know he missed at least one negative result, it is difficult to have confidence in his overall conclusions. As a recent review of guidelines for conducting meta-analyses

emphasized, “The belief that all relevant studies have been ‘comprehensively’ identified, and that this process has been ‘transparently’ reported, increases confidence in the estimate of effect and the conclusions that can be drawn.” (Cooper, et al, 2018) Jackson’s review fails to meet this standard.

B) **The significant asymmetry of results in Jackson’s list is suspicious and indicates the presence of “file drawer” bias. Estimating missing findings to create a symmetrical pattern suggests that a complete and unbiased set of findings would likely yield a result that is indistinguishable from zero.**

While finding the complete set of published and unpublished studies is very challenging, it is literally impossible to find the set of findings that researchers never completed or never even posted as working papers. Nevertheless, the absence of those “invisible” studies from systematic reviews can significantly distort conclusions. Researchers who conduct meta-analyses are aware of this problem and have developed techniques for identifying the results that are likely missing and making adjustments for their absence. One of the main tools researchers use to identify and correct for this problem is called a “funnel plot.” (See, for example, Egger, Smith, Schneider, & Minder, 1997) It is based on the observation that in a complete set of research findings, results should be clustered around the true effect and distributed symmetrically, with the dispersion of results growing wider when estimates are less precise. The reason results should be distributed symmetrically is that if the meta-analysis’ conclusion is the true effect and errors are random, then there should be as many results that err by finding larger effects as those that err by finding smaller effects. If results are not distributed symmetrically, then we can impute the findings of the results that are likely missing and would make the pattern symmetrical.

Jackson does not provide sufficient detail for me to construct a funnel plot, but I have created a histogram of the estimated effect sizes of the studies in Jackson’s list that similarly allows us to assess whether the studies Jackson considers are symmetrically distributed. Because Jackson does not describe the exact effect size for each study, I had to estimate them from the bar graphs in his figures. If a study appeared more than once in Jackson’s figures, I took an average of the effects. (See Exhibit 3) Also, because only 31 of the 33 studies in Jackson’s list are included in his figures and appear to be included in his estimates of combined effects, I am limited to those 31 results.

As can be seen in Exhibit 8, it is clear that the findings in Jackson’s list are not symmetrically distributed. The largest group (13 studies) have effect sizes that are between 0 and .05. The next largest groups of studies (7 each) have effect sizes between .05 and .1 and between .1 and .15. Two studies have effects between .15 and .2, one is between .2 and .25, and one study has an effect of .35. The large difference

between the mean and median effects of the combined research literature that Jackson reports gives us some inkling of this skewed distribution, but Exhibit 9 shows that the asymmetry of results is quite stark.

If Jackson's list of studies were complete and unbiased, it would be extremely unlikely to find so many studies with results that are just barely on the positive side of zero and to find none on the negative side. Just by chance, some of the findings should be to the left of the main grouping of results, which means that they would report negative results. Jackson seems to think that the complete absence of negative findings is proof of how right he is -- "one in 8,589,934,592." (p. 12) Ironically, the complete absence of negative results, while the bulk of findings is barely on the positive side of zero, suggests the opposite.

The pattern of findings illustrated in Exhibit 9 looks like a normal, or bell-shaped, curve that is truncated at zero. This is a highly suspicious pattern given that we should expect results to be normally distributed around the true effect. The asymmetry of results that produces the absence of negative findings strongly suggests that results are missing from Jackson's review. Those results could be missing because Jackson failed to conduct an exhaustive search to find all published and unpublished findings. Those results could be missing because of inconsistencies in how Jackson classifies the results of studies. Or those results could be missing because journals are less likely to publish and researchers are less likely to report negative results. Regardless of why those studies are missing, their absence from his list significantly distorts his conclusions. We can roughly estimate what the distribution of results would look like if it were symmetrical. (See Exhibit 10) That estimated symmetrical pattern of results suggests that the complete and unbiased set of findings would likely produce a combined result that was indistinguishable from zero.

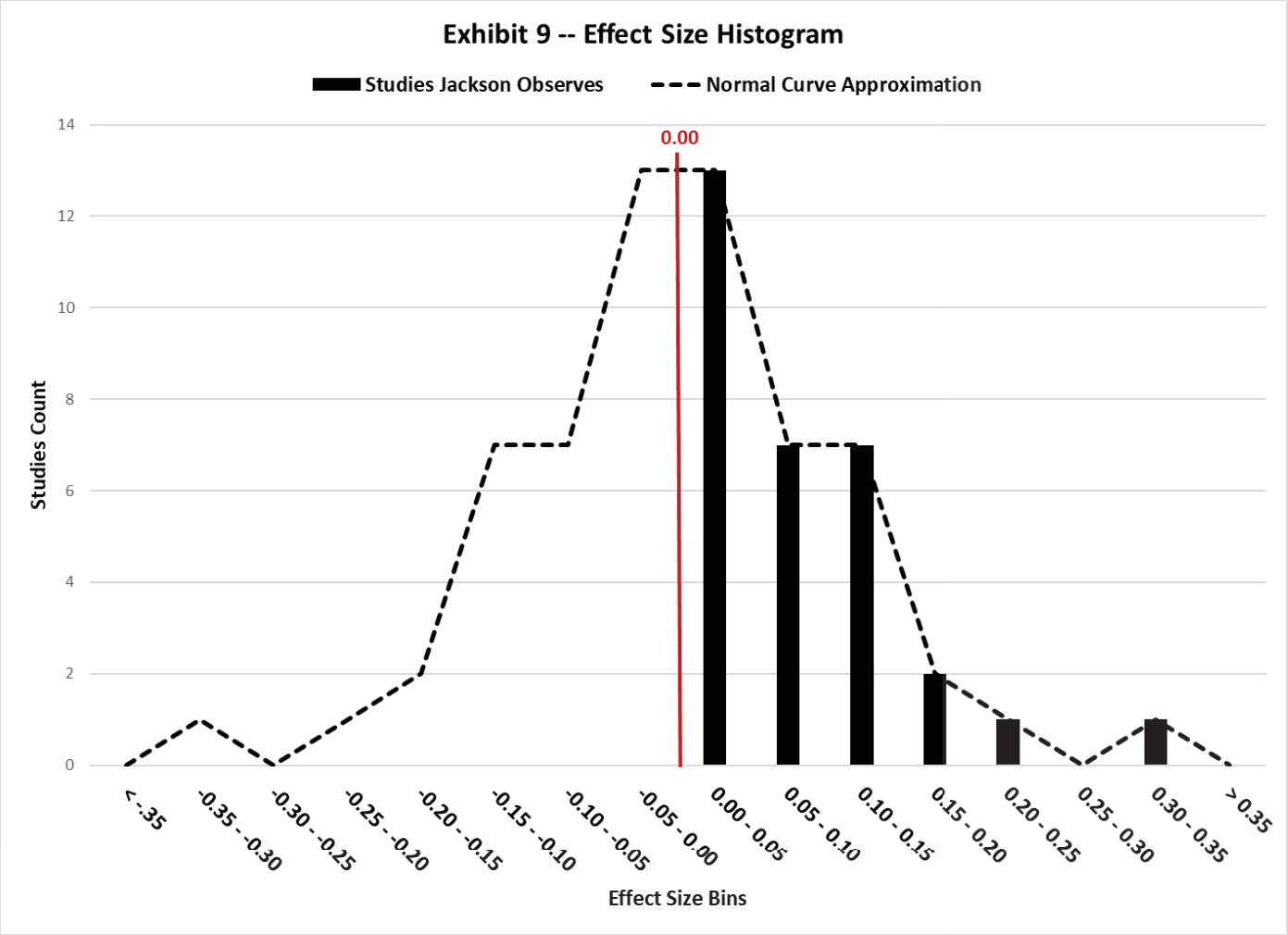


Exhibit 9 -- Effect Size Histogram
Studies Jackson Observes
Normal Curve Approximation
0.00
Studies Count
14
12
10
8
6
4
2
0
< -.35
-0.35 - -0.30
-0.30 - -0.25
-0.25 - -0.20
-0.20 - -0.15
-0.15 - -0.10
-0.10 - -0.05
-0.05 - 0.00
0.00 - 0.05
0.05 - 0.10
0.10 - 0.15
0.15 - 0.20
0.20 - 0.25
0.25 - 0.30
0.30 - 0.35
> 0.35
Effect Size Bins

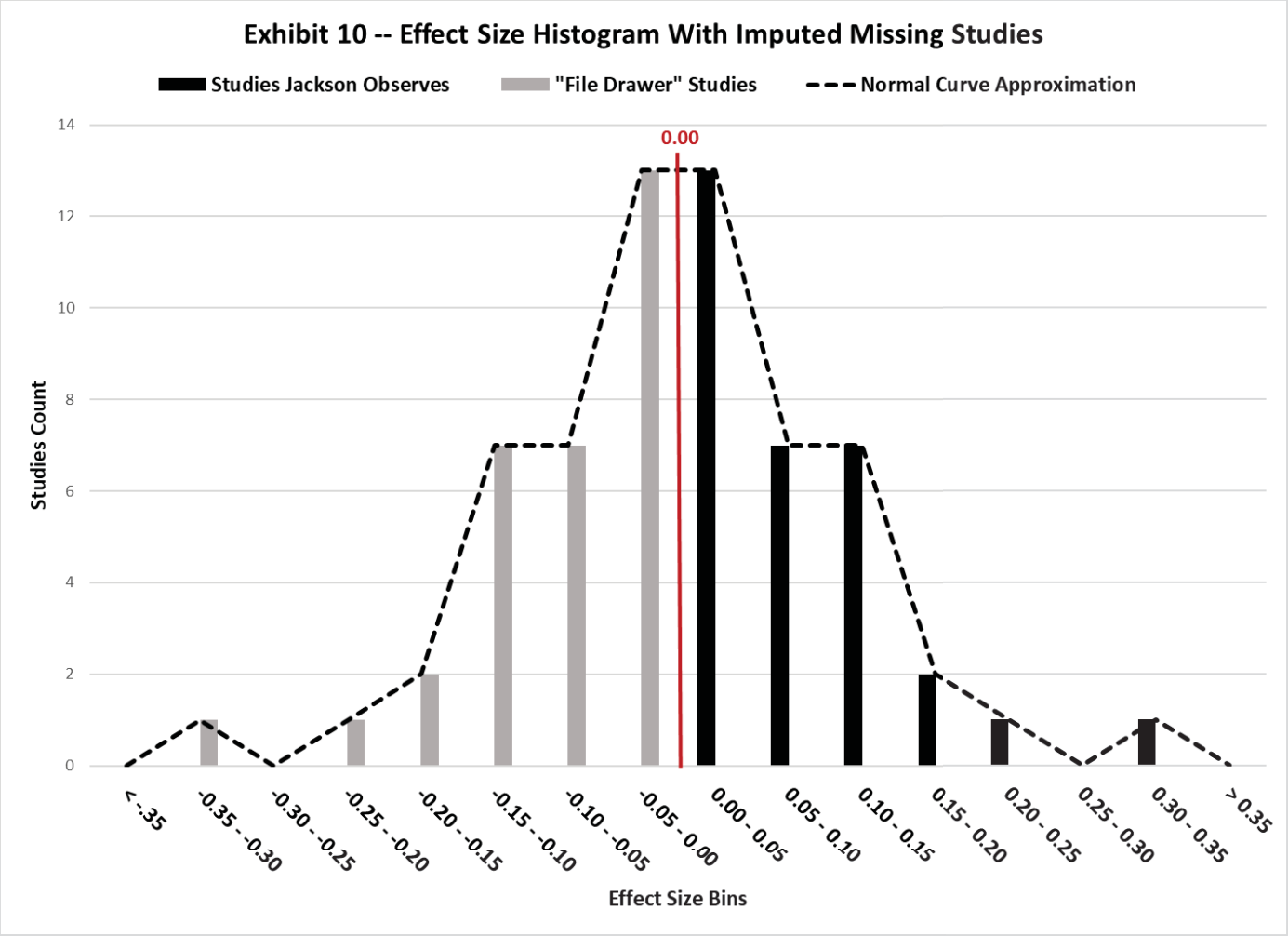


Exhibit 10 -- Effect Size Histogram With Imputed Missing Studies
Studies Jackson Observes
"File Drawer" Studies
Normal Curve Approximation
Studies Count
0.00
Effect Size Bins
< -.35
-0.35 - -0.30
-0.30 - -0.25
-0.25 - -0.20
-0.20 - -0.15
-0.15 - -0.10
-0.10 - -0.05
-0.05 - 0.00
0.00 - 0.05
0.05 - 0.10
0.10 - 0.15
0.15 - 0.20
0.20 - 0.25
0.25 - 0.30
0.30 - 0.35
> 0.35

Key Claim 5 – In his Report, Jackson provides specific claims about the extent to which student outcomes would improve if school spending were increased. These claims are highly implausible given the experience with past changes in school spending and student outcomes.

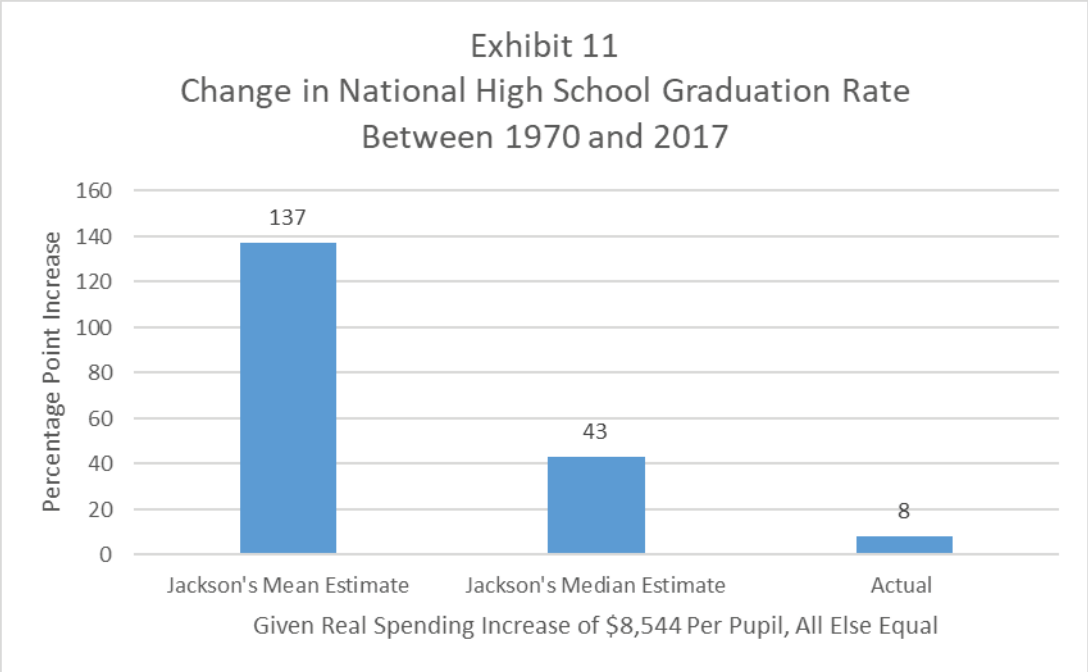
A) Jackson's claimed benefit for additional spending on high school graduation rates is implausible given the national experience between 1970 and 2017.

Based on his review of what he deems to be methodologically appropriate studies, Jackson claims that "a policy that increases per-pupil spending by $1000 (that is sustained for four years) increases educational attainment on average by about 45 percent of a standard deviation. This corresponds to being about 16 percentage points more likely to graduate high school, and 20 percentage points more likely to attend college." (p. 6) If this were true and all else were equal, the national high school graduation rate should have increased by about 137 percentage points between 1970 and 2017.

According to the US Department of Education's Digest of Education Statistics, total per pupil spending in 1969-70 on average was $5,895, adjusted into 2018-19 dollars. (2019, Table 236.55) In 2016-17 total per pupil spending rose to $14,439 in 2018-19 dollars, an increase of $8,544 in real spending. In addition, the Digest of Education Statistics provides a reasonable longitudinal estimate of national graduation rates by dividing the number of diplomas awarded by the 17 year old population. (2018, Table 219.10) The estimated graduation rate for 1970 was 77%. If we can expect a 16 percentage point gain per $1,000 increase in spending, as Jackson claims, and if total spending per pupil increased by $8,544, then we would expect high school graduation rates to have increased by 137 percentage points, all else being equal.[5] According to the Digest of Education Statistics' estimate, the actual change in high school graduation rates was 8 percentage points between 1970 and 2017, not 137 percentage points. Jackson's estimate overstated the increase by more than 16 times the actual gain. (See Exhibit 11)

[5] Starting at a rate of 77%, an increase of 137 percentage points would lead us to expect a high school graduation rate of 214% in 2017. Since schools cannot graduate more than 100% of their students, the magnitude of the improvement in graduation rates Jackson claims we can expect from $1,000 in additional spending is simply not plausible. Technically, the standard deviation of graduation rates would shrink as the level approached 100%. If we applied a dynamic process, the expected increase in graduation rate per $1,000 of additional spending would become smaller as the graduation rate approached 100% and the standard deviation became smaller and could never exceed 100%. But even a graduation rate of 100% did not occur and is implausible. Since Jackson does not suggest this dynamic process and since the variation in graduation rates did not dramatically change between 1970 and 2017, extrapolating his prediction to its absurd conclusion is still informative.

Jackson seems aware that the average gains he expects from additional spending may be too large and offers a lower alternative: "These large average effects are potentially affected by two studies (Cascio, Gordon & Reber 2013, and Johnson 2015) that find very large effects of Title 1 spending on educational attainment after the inception of the program. As such, one may be interested in the more conservative median effects." (p. 31) It should be noted, however, that even Jackson's lower median estimate of a 5 percentage point increase per $1,000 in additional spending would result in an expected 43 percentage point increase in high school graduation rates nationwide between 1970 and 2017, all else being equal. That figure is also not realistic. (See Exhibit 11)



B) It is implausible that students could have become so much more challenging to educate that it largely offset the gains Jackson would have expected.

If we simply looked at the changes in student outcomes relative to the increases in school spending, we do not see anything approximating the types of gains that Jackson claims should have been produced. The only way Jackson's expected improvements could be this far off what actually happened is if students became so much more difficult to educate over time that it almost completely offset the large gains extra spending should have yielded. The suggestion that students in the US have become so much more challenging to educate since 1970 that it would largely erase an expected gain in graduation rates of 137 percentage points is at odds with the evidence and highly implausible.

Consider that real per capita GDP in the US more than doubled between 1970 and 2017, going from $23,944 to $56,238 in 2012 constant dollars. (US Real GDP Per Capita by Year, n.d.) Even households in the bottom quintile experienced an increase in their mean income from $11,527 in 1970 to $13,647 in 2017, adjusted for inflation. (U.S. Census Bureau, Table H-3) Before counting the receipt of government benefits, the percentage of families below the poverty level was 10.9% in 1970 and 10.3% in 2017. (U.S. Department of Education, 2018, Table 102.50) The percentage of 3 and 4 year old children in pre-school has increased from 20.5% in 1970 to 53.8% in 2017. (U.S. Department of Education, 2018, Table 202.10) As an indication of improvements in childhood health, in 1970 19.9 out of every 1,000 children born in the U.S. would die before their first birthday; in 2017 that figure dropped to 5.7. (Federal Reserve Bank of St. Louis, n.d.) On a whole host of indicators, children in the United States are better off and better prepared to learn in school than they were in 1970. It beggars belief that children are so much more difficult to educate that it would largely erase huge gains in student outcomes that additional spending should have produced if we believed Jackson's claims.

Jackson is right that simply comparing the improvements in student outcomes to increases in school spending does not provide us with causal estimates of that relationship. At the same time, if research claims to be rigorous but yields results that are just implausible given our experience, we should suspect that there is something seriously wrong with the research claims.

Key Claim 6 – Even if the studies in Jackson's list could truly be considered causal, and even if his characterization of that literature were complete, accurate, and unbiased, it is highly unlikely that the findings from this literature are applicable to the current circumstances in Delaware.

A) Studies showing the effects of additional funding when school spending is very low are unlikely to be applicable to Delaware given the state's already high level of spending.

Delaware currently spends far more per pupil, in both absolute and relative terms, than most of the states and districts did when they were examined by studies in Jackson's list. There is no reason to believe that any benefit observed from additional spending from such low levels would apply if Delaware were to add to its already high level of school spending.

For example, most of the multi-state studies of the effects of SFRs include examining the effect of Tennessee's reforms in the early 1990s. According to the Digest of Education Statistics, in 1990 Tennessee had current school expenditures of $6,791, adjusted into 2018-19 dollars. At that time, 43 states spent more than Tennessee, including

Delaware, which spent $10,624. Not only did Tennessee's school spending significantly lag Delaware and most other states in 1990, but the amount that Tennessee spent then is less than half as much as the $15,543 that Delaware spent in 2016-17. (U.S. Department of Education, 2019, Table 236.65) Even if students in places like Tennessee benefited from increasing such low levels of spending, there is no reason to expect that Delaware students would receive comparable benefits by adding to their much higher level of funding.

B) The current political, social, and educational context in Delaware often differs dramatically from the situations examined in Jackson's list of studies, making their results unlikely to be applicable to Delaware.

In addition to examining the effects of spending at times and in places where the level of spending was much lower than in Delaware, many of the studies in Jackson's list examine the effects of additional spending in circumstances dramatically different from those currently found in Delaware. For example, Cascio, Gordon, & Reber (2013) look at the effect of Title I funding in the southern states in the 1960s. They found that racial politics were so fraught in the South at that time that Title I funding only improved outcomes for white students but made no difference for black students. Is this what we would expect from additional spending in Delaware today?

Martorell, Stange, & McFarlin (2016) examine the effects of quirks in the Texas funding formula for very rural and sparsely populated districts. Delaware does not generally have such rural and sparsely populated districts, so it is highly doubtful that what they learned in Texas would apply to Delaware. Similarly, Holden (2016) examine the effect of school districts in California receiving a one-time payment of $96.90 per pupil to buy textbooks. By including a study like this in his list, Jackson is extrapolating from the effect of $96.90 for textbooks to the effect of an extra $1,000 per pupil for operating expenses.

Neilson & Zimmerman (2014) study the effects of a school construction program in New Haven, where "schools reported problems with more than half of basic service systems, such as heating, air conditioning, plumbing, and lighting." If those are not similarly extensive problems in Delaware, additional spending to repair basic service systems that are not broken are unlikely to yield the same effects as Neilson & Zimmerman claim from New Haven. Similarly, Lafortune & Schonholzer (2018) study the effects of a school construction program in Los Angeles where over-crowding was a serious issue. Again, if over-crowding is not a problem to the same extent in Delaware, there is no reason to expect that additional spending would produce the same effects.

The context examined in most of the studies in Jackson's list differs so dramatically from the current context in Delaware that extrapolating from those results to confidently predict what we should expect from additional spending in Delaware seems imprudent.

Key Claim 7 – There are important inconsistencies between claims made by plaintiffs' experts, Kirabo Jackson, Jesse Rothstein, Hunter Gehlbach, and Clive Belfield.

A) Hunter Gehlbach and Clive Belfield make claims in their Expert Reports that are based on observational research designs that Kirabo Jackson and Jesse Rothstein dismiss as uninformative.

Much of the evidence cited by Gehlbach in his Report relies on observational studies of the type that Jackson and Rothstein dismiss as uninformative. For example, to demonstrate the benefits of tutoring and small group interventions, Gehlbach draws upon a meta-analysis by Slavin, et al (2011) that includes observational studies with "well-matched control groups." (p. 1) To demonstrate benefits of increasing school counselors, Gehlbach draws upon a study by Carrell & Carrell (2006) that uses an observational research design controlling for school fixed-effects. To demonstrate the benefits of increased teacher salaries, Gehlbach draws upon Baker (2016), which contains a review of mostly observational studies on the issue.

Clive Belfield's Report similarly makes claims that are based on observational studies that use research designs dismissed by Jackson and Rothstein. For example, Belfield lists in Table A-6 of his Report (p. 43) a variety of interventions that he believes could improve student outcomes. One of the listed interventions is "Consultant teachers: Literacy Collaborative," support for which comes from an observational study by Biancarosa, et al (2010). Another intervention touted by Belfield, Project Lead the Way, is also supported by evaluations using observational research designs. (Van Overschelde, 2013)

Plaintiffs' experts cannot dismiss observational studies of the effects of school spending on student outcomes at the same time that they embrace those kinds of studies to support various programmatic interventions.

B) Jackson, Rothstein, and Gehlbach differ on what specific interventions are likely to be effective in improving student outcomes.

Jackson, Rothstein, and Gehlbach differ in what specific uses of additional funding are supported by research as likely to be effective. Jackson in the 2016 article he co-authored writes, "the results suggest that the positive effects are driven, at least in part, by… increases in instructional time." (p. 211) But in his Report, Gehlbach concludes, "Investing in longer school days, longer school years, or more evenly distributed school

years seems less likely to yield clear benefits to students." (p. 9) In his Report, Rothstein emphasizes the likely benefits of additional spending on school facilities (pp. 14-15), but in his 2018 review, Jackson writes, "the evidence is consistent with capital spending and Title I spending being less predictably effective than spending in general." (p. 13) If the scientific evidence were so overwhelming and unambiguous, as Jackson claims in his Report, we would not expect to see these disagreements over what specific policy interventions are best supported by that evidence.

C) In their research examining school finance reforms, Jackson and Rothstein have numerous inconsistencies in what they consider a school finance reform and when it occurred, calling into question the scientific rigor of using SFRs as the instrument for judging the effects of increased school spending.

School finance reforms (SFRs) are the basis for a large portion of the research claiming that additional spending improves student outcomes. However, there is considerable disagreement among the researchers who examine SFRs as to what constitutes a SFR and when the event occurred. For example, in the 2018 article Rothstein co-authored with Lafortune and Schanzenbach, they provide an online appendix (p. ix) documenting all of the differences between how they treated SFRs and how Jackson, Johnson, & Persico (2016) did. (See Exhibit 12) In total, there were 23 disagreements in 14 different states over what constituted a school finance reform and when it occurred. Obviously, the numerous studies relying on SFRs to identify the effects of additional spending depend to a large degree on the subjective judgment of the researchers about how and when events should be classified. Any enterprise that dependent on inconsistent human judgment is lacking in scientific precision.

Jackson wants to characterize SFRs as roughly equivalent to a random "money drop" (p. 46) that would isolate the causal effects of additional spending on student outcomes. But the researchers who study these SFRs, including two of the plaintiffs' experts, cannot even agree on where or when those events took place. If they differ so often on what constitutes a SFR and when it took place, it is unreasonable to think that these disputed events somehow consistently reveal the causal effects of extra funding. The plaintiffs' experts cannot agree on where or when SFRs took place, but they are confident that wherever or whenever these things occurred, they must have improved student outcomes.

Exhibit 12 – Screenshot of Portion of Online Appendix E from Lafortune, Rothstein, & Schanzenbach (2018) Listing Discrepancies in the Identification of School Finance Reforms Between Lafortune, Rothstein, & Schanzenbach (2018) and Jackson, Johnson, & Persico (2016)

The states and years for which the two tabulations disagree are:

- Alabama, 1993
- Arizona, 2007
- Connecticut, 1995 & 2010
- Idaho, 1993 & 1998
- Maryland, 1996 & 2005
- Michigan, 1997
- Montana, 1993 & 2008
- New Hampshire, 2006
- New Jersey, 1991, 1998 & 2000
- New Mexico, 1998 & 1999
- Oregon, 2009
- South Carolina, 2005
- Texas, 2004
- Washington, 1991, 2007 & 2010

This includes only cases in scope for both lists but coded differently. This in particular means that we do not discuss our tabulation of legislative school finance reforms, as these are out of JJP's scope. For each state, we discuss only the events where the two tabulations disagree; see Online Appendix Table A1 for a full listing of events in each state.

Bibliography

Abott, C., Kogan, V., Lavertu, S., & Peskowitz, Z. (2020). School district operational spending and student outcomes: Evidence from tax elections in seven states. *Journal of Public Economics*, *183*, 104142.

Angrist, J.D., Cohodes, S.R., Dynarski, S.M., Pathak, P.A., & Walters, C.R. (2016) Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice. *Journal of Labor Economics,* 34(2) 275-318.

Baker, B. (2016). Does Money Matter in Education? Albert Shanker Institute. Retrieved from https://www.shankerinstitute.org/resource/does-money-matter-second-edition

Baron, E. J. (2019). School Spending and Student Outcomes: Evidence from Revenue Limit Elections in Wisconsin. *Available at SSRN 3430766*.

Biancarosa, G., Bryk, A.S., & Dexter, E.R. (2010). Assessing the value-added effects of Literacy Collaborative professional development on student learning. *The Elementary School Journal*, 111(1), 7-34.

Biasi, B. (2019). *School Finance Equalization Increases Intergenerational Mobility: Evidence from a Simulated-Instruments Approach* (No. w25600). National Bureau of Economic Research.

Blume, H. (2019). L.A. school tax measure defeated, in blow to Garcetti and Beutner. *Los Angeles Times*, June 4. Retrieved from https://www.latimes.com/local/lanow/la-me-edu-lausd-parcel-tax-election-20190604-story.html

Brunner, E., Hyman, J., & Ju, A. (2019). School Finance Reforms, Teachers' Unions, and the Allocation of School Resources. *The Review of Economics and Statistics*, 1–47.

Candelaria, C. A., & Shores, K. A. (2019). Court-Ordered Finance Reforms in the Adequacy Era: Heterogeneous Causal Effects and Sensitivity. *Education Finance and Policy*, *14*(1), 31–60.

Card, D., & Payne, A. A. (2002). School finance reform, the distribution of school spending, and the distribution of student test scores. *Journal of Public Economics*, *83*(1), 49-82.

Carlson, D., & Lavertu, S. (2018). School improvement grants in Ohio: Effects on student achievement and school administration. *Educational Evaluation and Policy Analysis*, *40*(3), 287-315.

Carnoy, M., & Loeb, S. (2002). Does External Accountability Affect Student Outcomes? A Cross-State Analysis. *Educational Evaluation and Policy Analysis*, 24(4), 305–331.

Carrell, S., & Carrell, S. (2006). Do Lower Student to Counselor Ratios Reduce School Disciplinary Problems?. *Contributions to Economic Analysis & Policy*, 5(1), 1463-1463.

Cascio, E. U., Gordon, N., & Reber, S. (2013). Local Responses to Federal Grants: Evidence from the Introduction of Title I in the South. *American Economic Journal: Economic Policy*, *5*(3), 126–159.

Cellini, S. R., Ferreira, F., & Rothstein, J. (2010). The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design. *Quarterly Journal of Economics*, *125*(1), 215–261.

Clark, M. A. (2003). *Education Reform, Redistribution, and Student Achievement: Evidence from the Kentucky Education Reform Act*. Unpublished Manuscript.

Conlin, M., & Thompson, P. N. (2017). Impacts of new school facility construction: An analysis of a state-financed capital subsidy program in Ohio. *Economics of Education Review*, *59*, 13–28.

Cooper, C., Booth, A., Varley-Campbell, J., Britten, N., & Garside, R. (2018). Defining the process to literature searching in systematic reviews: a literature review of guidance and supporting studies. *BMC Medical Research Methodology, 18, 85*.

Cullen, J., & Loeb, S. (2004). "School Finance Reform in Michigan: Evaluating Proposal A." in John Yinger, editor *Helping Children Left Behind: State Aid and the Pursuit of Educational Equity*. Cambridge, MA: MIT Press.

Dee, T.S., & Jacob, B. (2011). The Impact of No Child Left Behind on Student Achievement. *Journal of Policy Analysis and Management*, 30(3), 418-446.

Downes, T. & Figlio, D. (1998). School Finance Reforms, Tax Limits, and Student Performance: Do Reforms Level Up or Dumb Down?. Institute for Research on Poverty Discussion Paper no. 1142-97.

Egger, M., Smith, G.D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ [British Medical Journal]*, 315(7109): 629–634.

Federal Reserve Bank of St. Louis. (n.d.). Infant Mortality Rate for the United States. Retrieved from https://fred.stlouisfed.org/series/SPDYNIMRTINUSA

Gigliotti, P., & Sorensen, L. C. (2018). Educational resources and student achievement: Evidence from the Save Harmless provision in New York State. *Economics of Education Review*, *66*, 167–182.

Goncalves, F. (2015). The Effects of School Construction on Student and District Outcomes: Evidence from a State-Funded Program in Ohio. *SSRN Electronic Journal*.

Gray, L. (2014). Katy superintendent fights for bond issue. *Houston Chronicle*, July 23. Retrieved from https://www.houstonchronicle.com/news/education/article/Katy-superintendent-fights-for-bond-issue-5300834.php

Guryan, J. (2001). Does Money Matter? Regression-Discontinuity Estimates from Education Finance Reform in Massachusetts. *NBER Working Paper*.

Holden, K. L. (2016). Buy the Book? Evidence on the Effect of Textbook Funding on School-Level Achievement. *American Economic Journal: Applied Economics*, *8*(4), 100–127.

Hong, K., & Zimmer, R. (2016). Does Investing in School Capital Infrastructure Improve Student Achievement. *Economics of Education Review*, 44.

Hoxby, C.M. (2001). All School Finance Equalizations are Not Created Equal. *The Quarterly Journal of Economics,* 116(4), 1189–1231.

Husted, T.A., & Kenny, L.W. (2000). Evidence on the Impact of State Government on Primary and Secondary Education and the Equity-Efficiency Trade-Off. *The Journal of Law & Economics*, 43(1), 285-308.

Hyman, J. (2017). Does Money Matter in the Long Run? Effects of School Spending on Educational Attainment. *American Economic Journal: Economic Policy*, *9*(4), 256–280.

Jackson, C. K. (2018) Does School Spending Matter? The New Literature on an Old Question (No. 25368). National Bureau of Economic Research.

Jackson, C. K., Johnson, R. C., & Persico, C. (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *The Quarterly Journal of Economics*, *131*(1), 157-218.

Jackson, C. K., Wigger, C., & Xiong, H. (2018). *Do School Spending Cuts Matter? Evidence from the Great Recession* (No. w24203). National Bureau of Economic Research.

Johnson, Rucker C. (2015). Follow the Money: School Spending from Title I to Adult Earnings. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *1*(3), 50.

Kogan, V., Lavertu, S., & Peskowitz, Z. (2017). Direct Democracy and Administrative Disruption. *Journal of Public Administration Research and Theory*, *27*(3), 381–399.

Kreisman, D., & Steinberg, M. P. (2019). *The Effect of Increased Funding on Student Achievement: Evidence From Texas's Small District Adjustment*. Unpublished Manuscript.

Lafortune, J., Rothstein, J., & Schanzenbach, D. W. (2018). School Finance Reform and the Distribution of Student Achievement. *American Economic Journal: Applied Economics*, *10*(2), 1–26.

Lafortune, J., & Schonholzer, D. (2018). Do School Facilities Matter? Measuring the Effects of Capital Expenditures on Student and Neighborhood Outcomes. *Working Paper*, 85.

Lee, K.-G., & Polachek, S. W. (2018). Do school budgets matter? The effect of budget referenda on student dropout rates. *Education Economics*, *26*(2), 129–144.

Martorell, P., Stange, K., & McFarlin, I. (2016). Investing in Schools: Capital Spending, Facility Conditions, and Student Achievement (Revised and Edited). *Journal of Public Economics*.

Matsudaira, J. D., Hosek, A., & Walsh, E. (2012). An integrated assessment of the effects of Title I on school behavior, resources, and student achievement. *Economics of Education Review*, *31*(3), 1–14.

Miller, C. L. (2018). *The Effect of Education Spending on Student Achievement: Evidence from Property Values and School Finance Rules*. Unpublished Manuscript.

Neilson, C. A., & Zimmerman, S. D. (2014). The effect of school construction on test scores, school enrollment, and home prices. *Journal of Public Economics*, *120*, 18–31.

Papke, L. E. (2008). The Effects of Changes in Michigan's School Finance System. *Public Finance Review*, *36*(4), 456–474.

Rauscher, E. (2019). Delayed Benefits: Effects of California School District Bond Elections on Achievement by Socioeconomic Status. *EdWorkingPaper*, *No. 19-18*.

Roy, J. (2011). Impact of School Finance Reform on Resource Equalization and Academic Performance: Evidence from Michigan. *Education Finance and Policy*, *6*(2), 137–167.

Sass, T.R., Zimmer, R.W., Gill, B.P., & Booker, T.K. (2016). Charter High Schools' Effects on Long-Term Attainment and Earnings. *Journal of Policy Analysis and Management*, 35(3), 683–706.

Slavin, R., Lake, C., Davis, S., & Madden, N. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, 6(1), 1-26.

Turner, E.H., Matthews, A.M., Linardatos, E., Tell, R.A., & Rosenthal, R. (2008). Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy. *New England Journal of Medicine*, 358(3), 252-60.

U.S. Census Bureau. (n.d.) Historical Income Tables: Households. Table H-3. Mean Household Income Received by Each Fifth and Top 5 Percent, All Races. Retrieved from https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-households.html

U.S. Department of Education, Digest of Education Statistics. (2018). Table 102.50. Poverty status of all persons, persons in families, and related children under age 18, by race/ethnicity: Selected years, 1960 through 2017. Retrieved from https://nces.ed.gov/programs/digest/d18/tables/dt18_102.50.asp

U.S. Department of Education, Digest of Education Statistics. (2018). Table 202.10. Enrollment of 3-, 4-, and 5-year-old children in preprimary programs, by age of child, level of program, control of program, and attendance status: Selected years, 1970 through 2017. Retrieved from https://nces.ed.gov/programs/digest/d18/tables/dt18_202.10.asp

U.S. Department of Education, Digest of Education Statistics. (2018). Table 219.10: High school graduates, by sex and control of school; public high school averaged freshman graduation rate

(AFGR); and total graduates as a ratio of 17-year-old population: Selected years, 1869-70 through 2028-29. Retrieved from https://nces.ed.gov/programs/digest/d18/tables/dt18_219.10.asp

U.S. Department of Education, Digest of Education Statistics. (2019). Table 236.55: Total and current expenditures per pupil in public elementary and secondary schools: Selected years, 1919-20 through 2016-17. Retrieved from https://nces.ed.gov/programs/digest/d19/tables/dt19_236.55.asp

U.S. Department of Education, Digest of Education Statistics. (2019). Table 236.65. Current expenditure per pupil in fall enrollment in public elementary and secondary schools, by state or jurisdiction: Selected years, 1969-70 through 2016-17. Retrieved from https://nces.ed.gov/programs/digest/d19/tables/dt19_236.65.asp

US Real GDP Per Capita by Year. (n.d.). Retrieved from https://www.multpl.com/us-real-gdp-per-capita/table/by-year

van der Klaauw, W.(2008). Breaking the link between poverty and low student achievement: An evaluation of Title I. *Journal of Econometrics*, 142(2), 731–756.

Van Overschelde, J.P. (2013). Project lead the way students more prepared for higher education. *American Journal of Engineering Education*, 4(1).

Weinstein, M. G., Stiefel, L., Schwartz, A. E., & Chalico, L. (2009). Doesn't Title I Increase Spending and Improve Performance? Evidence from New York City. *IESP Working Paper Series*.

Jay P. Greene

201 Graduate Education Building
Department of Education Reform
College of Education and Health Professions
University of Arkansas
Fayetteville, AR 72701

E-mail jpg@uark.edu
Phone (479) 575-3162
Fax (479) 575-3196

EMPLOYMENT

University of Arkansas, Department Head and 21st Century Chair in Education Reform, August 2005 - present
The Manhattan Institute for Policy Research, Senior Fellow, January 2000 - August 2005.
University of Texas at Austin, Assistant Professor, July 1997 - May 2000.
University of Houston, Assistant Professor, July 1994 - July 1997.

EDUCATION

Harvard University, Cambridge, Massachusetts, Ph.D. in Political Science in June 1995; A.M. in Political Science in November 1991
Tufts University, Medford, Massachusetts, B.A. in History, *Summa Cum Laude* in May, 1988; *Phi Beta Kappa* 1987

SELECTED AWARDS AND HONORS

Significant Research Awards, College of Education and Health Professions, **University of Arkansas,** 2008, 2012, 2017
Presidential Appointment as a Trustee for the James Madison Memorial Foundation, 2003
Distinguished Teaching Award, Harvard University, 1994
Mellon Dissertation Completion Grant, 1993
Harvard MacArthur Fellowship, 1991-92
Marshall Scholarship, 1988-90 (declined)
Truman Scholarship, 1986-90
National Endowment for the Humanities, Younger Scholars Research Grant, 1987
Truman Library Research Grant, 1987
National Merit Scholarship, 1984-88

BOOKS

Religious Liberty & Education, edited with Jason Bedrick and Matthew Lee, Rowman and Littlefield, 2020.

Failure Up Close, edited with Michael Q. McShane, Rowman and Littlefield, 2018.

Why America Needs School Choice, Encounter Broadside, 2011.

Education Myths, with Greg Forster and Marcus A. Winters, Rowman and Littlefield, 2005.

EDITING

Guest Editor, *Peabody Journal of Education*, Volume 82, Issue 4, 2007.

Contributing Editor, *Education Next*, 2001-Present

PEER-REVIEWED PUBLICATIONS

Kisida, B., Bowen, D. H., & Greene, J. P. (2018). Cultivating Interest in Art: Causal Effects of Arts Exposure During Early Childhood. *Early Childhood Research Quarterly*, 45(4), 197-203.

Greene, J.P., Erickson, H.H., Watson, A.R., & Beck, M.I. (2018). The Play's the Thing: Experimentally Examining the Social and Cognitive Effects of School Field Trips to Live Theater Performances. *Educational Researcher*, 47(4), 246-254.

Greene, J.P., & Kingsbury, I. (2017). The Relationship Between Public and Private Schooling and Anti-Semitism. *Journal of School Choice*, 11(1), 111-130.

Kisida, B., Bowen, D. H., & Greene, J. P. (2016). Measuring critical thinking: Results from an art museum field trip experiment. *Journal of Research on Educational Effectiveness*, 9(1), 171-182.

Egalite, A. J., Mills, J., & Greene, J. P. (2016). The softer side of learning: Measuring students' non-cognitive skills. *Improving Schools*, *19*(1), 27-40.

Greene, J. P., Hitt, C., Kraybill, A., & Bogulski, C. A. (2015). Learning from live theater. *Education Next*, *15*(1), 54-61.

Margulis, E. H., Kisida, B., & Greene, J. P. (2015). A knowing ear: The effect of explicit information on children's experience of a musical performance. *Psychology of Music*, *43*(4), 596-605.

Kisida, B., Greene, J. P., & Bowen, D. H. (2014). Creating cultural consumers: The dynamics of cultural capital acquisition. *Sociology of Education*, *87*(4), 281-295.

Greene, J. P., Kisida, B., & Bowen, D. H. (2014). The educational value of field trips. *Education Next*, *14*(1), 78-86.

Bowen, D. H., Greene, J. P., & Kisida, B. (2014). Learning to think critically: A visual art experiment. *Educational Researcher*, *43*(1), 37-44.

Bowen, D. H., & Greene, J. P. (2012). Does athletic success come at the expense of academic success?. *Journal of Research in Education*, *22*(2), 1-22.

Winters, M. A., & Greene, J. P. (2012). The medium-run effects of Florida's test-based promotion policy. *Education Finance and Policy*, *7*(3), 305-330.

Winters, M. A., Dixon, B. L., & Greene, J. P. (2012). Observed characteristics and teacher quality: Impacts of sample selection on a value added model. *Economics of Education Review*, *31*(1), 19-32.

Greene, J. P., & McGee, J. B. (2012). When the best is mediocre: Developed countries far outperform our most affluent suburbs. *Education Next*, *12*(1), 34-40.

Winters, M. A., & Greene, J. P. (2011). Public school response to special education vouchers: The impact of Florida's McKay Scholarship Program on disability diagnosis and student achievement in public schools. *Educational Evaluation and Policy Analysis*, *33*(2), 138-158.

Greene, J. P. (2011). The Big Rock Candy Mountain of education. In G. Forster & C. B. Thompson (Eds.), *Freedom and School Choice in American Education* (pp. 1-16). New York, NY: Palgrave Macmillan.

Greene, J. P., & Winters, M. A. (2011). Florida's program to end social promotion. In D. L. Leal & C. B. Thompson (Eds.), *The politics of Latino education* (pp. 58-71). New York, NY: Teachers College Press.

Winters, M. A., Trivitt, J. R., & Greene, J. P. (2010). The impact of high-stakes testing on student proficiency in low-stakes subjects: Evidence from Florida's elementary science exam. *Economics of Education Review*, *29*(1), 138-146.

Winters, M. A., Ritter, G. W., Greene, J. P., & Marsh, R. (2009). Student outcomes and teacher productivity and perceptions in Arkansas. In M. G. Springer (Ed.), *Performance incentives: Their growing impact on American K-12 education* (pp. 273-293). Washington, DC: Brookings Institution Press.

Greene, J. P., & Winters, M. A. (2009). The effects of exemptions to Florida's test-based promotion policy: Who is retained? Who benefits academically?. *Economics of Education Review*, *28*(1), 135-142.

Greene, J. P., Butcher, J., Jensen, L. I., & Shock, C. (2008). You can't choose if you don't know: The failure to properly inform parents about NCLB school choice. *Georgetown Journal of Law & Public Policy*, *7*, 7-25.

Greene, J. P., & Winters, M. A. (2007). Revisiting grade retention: An evaluation of Florida's test-based promotion policy. *Education Finance and Policy*, *2*(4), 319-340.

Greene, J. P. (2006). Fixing special education. In A. S. Gerber & E. M. Patashnik (Eds.), *Promoting the general welfare: New perspective on government performance* (pp. 128-148). Washington, DC: Brookings Institution Press.

Winters, M. A., & Greene, J. P. (2006). Getting ahead by staying behind: An evaluation of Florida's program to end social promotion. *Education Next*, *6*(2), 65-69.

Greene, J. P., Winters, M. A., & Forster, G. (2004). Testing high-stakes test: Can we believe the results of accountability tests?. *Teachers College Record*, *106*(6), 1124-1144.

Winters, M. A., & Greene, J. P. (2004). Competition passes the test: Vouchers improve public schools in Florida. *Education Next*, *4*(3), 66-71.

Greene, J. P. (2001). The looming shadow: Can the threat of vouchers persuade a public school to turn itself around?. *Education Next*, *1*(4), 76-82.

Greene, J. P. (2001). Vouchers in Charlotte: Vouchers and the test-score gap. *Education Next*, *1*(2), 55-60.

Greene, J. P. (2001). The hidden research consensus for school choice. In P. E. Peterson & D. E. Campell (Eds.), *Charters, vouchers, and public education* (pp. 83-101). Washington, DC: Brookings Institution Press.

Wolf, P. J., Greene, J. P., Kleitz, B., & Thalhammer, K. (2001). Private schooling and political tolerance. In P. E. Peterson & D. E. Campbell (Eds.), *Charters, vouchers, and public education* (pp. 268-289). Washington, DC: Brookings Institution Press.

Greene, J. P, Giammo, J., & Mellow, N. (1999). The effect of private education on political participation, social capital and tolerance: An examination of the Latino National Political Survey. *Georgetown Public Policy Review*, *5*(1), 53-67.

Greene, J. P. (1999). A meta-analysis of the effectiveness of bilingual education. *Bilingual Research Journal*, *21*(2-3), 103-122.

Greene, J. P., Mellow, N., & Giammo, J. (1999). Private schools and the public good: The effect of private education on political participation and tolerance in the Texas poll. *Catholic Education*, *2*(4), 429-443.

Greene, J. P., Peterson, P. E., & Du, J. (1999). Effectiveness of school choice: The Milwaukee experiment. *Education and Urban Society*, *31*(2), 190-213.

Greene, J. P., & Peterson, P. (1999). Vouchers and central city schools. In C. Foreman (Ed.), *The African-American predicament* (pp. 82-96). Washington, DC: Brookings Institution Press.

Greene, J. P. (1998). Civic values in public and private schools. In P. E. Peterson & B. C. Hassel (Eds.), *Learning from school choice* (pp. 83-106). Washington, DC: Brookings Institution Press.

Greene, J. P., Peterson, P. E., & Du, J. (1998). School choice in Milwaukee: A randomized experiment. In P. E. Peterson & B. C. Hassel (Eds.), *Learning from school choice* (pp. 335-356). Washington, DC: Brookings Institution Press.

Greene, J. P., Howell, W. G., & Peterson, P. E. (1998). Lessons from the Cleveland Scholarship Program. In P. E. Peterson & B. C. Hassel (Eds.), *Learning from school choice* (pp. 357-392). Washington, DC: Brookings Institution Press.

Peterson, P. E., & Greene, J. P. (1994). Questioning by the foreign policy committees. In P. E. Peterson (Ed.), *The president, the congress, and the making of foreign policy* (pp. 74-97). Norman, OK: University of Oklahoma Press.

Peterson, P. E., & Greene, J. P. (1993). Why executive-legislative conflict in the United States is dwindling. *British Journal of Political Science*, *24*(1), 33-55.

ARTICLES, BOOK CHAPTERS, and REVIEWS

Greene, J.P., & Kingsbury, I. (2020). The Relationship Between Public and Private Schooling and Anti-Semitism. In R. Maranto & M.D. Shakeel (Eds.), *Educating Believers: Religion and School Choice*. Routeledge.

Greene, J.P. (2019). What Is the Purpose of Education? In L. Burke & J. Butcher (Eds.), *The Not-So-Great Society*. Washington, D.C.: The Heritage Foundation.

Greene, J.P. (2019). The moral and religious roots of social and emotional learning. *American Enterprise Institute*, June 2019.

Greene, J.P., & Hess, F. (2019). Education Reform's Deep Blue Hue. *Education Next*, published online on March 11.

Greene, J.P., & McShane, M.Q. (2018). Learning from school reform failures. *Phi Delta Kappan*, 99(8), 46- 50

Greene, J. P. (2017). Uncommon Sense for Education Reformers [Review of the book Commitment and Common Sense, by David P. Driscoll]. *Education Next*, published online on October 16.

Greene, J.P. (2017) Arts Integration Is a Sucker's Game. *Education Week*, 37(7), 23.

Greene, J.P. (2017). The High Regulation Approach to School Choice. In J. Allen, C. Candal, & M. Eden (Eds.), *Charting a New Course: The Case for Freedom, Flexibility, and Opportunity Through Charter Schools* (pp. 91-97). Washington, D.C.: Center for Education Reform.

Greene, J. P., & Bedrick, J. (2017). On Coulson's Historical Perspective, In N. McCluskey & J. Bedrick (Eds.), *Educational Freedom* (pp. 59-64). Washington, D.C.: Cato Institute Press.

Greene, J. P. (2017). Futile Accountability Systems Should Be Abandoned. *Education Next*, *17*(3).

Greene, J. P. (2017). Undoing Irrational Thinking [Review of the book *The Undoing Project*, by Michael Lewis]. *Education Next*, published online on February 13.

Greene, J. P. (2017). Teaching character: Grit is critical to how and why people succeed [Review of the books *Grit: The Power of Passion and Perseverance*, by A. Duckworth & *Helping Children Succeed: What Works and Why*, by P. Tough]. *Education Next*, *17*(1).

Greene, J. P. (2016). The trouble with texting: Success in college requires more than completing forms [Review of the book *The 160-Character Solution*, by B. L. Castleman]. *Education Next*, *16*(3).

Greene, J. P. (2016). Truth telling is academia's privilege (and obligation). *Education Week*, *35*(17), 25.

Greene, J. P. (2015). Buckets into another sea. In F. M. Hess & J. R. Henig (Eds.), *The New Education Philanthropy: Politics, Policy, and Reform* (pp. 11-28). Cambridge, MA: Harvard Education Press.

Greene, J. P. (2015). History lessons from a policy insider: What should we be learning from past reform fights? [Review of the book *Presidents, Congress, and the Public Schools: The Politics of Education Reform*, by J. Jennings]. *Education Next*, *16*(1).

Greene, J. P. (2015). Wrong diagnosis on homework help from parents: Authors find correlation, mistake it for causation [Review of the book *The broken compass: Parental involvement with children's education*, by K. Robinson & A. L. Harris]. *Education Next*, *15*(2).

Greene, J. P., Kisida, B., Bogulski, C. A., Kraybill, A., Hitt, C., & Bowen, D. H. (2014). Arts education matters: We know, we measured it. *Education Week*, *34*(13), 24.

Greene, J. P., & Bowen, D. H. (2014, October 21). Sports and education work well together. *The New York Times*.

Greene, J. P., Kisida, B., & Bowen, D. H. (2014). Why field trips matter: New study shows visiting an art museum improves critical thinking skills and more. *Museum*, *93*(1), 32-41.

Kisida, B., Greene, J. P., & Bowen, D. H. (2013, November 23). Art makes you smart. *The New York Times*, SR12.

Greene, J. P. (2014). Historian Ravitch trades fact for fiction [Review of the book *Reign of error: The hoax of the privatization movement and the danger to America's public schools*, by D. Ravitch]. *Education Next*, *14*(2), 80-81.

Greene, J. P. (2013, July 10). NCTQ doesn't know what works. *Education Next*, Retrieved from http://educationnext.org

Greene, J. P. (2012). Best practices are the worst [Review of the book *Surpassing Shanghai: An agenda for American education built on the world's leading systems*, edited by M. Tucker]. *Education Next*, *12*(3), 72-73.

Greene, J. P. (2012). Unions and the public interest. *Education Next*, *12*(1), 60-68.

Buck, S., & Greene, J. P. (2011). Blocked, diluted, and co-opted. *Education Next*, *11*(2), 26-31.

Greene, J. P. (2011). The education reform book is dead. *Education Next*, *11*(2), 78-80.

Greene, J. P. (2011). How schools spend their money [Review of the book *Education economics: Where do school funds go?*, by M. Roza]. *Education Next*, *11*(1), 84-85.

Greene, J. P. (2010). Look in the mirror [Review of the book *Making the grade: The economic evolution of American school districts*, by W. A. Fischel]. *Education Next*, *10*(3), 76-77.

Buck, S., & Greene, J. P. (2010). The case for special education vouchers. *Education Next*, *10*(1), 37-43.

Greene, J. P. (2009, December 14). A world inside the Bible Belt [Review of the book *Boom town*, by M. Rosen]. *The Wall Street Journal*, Retrieved from http://online.wsj.com

Winters, M. A., & Greene, J. P. (2009, October 19). End the special-ed racket. *National Review*, Retrieved from http://www.manhattan-institute.org

Greene, J. P. (2009). The use of vouchers: Keeping promises. *The City Journal* [Special issue], Retrieved from http://www.city-journal.org

Greene, J. P. (2008, August 1). Abraham Lincoln and the virtue of incremental reform. *Learning from Lincoln: Principle and pragmatism, a symposium*, 20.

Greene, J. P. (2008). "To Read or Not to Read": Responses to the New NEA Study: Not so bad. *Academic Questions*, *21*(2), 195-220.

Greene, J. P., & Trivitt, J. R. (2008). Can judges improve academic achievement?. *Peabody Journal of Education*, *83*(2), 224-237.

Greene, J. P., & Butcher, J. (2008, May 8). Proximity and power. *The American*, Retrieved from http://www.american.com

Greene, J. P., & Shock, C. (2008). Adding up to failure. *The City Journal*, *18*(1), Retrieved from http://www.city-journal.org

Greene, J. P. (2007). Fixing special education. *Peabody Journal of Education*, *82*(4), 703-723.

Greene, J. P. (2007). Introduction to the special issue on the single best idea for improving K-12 education. *Peabody Journal of Education*, *82*(4), 549-550.

Greene, J. P. (2007). The odd couple: Murray and Rothstein find some unexpected common ground. *Education Next*, *7*(4), 75-80.

Greene, J. P., & Winters, M. A. (2007). An evaluation of the effect of DC's voucher program on public school achievement and racial integration after one year. *Catholic Education: A Journal of Inquiry and Practice*, *11*(1), 83-101.

Greene, J. P., Kisida, B., & Butcher, J. (2007). Four score and seven manatees ago. *The City Journal*, *17*(2), Retrieved from http://www.city-journal.org

Carr, M., Greene, J. P., & Holley, M. (2007). Out-in-left-field trips. *The City Journal*, *17*(2), Retrieved from http://www.city-journal.org

Winters, M. A., & Greene, J. P. (2007). Debunking a special education myth. *Education Next*, *7*(2), 67-71.

Greene, J. P. (2006). Do funding systems create a perverse incentive to place students in special education? Yes. In M. A. Byrnes (Ed.), *Taking sides: Clashing views in special education* (3rd ed., pp. 71-82). Dubuque, IA: McGraw-Hill.

Greene, J. P. (2006). Milton Friedman, vouchers, and civic values. In R. C. Enlow & L. T. Ealy (Eds.), *Liberty and learning: Milton Friedman's voucher idea at fifty* (pp. 49-56). Washington, DC: CATO Institute.

Greene, J. P. (2005, July/August). Education myths: What special interest groups want you to believe about our schools--and why it isn't so. *The American Enterprise*, Retrieved from The Manhattan Institute for Policy Research website: http://www.manhattan-institute.org/educationmyths/

Greene, J. P., & Winters, M. A. (2006). The effect of residential school choice on public high school graduation rates. *Peabody Journal of Education*, *81*(1), 203-216.

Greene, J. P. (2006). A "comprehensive" problem. *Education Next*, *6*(1), 23-26.

Greene, J. P. (2005). Buckets into the sea: Why philanthropy isn't changing schools, and how it could. In F. M. Hess (Ed.), *The best of intentions* (pp. 49-76).

Greene, J. P. (2005). Choosing integration. In J. T. Scott (Ed.), *School choice and diversity: What the evidence says* (pp. 27-41). New York, NY: Teachers College Press.

Greene, J. P., & Forster, G. (2003-4, Winter). Public high school graduation and college readiness rates in the United States. *Texas Education Review*.

Greene, J. P. (2002, Summer). The GED myth. *Texas Education Review*.

Greene, J. P. (2002). The business model: Value added analysis is a crucial tool in the accountability toolbox despite its flaws. *Education Next*, *2*(2), 20-22.

Greene, J. P. (2002). GEDs aren't worth the paper they're printed on. *City Journal*, *12*(1), Retrieved from http://www.city-journal.org

Greene, J. P. (2001). The surprising consensus on school choice. *Public Interest*, *144*, 19-35.

Greene, J. P. (2001). Do students learn more where parents have more educational choices?. *The Journal of Private Enterprise*, *16*(2), 40-47.

Greene, J. P. (2000). The Texas school miracle is for real. *City Journal*, *10*(3), 24-30.

Greene, J. P., & Peterson, P. E. (2000). Should public discussion of political science research be controlled?: Why interest group recommendations on the proper procedures for reporting research should be treated with skepticism. *PS: Political Science & Politics*, *33*(2), 221-224.

Greene, J. P., & Peterson, P. E. (2000). If the peer review attack fails, attack something else. *PS: Political Science & Politics*, *33*(2), 229-231.

Greene, J. P., & Mellow, N. (2000). Integration where it counts: A study of racial integration in public and private school lunchrooms. *Texas Education Review*, *1*(1), 15-26.

Greene, J. P. (1999). Review of *Presidential-congressional relations: Policy and time approaches* by Steven A. Shull, *Political Science Quarterly*, *114*(1), 171-172.

Peterson, P. E., & Greene, J. P. (1998, Spring). Race relations and central city schools. *Brookings Review*, 33-37.

Greene, J. P. (1998). School choice: What does the evidence say?. In M. J. Breger & D. M. Gordis (Eds.), *Vouchers for school choice: Challenge or opportunity? An American Jewish reappraisal* (pp. 37-41). Brookline, MA: Wilstein Institute of Jewish Policy.

Peterson, P. E., Greene, J. P., & Noyes, C. (1996). School choice in Milwaukee. *Public Interest*, *125*, 38-56.

Greene, J. P. (1996). Review of *The Speaker*, edited by Ronald M. Peters. Jr. *Journal of Legislative Studies*, *2*(2), 131.

Greene, J. P. (1995). Term limits: A measure of our ignorance. *Social Science Quarterly*, *76*(4), 717-719.

Greene, J. P. (1993). Forecasting follies: Why political scientists can't predict presidential elections. *The American Prospect*, *4*(15).

Greene, J. P. (1993). Forewarned before forecast: Presidential election forecasting models and the 1992 election. *PS: Political Science & Politics*, *26*(1), 17-21

REPORTS and WORKING PAPERS

"Does Art Make You Smart? A Longitudinal Experiment of the Effects of Multiple Arts- Focused Field Trips," Holmes Erickson, Heidi and Greene, Jay and Watson, Angela and Beck, Molly (February 22, 2019). EDRE Working Paper No. 2019-05. Available at SSRN: https://ssrn.com/abstract=3340171 or http://dx.doi.org/10.2139/ssrn.3340171

"Altered Attitudes and Actions: Social-Emotional Effects of Multiple Arts Field Trips," Watson, Angela and Greene, Jay and Holmes Erickson, Heidi and Beck, Molly (February 19, 2019). EDRE Working Paper 2019-04. Available at SSRN: https://ssrn.com/abstract=3340163 or http://dx.doi.org/10.2139/ssrn.3340163

"The Play's the Thing: Experimentally Examining the Social and Cognitive Effects of School Field Trips to Live Theater Performances," with Holmes Erickson, Heidi and Watson, Angela and Beck, Molly, (August 31, 2017). EDRE Working Paper No. 2017-13. Available at SSRN: https://ssrn.com/abstract=3030928

"The effect of public and private schooling on anti-Semitism," with Cari A. Bogulski, American Enterprise Institute, November 2015.

"The Global Report Card," with Josh McGee, The George W. Bush Institute, September 2010.

"Administrative Bloat at American Universities: The Real Reason for High Costs in Higher Education," with Brian Kisida and Jonathan Mills, The Goldwater Institute, Policy Report No. 239, August 2010.

"The Milwaukee Parental Choice Program's Effect on School Integration," with Jonathan N. Mills, and Stuart Buck, The School Choice Demonstration Project, University of Arkansas, April 2010.

"Expanding Choice in Elementary and Secondary Education: A Report on Rethinking the Federal Role in Education," with W. Bentley MacLeod, Thomas Nechyba, Paul Peterson, Meredith Rosenthal, and Michelle Croft, The Brookings Institution, February 2, 2010.

"How Special Ed Vouchers Keep Kids From Being Mislabeled as Disabled," with Marcus Winters, Manhattan Institute, Civic Report No. 58, August 2009.

"The Effect of Milwaukee's Parental Choice Program on Student Achievement in Milwaukee Public Schools," with Ryan H. Marsh, The School Choice Demonstration Project, University of Arkansas, March 2009.

"The Impact of Performance Pay for Public School Teachers: Theory and Evidence," with Marcus A. Winters, Gary W. Ritter, Ryan H. Marsh, and Marc J. Holley, The Program on Education Policy and Governance, Harvard University, PEPG 08-15, December 2008.

"Building on the Basics: The Impact of High-Stakes Testing on Student Proficiency in Low-Stakes Subjects," with Marcus A. Winters and Julie Trivitt, Manhattan Institute, Civic Report No. 54, July 2008.

"Evaluating the Impact of Special Education Vouchers on Public Schools," with Marcus A. Winters, Manhattan Institute, Civic Report No. 52, April 2008.

"Year Two Evaluation of the Achievement Challenge Pilot Project in the Little Rock Public School District," with Gary Ritter, Marc Holley, Nate Jensen, Brent Riffel, Marcus Winters, and Joshua Barnett, University of Arkansas Department of Education Reform, January 22, 2008.

"Second Year Evaluation of the Systemic Effects of the DC Voucher Program," with Marcus A. Winters, *The School Choice Demonstration Project*, July 19, 2007.

"What's in a Name? The Decline in the Civic Mission of School Names," with Brian Kisida and Jonathan Butcher, Manhattan Institute, Civic Report No. 51, July 2007.

"How Much Are Public School Teachers Paid?," with Marcus A. Winters, Manhattan Institute Civic Report No. 50, February 2007.

"Getting Farther Ahead by Staying Behind: A Second-Year Evaluation of Florida's Policy to end Social Promotion," with Marcus Winters, Manhattan Institute, Civic Report No. 49, September 2006.

"Leaving Boys Behind: Public High School Graduation Rates," with Marcus A. Winters, Manhattan Institute, Civic Report No. 48, April 2006.

"The School Performance Index in Arkansas," with Joshua H. Barnett; Gary W. Ritter, Ph.D.; and Marcus A. Winters; Department of Education Reform, University of Arkansas, January 2006.

"An Evaluation of the Effect of D.C.'s Voucher Program on Public School Achievement and Racial Integration After One Year," with Marcus A. Winters, Education Working Paper 10, January 2006.

"Helping Kids, Saving Money: How to Reform New York's Special Education System," with Marcus A. Winters and Greg Forster, Empire Center Report, August 2005.

"The Effect of Residential School Choice on Public High School Graduation Rates," with Marcus A. Winters, Manhattan Institute, Education Working Paper Number 9, April 2005.

"Public High School Graduation and College-Readiness Rates: 1991-2002," with Marcus A. Winters, Manhattan Institute, Education Working Paper Number 8, February 2005.

"An Evaluation of Florida's Program to End Social Promotion," with Marcus A. Winters, Manhattan Institute, Education Working Paper Number 7, December 2004.

"Graduation Rates for Choice and Public School Students in Milwaukee," School Choice Wisconsin, September 28, 2004.

"The Teachability Index: Can Disadvantaged Students Learn?," with Greg Forster, Manhattan Institute, Education Working Paper Number 6, September 2004.

"Pushed Out or Pulled Up? Exit Exams and Dropout Rates in Public High Schools," with Marcus A. Winters, Manhattan Institute, Education Working Paper Number 5, April 2004.

"Sex, Drugs, and Delinquency in Urban and Suburban Public Schools," with Greg Forster, Manhattan Institute, Education Working Paper Number 4, January 2004.

"No Exit: How Lack of Educational Choice Shortchanges Hispanic Students," with Marcus A. Winters and Greg Forster, Hispanic Council for Reform and Educational Options, October 9, 2003.

"Public High School Graduation and College Readiness Rates in the United States," with Greg Forster, Manhattan Institute, Education Working Paper Number 3, September 2003.

"When Schools Compete: The Effects of Vouchers on Florida Public School Achievement," with Marcus A. Winters, Manhattan Institute, Education Working Paper Number 2, August 2003.

"Apples to Apples: An Evaluation of Charter Schools Serving General Population Students," with Greg Forster and Marcus A. Winters, Manhattan Institute, Education Working Paper Number 1, July 2003.

"Vouchers for Special Education Students: An Evaluation of Florida's McKay Scholarship Program," with Greg Forster, Manhattan Institute Civic Report 38, June 2003.

"This Works: Improving Urban Education," Civic Bulletin 34, March 2003.

"Testing High Stakes Tests: Can We Believe the Results of Accountability Tests?," with Marcus A. Winters and Greg Forster, Manhattan Institute Civic Report 33, February 2003.

"Effects of Funding Incentives on Special Education Enrollment," Manhattan Institute Civic Report 32, with Greg Forster, December 2002.

"Public School Graduation Rates in the United States," Civic Report 31, with Marcus Winters, Manhattan Institute Civic Report 31, November 2002.

"Rising to the Challenge: The Effect of School Choice on Public Schools in Milwaukee and San Antonio," Manhattan Institute Civic Bulletin 27, with Greg Forster, October 2002.

"High School Graduation Rates in Ohio," with Joshua Hall, The Buckeye Institute, September 2002.

"Graduation Rates in Washington State," Manhattan Institute Civic Report 27, August 2002.

"2001 Education Freedom Index," Manhattan Institute Civic Report 24, January 2002.

"High School Graduation Rates in the United States," Black Alliance for Educational Options and The Manhattan Institute for Policy Research.

"The CEO Horizon Scholarship Program: A Case Study of School Vouchers in the Edgewood Independent School District, San Antonio, Texas," Mathematica Policy Research, May 23, 2001.

"An Evaluation of the Florida A-Plus Accountability and School Choice Program," Florida State University, The Manhattan Institute, and The Harvard Program on Education Policy and Governance, February 2001.

"The Education Freedom Index," The Manhattan Institute Center for Civic Innovation, Civic Report Number 14, September 2000.

"The Cost of Remedial Education," The Mackinac Center for Public Policy, August 2000.

"The Effect of School Choice: An Evaluation of the Charlotte Children's Scholarship Fund," The Manhattan Institute Center for Civic Innovation, Civic Report Number 12, August 2000.

"A Survey of Results from Voucher Experiments: Where We Are and What We Know," The Manhattan Institute Center for Civic Innovation, Civic Report Number 11, July 2000.

"Community and Choice: The Racial, Economic, and Religious Context of Parental Choice in Cleveland," The Buckeye Institute, November, 1999.

"An Evaluation of the Cleveland Voucher Program After Two Years," with Paul Peterson and William Howell, Harvard Program on Education Policy and Governance Working Paper, June, 1999.

"Initial Findings from an Evaluation of School Choice Programs in Washington, D. C.," with Paul E. Peterson, William G. Howell, and William McCready," Occasional Paper, Program on Education Policy and Governance, Harvard University, September, 1998.

"Initial Findings from an Evaluation of School Choice Programs in Washington, D.C. and Dayton, Ohio," with Paul E. Peterson, William G. Howell, and William McCready," Occasional Paper, Program on Education Policy and Governance, Harvard University, October, 1998.

"A Meta-Analysis of the Effectiveness of Bilingual Education," Tomas Rivera Policy Institute, March 2, 1998.

"An Evaluation of the Cleveland Scholarship Program," with Paul E. Peterson and William G. Howell, Program on Education Policy and Governance, Harvard University, September, 1997.

"Methodological Issues in Evaluation Research: The Milwaukee School Choice Plan," with Paul E. Peterson, Occasional Paper, Program on Education Policy and Governance, Harvard University, August 29, 1996.

OP-EDS

"America's Students Flounder while Education Reformers Virtue Signal," with Frederick M. Hess, *National Review*, November 5, 2019.

"States have lousy records when they take over local school districts," *Houston Chronicle*, September 5, 2019.

"Democrats Have Taken Over Education Reform," *The Wall Street Journal*, March 7, 2019

"Vouching for Tolerance at Religious Schools," The Wall Street Journal, July 11, 2013.

"The Imaginary Teacher Shortage," The Wall Street Journal, October 8, 2012.

"Should All U.S. Students Meet a Single Set of National Proficiency Standards?," *Wall Street Journal*, June 22, 2012.

"Results Prove Charter Schools' Effectiveness," *Gainesville Times*, May 13, 2012.

"What Research Says about School Choice," with Kenneth Campbell, Paul Diperna, Robert C. Enlow, Greg Forster, Frederick M. Hess, Matthew Ladner, Michael J. Petrilli, and Patrick J. Wolf, *Education Week*, February 21, 2012.

"R.I. School Reform No Just Urban Issue," *Providence Journal*, January 22, 2012.

"Cobb, Fulton, and Forsyth Schools Lose Luster in Global Comparisons," *Atlanta Journal Constitution*, October 21, 2011.

"Suburban School Districts Falling Behind," *The Statesman*, October 16, 2011.

"Suburban Schools Post Low Global Grades," *Hartford Courant*, October 9, 2011.

"Suburban Schools Can Use Improvement, Too," *The Oklahoman*, October 9, 2011.

"Suburban Schools Fall Short Globally," *The Sacramento Bee*, October 7, 2011.

"Diane Ravitch's Credibility in Dispute," *National Review Online*, May 26, 2011.

"Small gains add up," with Stuart Buck, *New York Post*, November 9, 2010.

"The higher price for higher ed," with Brian Kisida and Jonathan Mills, *Austin American-Statesman*, September 5, 2010.

"Colleges feeding administrative bloat," with Brian Kisida and Jonathan Mills, *Atlanta Journal- Constitution*, August 25, 2010.

"Colleges' bloated bureaucracy," with Brian Kisida and Jonathan Mills, *The Baltimore Sun*, August 17, 2010.

"Indiana universities suffer from administrative bloat," with Brian Kisida and Jonathan Mills, *Indianapolis Star*, August 14, 2010.

"One size fits none," *Arkansas Democrat-Gazette*, April 11, 2010.

"Retaining students forces schools to act," *Indianapolis Star*, February 21, 2010.

"So Much for the Evidence," *City Journal*, February 3, 2010.

"Obama's education rhetoric doesn't come close to reality," *The Washington Examiner*, February 3, 2010.

"Head Start basically has no effect," *White Mountain Independent*, January 22, 2010.

"Where the Unemployment Isn't," with Stuart Buck, *National Review Online*, January 6, 2010.

"Funding may push special ed labeling," with Marcus A. Winters, *Atlanta Journal Constitution*, September 23, 2009.

"Special-Education Needs Help," *National Public Radio*, September 14, 2009.

"The Problems with Special Ed," *National Review Online*, September 14, 2009.

"Vouchers control special ed growth," with Marcus A. Winters, *The Tampa Tribune*, September 3, 2009.

"Special-education vouchers work for everybody," with Marcus A. Winters, *The Examiner*, August 19, 2009.

"Special-Education Stigmatization," with Marcus A. Winters, *Forbes*, August 18, 2009.

"Marcus A. Winters and Jay P. Greene: Special ed vouchers work for everybody," with Marcus A. Winters, *Washington Examiner*, August 18, 2009.

"The Union War on Charter Schools," *Wall Street Journal*, April 16, 2009.

"Vouchers: Not Dead Yet," *National Review Online*, April 8, 2009.

"Change that Won't Happen... Thankfully," *The Washington Times*, February 13, 2009.

"Change that Won't Happen... Thankfully," *National Review Online*, February 4, 2009.

"Schools Don't Need Snake Oil," *The Morning News*, February 2, 2009.

"A Special Plan for Palin," *National Review Online*, September 9, 2008.

"Consolidating School Districts Not Always Beneficial," *Arkansas Democrat-Gazette*, July 20, 2008.

"Pro-Con on Social Promotion: No holding back: All we need to do is follow the law," *Atlanta Journal-Constitution*, July 13, 2008.

"Yes, Virginia, There Is a Special Ed Bounty," with Greg Forster, *PajamasMedia*, July 8, 2008.

"The Special Education Epidemic," with Greg Forster, *PajamasMedia*, June 16, 2008.

"School Choice Helps Disabled Ariz. Students," with Marcus A. Winters, The East Valley Tribune, May 3, 2008.

"The Politics of Special-Ed Vouchers – Part III," with Marcus A. Winters, *The Washington Times*, May 1, 2008.

"A Special-Ed Fix," with Marcus A. Winters *The New York Post*, April 30, 2008.

"Vouchers for Special-Ed Students – Part II," with Marcus A. Winters, *The Washington Times*, April 30, 2008.

"Vouchers and Special Education," with Marcus A. Winters, *The Washington Times*, April 29, 2008.

"Reforming drop-out rates," with Marcus A. Winters, *The New York Sun*, April 8, 2008.

"Charter Schools Get Top Marks," *BusinessWeek*, March 17, 2008.

"Students Lose When Diversity Is Main Focus," with Catherine Shock, *Investors Business Daily*, January 8, 2008.

"Freston as Scapegoat," with Marcus A. Winters, *The New York Sun*, November 8, 2007.

"Union Days," *National Review Online*, September 12, 2007.

"Future raises for Oklahoma teachers should be tied to classroom performance," with Marcus A. Winters, *Muskogee Phoenix*, August 8, 2007.

"Merit Pay Could Boost Learning," with Marcus A. Winters, *The Oklahoman*, August 5, 2007.

"Steve Jobs Has Guts," *The New York Sun*, February 21, 2007.

"Teachers are better paid than you think," with Marcus A. Winters, *Providence Journal*, February 9, 2007.

"The myth of the underpaid teacher," with Marcus A. Winters, *The Philadelphia Daily News*, February 7, 2007.

"Is $34.06 Per Hour 'Underpaid?"' with Marcus Winters, The Wall Street Journal, February 2, 2007.

"Let Griffen Speak," with Brent Riffel, Arkansas Democrat-Gazette, January 25, 2007.

"San Antonio Demonstrates the Power of School Vouchers," with Greg Forster, *The Amarillo Globe News*, January 5, 2007.

"San Antonio's Proven that School Vouchers Work Well," with Greg Forster, *The Houston Chronicle*, December 25, 2006.

"Reason to rethink social promotion," with Marcus A. Winters, *The Atlanta Journal-Constitution*, September 30, 2006.

"Promoting the end of social promotion: Students should be held back if they can't pass the test, says a new study," with Marcus A. Winters, *Chicago Sun-Times*, September 17, 2006.

"School officials shouldn't go wobbly on social promotion," with Marcus A. Winters, *The Baltimore Examiner*, September 14, 2006.

"Social Promotion is Running Before Walking," with Marcus A. Winters, *The Tampa Tribune*, September 14, 2006.

"Advancement: Promoting the End of Social Promotion," with Marcus A. Winters, *National Review Online*, September 14, 2006.

"Separate Fact from Fiction in Education," *The Free Lance-Star*, August 27, 2006.

"Testing Takes Teachers to Task," with Marcus A. Winters, *CBS News*, July 31, 2006.

"'Losing Ground' in Education Reform," with Marcus A. Winters, *National Review Online*, July 27, 2006.

"Counting Diplomas and 9th-Graders," with Marcus A. Winters, *The Washington Post*, May 23, 2006.

"Try Altering Incentives for Teachers," with Marcus A. Winters, *The Philadelphia Inquirer*, May 4, 2006.

"Try Altering Incentives for Teachers," with Marcus A. Winters, *The Centre Daily*, May 4, 2006.

"There's a Dropout Crisis in Dallas Public Schools," *The Dallas Morning News*, May 3, 2006.

"Spending More on Education but Receiving Less," *The Arkansas Democrat-Gazette*, April 30, 2006.

"Educating from the Bench," *The Wall Street Journal*, April 27, 2006.

"Merit Pay: A Tool for Teaching," *Arkansas Democrat-Gazette*, April 23, 2006.

"The Dropout Crisis," with Marcus A. Winters, *The Baltimore Examiner*, April 19, 2006.

"Chicago Must Find a Way for Students to Finish High School," with Marcus A. Winters, *The Chicago Sun-Times*, April 19, 2006.

"The Boys Left Behind," with Marcus A. Winters, *National Review*, April 19, 2006.

"Making the Grade," *New York Daily News*, April 15, 2006.

"Stop the Train," *Arkansas Democrat-Gazette*, April 4, 2006.

"Missing the Mark on Graduation Rates," with Marcus A. Winters and Christopher B. Swanson, *Education Week*, March 29, 2006.

"The 65-Cent Delusion," with Jonathan Butcher, *National Review Online*, March 2, 2006.

"A New Yardstick for Education," with Gary Ritter, *Arkansas Democrat-Gazette*, January 15, 2006.

"Florida Voucher Ruling Shows Courts Go Too Far," with Marcus A. Winters, *The Philadelphia Inquirer*, January 13, 2006.

"Stupid Judge Tricks," with Marcus A. Winters, *National Review Online*, January 11, 2006.

"Benefits of Educational Options Clear," *Arizona Republic*, December 15, 2005.

"Five Myths…," with Marcus A. Winters, *National Review*, October 24, 2005.

"Ever Larger Spending Won't Improve Our Schools," *Arkansas Democrat-Gazette*, October 23, 2005.

"Another Option," with Greg Forster, *Baltimore Sun*, October 3, 2005.

"The Teacher Pay Myth," with Marcus A. Winters, *New York Post*, September 22, 2005.

"Education Myths: Why So Many?," with Marcus A. Winters, *National Review Online*, September 7, 2005.

"Tougher Standards Opening More Doors," with Marcus A. Winters, *Philadelphia Inquirer*, June 19, 2005.

"Tougher Standards Are Already Yielding Results," with Marcus A. Winters, *USA Today*, May 31, 2005.

"Drop the Skepticism: Cleveland's high school graduation rate is up," *Cleveland Plain Dealer*, March 29, 2005.

"Proficiency 101: Ending 'social promotion' could help close gaps," with Marcus A. Winters, *Philadelphia Inquirer*, February 28, 2005.

"Wasting the Windfall," *New York Post*, February 1, 2005.

"Judge: Children Must Fail," *New York Sun*, January 19, 2005.

"California Should End Social Promotion," with Marcus A. Winters, *San Diego Union Tribune*, December 9, 2004.

"Yanking Schools Back from Oz," with Marcus A. Winters, *New York Post*, December 8, 2004.

"Put to the Test, Unearned Passes Don't Help Kids," with Marcus A. Winters, *Chicago Sun-Times*, December 8, 2004.

"Ending Social Promotion Works," with Marcus A. Winters, *Vero Beach Press-Journal*, December 8, 2004.

"Push Forward Schools, Leave Behind Skeptics," *Newsday*, November 21, 2004.

"Choice Battle Needs a Real Truce," *Milwaukee Journal Sentinel*, November 20, 2004.

"Special Ed Inflates State School Costs," with Greg Forster, *Detroit News*, October 10, 2004.

"GEDs Just Don't Cut It," with Greg Forster, *New York Post*, October 1, 2004.

"A Long Term Choice: Vouchers Keep Kids in School," *National Review Online*, September 30, 2004.

"Another View of Arizona Students," with Greg Forster, *Arizona Republic*, September 14, 2004.

"Quit Blaming the Kids," with Greg Forster, *National Review Online*, September 13, 2004.

"Colorado Students Soaring Over Obstacles," with Greg Forster, *Rocky Mountain News*, September 13, 2004.

"Connecticut's Schools are Worse than They Look," with Greg Forster, *Hartford Courant*, September 9, 2004.

"Yardstick Lifts School Ranking," with Greg Forster, *Oklahoma City Oklahoman*, September 8, 2004.

"Wasting Away," *New York Sun*, September 8, 2004.

"Kerry Plan Can't Solve Problem that Keeps Kids Out of College," with Greg Forster, *Canton Repository*, August 31, 2004.

"Kerry's Plan Won't Put More Americans in College," with Greg Forster, *The New Hampshire Union Leader*, August 21, 2004.

"Kerry Plan Won't Aid Access to College," with Greg Forster, *Myrtle Beach Sun News*, August 21, 2004.

"Kerry's Plan Won't Help Poor Students," with Greg Forster, *Tallahassee Democrat*, August 20, 2004.

"No Comparison," *New York Sun*, August 19, 2004.

"Killing Opportunity," with Marcus A. Winters, *National Review Online*, August 19, 2004.

"Kerry's College Plan Fails Poor Kids," with Greg Forster, *Los Angeles Times*, August 18, 2004.

"Charters Ease Florida Public-School Crowding," with Marcus A. Winters, *Vero Beach Press-Journal*, August 6, 2004.

"Learning How Vouchers Help Disabled Children," with Greg Forster, *Trenton Times*, July 23, 2004.

"The Public Schools' Dirty Little Secret," with Marcus A. Winters, *New York Post*, July 2, 2004.

"Dollars in the Classroom," with Greg Forster, *World and I*, June 25, 2004.

"It's Elementary," with Marcus A. Winters, *National Review Online*, May 12, 2004.

"The Voucher Challenge," with Marcus A. Winters, *National Review Online*, May 11, 2004.

"The Positive Effects of Exit Exams," with Marcus A. Winters, *Indianapolis Star*, May 2, 2004.

"Exit Exams Don't Cause Increase in Dropout Rate," with Marcus A. Winters, *Arkansas Democrat-Gazette*, May 2, 2004.

"Surprise! The FCAT Doesn't Cause Dropouts," with Marcus A. Winters, *Tampa Tribune*, April 29, 2004.

"Testing Our Resolve," *New York Sun*, April 28, 2004.

"What D.C. Can Expect from Vouchers," with Marcus A. Winters *Washington Post*, March 7, 2004.

"Give Klein a Break," *New York Post*, February 27, 2004.

"High-Stakes Editorializing," with Marcus A. Winters, *National Review Online*, February 23, 2004.

"The Myth of the Wholesome Suburban High School," with Greg Forster, *Hartford Courant*, January 29, 2004.

"College diversity: Fix the K-12 pipeline first," with Greg Forster, *Salt Lake Tribune*, January 18, 2004.

"Fixing the Diversity Pipeline," with Greg Forster, *Daily Camera* (CO), January 18, 2004.

"Let vouchers help special-education students," with Greg Forster, Rocky Mountain News, January 11, 2004.

"College Diversity: Fix the Pipeline First," with Greg Forster, *Washington Post*, January 7, 2004.

"Support multiple school districts," with Marcus A. Winters, *Honolulu Advertiser*, December 23, 2003.

"Bad-Mouthing Texas," *National Review Online*, December 19, 2003.

"Vouchers Do Help Disabled Students," with Greg Forster, *Tampa Tribune*, December 11, 2003.

"Cooking the Graduation Numbers," with Greg Forster, *Los Angeles Times*, December 11, 2003.

"S.C. Schools Need Vouchers," with Marcus A. Winters, *Charleston Post and Courier*, December 8, 2003.

"Use Vouchers for Integration," *Hartford Courant*, November 30, 2003.

"Falling Behind West Virginia," with Greg Forster, *New York Post*, November 16, 2003.

"Smaller Classes Means Less Qualified Teachers," with Greg Forster, *National Post*, October 2, 2003.

"Indiana's Graduation Rate Doesn't Add Up," with Greg Forster, *Indianapolis Star*, September 21, 2003.

"Connecticut's Unbelievable Graduation Rate," with Greg Forster, *Hartford Courant*, September 21, 2003.

"Houston Isn't Alone When It Comes to Dropouts," with Greg Forster, *Houston Chronicle*, September 17, 2003

"New York Dropouts Uncounted," with Greg Forster, *New York Sun*, September 17, 2003.

"Vouchers Will Help Public Schools," with Marcus A. Winters, *Sacramento Bee*, September 1, 2003.

"Vouchers Will Help Public Schools," with Marcus A. Winters, *Raleigh News-Observer*, August 28, 2003.

"Vouchers Will Help Public Schools," with Marcus A. Winters, *Christian Science Monitor*, August 28, 2003.

"Apples to Apples on Charter Schools," *New York Sun*, July 17, 2003.

"Shorter School Day Hurts Teachers, Students," with Greg Forster, *South Florida Sun-Sentinel*, June 23, 2003.

"A New Model for Special Ed," with Greg Forster, *Boston Herald*, June 15, 2003.

"Small Classes: Union Scam," with Greg Forster, *New York Post*, June 12, 2003.

"Books and Bombers: Who Really Needs a Bake Sale?" *National Review Online*, May 22, 2003.

"An Unfair Grade for Vouchers," *Wall Street Journal*, May 16, 2003.

"Forcing the FCAT on Voucher Schools is a Bad Idea," with Marcus Winters, *Tallahassee Democrat*, March 31, 2003.

"Teachers Unions v. the Teachers," *New York Sun*, Feb.21, 2003.

"FCAT: Why teaching to the test isn't necessarily a bad thing" with Marcus A. Winters, *Tallahassee Democrat*, February 19, 2003.

"Does FCAT Pass the Test?" with Marcus A. Winters, *Tampa Tribune*, February 12, 2003.

"What Every Student Needs to Know: Does school testing make the grade?" with Marcus A. Winters, *San Francisco Chronicle*, February 11, 2003.

"Widespread Exploitation: How the teachers' unions take advantage of their own members," with Greg Forster, *National Review Online*, February 10, 2003.

"School Choice Is Good Move for All Students," with Greg Forster, *Austin American-Statesman*, January 14, 2003.

"Choice Helps Public Schools," *New York Sun*, January 14, 2003.

"Burning High-Stakes Testing at the Stake," with Greg Forster, *The Education Gadfly*, Volume 3, Number 1. January 9, 2003.

"The Money Threat to Special Education" *New York Sun*, December 16, 2002.

"Florida's miserable graduation rate," with Marcus A. Winters, *Tallahassee Democrat*, November 25, 2002.

"Choice proves beneficial for public schools, too," with Greg Forster, *Milwaukee Journal-Sentinel*, October 17, 2002.

"Raise Standards, Not Money" *Wall Street Journal*, October 7, 2002.

"Get more local control with smaller school districts," with Marcus A. Winters, *Sun Sentinel*, September 23, 2002.

"Choice Proves Beneficial for Public Schools, Too," with Greg Forster, *Milwaukee Journal-Sentinel*, October 23, 2002.

"Raise Standards, Not Money," *Wall Street Journal*, October 7, 2002.

"Get More Local Control with Smaller School Districts," with Marcus Winters, *South Florida Sun*

"Let's Not Pay More for Less," with Greg Forster, *Florida Times-Union*, September 3, 2002.

"The City's Special-Ed Crisis," *The New York Sun*, August 6, 2002.

"The Special Ed Hoax," *City Journal Online*, August 2, 2002.

"Choosing Integration," *The Wall Street Journal*, July 8, 2002.

"The Myth of the Special Education Burden," *Education Week*, June 12, 2002.

"Blaming Special Ed," *National Review Online*, May 23, 2001.

"The Myth of the Special Ed Burden," *Education Gadfly*, volume 2, number 19, May 9, 2002.

"Not-Quite High School: GED exams need toughening," *National Review Online*, April 24, 2002.

"GED Proves It's Not Equivalent to High School Degree," *Detroit News*, April 7, 2002.

"Why Vouchers are Constitutional," *City Journal On-Line*, March 14, 2002.

"The New *Brown*" *National Review Online*, February 21, 2002.

"School Choice = Higher Test Scores," *The Wall Street Journal*, January 23, 2002.

"Double Standard in Voucher Research," *The Education Gadfly*, volume 2, number 3, January 16, 2002.

"Graduation Statistics: Caveat Emptor," *Education Week*, January 16, 2002.

"Graduation Statistics: Caveat Emptor," *The Education Gadfly*, volume 1, number 27, November 29, 2001.

"Rising choice lifts all pupils," with Chester E. Finn, Jr., *The Christian Science Monitor*, October 5, 2000.

"Low-income students make academic gains: Study of four cities links scholarships to private schools with higher performance on standardized tests," *The Charlotte Observer*, September 22, 2000.

"Grading States on Their Education Freedom," with Chester E. Finn, Jr., *Wall Street Journal*, September 19, 2000.

"Fixing the N.Y.C. Schools," *The New York Post*, July 30, 2000.

"Why School Choice Can Promote Integration," *Education Week*, April 12, 2000, p. 72.

"AISD Arithmetic on Student Dropout Rates Requires Recalculation," *Austin American Statesman*, March 12, 1999, p. A15.

"Texans Deserve Honest Answers on Drop-Outs," *The Houston Chronicle*, October 22, 1998.

"Private Schools Better Promote Civic Values," *Intellectual Ammunition*, Volume 7, Number 4, September/October, 1998, p. 15.

"What Cleveland Teaches About School Vouchers," with Paul E. Peterson and William G. Howell, *The Wall Street Journal*, May 7, 1998, p. A22.

"Give Vouchers Time," with Paul E. Peterson, *The Washington Post*, May 1, 1998.

"Rescuing Education Research," *Education Week*, April 29, 1998, p. 52.

"Cleveland Shatters Myths about School Choice," with Paul E. Peterson and William G. Howell, *The Wall Street Journal*, September 18, 1997, p. A 14.

"All Tests Are Not Created Equal," *The Report Card*, Volume 2, Number 7, March/April, 1997, p. 12.

"Review of *The Case Against School Choice* by Edd Doerr et al," *The Report Card*, Volume 2, Number 6, January/February, 1997, p. 16.

"School Choice Data Rescued from Bad Science," with Paul E. Peterson, *The Wall Street Journal*, August 14, 1996, p. A14.

Appendix 1 – Quotations from Studies Listed in Jackson's Report Describing the Research Literature as "Mixed," "Inconclusive," or "Contradictory"

- Jackson, Johnson, Persico (2016): "Overall, the evidence on the effects of SFRs [school finance reforms] on academic outcomes is mixed, and the effects on long-run economic outcomes is unknown." (p. 160)

- Lafortune, Rothstein, and Schanzenbach (2018): "SFRs are arguably the most substantial national policy effort aimed at promoting equality of educational opportunity since the turn away from school desegregation in the 1980s. But there is little evidence about their effects on student achievement…. The literature regarding whether 'money matters' in education (Hanushek 1986, 2003, 2006; Card and Krueger 1992a; Burtless 1996) is contentious and does not offer clear guidance." (pp. 2-3)

- Cellini, Ferreira, and Rothstein (2010): "Despite the importance of capital spending, little is known about the overall impact of public infrastructure investment on economic output, and even less is known about the effects of school facilities investments…. Also closely related is the long literature on the effects of school spending more generally. Hanushek (1996) reviews more than ninety studies and concludes that '[s]imple resource policies hold little hope for improving student outcomes,' but Card and Krueger (1996) dispute Hanushek's interpretation of the literature…. Angrist and Lavy (2002) and Goolsbee and Guryan (2006) exploit credibly exogenous variation in school technology investments. Neither study finds shortrun effects on student achievement." (p. 216)

- Lafortune and Schonholzer (2018): "The empirical literature on capital expenditures offers little guidance with regard to these questions. Several studies find no or imprecise effects of capital expenditures on student achievement (see Cellini et al., 2010; Bowers and Urick, 2011; Goncalves, 2015; Martorell et al, 2016), while others find some evidence of positive impacts on student achievement, often only in reading and English-language arts (Welsh et al., 2012; Neilson and Zimmerman, 2014; Hong and Zimmer, 2016; Conlin and Thompson, 2017; Hashim et al., 2018). Other studies have looked at longer-run impacts of school construction programs in other countries that expand access to education (e.g. Duflo, 2001, 2004), measuring the effects of more general increases in human capital accumulation. Despite inconclusive evidence in the literature and general skepticism among economists, resource-based capital expenditure programs continue to be used by policymakers at the state and local level as tools to improve schools and reduce achievement gaps." (pp. 1-2)

- Lee and Polachek (2018): “Current analyses find contradictory evidence of the effect of school expenditures on dropout and graduation rates.” (p. 131)

- Rauscher (2019): “Debates about the efficiency of education funding for student achievement have continued at least since the 1966 Coleman Report (e.g., Hanushek 1989, 1996; Burtless 1996; Greenwald et al. 1996; see Biddle and Berliner 2002 and Baker 2016 for reviews), including contemporary evidence of no relationship between funding and achievement (Morgan and Jung 2016)…. Existing research provides contradictory evidence about the effects of education funding on student achievement (e.g., Jackson, Johnson and Persico 2016; Morgan and Jung 2016).” (pp. 1, 27)

- Johnson (2015): “Despite its fiscal importance, evidence on the effectiveness of Title I is mixed (Matsudaira, Hosek, and Walsh 2012; Cascio, Gordon, and Reber 2013; Van der Klaauw 2008).” (p. 50)

- Papke (2008): “Yinger (2004) discusses education finance litigation and resulting reforms to state finance systems. He concludes that, while some of the evidence indicates that state aid can boost student performance, none of the findings is definitive and some are quite ambiguous.” (p. 466)

- Hyman (2017): “However, it is less clear whether the changes in spending affected student achievement, with some studies finding positive effects and others finding no effects.”

- Conlin and Thompson (2017): “Recent literature has focused on using quasi-experimental designs to identify the causal effect of capital investment on student outcomes and housing prices. A set of quasi-experimental papers (Cellini et al., 2010; Hong & Zimmer, 2016; Kogan, Lavertu, & Peskowitz, 2017; Martorell, Stange, & McFarlin, 2016) estimate regression discontinuity designs using the majority rule cutoff in school bond referendum elections to compare outcomes (test scores and/or housing prices) for districts that just pass a bond referendum to fund additional capital expenditures to those that just fail to pass a bond referendum and generally find mixed evidence on the role of capital investments on student achievement.” (p. 14)

Appendix 2 -- Exhibits Presented in this Report

Exhibit 1: Jackson's Classification of Studies in Expert Report (2020)

Table 1: *List of Studies Meeting the Inclusion Criteria*

Study	pos	pos. & sig.	Outcome	Strategy	$ type	Low-I
Multi-State Studies						
Abott Kogan Lavertu Peskowitz, 2019	Yes	Yes	Test Scores, Grad. Rates	Regression Discontinuity	Operational	No
Biasi, 2019	Yes	Yes	Enroll Postsecondary, Income Mobility	Instrumental Variables	Any	Yes
Brunner Hyman Ju, 2019	Yes	No	Test Scores	Instrumental Variables	Any	n/a
Candelaria Shores, 2019	Yes	No	Grad. Rates	Event-Study DiD	Any	Yes
Card Payne, 2002	Yes	Yes	Test Scores	CO-SFR	Any	Yes
Cascio Gordon Reber, 2013	Yes	No	Dropout Rates	Instrumental Variables	Title I	n/a
Jackson Johnson Persico 2015, Johnson Jackson, 2019	Yes	Yes	Grad. Rates, Years of Ed., Wages	Event-Study DiD, Instrumental Variables	Any	Yes
Jackson Wigger Xiong, 2020	Yes	Yes	Test Scores, Enroll Postsecondary	Instrumental Variables	Any	Yes
Johnson, 2015	Yes	Yes	Grad. Rates, Other Ed. Outcomes, Wages	Event-Study DiD	Title I	Yes
Lafortune Rothstein Schanzenbach, 2018	Yes	No	Test Scores	Event-Study DiD	Any	Yes
Miller, 2018	Yes	Yes	Test Scores, Grad. Rates	Instrumental Variables	Any	n/a
Non-Multi-State Studies						
Baron, 2019	Yes	Yes	Test Scores	Regression Discontinuity	Operational	n/a
Carlson Lavertu, 2018	Yes	No	Test Scores	Regression Discontinuity	Any	n/a
Cellini Ferreira Rothstein, 2010	Yes	Yes	Test Scores	Regression Discontinuity	Capital	n/a
Clark, 2003	Yes	No	Test Scores	Event-Study DiD	Any	n/a
Conlin Thompson, 2017	Yes	No	Test Scores	Instrumental Variables	Capital	n/a
Gigliotti Sorensen, 2018	Yes	Yes	Test Scores	Instrumental Variables	Any	n/a
Goncalves, 2015	Yes	No	Test Scores	Event-Study DiD	Capital	Yes
Guryan, 2001	Yes	Yes	Test Scores	Instrumental Variables	Any	n/a
Holden, 2016	Yes	No	Test Scores	Regression Discontinuity	Operational	n/a
Hong Zimmer, 2016	Yes	No	Test Scores	Regression Discontinuity	Capital	n/a
Hyman, 2017	Yes	Yes	Enroll Postsecondary	Instrumental Variables	Any	No
Kogan Lavertu Peskowitz, 2017	Yes	Yes	Test Scores	Regression Discontinuity	Any	n/a
Kreisman Steinberg, 2019	Yes	Yes	Test Scores, Grad. Rates, Enroll Postsecondary	Instrumental Variables	Any	Yes
Lafortune Schonholzer, 2018	Yes	Yes	Test Scores	Event-Study DiD	capital	n/a
Lee Polachek, 2018	Yes	Yes	Dropout Rates	Regression Discontinuity	Any	n/a
Martorell Stange McFarlin, 2016	Yes	No	Test Scores	Regression Discontinuity	Capital	n/a
Matsudaira Hosek Walsh, 2012	Yes	No	Test Scores	Regression Discontinuity	Title I	No
Neilson Zimmerman, 2014	Yes	No	Test Scores	Event-Study DiD	Capital	n/a
Papke, 2008	Yes	Yes	Test Scores	Instrumental Variables	Any	n/a
Rauscher, 2019	Yes	No	Test Scores	Regression Discontinuity	Capital	Yes
Roy, 2011	Yes	Yes	Test Scores	Instrumental Variables	Any	Yes
Weinstein Stiefel Schwartz Chalico, 2009	Yes	No	Grad. Rates	Regression Discontinuity	Title I	n/a

Exhibit 2: Jackson's Classification of Studies in 2018 Review

Study	Pos and Sig	Neg and Sig	Not Sig	Outcomes	Variation	State	Type of Spending
Multi-State Studies							
Jackson Johnson and Persico (2015)	Y			Education, Wages	CO-SFR	ALL	Any
Johnson and Jackson (2018)	Y			Education, Wages, other	CO-SFR	ALL	Any
Lafortune, Rothstien, Shanzenbach (2018)	Y			Test Scores	SFR	ALL	Any
Candelaria and Shores (2018)	Y			Graduation Rates	CO-SFR	ALL	Any
Brunner, Hyman, and Ju (2018)	Y			Test Scores	SFR	ALL	Any
Biasi (2018)	Y			Income mobility	SFR	ALL	Any
Card and Payne (2002)	Y			SAT score inequality	CO-SFR	ALL	Any
Hoxby (2001)			Y	Dropout rates	SFR	ALL	Any
Downes and Figlio (1997)	Y			Test Scores	DiD-Tax Limit	ALL	Any
Johnson (2015)	Y			Graduation Rates	DiD	ALL	Title I
Jackson, Wigger and Xiong (2018)	Y			Test Scores	IV-Recession	ALL	Any
Miller (2018)	Y			Test Scores	IV-House Values	ALL	Any
Cascio, Gordon, and Reber (2013)	Y			Dropout	Event Study	SOUTH	Title I
Single-State Studies							
Hyman (2017)	Y			College-going	Rules-Based IV	MI	Any
Gigliotti (2018)	Y			Test Scores	Rules-Based IV	NY	Any
Papke (2008)	Y			Test Scores (pass rates)	IV-SFR	MI	Any
Roy (2011)	Y			Test Scores	Policy (SFR)	MI	Any
Guryan (2001)	Y			Test Scores	Rules-Based IV	MA	Any
Clark (2003)			Y	Test Scores	Rules-Based IV	KY	Any
Lee and Polachek (2018)	Y			Graduation Rates	RD-Referenda	NY	Any
Holden (2016)	Y			Test scores	RD	CA	Textbooks
Husted and Kenny (2000)	Y			SAT Scores	DiD	CA	Any
Cellini, Ferreira, and Rothstein (2010)			Y	Test Scores	RD-Bonds	CA	Capital
Lafortune and Schonholzer (2018)	Y			Test Scores	Event-Study	CA	Capital
Martorell, Stang, McFarlin (2016)			Y	Test Scores	RD-Bonds	TX	Capital
Conlin and Thompson (2017)	Y			Test Scores	IV	OH	Construction
Goncalves (2015)			Y	Test Scores	Event-Study	OH	Construction
Hong and Zimmer (2016)	Y			Test Scores	RD-Bonds	MI	Capital
Kogan, Lavertu, and Peskowitz (2017)	Y			Test Scores	RD-Referenda	OH	Any
Zimmerman and Neilson (2014)	Y			Test Scores	Event-study	New Haven, CT	Construction
Van Der Klaue (2008)			Y	Test Scores	RD Title I	NY City	Title-I
Matsudaira, Hosek, Walsh (2012)			Y	Test Scores	RD-Title I	NYC*	Title-I
Weinstein, M. G., Stiefel, et al (2009)			Y	Test Scores	RD-Title I	NYC*	Title-I

Exhibit 3: Figures in Jackson's Expert Report of Results by Study with 95% Confidence Intervals

Figure 2: Estimated School Spending Effect on Test Scores (Operational Spending)

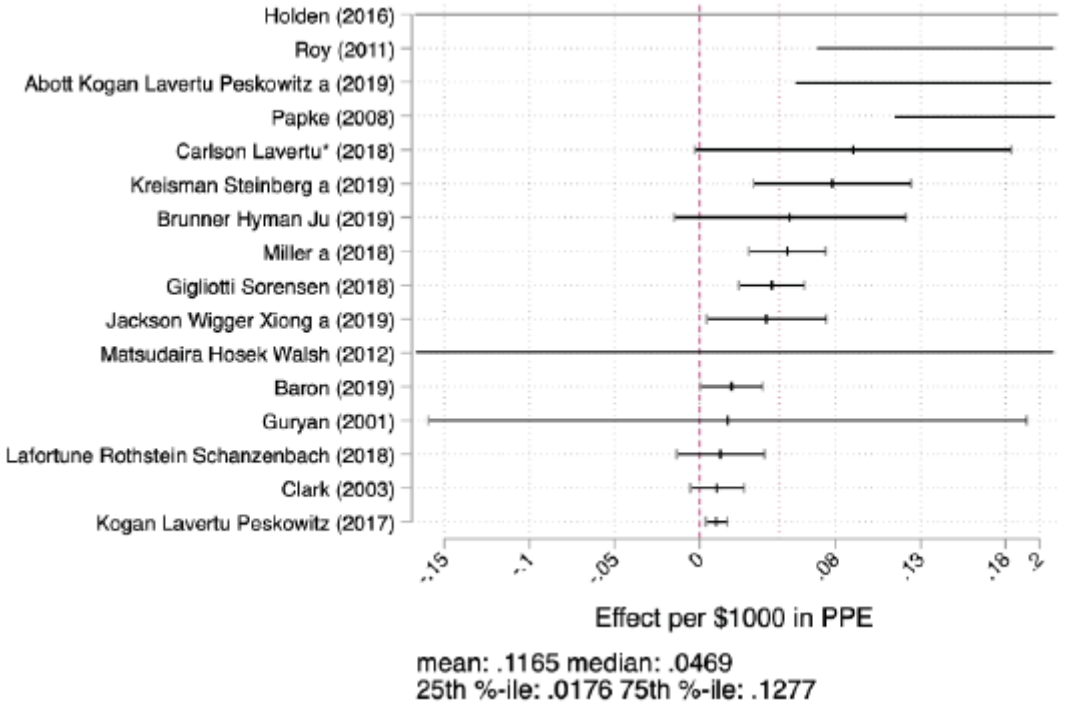


Figure 6: *Estimated School Spending Effects on Test Scores (capital and operational spending)*

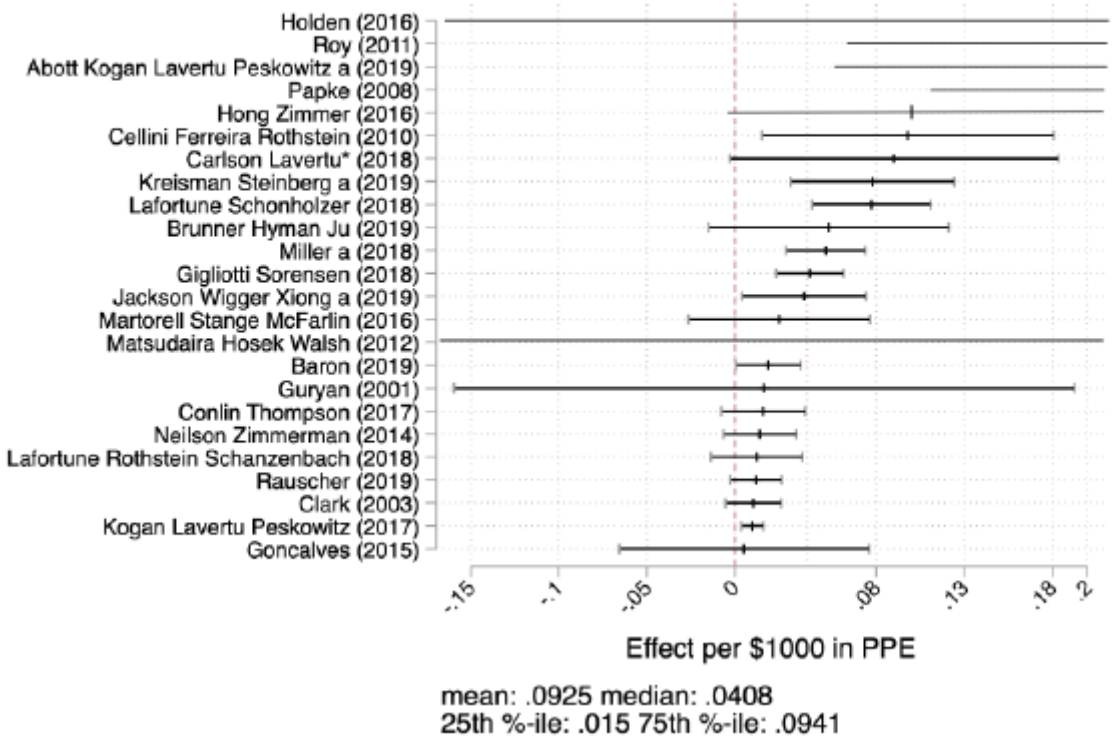


Figure 7: *Estimated School Spending Effects on Educational Attainment*

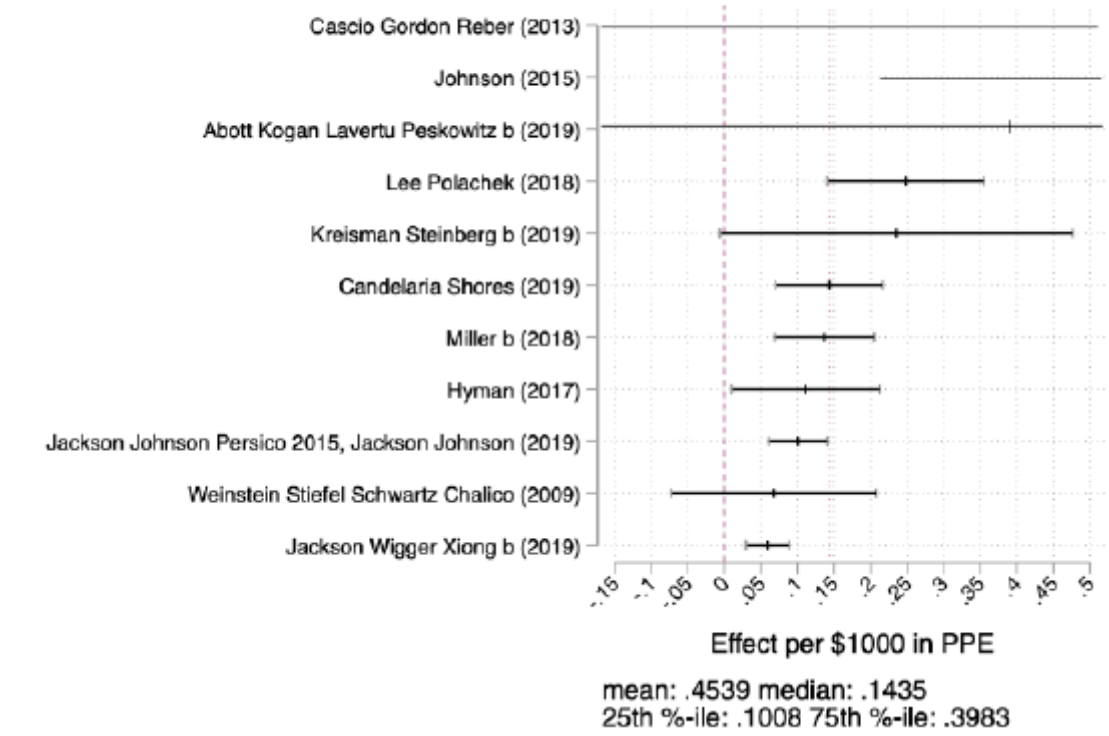


Exhibit 4: Figure in Jackson, Wigger, and Xiong (2018) of Results by Study with 95% Confidence Intervals

Figure 5. *Forrest Plot of Existing Studies*

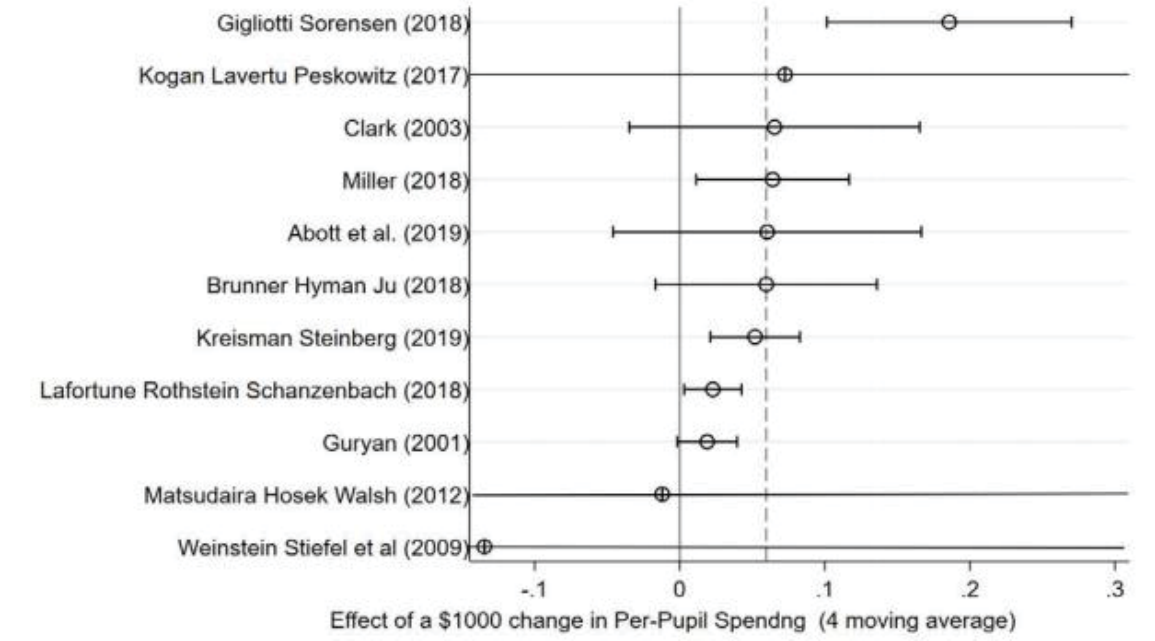


Exhibit 5: Results from Kogan, et al (2017)

Table 6. Impact of Tax Levy Failure on Student Achievement

	State "Value Added" Estimate (District SDs)				State Performance Index (District SDs)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2 years prior	−0.102	−0.00960	−0.144	−0.00754	−0.00593	−0.0127	−0.0118	−0.00703
	(0.129)	(0.0909)	(0.167)	(0.0577)	(0.0217)	(0.0174)	(0.0230)	(0.0121)
1 year prior	—	—	—	—	—	—	—	—
Election year	−0.0926	−0.0242	−0.155	−0.0429	0.0295	−0.00163	0.0271	−0.0141
	(0.113)	(0.0831)	(0.140)	(0.0552)	(0.0197)	(0.0150)	(0.0212)	(0.00998)
1 year after	−0.0894	−0.0342	−0.0666	−0.0486	−0.0187	−0.0351*	−0.0299	−0.0356**
	(0.117)	(0.0873)	(0.147)	(0.0548)	(0.0220)	(0.0167)	(0.0240)	(0.0123)
2 years after	−0.199^	−0.124	−0.178	−0.0913^	−0.0198	−0.0419*	−0.0251	−0.0312*
	(0.103)	(0.0782)	(0.146)	(0.0524)	(0.0245)	(0.0190)	(0.0265)	(0.0138)
3 years after	−0.179^	−0.126	−0.168	−0.00385	−0.0274	−0.0630**	−0.0317	−0.0281^
	(0.108)	(0.0826)	(0.148)	(0.0551)	(0.0273)	(0.0204)	(0.0290)	(0.0145)
4 years after	−0.0195	0.0722	0.0678	0.0115	0.0123	−0.0281	−0.0160	−0.0181
	(0.108)	(0.0868)	(0.154)	(0.0585)	(0.0296)	(0.0220)	(0.0309)	(0.0156)
5 years after	−0.120	−0.0701	−0.0344	−0.00224	−0.00610	−0.0394	−0.0364	−0.0148
	(0.117)	(0.0955)	(0.156)	(0.0614)	(0.0326)	(0.0246)	(0.0341)	(0.0173)
6 years after	−0.150	0.00258	−0.0456	−0.0312	−0.00776	−0.0213	−0.0339	−0.0126
	(0.123)	(0.0975)	(0.156)	(0.0667)	(0.0354)	(0.0270)	(0.0363)	(0.0194)
N	24,796	24,796	10,936	24,796	33,199	33,199	21,660	33,199
District count	571	571	509	571	571	571	541	571
Levy count	4,324	4,324	1,916	4,324	4,324	4,324	2,812	4,324
Mean dependent variable	0.03	0.03	0.03	0.03	0.09	0.09	0.06	0.09
Model	RD	RD	RD	Differences-in-Differences	RD	RD	RD	Differences-in-Differences
Specification	Quad.	Linear	Linear	N/A	Quad.	Linear	Linear	N/A
Restricted bandwidth	No	No	Yes	N/A	No	No	Yes	N/A
Levy type	Op. & Cap.	Op. & Cap.	Op. & Cap.	Op. & Cap.	Op. & Cap.	Op. & Cap.	Op. & Cap.	Op. & Cap.

Note: The results above are from models estimating the impact of levy failure (as opposed to passage) on district performance measures standardized by year. SEs clustered by district are presented in parentheses below the estimated coefficients. *p*-values were calculated using a two-tailed test.

^$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

Exhibit 6: Results from Weinstein, et al (2009)

Table 7: Academic Outcomes, RD Estimates, Cubic Polynomial, Elementary and Middle Schools 1997-2003

	Math		Reading	
	(1)	(2)	(3)	(4)
Title I	-0.035*	-0.011	-0.037*	-0.031*
	(0.020)	(0.017)	(0.019)	(0.016)
%Poverty	-0.005**	-0.002	-0.002	-0.000
	(0.002)	(0.003)	(1.979)	(0.002)
%PovertySq	-0.000	-0.000	0.000*	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
%Poverty Cub	-1.37	-4.85	8.86***	-4.72
	(9.74)	(1.27)	(9.31)	(1.23)
TitleI*%Poverty	-0.004	0.005	-0.012***	0.000
	(0.004)	(0.004)	(0.004)	(0.003)
TitleI*%PovertySq	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
TitleI*%PovertyCub	1.56	5.76*	-3.65	3.26
	(3.31)	(3.07)	(3.12)	(2.89)
Total Registration	0.000***	-4.29	0.000	-7.35
	(7.48)	(0.000)	(7.27)	(0.000)
%ELL	-0.008***	-0.007***	-0.007***	-0.004***
	(0.000)	(0.001)	(0.000)	(0.001)
%Recent Immigrant	-0.005***	-0.000	0.007***	-0.002
	(0.000)	(0.001)	(0.001)	(0.001)
%Special Ed	-0.001***	0.000	-0.001**	0.001
	(0.001)	(0.001)	(0.000)	(0.001)
% black	-0.008***	-0.010***	-0.006***	-0.006***
	(0.000)	(0.001)	(0.000)	(0.001)
% Hispanic	-0.006***	-0.010***	-0.004***	-0.008***
	(0.000)	(0.001)	(0.000)	(0.001)
% Asian	0.004***	-0.001	0.001***	0.002***
	(0.000)	(0.001)	(0.000)	(0.001)
% female	0.018***	0.003***	0.019***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Constant	-0.237***	0.639***	-0.434***	0.270**
	(0.094)	(0.109)	(0.089)	(0.115)
Borough Effects	Yes	No	Yes	No
School Effects	No	Yes	No	Yes
N Schools	776	776	776	776
Observations	5432	5432	5432	5432
R-squared	0.66	0.83	0.63	0.81

(i) Robust standard errors in parentheses
(ii) * significant at 5%; ** significant at 1%
(iii)Year, middle school, and grade dummies are included but not shown.

Exhibit 7: Results for Gonclaves (2015)

	(1) Math	(2) Reading	(3) Math	(4) Reading	(5) Math	(6) Reading
1 yr. Construction Exposure	-0.527	-0.651*	-0.525	-0.878	-0.698	-1.426**
	(0.376)	(0.322)	(0.541)	(0.489)	(0.620)	(0.518)
2 yr. Construction Exposure	-1.977***	-1.308**	-2.620***	-2.115*	-3.460***	-2.767***
	(0.442)	(0.412)	(0.737)	(0.835)	(0.727)	(0.697)
3 yr. Construction Exposure	-2.501***	-1.680**	-3.723***	-3.364**	-3.438***	-3.346***
	(0.596)	(0.566)	(0.979)	(1.150)	(0.862)	(0.872)
4 yr. Construction Exposure	-2.553**	-1.720*	-4.273**	-3.143*	-4.610**	-4.089***
	(0.945)	(0.765)	(1.437)	(1.454)	(1.383)	(1.070)
5 yr. Construction Exposure	-2.399*	-0.994	-4.001	-2.043	-4.280**	-3.630**
	(1.010)	(0.854)	(2.104)	(1.832)	(1.352)	(1.182)
6+ yr. Construction Exposure	-2.269	-1.076	-7.408	-4.123	-4.260	-4.682*
	(2.465)	(1.975)	(5.011)	(5.054)	(2.586)	(2.283)
1 yr. Completion Exposure	0.00874	0.129	0.689	0.120	0.635	0.522
	(0.587)	(0.254)	(0.589)	(0.484)	(0.687)	(0.324)
2 yr. Completion Exposure	-0.377	0.0665	-0.0853	-0.273	0.834	0.511
	(0.630)	(0.441)	(1.071)	(0.804)	(0.884)	(0.676)
3 yr. Completion Exposure	-0.135	-0.310	0.508	-0.772	1.386	0.0463
	(0.811)	(0.704)	(1.289)	(1.118)	(1.261)	(1.044)
4 yr. Completion Exposure	0.188	-0.186	0.652	-0.827	2.024	0.225
	(1.013)	(0.856)	(1.532)	(1.300)	(1.537)	(1.245)
5 yr. Completion Exposure	0.386	-0.123	0.948	-0.568	2.688	-0.169
	(1.144)	(1.010)	(1.739)	(1.584)	(1.754)	(1.511)
6+ yr. Completion Exposure	1.266	-1.442	2.216	-1.931	4.573*	-1.107
	(1.347)	(1.072)	(2.004)	(1.669)	(2.066)	(1.558)
N	14569	14570	7090	7091	5720	5720
R2	0.805	0.793	0.799	0.782	0.813	0.802
Avg. Proficiency	80.122	80.122	80.006	80.006	74.913	74.913
Districts	All	All	Poorest 25%	Poorest 25%	Lunch > 25%	Lunch > 25%

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Exhibit 8 -- Discrepancies And Differences in Classification of Studies

	Jackson (2018)				Jackson, Wigger, & Xiong (2018)			Jackson Report			Greene Re-Analysis			
Study	Pos. & Sig.	Not Sig.	Neg. & Sig.	Meets Criteria for Inclusion	Pos. & Sig.	Pos. & Not Sig.	Neg & Not Sig.	Pos. & Sig.	Pos. & Not Sig.	Meets Criteria for Inclusion	Pos. & Sig.	Not Sig.	Neg. & Sig.	Meets Criteria for Inclusion
Abott, Korgan, Lavertu, & Peskowitz (2019)				NA		Yes		Yes		Yes	Yes			Yes
Biasi (2019)	Yes			Yes				Yes		Yes	Yes			No
Brunner, Hyman, & Ju (2019)	Yes			Yes		Yes			Yes	Yes	Yes			No
Candelaria & Shores (2019)	Yes			Yes					Yes	Yes	Yes			Yes
Card & Payne (2002)	Yes			Yes				Yes		Yes	Yes			Yes
Cascio, Gordon, & Reber (2013)	Yes			Yes					Yes	Yes		Yes		Yes
Downes & Figio (1997)	Yes			Yes						No		Yes		Yes
Hoxby (2001)		Yes		Yes						No		Yes		Yes
Jackson, Johnson, & Persico (2016) Johnson & Jackson (2019)	Yes			Yes				Yes		Yes	Yes			Yes
Jackson, Wigger, & Xiang (2018)	Yes			Yes				Yes		Yes	Yes			Yes
Johnson (2015)	Yes			Yes				Yes		Yes	Yes			Yes
LaFortune, Rothstein, & Schanzenbach (2018)	Yes			Yes	Yes				Yes	Yes		Yes		No
Miller (2018)	Yes			Yes	Yes			Yes		Yes	Yes			No
Baron (2019)				NA				Yes		Yes	Yes			No
Carlson & Lavertu (2018)				NA					Yes	Yes	Yes			No
Cellini, Ferreira, & Rothstein (2010)		Yes		Yes				Yes		Yes		Yes		No
Clark (2003)		Yes		Yes		Yes			Yes	Yes		Yes		Yes
Conlin & Thompson (2017)	Yes			Yes					Yes	Yes		Yes		Yes
Gigliotti & Sorensen (2018)	Yes			Yes	Yes			Yes		Yes	Yes			No
Goncalves (2015)		Yes		Yes					Yes	Yes			Yes	Yes
Guryan (2001)	Yes			Yes		Yes		Yes		Yes	Yes			Yes
Holden (2016)	Yes			Yes					Yes	Yes		Yes		Yes
Hong & Zimmer (2016)	Yes			Yes					Yes	Yes		Yes		Yes
Husted & Kenny (2000)	Yes			Yes						No		Yes		No
Hyman (2017)	Yes			Yes				Yes		Yes	Yes			Yes
Kogan, Lavertu, & Peskowitz (2017)	Yes			Yes		Yes		Yes		Yes		Yes		No
Kreisman & Steinberg (2019)				NA	Yes			Yes		Yes	Yes			No
LaFortune & Schonholzer (2018)	Yes			Yes				Yes		Yes	Yes			Yes
Lee & Polachek (2018)	Yes			Yes				Yes		Yes	Yes			Yes
Lee & Polachek (2014)				No						No		Yes		NA
Martorell, Stange, & McFarlin (2016)		Yes		Yes					Yes	Yes		Yes		No
Matsudaira, Hosek, & Walsh (2012)		Yes		Yes			Yes		Yes	Yes		Yes		Yes
Neilson & Zimmerman (2014)	Yes			Yes					Yes	Yes	Yes			Yes
Papke (2008)	Yes			Yes				Yes		Yes	Yes			Yes
Rauscher (2019)				NA					Yes	Yes		Yes		No
Roy (2011)	Yes			Yes				Yes		Yes	Yes			Yes
Van der Klaauw (2008)		Yes		Yes						No			Yes	Yes
Weinstein, Stiefel, Schwartz, & Chalico (2009)		Yes		Yes			Yes		Yes	Yes			Yes	Yes
								4	10	4	4	6	3	15
								Total	18		Total	28		
						Discrepancies from Jackson (2018)					Differences from Jackson Report			
						Discrepancies from Jackson, Wigger, & Xiong (2018)								

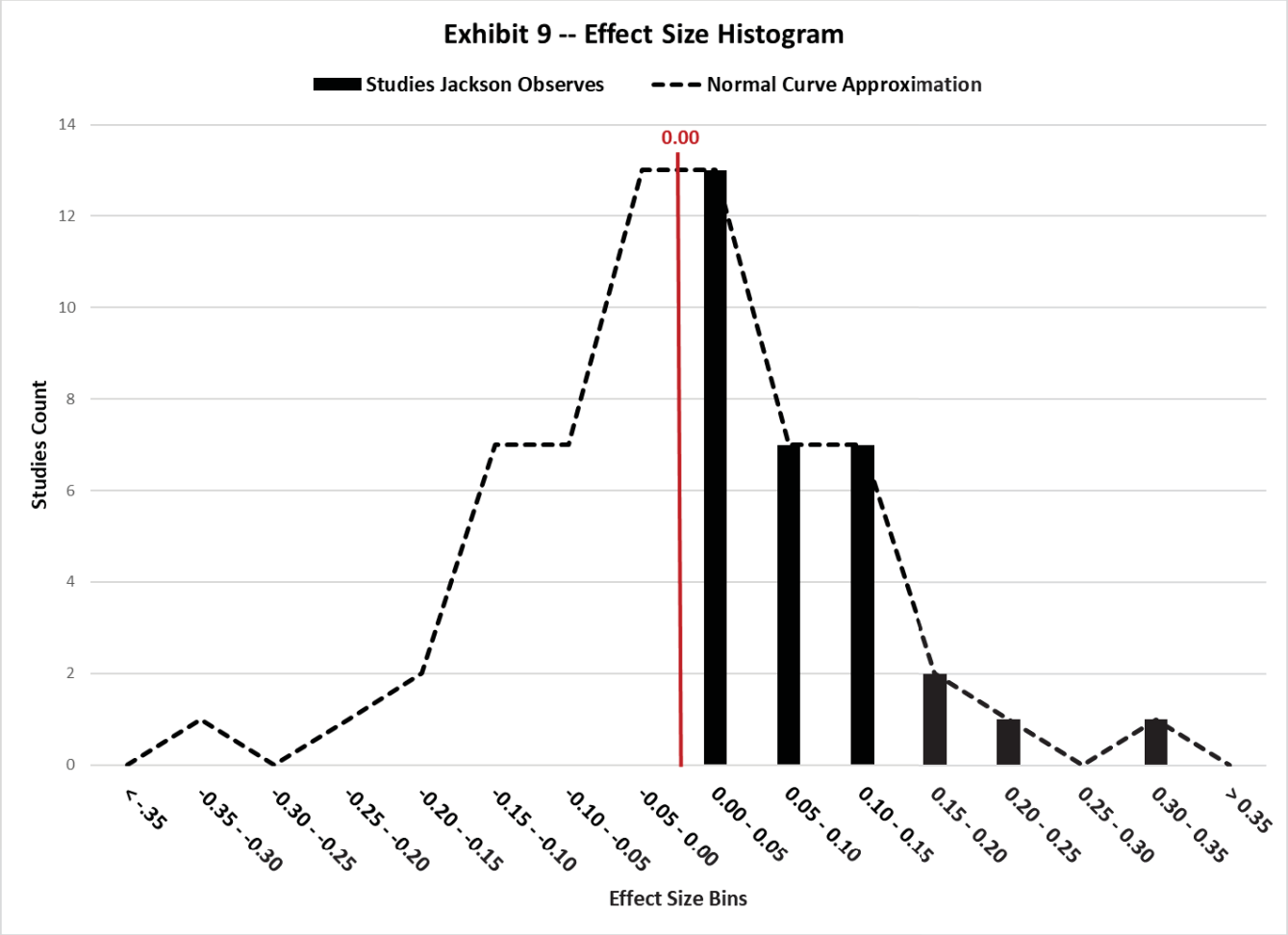


Exhibit 9 -- Effect Size Histogram
Studies Jackson Observes
Normal Curve Approximation
Studies Count
14
12
10
8
6
4
2
0
0.00
< -.35
-0.35 - -0.30
-0.30 - -0.25
-0.25 - -0.20
-0.20 - -0.15
-0.15 - -0.10
-0.10 - -0.05
-0.05 - 0.00
0.00 - 0.05
0.05 - 0.10
0.10 - 0.15
0.15 - 0.20
0.20 - 0.25
0.25 - 0.30
0.30 - 0.35
> 0.35
Effect Size Bins

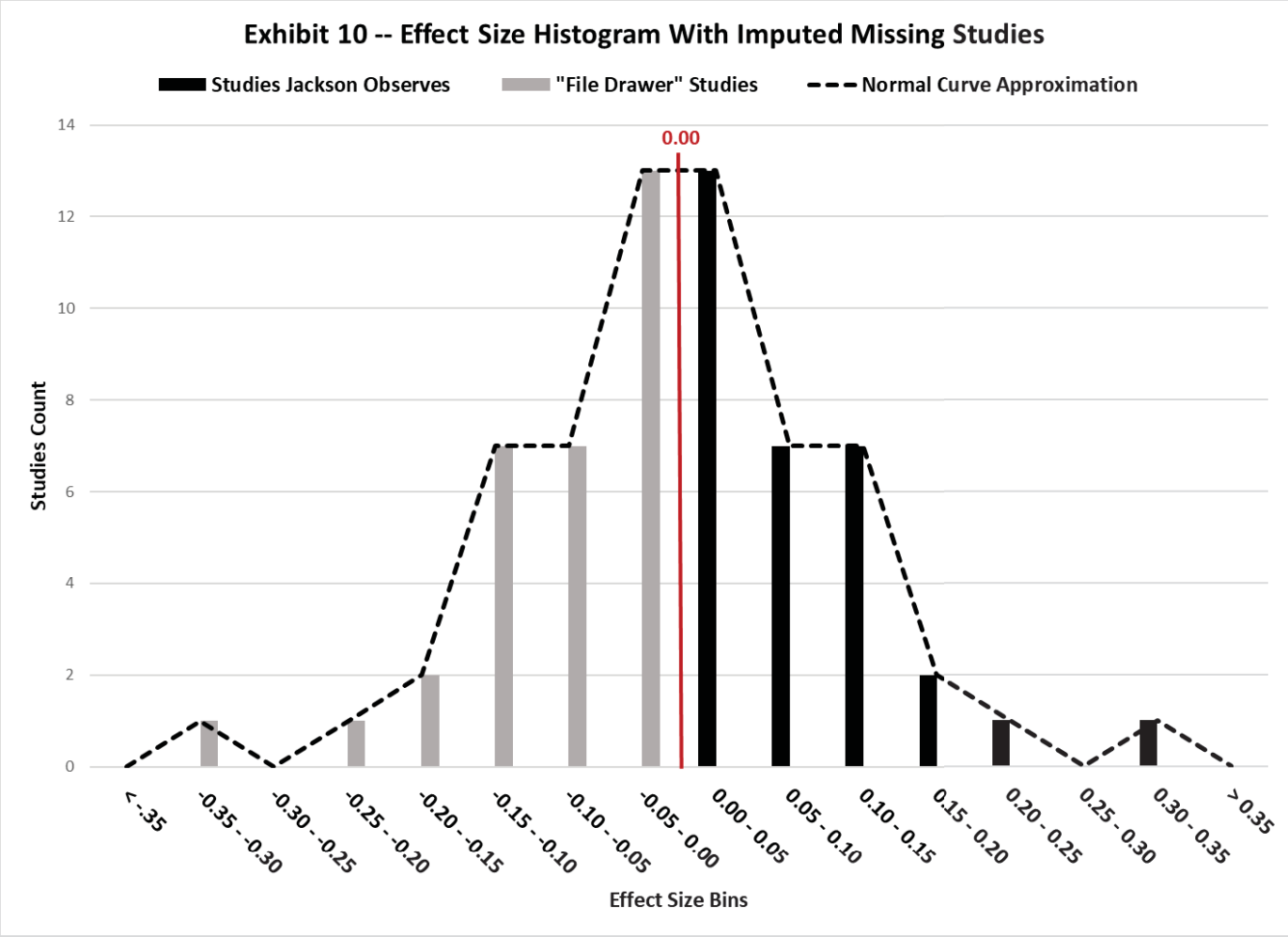


Exhibit 10 -- Effect Size Histogram With Imputed Missing Studies
Studies Jackson Observes
"File Drawer" Studies
Normal Curve Approximation
Studies Count
0.00
Effect Size Bins
< -.35
-0.35 - -0.30
-0.30 - -0.25
-0.25 - -0.20
-0.20 - -0.15
-0.15 - -0.10
-0.10 - -0.05
-0.05 - 0.00
0.00 - 0.05
0.05 - 0.10
0.10 - 0.15
0.15 - 0.20
0.20 - 0.25
0.25 - 0.30
0.30 - 0.35
> 0.35

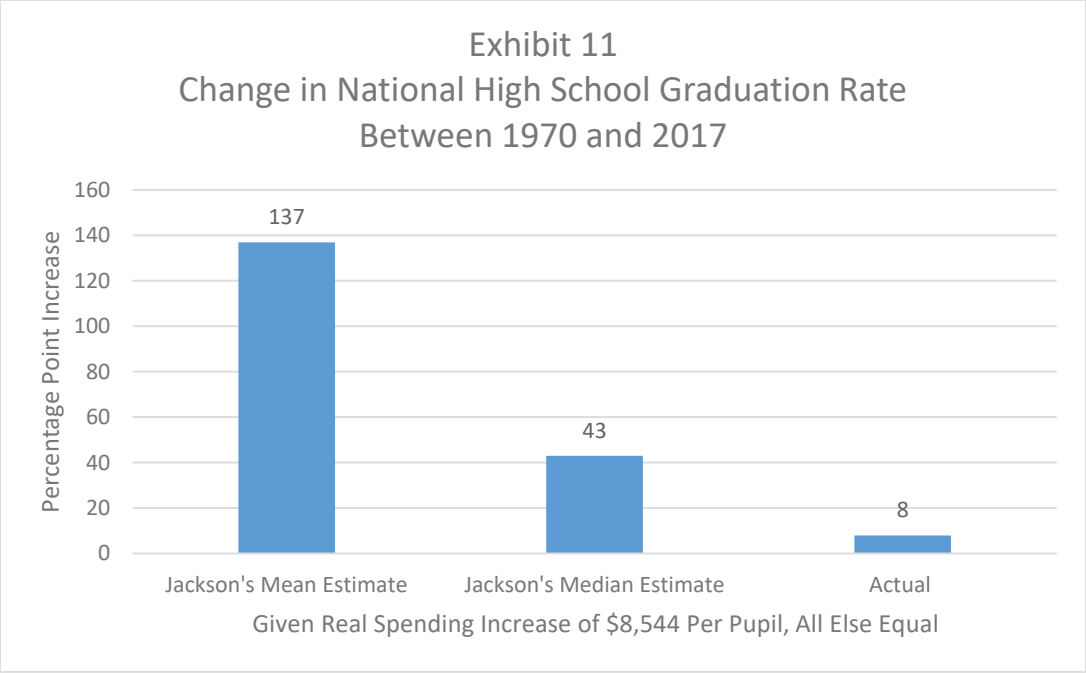


Exhibit 11
Change in National High School Graduation Rate
Between 1970 and 2017
Percentage Point Increase
137
43
8
Jackson's Mean Estimate
Jackson's Median Estimate
Actual
Given Real Spending Increase of $8,544 Per Pupil, All Else Equal

Exhibit 12 – Screenshot of Portion of Online Appendix E from Lafortune, Rothstein, & Schanzenbach (2018) Listing Discrepancies in the Identification of School Finance Reforms Between Lafortune, Rothstein, & Schanzenbach (2018) and Jackson, Johnson, & Persico (2016)

The states and years for which the two tabulations disagree are:

- Alabama, 1993
- Arizona, 2007
- Connecticut, 1995 & 2010
- Idaho, 1993 & 1998
- Maryland, 1996 & 2005
- Michigan, 1997
- Montana, 1993 & 2008
- New Hampshire, 2006
- New Jersey, 1991, 1998 & 2000
- New Mexico, 1998 & 1999
- Oregon, 2009
- South Carolina, 2005
- Texas, 2004
- Washington, 1991, 2007 & 2010

This includes only cases in scope for both lists but coded differently. This in particular means that we do not discuss our tabulation of legislative school finance reforms, as these are out of JJP's scope. For each state, we discuss only the events where the two tabulations disagree; see Online Appendix Table A1 for a full listing of events in each state.