What Do We Know About the Long-Term Impacts of Teacher Value-Added?

Stephen W. Raudenbush

University of Chicago | Apr 2014

CARNGEGIE KNOWLEDGE NETWORK

***What We Know Series*:**

Value –Added Methods and Applications

Should We Use VA?

- **Social Welfare**
- **Instructional Sensitivity**
- **Fairness**

From MET

- **Value Added**
- **Classroom Observations**
- **Student Perceptions**

But

- **Does VA last ?**
- **Does high VA predict things that matter?**

Highlights

- **Two recent studies provide evidence that attending a high value-added classroom predicts college attendance and earnings.**

- **In one study, part of the impact of attending an effective classroom may have been attributable to small class size; in the other, part of the effect may be attributable to the effectiveness of the school.**

- **Teacher value-added scores “fade out” over time.**

How Big is “Initial” VA

Teacher 1	**70th percentile**
Teacher 2	**30th percentile**
Teacher 1’s kids	**53rd percentile**
Teacher 2’s kids	**47th percentile**

Long Term Impact

Chetty et al. 2011 Tennessee STAR

- **Random Assignment k Teachers**

- **Random Assignment of teachers to small class size**

Chetty et al. 2013

- **2.5 million kids in NY**

- **Grade 3 – 8**

Table 1: Impacts of Value-Added on Adult Outcomes

	Impact of classroom quality overall (Chetty et al. 2011)	Impact of classroom value added (Chetty et al., 2011)	Impact of teacher value added (Chetty et al., 2013)
Initial test scores	8.8 percentiles (.32 sd)		
College Attendance		0.28% above mean of 45.5%	0.82% above mean of 37.22%
College Quality index		0.06 sd	0.02 sd
Earnings	$1520 =8.8% above mean	$1619 =11.1% above mean	$350 =1.65% above mean
Teen parenthood			0.61% below mean of 14.3%
Other outcomes			Increases in neighborhood quality, saving with 401K

Table 1: Impacts of Value-Added on Adult Outcomes

	Impact of classroom quality overall (Chetty et al. 2011)	Impact of classroom value added (Chetty et al., 2011)	Impact of teacher value added (Chetty et al., 2013)
Initial test scores	8.8 percentiles (.32 sd)		
College Attendance		0.28 percent above mean of 45.5%	0.82% above mean of 37.22%
College Quality index		0.06 sd	0.02 sd
Earnings	$1520 = 8.8% above mean	$1619 = 11.1% above mean)	$350 = 1.65% above mean)
Teen parenthood			0.61% below mean of 14.3%
Other outcomes			Increases in neighborhood quality, saving with 401K

Table 1: Impacts of Value-Added on Adult Outcomes

	Impact of classroom quality overall (Chetty et al. 2011)	Impact of classroom value added (Chetty et al., 2011)	Impact of teacher value added (Chetty et al., 2013)
Initial test scores	8.8 percentiles (.32 sd)		
College Attendance		0.28 percent above mean of 45.5%	0.82% above mean of 37.22%
College Quality index		0.06 sd	0.02 sd
Earnings	$1520=8.8% above mean	$1619 (11.1% above mean)	$350=1.65% above mean
Teen parenthood			0.61% below mean of 14.3%
Other outcomes			Increases in neighborhood quality, saving with 401K

Table 2: Persistence of Value-Added After Initial Year as Fraction of Value-Added During Initial Year

Study	Sample	Year 1	Year 2	Year 3	Year > 3
Kinsler (2012)	N=689,641 students, grades 3-5, 1998-2005, in North Carolina	.24 (math) .14 (reading)			
Master, Loeb, and Wycoff, 2014	N=700,000 students, grades 3-8, 2005-2226 in New York City	.19 (math) .21 (language arts)			
McCaffrey et al (2004)	N=678, grades 3-5, large suburban district	.25	.15	--	--
Lockwood et al	N=10,000, Grades 1-5, large urban district	.18	.15	.14	.12
Kane and Staiger (2008)	97 pairs of teachers, grades 2-5, randomization to students to teachers within pairs	.50			
Jacob, Lefgren, and Sims (2010)	n=18,240, grades 4-15, mid-size Western District	.20			
Rothstein (2010)	n=99,071, grades 3-5, North Carolina statewide	.27 (math) .33 (reading)			
Measurement of Effective Teaching (2012)	1811 teachers randomized within schools to student rosters, grades 4-8 in 6 school districts	.45			
Chetty et al. (2012)	10,992 students randomized to classes within 79 schools in Tennessee				0
Chetty et al. (2013)	2.5 million children grades 3-8 in NY	.50	.40	.20	.20

Table 2: Persistence of Value-Added After Initial Year as Fraction of Value-Added During Initial Year

Study	Sample	Year 1	Year 2	Year 3	Year > 3
Kinsler (2012)	N=689,641 students, grades 3-5, 1998-2005, in North Carolina	.24 (math) .14 (reading)			
Master, Loeb, and Wycoff, 2014	N=700,000 students, grades 3-8, 2005-2226 in New York City	.19 (math) .21 (language arts)			
McCaffrey et al (2004)	N=678, grades 3-5, large suburban district	.25	.15	--	--
Lockwood et al	N=10,000, Grades 1-5, large urban district	.18	.15	.14	.12
Kane and Staiger (2008)	97 pairs of teachers, grades 2-5, randomization to students to teachers within pairs	.50			
Jacob, Lefgren, and Sims (2010)	n=18,240, grades 4-15, mid-size Western District	.20			
Rothstein (2010)	n=99,071, grades 3-5, North Carolina statewide	.27 (math) .33 (reading)			
Measurement of Effective Teaching (2012)	1811 teachers randomized within schools to student rosters, grades 4-8 in 6 school districts	.45			
Chetty et al. (2012)	10,992 students randomized to classes within 79 schools in Tennessee				0
Chetty et al. (2013)	2.5 million children grades 3-8 in NY	.50	.40	.20	.20

Table 2: Persistence of Value-Added After Initial Year as Fraction of Value-Added During Initial Year

Study	Sample	Year 1	Year 2	Year 3	Year > 3
Kinsler (2012)	N=689,641 students, grades 3-5, 1998-2005, in North Carolina	.24 (math) .14 (reading)			
Master, Loeb, and Wycoff, 2014	N=700,000 students, grades 3-8, 2005-2226 in New York City	.19 (math) .21 (language arts)			
McCaffrey et al (2004)	N=678, grades 3-5, large suburban district	.25	.15	--	--
Lockwood et al	N=10,000, Grades 1-5, large urban district	.18	.15	.14	.12
Kane and Staiger (2008)	97 pairs of teachers, grades 2-5, randomization to students to teachers within pairs	.50			
Jacob, Lefgren, and Sims (2010)	n=18,240, grades 4-15, mid-size Western District	.20			
Rothstein (2010)	n=99,071, grades 3-5, North Carolina statewide	.27 (math) .33 (reading)			
Measurement of Effective Teaching (2012)	1811 teachers randomized within schools to student rosters, grades 4-8 in 6 school districts	.45			
Chetty et al. (2012)	10,992 students randomized to classes within 79 schools in Tennessee				0
Chetty et al. (2013)	2.5 million children grades 3-8 in NY	.50	.40	.20	.20

Table 2: Persistence of Value-Added After Initial Year as Fraction of Value-Added During Initial Year

Study	Sample	Year 1	Year 2	Year 3	Year > 3
Kinsler (2012)	N=689,641 students, grades 3-5, 1998-2005, in North Carolina	.24 (math) .14 (reading)			
Master, Loeb, and Wycoff, 2014	N=700,000 students, grades 3-8, 2005-2226 in New York City	.19 (math) .21 (language arts)			
McCaffrey et al (2004)	N=678, grades 3-5, large suburban district	.25	.15	--	--
Lockwood et al	N=10,000, Grades 1-5, large urban district	.18	.15	.14	.12
Kane and Staiger (2008)	97 pairs of teachers, grades 2-5, randomization to students to teachers within pairs	.50			
Jacob, Lefgren, and Sims (2010)	n=18,240, grades 4-15, mid-size Western District	.20			
Rothstein (2010)	n=99,071, grades 3-5, North Carolina statewide	.27 (math) .33 (reading)			
Measurement of Effective Teaching (2012)	1811 teachers randomized within schools to student rosters, grades 4-8 in 6 school districts	.45			
Chetty et al. (2012)	10,992 students randomized to classes within 79 schools in Tennessee				0
Chetty et al. (2013)	2.5 million children grades 3-8 in New York	.50	.40	.20	.20

Table 2: Persistence of Value-Added After Initial Year as Fraction of Value-Added During Initial Year

Study	Sample	Year 1	Year 2	Year 3	Year > 3
Kinsler (2012)	N=689,641 students, grades 3-5, 1998-2005, in North Carolina	.24 (math) .14 (reading)			
Master, Loeb, and Wycoff, 2014	N=700,000 students, grades 3-8, 2005-2226 in New York City	.19 (math) .21 (language arts)			
McCaffrey et al (2004)	N=678, grades 3-5, large suburban district	.25	.15	--	--
Lockwood et al	N=10,000, Grades 1-5, large urban district	.18	.15	.14	.12
Kane and Staiger (2008)	97 pairs of teachers, grades 2-5, randomization to students to teachers within pairs	.50			
Jacob, Lefgren, and Sims (2010)	n=18,240, grades 4-15, mid-size Western District	.20			
Rothstein (2010)	n=99,071, grades 3-5, North Carolina statewide	.27 (math) .33 (reading)			
Measurement of Effective Teaching (2012)	1811 teachers randomized within schools to student rosters, grades 4-8 in 6 school districts	.45			
Chetty et al. (2012)	10,992 students randomized to classes within 79 schools in Tennessee				0
Chetty et al. (2013)	2.5 million children grades 3-8 in NY	.50	.40	.20	.20

Questions for Future

- **Why fade out?**
- **Skills not measured on achievement tests**
  - **Academic**
  - **Non-academic**

Key Caveats

- **Precision**
- **Partial coverage of “Social Welfare”**
- **Role of school as organization**