In the Name of the Father: Marriage and Intergenerational Mobility in the United States, 1850-1930.

Claudia Olivetti and M. Daniele Paserman

BU and NBER

November 2012

Introduction

- Why intergenerational mobility?
  - Key for understanding the importance of family background in determining economic outcomes.
  - In the US, tolerance for high inequality sometimes explained by the belief that mobility is also high (Alesina et al., 2004)
  - American exceptionalism?
- In fact, mobility in the US is among the lowest in OECD (Corak, 2011).
- Has it always been this way?

Introduction

- Ferrie (2005) and Long and Ferrie (2007, forthcoming) establish that mobility in the US was higher in the 19th Century.
- Typically, people look at the relationship between father's and son's economic standing.
- This misses part of the picture: daughters.
    - How is the *average* status of one generation (both sons and daughters) related to that of their parents?
    - Daughters can have an important role in transmitting status over to the next generation. Mobility over three generations?

Intergenerational Mobility Literature

- Vast literature based on modern linked longitudinal data sets (Solon, 1999, Black and Devereux, 2010):
  - Mostly focused on father/son correlations.
  - Few studies on father/son-in-law correlation: evidence that it is smaller than father/son correlation.
- Historical literature based on data obtained linking individuals by *first and last* name across Census decades (Ferrie, 2005).
  - Can construct father/son links and estimate father/son correlations.
  - **But** impossible to construct father/daughter links because daughters change last name upon marriage.

Our Contribution

- Develop methodology that allows estimation of intergenerational elasticities even without individually linked data.
  - Construct synthetic cohorts using information on first names
  - Can be applied equally to sons and daughters

- Investigate gender differentials in intergenerational mobility during 1850-1930 period by calculating elasticities in occupational income:
  - Father/son
  - Father/son-in-law

Basic Idea

- First names contain information about economic status.
- Suppose that in generation t high SES adults call their sons Adam, low SES call their sons Zachary.
- What happens in generation $t+1$? Are the Adam still higher SES than the Zacharys?
    - If yes, we would say that there is relatively little mobility. If no, high mobility.
- Nice feature of this methodology: can be applied just as easily to Abigails and Zoës.

Preview of the Findings

- Intergenerational elasticity between fathers and sons (η_{SON}) shows a 30% increase between 1870 and 1930.
  - Consistent with "the end of American exceptionalism" (Ferrie, 2005, Long and Ferrie, 2007, forthcoming).
- Intergenerational elasticity between fathers and sons-in-law (η_{SIL}):
  - Trend similar to that of η_{SON}, although timing of the increase slightly different.
  - By the end of the sample period is lower than η_{SON} in most specifications – similar to results for modern studies.
- Results likely driven by changes in the parameters of the income transmission process, not changes in the distribution of names.
- Results robust to different imputations of occupational income, name coding and treatment of: farmers, immigrants, child mortality.

Outline

- Introduction
- Illustrative Model
- Econometric Methodology
- Data
- Results
  - Benchmark
  - Simulations
- Discussion

An Illustrative Model of Marriage and Mobility

- Families containing 2 parents and 2 children: one male, one female.
- Only men work.
- Altruistic parents with consensus utility choose how to optimally allocate lifetime earnings, y_{t-1}, between own consumption and investment in children's human capital.
- Parents investment in children's human capital determines:
  - Son's earnings on the labor market, y_t
  - Daughter's *spouse* earnings through the marriage market, $y_{SIL,t}$
- Optimal human capital investment proportional to y_{t-1}.

Fathers and Sons

- Reduced form earnings equation:

$$\log y_t = \gamma_1 \log y_{t-1} + e_t + u_t$$
$$e_t = \lambda e_{t-1} + v_t$$

  - γ_1 = rate of return to human capital, $\gamma_1 \in (0,1)$
  - e_t = child's "endowment", $0 \leq \lambda < 1$, v_t i.i.d. with variance σ_v^2.
  - u_t = "labor market luck" i.i.d. with variance σ_u^2.

- Father/son *intergenerational elasticity*, estimated by OLS:

$$\eta_{SON} \equiv p\lim \frac{\widehat{Cov(y_t, y_{t-1})}}{\widehat{Var(y_{t-1})}} = \gamma_1 + \frac{\lambda\left(1-\gamma_1^2\right)}{(1+\gamma_1\lambda) + (1-\gamma_1\lambda)\left(\sigma_u^2/\sigma_e^2\right)}$$

Fathers and Sons-in-law

- Reduced form earnings equations:

$$\log y_{SIL,t} = \alpha_1 \log y_{t-1} + \theta e_t + \mu_t$$

  - α_1 = rate of return to female human capital in marriage market
  - μ_t = luck in the marriage market, i.i.d. with variance σ^2_μ.
  - θ : relative importance of family endowment for daughters.

- Father/son-in-law *intergenerational elasticity*:

$$\eta_{SIL} \equiv p \lim \frac{\widehat{Cov(y_{SIL,t}, y_{t-1})}}{\widehat{Var(y_{t-1})}}$$

$$= \alpha_1 + \theta \left(\frac{\lambda \left(1 - \gamma_1^2\right)}{\left(1 + \gamma_1 \lambda\right) + \left(1 - \gamma_1 \lambda\right) \left(\sigma_u^2 / \sigma_e^2\right)} \right)$$

Econometric Methodology

- With individually linked data:
  - Both y_{it} and y_{it-1} observed
  - Intergenerational elasticity obtained by regressing y_{it} on y_{it-1}
  - *Linked* estimator: $\hat{\eta}_{LINKED}$

- In our data it is impossible to link individuals across cross-sections t and $t-1$ but information on first names is available.

Econometric Methodology: Pseudo panel

- Define:
  - $\tilde{y}_{j,t-1}$ = average log earnings of *fathers* of children named j in Census year $t-1$
  - $\tilde{y}_{jt}$ = average log earnings (as *adults*) of children named j in Census year t

- *Pseudo-panel* estimator, $\hat{\eta}_{PSEUDO}$, obtained by:
  - Merging two cross sections by first names
  - Regressing $\tilde{y}_{jt}$ on $\tilde{y}_{jt-1}$ (weighted by name frequency)

Econometric Methodology: Pseudo panel

- Estimator is equivalent to Two Sample IV (2SIV or 2S2SLS)
    - First stage: Regress father's income on matrix of sons' first name dummies (sample 1)
    - Second stage: Regress son's income on fitted values from first stage (sample 2)
- Alternative interpretation: father's income is "generated regressor"
    - Actual father's income replaced by predicted income by son's first name.

Econometric Methodology: Pseudo panel

Key requirement:
Names carry information about socioeconomic status.

- If not:
    - Zero first stage
    - "Generated regressor" is just noise.

Data

- Main Analysis: full US Census 1% samples from IPUMS, 1850-1930.
    - Measure of income: median 1950 income in occupation (OCCSCORE).
- In addition: IPUMS Linked Representative samples 1850-1930 (available for father-son pairs but not for *married* daughters).
    - People observed in 1880 census (100% sample) and one other census between 1850 and 1930.

Summary Statistics

Table 1. Summary Statistics for Children's Names: 1850-1910

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Number of children ages 0-15	Number of distinct names	Mean number of observations per name	Percent of names that are singletons	Percent of children with unique names	Percent of children with names linked 20 years later	Share with top-50 name	Share of total variation in log earnings explained by between name variation
Year				*Males*				
1850	35,597	3,524	10.1	71.9	7.1	92.6	0.6919	0.1343
1860	48,114	4,083	11.8	70.5	6.0	93.7	0.6946	0.1108
1870	58,039	4,582	12.7	69.4	5.5		0.6978	0.1053
1880	75,004	6,589	11.4	69.4	6.1	92.9	0.6529	0.1119
1900	103,817	9,696	10.7	71.0	6.6	92.8	0.5638	0.1265
1910	117,612	9,818	12.0	69.5	5.8	94.1	0.5342	0.1256
				Females				
1850	34,272	3,442	10.0	71.9	7.2	92.4	0.6984	0.1357
1860	46,874	4,488	10.4	70.7	6.8	92.8	0.6573	0.1320
1870	55,739	5,206	10.7	71.1	6.6		0.6193	0.1356
1880	72,160	7,161	10.1	69.0	6.8	92.0	0.5475	0.1331
1900	101,516	10,081	10.1	70.9	7.0	92.3	0.4744	0.1526
1910	114,074	10,103	11.3	69.3	6.1	93.5	0.4726	0.1545

Ranking of Names

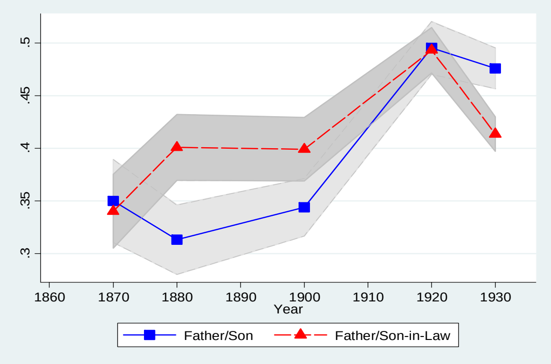
Table 2: Common Names Given to Children, Ranked by Mean Father's Occupational Income 1850-1930.

	1850	1860	1870	1880	1900	1910	1920	1930
				Males				
Rank:	*Most Prestigious*							
1	Edward	Walter	Harry	Paul	Donald	Abraham	Jerome	Irving
2	Frederick	Frank	Walter	Harry	Kenneth	Max	Irving	Frederick
3	Edwin	Willie	Herbert	Frederick	Harold	Nathan	Jack	Richard
4	Charles	Louis	Theodore	Ralph	Morris	Vincent	Nathan	Roger
5	Franklin	Fred	Edward	Philip	Max	Edmund	Abraham	Robert
	Least Prestigious							
1	Jesse	Levi	Jesse	Luther	Luther	Jessie	Willie	Jose
2	Hiram	Isaac	Franklin	Ira	Dewey	Otis	Loyd	Loyd
3	Isaac	Benjamin	Isaac	Isaac	Perry	Luther	Luther	Willie
4	Daniel	Andrew	Hiram	Willis	Virgil	Eddie	Jessie	Ervin
5	David	Jacob	Martin	Charley	Ira	Charley	Otis	Archie
				Females				
Rank:	*Most Prestigious*							
1	Emma	Ada	Bertha	Bessie	Dorothy	Eleanor	Betty	Jeanne
2	Alice	Kate	Jessie	Mabel	Marion	Marian	Jean	Jane
3	Anna	Lizzie	Grace	Helen	Helen	Dorothy	Jane	Carolyn
4	Isabella	Clara	Carrie	Ethel	Louise	Marion	Kathryn	Ann
5	Josephine	Fanny	Helen	Blanche	Marie	Virginia	Muriel	Joan
	Least Prestigious							
1	Sally	Amanda	Nancy	Nancy	Nancy	Sallie	Lela	Eula
2	Nancy	Nancy	Lucinda	Viola	Ollie	Addie	Maggie	Lorene
3	Lucinda	Rachel	Rebecca	Martha	Nannie	Ollie	Ollie	Dortha
4	Martha	Lucinda	Amanda	Rachel	Sallie	Mattie	Effie	Willie
5	Lydia	Martha	Martha	Amanda	Alta	Iva	Eula	Opal

(light grey) Exact name, nickname or alternative spelling appears more than once (most prestigious).
(dark grey) Exact name, nickname or alternative spelling appears more than once (least prestigious).

Benchmark Results

Figure 1: Father/Son and Father/Son in Law Elasticities in Occupational Income



Benchmark Results

Table 3. Intergenerational Elasticities in Occupational Income, 1850-1930.

	(1)	(2)	(3)	(4)	(5)
	1850-1870	1860-1880	1880-1900	1900-1920	1910-1930
Sample:					
Sons: baseline	0.3500	0.3133	0.3440	0.4953	0.4760
	(0.0239)	(0.0200)	(0.0166)	(0.0152)	(0.0118)
	[37077, 1182]	[50847, 1478]	[80255, 2234]	[109079, 3253]	[122468, 3720]
Son's Age 5-15	0.3286	0.3050	0.3574	0.4527	0.4199
	(0.0293)	(0.0243)	(0.0203)	(0.0173)	(0.0134)
	[24336, 984]	[32657, 1257]	[53629, 1860]	[76365, 2782]	[83920, 3257]
Married Sons	0.2868	0.3433	0.3805	0.4715	0.4428
	(0.0312)	(0.0260)	(0.0223)	(0.0178)	(0.0133)
	[17912, 891]	[24510, 1155]	[36521, 1641]	[57570, 2586]	[67137, 3051]
Sons in law: baseline	0.3402	0.4009	0.3992	0.4932	0.4136
	(0.0213)	(0.0191)	(0.0183)	(0.0131)	(0.0100)
	[23280, 976]	[30081, 1376]	[45804, 2063]	[68439, 2888]	[79314, 3326]
Daughter's Age 5-15	0.3440	0.3991	0.3918	0.5013	0.4186
	(0.0256)	(0.0232)	(0.0214)	(0.0152)	(0.0116)
	[17019, 839]	[22037, 1203]	[34712, 1825]	[52967, 2565]	[61308, 2979]
Sons in law 20-35	0.3283	0.4394	0.3860	0.4889	0.4143
	(0.0250)	(0.0224)	(0.0218)	(0.0151)	(0.0116)
	[15404, 840]	[20383, 1197]	[30533, 1712]	[46762, 2479]	[54600, 2885]
Sons: Individually linked data		0.4654	0.4751		
		(0.0175)	(0.0120)		
		3947	8847		

Basic findings

- Father/son intergenerational elasticity increases over time.
- Father/son-in-law elasticity also increases, but timing is slightly different.
    - Increase happens earlier, but η_{SIL} lower than η_{SON} at the end of the period.
    - Results almost identical when we make sons and sons-in-law samples comparable.
- Pseudo-panel estimator lower than individually-linked estimator by about 28-33%.

Robustness Checks

- Results are robust to:
  - Imputation of farmer's income (use 1901 wage distribution, exclude farmers, etc.).
  - Alternative measures of log occupational income (income rank, 1990 distribution, SEI).
  - Controls for age (both fathers and sons).

Can trends be explained by changing name distribution?

- η_{SON} goes from 0.31 to 0.48 between 1860-1880 and 1910-1930.
- Can this be driven by changes in name distribution?
- Numerical exercise:
  - Simulate income and name generating process
  - Set model parameters to match simulated moments to their data counterparts for 1860-1880.
  - How does the estimate of η_{SON} change as we vary the parameters governing:
    1. The name distribution?
    2. The income process?

Numerical Simulations

- Income generating process:

$$\log y_t = \gamma_1 \log y_{t-1} + e_t + u_t$$

  - $e_t = \lambda e_{t-1} + v_t$
  - $u_t \sim N\left(0, \sigma_u^2\right); \quad v_t \sim N\left(0, \sigma_v^2\right)$

- Name assignment process:

$$P\left(Name = j\right) = \frac{\exp\left(\delta_{CON,j} + \delta_{SES,j} e_{t-1}\right)}{\sum_j \exp\left(\delta_{CON,j} + \delta_{SES,j} e_{t-1}\right)}.$$

  - $\delta_{CON,j} \sim N\left(0, \sigma_{CON}^2\right); \quad \delta_{SES_j} \sim N\left(0, \sigma_{SES}^2\right).$
  - σ_{CON}^2 = concentration of names: high σ_{CON}^2, high concentration.
  - σ_{SES}^2 = sensitivity of names to SES

Numerical Simulations

- Generate population of $N = 500,000$ families, $J = 1,500$ names.
- Generate income and assign names.
- Create:
  - 10% individually linked father/son sample
  - 10% father-son pseudo-panel linked by first names
  - N and J chosen so that extracts match 1860 Census data.
- Estimate $\psi = \left(\gamma_1, \lambda, \sigma_u^2, \sigma_v^2, \sigma_{CON}^2, \sigma_{SES}^2\right)$ by SMM.

Moments and Parameters

Table 7. Moments and Parameters Used in the Simulations

Moments	Simulation	Data	Source
$Cov(y_t, y_{t-1})/V(y_{t-1})$	0.464	0.465	1860-1880 Linked sample
$V(y_{t-1})$	0.158	0.160	1860-1880 Linked sample
$Cov_{PS}(y_t, y_{t-1})/V_{PS}(y_{t-1})$	0.314	0.313	1860 and 1880 1% samples
$V_{PS}(y_{t-1})$	0.011	0.011	1860 1% sample
Share of top 50 names	0.695	0.695	1860 1% sample
R-squared	0.105	0.111	1860 1% sample

Distance minimizing parameters

γ_1	λ	σ^2_u	σ^2_v	σ^2_{CON}	σ^2_{SES}
0.421	0.191	0.092	0.031	7.833	5.958

Sensitivity to the name distribution

Table 8. The Effects of the Features of the Name Distribution on Estimated Elasticities
Simulation Results.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Concetration of the name distribution (σ^2_{con})	Socio-economic content of names (σ^2_{ses})						
	0	1	3	5.958	10	20	30
2.5	η=0.0345	0.1131	0.2301	0.3107	0.3735	0.4343	0.4662
	[share50= 0.3444]	[0.344]	[0.3437]	[0.3452]	[0.3468]	[0.3542]	[0.3651]
	(R^2=0.1078)	(0.1139)	(0.1269)	(0.1421)	(0.1592)	(0.1897)	(0.209)
5	0.0275	0.1073	0.2203	0.3087	0.3757	0.4385	0.4616
	[0.5526]	[0.5524]	[0.5521]	[0.5517]	[0.552]	[0.5542]	[0.5584]
	(0.0894)	(0.0967)	(0.1084)	(0.1232)	(0.1406)	(0.1718)	(0.1901)
7.833	0.0139	0.1160	0.2246	**0.3144**	0.3794	0.4494	0.4746
	[0.6976]	[0.6965]	[0.6958]	**[0.6952]**	[0.6949]	[0.6947]	[0.6972]
	(0.0713)	(0.0774)	(0.0898)	**(0.1053)**	(0.1215)	(0.1519)	(0.1716)
10	0.0146	0.1169	0.2324	0.3148	0.3890	0.457	0.48
	[0.7638]	[0.7638]	[0.7636]	[0.7623]	[0.7615]	[0.7609]	[0.761]
	(0.0605)	(0.0666)	(0.0774)	(0.0922)	(0.1098)	(0.138)	(0.1596)
15	0.0122	0.1209	0.2419	0.3385	0.4009	0.4703	0.4892
	[0.8444]	[0.8447]	[0.8438]	[0.8428]	[0.842]	[0.8408]	[0.8396]
	(0.0441)	(0.0498)	(0.0599)	(0.0736)	(0.09)	(0.1191)	(0.1394)

Can trends be explained by changing name distribution?

- To explain observed increase in η_{SON} we need massive increase in σ^2_{SES}, approximately from 2 to 20.
- This implies that R^2 in regression of father's income on son's name fixed effects also increases dramatically, from 0.12 to 0.20.
- In practice, R^2 constant around 0.12 over the period (Table 1)

Can this be driven by changes in the income process?

Table 9. The Effects of Changes in the Income Generating Process on Intergenerational Pseudo-Elasticities
Simulation Results.

	(1)	(2)	(4)	(5)	(6)	(7)
Persistence of income (γ_1):	Persistence of income shock (λ):					
	0	0.1	0.191	0.3	0.4	0.5
0.1	η=0.0502	0.1070	0.1543	0.2239	0.2931	0.3763
	[share50=0.6953]	[0.6952]	[0.6952]	[0.695]	[0.6956]	[0.6948]
	(R^2=0.105)	(0.1057)	(0.1081)	(0.1109)	(0.1168)	(0.1268)
0.2	0.1024	0.1591	0.2080	0.2796	0.3496	0.4340
	[0.6953]	[0.6952]	[0.6952]	[0.695]	[0.6956]	[0.6948]
	(0.1039)	(0.1053)	(0.1081)	(0.1115)	(0.1182)	(0.1292)
0.3	0.1518	0.2084	0.2591	0.3330	0.4039	0.4897
	[0.6953]	[0.6952]	[0.6952]	[0.695]	[0.6956]	[0.6948]
	(0.1022)	(0.1041)	(0.1074)	(0.1113)	(0.1186)	(0.1305)
0.421	0.2049	0.2613	**0.3144**	0.3915	0.4641	0.5522
	[0.6953]	[0.6952]	**[0.6952]**	[0.695]	[0.6956]	[0.6948]
	(0.0992)	(0.1016)	**(0.1053)**	(0.1097)	(0.1175)	(0.1303)
0.5	0.2331	0.2892	0.3438	0.4233	0.4975	0.5878
	[0.6953]	[0.6952]	[0.6952]	[0.695]	[0.6956]	[0.6948]
	(0.0967)	(0.0993)	(0.1031)	(0.1077)	(0.1156)	(0.1289)
0.6	0.2582	0.3137	0.3701	0.4526	0.5291	0.6236
	[0.6953]	[0.6952]	[0.6952]	[0.695]	[0.6956]	[0.6948]
	(0.0928)	(0.0956)	(0.0994)	(0.104)	(0.1118)	(0.1251)

- Much more likely that this is driven by real changes in the parameters governing the income process, γ_1 and λ.

What factors can explain the trends?

- Period under examination characterized by major economic and demographic changes.
  - Dramatic drop in fertility and family size.
  - Migration – international and internal.
  - Regional differences in industrialization and economic development.
  - Investments in public education.

Fertility?

- Increase in elasticity possible if fertility decline occurs earlier for high income group.
  - High income parents can divide same wealth among fewer children.
- Jones and Tertilt (2008): smooth fertilty transition for high-income groups, more abrupt for low-income groups.
- But the timing is off: increase in elasticity should have occurred at the end of the 19th Century.
- Directly controlling for number of siblings and fort birth order has no effect on the results (Table 10).

International Migration?

- Common belief that migration can serve as one of the main engines of social mobility.
- Age of mass migration: 1880-1920.
- But this implies that mobility should have *increased* during this period, contrary to what we observe
- Directly controlling for immigrant status (of both fathers and sons), has almost no effect on the coefficients (Table 11)

Internal Migration?

- Long and Ferrie (2012) argue that internal migration is responsible for high intergenerational mobility: a form of investment in children's human capital.
- Timing is more plausible: internal migration peaked in the middle of the 19th Century, then flat.
- But directly controlling for internal mobility has no effect on the estimates (Table 11).
    - Caveat: we may not be able to control properly for internal mobility, no measure of within-state mobility.

Regional Differences?

- Period of increase in intergenerational elasticity coincides with period of economic divergence between regions.
  - Northeast and Midwest complete industrial transition, South still agricultural and lags behind.
- If relatively low mobility across regions, regional differences in economic development could explain the decline in mobility.
- Controls for region fixed effects or state-level measures of development: upward trend in elasticity all but disappears (Table 12).

Regional Differences?

Table 12. Intergenerational Elasticities 1850-1930.
By Region of Birth.

	(1)	(2)	(3)	(4)	(5)
	1850-1870	1860-1880	1880-1900	1900-1920	1910-1930
			A: Fathers-Sons		
All	0.3500	0.3133	0.3440	0.4953	0.4760
	(0.0239)	(0.0200)	(0.0166)	(0.0152)	(0.0118)
Control for state of residence	0.2765	0.1943	0.2108	0.2746	0.2799
	(0.0228)	(0.0189)	(0.0156)	(0.0142)	(0.0111)
Control for indicators of economic develo|	0.2784	0.1975	0.2013	0.2633	0.2656
	(0.0228)	(0.0188)	(0.0156)	(0.0142)	(0.0110)
N, no. names (all)	[37077, 1182]	[50847, 1478]	[80255, 2234]	[109079, 3253]	[122468, 3720]
			B: Fathers-Sons in Law		
All	0.3402	0.4009	0.3992	0.4932	0.4136
	(0.0213)	(0.0191)	(0.0183)	(0.0131)	(0.0100)
Control of region of residence	0.2474	0.2947	0.2509	0.3199	0.2600
	(0.0205)	(0.0182)	(0.0175)	(0.0127)	(0.0099)
Control for indicators of economic develo|	0.2513	0.2988	0.2517	0.3177	0.2550
	(0.0204)	(0.0181)	(0.0174)	(0.0127)	(0.0098)
N, no. names (all)	[23280, 976]	[30081, 1376]	[45804, 2063]	[68439, 2888]	[79314, 3326]

Public Schooling?

- Increased investment in public schooling should increase mobility.
- Timing is inconsistent with observed trends.
- But analysis within regions does show that mobility was higher in regions with higher scholarization rates (Northeast and Midwest) than in the South. (Table 13).

Increase in returns to human capital

- Trends in η_{SON} consistent with improvements in men's labor market outcomes that increase γ_1:
  - Rise in returns to education (Goldin, 1999; Margo, 2000)
  - Improved men's career prospects (Cverk, 2011)
- Trends in η_{SIL}:
  - With positive assortative mating, increases in γ_1 also lead to increases in the return to human capital in the marriage market.
  - Also, imbalanced sex ratios induced by war and immigration could affect the returns to female human capital.

Conclusion

- Propose a new method for estimating intergenerational elasticities in the US in the late 19th-early 20th Century.
    - Applicable to both sons and daughters.
- Large increases in both father/son and father/son-in-law elasticity.
- Our preferred explanation: regional differences in economic development.

Thank you!

Robustness: Sensitivity to Farmer's Income Imputations

Table 4. Intergenerational Elasticities 1850-1930.
Sensitivity to Farmers' Income Imputations.

	(1)	(2)	(3)	(4)	(5)
	1850-1870	1860-1880	1880-1900	1900-1920	1910-1930
Log occupational income in:			**A: Fathers-Sons**		
1950	0.3500	0.3133	0.3440	0.4953	0.4760
	(0.0239)	(0.0200)	(0.0166)	(0.0152)	(0.0118)
1900	0.3502	0.3542	0.3823	0.4471	0.4436
	(0.0222)	(0.0189)	(0.0155)	(0.0121)	(0.0101)
1900, imputed farmer wage	0.3467	0.2879	0.3634	0.4660	0.4701
	(0.0284)	(0.0229)	(0.0196)	(0.0150)	(0.0127)
1950 ex. farmers	0.1899	0.1561	0.1463	0.2540	0.2922
	(0.0476)	(0.0359)	(0.0280)	(0.0322)	(0.0277)
1900 ex. farmers	0.2487	0.2075	0.2320	0.2992	0.2954
	(0.0460)	(0.0374)	(0.0329)	(0.0312)	(0.0259)
1950 ex. farmers (linked sample)		0.2860	0.3266		
		(0.0495)	(0.0340)		
N, no. of names: 1950	[37077, 1182]	[50847, 1478	[80255, 2234	109079, 3253	[122468, 3720]
N, no. of names: 1950 ex. Farmers	[26988, 741]	[36460, 943]	[65726, 1529	[92664, 2337]	[109830, 2845]

Robustness: Sensitivity to Farmer's Income Imputations

	B: Fathers-Sons in Law				
1950	0.3402	0.4009	0.3992	0.4932	0.4136
	(0.0213)	(0.0191)	(0.0183)	(0.0131)	(0.0100)
1900	0.3115	0.4229	0.4120	0.4900	0.4387
	(0.0203)	(0.0192)	(0.0182)	(0.0126)	(0.0100)
1900, imputed farmer wage	0.2509	0.3161	0.3166	0.4415	0.4221
	(0.0242)	(0.0205)	(0.0208)	(0.0146)	(0.0120)
1950 ex. Farmers	0.2150	0.2003	0.1802	0.3270	0.3220
	(0.0465)	(0.0303)	(0.0284)	(0.0288)	(0.0227)
1900 ex. Farmers	0.1986	0.2290	0.2224	0.3490	0.3744
	(0.0403)	(0.0316)	(0.0297)	(0.0289)	(0.0248)
N, no. of names: 1950	[23280, 976]	[30081, 1376	[45804, 2063	[68439, 2888]	[79314, 3326]
N, no. of names: 1950 ex. Farmers	[22586, 697]	[29344, 1004	[44917, 1547	[67488, 2313]	[78026, 2724]

Robustness: Sensitivity to income measures

Table 5. Intergenerational Elasticities 1850-1930.
Alternative Measures of Log Occupational Income.

	(1)	(2)	(3)	(4)	(5)
	1850-1870	1860-1880	1880-1900	1900-1920	1910-1930
			A: Fathers-Sons		
1950	0.3500	0.3133	0.3440	0.4953	0.4760
	(0.0239)	(0.0200)	(0.0166)	(0.0152)	(0.0118)
Rank regression	0.2896	0.3001	0.2879	0.3384	0.3510
(rank sample only)	(0.0152)	(0.0137)	(0.0112)	(0.0092)	(0.0080)
Rank regression	0.3161	0.3637	0.3621	0.4250	0.4033
(rank all working age males	(0.0165)	(0.0167)	(0.0137)	(0.0110)	(0.0088)
1990	0.2571	0.2069	0.2388	0.3585	0.4159
	(0.0260)	(0.0217)	(0.0187)	(0.0163)	(0.0140)
ERSCOR50	0.2870	0.3584	0.3427	0.4154	0.4005
	(0.0197)	(0.0203)	(0.0142)	(0.0115)	(0.0091)
SEI	0.2695	0.2979	0.3062	0.4597	0.4684
	(0.0204)	(0.0189)	(0.0157)	(0.0135)	(0.0118)
N, no. of names	[37077, 1182]	[50847, 1478]	[80255, 2234]	[109079, 3253]	[122468, 3720]

Robustness: Sensitivity to income measures

	B: Fathers-Sons in Law				
1950	0.3402	0.4009	0.3992	0.4932	0.4136
	(0.0213)	(0.0191)	(0.0183)	(0.0131)	(0.0100)
Rank regression (rank sample only)	0.3301	0.4405	0.3975	0.4275	0.3700
	(0.0163)	(0.0165)	(0.0143)	(0.0102)	(0.0085)
Rank regression (rank all working age males	0.3087	0.4429	0.4266	0.4902	0.4074
	(0.0157)	(0.0171)	(0.0160)	(0.0118)	(0.0092)
1990	0.2137	0.2685	0.2586	0.4418	0.3997
	(0.0229)	(0.0211)	(0.0218)	(0.0161)	(0.0128)
ERSCOR50	0.3031	0.4746	0.4228	0.4934	0.4105
	(0.0196)	(0.0218)	(0.0175)	(0.0123)	(0.0096)
SEI	0.1887	0.3243	0.3244	0.5097	0.4879
	(0.0200)	(0.0203)	(0.0213)	(0.0147)	(0.0124)
N, no. of names	[23280, 976]	[30081, 1376]	[45804, 2063]	[68439, 2888]	[79314, 3326]

Robustness: Controls for Age

Table 6. Intergenerational Elasticities 1850-1930.
Age Controls.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	1850-1870		1860-1880		1880-1900		1900-1920		1910-1930	
Variable:					**A: Fathers-Sons**					
Father's Income	0.3500 (0.0239)	0.3523 (0.0240)	0.3133 (0.0200)	0.3307 (0.0199)	0.3440 (0.0166)	0.3466 (0.0164)	0.4953 (0.0152)	0.4855 (0.0151)	0.4760 (0.0118)	0.4605 (0.0117)
Father's age		0.0096 (0.0093)		0.0009 (0.0080)		0.0289 (0.0060)		0.0196 (0.0055)		0.0183 (0.0043)
Father's age squared		-0.0001 (0.0001)		-0.0001 (0.0001)		-0.0004 (0.0001)		-0.0002 (0.0001)		-0.0002 (0.0001)
Son's age		0.1075 (0.0069)		0.0879 (0.0058)		0.1014 (0.0048)		0.0907 (0.0044)		0.1174 (0.0039)
Son's age squared		-0.0017 (0.0001)		-0.0013 (0.0001)		-0.0015 (0.0001)		-0.0014 (0.0001)		-0.0018 (0.0001)
N, no. of names	[37077, 1182]		[50847, 1478]		[80255, 2234]		[109079, 3253]		[122468, 3720]	

Robustness: Controls for Age

	B: Fathers-Sons in Law									
Father's Income	0.3402	0.3330	0.4009	0.3873	0.3992	0.3987	0.4932	0.4869	0.4136	0.4077
	(0.0213)	(0.0219)	(0.0191)	(0.0192)	(0.0183)	(0.0183)	(0.0131)	(0.0134)	(0.0100)	(0.0102)
Father's age		0.0062		0.0106		0.0016		0.0093		0.0046
		(0.0100)		(0.0085)		(0.0073)		(0.0059)		(0.0040)
Father's age squared		-0.0001		-0.0002		-0.0000		-0.0001		-0.0001
		(0.0001)		(0.0001)		(0.0001)		(0.0001)		(0.0000)
Son's age		0.0447		0.0328		0.0282		0.0179		0.0249
		(0.0029)		(0.0020)		(0.0018)		(0.0013)		(0.0013)
Son's age squared		-0.0006		-0.0004		-0.0004		-0.0002		-0.0003
		(0.0000)		(0.0000)		(0.0000)		(0.0000)		(0.0000)
N, no. of names	[23280, 976]		[30081, 1376]		[45804, 2063]		[68439, 2888]		[79314, 3326]	

Robustness: Sensitivity to name coding schemes

Table A5. Intergenerational Elasticities 1850-1930.
Sensitivity to Different Name Coding Schemes.

	(1)	(2)	(3)	(4)	(5)
	1850-1870	1860-1880	1880-1900	1900-1920	1910-1930
	A: Fathers-Sons				
Name concept:					
All	0.3500	0.3133	0.3440	0.4953	0.4760
	(0.0239)	(0.0200)	(0.0166)	(0.0152)	(0.0118)
Middle initials	0.3400	0.3112	0.3291	0.4189	0.4389
	(0.0230)	(0.0191)	(0.0156)	(0.0136)	(0.0111)
Nicknames	0.3673	0.3310	0.3412	0.4489	0.4268
	(0.0246)	(0.0207)	(0.0176)	(0.0159)	(0.0123)
Soundex codes	0.4212	0.4041	0.4771	0.5571	0.5530
	(0.0304)	(0.0250)	(0.0223)	(0.0184)	(0.0155)
N, no. names (All)	[37077, 1182]	[50847, 1478]	[80255, 2234]	[109079, 3253]	[122468, 3720]
N, no. names (M.I.)	[36685, 1419]	[50243, 1789]	[79227, 2676]	[107721, 3910]	[120706, 4605]
N, no. names (Nicknames)	[37172, 1138]	[50947, 1415]	[80315, 2107]	[109098, 3111]	[122501, 3581]
N, no. names (Soundex)	[39262, 887]	[54941, 995]	[84686, 1248]	[116154, 1595]	[130274, 1623]

Robustness: Sensitivity to name coding schemes

	B: Fathers-Sons in Law				
All	0.3402	0.4009	0.3992	0.4932	0.4136
	(0.0213)	(0.0191)	(0.0183)	(0.0131)	(0.0100)
Middle initials	0.3441	0.3619	0.3771	0.4249	0.3834
	(0.0208)	(0.0179)	(0.0170)	(0.0122)	(0.0096)
Nicknames	0.4360	0.4152	0.4135	0.4551	0.3882
	(0.0258)	(0.0204)	(0.0189)	(0.0140)	(0.0107)
Soundex codes	0.5907	0.5543	0.5570	0.6122	0.4944
	(0.0305)	(0.0257)	(0.0256)	(0.0176)	(0.0134)
N, no. names (All)	[23280, 976]	[30081, 1376]	[45804, 2063]	[68439, 2888]	[79314, 3326]
N, no. names (M.I.)	[22954, 1142]	[29682, 1644]	[45239, 2459]	[67637, 3496]	[77963, 4083]
N, no. names (Nicknames)	[23627, 945]	[30152, 1309]	[45814, 1958]	[68445, 2787]	[79322, 3227]
N, no. names (Soundex)	[25482, 566]	[32626, 705]	[48695, 855]	[72906, 1113]	[84541, 1198]

Robustness: 30-year Elasticities

Appendix Table 4: 30-year elasticities

	(1)	(2)	(3)	(4)
	1850-1880	1870-1900	1880-1910	1900-1930
Sample:				
Sons: baseline	0.2311	0.3108	0.3189	0.3871
	(0.0185)	(0.0165)	(0.0156)	(0.0123)
N, no. names	[37778, 1240]	[64972, 1645]	[83447, 2240]	[115713, 3313]
Sons in law: baseline	0.2913	0.3315	0.3726	0.4144
	(0.0189)	(0.0167)	(0.0174)	(0.0108)
N, no. names	[26311, 1093]	[43954, 1655]	[56494, 2105]	[87271, 3152]

Mechanisms: Fertility and Birth Order

Table 10. Fertility and Birth order

	(1)	(2)	(3)	(4)	(5)
	1850-1870	1860-1880	1880-1900	1900-1920	1910-1930
	A: Fathers-Sons				
Baseline	0.3500	0.3133	0.3440	0.4953	0.4760
	(0.0239)	(0.0200)	(0.0166)	(0.0152)	(0.0118)
Control for number of siblings	0.2836	0.2735	0.3444	0.5024	0.4740
	(0.0255)	(0.0214)	(0.0168)	(0.0157)	(0.0121)
Control for birth order	0.3277	0.2860	0.3433	0.4974	0.4642
	(0.0247)	(0.0207)	(0.0166)	(0.0154)	(0.0119)
N, no. names (baseline)	[37077, 1182]	[50847, 1478]	[80255, 2234]	[109079, 3253]	[122468, 3720]
	B: Fathers-Sons in Law				
Baseline	0.3402	0.4009	0.3992	0.4932	0.4136
	(0.0213)	(0.0191)	(0.0183)	(0.0131)	(0.0100)
Control for number of siblings	0.2920	0.3044	0.3949	0.4651	0.3815
	(0.0239)	(0.0210)	(0.0190)	(0.0140)	(0.0109)
Control for birth order	0.3289	0.3659	0.3962	0.4734	0.3951
	(0.0215)	(0.0197)	(0.0184)	(0.0133)	(0.0104)
N, no. names (baseline)	[23280, 976]	[30081, 1376]	[45804, 2063]	[68439, 2888]	[79314, 3326]

Mechanisms: Migration

Table 11. Immigration and Internal Migration

	(1)	(2)	(3)	(4)	(5)
	1850-1870	1860-1880	1880-1900	1900-1920	1910-1930
	A: Fathers-Sons				
Baseline	0.3500	0.3133	0.3440	0.4953	0.4760
	(0.0239)	(0.0200)	(0.0166)	(0.0152)	(0.0118)
Control for immigrant status	0.2992	0.2769	0.3247	0.4705	0.4659
	(0.0235)	(0.0198)	(0.0165)	(0.0151)	(0.0118)
Control for internal migrant status	0.2984	0.2766	0.3249	0.4708	0.4667
	(0.0235)	(0.0198)	(0.0164)	(0.0151)	(0.0118)
Control for immigrant status and father's		0.2367	0.2883	0.4420	0.4368
		(0.0195)	(0.0163)	(0.0150)	(0.0117)
Control for internal migrant status and father's		0.2328	0.2862	0.4387	0.4342
		(0.0195)	(0.0163)	(0.0150)	(0.0117)
N, no. names (baseline)	[37077, 1182]	[50847, 1478]	[80255, 2234]	[109079, 3253]	[122468, 3720]

Mechanisms: Migration

	B: Fathers-Sons in Law				
Baseline	0.3402 (0.0213)	0.4009 (0.0191)	0.3992 (0.0183)	0.4932 (0.0131)	0.4136 (0.0100)
Control for immigrant status	0.2720 (0.0211)	0.3625 (0.0190)	0.3676 (0.0182)	0.4773 (0.0131)	0.4086 (0.0101)
Control for internal migrant status	0.2722 (0.0211)	0.3619 (0.0190)	0.3640 (0.0182)	0.4733 (0.0131)	0.4043 (0.0100)
Control for immigrant status and father's		0.3254 (0.0188)	0.3122 (0.0180)	0.4433 (0.0131)	0.3815 (0.0101)
Control for internal migrant status and father's		0.3215 (0.0188)	0.3051 (0.0180)	0.4372 (0.0130)	0.3743 (0.0100)
N, no. names (baseline)	[37077, 1182]	[50847, 1478]	[80255, 2234]	[109079, 3253]	[122468, 3720]

Mechanisms: Within-Region Analysis

Table 13. Intergenerational Elasticities 1850-1930.
By Region of Birth.

	(1)	(2)	(3)	(4)	(5)
	1850-1870	1860-1880	1880-1900	1900-1920	1910-1930
			A: Fathers-Sons		
All	0.3500	0.3133	0.3440	0.4953	0.4760
	(0.0239)	(0.0200)	(0.0166)	(0.0152)	(0.0118)
Northeast	0.2948	0.2539	0.1677	0.2187	0.1918
	(0.0383)	(0.0337)	(0.0310)	(0.0279)	(0.0224)
Midwest	0.1499	0.2521	0.2677	0.2771	0.2701
	(0.0468)	(0.0368)	(0.0315)	(0.0279)	(0.0230)
South	0.4593	0.1591	0.2878	0.3081	0.3641
	(0.0564)	(0.0337)	(0.0311)	(0.0293)	(0.0229)
N, no. names (all)	[37077, 1182]	[50847, 1478]	[80255, 2234]	[109079, 3253]	[122468, 3720]
N, no. names (northeast)	[11461, 580]	[14846, 672]	[19327, 727]	[23818, 891]	[29959, 1040]
N, no. names (midwest)	[7091, 442]	[12713, 629]	[25372, 1039]	[35418, 1406]	[38069, 1589]
N, no. names (south)	[7709, 474]	[11481, 607]	[16570, 973]	[23490, 1558]	[30305, 1965]

Mechanisms: Within-Region Analysis

	B: Fathers-Sons in Law				
All	0.3402	0.4009	0.3992	0.4932	0.4136
	(0.0213)	(0.0191)	(0.0183)	(0.0131)	(0.0100)
Northeast	0.2014	0.2221	0.3111	0.2743	0.2100
	(0.0380)	(0.0382)	(0.0409)	(0.0333)	(0.0261)
Midwest	0.3471	0.3811	0.3289	0.3371	0.3015
	(0.0520)	(0.0353)	(0.0337)	(0.0238)	(0.0183)
South	0.3975	0.3303	0.3192	0.4649	0.3791
	(0.0478)	(0.0286)	(0.0306)	(0.0252)	(0.0178)
N, no. names (all)	[23280, 976]	[30081, 1376]	[45804, 2063]	[68439, 2888]	[79314, 3326]
N, no. names (northeast)	[6602, 448]	[8102, 559]	[9741, 602]	[12819, 769]	[16865, 923]
N, no. names (midwest)	[4877, 354]	[7883, 586]	[14957, 964]	[22529, 1340]	[24911, 1457]
N, no. names (south)	[5337, 408]	[7200, 587]	[10413, 926]	[16556, 1335]	[21104, 1625]