Appendix 1: Additional Tables

The following tables present additional summary statistics and robustness checks for the article, "American Mobility and the Expansion of Public Education." Descriptions of the data sources used for the article are given in Appendix 2. Appendix 3 presents a discussion of the effects of mismatches between the Iowa state census and the federal census when constructing the intergenerational data set used in the article.

Appendix Table 1 provides summary statistics for the sons in the sample by distance moved. The purpose of the table is to assess the selection issues caused by losing the most geographically mobile individuals in the data construction process. The statistics in Appendix Table 1 reveal that the more geographically mobile individuals that remain in the data set tend to have higher incomes, more years of education, and fewer months of unemployment than their less geographically mobile counterparts. This suggests that the limitations of the linking process may lead to a data set that is underrepresentative of highly educated and successful sons. However, it should be noted that the differences in means between the highly geographically mobile sons and the stationary sons are quite small relative to the standard deviations of the variables within each group.

Appendix Table 2 presents several alternative measures of income mobility to supplement the intergenerational income elasticities in Table 2 of the article. As with the intergenerational income elasticity, these alternative measures all show greater income mobility in Iowa in 1915 than in modern times. Goodman and Kruskal's gamma and Kendall's tau-b are both measures of the number of concordant pairs relative to discordant pairs in the data.[1] A pair of observations A and B is concordant if the income of son A is greater than the income of son B and the income of father A is greater than the income of father B. A discordant pair would be a pair of observations A and B for which son A has a greater income than son B, but father B has a greater income than father A. A larger number of concordant pairs produces larger values for gamma and tau-b and is interpreted as less income mobility. Thus the significantly larger values of these measures for 2001 relative to 1915 suggests that the Iowa sample exhibits substantially more income mobility than the modern sample. The quintile measures show that sons were more likely to move larger distances in the income distribution in 1915 and that there was less persistance in each quintile of the income distribution in 1915, particularly in the bottom and top income quintiles. These results are consistent with intergenerational income elasticities in Table 2 of the article as well as the increasing persistence in the tails of the income distribution with improving school access and quality found in the logit estimates in Table 5 of the article.

Appendix Table 3 addresses the issue of the endogeneity of school district characteristics. This table shows statistics of the father's income distribution by school district access and quality. Graded classrooms per square mile is used as a proxy for school access and spending per student is used as a proxy for school quality. The table reveals no discernible trends in the mean or variance of the income distribution by school district access or quality suggesting that the better school districts were not systematically located in richer communities.

Appendix Table 4 examines whether the quality of a school district was correlated with the geographical mobility of its residents and consequently the likelihood of

[1] Goodman and Kruskal, "Measures of Association"; and Kendall, "New Measure of Rank Correlation."

APPENDIX TABLE 1
SUMMARY STATISTICS FOR THE IOWA SAMPLE BY DISTANCE MOVED BETWEEN 1900 AND 1915

	All Sons	Sons Moving More than 20 Miles	Sons that did not Move
Son's earnings	652	893	700
	(488)	(682)	(579)
Father's earnings	1039	1052	1186
	(1054)	(774)	(2041)
Son's months unemployed	1.0	0.1	1.2
	(2.2)	(0.6)	(2.4)
Father's months unemployed	0.7	0.7	0.9
	(1.9)	(1.5)	(2.3)
Son's total years of education	9.1	10.2	9.4
	(2.5)	(3.6)	(2.5)
Father's total years of education	7.9	8.6	8.3
	(2.7)	(2.6)	(2.7)
Son's years of common school	4.0	2.1	2.7
	(4.3)	(3.3)	(3.8)
Son's years of grammar school	4.0	5.6	5.3
	(4.0)	(4.1)	(3.9)
Son's years of high school	0.8	1.7	1.1
	(1.4)	(1.8)	(1.7)
Son's years of college	0.2	0.6	0.3
	(1.4)	(1.4)	(1.7)
Number of observations	1094	28	158

Notes: Moving distances are based on the distance between the son's location in 1900 and in 1915. "Did not move" refers to sons that stayed in the same township, not necessarily the same dwelling.
Sources: Data are from the linked father-son sample based on the 1915 Iowa state census and the 1900 federal census.

those individuals being successfully linked. The data for the table come from a random sample of individuals from the 1900 federal census living in the sample counties matched to the 1920 federal census. The probability of being found in 1920 was very weakly correlated with school quality and school access. The probability of still living in a sample county was negatively correlated with the number of graded classrooms per square mile and only very weakly correlated with school quality. These results suggest that communities with differing school quality did not have significantly different levels of geographical mobility, but that communities with better school access may have had more geographically mobile individuals. If more geographically mobile individuals also exhibit more income mobility, the marginal effect of school access on the intergenerational income elasticity estimated in the article will be overstated due to geographically mobile individuals being underrepresented in the sample.

APPENDIX TABLE 2
ALTERNATIVE INCOME MOBILITY MEASURES

	Iowa, 1915	PSID, 2001
	(1)	(2)
Goodman and Kruskal's gamma	0.082	0.258
	(0.031)	(0.027)
Kendall's tau-b	0.065	0.208
	(0.025)	(0.022)
Mean change in quintile	1.51	1.31
	(1.18)	(1.13)
Percentage persisting in father's income quintile:		
From bottom quintile	26.9	34.3
From 2nd quintile	20.2	23.0
From 3rd quintile	19.6	26.0
From 4th quintile	20.9	23.3
From top quintile	25.2	32.5

Notes: All gamma and tau-b values are calculated using income quintiles. Values in parentheses are asymptotic standard errors for gamma and tau-b and the standard deviation for the mean change in quintile. All income quintiles are based on age-adjusted incomes.
Sources: Iowa data are from the linked father-son sample created from the 1915 Iowa state census and the 1900 federal census. The PSID data are from the 2001 Panel Study of Income Dynamics.

APPENDIX TABLE 3
SUMMARY STATISTICS FOR THE DISTRIBUTION OF FATHER'S LOG EARNINGS BY SCHOOL DISTRICT CHARACTERISTICS

	Mean	Standard Deviation	10th Percentile	50th Percentile	90th Percentile
School access quartile:					
1	6.69	0.82	5.99	6.68	7.60
2	6.80	0.78	5.52	6.91	7.70
3	6.69	0.89	5.70	6.68	7.60
4	6.80	0.72	5.93	6.73	7.60
School quality quartile:					
1	6.63	0.87	5.30	6.68	7.60
2	6.80	0.62	5.70	6.91	7.82
3	6.81	0.73	5.52	6.80	7.60
4	6.62	0.81	5.86	6.57	7.50

Notes: School access is measured as the number of graded classrooms per square mile. School quality is measured as the spending per student. All statistics refer to the distribution of father's log earnings, where earnings are measured in 1915 dollars.
Sources: Earnings data are from the 1915 Iowa state census. School access and quality data are from the 1900 annual report of the superintendent of public instruction for the state of Iowa.

APPENDIX TABLE 4
CORRELATIONS BETWEEN GEOGRAPHICAL MOBILITY AND CHILDHOOD SCHOOL DISTRICT CHARACTERISTICS

		Conditional on Being Found in 1920:	
	Found in 1920 (1 = found, 0 = not found)	Living in Iowa (1 = in Iowa)	Living in a Sample County (1 = in sample county)
Mean	0.30	0.57	0.42
Correlation with number of graded classrooms per square mile	0.019	–0.126	–0.086
Correlation with spending per student	–0.023	0.025	–0.010
Number of observations	1719	510	510

Notes: School district characteristics correspond to the year 1900 for the district the child was living in during that year.
Sources: The geographical mobility measures are based on a sample of individuals matched from the 1900 federal census to the 1920 federal census. School district data are from the 1900 annual report of the superintendent of public instruction for the state of Iowa.

Appendix Tables 5 through 9 present the complete regression results for the estimates of intergenerational income elasticities conditional on school district quality. The coefficients for the interaction terms between father's log earnings and the schooling measure from these regressions are reported in Table 4 of the article. Each table corresponds to a different sample (urban, rural, rural outside of a town, rural inside of a town). It is important to note that in all tables son's age is defined as age minus 30. Therefore, coefficients on father's log earning and the interaction term between father's log earnings and the schooling measure give the intergenerational income elasticity for a 30-year-old son. Computing the income elasticity at any other age requires using all of the son's age—father's log earnings interaction coefficients as well.

Appendix 2: Iowa Schools Data

The schooling data used throughout the article come from the Annual Reports of the Superintendent of Public Instruction. Each year, the superintendent of schools for each county in Iowa would submit a report to the superintendent of public instruction for the state. Included in this report was detailed information on the finances, enrollments, teachers, and miscellaneous details of the schools in the county.

APPENDIX TABLE 5
INTERGENERATIONAL INCOME ELASTICITY ESTIMATES WITH SCHOOLING MEASURE INTERACTIONS FOR SONS FROM URBAN TOWNSHIPS, SON'S LOG EARNINGS AS DEPENDENT VARIABLE

Schooling Measure:	Teacher-Student Ratio	Graded Teacher-Student Ratio	Classrooms per Square Mile	Graded Classrooms per Square Mile	Spending per Student	Taxes per Student
	(1)	(2)	(3)	(4)	(5)	(6)
Schooling measure	140.960	–196.285**	0.238***	0.207**	–0.159	–0.466*
	(140.784)	(46.307)	(0.048)	(0.046)	(0.441)	(0.196)
Father's log earnings x schooling measure	–23.909	26.402**	–0.032**	–0.027**	0.025	0.067*
	(19.067)	(8.066)	(0.009)	(0.008)	(0.068)	(0.029)
Father's log earnings	0.968	–0.408	0.386***	0.349***	–0.093	–0.581
	(0.584)	(0.203)	(0.050)	(0.057)	(0.940)	(0.339)
Father's age	–0.044	–0.038	–0.042	–0.046	–0.044	–0.038
	(0.097)	(0.099)	(0.096)	(0.094)	(0.094)	(0.095)
(Father's age)^2	0.000	0.000	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Son's age	0.046	–0.002	0.002	0.010	0.034	–0.025
	(0.155)	(0.159)	(0.148)	(0.146)	(0.163)	(0.173)
(Son's age)^2	–0.001	–0.007	–0.006	–0.006	–0.003	–0.007
	(0.009)	(0.011)	(0.010)	(0.010)	(0.009)	(0.010)
Son's age x Father's log earnings	–0.002	0.005	0.004	0.003	–0.000	0.008
	(0.023)	(0.024)	(0.022)	(0.022)	(0.024)	(0.026)
(Son's age)^2 x Father's log earnings	–0.000	0.000	0.000	0.000	–0.000	0.000
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Constant	2.234	11.261***	5.449	5.803*	8.710	12.169**
	(5.949)	(1.881)	(2.672)	(2.557)	(5.162)	(3.554)
Observations	197	197	197	197	197	197
R-squared	0.41	0.41	0.41	0.41	0.39	0.40

* significant at 10 percent.
** significant at 5 percent.
*** significant at 1 percent.
Notes: Standard errors, clustered by township, are given in parentheses. Graded schools dummy equals one if district has graded schools and zero otherwise. Subsidy per student is defined as spending per student minus annual tuition per student.
Sources: See the text.

APPENDIX TABLE 6
INTERGENERATIONAL INCOME ELASTICITY ESTIMATES WITH SCHOOLING MEASURE INTERACTIONS FOR SONS FROM RURAL TOWNSHIPS, SON'S LOG EARNINGS AS DEPENDENT VARIABLE

Schooling Measure:	Teacher-Student Ratio	Graded Teacher-Student Ratio	Classrooms per Square Mile	Graded Classrooms per Square Mile	Spending per Student	Taxes per Student
	(1)	(2)	(3)	(4)	(5)	(6)
Schooling measure	–23.638*	–124.932	–1.410*	–1.673***	–0.088**	–0.097**
	(12.746)	(87.624)	(0.780)	(0.589)	(0.037)	(0.044)
Father's log earnings x Schooling measure	2.990	18.301	0.229*	0.276***	0.012**	0.014**
	(1.904)	(13.684)	(0.125)	(0.079)	(0.006)	(0.007)
Father's log earnings	–0.111	–0.426*	–0.029	–0.002	–0.085	–0.067
	(0.122)	(0.198)	(0.074)	(0.060)	(0.083)	(0.078)
Father's age	–0.097*	0.037	–0.073	–0.077	–0.073	–0.071
	(0.050)	(0.238)	(0.052)	(0.051)	(0.052)	(0.051)
(Father's age)^2	0.001*	–0.000	0.001	0.001	0.001	0.000
	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)
Son's age	–0.046	–0.052	–0.073	–0.075	–0.085	–0.087
	(0.062)	(0.180)	(0.057)	(0.059)	(0.055)	(0.057)
(Son's age)^2	–0.010	–0.033	–0.009	–0.009	–0.007	–0.007
	(0.007)	(0.020)	(0.007)	(0.007)	(0.007)	(0.007)
Son's age x Father's log earnings	0.014	0.013	0.018**	0.019**	0.020**	0.020**
	(0.010)	(0.031)	(0.009)	(0.009)	(0.009)	(0.009)
(Son's age)^2 x Father's log earnings	0.001	0.005	0.001	0.001	0.001	0.001
	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	10.518***	8.392	9.130***	9.070***	9.609***	9.417***
	(1.623)	(7.093)	(1.463)	(1.469)	(1.420)	(1.374)
Observations	468	70	469	469	469	469
R-squared	0.20	0.23	0.19	0.19	0.19	0.19

* significant at 10 percent.
** significant at 5 percent.
*** significant at 1 percent.
Notes: Robust standard errors, clustered by township, are given in parentheses. Graded schools dummy equals one if district has graded schools and zero otherwise. Subsidy per student is defined as spending per student minus annual tuition per student.
Sources: See the text.

APPENDIX TABLE 7
INTERGENERATIONAL INCOME ELASTICITY ESTIMATES WITH SCHOOLING MEASURE INTERACTIONS FOR SONS FROM RURAL TOWNSHIPS LIVING OUTSIDE OF TOWNS, SON'S LOG EARNINGS AS DEPENDENT VARIABLE

Schooling Measure:	Teacher-Student Ratio	Graded Teacher-Student Ratio	Classrooms per Square Mile	Graded Classrooms per Square Mile	Spending per Student	Taxes per Student
	(1)	(2)	(3)	(4)	(5)	(6)
Schooling measure	–11.046	–291.061***	–2.648*	–2.774	–0.069	–0.073
	(19.374)	(59.223)	(1.436)	(1.947)	(0.054)	(0.056)
Father's log earnings x schooling measure	1.528	46.903***	0.381*	0.419	0.009	0.009
	(2.815)	(5.829)	(0.222)	(0.310)	(0.008)	(0.009)
Father's log earnings	–0.056	–1.806***	–0.049	0.005	–0.055	–0.031
	(0.148)	(0.134)	(0.053)	(0.049)	(0.075)	(0.061)
Father's age	–0.080	0.222	–0.065	–0.071	–0.078	–0.079
	(0.063)	(0.894)	(0.067)	(0.065)	(0.061)	(0.061)
(Father's age)^2	0.001	–0.002	0.000	0.000	0.001	0.001
	(0.001)	(0.008)	(0.001)	(0.001)	(0.001)	(0.001)
Son's age	–0.075	2.664***	–0.111**	–0.099*	–0.104**	–0.105**
	(0.052)	(0.382)	(0.047)	(0.052)	(0.043)	(0.043)
(Son's age)^2	–0.018**	0.193***	–0.017**	–0.016**	–0.017**	–0.017**
	(0.008)	(0.030)	(0.008)	(0.008)	(0.008)	(0.008)
Son's age x Father's log earnings	0.017*	–0.388***	0.023***	0.021**	0.022***	0.022***
	(0.009)	(0.049)	(0.008)	(0.009)	(0.007)	(0.007)
(Son's age)^2 x Father's log earnings	0.002*	–0.029***	0.002	0.002	0.002	0.002
	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	9.554***	10.956	9.130***	8.890***	9.607***	9.496***
	(2.257)	(23.359)	(1.938)	(1.973)	(1.675)	(1.682)
Observations	175	26	178	178	178	178
R-squared	0.25	0.52	0.27	0.26	0.27	0.27

Notes: Robust standard errors, clustered by township, are given in parentheses. Subsidy per student is defined as spending per student minus annual tuition per student.
Sources: See the text.

APPENDIX TABLE 8
INTERGENERATIONAL INCOME ELASTICITY ESTIMATES WITH SCHOOLING MEASURE INTERACTIONS FOR SONS FROM RURAL TOWNSHIPS LIVING IN TOWNS (TOWNSHIP-LEVEL SCHOOL DATA), SON'S LOG EARNINGS AS DEPENDENT VARIABLE

Schooling Measure:	Teacher-Student Ratio	Graded Teacher-Student Ratio	Classrooms per Square Mile	Graded Classrooms per Square Mile	Spending per Student	Taxes per Student
	(1)	(2)	(3)	(4)	(5)	(6)
Schooling measure	–27.674*	–366.864	–1.057	–1.841**	–0.121***	–0.150***
	(14.466)	(223.175)	(1.193)	(0.770)	(0.042)	(0.051)
Father's log earnings x schooling measure	3.353	52.079	0.173	0.300**	0.018**	0.023***
	(2.051)	(29.397)	(0.191)	(0.116)	(0.007)	(0.007)
Father's log earnings	–0.153	–1.311**	–0.072	–0.071	–0.212	–0.224*
	(0.136)	(0.573)	(0.120)	(0.099)	(0.130)	(0.119)
Father's age	–0.056	0.005	–0.079	–0.085	–0.077	–0.078
	(0.067)	(0.137)	(0.072)	(0.071)	(0.072)	(0.071)
(Father's age)^2	0.000	–0.000	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Son's age	0.069	–0.309	0.063	0.046	0.028	0.023
	(0.104)	(0.228)	(0.108)	(0.111)	(0.102)	(0.108)
(Son's age)^2	0.003	–0.064**	0.004	0.003	0.007	0.005
	(0.010)	(0.026)	(0.011)	(0.011)	(0.010)	(0.010)
Son's age x Father's log earnings	–0.004	0.054	–0.004	–0.001	0.001	0.002
	(0.016)	(0.040)	(0.017)	(0.017)	(0.016)	(0.017)
(Son's age)^2 x Father's log earnings	–0.001	0.010*	–0.001	–0.001	–0.001	–0.001
	(0.002)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	9.716***	15.939**	9.508***	9.724***	10.483***	10.500***
	(2.244)	(6.608)	(2.249)	(2.211)	(2.340)	(2.193)
Observations	293	44	268	268	268	268
R-squared	0.20	0.28	0.20	0.20	0.21	0.21

* significant at 10 percent.
** significant at 5 percent.
*** significant at 1 percent.
Notes: Robust standard errors, clustered by township, are given in parentheses. Subsidy per student is defined as spending per student minus annual tuition per student.
Sources: See the text.

APPENDIX TABLE 9
INTERGENERATIONAL INCOME ELASTICITY ESTIMATES WITH SCHOOLING MEASURE INTERACTIONS FOR SONS FROM RURAL TOWNSHIPS LIVING IN TOWNS (TOWN LEVEL SCHOOL DATA), SON'S LOG EARNINGS AS DEPENDENT VARIABLE

Schooling Measure:	Teacher-Student Ratio	Graded Teacher-Student Ratio	Spending per Student	Taxes per Student
	(1)	(2)	(3)	(4)
Schooling measure	–2.747	–300.019	–0.034	–0.122**
	(10.127)	(171.821)	(0.061)	(0.055)
Father's log earnings x schooling measure	0.577	42.964	0.007	0.022***
	(1.422)	(24.389)	(0.009)	(0.008)
Father's log earnings	0.067	–0.787	0.012	–0.115
	(0.139)	(0.513)	(0.170)	(0.148)
Father's age	–0.069	–0.085	–0.048	–0.041
	(0.073)	(0.107)	(0.074)	(0.076)
(Father's age)^2	0.000	0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Son's age	–0.012	–0.416**	–0.009	–0.027
	(0.111)	(0.166)	(0.105)	(0.104)
(Son's age)^2	0.003	–0.039**	0.006	0.006
	(0.009)	(0.017)	(0.010)	(0.009)
Son's age x Father's log earnings	0.009	0.074**	0.009	0.012
	(0.017)	(0.029)	(0.016)	(0.016)
(Son's age)^2 x Father's log earnings	–0.001	0.006*	–0.001	–0.001
	(0.001)	(0.003)	(0.001)	(0.001)
Constant	8.476***	15.464***	8.004***	8.645***
	(2.365)	(4.574)	(2.420)	(2.495)
Observations	273	70	281	281
R-squared	0.20	0.36	0.20	0.21

* significant at 10 percent.
** significant at 5 percent.
*** significant at 1 percent.
Notes: Robust standard errors, clustered by township, are given in parentheses. Subsidy per student is defined as spending per student minus annual tuition per student.
Sources: See the text.

AVAILABILITY OF TOWNSHIP AND DISTRICT-LEVEL DATA

Within a report for a single county, the information is broken down by either township or school district. For most rural townships, information is given for the township as a whole. The number of subdistricts in the township (typically corresponding to the number of common schools in the district) is listed but all of the data are reported for the township as a whole, not for each individual subdistrict. In some cases, data for each subdistrict was listed separately but this was rare.

In cases where there were independent towns or villages within a township, these towns or villages would have their own independent school district with its own information listed in the schools report. Consequently, for individuals living in rural counties, if they live in the open country the schools report would provide data at the township level while if they lived in a town, the schools report would provide data at the school district level.

These independent school districts typically contained the graded schools if a rural township had graded schools. Individuals living outside of the town had access to these schools, but typically had to pay an out-of-district tuition fee (the amount of which is given in the schools report). Given that a person living outside of a town with graded schools has different educational resources available to them than a person living outside of a town without graded schools, I combine the township school data with the data for the independent districts within a township to construct an overall measure of the educational resources in a township.

For urban counties (containing the cities of Davenport, Des Moines, and Dubuque), there are a large number of independent districts listed in the reports creating the possibility that individuals could be matched to their school districts. However, in these urban counties, the federal census records do not have sufficiently detailed location information to determine which school district an individual lives in (the location information given is typically the township name and then the name of the larger city). Consequently, it is necessary to aggregate these district data into township level data for the urban individuals.

VARIABLES REPORTED AT THE TOWNSHIP AND INDEPENDENT DISTRICT LEVEL

The information in the reports for the individual townships or independent districts is divided into several sections: schools, teachers and pupils, general, teachers' fund, schoolhouse fund, and contingent fund. Listed below are the items reported by section. Item names are given exactly as they are printed in the annual reports. The spending per student measures were calculated by summing all of the payments listed in the various fund sections. Taxes per student measures were calculated by summing the amounts received from district taxes in the various fund sections.

Schools — number ungraded, number rooms in the graded schools, and average number months taught.

Teachers and pupils — number employed males, number employed females, average compensation per month males, average compensation per month females, number of persons between the ages of 5 and 21 years males, number of persons between the ages of 5 and 21 years females, number enrolled in each district, total average attendance in the whole district, and average cost of tuition per month for each pupil.

General — number of schoolhouses, value of schoolhouses, value of apparatus, number of volumes in the libraries, number of trees in thrifty condition on the schoolhouse sites, and number of schoolrooms in which effects of stimulants and narcotics are taught.

Teachers' fund (debit section) — on hand last report, received from district tax, received from semiannual apportionment, received by transfer from schoolhouse fund, received from other sources.

Teachers' fund (credit section) — paid teachers since last report, paid other districts for tuition, paid for other purposes, on hand.

Schoolhouse fund (debit section) — on hand last report, received from district tax, received from other sources.

Schoolhouse fund (credit section) — paid for schoolhouses and sites, paid on bonds and interest, paid for library books, transferred to other funds, paid for other purposes, on hand.

Contingent fund (debit section) — on hand last report, received from district tax, received from sale of textbooks and supplies, received from schoolhouse fund and other sources.

Contingent fund (credit section) — Paid for fuel, rent, repairs, insurance and janitors, paid secretaries and treasurers, paid for records and apparatus, paid for library books and dictionaries, paid for free textbooks, paid for textbooks and general supplies, and paid for other purposes, on hand.

VARIABLES AVAILABLE AT THE COUNTY LEVEL

In addition to the information listed by township or independent district in the reports, there is a section of additional details given for the county as a whole. Included in this section is a listing of textbooks selected by the County Board of Education and the contract prices for those textbooks, responses to questions about the condition of country schools, reports of the examinations of teachers, statistics on the blind, deaf, dumb, and feeble-minded students, listings of colleges, academies, and private schools, and a report on the teachers' normal institute.

COVERAGE AND PUBLIC AVAILABILITY OF THE SCHOOLS DATA

These reports of the superintendents of schools were published annually and are available on microfilm from the Iowa State Historical Library. Scanned copies of the reports for the year 1900 are available for all Iowa counties from the author. Fully transcribed versions of the reports for the year 1900 for the counties used in this article will be made available for download.

Appendix 3: Mismatches and Mobility Estimation

The process of linking the 1915 Iowa census and 1900 federal census presents the possibility that some father-son pairs are incorrectly matched. Relying on name, age, and birthplace rather than a truly unique identifier makes the possibility of mismatches unavoidable. While every effort has been made to maintain strict criteria for matches, including discarding observations for which multiple individuals met the match criteria, the possibility of mismatches still looms. This section of the appendix offers a brief discussion of the estimation issues these mismatches create and an assessment of how common mismatches would have to be to account for the estimated difference in mobility rates between the 1915 Iowa and modern mobility estimates.

In its simplest incarnation, this mismatch error could be characterized as a son being paired with his correct father with probability π and incorrectly with a father drawn at random from the population with probability $1-\pi$. The value of the father's income used in the intergenerational income elasticity regressions is then given by

$$y^*_{i,f} = \begin{cases} y_{i,f} + u_i, & \textit{with probability } \pi \\ \tilde{y}, & \textit{with probability } 1-\pi \end{cases}$$

where $y_{i,f}$ is the father's true income, u_i is a classical measurement error term, and $\tilde{y}$ is a randomly drawn income corresponding to some other father in the population. The distribution from which $\tilde{y}$ is drawn can be assumed to be the income distribution of the entire population. Alternatively, and perhaps more realistically, the income could come from a distribution of father's income conditional on the son's age. If the son and father are correctly matched, the income observation of the father will be his true income with some classical measurement error u_i having mean zero and uncorrelated with the true income. The measurement error for the mismatched fathers will be equal to the difference between the randomly drawn income and his true income: $\tilde{y} - y_{i,f}$. Letting ε_i represent the measurement error for any given individual i in the sample and assuming that $E(\tilde{y})$ is equal to $E(y_{i,f})$, the measurement error for a sample containing some mismatched individuals can be characterized as follows,

$$\varepsilon_i = \begin{cases} u_i, & \textit{with probability } \pi \\ \tilde{y} - y_{i,f}, & \textit{with probability } 1-\pi, \end{cases}$$

$$E(\varepsilon_i) = 0,$$

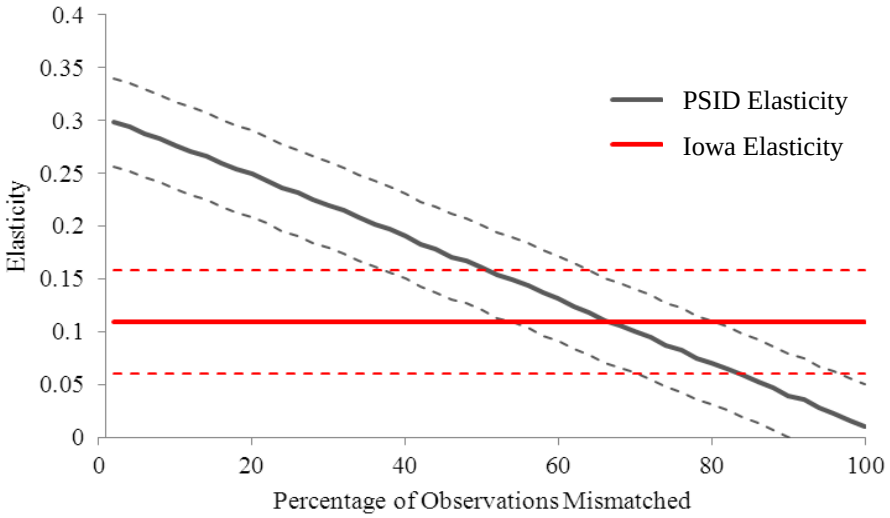
$$Cov(\varepsilon_i, y_{i,f}) = -(1-\pi)Var(y_{i,f}).$$

This negative correlation between the measurement error and the true value of the father's income implies that the measurement error introduced by mismatching is mean reverting.

The problem of mean-reverting measurement error is not uncommon in the labor literature, especially in studies using various measures of income. Bonggeun Kim and Gary Solon outline the dramatic effects that mean-reverting measurement error can have on the economic interpretation of wage data.[2] Arie Kapteyn and Jelmer Ypman specifically consider the case of mismatched administrative income data and show that in the simple case where the mismeasured income variable is the single independent variable in a linear regression and mismatches are drawn from the same population as the correctly matched individuals, the estimated coefficient on income is biased downwards by an amount proportional to the rate of mismatches.[3]

The bias introduced in the intergenerational income elasticity estimates by mismatched data is not easily characterized. The estimation equation includes both the mismatched variable and interactions of the mismatched variable with the correctly measured age of the son as regressors. Furthermore, the distribution from which a mismatch is drawn is dependent on the true value of the son's age. A final complication is that the likelihood of a son and father being mismatched may be correlated with

[2] Kim and Solon, "Implications."
[3] Kapteyn and Ypma, "Measurement."



APPENDIX FIGURE 1
INTERGENERATIONAL INCOME ELASTICITY ESTIMATES FOR THE 1915 IOWA SAMPLE AND THE 2001 PSID SAMPLE WITH RANDOM MISMATCHES IN THE PSID DATA

characteristics of the son and father including age, income, location, literacy, and so on that enter the income elasticity regressions either directly or through the error term. All of these factors make it difficult to assess how large a problem mismatches are for the Iowa data. Unlike the classical measurement error for the income variables discussed in the article common to both the Iowa and modern data, this source of error is specific to the Iowa sample and consequently could lead to a bias that generates the observed difference in intergenerational income elasticities between the Iowa sample and modern data even if the true elasticities are the same.

While there is no way to confidently state the number of mismatches in the linked Iowa sample, it is possible to introduce mismatches in a sample of observations from the Panel Study of Income Dynamics (PSID) and determine the level of mismatches required to obtain similar elasticity estimates for both the Iowa and PSID data. To generate random mismatches in the PSID data, an appropriate number of father-son observations are chosen at random to be mismatched. The father's income and age information is discarded. A new age for the father is randomly drawn from the distribution of father ages conditional on the son's age. The father's income is then randomly drawn from the distribution of income conditional on the father's newly chosen age. The new sample of individuals is then used to estimate the intergenerational income elasticity. The original data set is restored and then the entire process is repeated with new random number seeds.

Appendix Figure 1 depicts the results from simulating mismatches in the PSID sample. Mismatch rates of 2 percent to 100 percent are simulated, with 1,000 iterations of the mismatch and estimation procedure completed for each rate. The figure demonstrates that a mismatch rate approaching 50 percent would be required to account for the observed difference in 1915 and 2001 elasticities if the true elasticities are actually the same. Given that matches are identified by name, age, state of birth, father's state of birth, and mother's state of birth, the rate of mismatch in the linked Iowa data is likely far lower than this 50 percent figure.

REFERENCES

Goodman, Leo, and William Kruskal. “Measures of Association for Cross Classifications.” *Journal of the American Statistical Association* 49, no. 268 (1954): 732–64.

Kapteyn, Arie, and Jelmer Ypma. “Measurement Error and Misclassification: A Comparison of Survey and Administrative Data.” *Journal of Labor Economics* 25, no. 3 (2007): 513–51.

Kendall, Maurice. “A New Measure of Rank Correlation.” *Biometrika* 30, no. 1–2 (1938): 81–93.

Kim, Bonggeun, and Gary Solon. “Implications of Mean-Reverting Measurement Error for Longitudinal Studies of Wages and Employment.” *Review of Economics and Statistics* 87, no. 1 (2005): 193–96.