Online Appendix:

An Alternative Measure of Intergenerational Income Mobility

Based on a Random Coefficient Model

Irina Murtazashvili

January 15, 2011

Sensitivity Analysis for the IV and CF Estimates of Intergenerational Income Mobility

When applying IV to the estimation of intergenerational income mobility, the traditional approach is to start with the assumption that the sons' long-run income, y_{Si}, is determined by

$$y_{Si} = \gamma_1 y_{Fi} + \gamma_2 E_i + \varepsilon_i^*, \tag{1}$$

where y_{Fi} is the fathers' long-run income, E_i is the fathers' education, and ε_i^* is the idiosyncratic error.

When attempting a sensitivity analysis of our IV and CF estimates of intergenerational income mobility, we first searched for values of γ_2 obtained in other empirical studies. When doing so, we ran into several problems:

(1) We did not find any studies, which used specification (6) from our paper. This, of course, was not surprising. Other empirical studies differed from our specification (6) in four directions. First, all of the other papers (obviously) added more control variables to model (6). Second, some

other papers used a different dependent variable (for example, log of hourly wage (Corcoran et al. (1992), Datcher (1982), Hill and Duncan (1987), Zimmerman (1992)), log of family income (Corocoran et al. (1992), Peters (1992)), the Duncan index of socioeconomic status (Zimmerman (1992)), or income in levels (Das and Sjogren (2002))). Third, some of the other empirical studies employed different measures of education. For example, some studies used a dummy for whether the father had a college degree only and included the grandfather's education as an additional regressor (Peters (1992)). Finally, some of the existing studies estimated (augmented) model (6) for subsets of the population (for example, for black and white individuals (Datcher (1982)), or for adopted and biological children (Das and Sjogren (2002)) separately).

(2) Some empirical studies discussed the significance of their estimation results but did not report them in full, thus, making it impossible for us to know their estimates of γ_2 (for example, Mazumder (2005)).

(3) More importantly, some of the reported significant estimates of γ_2 from other empirical studies had a WRONG sign (i.e., negative sign). See, for example, Das and Sjogren (2002) and Datcher (1982) among others.

(4) There are studies of intergenerational mobility in countries other than the USA, of course. In particular, Lillard and Kilburn (1995) study intergenerational mobility in Malaysia. Even though the latter study is devoted to Malaysia we decided to use the estimate of γ_2 from that paper because Lillard and Kilburn pointed out that the "data from Malaysia yield surprisingly similar estimates to those obtained for the U.S." We assumed that this statement was true about all the estimates they reported given that they did point out similarity of their results with the ones for the U.S. on several occasions.

The First Estimation Strategy from Conley et al. (2007)

Clearly, all the issues discussed above affect the estimates of γ_2. However, we decided to ignore those issues and used all the estimates of γ_2 that we found for our sensitivity analysis. Table

A reports the results of sensitivity analysis using the first approach from Conley et al. (2007). Specifically, Table A contains the IV and CF estimates of ρ_o when the dependent variables used for these methods were constructed by subtracting $\hat{\gamma}_2$ times father's education from the log earnings of sons. Columns (3)-(5) of Table A report the results for the IV method, while columns (6)-(8) report the results for the CF method.

Given that the only significant result with the expected sign is from Lillard and Kilburn (1995), we use their estimate of γ_2 to interpret the results from Table A. The 95% confidence intervals for the IV estimate of ρ_o is $(0.333; 0.913)$ and $(0.369; 0.863)$ based on the sample of oldest and multiple sons, respectively. Similarly, the 95% confidence intervals for the CF estimate of ρ_o is $(0.366; 0.982)$ and $(0.407; 0.917)$ for oldest and multiple sons, respectively.

"Back of the Envelope Approach"

Finally, we attempted to follow the "back of the envelope" approach from Bound et al. (1995) to assess whether the estimate of γ_2 from Lillard and Kilburn (1995) was appropriate for the U.S. Behrman and Taumban (1985) find the intergenerational correlation in schooling between fathers and sons in the U.S. to range from 0.17 to 0.33. Behrman and Taubman (1985) estimate the effect of educational attainment on log earnings for a particular generation in the U.S. to be 0.082, while Grawe (2004) finds it to be 0.089 (using the U.S. data, as well). Given these results, the direct effect of fathers' education on sons' earnings is expected to be between 0.014 and 0.029, making the estimate of γ_2 from Lillard and Kilburn (1995) reasonable for the U.S.

Note that the estimates of ρ_o that are based on the insignificant estimates of γ_2 from Hill and Duncan (1987) and Corcoran et al. (1992) are within one standard deviation of the estimates of ρ_o reported in Table 4 of our paper.

Overall, the two exercises above suggest that our estimates of ρ_o in the main paper are within the interval obtained based on sensitivity analysis. However, we admit that this interval is not quite reliable due to its width.

References not in the Main Paper

Bond J, Jaerger DA, Baker RM. 1995. Problems with Instrumental Variables When the Correlation Between the Instruments and the Exogenous Explanatory Variable is Weak. *Journal of the American Statistical Association* **90**: 443-450.

Conley T, Hansen C, Rossi P. 2007. Plausibly Exogenous. Mimeo, Chicago Graduate School of Business.

Das M, Sjogren T. 2002. The Intergenerational Link in Income Mobility: Evidence from Adoptions. *Economics Letters* **75**: 55-60.

Grawe ND. 2004. Intergenerational Mobility for Whom? The Experience of High- and Low-Earnings Sons in International Perspective. In *Generational Income Mobility in North America and Europe*, Miles Corak (ed). Cambridge University Press: Cambridge, MA.

Hahn J, Hausman J. 2005. Instrumental Variable Estimation with Valid and Invalid Instruments. *Annales d'Economie et Statistique* **79/80**: 25-57.

Lillard LA, Kilburn MR. 1995. Intergenerational Earnings Links: Sons and Daughters. RAND Labor and Population Program Working Paper DRU-1125-NIA.

Table A. Results for Sensitivity Analysis.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Source	$\hat{\gamma}_2$	$\hat{\rho}_o^{IV}$	SE	$P-value$	$\hat{\rho}_o^{CF}$	SE	$P-value$
				Oldest Sons			
Das & Sjogren (2002)	-0.098*	-0.487	0.166	0.004	-0.436	0.188	0.020
Datcher (1982)	-0.061*	-0.137	0.151	0.363	-0.086	0.169	0.610
Hill & Duncan (1987)	-0.012	0.326	0.143	0.024	0.377	0.156	0.016
Corcoran et al. (1992)	-0.011	0.336	0.143	0.020	0.386	0.156	0.013
Lillard & Kilburn (1995)	0.0194*	0.623	0.148	0.000	0.674	0.157	0.000
				Multiple Sons			
Das & Sjogren (2002)	-0.098*	-0.487	0.149	0.001	-0.441	0.152	0.004
Datcher (1982)	-0.061*	-0.139	0.132	0.289	-0.094	0.137	0.495
Hill & Duncan (1987)	-0.012	0.321	0.122	0.009	0.369	0.129	0.004
Corcoran et al. (1992)	-0.011	0.331	0.122	0.007	0.376	0.129	0.003
Lillard & Kilburn (1995)	0.0194*	0.616	0.126	0.000	0.662	0.130	0.000

*Significant (at least) at one of the three conventional levels (0.01, 0.05, or 0.10).

Notes: Columns (3)-(5) report the results for the IV method, while columns (6)-(8) report the results for the CF method. The sample from Das and Sjogren (2002) contains biological sons only. The sample from Datcher (1982) contains black individuals only. Data used in Lillard and Kilburn (1995) are Malaysian data. Standard errors for the IV method are robust SEs for the sample of oldest sons. Standard errors for the IV method are robust to sibling's correlation for the sample of multiple sons. Standard errors for the CF method are bootstrapped SEs using 1000 replications.