Intergenerational occupational mobility in Norway, 1865-2011

Jørgen Modalsli

Research Department, Statistics Norway

Conference on Social Mobility, University of Chicago

Motivation: Social mobility in the long run

- Large changes in the structure of the economy over the last 150 years
  - The cross-section distribution of activity has changed substantially — example Norway: Farmers 1865: $>$ 40 per cent; 2011: $\approx$ 1 per cent
  - Large growth in aggregate productivity
  - Within-country convergence between regions
  - Income inequality fell from the nineteenth century to the late twentieth century in most countries
- **How did this transition affect social mobility?**
  - While income inequality has fallen, we do not know whether social mobility has increased

Motivation: Why Norway

- Comparable evidence on long-run mobility mainly from Great Britain and US
- Both have special development paths
  - Britain: Early industrialization, low share of farmers in nineteenth century
  - United States: "New" society, extensive immigration, low population to land ratio
- Not clear how this translates to other countries
- Norway is an interesting case also for other reasons
  - Transition from rural, remote society to one of the most well-off in Europe
  - From (relatively) inegalitarian to egalitarian
  - Consensus among historians: Low mobility in nineteenth century
  - Late industrializing, early state-building: good data for most of the industrializing period

This presentation

- Newly-compiled data on occupation distribution and intergenerational occupational mobility in Norway between 1865 and 2011
- Preview of findings:
  - Nineteenth-century Norwegian mobility low: at similar level to United Kingdom
  - Large increase in social mobility in Norway over time, except for farm sector

This presentation

- Newly-compiled data on occupation distribution and intergenerational occupational mobility in Norway between 1865 and 2011
- Preview of findings:
  - Nineteenth-century Norwegian mobility low: at similar level to United Kingdom
  - Large increase in social mobility in Norway over time, except for farm sector
  - Increase not driven by geographic differences / convergence
  - Modest contributions from mobility to between-occupation "dynastic" income inequality

Literature: Intergenerational mobility over long time periods

- Long and Ferrie (2013):
  - United States and Great Britain, 19th and 20th century
  - Mobility used to be higher in the US; not so any more
- Clark and Cummins (2014) on wealth mobility: no large changes over time in UK
- Local areas in Sweden: no big changes over time (Lindahl et al, 2012; Dribe et al, 2012)

Literature

- High social mobility in Scandinavia today, at least compared to UK and US
  - Intergenerational income elasticities low in Scandinavian countries (Jäntti et al 2006, Raaum et al 2007)
  - Small changes in income mobility for cohorts born 1950-1965 in Norway (Bratberg et al 2005)
  - Increase in intergenerational mobility if one starts with the 1930s cohorts (Salvanes, this workshop)
  - Sweden: Fall in sibling correlations for cohorts born 1932-1950; increase thereafter (Björklund et al 2009)
  - Occupational mobility: Standardized studies on post-1970 data (Breen 2004) show moderately increasing mobility in several European countries (incl Scandinavia), but not in Great Britain
- Geographical heterogeneity in mobility (Chetty et al 2014)

Data sources

- Full-count data from Norwegian censuses of 1865, 1900, 1910, 1960, 1970, 1980, 2011
- 1865-1910: digitized by Norwegian National Archives and partners from 1990s until today
  - Occupation and geographical covariates coded
  - Used in some economic research (eg Abramitzky et al 2012)
  - No information on income or education
- 1960-2011: based on original working files from Statistics Norway
  - Norwegian population register with individual IDs originate from 1964; 1960 census later added
- Occupation mean incomes: varying quality. Will use for some interpretation of results
- Covariates: some economic variables at the municipality level

Observation structure

- Intergenerational occupation pair always comes from two different censuses

Year	Individual A		Individual B
1865	Adult	$\leftarrow$*(father-son link)*$\rightarrow$	Child
			$\uparrow$
			(person match)
			$\downarrow$
1900			Adult

Observation structure

- Intergenerational occupation pair always comes from two different censuses

Year	Individual A		Individual B
1865	Adult	$\leftarrow$*(father-son link)*$\rightarrow$	Child
	(Father's occ.)		$\uparrow$
			(person match)
			$\downarrow$
1900			Adult
			(Son's occ.)

Observation structure

- Intergenerational occupation pair always comes from two different censuses

Year	Individual A		Individual B
1865	Adult	$\leftarrow$*(father-son link)*$\rightarrow$	Child
	(Father's occ.)		$\uparrow$
			(person match)
			$\downarrow$
1900			Adult
			(Son's occ.)

- Before 1960: No individual ID numbers
- Individuals matched on names, time of birth and place of birth
  - Allow for differences in spelling, minor misreporting of times/places
  - No matching on address / household composition
- Father-son identification based on co-residence at time of census

Matching: Results

Matchable: "Son" is age 30-60 at t_1, alive at t_0, born in Norway.

t_0-t_1	Match-able in t_1	Share found in t_0	Known father in t_0	Matched pop.	Father age 30-60	Both have occ.	Final sample
1865-1900	246,373	37.7%	71.9%	66,790	91.4%	98.1%	59,896
1910-1960	246,911	45.4%	77.8%	87,188	88.8%	89.6%	69,356
1960-1980	717,678	100.0%	40.3%	289,040	82.3%	84.6%	201,297
1980-2011	883,951	100.0%	93.6%	827,210	80.8%	75.6%	505,441

Matching: Results

Matchable: "Son" is age 30-60 at t_1, alive at t_0, born in Norway.

t_0-t_1	Match-able in t_1	Share found in t_0	Known father in t_0	Matched pop.	Father age 30-60	Both have occ.	Final sample
1865-1900	246,373	37.7%	71.9%	66,790	91.4%	98.1%	59,896
1910-1960	246,911	45.4%	77.8%	87,188	88.8%	89.6%	69,356
1960-1980	717,678	100.0%	40.3%	289,040	82.3%	84.6%	201,297
1980-2011	883,951	100.0%	93.6%	827,210	80.8%	75.6%	505,441
	Alternative sample: age 0-15 at t_0 only						
1865-1900	159,850	38.1%	82.9%	50,490	92.5%	98.1%	45,835
1910-1960	246,911	45.4%	77.8%	87,188	88.8%	89.6%	69,356
1960-1980	154,901	100.0%	80.3%	124,437	97.5%	86.0%	104,401
1980-2011	455,843	100.0%	97.4%	444,175	81.0%	78.5%	282,613

Matching: Results

Matchable: "Son" is age 30-60 at t_1, alive at t_0, born in Norway.

t_0-t_1	Match-able in t_1	Share found in t_0	Known father in t_0	Matched pop.	Father age 30-60	Both have occ.	Final sample
1865-1900	246,373	37.7%	71.9%	66,790	91.4%	98.1%	59,896
1910-1960	246,911	45.4%	77.8%	87,188	88.8%	89.6%	69,356
1960-1980	717,678	100.0%	40.3%	289,040	82.3%	84.6%	201,297
1980-2011	883,951	100.0%	93.6%	827,210	80.8%	75.6%	505,441
	Alternative sample: age 0-15 at t_0 only						
1865-1900	159,850	38.1%	82.9%	50,490	92.5%	98.1%	45,835
1910-1960	246,911	45.4%	77.8%	87,188	88.8%	89.6%	69,356
1960-1980	154,901	100.0%	80.3%	124,437	97.5%	86.0%	104,401
1980-2011	455,843	100.0%	97.4%	444,175	81.0%	78.5%	282,613
	Other studies						
1850-1880	62,811	21.9%	74.2%	9,497		US 1% (1)	
1851-1881		20.3%		14,191		UK 2% (1)	
1865-1900		$\approx$ 5%		20,446		NO/US (2)	

(1): Long and Ferrie 2013; (2): Abramitzky et al 2012

Occupations

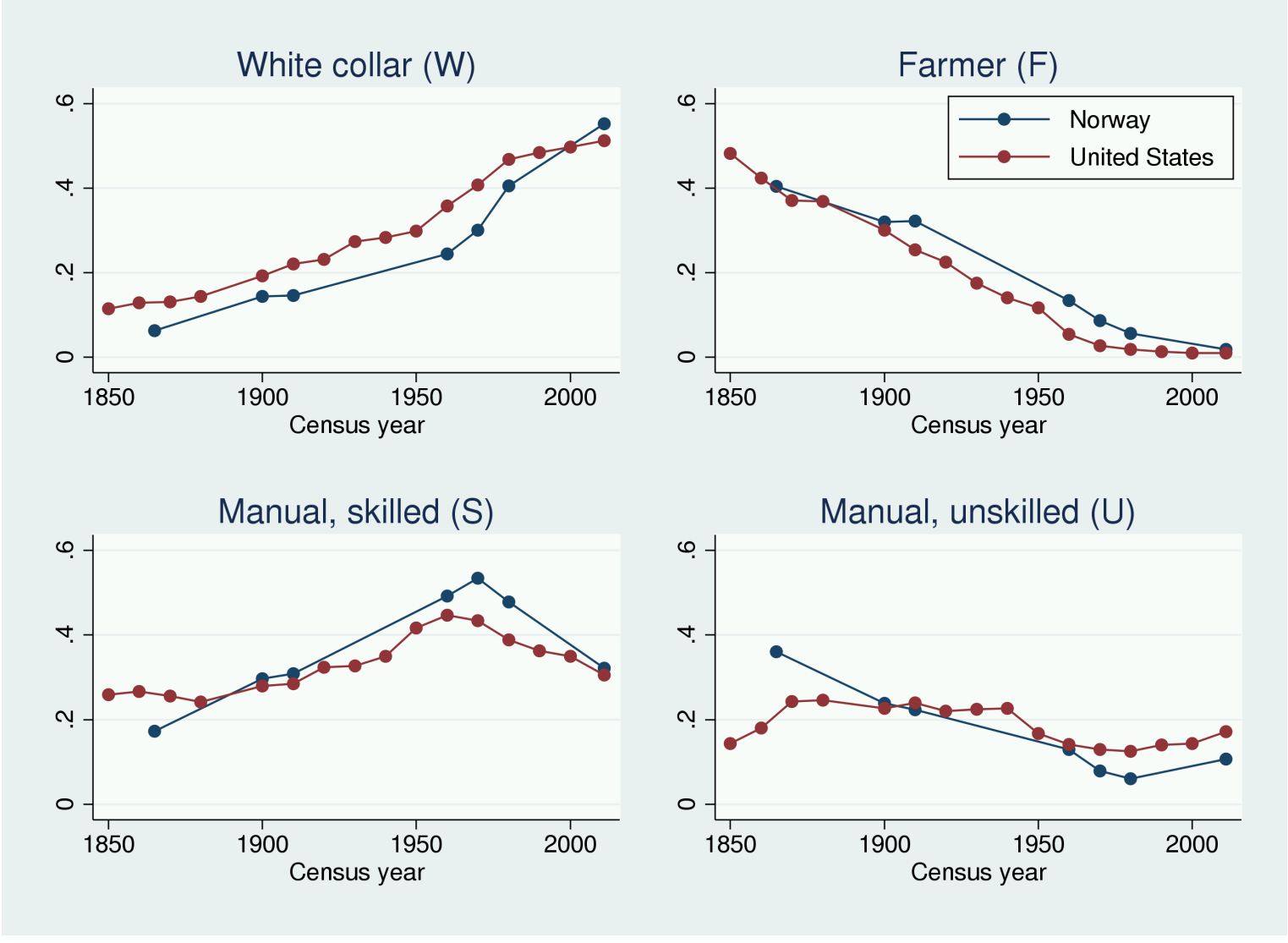
- Occupations present the longest-running consistent information on individual economic conditions
- Changes in occupation reporting standards over time
  - Here: standardize to four occupation categories similar to Long and Ferrie (2013)
- Population universe for this talk: Men 30-60 years old at time of observation

Occupations

Category	Examples of subgroups
White-collar	Teachers, Merchants, Captains Lawyers, Managers, Office clerks Shop clerks, Salesmen
Farmer	Farmers, Farmer-fishermen, Farm managers
Manual, skilled	Carpenters, Electricians, Welders, Car mechanics, Butchers
Manual, unskilled	Husbandmen/Cottars, Fishermen Farm workers, Day laborers, Loggers

(Compatibility: Long and Ferrie (2013))

Occupations over time



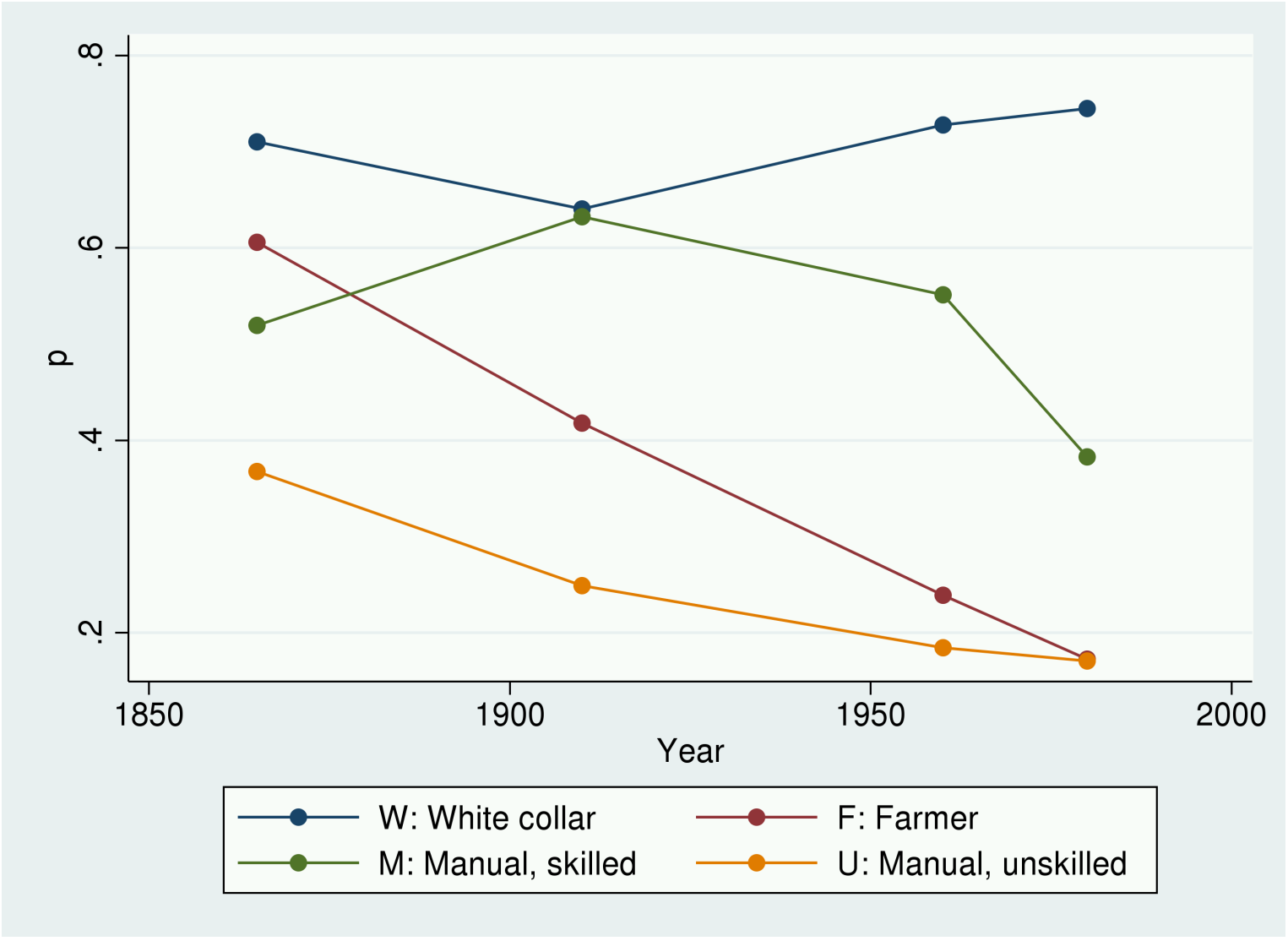
Unit of observation: 4×4 matrices

Example: 1865-1900

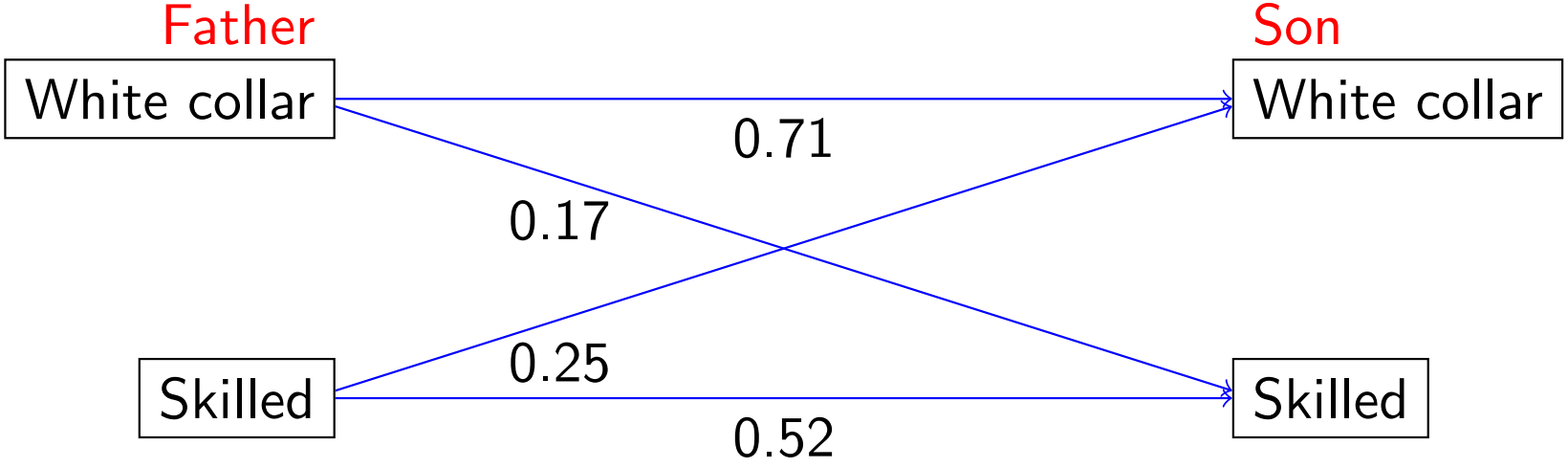
		Son's occupation			
		W	F	S	U
Father's occupation	White-collar (W)	2,277	189	541	187
	Farmer (F)	3,233	21,118	5,247	5,287
	Manual, skilled (S)	1,622	519	3,304	935
	Manual, unskilled (U)	1,060	4,006	4,892	5,793

Transition probabilities

Probability of son having same occupation as father

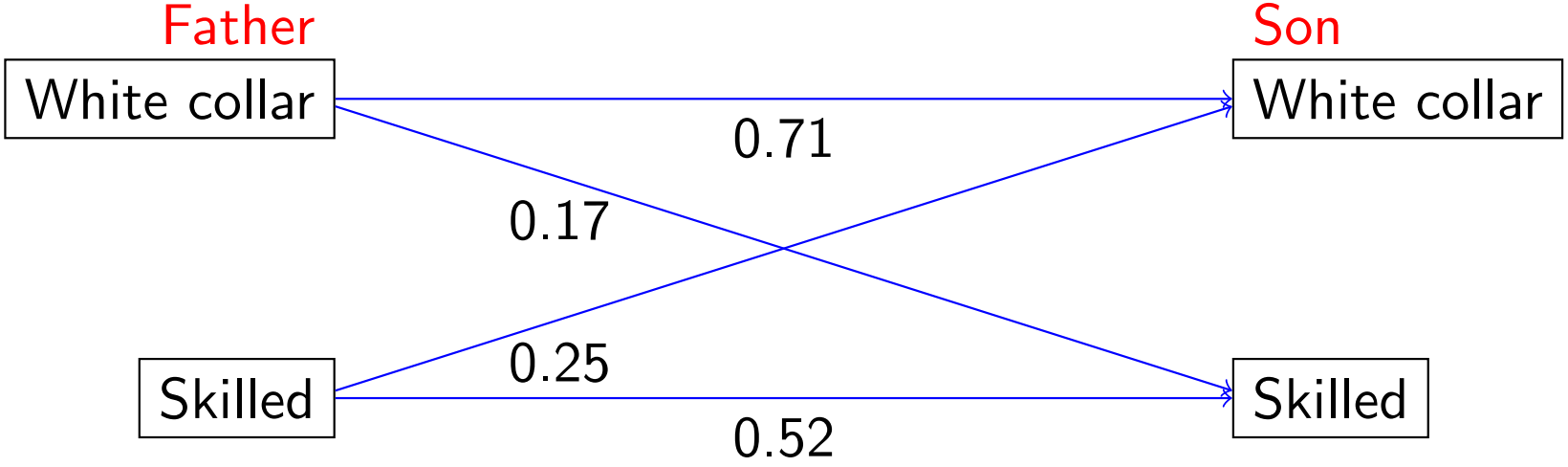


Transition probabilities and odds ratios



$$\Theta_{wwss} = \log\left(\frac{0.71/0.17}{0.25/0.52}\right) = \log\left(\frac{4.16}{0.48}\right) = \log(8.60) = 2.15$$

Transition probabilities and odds ratios



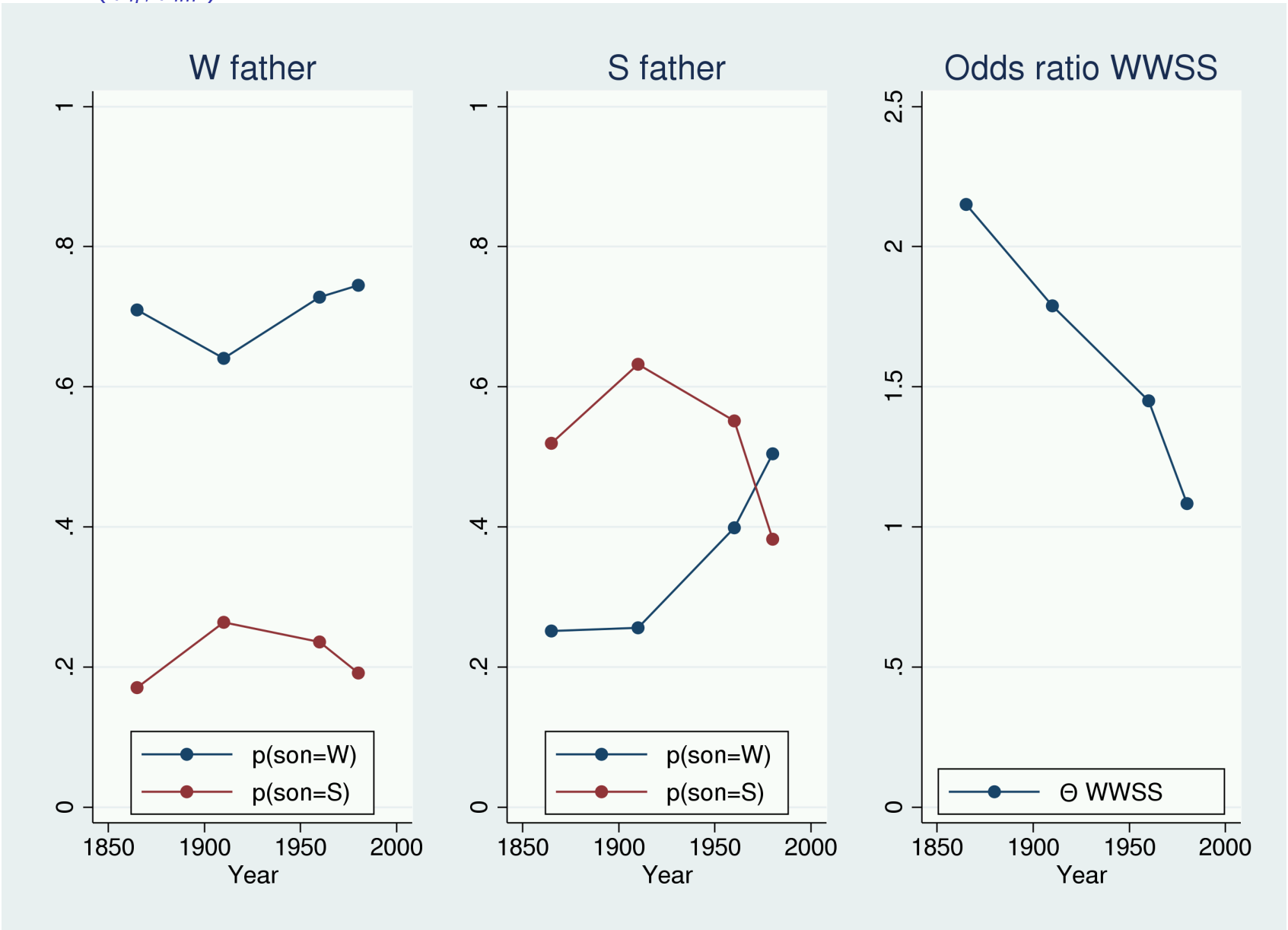
$$\Theta_{wwss} = \log\left(\frac{0.71/0.17}{0.25/0.52}\right) = \log\left(\frac{4.16}{0.48}\right) = \log(8.60) = 2.15$$

Aggregate mobility: use Altham statistic (Altham 1970; Long and Ferrie 2013):

$$d(P,J) = \left(\sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{l=1}^{r}\sum_{m=1}^{s}\left[\log\left(\frac{p_{ij}/p_{im}}{p_{lj}/p_{lm}}\right)\right]^2\right)^{1/2}$$

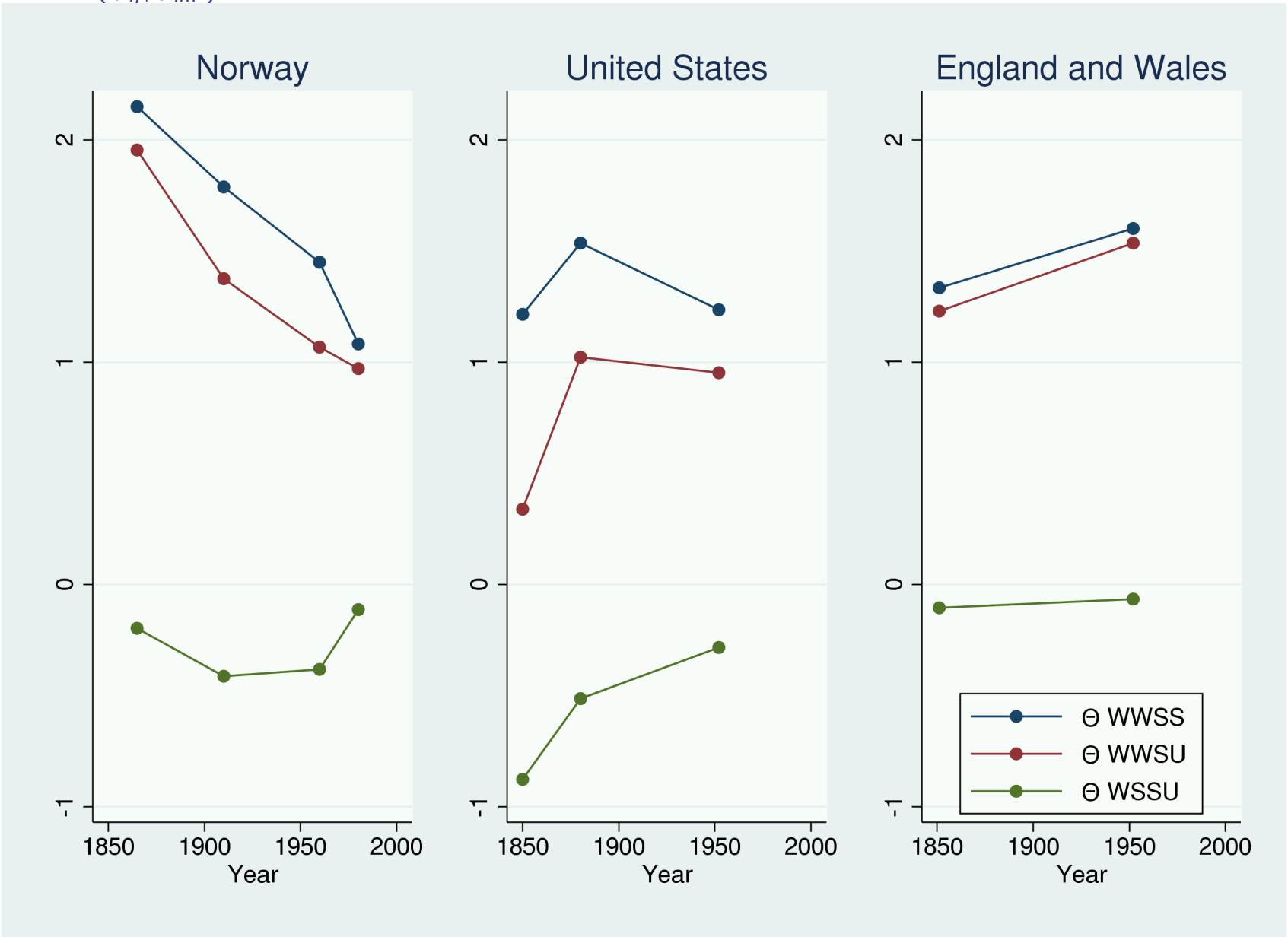
Probabilities and odds ratios over time

$$\theta_{ijlm} = \log\left(\frac{p_{ij}/p_{im}}{p_{lj}/p_{lm}}\right)$$



Odds ratios: Norway, UK and US

$$\theta_{ijlm} = \log\left(\frac{p_{ij}/p_{im}}{p_{lj}/p_{lm}}\right)$$



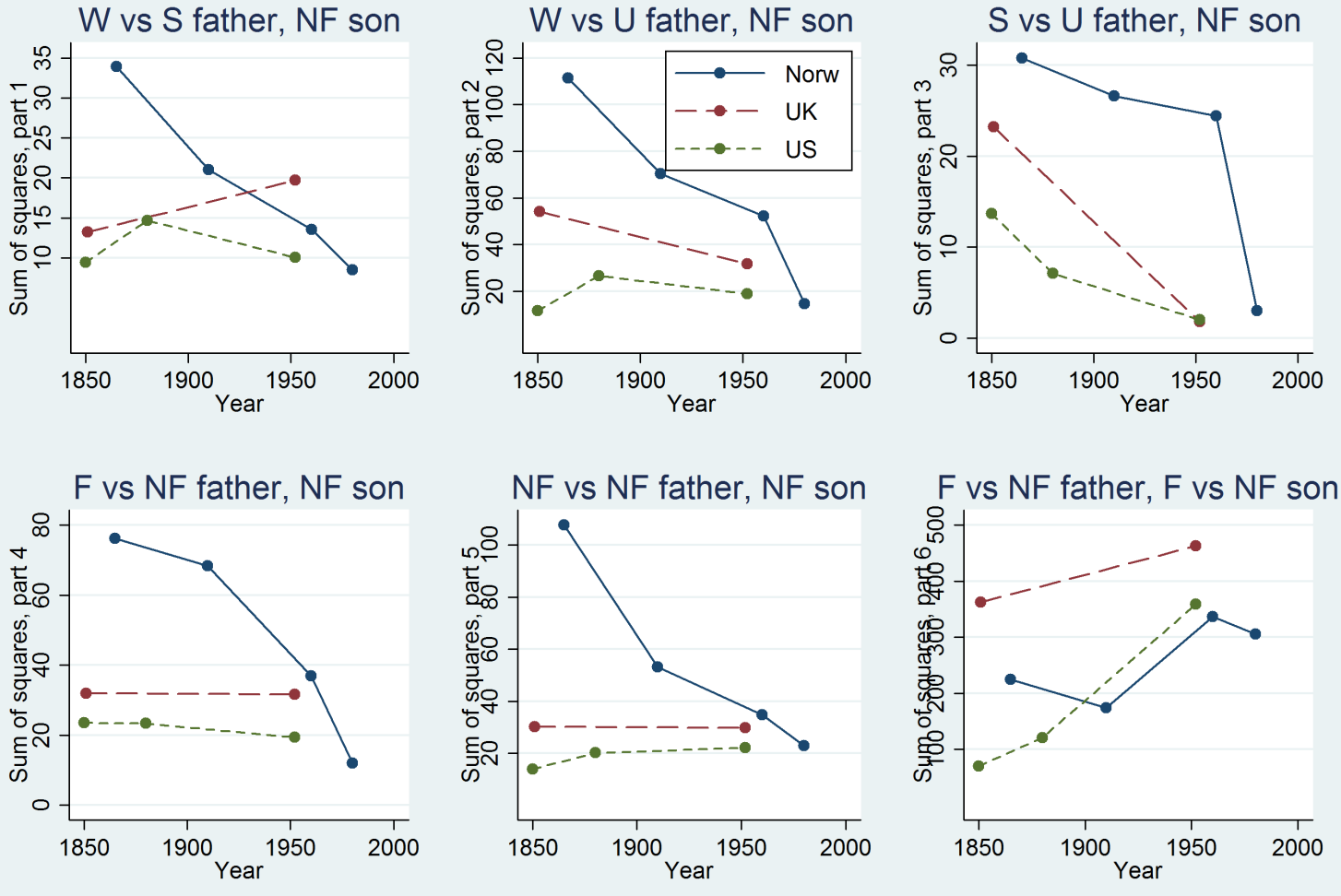
Components of the Altham statistic

$$d(P,J)=\left(\sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{l=1}^{r}\sum_{m=1}^{s}\left[\log\left(\frac{p_{ij}/p_{im}}{p_{lj}/p_{lm}}\right)\right]^2\right)^{1/2}$$

Father's occupation	Son's occupation: WS	WU	SU	FW	FS	FU
WS	(1) W vs S father, nonfarmer son			(5) Nonfarmer vs nonfarmer father, farmer vs nonfarmer son		
WU	(2) W vs U father, nonfarmer son					
SU	(3) S vs U father, nonfarmer son					
FW	(4) Farmer vs nonfarmer father, nonfarmer son			(6) Farmer vs nonfarmer father, farmer vs nonfarmer son		
FS						
FU						

Odds ratio aggregates

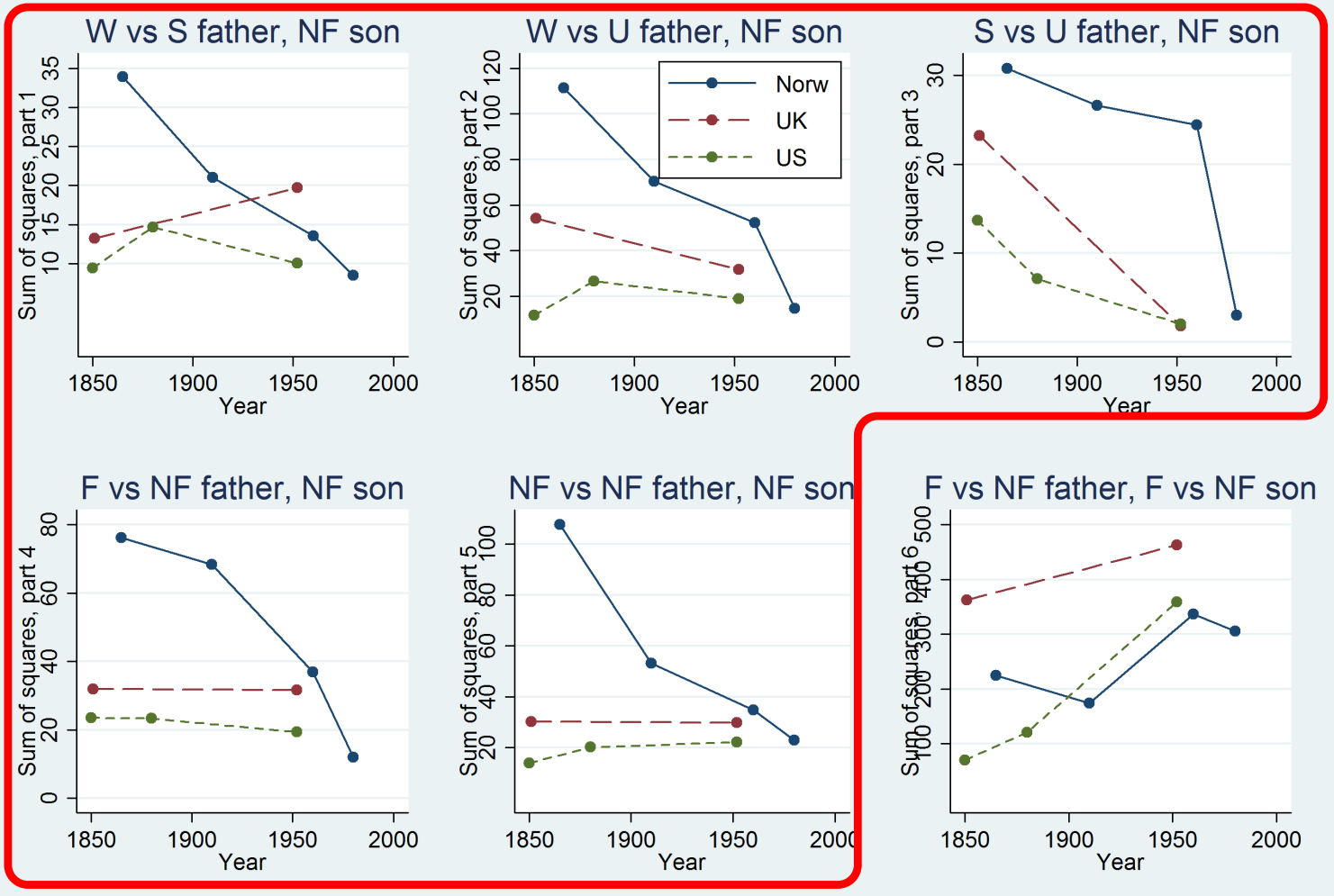
Six subgroups that sum to the Altham statistic



W=White collar, S=Skilled, U=Unskilled, F=Farmer, NF=Not Farmer (=W, S or U)

Odds ratio aggregates

Six subgroups that sum to the Altham statistic

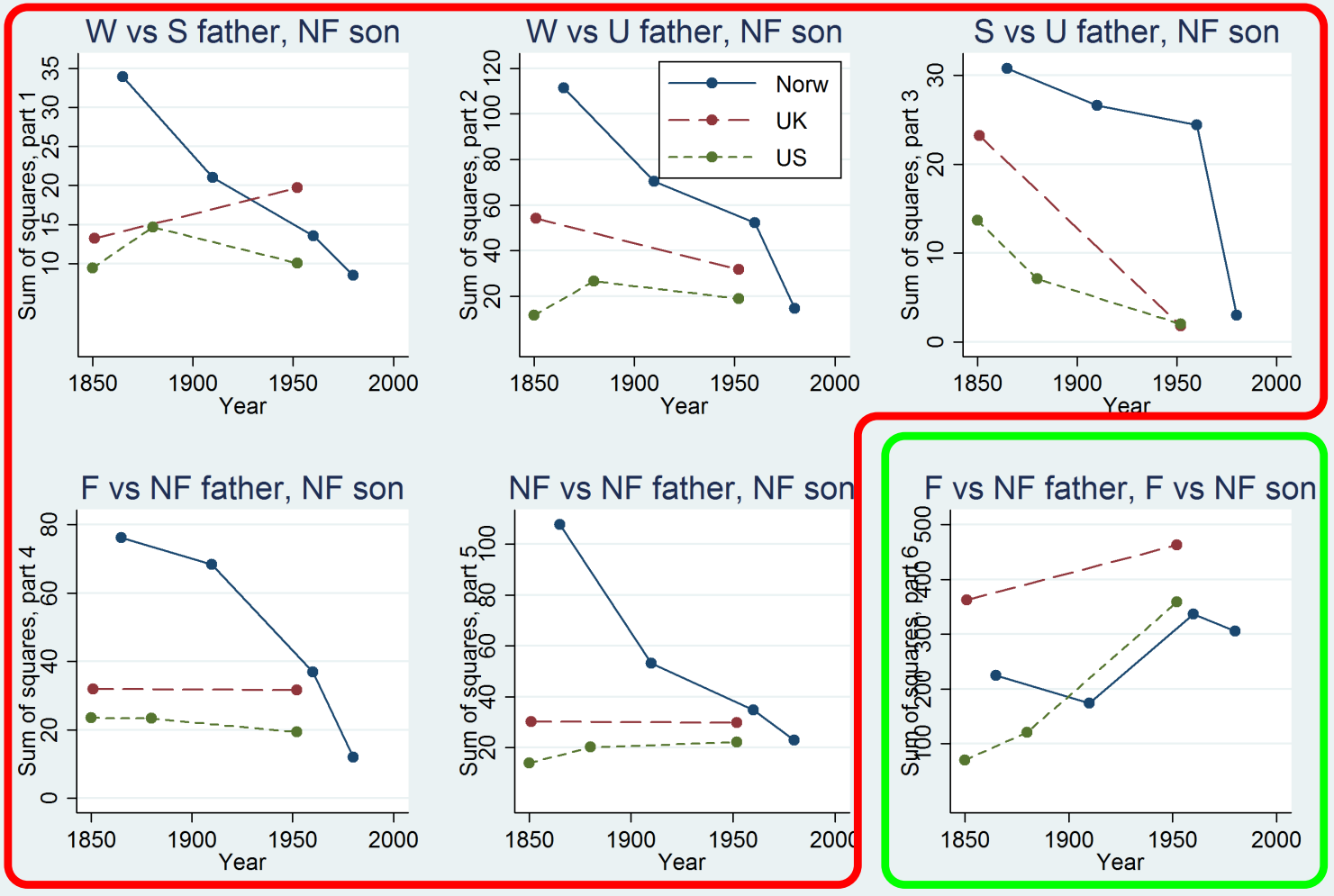


Nonfarm and farm-nonfarm

W=White collar, S=Skilled, U=Unskilled, F=Farmer, NF=Not Farmer (=W, S or U)

Odds ratio aggregates

Six subgroups that sum to the Altham statistic



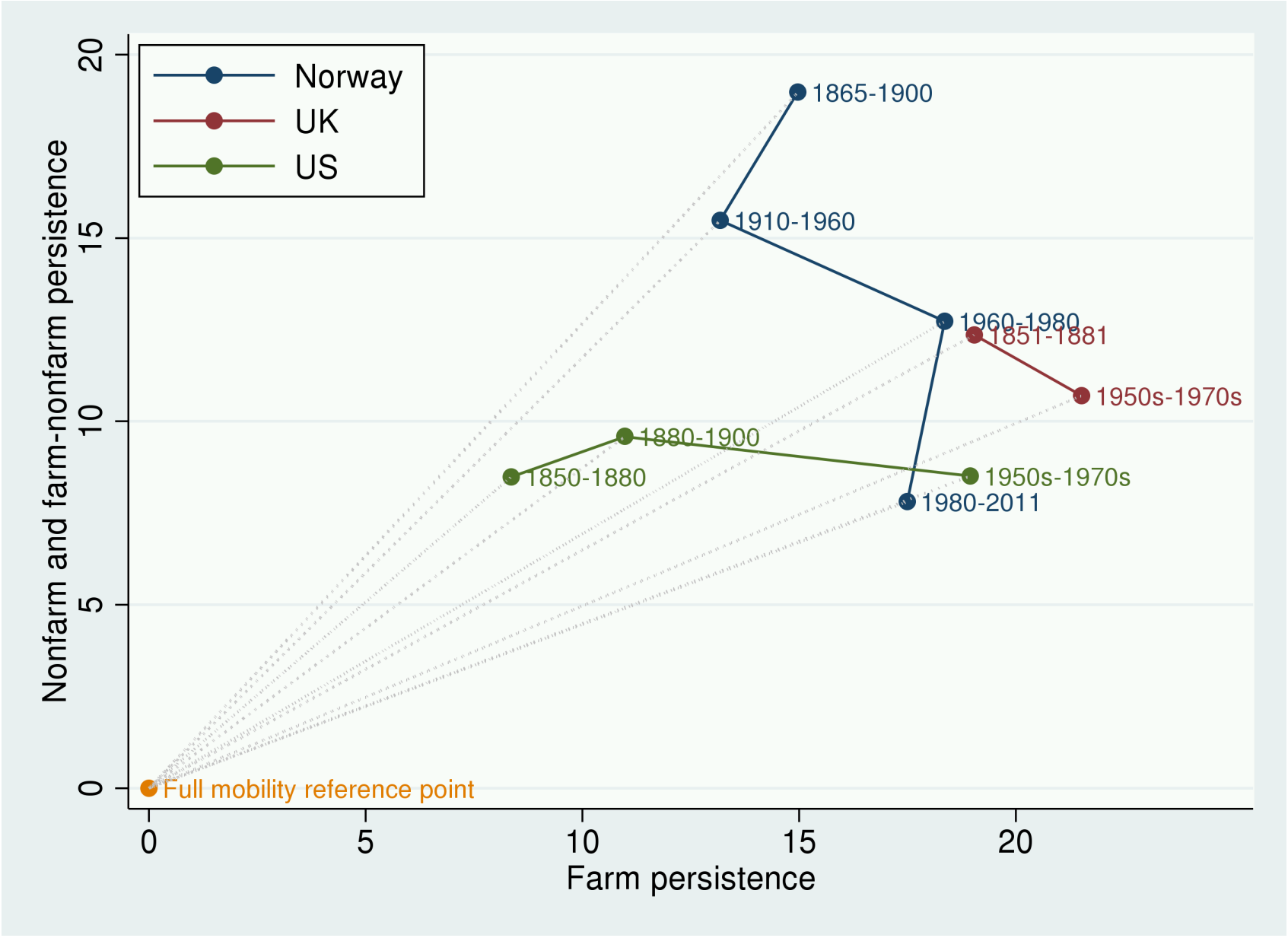
Nonfarm and farm-nonfarm

Farm

W=White collar, S=Skilled, U=Unskilled, F=Farmer, NF=Not Farmer (=W, S or U)

Different trends in farm mobility and nonfarm mobility

Norway, UK, US: Twoway decomposition of social mobility



Different trends in farm mobility and nonfarm mobility

		$d(P, J)$= Distance from full mobility	Nonfarm and farm-nonfarm component	Farm component
Norway	1865-1900	24.2	19.0	15.0
	1910-1960	20.3	15.5	13.2
	1960-1980	22.3	12.7	18.4
	1980-2011	19.2	7.8	17.5
UK	1851-1881	22.7	12.4	19.0
	1952	24.0	10.7	21.5
US	1850	11.9	8.5	8.4
	1880	14.6	9.6	11.0
	1952	20.8	8.5	18.9

Mobility differences across subregions

- Calculate the farm and nonfarm components for $d(P, J)$ for subpopulations
- Regional comparisons: Rural vs. urban, High vs. low economic growth
  - High-growth areas more "dynamic"?
- Individual-group comparisons: Movers vs. nonmovers
  - Hypothesis: Correlation between moving location and moving occupation: movers have higher social mobility
- The emigration question: Areas with high and low emigration
  - Hypothesis: depends on selection of immigrants
  - Poor but industrious more likely to emigrate: high emigration $\rightarrow$ low social mobility
  - Well-off more likely to emigrate: high emigration $\rightarrow$ high social mobility

Mobility differences across subregions

Nonfarm and farm-nonfarm components

		1865-1900	1910 - 1960	1960-1980	1980-2011
Reference		19.0	15.5	12.7	7.8
Rural / Urban	Rural	17.4	15.8	13.3	7.9
	Urban	18.3	14.8	12.3	7.7
Local inc. growth	Below mean	19.5	15.1	12.5	7.8
	Above mean	18.3	15.0	12.5	7.8
Mover / Nonmover	Rural nonmover	18.8	16.7	14.6	8.7
	Urban nonmover	18.9	15.9	13.3	8.5
	Mover ($R \to R$)	18.3	16.0	10.9	6.0
	Mover ($R \to U$)	14.8	12.0	10.1	5.7
	Mover ($U \to R$)	15.9	11.9	9.7	5.7
	Mover ($U \to U$)	16.8	11.4	8.4	5.4
Local emigration rate	Low	18.6	15.7		
	High	19.3	14.4		

Childhood environment over and above father's occupation

- Social mobility: the extent to which son's occupation depends on father's occupation
- However, also neighborhood effects
- Son of farmer father more likely to live in farming area $\rightarrow$ more likely to be a farmer himself
- Hypothesis: Measures of intergenerational occupational mobility goes down when we control for region of origin

Correcting for region of origin

To add control variables, we can consider a multinomial logit system of three equations (W as reference category):

$$\log\left(\frac{Pr(Occ_q = k)}{Pr(Occ_q = W)}\right) = \alpha_k + \boldsymbol{\beta}_k' \mathbf{D}_q + \boldsymbol{\gamma}_k' \mathbf{X}_q + \epsilon_{k,q} \qquad k = F, S, U$$

- $\mathbf{D_q} = \{D_F, D_S, D_U\}$ characterizes father's occupation
- $\beta_k = \{\beta_k^F, \beta_k^S, \beta_k^U\}$ is the associated parameter vector
- $\mathbf{X_q}$ is a vector of other individual covariates with associated parameters γ_k

The Altham statistic depends only on the β's:

$$d(P, J) = \left(\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{l=1}^{N}\sum_{m=1}^{N}\left[(\beta_j^i - \beta_m^i) - (\beta_j^l - \beta_m^l)\right]^2\right)^{1/2} \qquad (1)$$

Can bootstrap confidence intervals using parameter standard errors

Correcting for region of origin: Results

Time period	No controls	No controls, clustered SE	Local mean income	Employment shares	Regional dummies
1865-1900	24.2	24.2	22.1	20.8	21.0
	(23.7 – 24.7)	(23.0 – 25.4)	(21.1 – 23.1)	(19.4 – 22.3)	(19.7 – 22.3)
1910-1960	20.3	20.3	18.1	17.7	17.7
	(20.0 – 20.7)	(19.2 – 21.6)	(17.5 – 18.8)	(17.0 – 18.4)	(17.0 – 18.4)
1960-1980	22.3	22.3	21.2	20.0	19.9
	(22.1 – 22.6)	(21.2 – 23.6)	(20.5 – 22.0)	(19.3 – 20.8)	(19.1 – 20.7)
1980-2011	19.2	19.2	18.1	17.1	16.9
	(18.9 – 19.4)	(18.3 – 20.1)	(17.4 – 18.9)	(16.4 – 17.8)	(16.3 – 17.6)

Interpreting occupational mobility as welfare changes

Using mean income data

- Before 1967, no individual data on income
- However, can piece together mean incomes by occupation
  - 1980 and 2011: from tax micro data
  - 1960: using tax micro data from 1967 on individuals in 1960
  - 1910 and 1900: using tabulations of income by occupation, age and gender from 1911
  - 1865: using tabulations of mean taxes paid by occupation group from 1868 (different population definition: all men age 25+)

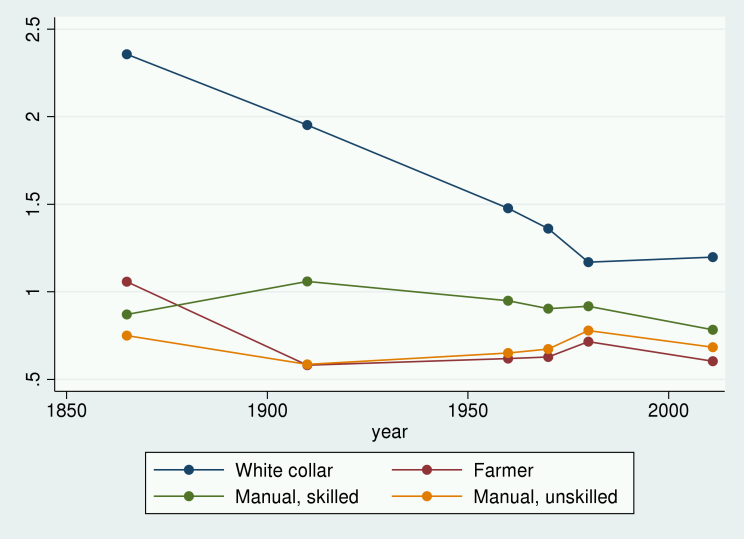


Figure: Mean income of occupation group relative to population mean, by year

The contribution of occupational mobility to dynastic income equalization

- Consider the two-generation "dynastic income" of fathers and sons:

$$Y_{\text{dynasty},1865-1900} = Y_{\text{father},1865} + \frac{1}{g} Y_{\text{son},1900}$$

  using the mean incomes of occupation groups.
- Question: How much has increased social mobility contributed to equalization of dynastic incomes?
- Decomposition analysis a la DiNardo et al (1996) / Eika et al (2014):
  - Contribution of marginal income distributions *vs* contributions of intergenerational occupational mobility
- Here: Fix the marginal income distributions, consider counterfactual intergenerational mobility matrices

Constructing a counterfactual 4×4 income matrix

- Adjustment procedure of Mosteller (1968) used to construct matrix M_t^k matching marginal distributions of time interval t and odds ratios (and hence $d(P, J)$) of time k
  - The matrix has 16 degrees of freedom
  - 7 degrees needed to match father and son population distributions
  - 9 remaining degrees correspond to the 9 independent odds ratios (for example WF, WU, WF for father and son)
- For t, use the actual father-son distributions of individuals, and the actual income distributions
- For k, consider actual matrices and (a) full mobility, ie all odds ratios are 1, and (b) lowest constructable mobility respecting marginal distributions

Results from the counterfactual analysis

Gini coefficient of occupation mean incomes

	Gini coefficients (no counterfactual)		
t	Fathers+Sons	Fathers	Sons
1865-1900	15.7	12.8	24.5
1910-1960	19.0	24.5	17.2
1960-1980	11.0	16.6	7.9
1980-2011	8.3	8.0	10.9

Results from the counterfactual analysis

Gini coefficient of occupation mean incomes

	Gini coefficients (no counterfactual)		
t	Fathers+Sons	Fathers	Sons
1865-1900	15.7	12.8	24.5
1910-1960	19.0	24.5	17.2
1960-1980	11.0	16.6	7.9
1980-2011	8.3	8.0	10.9

	F+S Gini coefficients with mobility counterfactuals					
t	Lowest mob.	1865-1900	1910-1960	1960-1980	1980-2011	Highest mob. (J)
1865-1900	18.2	15.7	15.4	15.2	15.0	15.4
1910-1960	20.8	19.2	19.0	18.9	18.3	16.2
1960-1980	12.2	11.2	11.0	11.0	10.6	9.7
1980-2011	9.5	8.9	8.7	8.6	8.3	7.5

Results from the counterfactual analysis

- High dynastic income inequality before 1960:
  - 1865-1900: *no* contribution from mobility (high change in income distribution gives lower father-son correlations anyway)
  - 1910-1960: Some contribution from mobility (actual: Gini=19, full-mobility Gini 16)
- After 1960: less variation between occupation mean incomes; lower span of potential Gini coefficient
- Despite dissimilarity of mobility matrices, replacing one with another does not greatly affect the dynastic income Gini in any of the time periods
- Results must be cautiously interpreted; only between-occupation inequality measured

Change in mobility over time

- Intergenerational mobility has increased over time in Norway, unlike in Great Britain
- However, composition of decrease different from that in the United States
- Increase in mobility mid-20th century also found for earnings in Norway (Salvanes) and Sweden (Björklund et al)
- However, no trace of increased inequality (and possible decrease in mobility, cf. Swedish results) for Norway
  - Split of white-collar into two occupation groups only amplifies increase in mobility over time
  - However, still plenty of scope for within-occupation inequality; more individual-based rather than class-based rewards etc.

Geographic determinants

- Increased intergenerational mobility not primarily driven by regional convergence
- ...or by transatlantic migration
- High-income areas experience slightly higher intergenerational mobility in the early period
- Geographic mobility correlated with occupational mobility

Future work

- Mobility over several generations
- More case-specific analyses (technology etc)
  - Mobility by industry?
- Interaction between geographical and occupational mobility
- More family background: mother's coded occupation, siblings