Little School on the Prairie: A Push for Structural Transformation

Devin Bissky Dziadyk[a]

University of Toronto

July 1, 2025

[a] Supported in part by funding from the Social Sciences and Humanities Research Council. All errors are my own.

The One Room School

- One room schools loom large in the American mythos and in the development of education systems around the world.
  - Common in the United States, Canada, United Kingdom, Australia, and Switzerland.
  - Formed an input for the US becoming a world leader in educating its population (Goldin & Katz, 2008).
  - Rural, ungraded, between ages of 5-14, teach children the 3R's: Reading, Writing, and Arithmetic.
- This project: What was the impact of one room schools on children and their communities?
  - What is the role of changing labor supply in structural transformation out of agriculture?

Research Question: What was the impact of one room schools?

- Rural one room schools were not known for high quality education:
  - Low educational quality, poor resources, difficulty recruiting teachers, and lower educational returns (Goldin & Katz, 2000; Lachanski, 2024).
- But early forms of education were important:
  - Decentralization allowed communities to custom tailor education (Dippel et al., 2020).
  - Widespread schooling raised intergenerational mobility, decreasing correlation with mother's human capital (Althoff et al., 2025).
  - Schooling as a push factor out of agriculture (Caselli & Coleman II, 2001; Porzio et al., 2022).

Schools Push People out of Agriculture

- Utilize the rollout of rural schools in the western Canadian prairies at the start of the 20th century.
  - Combine census records with the near universe of school openings from 1871-1912.
  - Geolocate schools and households to exact land parcel (unique feature to Canada).
- Distance to school is likely endogenous, could reflect a variety of factors.
- Plausibly exogenous timing of school construction relative to a child's age.
  - Counterfactual is having a school versus not.

School Example: MacDowall/Cecil School #130

School Example: MacDowall/Cecil School #130

School Example: MacDowall School #130

Preview of Results

- Estimate a difference-in-differences finding that exposure to one room schools:
  - Decreases the likelihood of being a farmer 3.4pp (6%), increases incomes 12.3%, and increases distance from childhood home by 12.7%.
  - Increases likelihood of becoming teachers (1.1pp), managers (0.9pp), and agents (0.5pp), decreases likelihood of becoming retail clerks (-0.7pp).
  - Increases farm sizes by 4.8% in the areas of construction.
- Despite low returns to education from one room schools, they were an important institution to sort children to other occupations and encourage regional development. Two possible mechanisms:
  1. Higher returns to human capital outside of agriculture.
  2. One room schools reveal underlying ability in non-agricultural sectors.

Related Literature

- Early Education in United States & Canada: Goldin and Katz (1998), Goldin (1998), Goldin and Katz (2000), MacKinnon and Minns (2009), Dippel et al. (2020), Schaede (2021), Card et al. (2022a), Card et al. (2022b), Lachanski (2024) Althoff et al. (2025).
  - Contribution: Plausibly exogenous availability of schooling.
- Education & Structural Change: Caselli and Coleman II (2001), Porzio et al. (2022), Budı-Ors (2023), Gauthier et al. (2025).
  - Contribution: Causal evidence that education reallocates labour away from agriculture.
- Western Settlement and Development: Mattheis and Raz (2019), Smith (2022), Leonard and Kogelmann (2022), French (2022), Nagy (2023), Bagagli (2023).
  - Contribution: First to geolocate entire rural populations.
- Census Linking: Abramitzky et al. (2012), Antonie et al. (2014), Abramitzky et al. (2014), Feigenbaum (2016), Antonie et al. (2020), Abramitzky et al. (2020), Bailey et al. (2020), Price et al. (2021) Helgertz et al. (2022), Feigenbaum et al. (2023), Buckles et al. (2023), Abramitzky et al. (2024).
  - Contribution: First to link Canadian census in 20th century, including waves of prairie census.

Background & Data

Schools in the Prairies

- The early need for schools was filled by single one room schoolhouses which covered Alberta, Saskatchewan, and Manitoba.
- The rollout of these schools was not uniform and required a petition by the community to the provincial government. ▸ Details
  - Created a logistical problem in the 1910's as new schools were needed to literally fill in the gaps between existing schools.
- Schools would accommodate the local religion, with no difference in teacher certification by religion.
- Attendance by children was inconsistent.
  - *"Parents who keep children out of school because of distance or danger, appear to have no hesitation in sending the boys to market"* - A. Kennedy, Inspector of Schools, 1910
  - ▸ More

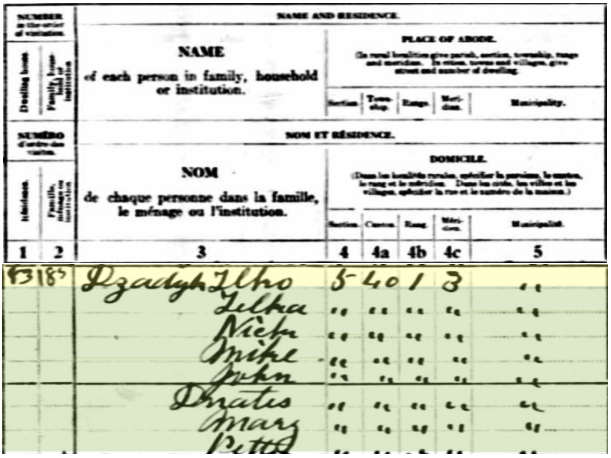
Data

- Historical Canadian Census: 1901, 1906 (Prairies), 1911, 1916 (Prairies), 1921, 1926 (Prairies), 1931.
  - Primary focus: 1906, 1911, and 1931.
- Historical school locations and founding dates from 1871 to 1912. ▸ Details
- Historical railroad line locations.
- FAO estimates for agricultural yields.
- ▸ Other Data

Geolocation

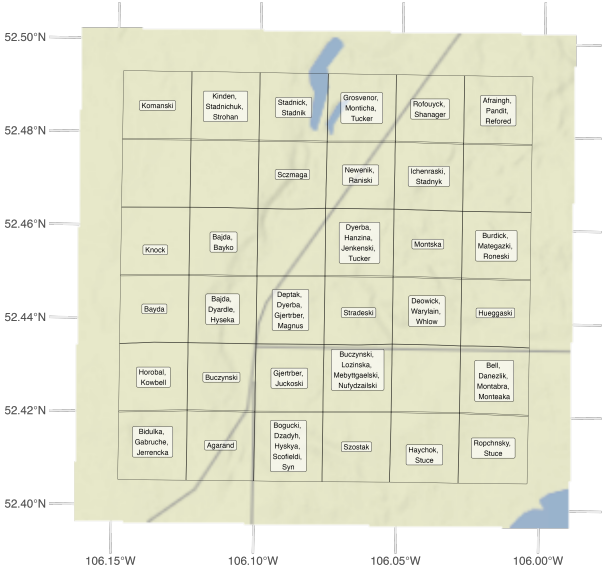
- The widespread use of the survey boundaries as addresses is unique to the Canadian prairies. ▸ Details
- In other projects (UK, US) it's typical to be able to locate 20-30% of rural households, and then only to the nearest town.
- This project: Geolocate 72% of rural households to the section number (1 mile square) in 1906 census, 84% in 1921, and 71% in 1931. ▸ Household Rates
- Schools are located in a similar fashion, with some geolocations being less precise.
  - I'm able to locate 91.8% of schools relative to the number listed as operating. ▸ School Counts

An Example Family



NUMBER	NAME AND RESIDENCE.						
Dwelling house	Family or institution	NAME of each person in family, household or institution.	PLACE OF ABODE. (In rural localities give parish, section, township, range and meridian. In cities, towns and villages, give street and number of dwelling.)				
			Section	Township	Range	Meridian	Municipality.
NUMÉRO		NOM ET RÉSIDENCE.					
Habitation	Famille, institution	NOM de chaque personne dans la famille, le ménage ou l'institution.	DOMICILE. (Dans les localités rurales, subdiviser la paroisse, la section, le rang et le méridien. Dans les cités, les villes et les villages, spécifier la rue et le numéro de la maison.)				
			Section.	Canton.	Rang.	Méridien.	Municipalité.
1	2	3	4	4a	4b	4c	5
83	83	Dzadyk Ilko	5	40	1	3	"
		Zelka	"	"	"	"	"
		Nicia	"	"	"	"	"
		Mike	"	"	"	"	"
		John	"	"	"	"	"
		Dratis	"	"	"	"	"
		Mary	"	"	"	"	"
		Pete	"	"	"	"	"

Family and Surrounding Township



Canadian Census Linking

- Need to link census observations in 1906 with later waves in 1911 and 1931.
  - Unlike the US there are no standard set of links available for these census waves.
- Implement the MLP method used at IPUMS to link US census records.
  - Use US training data, training on the set of features that can be recreated with the Canadian data.
  - Match 26% of children from 1906-1911, and 16% from 1906-1931. ▸ Rates ▸ Methodology
  - Lower than MLP results for the United States (46.3%), similar to ABE (26.5%)
  - Expected as the Canadian immigration is much higher at this time.

Bias From Linking - Makes Estimates More Conservative

- Earlier linking methods used on immutable characteristics.
- Newer methods like MLP and the newest ABE use mutable characteristics to differentiate observations that are otherwise observationally equivalent (Abramitzky et al., 2024).
  - John Smith in Ontario in 1911 is more likely to be in Ontario in 1921 than in Nova Scotia.
  - Get linking performance improvements for shorter distances ▸ MLP-ABE Distance.
  - Bias works against effects on occupation changes and migration.
- More likely to match people who are younger ▸, men ▸, wealthier ▸, and native to Ontario or the United Kingdom ▸.

.

Motivating Facts

How does distance affect schooling?

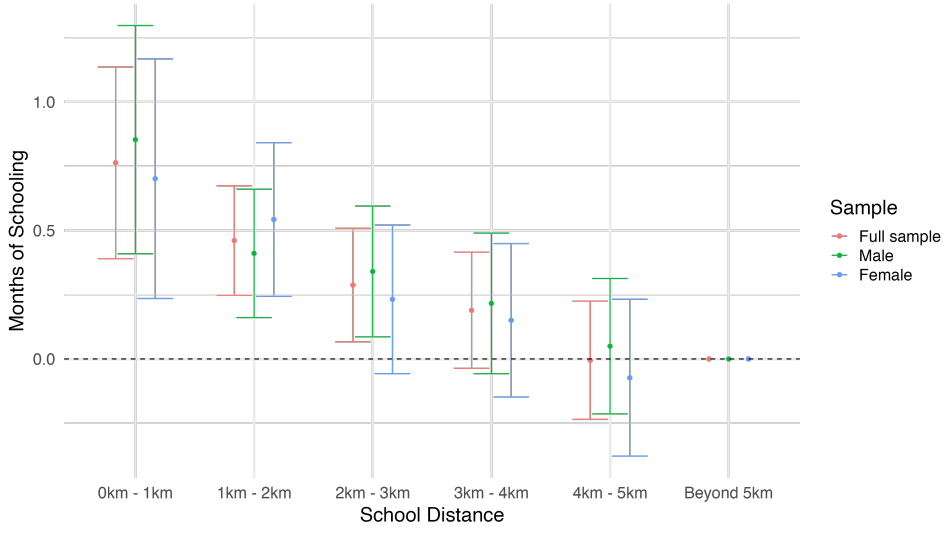
- Distance is mentioned repeatedly as a barrier to school attendance, and as one of the arguments for one room schools.
  - School locations don't overlap with railways, towns, post offices, or other amenities.

  ▸ Map
- Link 1906 (have locations) with 1911 (have school attendance), restrict sample to households that didn't move.
- Estimate effect of distance on school attendance in 1km bins:

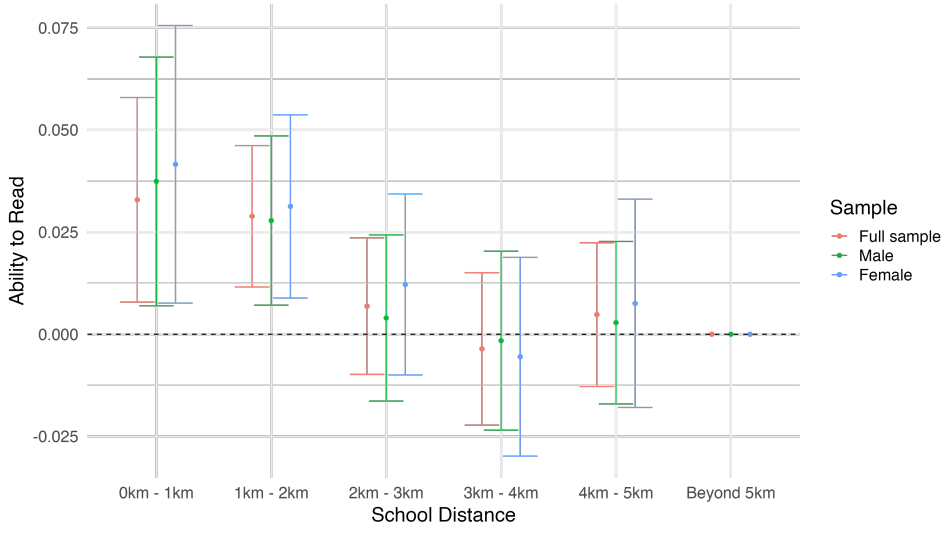
$$y_i = \sum_{k=1}^{5} \beta_k \mathbb{1}\{d \in k_{1km}\} + X_i + p_i + \epsilon_i \tag{1}$$

- Include sex by age by birth order, birthplace, and enumeration subdistrict fixed effects, family economic and railway distance controls.

Schooling vs Distance



Ability to Read vs Distance



Distance Decreases Schooling

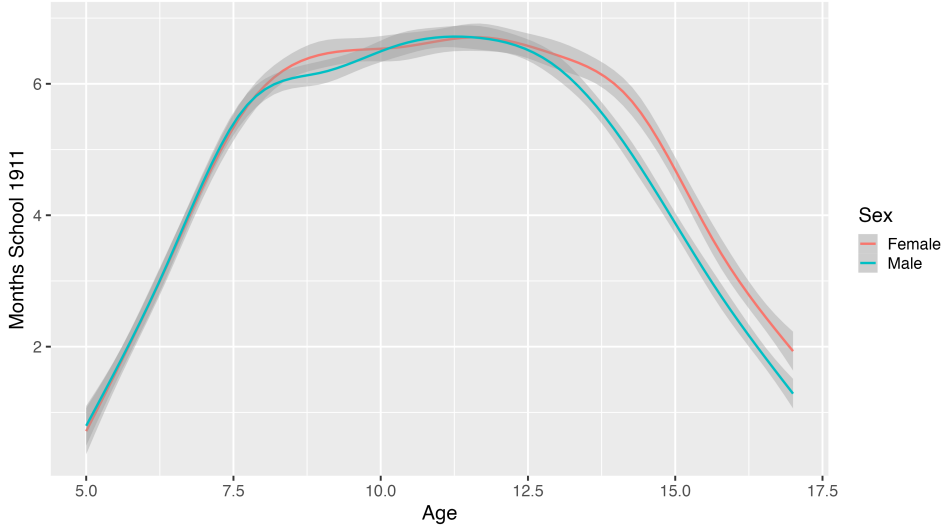
- Similar effects on both intensive and extensive margins ▸ Intensive ▸ Extensive.
- In regressions, find an elasticity of -0.039 between log months of school and log distance to school, and coefficient of -0.027 between likelihood of attending school and distance to school.
- Note that the effects dissipate to zero around 5km mark.
  - Policy was to make school districts no larger than 20 square miles, or $\approx 4km$ in radius.

Impact of School Construction on Later Life

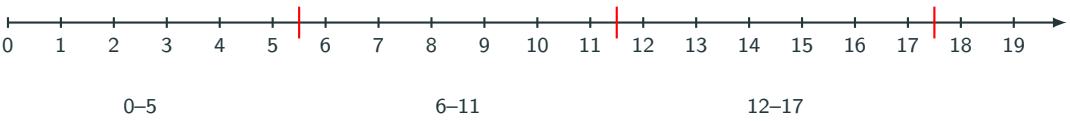
Difference in Differences

- School location is likely endogenous with a variety of factors.
- Solution: Exploit the timing of schools, comparing children who had schools constructed nearby (treated) to those who didn't (control) in younger versus older cohorts as in Duflo (2001).
  - Difference: treatment is defined at the individual instead of region.
- School construction was rapid, with a new school district being established every day by 1905 ▸ Graph.
- Assumptions:
  1. Children do not benefit from school construction after a certain age.
  2. Parents cannot manipulate when the school is constructed relative to the age of their children.
- Link 1906 census (boys with school distances) to 1931 (adults).

Schooling vs Age (Raw Data)

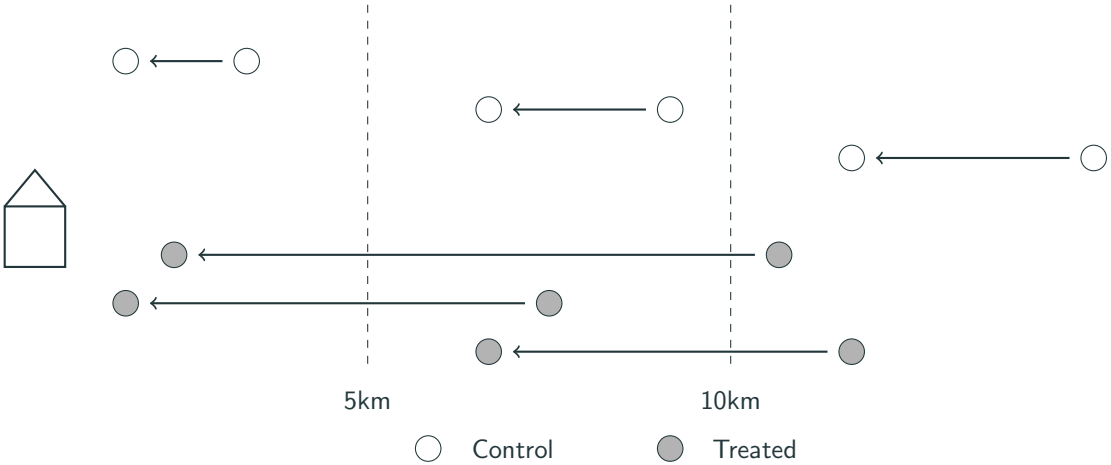


Difference in Differences



- Comparison between those in the 0-5 cohort ($t_c = \mathbb{1}\{age_{1906} <= 5\}$) in 1906, and those who are 12-17 in 1906.
- Considered treated if the distance to school decreased by 5km or more between 1900 to 1905, and the final distance is within 10km ($s_i = \mathbb{1}\{\Delta dist_{i,1900,1905} > 5km | dist_{i,1905} < 10km\}$).
- Include geographic fixed effects α_d, probability of link p_i and weight by observations by p_i.

Treated vs Control



Dependent Variable:	Family in Ag				
Model:	(1)	(2)	(3)	(4)	(5)
School Constr. × Cohort	-0.0409***	-0.0317**	-0.0331**	-0.0342**	-0.0346**
	(0.0151)	(0.0154)	(0.0154)	(0.0153)	(0.0153)
School Constr.	0.0112	0.0057	0.0060	0.0047	0.0049
	(0.0162)	(0.0164)	(0.0164)	(0.0165)	(0.0163)
Cohort	0.0394***	0.0452***	0.0626***	0.0626***	0.0618***
	(0.0091)	(0.0095)	(0.0093)	(0.0092)	(0.0092)
Log RR Distance (1906)					0.0210***
					(0.0045)
Link Prob.	0.4399***	0.4261***	0.4149***	0.4146***	0.4156***
	(0.0245)	(0.0245)	(0.0244)	(0.0245)	(0.0244)
Enumeration Subdistrict (1906)	Yes	Yes	Yes	Yes	Yes
Fam. Bpl.		Yes	Yes	Yes	Yes
Homestead			Yes	Yes	Yes
Livestock			Yes	Yes	Yes
Ag. Productive.				Yes	Yes
Observations	14,944	14,944	14,944	14,944	14,944
R^2	0.07873	0.08514	0.09582	0.09796	0.09986
Dependent variable mean	0.69279	0.69279	0.69279	0.69279	0.69279

Clustered (Enumeration Subdistrict (1906)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

▸ Farm Work ▸ Event Study

Dependent Variable:	Log Income				
Model:	(1)	(2)	(3)	(4)	(5)
School Constr. × Cohort	0.1012*	0.1108*	0.1160**	0.1213**	0.1226**
	(0.0566)	(0.0581)	(0.0587)	(0.0590)	(0.0591)
School Constr.	0.0130	0.0083	0.0052	0.0058	0.0050
	(0.0541)	(0.0552)	(0.0558)	(0.0549)	(0.0546)
Cohort	-0.3057***	-0.2693***	-0.2639***	-0.2594***	-0.2585***
	(0.0276)	(0.0294)	(0.0306)	(0.0307)	(0.0308)
Log RR Distance (1906)					-0.0318**
					(0.0123)
Link Prob.	-0.2310***	-0.1887***	-0.1915***	-0.1966***	-0.2018***
	(0.0620)	(0.0620)	(0.0618)	(0.0620)	(0.0621)
Enumeration Subdistrict (1906)	Yes	Yes	Yes	Yes	Yes
Fam. Bpl.		Yes	Yes	Yes	Yes
Homestead			Yes	Yes	Yes
Livestock			Yes	Yes	Yes
Ag. Productive.				Yes	Yes
Observations	5,750	5,750	5,750	5,750	5,750
R^2	0.15606	0.16376	0.16727	0.17174	0.17288
Dependent variable mean	6.5231	6.5231	6.5231	6.5231	6.5231

Clustered (Enumeration Subdistrict (1906)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

▸ Event Study

Migration Out of Agriculture

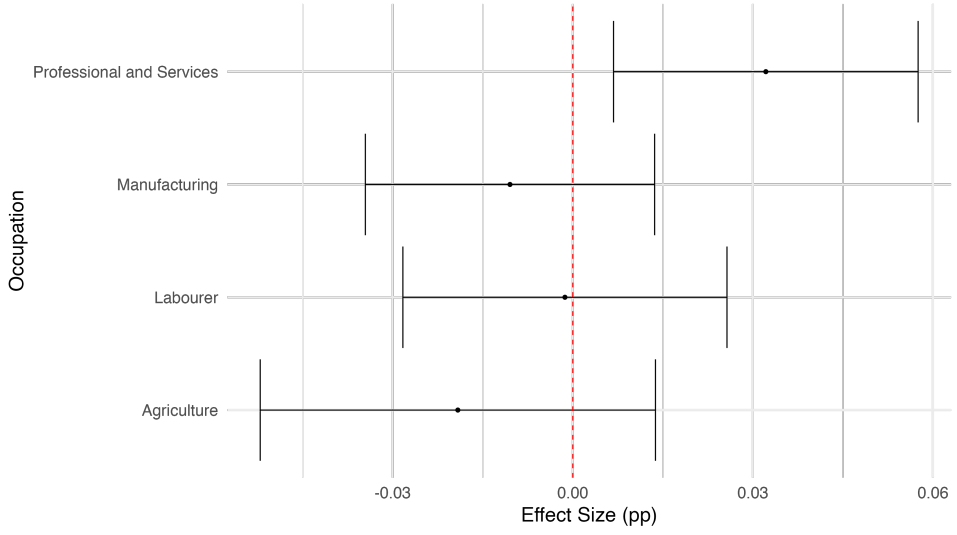
- Treated cohorts are also likely to live further away from home.
- Some evidence that they are more likely to be employees.
- These results hold while varying the treatment cutoff, how treatment is defined, and the cohorts used comparison.
- Results are concentrated in those that are furthest from the railway. Robustness
- If people are migrating out of agriculture, what are they doing?

Changes in Occupations

Switch to Services

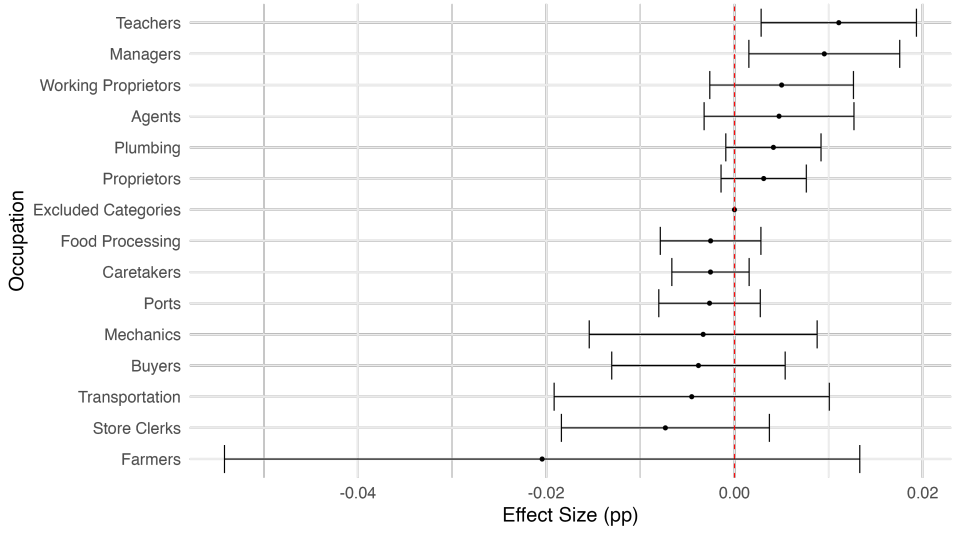
- Can use occupation codes to determine what types of occupations people are switching into.
- Start by looking at four primary industries: Agriculture, Manufacturing, Services, and General Labour.
- Then further divide into 2 digit OCCHISCO codes.

Primary Industries: Bypassing Manufacturing



Switching to High Skill Services

All



But in Different Industries than Urbanites

Agents who grew up in Rural Areas

Industry	n
Unknown	131
Grain Elevator	56
Insurance	47
General	12
Real Estate	9
Steam Railway	9
General Farm	6
Lumber Mill	5
Retail Hardware	5
Steam Railroad	5
Other	28

Agents who grew up in Urban Areas

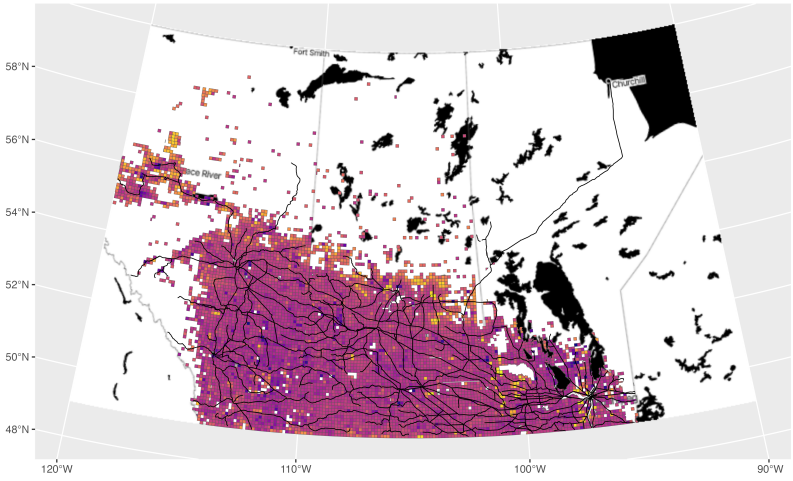
Industry	n
Unknown	97
Insurance	29
Real Estate	10
Grain Elevator	7
Steam Railway	5
Mill Farm	4
Wholesale Grocery	3
Financial	2
Automobile	1
Barber Shop	1
Other	15

Land Consolidation

Land Consolidation

- If people are leaving agriculture and farms are primarily owner operated, then we would expect farm sizes to increase.
- Don't observe farm sizes directly, only the location of the farm.
- Can infer farm sizes by looking at the density of neighbors.
  - A farmer with fewer neighbors is likely to have a larger farm.
- Imperfect measure with a lot of noise.
  - Validate by looking at correlations with other farm characteristics and persistence in farm size. ▸ Characteristics ▸ Persistance

Population Change from 1921 to 1931



Land Consolidation - Triple Differences

- Can't use linked census waves because cohort timing doesn't apply to land.
- Instead, look at the entire set of farms observable in 1906, 1921, and 1931, and use the set of schools open in 1912, with opening dates between 1895 and 1912.
- Define a close school if it was constructed within 5km of the parcel ($p_i = \mathbb{1}\{dist_i < 5km\}$).
- Define early school construction as areas that got a school before 1906 ($s_j = \mathbb{1}\{date_j < 1906\}$).
- Use 1906 as the reference period for 1921 and 1931 (T_t).
- Include school fixed effects α_j

$$y_{ijt} = \beta p_i s_j T_t + \text{Interactions}_{i,j,t} + \alpha_j + \epsilon_{i,j,t} \tag{3}$$

Dependent Variable:	Ln Farm Area				
Model:	(1)	(2)	(3)	(4)	(5)
Early × Close × Year = 1931	0.0612*	0.0415	0.0453*	0.0477*	0.0484*
	(0.0325)	(0.0260)	(0.0258)	(0.0260)	(0.0260)
Early × Close × Year = 1921	0.0223	0.0104	0.0117	0.0104	0.0106
	(0.0312)	(0.0258)	(0.0256)	(0.0257)	(0.0257)
Early × Close	-0.0446*	-0.0298	-0.0343	-0.0342	-0.0340
	(0.0265)	(0.0220)	(0.0219)	(0.0221)	(0.0221)
Homestead			-0.2121***	-0.2135***	-0.2135***
			(0.0050)	(0.0050)	(0.0050)
RR Dist in Year				0.0407***	0.0407***
				(0.0035)	(0.0035)
Family Size					-0.0006
					(0.0005)
Fam. Adult Males					0.0045***
					(0.0016)
Other Terms	Yes	Yes	Yes	Yes	Yes
School District		Yes	Yes	Yes	Yes
Observations	439,027	439,027	439,027	439,027	439,027
R^2	0.00614	0.13406	0.15113	0.15336	0.15384

Clustered (Enum. Sub.) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Land Consolidation - Not Driven by Settlement

- This result is robust to changing the cutoff boundary and early versus late year selection. ▸ Boundary ▸ Year
- Importantly find the same result if looking only at farms in 1921 and 1931, after the majority of settlement occurs. ▸ Year
  - Possible concern is that 1906 is in the middle of the settlement period so could be picking up general geographic expansion.

Conclusion

Conclusion

- This project has focused on the role that one room schools play in the development of rural areas.
- Exposure to these schools lead to agricultural outmigration, and switching towards high skill services.
- This encouraged the consolidation of land into larger farms.
- Want to think about mechanisms next. Two possibilities:
  1. Higher returns to human capital outside of agriculture.
  2. One room schools reveal underlying ability in non-agricultural sectors.

Thank you!

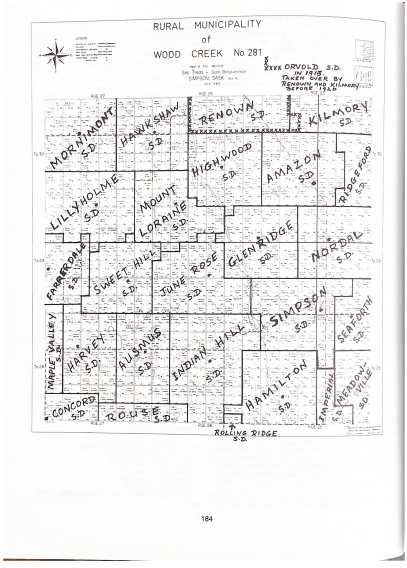
School Formation

Back

- Establishing a school district required petitioning the provincial government and a ratepayer vote.
  - Sometimes, votes would fail due to fear of higher taxes.
- If a school district was formed, the location of the school would be located in the middle of the district. Example
  - Some minor exceptions if land could not be acquired.
  - Settlers could not to anticipate the location of the school when settling.
- Distance was a key concern for school attendance.
  - Students were exempt from mandatory schooling in Alberta (1910) if they lived further than 2 miles from the nearest school.
  - Same in Saskatchewan after 1917, with a distance of 2.5 miles.

School Board Layout

Back



Quotes

Back

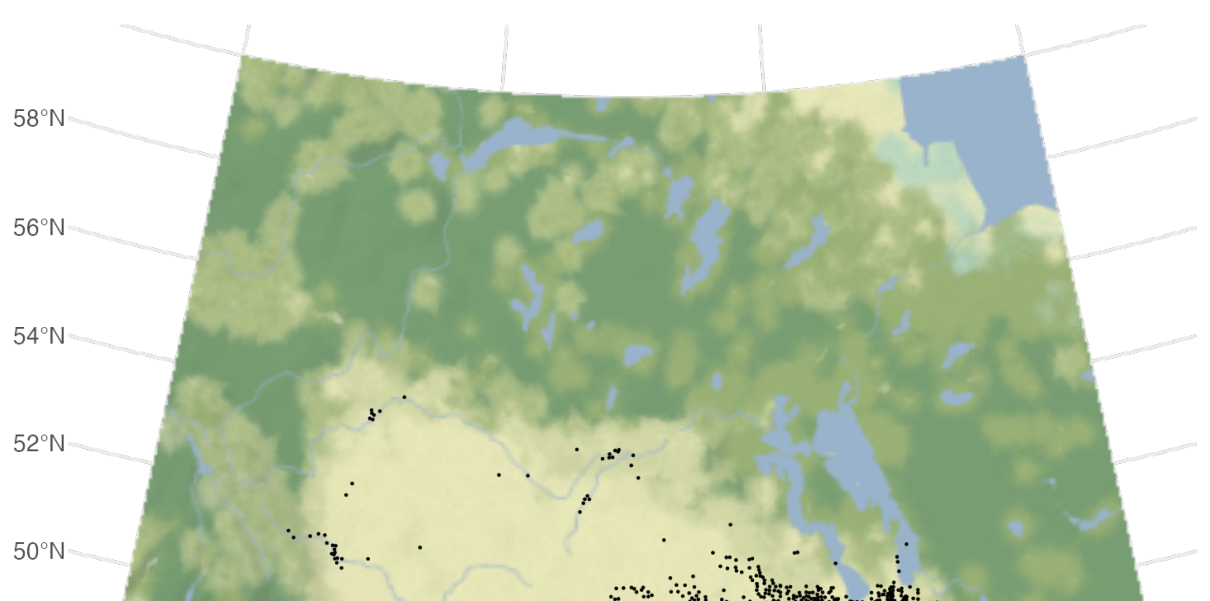
- *"The attendance in rural districts is not so regular as one would expect or could hope for. Undoubtably, it is in part due to the scarcity of farm help and the long distances some children have to go; but I fear it is also caused by a lack of interest on the part of many parents"*, H.H. Smith, Inspector of Schools, Saskatoon, December 31, 1909
- *"When a boy is old enough to work an outfit on a farm, his school days are over"*, John S. Huff, Davidson, Saskatchewan., March 1st, 1912

Historical School Locations

Back to Data

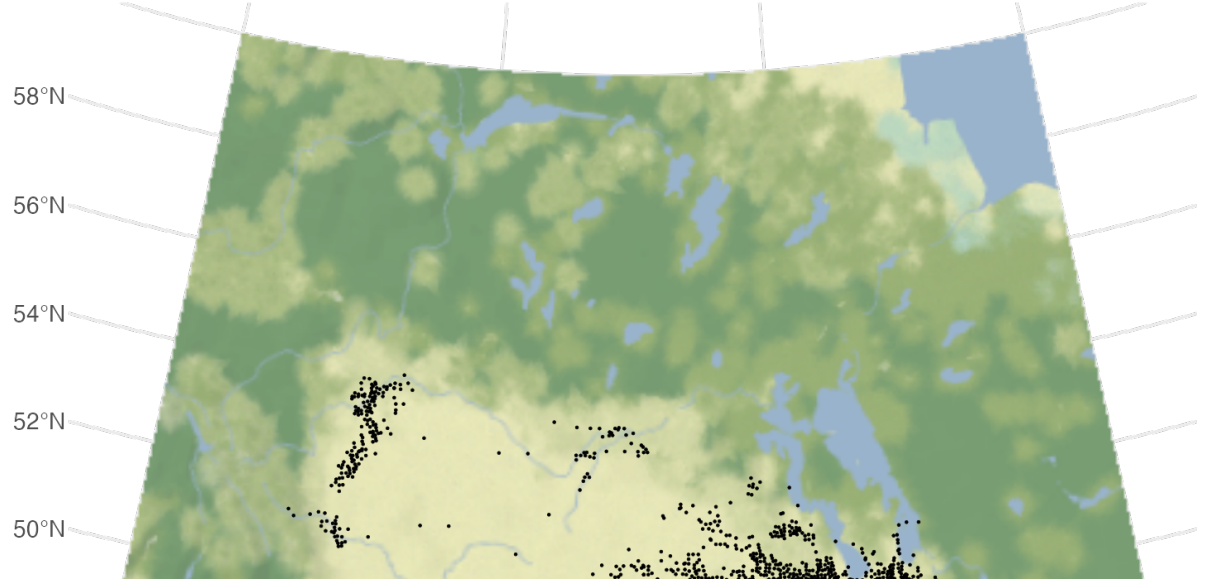
- Panel of school openings from 1871 to 1912 covering Alberta, Saskatchewan and Manitoba.
- Merge four independent sources of data for schools in the prairies:
  - Alberta: Glenbow Museum Archives
  - Saskatchewan: One Room School Project, Annual reports of the Department of Education
  - Manitoba: Manitoba Historical Society
- Saskatchewan data doesn't include the opening date for each school, but can infer date using Alberta school opening dates.
  - Alberta and Saskatchewan were both part of the Northwest Territories until 1905, so share a school numbering system until that time.

Schools 1890



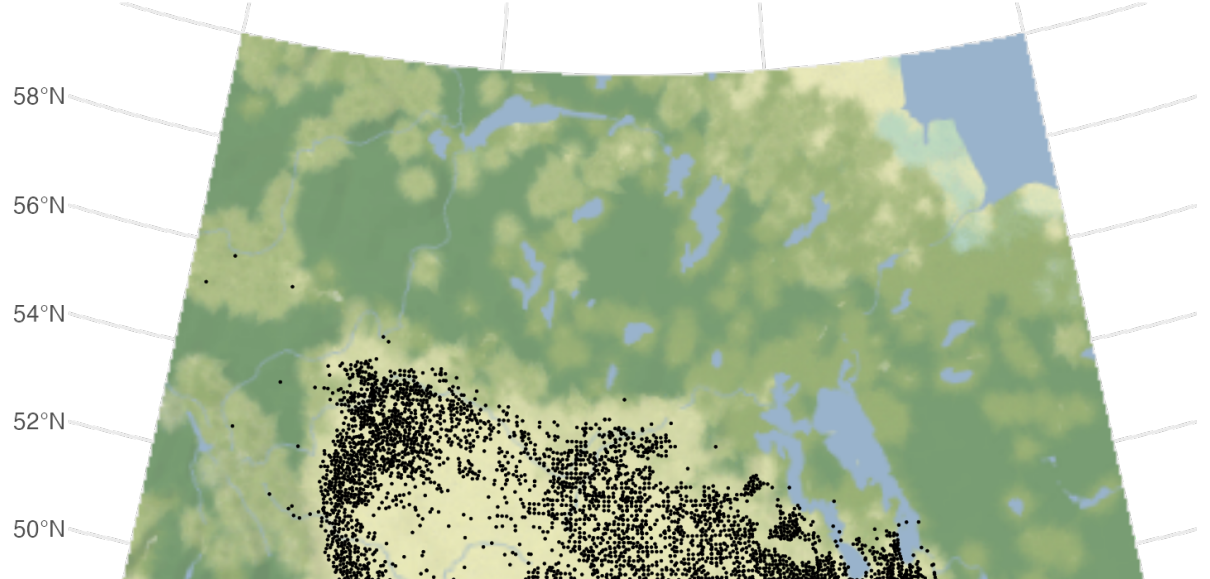
Schools 1900

Back to Data



Schools 1910

Back to Data



Other Data

Back to Data

- Cummins Maps. Example
  - Alberta, Saskatchewan, and Manitoba property ownership maps (approximately 1915 - 1930).
  - Some have been digitized, currently negotiating to get labelled versions for Manitoba and Saskatchewan.
- Land Grants of Western Canada, 1870-1930, CPR Land Sales Records, Hudsons Bay Company Land Records.
- Massey-Harris dealer records.

Household Geolocation Rates

Back to Geolocation

Table 1: Geolocation rates by census year for Alberta, Saskatchewan, and Manitoba. Exact geolocations are at the section level or in CSDs that are less than 1.5 miles in area. General geolocations are those that are at the township level or finer. CSD geolocations are at the CSD level. Splitting the geolocation rate by urban and rural households reveals that the majority of geolocations in both census waves are for the rural households.

	Total	Exact	General	CSD	Exact (Rural)	Exact (Urban)
1906	100.00	55.57	84.41	100.00	71.94	0.00
1921	100.00	65.97	69.61	100.00	83.89	36.35
1931	100.00	42.76	86.57	100.00	70.81	0.00

School Geolocation Rates

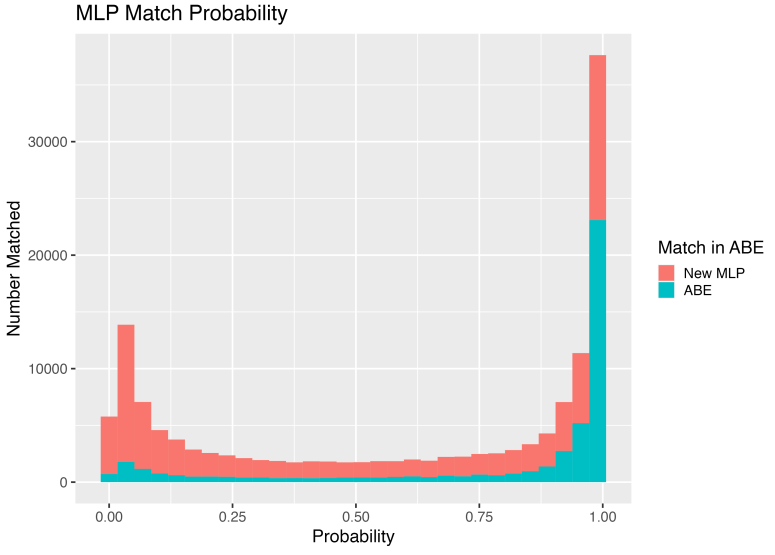
Back to Geolocation

Table 2: Summary statistics by province for the number of schools in 1911. The first row lists the number of schools reports in the 1911 report from the department of education for Alberta (Government of Alberta, 1912), Saskatchewan (Saskatchewan Department of Education, 1906, 1907, 1908, 1910, 1911, 1912, 1914), and Manitoba (Government of Manitoba, 1911, 1912). This figure includes both public and separate schools in Alberta and Saskatchewan while the data contain only public schools, so is a slight overestimate. The second row, "Observed" lists the number of schools I find in the data in 1911, which requires that they have an opening date before 1911. Any schools with an unknown opening date are dropped. "Geolocated" reports the number of schools, of those which are observed, which I am able to geolocate. "Geolocated (Section)" reports the number of schools where the geolocation is at the section level (1 mile square) or finer.

Type	Total	Alberta	Saskatchewan	Manitoba
Actual	5955	1784	2573	1598
Observed	5618	1750	2370	1498
Geolocated	5472	1750	2227	1495
Geolocated (Section)	4141	1619	1027	1495

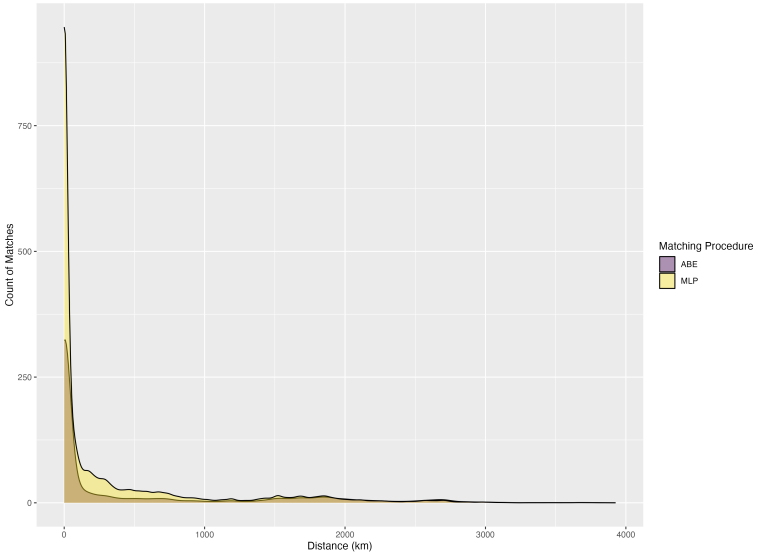
Link Probability MLP 1906-1921

Back to Linking



Distance from 1906-1911 by Linking Method

Back to Linking



Matching Rates

Back to Linking | By Province | Linking Counts

	Total	Matched	Children	Adults	Male	Female
1906-1911	100.00	22.80	26.69	21.43	26.75	17.76
1906-1921	100.00	19.73	20.59	20.12	25.62	12.02
1906-1931	100.00	16.20	16.27	16.91	23.02	7.18
1911-1921	100.00	19.95	23.06	18.25	24.69	14.67
1911-1931	100.00	13.37	13.92	13.19	18.86	7.21

Linking Details

Back to Linking

- MLP works in two stages:
  1. Generate all possible matches between two years. Use logit/machine learning to score matches. Take matches above a minimum threshold, and above a threshold for the next best match.
  2. Use these matches to identify families, then do a second stage of matching within families.
- Currently using the training data from the MLP replication package to train a model for the Canadian data.

Matching Counts

Back to Linking Rates Back to Linking

	Total	Matched	Children	Adults	Male	Female
1906-1911	802940	183049	80170	102879	122854	60112
1906-1921	802940	158416	61840	96576	117681	40694
1906-1931	802940	130060	48864	81196	105740	24303
1911-1921	7197003	1436085	638288	797797	940674	495213
1911-1931	7197003	961982	385183	576799	718432	243467

Match Rate by Province

Back to Linking Rates | Back to Linking

	Canada	AB	SK	MB	ON	QC
1906-1911	22.80	22.95	21.36	23.75		
1906-1921	19.73	18.98	19.17	20.51		
1906-1931	16.20	16.01	16.05	16.40		
1911-1921	19.95	18.22	18.41	19.92	21.21	18.19
1911-1931	13.37	14.40	14.28	14.74	13.60	12.21

Match Rate by Age

Back to Linking Bias

Dependent Variable:	Matched				
Model:	(1)	(2)	(3)	(4)	(5)
	1906-1911	1906-1921	1906-1931	1911-1921	1911-1931
Age = 1	0.0730***	0.0786***	0.0673***	0.0686***	0.0367***
Age = 2	0.0074*	0.0155***	0.0583***	-0.0178***	0.0132***
Age = 3	-0.0355***	0.0100***	0.0521***	-0.0302***	0.0088***
Age = 4	-0.0083**	0.0209***	0.0443***	0.0052***	0.0201***
Age = 6	-0.0268***	-0.0569***	-0.0872***	-0.0325***	-0.0582***
Age = 7	-0.0458***	-0.1299***	-0.1473***	-0.0806***	-0.1143***
Age = 8	-0.0642***	-0.1663***	-0.1502***	-0.1197***	-0.1272***
Age = 9	-0.1267***	-0.1751***	-0.1359***	-0.1434***	-0.1211***
Age = 10	-0.1999***	-0.1869***	-0.1474***	-0.1499***	-0.1134***
Age = 11				-0.1073***	-0.0706***
Age = 12				-0.1558***	-0.1033***
Sex	0.1307***	0.1465***	0.1467***	0.1197***	0.1084***
Family Size	0.0041***	0.0010***	0.0001	-5.8×10^{-5}	-7.95×10^{-5}***
Birthplace	Yes	Yes	Yes	Yes	Yes
Relation	Yes	Yes	Yes	Yes	Yes
Enum. Subdist.	Yes	Yes	Yes	Yes	Yes
Homestead	Yes	Yes	Yes		
Milk Cows	Yes	Yes	Yes		
Horses	Yes	Yes	Yes		
Wheat Prod.	Yes	Yes	Yes		
Religion				Yes	Yes

Match Rate by Relation

Back to Linking Bias

Dependent Variable:	Matched				
Model:	(1)	(2)	(3)	(4)	(5)
	1906-1911	1906-1921	1906-1931	1911-1921	1911-1931
Family Size	0.0043***	0.0013***	0.0004	-4.14×10^{-5}	-6.45×10^{-5}***
Relation = Boarder	-0.1487***	-0.1326***	-0.1166***	-0.1034***	-0.0951***
Relation = Daughter	-0.1340***	-0.2163***	-0.2192***	-0.1301***	-0.1627***
Relation = Lodger	-0.1398***	-0.1533***	-0.1187***	-0.1049***	-0.0944***
Relation = Other	-0.1284***	-0.1156***	-0.0941***	-0.0972***	-0.0842***
Relation = Sister	-0.1942***	-0.2063***	-0.1794***	-0.1756***	-0.1637***
Relation = Son	0.0110***	0.0118***	0.0425***	0.0319***	0.0435***
Relation = Wife	-0.0379***	-0.0605***	-0.0819***	-0.0285***	-0.0458***
Age	Yes	Yes	Yes	Yes	Yes
Birthplace	Yes	Yes	Yes	Yes	Yes
Enum. Subdist.	Yes	Yes	Yes	Yes	Yes
Homestead	Yes	Yes	Yes		
Milk Cows	Yes	Yes	Yes		
Horses	Yes	Yes	Yes		
Wheat Prod.	Yes	Yes	Yes		
Religion				Yes	Yes
Observations	430,101	430,101	430,101	1,191,247	1,191,247
R^2	0.08044	0.08238	0.09728	0.05538	0.06190

Clustered (Enum. Subdist.) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Match Rate by Livestock

Back to Linking Bias

Dependent Variable:	Matched		
Model:	(1)	(2)	(3)
	1906-1911	1906-1921	1906-1931
Sex	0.1313***	0.1467***	0.1466***
Family Size	0.0041***	0.0011***	9.13×10^{-5}
Milk Cows = 1-4	0.0266***	0.0266***	0.0199***
Milk Cows = 5-9	0.0537***	0.0475***	0.0372***
Milk Cows = 10-49	0.0618***	0.0495***	0.0403***
Milk Cows = 50+	0.0173	0.0080	-0.0086
Horses = 1-4	0.0311***	0.0280***	0.0205***
Horses = 5-9	0.0547***	0.0499***	0.0385***
Horses = 10-49	0.0634***	0.0559***	0.0436***
Horses = 50+	0.0406**	0.0231*	0.0205**
Age	Yes	Yes	Yes
Birthplace	Yes	Yes	Yes
Relation	Yes	Yes	Yes
Province	Yes	Yes	Yes
Homestead	Yes	Yes	Yes
Wheat Prod.	Yes	Yes	Yes
Observations	430,101	430,101	430,101
R^2	0.05674	0.06955	0.08882

Clustered (Enum. Subdist.) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Match Rate by Homestead

Back to Linking Bias

Dependent Variable:		Matched	
Model:	(1)	(2)	(3)
	1906-1911	1906-1921	1906-1931
Sex	0.1313***	0.1467***	0.1466***
Family Size	0.0041***	0.0011***	9.13×10^{-5}
Homestead = HBC	0.0059	-0.0019	-0.0015
Homestead = Railway	0.0004	0.0018	0.0012
Homestead = School	0.0079	-0.0006	0.0009
Age	Yes	Yes	Yes
Birthplace	Yes	Yes	Yes
Relation	Yes	Yes	Yes
Province	Yes	Yes	Yes
Milk Cows	Yes	Yes	Yes
Horses	Yes	Yes	Yes
Wheat Prod.	Yes	Yes	Yes
Observations	430,101	430,101	430,101
R^2	0.05674	0.06955	0.08882

Clustered (Enum. Subdist.) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Match Rate by Province

Back to Linking Bias

Dependent Variable:	Matched				
Model:	(1)	(2)	(3)	(4)	(5)
	1906-1911	1906-1921	1906-1931	1911-1921	1911-1931
Sex	0.1313***	0.1467***	0.1466***	0.1196***	0.1083***
Family Size	0.0041***	0.0011***	9.13×10^{-5}	-0.0001***	-0.0001***
Province = AB	0.0034	-0.0093	-0.0044	-0.0145***	-0.0003
Province = SK	-0.0114	-0.0077	-0.0072	-0.0099**	-0.0011
Age	Yes	Yes	Yes	Yes	Yes
Birthplace	Yes	Yes	Yes	Yes	Yes
Relation	Yes	Yes	Yes	Yes	Yes
Homestead	Yes	Yes	Yes		
Milk Cows	Yes	Yes	Yes		
Horses	Yes	Yes	Yes		
Wheat Prod.	Yes	Yes	Yes		
Religion				Yes	Yes
Observations	430,101	430,101	430,101	1,191,247	1,191,247
R^2	0.05674	0.06955	0.08882	0.04816	0.05726

Clustered (Enum. Subdist.) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

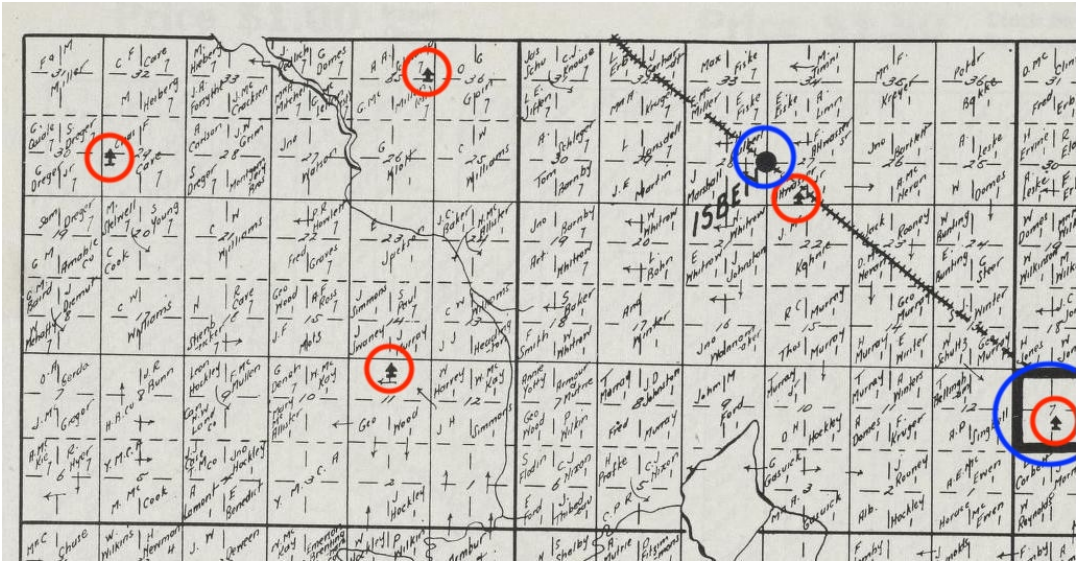
Match Rate by Birthplace

Back to Linking Bias

Dependent Variable:	Matched				
Model:	(1) 1906-1911	(2) 1906-1921	(3) 1906-1931	(4) 1911-1921	(5) 1911-1931
Sex	0.1307***	0.1465***	0.1467***	0.1197***	0.1084***
Family Size	0.0041***	0.0010***	0.0001	-5.8×10^{-5}	-7.95×10^{-5}***
Birthplace = UnitedStates	-0.0347***	-0.0533***	-0.0462***	-0.0260***	-0.0283***
Birthplace = Canada+BC	-0.0491***	-0.0425***	-0.0391***	-0.0131***	-0.0101***
Birthplace = MB	-0.0378***	-0.0389***	-0.0417***	0.0113***	0.0062***
Birthplace = Maritimes	-0.0322***	-0.0213***	-0.0155**	-0.0111***	-0.0132***
Birthplace = QC	-0.0563***	-0.0268***	-0.0199***	0.0067	0.0032
Birthplace = Caribbean	-0.1479***	-0.0886***	-0.1132***	-0.0441**	-0.0571***
Birthplace = Nordic	-0.0760***	-0.0547***	-0.0334***	-0.0327***	-0.0133***
Birthplace = UnitedKingdom	0.0319***	0.0322***	0.0358***	0.0003	0.0090***
Birthplace = Ireland	-0.0157*	-0.0178**	-0.0030	-0.0009	0.0065*
Birthplace = RestofEurope	-0.1094***	-0.0683***	-0.0469***	-0.0306***	-0.0115***
Birthplace = ROW	-0.0722***	-0.1175***	-0.0977***	-0.0556***	-0.0498***
Age	Yes	Yes	Yes	Yes	Yes
Relation	Yes	Yes	Yes	Yes	Yes
Enum. Subdist.	Yes	Yes	Yes	Yes	Yes
Homestead	Yes	Yes	Yes		
Milk Cows	Yes	Yes	Yes		
Horses	Yes	Yes	Yes		
Wheat Prod.	Yes	Yes	Yes		
Religion				Yes	Yes

Cummins Map Example

Back to Data | Back to Schooling Choice



Townships and Sections

Back to Geolocations

- We can geolocate households throughout the prairies because of how land was surveyed.
- Canadian prairies were surveyed in a grid, with 6 mile by 6 mile Townships, and 1 mile by 1 mile sections.
- In census records, administrative documents, and colloquial use, addresses were expressed using the survey boundaries of land.
    - These correspond to the modern day land divisions, allowing for precise geolocations.

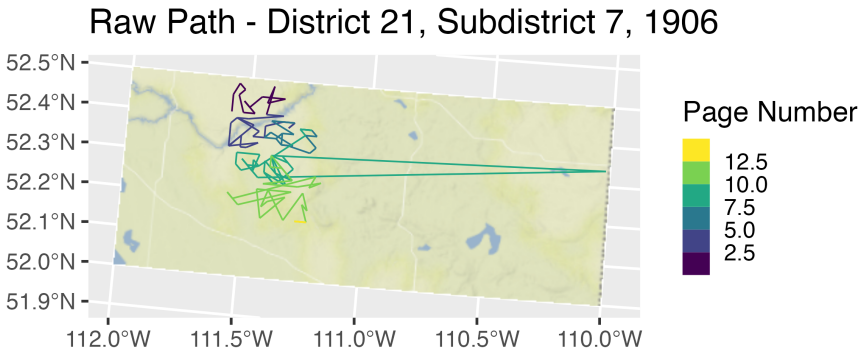
Townships

Back to Geolocations

PUBLIC ARCHIVES CANADA
NATIONAL MAP COLLECTION
NO REPRODUCTION WITHOUT PERMISSION
NMC 0043265
ARCHIVES PUBLIQUES CANADA
COLLECTION NATIONALE DE CARTES ET PLANS
REPRODUCTION INTERDITE SANS PERMISSION

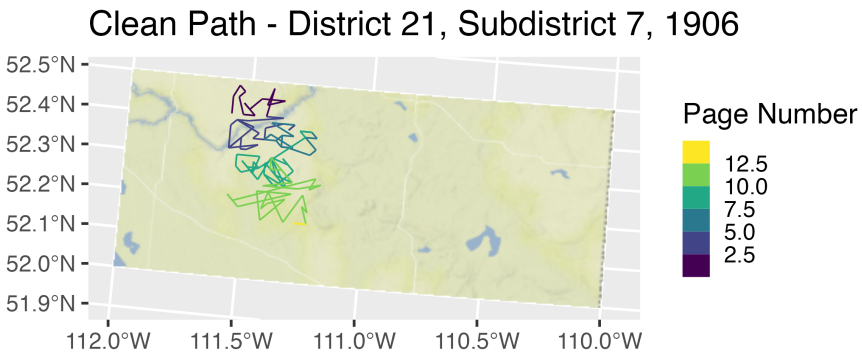
Enumerator Paths (Raw)

Back to Geolocations



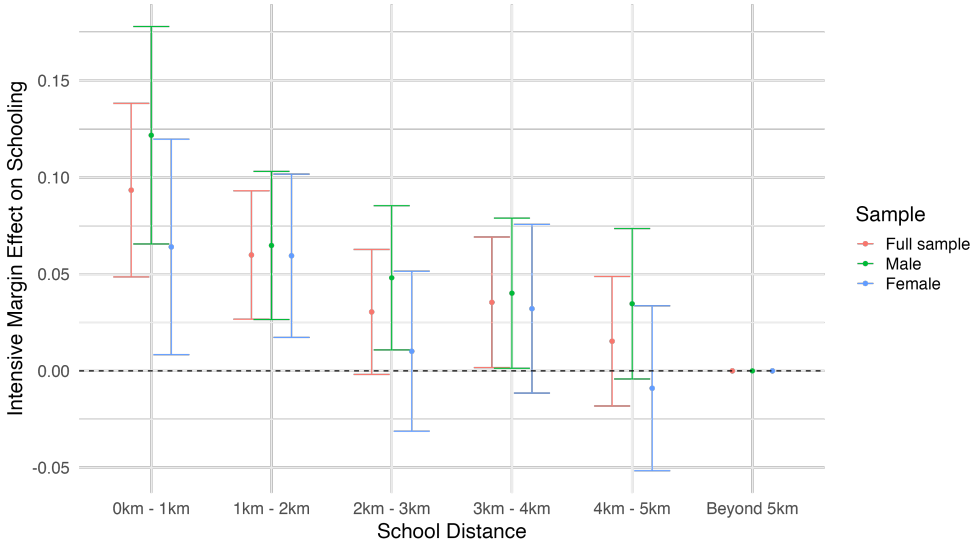
Enumerator Paths (Cleaned)

Back to Geolocations



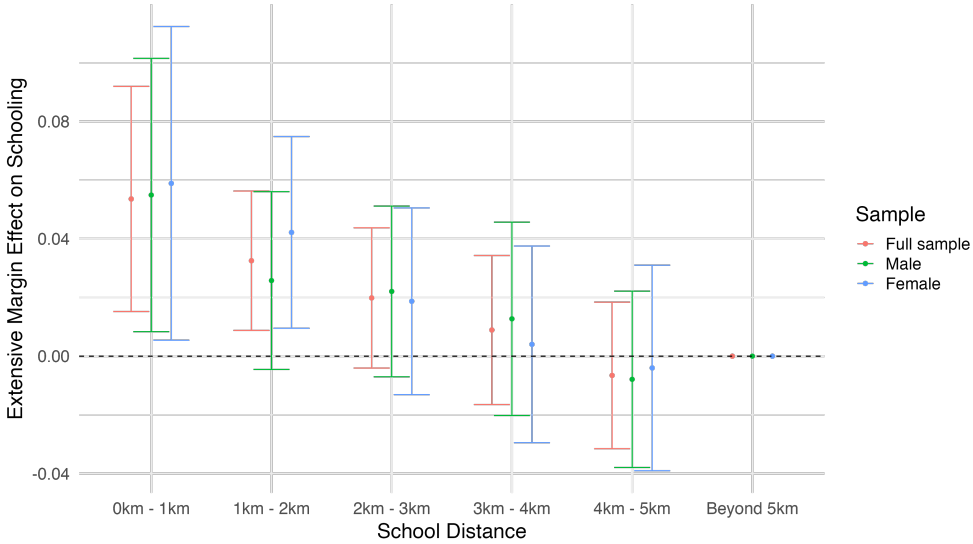
Schooling vs Distance: Intensive Margin

Back



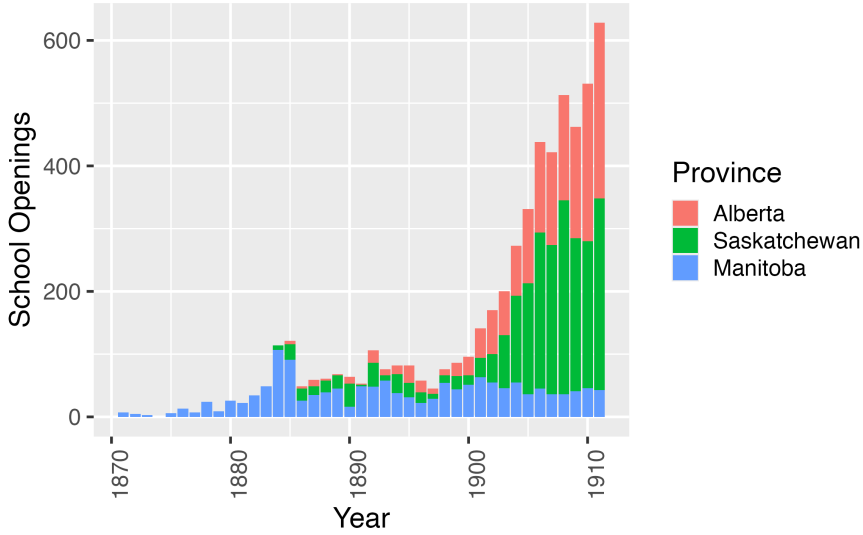
Schooling vs Distance: Extensive Margin

Back



Annual School Openings

Back to DiD



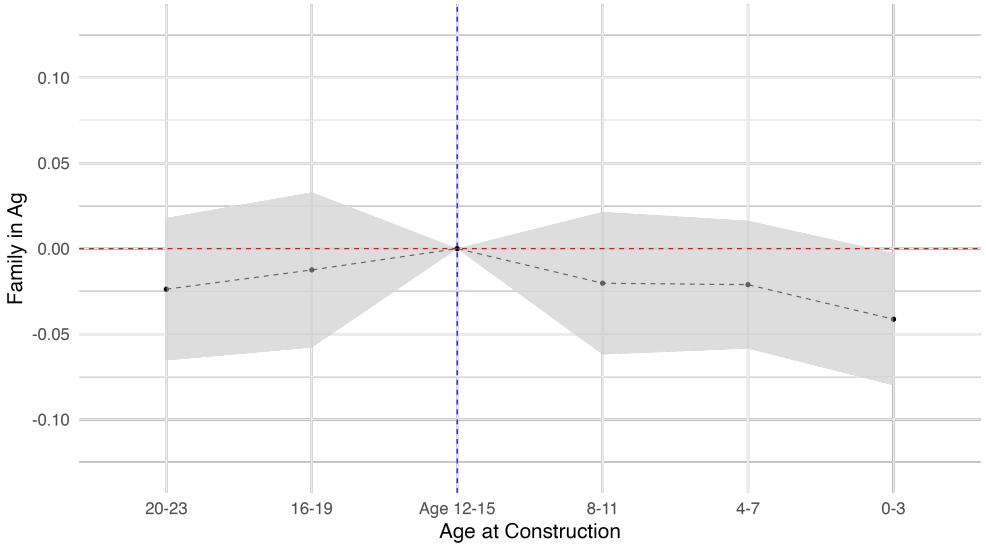
Dependent Variable:	Farm Work				
Model:	(1)	(2)	(3)	(4)	(5)
School Constr. × Cohort	-0.0405**	-0.0304*	-0.0318*	-0.0322*	-0.0326*
	(0.0167)	(0.0169)	(0.0169)	(0.0169)	(0.0169)
School Constr.	0.0108	0.0047	0.0055	0.0040	0.0042
	(0.0172)	(0.0174)	(0.0174)	(0.0175)	(0.0173)
Cohort	-0.0191**	-0.0115	0.0085	0.0084	0.0076
	(0.0092)	(0.0097)	(0.0098)	(0.0098)	(0.0098)
Log RR Distance (1906)					0.0205***
					(0.0046)
Link Prob.	0.4780***	0.4656***	0.4533***	0.4539***	0.4549***
	(0.0247)	(0.0251)	(0.0248)	(0.0249)	(0.0249)
Enumeration Subdistrict (1906)	Yes	Yes	Yes	Yes	Yes
Fam. Bpl.		Yes	Yes	Yes	Yes
Homestead			Yes	Yes	Yes
Livestock			Yes	Yes	Yes
Ag. Productive.				Yes	Yes
Observations	14,944	14,944	14,944	14,944	14,944
R^2	0.07941	0.08531	0.09614	0.09733	0.09893
Dependent variable mean	0.62018	0.62018	0.62018	0.62018	0.62018

Clustered (Enumeration Subdistrict (1906)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Back

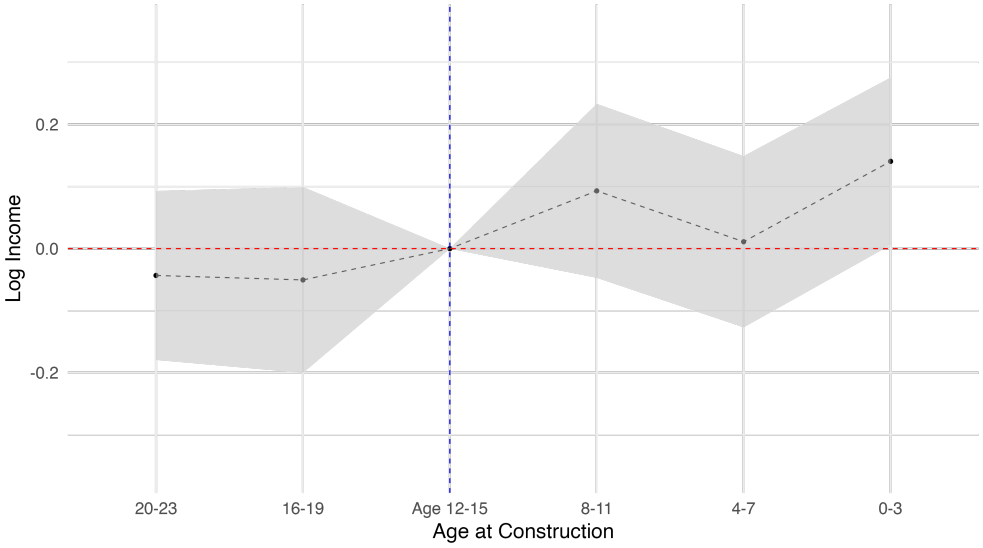
Family in Agriculture Event Study

Back



Income Event Study

Back



Dependent Variables:	Migration Dist. (km)	Asinh Migration. Dist.	Log Migration Dist.	On Family Farm
Model:	(1)	(2)	(3)	(4)
School Constr. × Cohort	28.62***	0.1542*	0.1268*	-0.0129
	(10.18)	(0.0850)	(0.0729)	(0.0174)
School Constr.	-2.942	-0.0327	-0.0290	0.0004
	(8.529)	(0.0789)	(0.0712)	(0.0157)
Cohort	-4.966	-0.1574***	0.0155	0.0385***
	(6.259)	(0.0500)	(0.0447)	(0.0095)
Log RR Distance (1906)	-3.407	-0.0244	-0.0044	0.0108***
	(2.448)	(0.0227)	(0.0197)	(0.0038)
Link Prob.	-971.6***	-5.512***	-4.698***	0.5013***
	(27.55)	(0.1231)	(0.1039)	(0.0188)
Enumeration Subdistrict (1906)	Yes	Yes	Yes	Yes
Fam. Bpl. + Homestead	Yes	Yes	Yes	Yes
Livestock + Ag.	Yes	Yes	Yes	Yes
Observations	14,944	14,944	13,217	14,944
R^2	0.28142	0.22182	0.24087	0.12132
Dependent variable mean	224.04	4.2159	4.0652	0.23367

Clustered (Enumeration Subdistrict (1906)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Back

Dependent Variable:			Employee		
Model:	(1)	(2)	(3)	(4)	(5)
School Constr. × Cohort	0.0176	0.0103	0.0115	0.0120	0.0124
	(0.0179)	(0.0179)	(0.0178)	(0.0178)	(0.0177)
School Constr.	0.0050	0.0113	0.0096	0.0130	0.0129
	(0.0155)	(0.0154)	(0.0153)	(0.0155)	(0.0154)
Cohort	0.0970***	0.0878***	0.0710***	0.0711***	0.0717***
	(0.0101)	(0.0108)	(0.0108)	(0.0108)	(0.0108)
Log RR Distance (1906)					-0.0181***
					(0.0045)
Link Prob.	-0.3419***	-0.3348***	-0.3247***	-0.3260***	-0.3269***
	(0.0219)	(0.0221)	(0.0220)	(0.0221)	(0.0221)
Enumeration Subdistrict (1906)	Yes	Yes	Yes	Yes	Yes
Fam. Bpl.		Yes	Yes	Yes	Yes
Homestead			Yes	Yes	Yes
Livestock			Yes	Yes	Yes
Ag. Productive.				Yes	Yes
Observations	14,944	14,944	14,944	14,944	14,944
R^2	0.07386	0.07982	0.08833	0.09050	0.09181
Dependent variable mean	0.34475	0.34475	0.34475	0.34475	0.34475

Clustered (Enumeration Subdistrict (1906)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

▸ Employer ▸ Own Account ▸ Back

Dependent Variable:			Employer		
Model:	(1)	(2)	(3)	(4)	(5)
School Constr. × Cohort	-0.0035	-0.0026	-0.0021	-0.0028	-0.0028
	(0.0155)	(0.0156)	(0.0157)	(0.0156)	(0.0156)
School Constr.	0.0118	0.0102	0.0117	0.0108	0.0108
	(0.0139)	(0.0139)	(0.0139)	(0.0140)	(0.0140)
Cohort	-0.0802***	-0.0815***	-0.0697***	-0.0694***	-0.0695***
	(0.0066)	(0.0072)	(0.0073)	(0.0073)	(0.0073)
Log RR Distance (1906)					0.0016
					(0.0034)
Link Prob.	0.1212***	0.1207***	0.1130***	0.1143***	0.1144***
	(0.0164)	(0.0167)	(0.0165)	(0.0165)	(0.0165)
Enumeration Subdistrict (1906)	Yes	Yes	Yes	Yes	Yes
Fam. Bpl.		Yes	Yes	Yes	Yes
Homestead			Yes	Yes	Yes
Livestock			Yes	Yes	Yes
Ag. Productive.				Yes	Yes
Observations	14,944	14,944	14,944	14,944	14,944
R^2	0.07100	0.07377	0.07988	0.08248	0.08249
Dependent variable mean	0.13597	0.13597	0.13597	0.13597	0.13597

Clustered (Enumeration Subdistrict (1906)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

▸ Employee ▸ Own Account ▸ Back

Dependent Variable:	Own Account				
Model:	(1)	(2)	(3)	(4)	(5)
School Constr. × Cohort	-0.0065	0.0054	0.0053	0.0063	0.0061
	(0.0194)	(0.0198)	(0.0199)	(0.0199)	(0.0199)
School Constr.	-0.0166	-0.0280*	-0.0283*	-0.0292*	-0.0291*
	(0.0166)	(0.0165)	(0.0166)	(0.0167)	(0.0166)
Cohort	-0.1493***	-0.1304***	-0.1250***	-0.1255***	-0.1260***
	(0.0091)	(0.0098)	(0.0100)	(0.0100)	(0.0099)
Log RR Distance (1906)					0.0120***
					(0.0045)
Link Prob.	0.1872***	0.1922***	0.1876***	0.1879***	0.1885***
	(0.0222)	(0.0221)	(0.0221)	(0.0222)	(0.0222)
Enumeration Subdistrict (1906)	Yes	Yes	Yes	Yes	Yes
Fam. Bpl.		Yes	Yes	Yes	Yes
Homestead			Yes	Yes	Yes
Livestock			Yes	Yes	Yes
Ag. Productive.				Yes	Yes
Observations	14,292	14,292	14,292	14,292	14,292
R^2	0.09792	0.11031	0.11388	0.11487	0.11541
Dependent variable mean	0.35502	0.35502	0.35502	0.35502	0.35502

Clustered (Enumeration Subdistrict (1906)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

▸ Employee ▸ Employer ▸ Back

Dependent Variable:	Log Income				
Model:	(1) 3km	(2) 4km	(3) 5km	(4) 6km	(5) 7km
School Constr. × Cohort	0.1010*	0.0815	0.1226**	0.1299**	0.1316**
	(0.0592)	(0.0560)	(0.0591)	(0.0587)	(0.0632)
School Constr.	0.0134	0.0346	0.0050	-0.0259	-0.0106
	(0.0498)	(0.0517)	(0.0546)	(0.0560)	(0.0588)
Cohort	-0.2508***	-0.2490***	-0.2585***	-0.2603***	-0.2590***
	(0.0305)	(0.0312)	(0.0308)	(0.0307)	(0.0308)
Log RR Distance (1906)	-0.0314**	-0.0315**	-0.0318**	-0.0320***	-0.0323***
	(0.0123)	(0.0123)	(0.0123)	(0.0123)	(0.0124)
Link Prob.	-0.2032***	-0.2024***	-0.2018***	-0.2023***	-0.2022***
	(0.0621)	(0.0621)	(0.0621)	(0.0621)	(0.0621)
Enumeration Subdistrict (1906)	Yes	Yes	Yes	Yes	Yes
Fam. Bpl. + Homestead	Yes	Yes	Yes	Yes	Yes
Livestock + Ag.	Yes	Yes	Yes	Yes	Yes
Observations	5,750	5,750	5,750	5,750	5,750
R^2	0.17262	0.17262	0.17288	0.17257	0.17269
Dependent variable mean	6.5231	6.5231	6.5231	6.5231	6.5231

Clustered (Enumeration Subdistrict (1906)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent Variable:	Log Income			
Model:	(1) Baseline	(2) Vs Close	(3) Vs Far	(4) Vs Urban
School Constr. × Cohort	0.1226**	0.1725**	0.0988	0.1490**
	(0.0591)	(0.0720)	(0.0706)	(0.0624)
School Constr.	0.0050	-0.0245	0.0388	-0.4199***
	(0.0546)	(0.0682)	(0.0553)	(0.0582)
Cohort	-0.2585***	-0.3148***	-0.2112***	-0.3872***
	(0.0308)	(0.0396)	(0.0543)	(0.0336)
Log RR Distance (1906)	-0.0318**	-0.0482***	-0.0216	
	(0.0123)	(0.0167)	(0.0196)	
Link Prob.	-0.2018***	-0.1819**	-0.1329	0.2660***
	(0.0621)	(0.0759)	(0.0876)	(0.0674)
Enumeration Subdistrict (1906)	Yes	Yes	Yes	
Birthplace	Yes	Yes	Yes	Yes
Father Birthplace	Yes	Yes	Yes	Yes
Mother Birthplace	Yes	Yes	Yes	Yes
Homestead	Yes	Yes	Yes	
Enumeration District (1906)				Yes
Fam. Bpl. + Homestead	Yes	Yes	Yes	partial
Livestock + Ag.	Yes	Yes	Yes	
Observations	5,750	3,857	3,095	3,986
R^2	0.17288	0.19322	0.23183	0.12405
Dependent variable mean	6.5231	6.5445	6.5006	6.8325

Clustered (Enumeration Subdistrict (1906)) standard errors in parentheses

Dependent Variable:	Log Income					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
	0-5 vs 6-11	0-5 vs 12-17	0-5 vs 18-23	6-11 vs 12-17	6-11 vs 18-23	12-17 vs 18-23
Variables						
School Constr. × Cohort	0.0468	0.1226**	0.1831***	0.1101*	0.0947*	0.0663
	(0.0528)	(0.0591)	(0.0529)	(0.0618)	(0.0551)	(0.0627)
School Constr.	-0.0258	0.0050	-0.1015**	-0.0711	-0.1556***	-0.0158
	(0.0488)	(0.0546)	(0.0505)	(0.0577)	(0.0537)	(0.0555)
Cohort	-0.1857***	-0.2585***	-0.2225***	-0.0921***	-0.0480	0.0187
	(0.0273)	(0.0308)	(0.0338)	(0.0292)	(0.0313)	(0.0314)
Log RR Distance (1906)	-0.0251**	-0.0318**	-0.0177	-0.0332***	-0.0275**	-0.0372***
	(0.0120)	(0.0123)	(0.0125)	(0.0116)	(0.0118)	(0.0132)
Link Prob.	-0.1089*	-0.2018***	-0.1134*	-0.2240***	-0.1455**	-0.2348***
	(0.0632)	(0.0621)	(0.0627)	(0.0667)	(0.0641)	(0.0666)
Fixed-effects						
Enumeration Subdistrict (1906)	Yes	Yes	Yes	Yes	Yes	Yes
Fam. Bpl. + Homestead	Yes	Yes	Yes	Yes	Yes	Yes
Livestock + Ag.	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	6,223	5,750	6,091	5,253	5,594	5,121
R^2	0.15385	0.17288	0.17415	0.16089	0.15167	0.15958
Dependent variable mean	6.4950	6.5231	6.5400	6.6499	6.6607	6.7075

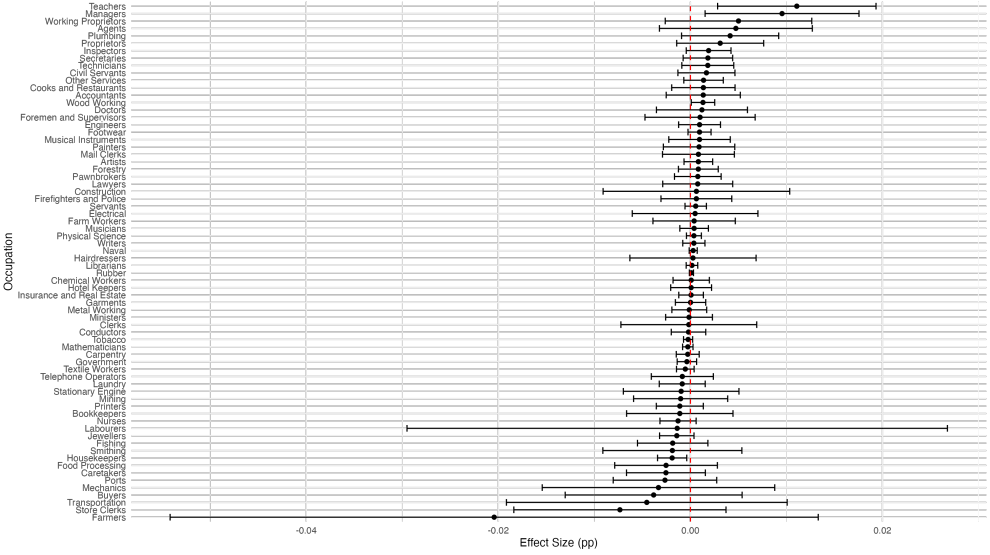
Clustered (Enumeration Subdistrict (1906)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent Variable:	Family in Ag		
Model:	(1)	(2)	(3)
	Full Sample	Within 5km of Railway	Beyond 5km of Railway
School Constr. × Cohort	-0.0342**	0.0108	-0.0503***
	(0.0153)	(0.0367)	(0.0191)
School Constr.	0.0047	0.0165	0.0047
	(0.0165)	(0.0406)	(0.0185)
Cohort	0.0626***	0.0537***	0.0684***
	(0.0092)	(0.0138)	(0.0127)
Link Prob.	0.4146***	0.3764***	0.4408***
	(0.0245)	(0.0406)	(0.0299)
Enumeration Subdistrict (1906)	Yes	Yes	Yes
Fam. Bpl. + Homestead	Yes	Yes	Yes
Livestock + Ag.	Yes	Yes	Yes
Observations	14,944	5,371	9,573
R^2	0.09796	0.16115	0.11051
Dependent variable mean	0.69279	0.65388	0.71461

Clustered (Enumeration Subdistrict (1906)) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Switching to High Skill Services

Back



Dependent Variable:	Ln Farm Area				
Model:	(1)	(2)	(3)	(4)	(5)
	1906	1921		1931	
Asinh Horses	0.0559***				
	(0.0049)				
Asinh Milk Cows	-0.0004				
	(0.0051)				
Asinh Hogs	-0.0112***				
	(0.0035)				
Asinh Cattle	0.0234***				
	(0.0032)				
Asinh Sheep	0.0057				
	(0.0059)				
Family Size	-0.0024	-0.0063***	-0.0044*	0.0030***	0.0012
	(0.0017)	(0.0009)	(0.0023)	(0.0008)	(0.0019)
Fam. Adult Males	0.0271***	0.0206***	0.0152**	0.0202***	0.0131***
	(0.0049)	(0.0028)	(0.0070)	(0.0026)	(0.0047)
Homestead	-0.3169***	-0.1730***	-0.1624***	-0.1699***	-0.1637***
	(0.0186)	(0.0072)	(0.0181)	(0.0062)	(0.0139)
Asinh Wheat	-0.0138	0.1437**	-0.0287	0.0074	0.0635
	(0.1390)	(0.0617)	(0.1968)	(0.0862)	(0.1766)
Asinh Oats	0.0756	-0.7318***	-0.6965**	-0.2634	-0.3276
	(0.2470)	(0.1506)	(0.3189)	(0.1626)	(0.2926)
Asinh Grass	-0.0049	-0.2574***	-0.1207*	0.0626	0.0454
	(0.0152)	(0.0599)	(0.0681)	(0.0581)	(0.0558)
Asinh Flax	0.0050	0.0970	0.0754	-0.1193	-0.1052
	(0.0450)	(0.0669)	(0.0757)	(0.0761)	(0.1003)
Asinh Barley	-0.1012	0.7859***	0.8071***	0.2907**	0.3371
	(0.2248)	(0.1405)	(0.3037)	(0.1320)	(0.2523)
Employer		0.0405***	-0.0096	0.0290***	0.0207
		(0.0137)	(0.0293)	(0.0049)	(0.0127)
Employee		-0.2295***	-0.2749***	-0.2082***	-0.2459***
		(0.0149)	(0.0336)	(0.0111)	(0.0160)
Log Income			-0.0181***		-0.0233***
			(0.0068)		(0.0084)

Dependent Variable:	Farm Area			
Model:	(1)	(2)	(3)	(4)
	Movers	On FF	Movers	On FF
	1906-1921		1906-1931	
Farm Area (1906)	0.0395***	0.0883***	0.0295*	0.0863***
	(0.0111)	(0.0112)	(0.0152)	(0.0160)
Homestead (1906)	0.0440*	0.0150	-0.0392	-0.0042
	(0.0233)	(0.0263)	(0.0322)	(0.0383)
Homestead (1921)	-0.2108***	-0.1662***		
	(0.0224)	(0.0243)		
Homestead (1931)			-0.1303***	-0.1251***
			(0.0274)	(0.0370)
Enum. Sub. (1921)	Yes	Yes		
Enum. Sub. (1921)			Yes	Yes
Observations	6,079	7,203	4,517	4,541
R^2	0.20421	0.21970	0.46946	0.43374

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent Variable:	Ln Farm Area				
Model:	(1)	(2)	(3)	(4)	(5)
	3km	4km	5km	6km	7km
Early × Close × Year = 1931	0.0226	0.0484*	0.0529*	0.0739*	0.1096**
	(0.0259)	(0.0260)	(0.0304)	(0.0407)	(0.0540)
Early × Close × Year = 1921	-0.0149	0.0106	0.0078	0.0225	0.0625
	(0.0259)	(0.0257)	(0.0288)	(0.0386)	(0.0571)
Early × Close	-0.0216	-0.0340	-0.0133	-0.0064	-0.0399
	(0.0223)	(0.0221)	(0.0243)	(0.0322)	(0.0457)
Homestead	-0.2125***	-0.2135***	-0.2131***	-0.2131***	-0.2130***
	(0.0050)	(0.0050)	(0.0050)	(0.0050)	(0.0050)
RR Dist in Year	0.0403***	0.0407***	0.0408***	0.0408***	0.0409***
	(0.0035)	(0.0035)	(0.0035)	(0.0035)	(0.0035)
Family Size	-0.0005	-0.0006	-0.0006	-0.0006	-0.0006
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Fam. Adult Males	0.0044***	0.0045***	0.0045***	0.0045***	0.0046***
	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)
Other Terms	Yes	Yes	Yes	Yes	Yes
School District	Yes	Yes	Yes	Yes	Yes
Observations	439,027	439,027	439,027	439,027	439,027
R^2	0.15431	0.15384	0.15382	0.15383	0.15378

Clustered (Enum. Sub.) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent Variable:	Ln Farm Area				
Model:	(1)	(2)	(3)	(4)	(5)
	1904	1905	1906	1907	1908
Early × Close × Year = 1931	0.0720*	0.0675*	0.0529*	0.0610**	0.0343
	(0.0408)	(0.0347)	(0.0304)	(0.0311)	(0.0317)
Early × Close × Year = 1921	-0.0117	0.0009	0.0078	0.0228	-0.0025
	(0.0348)	(0.0318)	(0.0288)	(0.0289)	(0.0303)
Early × Close	0.0060	-0.0089	-0.0133	-0.0234	0.0128
	(0.0286)	(0.0258)	(0.0243)	(0.0248)	(0.0274)
Homestead	-0.2125***	-0.2129***	-0.2131***	-0.2132***	-0.2132***
	(0.0050)	(0.0050)	(0.0050)	(0.0050)	(0.0050)
RR Dist in Year	0.0420***	0.0414***	0.0408***	0.0406***	0.0407***
	(0.0035)	(0.0035)	(0.0035)	(0.0035)	(0.0035)
Family Size	-0.0007	-0.0006	-0.0006	-0.0005	-0.0005
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Fam. Adult Males	0.0048***	0.0046***	0.0045***	0.0044***	0.0044***
	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)
Other Terms	Yes	Yes	Yes	Yes	Yes
School District	Yes	Yes	Yes	Yes	Yes
Observations	439,027	439,027	439,027	439,027	439,027
R^2	0.15403	0.15387	0.15382	0.15385	0.15390

Clustered (Enum. Sub.) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent Variable:	Ln Farm Area	
Model:	(1)	(2)
	Baseline	Drop 1906
Early × Close × Year = 1931	0.0529*	0.0459**
	(0.0304)	(0.0216)
Early × Close	-0.0133	-0.0045
	(0.0243)	(0.0149)
Early × Close × Year = 1921	0.0078	
	(0.0288)	
Homestead	-0.2131***	-0.1850***
	(0.0050)	(0.0047)
RR Dist in Year	0.0408***	0.0519***
	(0.0035)	(0.0032)
Family Size	-0.0006	-0.0014**
	(0.0005)	(0.0006)
Fam. Adult Males	0.0045***	0.0068***
	(0.0016)	(0.0017)
Other Terms	Yes	Yes
School District	Yes	Yes
Observations	439,027	358,129
R^2	0.15382	0.17134

Clustered (Enum. Sub.) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Why do Schools Open Sooner in Some Places?

- Rural common schools were constructed by local communities, who organized together.
- Goldin and Katz argue that smaller communities were more homogenous, had higher levels of social capital, and therefore constructed schools sooner (Goldin & Katz, 2008).
- Dippel and Ottinger makes a similar argument, that small decentralized school boards facilitated increased population heterogeneity (Dippel et al., 2020).
- Can test this by looking at the local determinants of school opening.
  - Look at the set of schools that opened between 1900-1912
  - Link schools to the characteristics of the people that lived nearby (in 1906 when observe locations).
- Standardize variables to compare, each variable is expressed in standard deviations.

	School Year Established		
	(1)	(2)	(3)
RR Distance 1899 (Log)	0.5537***	0.5375***	0.5328***
	(0.1440)	(0.1436)	(0.1441)
School Distance 1899 (Log)	0.2362*	0.2338*	0.2316*
	(0.1363)	(0.1363)	(0.1360)
Local Children 1906 (Count)	-0.4225***	-0.4704***	-0.4915***
	(0.1197)	(0.1201)	(0.1211)
Local Households 1906 (Count)	-0.1476	-0.1502	-0.1467
	(0.0998)	(0.0987)	(0.0991)
Cattle 1906 (Count)	0.0799	0.0762	0.0783
	(0.0618)	(0.0600)	(0.0601)
Hogs 1906 (Count)	-0.2889***	-0.2603**	-0.2613**
	(0.1092)	(0.1086)	(0.1089)
Milk Cows 1906 (Count)	-0.2925**	-0.2647**	-0.2560**
	(0.1175)	(0.1169)	(0.1177)
Area Entirely Homestead (Indicator)	0.1021*	0.0809	0.0821
	(0.0569)	(0.0570)	(0.0569)
Birthplace HHI (Percentage)	0.1668***	0.1691***	0.1568***
	(0.0596)	(0.0590)	(0.0593)
Religion HHI (Percentage)	0.1901***	0.1546**	0.1522**
	(0.0653)	(0.0651)	(0.0651)
Anglo-North American (Percentage)		-0.3393***	-0.3130***
		(0.0666)	(0.0699)
Methodist (Percentage)			-0.1027
			(0.0624)
Presbyterian (Percentage)			-0.1120*
			(0.0607)
Lutheran (Percentage)			-0.1215*
			(0.0676)
Observations	2,926	2,926	2,926
R^2	0.44815	0.45350	0.45483
Geographic Controls fixed effects	✓	✓	✓

References

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson (Aug. 1, 2012). **"Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration".** In: *American Economic Review* 102.5, pp. 1832–1856. ISSN: 0002-8282. DOI: 10.1257/aer.102.5.1832. URL: https://pubs.aeaweb.org/doi/10.1257/aer.102.5.1832 (visited on 09/04/2024).

— (June 2014). **"A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration".** In: *Journal of Political Economy* 122.3, pp. 467–506. ISSN: 0022-3808, 1537-534X. DOI: 10.1086/675805. URL: https://www.journals.uchicago.edu/doi/10.1086/675805 (visited on 01/22/2021).

Abramitzky, Ran, Roy Mill, and Santiago Pérez (Apr. 2, 2020). **"Linking Individuals across Historical Sources: A Fully Automated Approach*".** In: *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 53.2, pp. 94–111. ISSN: 0161-5440. DOI: 10.1080/01615440.2018.1543034. URL: https://doi.org/10.1080/01615440.2018.1543034 (visited on 06/20/2024).

Abramitzky, Ran et al. (2024). **"Finding John Smith: Using Extra Information for Historical Record Linkage".** In.

Althoff, Lukas, Harriet Brookes Gray, and Hugo Reichardt (2025). **"America's Rise in Human Capital Mobility".** In.

Antonie, Luiza et al. (Apr. 1, 2014). **"Tracking People over Time in 19th Century Canada for Longitudinal Analysis".** In: *Machine Learning* 95.1, pp. 129–146. ISSN: 1573-0565. DOI: 10.1007/s10994-013-5421-0. URL: https://doi.org/10.1007/s10994-013-5421-0 (visited on 06/20/2024).

Antonie, Luiza et al. (Oct. 2020). **"Selection Bias Encountered in the Systematic Linking of Historical Census Records".** In: *Social Science History* 44.3, pp. 555–570. ISSN: 0145-5532, 1527-8034. DOI: 10.1017/ssh.2020.15. URL: https://www.cambridge.org/core/journals/social-science-history/article/selection-bias-encountered-in-the-systematic-linking-of-historical-census-records/C1C07635B494232D7489B5D35230DBE1 (visited on 05/30/2024).

Bagagli, Sara (2023). **Once Upon a Time in the (Canadian) West: Railroads, Location Fundamentals, and the Growth of Cities.** Sara Bagagli. URL: https://sarabagagli.github.io/research/publication/ (visited on 03/21/2023).

Bailey, Martha J. et al. (Dec. 2020). **"How Well Do Automated Linking Methods Perform? Lessons from US Historical Data".** In: *Journal of Economic Literature* 58.4, pp. 997–1044. ISSN: 0022-0515. DOI: 10.1257/jel.20191526. URL: https://www.aeaweb.org/articles?id=10.1257/jel.20191526 (visited on 05/06/2025).

Buckles, Kasey et al. (Sept. 2023). **Breakthroughs in Historical Record Linking Using Genealogy Data: The Census Tree Project.** DOI: 10.3386/w31671. URL: https://www.nber.org/papers/w31671 (visited on 09/14/2023). Pre-published.

Budı-Ors, Tomas (2023). **"Rural-Urban Migration and Structural Change: A Reinterpretation".** In.

Card, David, Ciprian Domnisoru, and Lowell Taylor (Apr. 1, 2022a). **"The Intergenerational Transmission of Human Capital: Evidence from the Golden Age of Upward Mobility".** In: *Journal of Labor Economics* 40.S1, S39–S95. ISSN: 0734-306X, 1537-5307. DOI: 10.1086/718417. URL: https://www.journals.uchicago.edu/doi/10.1086/718417 (visited on 01/06/2025).

Card, David et al. (Sept. 2022b). **The Impact of Female Teachers on Female Students' Lifetime Well-Being.** w30430. Cambridge, MA: National Bureau of Economic Research, w30430. DOI: 10.3386/w30430. URL: http://www.nber.org/papers/w30430.pdf (visited on 01/09/2025).

Caselli, Francesco and Wilbur John Coleman II (2001). **"The U.S. Structural Transformation and Regional Convergence: A Reinterpretation".** In: *Journal of Political Economy* 109.3, pp. 584–616. ISSN: 0022-3808. DOI: 10.1086/321015. URL: https://www.jstor.org/stable/10.1086/321015 (visited on 11/20/2023).

Dippel, Christian, Dustin Frye, and Bryan Leonard (2020). **Property Rights without Transfer Rights: A Study of Indian Land Allotment.** DOI: 10.3386/w27479. URL: https://www.nber.org/papers/w27479 (visited on 09/18/2022). Pre-published.

Duflo, Esther (Sept. 1, 2001). **"Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment".** In: *American Economic Review* 91.4, pp. 795–813. ISSN: 0002-8282. DOI: 10.1257/aer.91.4.795. URL: https://pubs.aeaweb.org/doi/10.1257/aer.91.4.795 (visited on 06/27/2023).

Feigenbaum, James, Jonas Helgertz, and Joseph Price (2023). **Examining the Role of Training Data for Supervised Methods of Automated Record Linkage: Lessons for Best Practice in Economic History.** DOI: 10.2139/ssrn.4669534. URL: https://www.ssrn.com/abstract=4669534 (visited on 06/29/2024). Pre-published.

Feigenbaum, James J (2016). **"A Machine Learning Approach to Census Record Linking".** In.

French, Jacob (2022). **"Technological Change, Inequality, and Intergenerational Mobility: The Case of Early 20th Century Agriculture".** In: p. 61.

Gauthier, Jean-Francois, Teresa Molina, and Anant Nyshadham (2025). **"Learning and Selection in the Sorting of Households Across Sectors".** In.

Goldin, Claudia (1998). **"America's Graduation from High School: The Evolution and Spread of Secondary Schooling in the Twentieth Century".** In: *The Journal of Economic History* 58.2, pp. 345–374. ISSN: 0022-0507. URL: https://www.jstor.org/stable/2566738 (visited on 11/22/2023).

Goldin, Claudia and Lawrence F. Katz (1998). **"The Origins of Technology-Skill Complementarity".** In: *The Quarterly Journal of Economics* 113.3, pp. 693–732. ISSN: 0033-5533. URL: https://www.jstor.org/stable/2586871 (visited on 12/07/2023).

Goldin, Claudia and Lawrence F. Katz (Sept. 2000). **"Education and Income in the Early Twentieth Century: Evidence from the Prairies".** In: *The Journal of Economic History* 60.3, pp. 782–818. ISSN: 1471-6372, 0022-0507. DOI: 10.1017/S0022050700025766. URL: https://www.cambridge.org/core/journals/journal-of-economic-history/article/education-and-income-in-the-early-twentieth-century-evidence-from-the-prairies/0F5357717F81B36ADFDB9B8DC0C0A848 (visited on 12/09/2023).

— (2008). **The Race Between Education and Technology.** Cambridge, UNITED STATES: Harvard University Press. ISBN: 978-0-674-03773-1. URL: http://ebookcentral.proquest.com/lib/utoronto/detail.action?docID=3300082 (visited on 11/14/2024).

Government of Alberta (1912). **"Annual Report of the Department of Education of the Province of Alberta.".** In: *Annual report of the Department of Education of the Province of Alberta.*

Government of Manitoba (1911, 1912). **Report of the Department of Education.**

Helgertz, Jonas et al. (Jan. 2, 2022). **"A New Strategy for Linking U.S. Historical Censuses: A Case Study for the IPUMS Multigenerational Longitudinal Panel".** In: *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 55.1, pp. 12–29. ISSN: 0161-5440. DOI: 10.1080/01615440.2021.1985027. URL: https://doi.org/10.1080/01615440.2021.1985027 (visited on 05/24/2024).

Lachanski, Michael S. (2024). **The Elusive Rural U.S. Historical Educational Advantage: New Evidence for the 1881 to 1914 Cohorts.**

Leonard, Bryan and Brian Kogelmann (2022). **"Does Equality Persist? Evidence from the Homestead Act".** In: *Journal of Political Institutions and Political Economy* 3.2, pp. 215–241. ISSN: 2689-4823, 2689-4815. DOI: 10.1561/113.00000059. URL: http://www.nowpublishers.com/article/Details/PIP-0059 (visited on 02/02/2023).

MacKinnon, Mary and Chris Minns (2009). **"The Impact of School Provision on Pupil Attendance: Evidence From the Early 20th Century".** In.

Mattheis, Ross and Itzchak Tzachi Raz (2019). **"There's No Such Thing As Free Land: The Homestead Act and Economic Development".** In: p. 91.

Nagy, Dávidkrisztián (Jan. 31, 2023). **"Hinterlands, City Formation and Growth: Evidence from the U.S. Westward Expansion".** In: *The Review of Economic Studies*, rdad008. ISSN: 0034-6527, 1467-937X. DOI: 10.1093/restud/rdad008. URL: https://academic.oup.com/restud/advance-article/doi/10.1093/restud/rdad008/7017614 (visited on 03/20/2023).

Porzio, Tommaso, Federico Rossi, and Gabriella Santangelo (Aug. 2022). **"The Human Side of Structural Transformation".** In: *American Economic Review* 112.8, pp. 2774–2814. ISSN: 0002-8282. DOI: 10.1257/aer.20201157. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20201157 (visited on 11/20/2023).

Price, Joseph et al. (Apr. 1, 2021). **"Combining Family History and Machine Learning to Link Historical Records: The Census Tree Data Set".** In: *Explorations in Economic History* 80, p. 101391. ISSN: 0014-4983. DOI: 10.1016/j.eeh.2021.101391. URL: https://www.sciencedirect.com/science/article/pii/S0014498321000024 (visited on 08/31/2023).

Saskatchewan Department of Education (1906, 1907, 1908, 1910, 1911, 1912, 1914). **Annual Report of the Saskatchewan Department of Education.**

Schaede, Ursina (2021). **The Long Run Effects of Funding for Public Education: Evidence from Land Grants.** Ursina Schaede. URL: https://ursinaschaede.github.io/research/ (visited on 03/22/2023).

Smith, Cory (2022). **"Land Concentration and Long-run Development in the Frontier United States".** In.