Estimating Intergenerational and Assortative Processes in Extended Family Data

M. Dolores Collado*
Ignacio Ortuño-Ortín**
Jan Stuhler**

*Universidad de Alicante
**Universidad Carlos III de Madrid

23rd August 2022

Introduction

Two central questions:

1. How strongly are economic advantages transmitted across generations?
    - How do assortative or other processes contribute to this persistence?
    - Can we explain the pattern across different kins within a unified model?
2. Which mechanisms can explain this persistence?
    - Parents, peers, places, nature vs. nurture

Introduction

Q:

1. How strongly are economic advantages transmitted across generations?
2. Which mechanisms can explain this persistence?

This paper:

- Proposes a "horizontal approach", chain-linking (distant) siblings *in-law*
  Yields many more and better-measured moments
- Considers a comparatively general intergenerational/assortative model
  Account for latent advantages along several dimensions
- Fits the pattern of inequality across many types of kins
- Remains fairly agnostic about the exact underlying mechanisms, but tests some common interpretations

... and show some newer work-in-progress.

Introduction

Q:

1. How strongly are economic advantages transmitted across generations?
2. Which mechanisms can explain this persistence?

This paper:

- Proposes a "horizontal approach", chain-linking (distant) siblings *in-law*
  Yields many more and better-measured moments
- Considers a comparatively general intergenerational/assortative model
  Account for latent advantages along several dimensions
- Fits the pattern of inequality across many types of kins
- Remains fairly agnostic about the exact underlying mechanisms, but tests some common interpretations

... and show some newer work-in-progress.

Related literature

1. Evidence on intergenerational and sibling correlations, e.g. Solon (1999), Björklund and Salvanes (2011), Jäntti and Jenkins (2014), and assortative mating, e.g. Ermisch et al (2006), Fernandez and Rogerson (2011), Greenwood et al (2014)
2. Evidence on multigenerational persistence, e.g. Clark (2014), Lindahl et al (2014), Braun and Stuhler (2018), Barone and Mocetti (2020), Adermon et al (2021), Clark (2021) + many recent
3. Evidence from population genetics, e.g. Rice, Cloninger and Reich (1978), Cavalli-Sforza and Feldman (1981) and genome-wide association studies, e.g. Lee et al (2018), Domingue et al (2014), Robinson et al. (2017)
4. Theoretical models with latent intergenerational transmission in economics and sociology, e.g. Duncan (1969), Goldberger (1972), Becker and Tomes (1979, 1986)

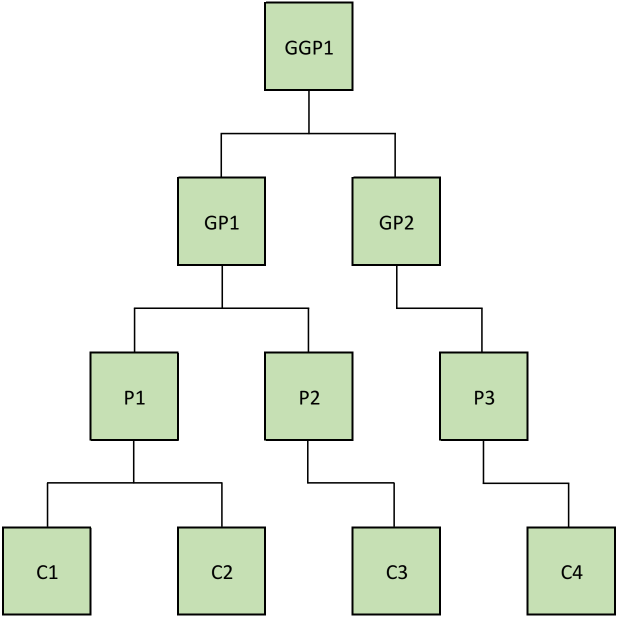
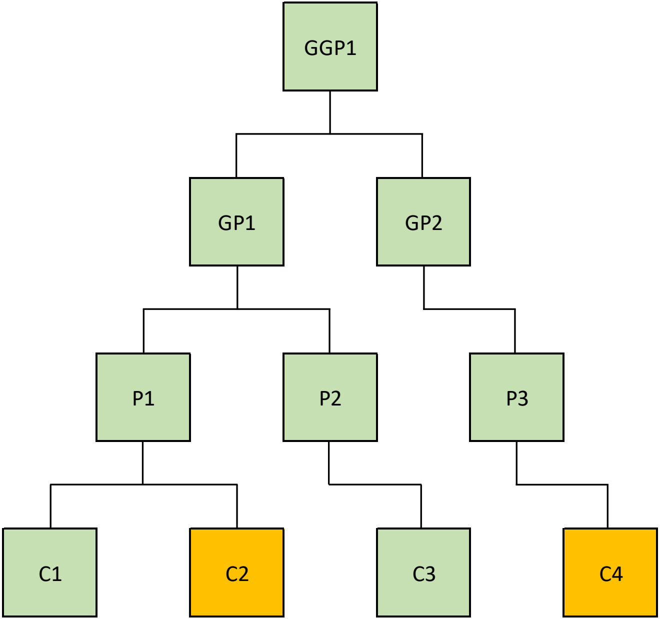
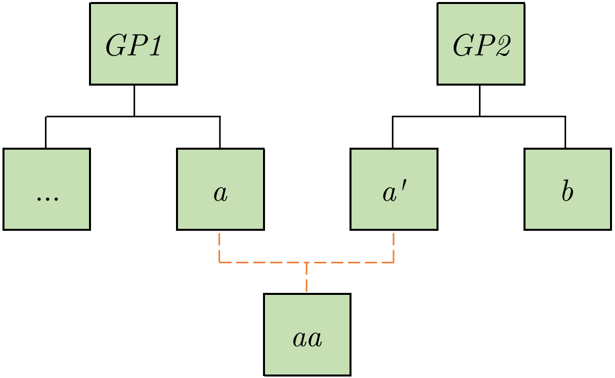

GGP1
GP1
GP2
P1
P2
P3
C1
C2
C3
C4

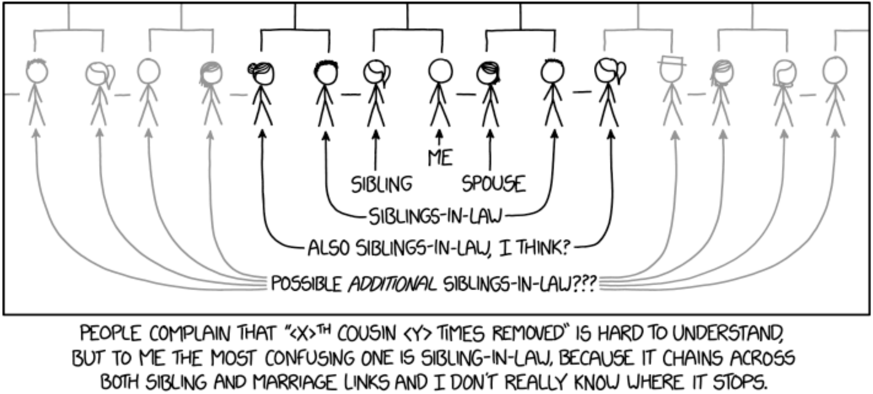
Figure: List of Kinship Types

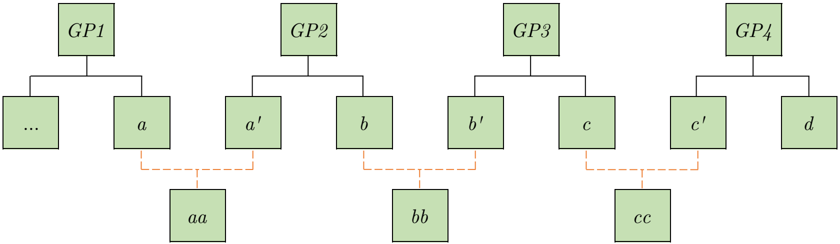
group	name	type	# correlations
a-a'	Spouses	Consanguine, horizontal	1
a'-b	Siblings	Consanguine, horizontal	3
aa-bb	Cousins	Consanguine, horizontal	10
aa-a	Parent-child	Consanguine, vertical	4
aa-b	Uncle/aunt-child	Consanguine, horizontal/vertical	8


GGP1
GP1
GP2
P1
P2
P3
C1
C2
C3
C4


GP1
GP2
...
a
a'
b
aa

Sibling-in-Law





Notes: Hypothetical family trees across child, parent, and grandparent (GP) generation. Consanguine (affine) relationships in black solid lines (orange dashed lines).

Figure: List of Kinship Types

group	name	class	# moments
a-a'	Spouses	consanguine, horizontal	1
a'-b	Siblings	consanguine, horizontal	3
aa-bb	Cousins	consanguine, horizontal	10
aa-a	Parent-child	consanguine, vertical	4
aa-b	Uncle/aunt-child	consanguine, horizontal & vertical	8
a-b	Siblings-in-law (degree 1)	affine, horizontal	4
a-b'	Spouse of sib-in-law (degree 1)	affine, horizontal	3
a'-c	Sibling of sib-in-law (degree 1)	affine, horizontal	4
aa–b'	Child-sibling in law (degree 1)	affine, horizontal & vertical	8
a-c	Siblings-in-law (degree 2)	affine, horizontal	8
...	...	...	
a-d	Siblings-in-law (degree 3)	affine, horizontal	16
...	...	...	...

- 205 moments (but some duplicates $\rightarrow$ 141 unique moments)
- Minimize difference between theoretical and sample moments ρ_i and $\hat{\rho}_i$, $min_{\{...\}} \sum_i w_i(\rho_i - \hat{\rho}_i)^2$

Swedish data: **Education**

Education (baseline)

Swedish register data:

- 1/3 of 1932-67 birth cohorts, plus their siblings, parents and children
- We use family links (outcomes) over three (two) generations

Education (years of schooling, demeaned by gender and cohort):

- 141 distinct moments (up to sibling-in-laws of 5th order)
- Correlations weighted by family size

Baseline specification:

- Include all correlations up to in-laws of 3rd order, exclude cousins
- 105 moments

Table: Estimated and Calibrated Moments in Swedish registers (Education)

Kinship			Data		Calibration	
group	#	name	# pairs	sample ρ	pred. ρ	% error
I	1	Spouses	413,062	0.491	0.489	-0.3
II	2	Brothers	387,028	0.432	0.431	-0.3
a'-b	3	Sisters	431,698	0.416	0.416	0.2
	4	Brother-Sister	800,127	0.375	0.376	0.4
III	5	Father-Son	396,304	0.380	0.381	0.2
aa-a	6	Father-Daughter	376,255	0.321	0.321	0.1
	7	Mother-Son	422,374	0.366	0.367	0.2
	8	Mother-Daughter	400,337	0.347	0.349	0.5
IV	9	Brothers in-law (MF-M)	602,262	0.302	0.296	-2.1
a-b	10	Brother-Sister in-law (FM-M)	578,269	0.296	0.304	2.7
	11	Brother-Sister in-law (MF-F)	650,127	0.298	0.307	3.0
	12	Sisters in-law (FM-F)	596,540	0.278	0.277	-0.2
V	13	Nephew-Uncle (BF)	280,067	0.254	0.249	-1.7
aa-b	14	Niece-Uncle (BF)	266,289	0.218	0.220	1.2
	15	Nephew-Uncle (BM)	312,019	0.241	0.238	-1.2
	16	Niece-Uncle (BM)	295,580	0.209	0.210	0.5
	17	Nephew-Aunt (SF)	285,618	0.234	0.229	-2.1
	18	Niece-Aunt (SF)	270,325	0.217	0.203	-6.7
	19	Nephew-Aunt (SM)	333,141	0.251	0.245	-2.3
	20	Niece-Aunt (SM)	316,625	0.234	0.218	-7.0
VI	21	Spouse of Sib-in-law (MF-FM)	252,232	0.252	0.246	-2.2
a-b'	22	Spouse of Sib-in-law (FM-MF)	226,795	0.229	0.232	1.1
	23	Spouse of Sib-in-law (MF-MF)	464,081	0.222	0.227	1.9
VII	24	Nephew-Aunt in-law (BF)	231,767	0.192	0.192	-0.3
aa-b'	25	Niece-Aunt in-law (BF)	221,287	0.172	0.171	-0.8
	26	Nephew-Aunt in-law (BM)	254,534	0.187	0.183	-2.2
	27	Niece-Aunt in-law (BM)	241,873	0.164	0.161	-2.0
	28	Nephew-Uncle in-law (SF)	227,403	0.190	0.188	-1.5
	29	Niece-Uncle in-law (SF)	215,068	0.163	0.167	2.2
	30	Nephew-Uncle in-law (SM)	264,524	0.197	0.198	0.8
	31	Niece-Uncle in-law (SM)	251,782	0.171	0.175	2.1
VIII	*32*	*Male Cousins (BF)*	*70,137*	*0.208*	*0.159*	*-23.8*
aa-bb	*33*	*Male Cousins (SM)*	*82,049*	*0.215*	*0.160*	*-25.4*
	34	*Male Cousins (SF)*	*156,747*	*0.202*	*0.152*	*-24.8*
	35	*Female Cousins (BF)*	*63,032*	*0.169*	*0.126*	*-25.5*
	36	*Female Cousins (SM)*	*73,649*	*0.197*	*0.124*	*-37.0*
	37	*Female Cousins (SF)*	*140,522*	*0.177*	*0.118*	*-33.2*

Kinship			Data		Calibration	
group	#	name	# pairs	sample ρ	pred. ρ	% error
	...					
XII	72	MF-FM-MF	299,602	0.138	0.135	-2.2
	73	FM-MF-MF	273,809	0.126	0.124	-1.4
XIII	74	M←MF-MF-MF	160,726	0.102	0.098	-3.5
aa-c'	75	M←MF-MF-FM	174,261	0.103	0.106	3.4
	76	M←MF-FM-MF	158,401	0.107	0.109	1.9
	77	M←MF-FM-FM	160,105	0.106	0.106	0.3
	78	M←FM-MF-MF	147,949	0.102	0.103	0.9
	79	M←FM-MF-FM	156,876	0.103	0.111	7.9
	80	M←FM-FM-MF	133,588	0.104	0.103	-1.0
	81	M←FM-FM-FM	131,756	0.101	0.101	-0.5
	82	F←MF-MF-MF	152,751	0.087	0.086	-0.1
	83	F←MF-MF-FM	165,828	0.091	0.093	3.1
	84	F←MF-FM-MF	151,100	0.094	0.095	1.7
	85	F←MF-FM-FM	153,065	0.089	0.093	4.9
	86	F←FM-MF-MF	140,585	0.093	0.092	-1.1
	87	F←FM-MF-FM	150,162	0.097	0.099	2.9
	88	F←FM-FM-MF	126,129	0.093	0.092	-1.2
	89	F←FM-FM-FM	124,968	0.085	0.090	5.3
XIV	*90*	*M←MF-MF-MF→M*	*84,025*	*0.094*	*0.082*	*-13.4*
aa-cc	*91*	*M←MF-MF-FM→M*	*100,261*	*0.101*	*0.086*	*-15.2*
	92	*M←MF-FM-MF→M*	*93,237*	*0.105*	*0.090*	*-13.9*
	93	*M←FM-MF-MF→M*	*80,486*	*0.097*	*0.085*	*-12.0*
	94	*M←MF-MF-MF→F*	*79,690*	*0.087*	*0.073*	*-16.7*
	95	*M←MF-MF-FM→F*	*95,733*	*0.094*	*0.076*	*-19.9*
	96	*M←MF-FM-MF→F*	*89,364*	*0.093*	*0.080*	*-13.6*
	97	*M←MF-FM-FM→F*	*95,020*	*0.095*	*0.075*	*-20.4*
	98	*M←FM-MF-MF→F*	*76,514*	*0.095*	*0.076*	*-20.0*
	99	*M←FM-MF-FM→F*	*89,054*	*0.088*	*0.079*	*-9.8*
	100	*M←FM-FM-MF→F*	*77,332*	*0.094*	*0.076*	*-19.0*
	101	*M←FM-FM-FM→F*	*80,067*	*0.082*	*0.072*	*-12.9*
	102	*F←MF-MF-MF→F*	*76,344*	*0.080*	*0.064*	*-20.3*
	103	*F←MF-MF-FM→F*	*91,080*	*0.090*	*0.066*	*-26.6*
	104	*F←MF-FM-MF→F*	*84,736*	*0.092*	*0.070*	*-23.4*
	105	*F←FM-MF-MF→F*	*72,410*	*0.082*	*0.068*	*-17.4*
XV	106	F-MF-MF-M	288,374	0.103	0.103	-0.2
a'-d	107	F-MF-MF-F	312,703	0.102	0.105	3.5

	38	*Male-Female Cousins (BF)*	*144,100*	*0.179*	*0.141*	*-20.9*
	39	*Male-Female Cousins (SM)*	*170,577*	*0.196*	*0.141*	*-28.2*
	40	*Male-Female Cousins (BM)*	*148,691*	*0.179*	*0.133*	*-25.5*
	41	*Male-Female Cousins (SF)*	*148,631*	*0.184*	*0.135*	*-26.5*
IX	42	F-MF-M	461,883	0.185	0.191	3.5
a′-c	43	F-MF-F	500,448	0.192	0.196	1.8
	44	F-FM-M	481,006	0.207	0.212	2.6
	45	M-MF-M	447,263	0.208	0.207	-0.6
X	46	MF-MF-M	362,409	0.156	0.157	0.8
a-c	47	MF-MF-F	393,579	0.156	0.161	3.4
	48	MF-FM-M	375,442	0.179	0.173	-3.4
	49	MF-FM-F	391,389	0.158	0.160	1.2
	50	FM-MF-M	353,470	0.160	0.161	0.8
	51	FM-MF-F	378,720	0.164	0.165	0.5
	52	FM-FM-M	341,316	0.157	0.160	2.1
	53	FM-FM-F	351,350	0.148	0.148	0.0
XI	54	M←MF-MF-M	202,632	0.119	0.127	6.7
aa-c	55	M←MF-MF-F	219,007	0.129	0.130	1.1
	56	M←MF-FM-M	192,819	0.134	0.140	4.6
	57	M←MF-FM-F	199,811	0.129	0.129	-0.3
	58	M←FM-MF-M	183,670	0.125	0.133	6.2
	59	M←FM-MF-F	196,631	0.132	0.136	3.6
	60	M←FM-FM-M	160,857	0.132	0.132	0.6
	61	M←FM-FM-F	164,528	0.122	0.122	-0.2
	62	F←MF-MF-M	192,818	0.104	0.112	7.1
	63	F←MF-MF-F	208,008	0.114	0.114	0.2
	64	F←MF-FM-M	183,929	0.116	0.123	6.1
	65	F←MF-FM-F	191,177	0.112	0.113	1.2
	66	F←FM-MF-M	175,507	0.110	0.119	8.0
	67	F←FM-MF-F	187,178	0.121	0.122	0.5
	68	F←FM-FM-M	151,606	0.112	0.118	5.5
	69	F←FM-FM-F	155,658	0.113	0.109	-3.7
XII	70	MF-MF-MF	278,938	0.122	0.122	-0.5
a-c′	71	MF-MF-FM	310,160	0.132	0.132	-0.6

	108	F-MF-FM-M	311,795	0.111	0.113	1.4
	109	F-MF-FM-F	162,928	0.099	0.104	5.4
	110	F-FM-MF-M	308,163	0.116	0.115	-1.5
	111	F-FM-MF-F	166,250	0.121	0.117	-2.6
	112	F-FM-FM-M	304,684	0.114	0.114	0.3
	113	M-MF-MF-M	278,416	0.117	0.111	-4.9
	114	M-MF-FM-M	149,478	0.131	0.122	-7.0
	115	M-FM-MF-M	143,733	0.115	0.112	-2.2
XVI	116	MF-MF-MF-M	230,313	0.087	0.085	-2.5
a-d	117	MF-MF-MF-F	251,223	0.087	0.087	0.2
	118	MF-MF-FM-M	248,811	0.097	0.093	-4.0
	119	MF-MF-FM-F	259,925	0.083	0.086	3.0
	120	MF-FM-MF-M	245,814	0.104	0.094	-9.7
	121	MF-FM-MF-F	265,220	0.101	0.096	-4.5
	122	MF-FM-FM-M	241,998	0.099	0.093	-6.0
	123	MF-FM-FM-F	248,449	0.084	0.086	1.7
	124	FM-MF-MF-M	224,873	0.091	0.087	-5.1
	125	FM-MF-MF-F	246,186	0.092	0.089	-3.9
	126	FM-MF-FM-M	237,791	0.100	0.095	-5.1
	127	FM-MF-FM-F	247,495	0.088	0.088	0.0
	128	FM-FM-MF-M	223,661	0.086	0.087	0.3
	129	FM-FM-MF-F	240,328	0.094	0.089	-5.8
	130	FM-FM-FM-M	213,155	0.091	0.086	-5.5
	131	FM-FM-FM-F	220,553	0.084	0.079	-4.9
XVII	*132*	*MF-MF-MF-MF*	*176,790*	*0.071*	*0.066*	*-8.3*
a-d′	*133*	*MF-MF-MF-FM*	*199,041*	*0.075*	*0.071*	*-6.0*
XVIII	*134*	*MF-MF-MF-MF-M*	*153,057*	*0.047*	*0.046*	*-3.3*
a-e	*135*	*FM-FM-FM-FM-F*	*144,976*	*0.054*	*0.043*	*-20.1*
XIX	*136*	*MF-MF-MF-MF-MF*	*117,473*	*0.047*	*0.035*	*-25.5*
a-e′	*137*	*MF-MF-MF-MF-FM*	*135,096*	*0.042*	*0.038*	*-9.9*
XX	*138*	*MF-MF-MF-MF-MF-M*	*106,844*	*0.031*	*0.025*	*-21.9*
a-f	*139*	*FM-FM-FM-FM-FM-F*	*100,871*	*0.043*	*0.023*	*-46.5*
XXI	*140*	*MF-MF-MF-MF-MF-MF*	*82,523*	*0.032*	*0.019*	*-40.0*
a-f′	*141*	*MF-MF-MF-MF-MF-FM*	*96,840*	*0.027*	*0.021*	*-24.4*

Figure: Baseline Fit in Swedish Registers

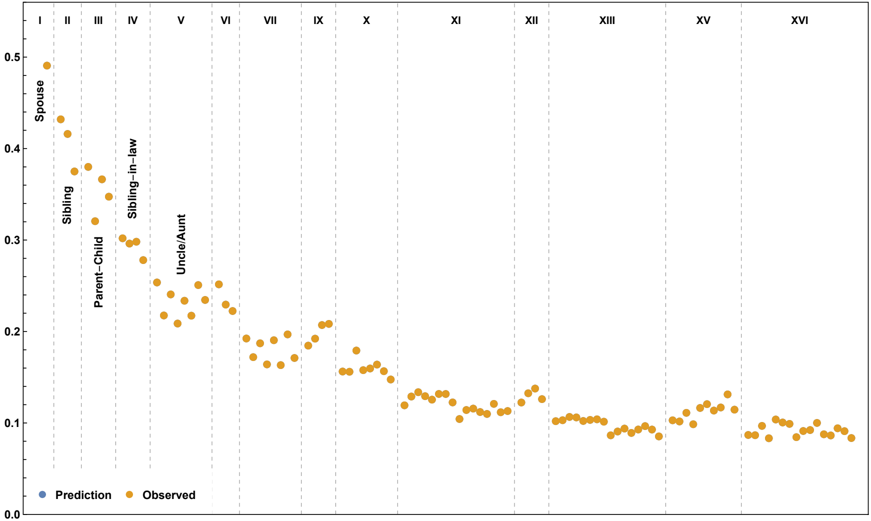
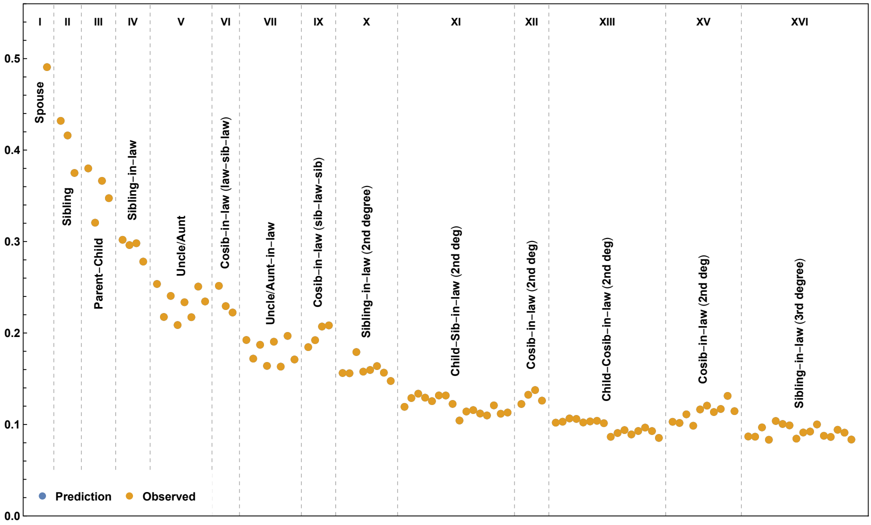


Figure: Baseline Fit in Swedish Registers



The model

Intergenerational: Outcome y_t^i of child i in generation t

$$y_t^k = \beta^k \tilde{y}_{t-1}^k + z_t^k + x_t^k + u_t^k$$
$$z_t^k = \gamma^k \tilde{z}_{t-1}^k + e_t^k + v_t^k$$

where $k = \{m, f\}$ denotes male or female children, and $\{\tilde{y}_{t-1}^k, \tilde{z}_{t-1}^k\}$ weighted parental averages,

$$\tilde{y}_{t-1}^k = \alpha_y^k y_{t-1}^m + (1 - \alpha_y^k) y_{t-1}^f$$
$$\tilde{z}_{t-1}^k = \alpha_z^k z_{t-1}^m + (1 - \alpha_z^k) z_{t-1}^f$$

Siblings: x_t^k and e_t^k are shared by siblings of the same gender, correlated between siblings of different genders.

Assortative mating

Assortative mating in both observed and latent variable, linear projections

$$z^f_{t-1} = r^m_{zz} z^m_{t-1} + r^m_{zy} y^m_{t-1} + w^m_{t-1} \quad (1)$$

$$y^f_{t-1} = r^m_{yz} z^m_{t-1} + r^m_{yy} y^m_{t-1} + \varepsilon^m_{t-1} \quad (2)$$

where w^f_{t-1} and ε^f_{t-1} are uncorrelated with z^f_{t-1} and y^f_{t-1}, and where $r^f_{sd}(s, d = y, z)$ are functions of correlations and std. dev. of z^f, z^m,y^f, y^m.

Model summary

The **baseline model** is comparatively general, allowing for:

1. Direct (β^k) and indirect (γ^k) transmission
2. Two-parent structure (*cannot* be reduced to one parent)
3. Assortative mating in two dimensions (in y_t and z_t)
4. Correlated shocks among siblings (x_t^k and e_t^k)
5. Gender asymmetries in all parameters

In total 21 unknown parameters.

Figure: Baseline Fit in Swedish Registers

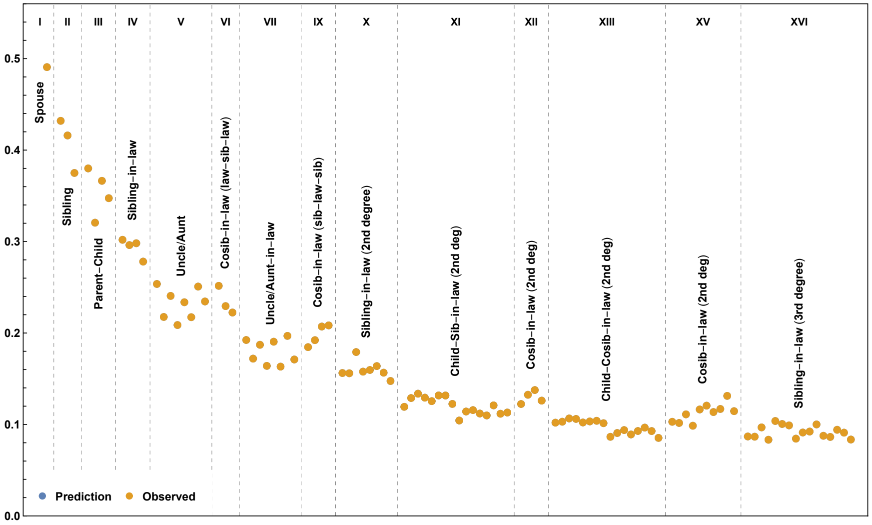
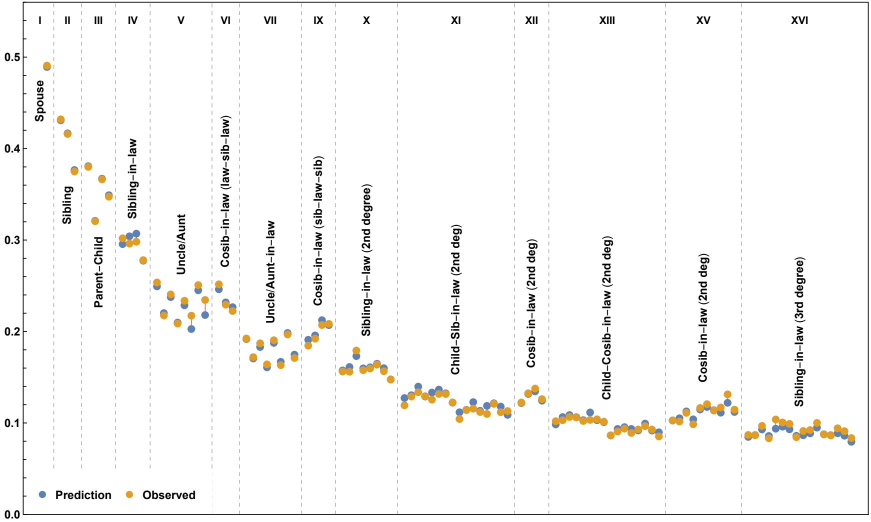


Figure: Baseline Fit in Swedish Registers



Baseline results

Intergenerational processes: ▸ Intergenerational

- Little direct ($\beta \approx 0.1$) and strong latent transmission ($\gamma \approx 0.6$), explaining 20-25% of variation in child education y and 35% in latent advantages z
- Conventional parent-child correlations: $< 14\%$

Sibling processes: ▸ Sibling

- Adding shared sibling influences, family background explains 42-43% of the variation in education and 67-82% in latent advantages.
- Conventional sibling correlations understate advantages shared by siblings

Assortative process: ▸ Assortative

- Strong sorting in latent factor, little additional sorting in education y
- Spousal are much more similar in latent than in educational advantages: $Corr(z^m, z^f) = 0.79$ vs. $Corr(y^m, y^f) = 0.49$

Baseline results

Intergenerational processes: ▸ Intergenerational

- Little direct ($\beta \approx 0.1$) and strong latent transmission ($\gamma \approx 0.6$), explaining 20-25% of variation in child education y and 35% in latent advantages z
- Conventional parent-child correlations: $< 14\%$

Sibling processes: ▸ Sibling

- Adding shared sibling influences, family background explains 42-43% of the variation in education and 67-82% in latent advantages.
- Conventional sibling correlations understate advantages shared by siblings

Assortative process: ▸ Assortative

- Strong sorting in latent factor, little additional sorting in education y
- Spousal are much more similar in latent than in educational advantages: $Corr(z^m, z^f) = 0.79$ vs. $Corr(y^m, y^f) = 0.49$

Baseline results

Intergenerational processes: ‣ Intergenerational

- Little direct ($\beta \approx 0.1$) and strong latent transmission ($\gamma \approx 0.6$), explaining 20-25% of variation in child education y and 35% in latent advantages z
- Conventional parent-child correlations: $< 14\%$

Sibling processes: ‣ Sibling

- Adding shared sibling influences, family background explains 42-43% of the variation in education and 67-82% in latent advantages.
- Conventional sibling correlations understate advantages shared by siblings

Assortative process: ‣ Assortative

- Strong sorting in latent factor, little additional sorting in education y
- Spousal are much more similar in latent than in educational advantages: $Corr(z^m, z^f) = 0.79$ *vs.* $Corr(y^m, y^f) = 0.49$

In-sample and out-of-sample fit

Good **in-sample** fit:

- Mean absolute error across 105 kinship types = 1.9 percent
- Fits vertical and horizontal, consanguine (“blood”) and affine (“in-law”) moments

Good **out-of-sample** fit:

- Similar results when reducing set of empirical moments from 105 to 35
- However, we understate the kinship correlations for cousins (measurement problem?) and very distant in-laws (model too simple?)

In the paper, but no time today:

- Can more restricted models fit the data? Restricted Models
- Other outcomes: Income, Swedish registers; Education, Spanish Census Other outcomes

In-sample and out-of-sample fit

Good **in-sample** fit:

- Mean absolute error across 105 kinship types = 1.9 percent
- Fits vertical and horizontal, consanguine (“blood”) and affine (“in-law”) moments

Good **out-of-sample** fit:

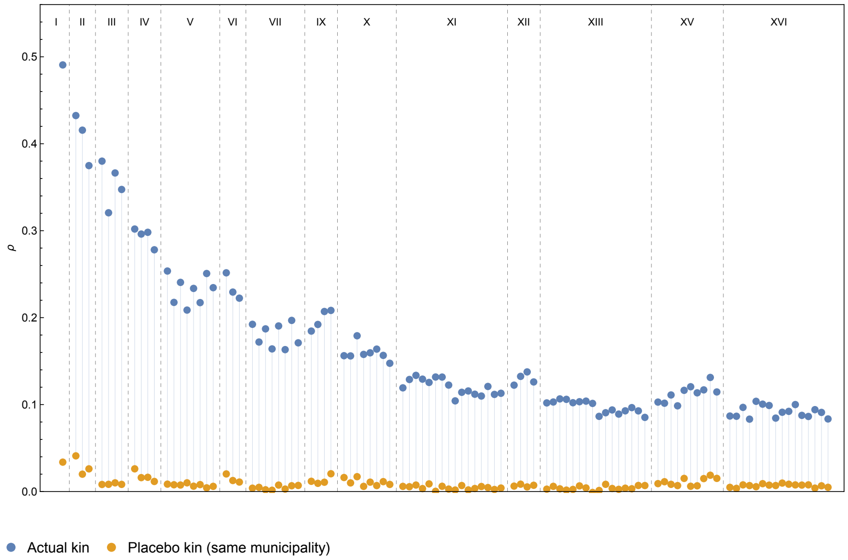
- Similar results when reducing set of empirical moments from 105 to 35
- However, we understate the kinship correlations for cousins (measurement problem?) and very distant in-laws (model too simple?)

In the paper, but no time today:

- Can more restricted models fit the data? ▸ Restricted Models
- Other outcomes: Income, Swedish registers; Education, Spanish Census ▸ Other outcomes

Swedish data: Mechanisms and Genetic Model(s)

Figure: Causal Pathways: Geography



Genetic and extended two-factor models

Are genes an important component of the latent advantages captured by our model? To address this question, we fit

1. Standard genetic model with sorting (nested by our baseline model)
2. Two-factor model (with latent genetic and sociocultural factors)

Results: Two Factor

Standard genetic model of genetic inheritance with assortative mating (Crow and Felsenstein, 1968) corresponds to our baseline model with restrictions:

- $\beta^m = \beta^f = 0$
- $\gamma^m = \gamma^f = 1$, $\alpha_z^m = \alpha_z^f = 0.5$, $\sigma(z^m) = \sigma(z^f)$
- Assortative mating in "phenotype" (i.e., education)
- Alternatively, assume assortative mating in "genotype" (manuscript)

Genetic and extended two-factor models

Are genes an important component of the latent advantages captured by our model? To address this question, we fit

1. Standard genetic model with sorting (nested by our baseline model)
2. Two-factor model (with latent genetic and sociocultural factors)
   ▸ Results: Two Factor

Standard genetic model of genetic inheritance with assortative mating (Crow and Felsenstein, 1968) corresponds to our baseline model with restrictions:

- $\beta^m = \beta^f = 0$
- $\gamma^m = \gamma^f = 1$, $\alpha_z^m = \alpha_z^f = 0.5$, $\sigma(z^m) = \sigma(z^f)$
- Assortative mating in "phenotype" (i.e., education)
- Alternatively, assume assortative mating in "genotype" (manuscript)

Figure: Using 105 Moments

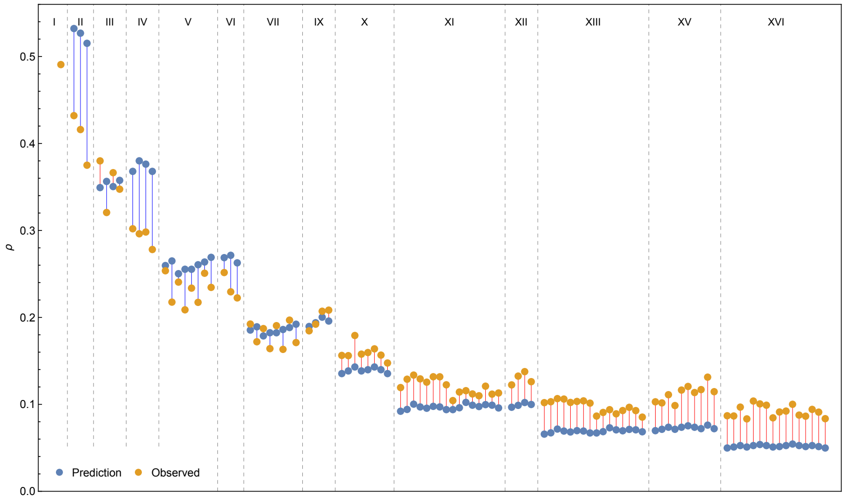
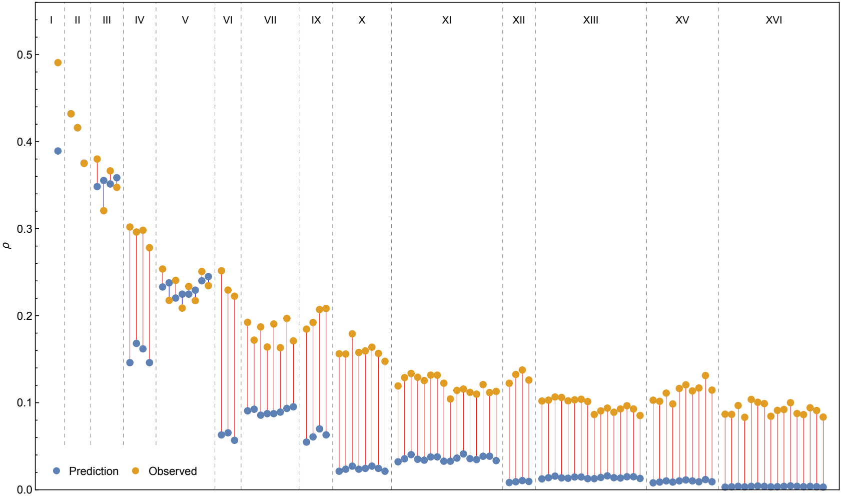


Figure: Using 15 Moments



Summary

Using an extensive set of kinship moments, we fit a model of intergenerational, sibling and assortative processes:

1. Parent-child correlations understate the transmission of socioeconomic advantages; distant kinships can be more informative
2. A "horizontal" approach that chain-links siblings in-law has important advantages ($\rightarrow$ more, more distant, better-measured moments)
3. Latent rather than observed characteristics play a key role in intergenerational, sibling and assortative processes
4. Strong intergenerational persistence, very strong assortative mating in latent advantages (even in Sweden!)
5. A purely genetic model cannot explain the kinship pattern; genes explain 7%, non-genetic advantages $>$ 30% of variance in education

Work-in-progress:

Work-in-progress:

1. Swiss population registers
   with Alexandre Jenni and Josef Zweimüller
2. Historical US Census data
   with Lola Collado and Ignacio Ortuño-Ortín

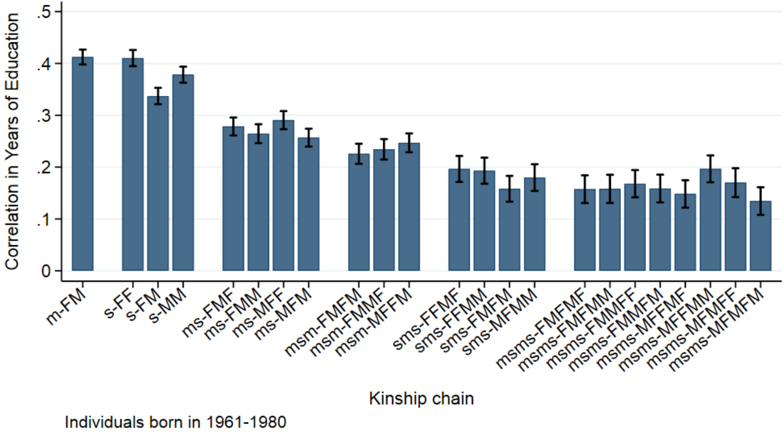
Swiss population registers

In work with Alexandre Jenni and Josef Zweimüller, we study kinship patterns in Switzerland.

Swiss population registers:

- We match population-wide registers with educational information from the Swiss "Strukturerhebung"
- Good match rate to parents, but match rate to grandparents <50%
- Measure education at age 30+ for cohorts born 1961-1980

Figure: Kinship moments in Swiss population registers (Education)



- We fit a simplified model with assortative mating only in z
- Implied spousal correlation in latent advantages $Corr(z^m, z^f)$ *is* 0.81

US historical Census data (MLP)

In work with Lola Collado and Ignacio Ortuño-Ortín, we use longitudinally linked US census data:

- Multigenerational Longitudinal Panel (IPUMS-MLP) contains linked records from the full-count 1850–1940 Censuses
- Educational information from the 1940 Census
- Longitudinal linking easier for males than females, affecting the representativeness of kinship moments involving the latter

We fit a simplified model

- No direct effects ($\gamma = 0$) and assortative mating only in latent factor z
- Use 15 moments based on brothers, sisters or mixed sibling links

Figure: Kinship moments in US Census (MLP)

Kinship type	Number of pairs	Sample correlation	Predicted correlation
Husband-Wife	26,142,900	0.615	0.599
Brothers	1,331,704	0.572	0.572
Father-Son	2,862,557	0.400	0.405
Father-Daughter	1,646,498	0.414	0.413
Mother-Son	3,514,216	0.410	0.413
Mother-Daughter	2,389,696	0.455	0.456
Brother in-law (Wife of the brother)	1,444,667	0.414	0.439
Uncle-Nephew (Brother of father)	34,006	0.314	0.297
Sisters in-law (Husbands are brothers)	494,422	0.395	0.391
Nephew-Wife of the uncle	29,224	0.282	0.264
Daughter-Father-in-law	1,266,734	0.340	0.360
Son-Father-in-law	418,537	0.362	0.363
Daughter-Mother-in-law	1,380,345	0.370	0.367
Son-Mother-in-law	691,363	0.401	0.400
Wife of the nephew-Uncle	18,850	0.262	0.264

Findings: US historical census data:

Preliminary findings based on the MLP data:

- Parent-child correlation in z is 0.60-0.67, depending on gender (vs. 0.51-0.60 in Swedish register data)
- Sibling correlation in z is 0.73-0.82 (vs. 0.68-0.83)
- Assortative mating $Corr(z^m, z^f)$ is 0.88 (vs. 0.79)
- Including shared sibling influences, family background explains 70-71% of the variation in education (vs. 42-43%)

Similar findings when using sister (e.g., AM=0.92) or brother-sister (AM=0.85) links.

The end.

Inter- vs. multigenerational mobility

Evidence on intergenerational transmission is primarily based on parent-child correlations. For example, slope coefficient in regression of child on parent income/occupation/education/...

$$y_{i,t} = \alpha + \beta y_{i,t-1} + \varepsilon_t$$

- Early U.S. literature (e.g. Becker and Tomes, 1986) found an intergenerational elasticity of income $\hat{\beta} \approx 0.15$.
- However, attenuation and lifecycle bias from use of short snapshots of income. More recent estimates for the U.S.: $\hat{\beta} \approx 0.4$ (Solon, 1990s) or $\hat{\beta} \approx 0.5$ (Mazumder, 2015).

Inter- vs. multigenerational mobility

How persistent are inequalities between more distant ancestors (e.g. grandparents $\longleftrightarrow$ grandchildren)?

Given the parent-child regression

$$y_{i,t} = \alpha + \beta y_{i,t-1} + \varepsilon_{i,t}, \tag{3}$$

we might iterate the coefficient β, i.e. the grandparent-grandchild correlation would be β^2 ($\rightarrow$ the iterated regression fallacy, Stuhler 2012)

However, recent multigenerational evidence find much higher persistence:

- Using surnames to proxy family links (e.g. Clark, 2014, Barone and Mocetti 2020)
- Using direct family links, but fewer generations
  Lindahl et al (2015), Braun and Stuhler (2018), Neidhöfer and Stockhausen (2018), Colagrossi, d'Hombres and Schnepf (2019), Adermon et al. (2020)

On the relation between inter- and multigenerational mobility

For illustration, consider the following simple latent factor model (with a one-parent one-child family structure):

$$y_{i,t} = \rho e_{i,t} + u_{i,t} \tag{4}$$
$$e_{i,t} = \lambda e_{i,t-1} + v_{i,t}, \tag{5}$$

where y and e are measured as trendless indices with mean zero and variance one:

- $y_{i,t}$: observed status in generation t of family i
- $e_{i,t}$: latent advantages or "endowments"
- $u_{i,t}$ and $v_{i,t}$: market and endowment luck

The latent factor model

- Given equations (4) and (5) the intergenerational coefficient equals

$$\beta_{-1} = \frac{Cov(y_t, y_{t-1})}{Var(y_{t-1})} = \rho^2\lambda \tag{6}$$

  and across three generations

$$\beta_{-2} = \frac{Cov(y, y_{t-2})}{Var(y_{t-2})} = \rho^2\lambda^2. \tag{7}$$

- The model can rationalize "excess persistence", as

$$\Delta = \beta_{-2} - (\beta_{-1})^2 = (1-\rho^2)\rho^2\lambda^2 > 0$$

  if latent $\neq$ observed status ($\rho < 1$).

Interpretation

Do these results matter?

- Contributes to the debate on capitalism and inequality (Friedman, Becker, Piketty, ...): Social mobility is much lower than what traditional parent-child estimates suggest.
- How to interpret prior comparative work based on parent-child evidence? Is mobility really higher in Canada than in the U.S., or higher in Sweden than in Spain?
- How much do policies and institutions affect *long-run* mobility across multiple generations (as compared to parent-child mobility in observable status?)
- Relates intergenerational with group-level inequality (Borjas 1992, Margo 2017). Our findings suggest that contrary to current beliefs, group-level inequalities might not be much more persistent than individual-level inequalities.

Outlook

Remaining questions:

1. Consider more alternative models, such as the “grandparent effect” model ...
2. How sensitive are our results to violations of the steady-state assumption?
3. To which degree can we abstract from vertical moments?

Intergenerational persistence

- Early literature (e.g. Becker and Tomes, 1986), estimates intergenerational elasticity of income

$$y_t = \beta y_{t-1} + \varepsilon_t$$

  with $\hat{\beta} \approx 0.15$ for U.S.

- However, these estimates turned out to be downward biased because of measurement error
  - Attenuation bias from classical measurement error (e.g. Atkinson 1980s, Solon, 1999) $\rightarrow \hat{\beta} \approx 0.4$ for U.S.
  - Lifecycle bias (e.g. Jenkins 1987, Nybom and Stuhler 2016, Mazumder 2016) $\rightarrow \hat{\beta} \approx 0.5$ for U.S. (?)

*Multi*generational persistence

More recent literature on persistence across multiple generations:

- High persistence of socioeconomic status on the **surname** level (e.g. Clark, 2014). For example, in historical data from Florence the average status of surnames still correlates across generations that are six centuries apart (Barone and Mocetti, 2019)
- Other studies observe direct family links, but fewer generations

Can be interpreted in **latent** transmission model in

$$y_t = \delta z_t + u_t$$
$$z_t = \gamma z_{t-1} + v_t$$

$\rightarrow \gamma \approx 0.75$ (Clark, 2014) or ≈ 0.6 (Braun and Stuhler, 2018)?

Panel A: Intergenerational Processes

Parameters:						
	β^m	β^f	γ^m	γ^f		
	0.144	0.129	0.664	0.565		
	σ^2_{ym}	σ^2_{yf}	σ^2_{zm}	σ^2_{zf}	σ^2_{um}	σ^2_{uf}
	4.648	4.465	2.070	1.560	1.978	2.329
	α_{ym}	α_{yf}	α_{zm}	α_{zf}		
	0.390	0.020	0.658	0.773		
Parent-child correlations in z:						
	Father-Son	Father-Dau	Mother-Son	Mother-Dau		
	0.586	0.600	0.527	0.508		
Ancestor correlations in y and z:						
	Father-Son	Grandf-...	GGrandf-...	GGGrandf-Son		
in y	0.381	0.209	0.121	0.071		
in z	0.586	0.343	0.201	0.118		

Panel A: Intergenerational Processes

Parameters:

	β^m	β^f	γ^m	γ^f		
	0.144	0.129	0.664	0.565		
	σ^2_{ym}	σ^2_{yf}	σ^2_{zm}	σ^2_{zf}	σ^2_{um}	σ^2_{uf}
	4.648	4.465	2.070	1.560	1.978	2.329
	α_{ym}	α_{yf}	α_{zm}	α_{zf}		
	0.390	0.020	0.658	0.773		

Parent-child correlations in z:

	Father-Son	Father-Dau	Mother-Son	Mother-Dau
	0.586	0.600	0.527	0.508

Ancestor correlations in y and z:

	Father-Son	Grandf-...	GGrandf-...	GGGrandf-Son
in y	0.381	0.209	0.121	0.071
in z	0.586	0.343	0.201	0.118

Panel A: Intergenerational Processes

Parameters:

β^m	β^f	γ^m	γ^f		
0.144	0.129	0.664	0.565		
σ^2_{ym}	σ^2_{yf}	σ^2_{zm}	σ^2_{zf}	σ^2_{um}	σ^2_{uf}
4.648	4.465	2.070	1.560	1.978	2.329
α_{ym}	α_{yf}	α_{zm}	α_{zf}		
0.390	0.020	0.658	0.773		

Parent-child correlations in z:

Father-Son	Father-Dau	Mother-Son	Mother-Dau
0.586	0.600	0.527	0.508

Ancestor correlations in y and z:

	Father-Son	Grandf-...	GGrandf-...	GGGrandf-Son
in y	0.381	0.209	0.121	0.071
in z	0.586	0.343	0.201	0.118

Panel B: Sibling Processes						
Parameters:						
	σ^2_{xm}	σ^2_{xf}	σ_{xmxf}	σ^2_{em}	σ^2_{ef}	σ_{emef}
	0.178	0.246	0.069	0.657	0.711	0.625
Variance Shares:						
in y	3.8%	5.5%	1.5%	14.1%	15.9%	13.7%
in z	-	-	-	31.7%	45.6%	34.8%
Sibling correlations in z:						
	Brothers	Sisters	Brother-Sister			
	0.678	0.824	0.711			

Panel C: Assortative Processes

Parameters:

r^{m}_{zz}	r^{m}_{zy}	r^{m}_{yz}	r^{m}_{yy}	$\sigma^{2}_{\omega m}$	$\sigma^{2}_{\epsilon m}$
0.663	-0.008	0.696	0.143	0.673	2.919
r^{f}_{zz}	r^{f}_{zy}	r^{f}_{yz}	r^{f}_{yy}		
0.747	0.112	0.662	0.249		

Spousal correlations in y and z:

ρ_{ymyf}	ρ_{zmzf}	ρ_{ymzf}	ρ_{zmyf}
0.489	0.754	0.540	0.580

Panel D: Variance Decomposition

%	y	z	x	Cov(y,z)
Male	0.013	0.445	0.038	0.038
Female	0.016	0.349	0.055	0.030

Back

Figure: Out-of-Sample Fit in Swedish Registers

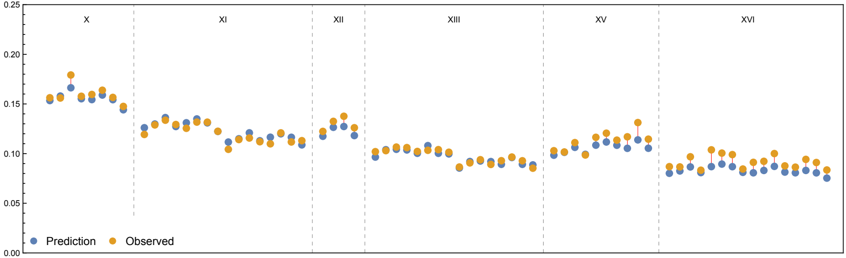


Figure: Restricted model without direct transmission ($\beta = 0$)

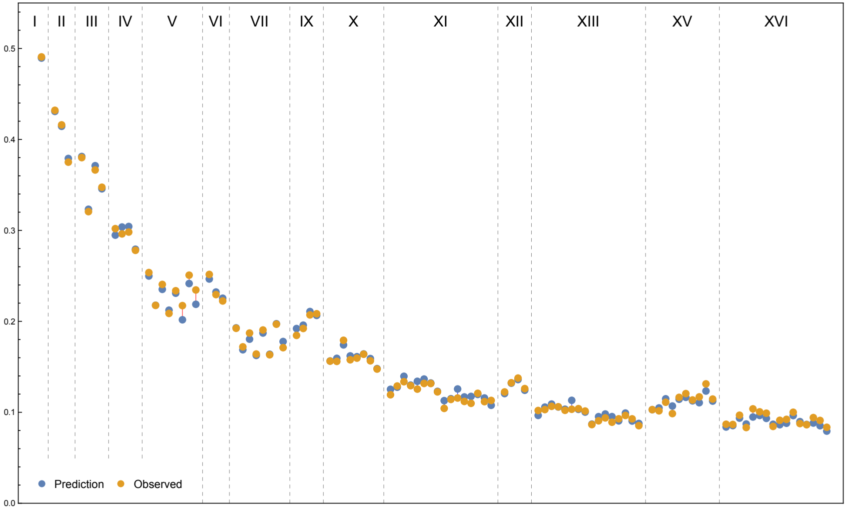


Figure: Restricted model without latent transmission ($\gamma = 0$)

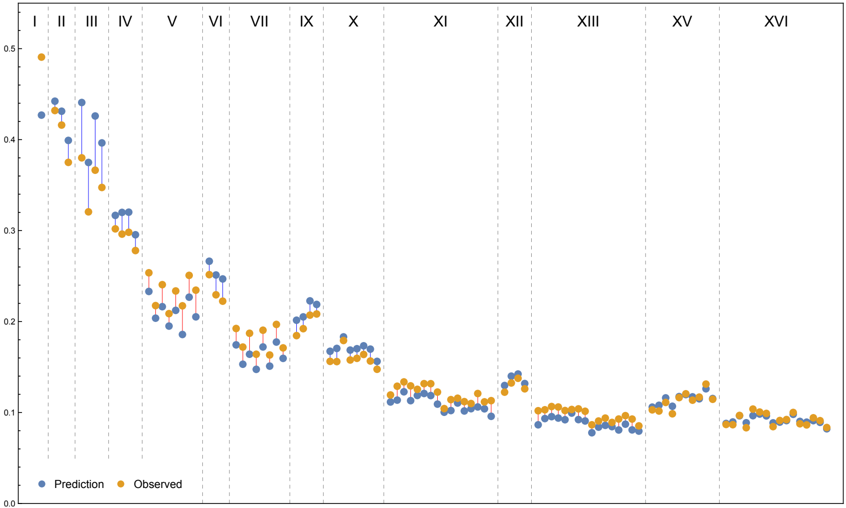


Figure: Restricted model without shared sibling component

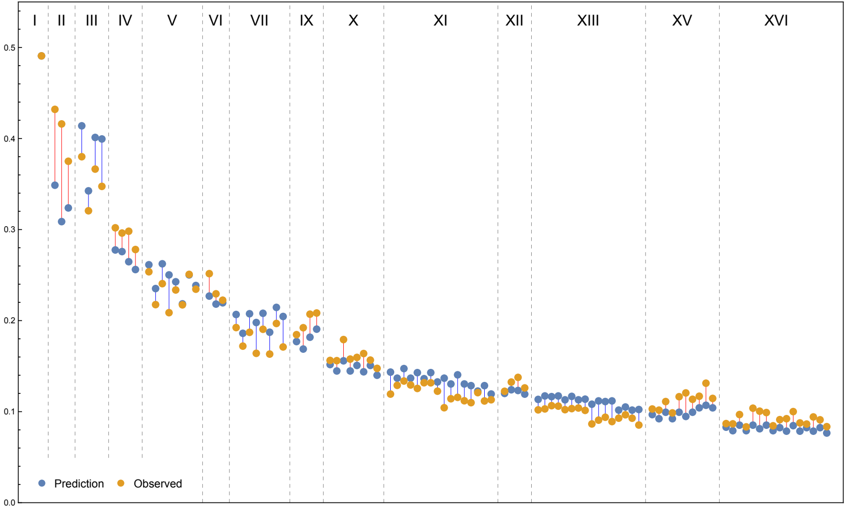
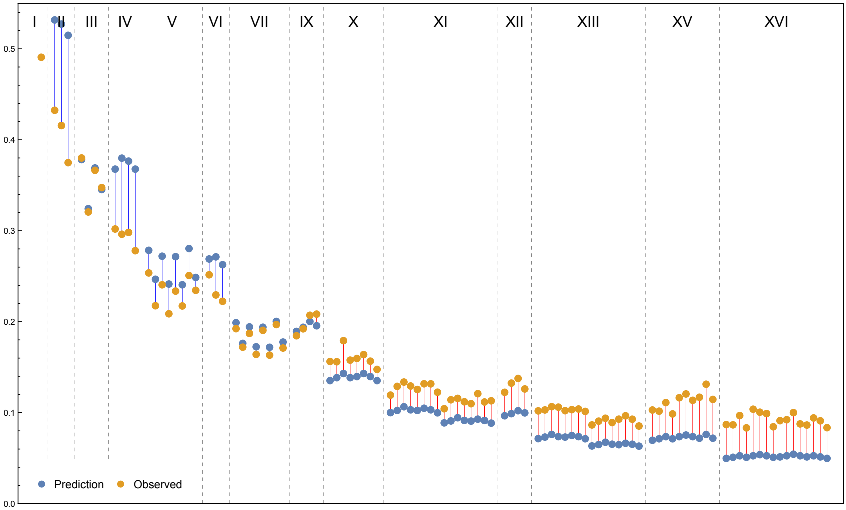


Figure: Assortative mating only in observables



Back

The standard genetic model and *body height*

Consider **body height** (from military enlistment). Interesting reference point:

1. Genes known to be important (Silventoinen, 2003). Can this be captured by our approach?
2. Height only observed for males. Can intergenerational, sibling and assortative process for females be *indirectly* identified?

We observe

- 37 (male) moments
- import spousal and sister correlation from external sources (Price and Vandenberg, 1980)Results for body height:

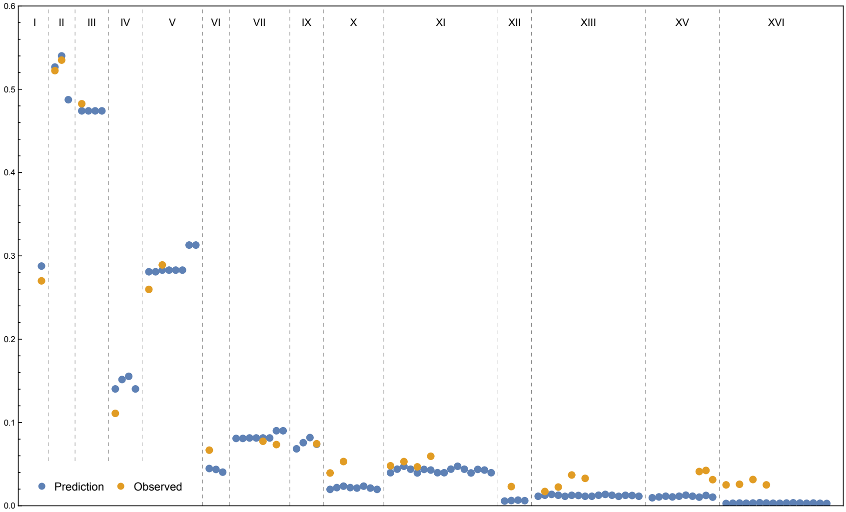
Genetic model fits body height comparatively well, with much fewer parameters than the baseline model

Table: Moments in Swedish Registers (Height)

Kinship			Data		Calibration	
group	#	name	# pairs	ρ	general	genetic
I	1	Spouses	External	0.270	0.270	0.288
II	2	Brothers	112,549	0.522	0.521	0.527
	3	Sisters	External	0.535	0.535	0.540
	4	Brother-Sister			0.458	0.487
III	5	Father-Son	46,441	0.483	0.488	0.474
	6	Father-Daughter			0.254	0.474
	7	Mother-Son			0.605	0.474
	8	Mother-Daughter			0.716	0.474
IV	9	Brothers in-law (MF-M)	135,006	0.111	0.114	0.140
	10	Brother-Sister in-law (FM-M)			0.198	0.152
	11	Brother-Sister in-law (MF-F)			0.158	0.155
	12	Sisters in-law (FM-F)			0.294	0.140
V	13	Nephew-Uncle (BF)	52,618	0.260	0.268	0.281
	...					
	15	Nephew-Uncle (BM)	66,270	0.289	0.295	0.283
VI	21	Spouse of Sib. in-law (MF)	44,034	0.067	0.053	0.045
VII	28	Nephew-Uncle in-law (SF)	28,901	0.077	0.075	0.081
	30	Nephew-Uncle in-law (SM)	39,634	0.073	0.077	0.090
VIII	32	Male Cousins (BF)	21,153	0.160	0.148	0.166
	33	Male Cousins (SM)	21,937	0.188	0.189	0.188
	34	Male Cousins (SF)	45,689	0.167	0.152	0.170
IX	45	M-MF-M	99,219	0.074	0.077	0.074
X	46	MF-MF-M	59,544	0.039	0.039	0.020
	48	MF-FM-M	63,798	0.053	0.045	0.024
XI	54	M←MF-MF-M	46,771	0.048	0.044	0.040

Kinship			Data		Calibration	
group	#	name	# pairs	ρ	general	genetic
	...					
	56	M←MF-FM-M	37,029	0.053	0.054	0.047
	58	M←FM-MF-M	43,068	0.047	0.053	0.039
	60	M←FM-FM-M	29,437	0.059	0.058	0.043
XII	71	MF-MF-FM	39,288	0.023	0.032	0.006
XIII	75	M←MF-MF-FM	29,311	0.017	0.027	0.013
	77	M←MF-FM-FM	21,488	0.022	0.029	0.013
	79	M←FM-MF-FM	26,207	0.037	0.037	0.013
	81	M←FM-FM-FM	16,828	0.033	0.036	0.011
XIV	90	M←MF-MF-MF→M	24,274	0.025	0.034	0.023
	91	M←MF-MF-FM→M	27,032	0.043	0.031	0.025
	92	M←MF-FM-MF→M	27,089	0.040	0.039	0.025
	93	M←FM-MF-MF→M	25,681	0.036	0.044	0.023
XV	113	M-MF-MF-M	50,330	0.041	0.035	0.010
	114	M-MF-FM-M	25,234	0.042	0.041	0.012
	115	M-FM-MF-M	29,208	0.031	0.029	0.010
XVI	116	MF-MF-MF-M	30,183	0.025	0.027	0.003
	118	MF-MF-FM-M	30,538	0.026	0.031	0.003
	120	MF-FM-MF-M	36,879	0.032	0.025	0.003
	122	MF-FM-FM-M	33,478	0.025	0.031	0.003
XVII	133	MF-MF-MF-FM	19,846	0.016	0.025	0.001
XVIII	134	MF-MF-MF-MF-M	16,980	0.020	0.021	0.000
XIX	137	MF-MF-MF-MF-FM	11,286	0.009	0.019	0.000
XX	138	MF-MF-MF-MF-MF-M	10,309	0.015	0.016	0.000
XXI	141	MF-MF-MF-MF-MF-FM	6,873	0.044	0.015	0.000

Figure: Fit for Height (Swedish register, genetic model)



Extended two-factor model

Extended two-factor model:

- Decompose z into genetic and non-genetic (''cultural'') factors z^G and z^C

$$y_t^k = \beta^k \widetilde{y}_{t-1}^k + z_t^{G,k} + z_t^{C,k} + x_t^k + u_t^k,$$

where

$$z_t^{G,k} = \frac{z_{t-1}^{G,m} + z_{t-1}^{G,f}}{2} + v_t^{G,k}$$

follows standard genetic model (Crow and Felsenstein, 1968).

- We do not need to impose that "environments" of parents and offspring are independent a $z_t^{C,k}$ captures shared environments.

Extended two-factor model

Extended two-factor model: ▸ Back

- Genetic factor explains $\sim 7\%$ of variance in years of schooling
  Broadly in line with heritability estimates from genome-wide association studies (e.g. Lee et al. 2018) but below twin estimates ($\rightarrow$ "missing heritability")
- Sociocultural factor explains $> 30\%$ of variance in years of schooling
- Only negligible correlation between genetic and latent non-genetic advantages ($\hat{\rho}_{z_c, z_g} = 0.03$).
- Little assortative mating in genes ($\hat{\rho}_{z_g^m, z_g^f} = 0.02$)
  Consistent with Domingue et al. (2014), Yengo et al. (2018)

Spanish data: **Education**

Education (Spain)

Education (years of schooling, demeaned by gender and cohort):

- Census from Cantabria
- 65 distinct moments: spouse, parent-child, siblings, nephew/niece-uncle/aunt, sibling-in-law up to second order
- # of moments should in principle suffice (-> robustness test in Swedish registers)

Main results:

- Results qualitatively similar as in Sweden
- But more persistence across all three dimensions: intergenerational, siblings, assortative
- Parent-child correlation in $z \approx 0.8$, spousal correlation ≈ 0.9

Figure: Fit in Spanish Census (Education)

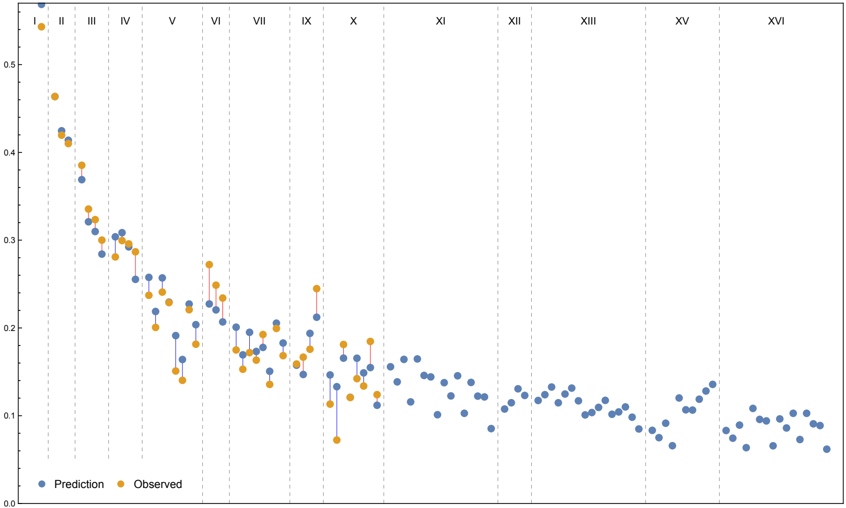


Table: Calibrated Parameters in Spanish Census

Panel A : Intergenerational Processes							
β^m	β^f	γ^m	γ^f	σ^2_{ym}	σ^2_{yf}	σ^2_{zm}	σ^2_{zf}
0.027	0.111	0.915	0.842	13.579	13.213	6.525	2.783
a_{ym}	a_{yf}	a_{zm}	a_{zf}	σ^2_{um}	σ^2_{uf}		
0.746	0.856	0.586	0.127	5.159	7.001		
Ancestor correlations in y and z:							
	Father-Son	Fr.-Daughter		Mother-Son	Mother-Daughter		
in z	0.760	0.827		0.732	0.883		
	Father-Son	Grandfr.-Son		GGrandfr.-Son	GGGrandfr-Son		
in y	0.369	0.271		0.205	0.156		
in z	0.760	0.594		0.451	0.343		
Panel B : Sibling Processes							
	σ^2_{xm}	σ^2_{xf}	σ_{xmxf}	σ^2_{em}	σ^2_{ef}	σ_{emef}	
	1.648	2.643	2.087	0.559	0.001	0.018	
Sibling correlations in y and z:							
	Brothers	Sisters	Mixed		Brothers	Sisters	Mixed
in y	0.464	0.425	0.414		0.674	0.784	0.666
Panel C : Assortative Processes							
r^{m}_{zz}	r^{m}_{zy}	r^{m}_{yz}	r^{m}_{yy}	$r^{\ddagger}_{zz}$	$r^{\ddagger}_{zy}$	$r^{\ddagger}_{yz}$	$r^{\ddagger}_{yy}$
0.732	-0.139	0.418	0.356	1.291	0.083	0.576	0.441
Spousal correlations in y and z:							
ρ_{ymyf}	ρ_{zmzf}	ρ_{ymzf}	ρ_{zmyf}				
0.569	0.903	0.483	0.549				
Panel D : Variance Decomposition of y							
	$\widetilde{y}_{t-1}$	$\widetilde{z}_{t-1}$	$\sigma_{\tilde{y}\tilde{z}}$	$\Sigma_{\tilde{y},\tilde{z}}$	x_t	e_t	$\Sigma_{\tilde{y},\tilde{z},x,e}$
male	0.1%	28.3%	1.8%	30.1%	12.1%	4.1%	46.4%
female	1.1%	16.5%	4.8%	22.5%	20.0%	0.0%	42.5%

Back

	Main parameters										
	Original	Mean	St.dev	Min	P05	P25	Median	P75	P95	Max	IQR
$\sigma^2_{z^m}$	2.070	1.965	0.197	1.175	1.585	1.860	1.997	2.101	2.235	2.468	0.240
$\sigma^2_{z^f}$	1.560	1.459	0.187	1.024	1.141	1.313	1.481	1.602	1.738	2.039	0.289
$\rho_{z^m z^f}$	0.754	0.726	0.049	0.561	0.638	0.691	0.734	0.763	0.794	0.844	0.071
$\rho_{z^m y^f}$	0.580	0.570	0.017	0.474	0.535	0.563	0.574	0.582	0.590	0.601	0.019
$\rho_{y^m z^f}$	0.540	0.527	0.024	0.480	0.487	0.508	0.530	0.546	0.564	0.594	0.038
$\rho_{y^m y^f}$	0.489	0.489	0.007	0.476	0.478	0.483	0.489	0.495	0.500	0.502	0.012
γ^m	0.664	0.663	0.015	0.604	0.637	0.653	0.663	0.674	0.688	0.711	0.021
γ^f	0.565	0.564	0.015	0.510	0.538	0.553	0.564	0.574	0.589	0.618	0.022
β^m	0.144	0.162	0.034	0.069	0.113	0.138	0.157	0.181	0.228	0.289	0.044
β^f	0.129	0.146	0.029	0.085	0.110	0.125	0.138	0.163	0.203	0.270	0.038
$\sigma^2_{x^m}$	0.178	0.175	0.038	0.019	0.113	0.147	0.176	0.204	0.233	0.288	0.057
$\sigma^2_{x^m}$	0.246	0.249	0.034	0.152	0.195	0.222	0.249	0.274	0.303	0.352	0.052
$\sigma_{x^m x^f}$	0.069	0.065	0.031	0.000	0.016	0.042	0.065	0.089	0.115	0.163	0.047
$\sigma^2_{e^m}$	0.657	0.675	0.035	0.573	0.625	0.650	0.670	0.694	0.742	0.858	0.044
$\sigma^2_{e^f}$	0.711	0.720	0.023	0.637	0.686	0.705	0.718	0.734	0.763	0.843	0.029
$\sigma_{e^m e^f}$	0.625	0.642	0.030	0.564	0.602	0.621	0.637	0.657	0.698	0.788	0.036
α_{z^m}	0.658	0.641	0.053	0.440	0.550	0.607	0.643	0.680	0.724	0.825	0.073
α_{y^m}	0.390	0.407	0.085	0.003	0.251	0.355	0.418	0.469	0.527	0.691	0.114
α_{z^f}	0.773	0.748	0.054	0.518	0.647	0.715	0.754	0.787	0.826	0.891	0.072
α_{y^f}	0.020	0.082	0.102	0.000	0.000	0.000	0.011	0.164	0.277	0.407	0.164

	Long-run correlations			
	y (male)	y (female)	z (male)	z (female)
Parent	0.381	0.349	0.586	0.508
Mean	0.381	0.349	0.574	0.498
Std. dev.	0.005	0.005	0.022	0.018
Grand-Parent	0.209	0.163	0.343	0.263
Mean	0.204	0.161	0.331	0.255
Std. dev.	0.010	0.004	0.024	0.016
Great-Grand-Parent	0.121	0.082	0.201	0.137
Mean	0.116	0.080	0.191	0.131
Std. dev.	0.010	0.004	0.020	0.012
GGreat-Grand-Parent	0.071	0.043	0.118	0.071
Mean	0.067	0.041	0.110	0.067
Std. dev.	0.008	0.003	0.015	0.008

	Sibling correlations	
	CorrY	CorrZ
Brothers	0.431	0.678
Mean	0.431	0.700
Std. dev.	0.006	0.040
Sisters	0.417	0.824
Mean	0.417	0.866
Std. dev.	0.006	0.079
Brother-sister	0.377	0.711
Mean	0.377	0.741
Std. dev.	0.005	0.050

Swedish data: **Income**

Income (Sweden)

Educational attainment is key mediator for transmission of socio-economic advantages (“OED triangle”, Goldthorpe 2014).

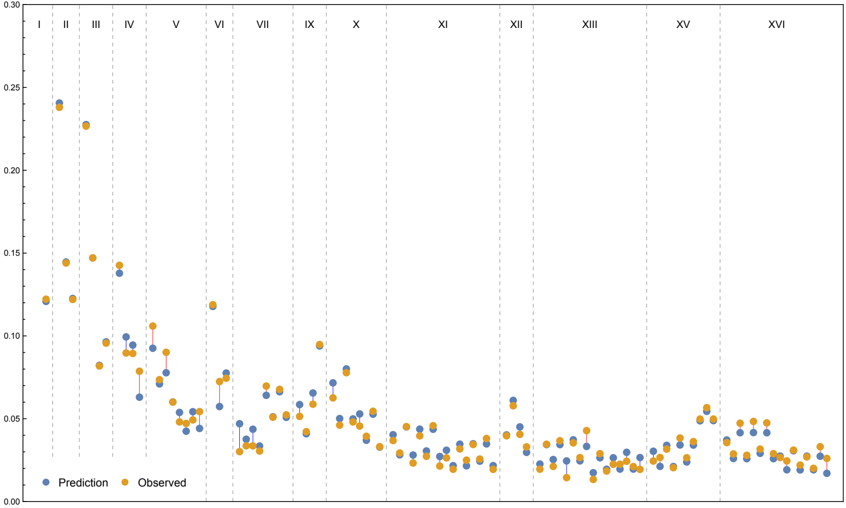
But do results generalize to other socioeconomic outcomes?

- *Ten-year* average of annual pre-tax **income**
- Measured around age 35 for children and around age 45 for parents
- 141 distinct moments, using 129 moments for calibration

Issues:

- Income correlations systematically lower for mixed or female pairs
- We do not model labor supply decisions, but the model is flexible enough to capture gender asymmetries

Figure: Sample and Predicted Moments (Income)



Income (Sweden): Results

Findings are qualitatively similar, but differ in magnitude:

Latent advantages more strongly transmitted than income itself, in all intergenerational, sibling and assortative processes:

- Father-son correlation in latent factor twice as large as in log income
- Sibling correlation in latent factor ≈ 0.8
- Spousal correlation in latent factor ≈ 0.65 (vs. ≈ 0.12 in log income)

However, the latent factors that determine educational attainment appear more strongly transmitted from one generation to the next than the latent factors that influence income. Back

Empirical Application: Spanish Census

Spanish Population Census 2001:

For the region of Cantabria we observe the full name of each person, and can use this information to identify kinship:

- Child generation born 1956-1976 (71,479 males, 68,830 females)
- A newborn in Spain receives two surnames, the first is the father's and the second the mother's (first) surname.
- Set of potential parents: couples born<1956, husband's and wife's surnames fit, age difference between parents and son>=16 years. Parents identified if only one couple in set (35% of cases).
- Siblings in child generation are identified.
- Set of potential siblings in parent generation: individuals sharing the same two surnames. Siblings identified if two individuals in set.
- Uncles nephews, and cousins are identified. Back

Table: Calibrated Parameters in Swedish Registers (Height, Baseline Model)

Panel A: Intergenerational Processes							
β^m	β^f	γ^m	γ^f	σ^2_{ym}	σ^2_{yf}	σ^2_{zm}	σ^2_{zf}
0.954	0.769	1.733	0.051	1.000	1.000	0.099	0.000
α_{ym}	α_{yf}	α_{zm}	α_{zf}	σ^2_{um}	σ^2_{uf}		
0.317	0.082	0.410	0.853	0.439	0.465		
Ancestor correlations in y and z:							
	Father-Son	Fr.-Daughter		Mother-Son		Mother-Daughter	
in z	0.760	0.807		0.700		0.735	
	Father-Son	Grandfr.-Son		GGrandfr.-Son		GGGrandfr-Son	
in y	0.488	0.238		0.115		0.056	
in z	0.760	0.580		0.443		0.339	
Panel B: Sibling Processes							
	σ^2_{xm}	σ^2_{xf}	σ_{xmxf}	σ^2_{em}	σ^2_{ef}	σ_{emef}	
	0.001	0.019	0.000	0.002	0.000	0.001	
Sibling correlations in y and z:							
	Brothers	Sisters	Mixed		Brothers	Sisters	Mixed
in y	0.521	0.535	0.458	*in z*	0.595	0.857	0.785
Panel C: Assortative Processes							
r^m_{zz}	r^m_{zy}	r^m_{yz}	r^m_{yy}	r^f_{zz}	r^f_{zy}	r^f_{yz}	r^f_{yy}
0.050	-0.004	-1.816	0.302	15.123	-0.098	-5.590	0.242
Spousal correlations in y and z:							
ρ_{ymyf}	ρ_{zmzf}	ρ_{ymzf}	ρ_{zmyf}				
0.270	0.908	-0.166	-0.553				
Panel D: Variance Decomposition of y							
	$\widetilde{y}_{t-1}$	$\widetilde{z}_{t-1}$	$\sigma_{\tilde{y}\tilde{z}}$	$\Sigma_{\tilde{y},\tilde{z}}$	x_t	e_t	$\Sigma_{\tilde{y},\tilde{z},x,e}$
male	62.3%	5.7%	-16.2%	51.9%	0.1%	0.2%	52.1%
female	52.7%	0.0%	-1.1%	51.6%	1.9%	0.0%	53.5%

Table: Calibrated Parameters in Swedish Registers (Height, Genetic Model)

Panel A: Intergenerational Processes							
β^m	β^f	γ^m	γ^f	σ^2_{ym}	σ^2_{yf}	σ^2_{zm}	σ^2_{zf}
0.954	0.769	1.733	0.051	1.000	1.000	0.099	0.000
α_{ym}	α_{yf}	α_{zm}	α_{zf}	σ^2_{um}	σ^2_{uf}		
0.317	0.082	0.410	0.853	0.439	0.465		
Ancestor correlations in y and z:							
	Father-Son		Fr.-Daughter		Mother-Son		Mother-Daughter
in z	0.760		0.807		0.700		0.735
	Father-Son		Grandfr.-Son		GGrandfr.-Son		GGGrandfr-Son
in y	0.488		0.238		0.115		0.056
in z	0.760		0.580		0.443		0.339
Panel B: Sibling Processes							
	σ^2_{xm}	σ^2_{xf}	σ_{xmxf}	σ^2_{em}	σ^2_{ef}	σ_{emef}	
	0.001	0.019	0.000	0.002	0.000	0.001	
Sibling correlations in y and z:							
	Brothers	Sisters	Mixed		Brothers	Sisters	Mixed
in y	0.521	0.535	0.458	*in z*	0.595	0.857	0.785
Panel C: Assortative Processes							
r^m_{zz}	r^m_{zy}	r^m_{yz}	r^m_{yy}	r^f_{zz}	r^f_{zy}	r^f_{yz}	r^f_{yy}
0.050	-0.004	-1.816	0.302	15.123	-0.098	-5.590	0.242
Spousal correlations in y and z:							
ρ_{ymyf}	ρ_{zmzf}	ρ_{ymzf}	ρ_{zmyf}				
0.270	0.908	-0.166	-0.553				
Panel D: Variance Decomposition of y							
	$\widetilde{y}_{t-1}$	$\widetilde{z}_{t-1}$	$\sigma_{\tilde{y}\tilde{z}}$	$\Sigma_{\tilde{y},\tilde{z}}$	x_t	e_t	$\Sigma_{\tilde{y},\tilde{z},x,e}$
male	62.3%	5.7%	-16.2%	51.9%	0.1%	0.2%	52.1%
female	52.7%	0.0%	-1.1%	51.6%	1.9%	0.0%	53.5%

Back