EC607 Public Presentation

Melissa Wilson

University of Oregon

May 13, 2018

"The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit"

Gordan B. Dahl and Lance Lochner (American Economic Review, 2012)

Introduction I

- This paper analyzes the effects of income on children's math and reading achievement by exploiting large, nonlinear changes in the Earned Income Tax Credit (EITC) using a panel of 4,412 children matched to their mothers
- Expansions of the EITC in the late 1980s and 1990s provide an exogenous source of variation in income, so the authors use this to overcome the problems caused by the endogeneity of income and identify true causal effects
- Use a first-differenced child outcome equation and implement IV approach outlined later
- Affirms importance of family income effects on child academic achievement and shows it is heterogeneous across family income levels

Background

- Prior literature shows significant variation in estimated effects of family income on child development
- Estimates are mostly positive and significant, but magnitude changes
- The direction of the relationship may be clear, but the causal relationship has not been identified due to omitted heterogeneous unobservable home-life characteristics
- Some have used FE in order to overcome this, but they do not control for endogenous transitory shocks
  - Eg. job loss, promotion, family illness, moving, etc.
  - Likely suffer from attenuation bias because income growth is measured noisily

Methodology I

- Child outcome equation:

$$y_{ia} = x_i^{'}\alpha_a + w_{ia}^{'}\beta + I_{ia}\delta_0 + I_{i,a-1}\delta_1 + \cdots + I_{i,a-L}\delta_L + \mu_i + \epsilon_{ia} \quad (1)$$

  for child i at age a with L lags

- Taking first differences and setting $L = 0$ gives the baseline equation:

$$\Delta y_{ia} = x_i^{'}\alpha + \Delta w_{ia}^{'}\beta + \Delta I_{ia}\delta_0 + \Delta\epsilon_{ia} \quad (2)$$

  where $\alpha \equiv \alpha_a - \alpha_{a-1}$ is the effect of x_i on achievement growth

- Assuming $L = 0$ gives the "contemporaneous effects" of family income on children, ignoring long-run effects
- Authors will allow one- and two-year lags, as well after baseline model

Methodology II

- Primary concern with OLS is that $\Delta\epsilon_{ia}$ may be correlated with entire history of family income
- Employ the following instrumental variable approach
- EITC income is a function of pretax income and taxes:

$$I_{ia} = P_{ia} + \chi_a^{s_{ia}}(P_{ia}) - \tau_a^{s_{ia}}(P_{ia})$$

where $\chi_a^{s_{ia}}(P_{ia})$ is a flexible function of lagged pretax income

- Use IV:

$$\Delta\chi_a^{IV}(P_{i,a-1}) \equiv \chi_a^{s_{i,a-1}}(\hat{E}\left[P_{i,a}|P_{i,a-1}\right]) - \chi_a^{s_{i,a-1}}(P_{i,a-1})$$

where $\hat{E}\left[P_{i,a}|P_{i,a-1}\right]$ is an estimate of pretax income given lagged pretax income

- This builds on IV used in Gruber and Saez (2002)

Methodology III

Thus, IV estimation is of the following equation

$$\Delta y_{ia} = x_i^{'}\alpha + \Delta w_{ia}^{'}\beta + \Delta I_{ia}\delta_0 + \Phi(P_{i,a-1}) + \eta_{ia} \tag{3}$$

where $\Phi()$ is a very flexible polynomial in lagged pretax income

Methodology IV

- Assumes the relationship between child development shocks and lagged pretax income must be stable over time
- Identification from differential changes in EITC schedule over time
- Minor issues with data and model:
  (1) Vast majority of EITC recipients receive their credit after filing taxes the following year; thus, authors link test scores with income earned in previous year
  (2) Only observe child achievement scores every other year; thus, authors use two-year differences

Data I

- Use data from National Longitudinal Survey of Youth (NLSY)
- Links children to their mothers and follows families over time, allowing use of child FEs
- Repeated measures of academic achievement and family income
- Oversamples minority families, providing a larger sample of families eligible for EITC
- Note that the NLSY does not provide information on how much a family receives from EITC, so authors must approximate this based on IRS (2002) and Scholz (1994) estimates that 80-87% of eligible HH receive EITC
- Authors assume full take-up and impute each family's state and federal EITC payment and tax burden using the TAXSIM program from the NBER

Data II

- Restrict sample to children observed in at least two consecutive survey years, since using FEs
- Restrict sample to those children whose mothers did not change marital status between test score measures because changes in family income are likely to mean something very different when there is a change of marital status relative to when there is not
- There is a noticeable income increase in the sample (increases faster than price levels) and authors show it is largely attributable to mothers in the sample aging
- Average child age in sample is 11
- Over half of sample are minorities due to oversampling of minorities in NSLY

OLS Results I

TABLE 2—OLS ESTIMATES OF THE EFFECT OF FAMILY INCOME ON MATH-READING ACHIEVEMENT

	(1)	(2)	(3)	(4)
Panel A. Estimated in levels				
Current income	0.0047**	0.0031**	0.0022	0.0023
	(0.0011)	(0.0014)	(0.0016)	(0.0015)
Lagged income ($a-1$)		0.0022	0.0019	
		(0.0016)	(0.0024)	
Lagged income ($a-2$)			0.0015	
			(0.0019)	
Sum of ($a-1$) and ($a-2$) lagged income				0.0017*
				(0.0009)
Medium-term effect of increasing income by \$1,000/year for 3 years	0.0047**	0.0053**	0.0056**	0.0056**
	(0.0011)	(0.0013)	(0.0015)	(0.0015)
Panel B. Estimated in differences				
Current income	0.0011	0.0015*	0.0011	0.0016*
	(0.0007)	(0.0008)	(0.0010)	(0.0009)
Lagged income ($a-1$)		0.0005	0.0012	
		(0.0009)	(0.0011)	
Lagged income ($a-2$)			−0.0007	
			(0.0009)	
Sum of ($a-1$) and ($a-2$) lagged income				0.0001
				(0.0005)
Medium-term effect of increasing income by \$1,000/year for 3 years	0.0010	0.0020*	0.0015	0.0018
	(0.0007)	(0.0010)	(0.0013)	(0.0013)
Sample size (for both panels)	8,609	6,543	5,019	5,019

OLS Results II

- Possible reasons for discrepancy when including lags:
  (1) Measurement error is greater for those measured in differences, so attenuation bias is greater for differenced estimates
  (2) Correlation between unobserved FEs (μ_i) and family income biasing cross-sectional OLS estimates
- Both suffer from OVB due to transitory shocks

IV Estmates: Contemporaneous Effects

TABLE 3—BASELINE IV ESTIMATES OF "CONTEMPORANEOUS EFFECTS" MODEL

	Combined math and reading (1)	Reading recognition (2)	Reading comprehension (3)	Math (4)
Current income	0.0610**	0.0359*	0.0613**	0.0582**
	(0.0231)	(0.0195)	(0.0273)	(0.0273)
First stage coefficient on instrument	1.270**	1.270**	1.270**	1.270**
	(0.381)	(0.381)	(0.381)	(0.381)

IV Results: Contemporaneous Effects

- Table 3 shows raising income by $1000 increases math-reading achievement by 6 percent of a standard deviation (not very big, but larger than OLS)
- Table 4 takes into account national time trends and changes in state-level school accountability and welfare policies
- Authors use
  A. year dummies to allow for average test scores to vary year to year, so identification of IV estimate comes from differences in predicted EITC across individuals
  B. linear time trends
  C. linear time trend and interaction of trend with control function $\Phi()$ to account for relationship between child outcomes and pretax income changing over time
  D. and
  E. Address changes in state policies

IV Estimates: Contemporaneous Effects and Controls

TABLE 4—IV ESTIMATES OF "CONTEMPORANEOUS EFFECTS" MODEL ACCOUNTING FOR TIME TRENDS AND TIME-VARYING STATE POLICIES
(*Math-Reading Achievement*)

	Effect of current income	First stage coefficient on instrument
A. Year dummies	0.0694*	0.745**
	(0.0390)	(0.348)
B. Linear time trend	0.0863**	0.847**
	(0.0379)	(0.334)
C. Linear time trend interacted with control function	0.0805**	1.115**
	(0.0399)	(0.485)
D. State school accountability policies interacted with control function	0.0533**	1.299**
	(0.0221)	(0.406)
E. State welfare policies interacted with control function	0.0670**	1.311**
	(0.0268)	(0.436)
F. Time trend, accountability, and welfare policies interacted with control function	0.0630*	1.193**
	(0.0338)	(0.513)

IV Estimates: Contemporaneous and Lasting Effects

TABLE 5—IV ESTIMATES OF ACHIEVEMENT MODELS WITH LASTING INCOME EFFECTS

	(1)	(2)	(3)
Current income	0.0436*	0.0551	0.0515**
	(0.0236)	(0.0478)	(0.0226)
Lagged income (a-1)	0.0216	0.0135	
	(0.0408)	(0.0733)	
Lagged income (a-2)		0.0206	
		(0.0381)	
Sum of (a-1) and (a-2) lagged income			0.0186
			(0.0254)
Medium-term effect of increasing income by $1,000/year for three years	0.0651*	0.0892	0.0888
	(0.0349)	(0.0604)	(0.0598)
F-statistics from first stage	6.17, 3.59	3.98, 1.39, 2.16	5.53, 1.77
Sample size	6,543	5,019	5,019

IV Results: Contemporaneous and Lasting Effects

- In table 5, the authors allow for lasting effects of income changes
- The results are largely consistent, but more precise
- Larger contemporaneous effects than lasting effects
- The simple "contemporaneous effects" model provides reasonably good estimates of short-run effects of income

IV Estimates: Heterogeneous Effects

TABLE 6—IV ESTIMATES OF "CONTEMPORANEOUS EFFECTS" MODEL FOR VARIOUS SUBGROUPS

	Mother's education	Race	Mother's marital status	Mother's AFQT	Child's age	Child's gender
	High school or less	Black or Hispanic	Not married	Low AFQT	Age < 12	Male
Effect of current income	0.0536** (0.0211)	0.0800** (0.0304)	0.0806* (0.0463)	0.0708** (0.0340)	0.0765* (0.0436)	0.0879** (0.0446)
First-stage coefficient on instrument	1.386** (0.402)	1.281** (0.428)	0.808** (0.389)	1.089** (0.433)	1.050** (0.495)	1.056** (0.472)
"Percent in EITC Range"	56.4	62.8	90.1	64.9	46.4	49.6
Sample size	6,253	4,602	2,977	4,311	4,654	4,261
	Some college or more	White (not Hisp.)	Married	High AFQT	Age ≥ 12	Female
Effect of current income	0.0163 (0.0107)	0.0146 (0.0295)	0.0434* (0.0248)	0.0486 (0.0361)	0.0516** (0.0235)	0.0399* (0.0221)
First-stage coefficient on instrument	0.086 (1.123)	1.265 (0.798)	2.153** (0.907)	1.466* (0.802)	1.460** (0.452)	1.479** (0.489)
"Percent in EITC Range"	30.8	34.1	28.0	33.3	53.0	49.3
Sample size	2,356	4,007	5,632	4,040	3,955	4,348

IV Results: Heterogeneous Effects

- The extent of the impact of a $1000 increase in current income reported as "Percent in EITC Range" for each subgroup
- Higher socioeconomic status (SES) groups have lower PEITCR
- This is reflected in higher SEs for those groups' point estimates

Robustness I

- Include additional controls for parents and families
- Remove all controls from baseline
- Interact all baseline controls with pretax income and the polynomial lagged pretax income
- Show inclusion of state FEs has little impact

Robustness II

- NLSY-created weights for initial sample of mothers to weight observations
  - Smaller point estimate
  - Larger SEs
- Use log total family income as RHS, rather than levels
- Add changes to maternal LFP and hours worked to baseline

Conclusion

- Exploiting exogenous variation in the EITC, the authors estimate that income has a modest positive contemporaneous effect on child academic achievement
- Baseline estimates imply that a $1000 increase in family income raises math and reading comprehension scores by 6 percent of a standard deviation
- For the whole sample, this implies an average test score increase of 10 percent of a standard deviation based on EITC payments increasing income of two-child families by $1670, on average
- Effects are larger for disadvantaged families, younger children, and boys