Online Appendix to

“Who benefited from industrialization? The local effects of hydropower technology adoption in Norway”

by Stefan Leknes and Jørgen Modalsli

This version: August 25, 2019

A Data details

A.1 Hydropower production

The data on hydroelectric power plants is mainly taken from a detailed tabulation published by the Norwegian Water Resources and Energy Directorate (1946). The source does not distinguish between mechanical and electrical generators. In general, this would give a too early start year of hydropower adoption since mechanical generators were already in use when electrical hydropower generators were introduced. To ensure that the source is reliable we cross-check the information against other historical accounts. The following supplementary sources were used:

Aalholm, O. A. (1983). *Handelshuset Thommesen-Smith : T. Thommesen & Søn - Smith & Thommesen*. Arendal: Rygene-Smith & Thommesen

Aktieselskapet Tou (1905). *1855-1905 Aktieselskapet Tou: Tou Brug*. Stavanger: Aktieselskapet Tou

Eek, B. (1998). *Fabrikken ved Hellefossen: Borregaard Hellefoss 1818-1998*. Hokksund: Borregaard

Eek, B. “En kort historikk: Vestfos Cellulosefabrik”. Last modified 8th of October 1991. http://eiker.org/Artikler/be/be-1991-10-08-VestfosCellulosefabrik.html

Fageraas, K. B., B. Bækkelund, C. Nilsson, and E. Bagle (2006). *Masse papir: Norsk papir- og massefabrikker gjennom 150 år*. Elverum: Norsk skogmuseum

Gervin, E. (1973). *A/S Follum fabrikker: et hundre år: 1873-1973*. Oslo: Follum fabrikker

Gierløff, C. (1959). *Sævareid: En vestlandsk treforedlingsbedrift og kultursaga*. Bergen: Sævareid

Grieg, S. (1946). *AS Arne fabrikker: 1846-1946.* Bergen: Arne fabrikkers direksjon

Hauge, Y. (1957). *Ulefos jernværk: 1657-1957.* Oslo: Aschehoug

Hunsfos historielag. "Otterelvens Papirfabrik/Hunsfos Fabrikker 1886". Accessed 20th of October 2015. http://hist.hunsfos.no/historie/

Iversen, K. P. (1991). *100 år med lys og varme: Hammerfest elektrisitetsverk 1891-1991.* Hammerfest: Verket

Kaldal, I. (1994). "Arbeid og miljø ved Follafoss tresliperi og Ranheim papirfabrikk 1920-1970". PhD diss. Trondheim: Historisk institutt

Kittilsen, I. (1953). *Union co.: en norsk storbedrifts historie gjennem 80 år: 1873-1953.* Oslo: Universitetsforlaget

Kjosbakken, E. (1973). *Mesna: kraftkilde, industriåre, kunstnermotiv, vannkilde, fiskeelv: utgitt ved Mesna kraftselskaps 50 års jubileum 1973.* Lillehammer

Kvinlaug, S. (1998). *Trælandsfoss 100 år: 1898-1998.* Kvinesdal: Trælandsfos A/S

Lange, E. (1985). *Fra Linderud til Eidsvold Værk IV. Treforedlingens epoke 1895-1970.* Oslo: Dreyers forlag

Lorentzen, B. (1966). *Vaksdal Mølle 1866-1966.* Bergen: J. W. Eide

Lund, T. (1991). *Elkrafta i Modum: Modum elverk 80 år: 1913-1993.* Vikersund: Elverket

Myrvang, C. (2014). *Troskap og flid. Kongsberg våpenfabrikks historie 1814-1945.* Oslo: Pax

Møller, I. (2002). *Norske vannkraftverk*, Vol. 1. Lysaker: Energi Forlag

Møller, I. (2003). *Norske vannkraftverk*, Vol. 2. Lysaker: Energi Forlag

Omang, R. (1935). *Fritzøe i slekten Treschows eie: 1835-1935.* Oslo: Aschehoug

Schwartz, J. J. (1914). *Kongsberg Vaabenfabrik: 1814-1914.* Kristiania: Grøndahl

Solem, A. (1954). *Norske kraftverker, Teknisk ukeblad 100 års jubileum.* Oslo: Teknisk ukeblads forlag

Sælen, F. (1961). *Fossen og fabrikken: litt om sævereid og virksomheten der.* Bergen

Throndsen, L. (1968). *A.S. Solberg Spinderi 150 år.* Drammen: Solberg Spinderi

Vevstad, A. (1988). *AS Egelands verk: Tresliperi 1888-1988.* Søndeled: Egelands Verk

Fosselv power stations 1 and 2 are counted as one plant. The two power stations have the same owners and start-up year. The same applies to the upper and lower power stations at Hønefoss.

In the analysis in Table 3 of the main paper, the relationships between the instrument and changes between 1891 and 1900 to labor force size and sector employment shares, the municipality of Askim is excluded. This is one of the municipalities where year of construction is available (1900-1903). The information is available in the following publication:

Norges vassdrags- og elektrisitetsvesen (1922). *Utbygget vannkraft i Norge: En forelbig oversikt.* Kristiania: H. Aschehoug & co

A.2 Historical infrastructure data on municipalities

The data on infrastructure up until 1880 is taken from the collection "Norwegian Ecological Data, 1868-1903", compiled by Frank H. Aarebrot. This collection is available as data set 41 from the Inter-university Consortium for Political and Social Research (ICSPR), `https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/41`. The infrastructure data are in Part 4 of the collection and refer to 1880 (though using the 1868 municipality structure, which we convert to the 1900 structure used in our paper). The variables used, and their description in the data set (Aarebrot uses "commune" for municipality, and "diligence" for coach), are:

- Existence of a pier where steamships would stop on a regularly scheduled basis in the commune
- Existence of a railway station in the commune
- Communes having one stop on steamship route 6 (the main coastal steamship route)
- Existence of a diligence stop in the commune

We update this information to 1890 using information found in the following publications:

Aspenberg, N. C. (1994). *Glemte spor: Boken om sidebanenes tragiske liv.* Mesna: Mesna trykk

Bergh, T. (2004). *Jernbanen in Norge 1854-2004: Nye spor og nye muligheter 1854-1940.* Bergen: Vigemostad & Bjørke AS

Bjerke, T. and F. Holom (2004). *Banedata 2004: Data om infrastrukturen til jernbanene i Norge.* Trondheim: Skipnes AS

Gibberud, I. J. and H. Sunde (1992). *Flåmsbana: Historien om en av verdens bratteste jernbaner.* Bergen: John Grieg Forlag

Hartmann, E., Ø. Mangset and Ø. Reisegg (1997). *Neste stasjon: En guide til jernbanens arkitekturhistorie*. Oslo: Gyldendal norsk forlag ASA

A.3 Sector composition data for municipalities

The data on sector composition between 1891 and 1920 are taken from the Norwegian Center for Research Data, NSD (*Kommunedatabasen*, `http://www.nsd.uib.no/kdb`) and contains a transcription of municipality-level results published in the original census reports.

The data collection and reporting become more detailed with each census. For instance, the 1910 census differentiates between rural and urban municipalities, while the 1920 census also distinguishes between the sexes. Consequently, the categories for the oldest census in 1891 to a large extent determine the grouping of professions in each sector. The data are reported for individuals aged 15 years or older. We distinguish between three sectors: primary sector, manufacturing and services.

The data consist of variables where workers are allocated to subsectors on the basis of occupation. The categories that comprise the primary sector, manufacturing and services are given in Tables A.1, A.2 and A.3, respectively. The categories in the four censuses are alike with some minor exceptions, and therefore the baseline specification uses variation between all four censuses. The importance of the small discrepancies between censuses is evaluated through robustness tests. The categories are chosen to maximize comparability. In the baseline we aggregate to larger sectors to minimize the potential changes in categories and recording practice between censuses.

In the 1920 census, sector affiliation is based on belonging to a household. Moreover, while the 1891-1910 censuses base their tabulations on individuals present at the day of count (de facto population), the 1920 census tabulations are rather based on individuals' registered residency municipalities (de jure population). According to the documentary material from the 1920 census, the difference between the two definitions should be negligible. Nevertheless, to deal with this we exclude occupational groups whose geographical work location may create large discrepancies between the two count systems, for instance people like sailors who work in maritime sectors. In robustness tests we exclude the data from the 1920 census altogether with small changes to the overall conclusions (see Table D.6). Defining the manufacturing sector narrowly still yields significant results.

The primary sector consists of occupations in the following areas: farming and animal husbandry, horticulture, forestry and hunting, and fisheries. The manufacturing industry consists of factory industry, mining and quarrying, artisan industries and other smaller industries (works, construction and communications). Workers in smaller works and construction of communications are not included in a separable category in 1891 census. The census fixed effects should absorb this difference, as long as the geographical distribution of this category in 1891 is uncorrelated with hydropower production. The IV approach also assists as long as the locations of the omitted group are not correlated with the instrument. As a robustness test we let manufacturing industries consist of factory industry, mining and quarrying. This is a definition that may be more stable across censuses. The conclusions are robust to this alteration of definition (see Tables D.6 and D.5).

The service sector consists of commerce, trade, banking, the running of hotels and restaurants, and transportation. Profession work (civil administration, defense, the courts, teaching, health, art and literary work, and religious professions) are available from 1900. As can be seen in Table D.7, the profession work share is positively related to hydropower adoption, but not significant in the FE+IV specification. It is however significant when we drop 1920 also. There are also some issues concerning workers in the post and telephone sector in 1891. We therefore exclude the 1891 census also from the baseline specification in the same table without altering to the conclusions. Throwing out census years is a rather harsh robustness test as the categories missing in the variables probably represent small groups for which it is not likely that distribution is correlated with hydropower technology.

We also rerun the analyses without the census years 1891 and 1920, and without municipality fixed effects. The identification assumption must then be somewhat adjusted, stating that the instrument is also independent of municipality fixed effects. As can be seen from Table D.8, we obtain similar conclusions.

Table A.1: Primary sector variables

Census year	Rural/ urban	Category
1891		Farming and animal husbandry
		Horticulture
		Forestry and hunting
		Fisheries
		Log driving
1900		Sedentary agricultural sectors including forestry and hunting
		Fisheries
1910	Rural	Farming and animal husbandry: farmers, landowners
		Farming and animal husbandry: tenant farmers
		Farming and animal husbandry: children living at home, etc.
		Farming and animal husbandry: servants
		Farming and animal husbandry: other agricultural laborers
		Forestry and hunting: forest workers
		Farming and livestock breeding, forestry: others
		fisheries: independent fishers
		fisheries: others
	Urban	Farming, animal husbandry, forestry
		Fisheries: independent fishers
		Fisheries: others
1920	Rural	Farming, horticulture and forestry: farmers, landowners
		Farming, horticulture and forestry: tenant farmers
		Farming, horticulture and forestry: children living at home occupied by farming and livestock breeding
		Farming, horticulture and forestry: servants at farms
		Farming, horticulture and forestry: other independent laborers
		Farming, horticulture and forestry: clerks
		Farming, horticulture and forestry: forest workers, log drivers
		Farming, horticulture and forestry: other workers in farming and horticulture
		Fisheries
	Urban	Farming, horticulture and forestry
		Fisheries

Table A.2: Manufacturing sector variables

Census year	Rural/ urban	Category
1891		Manufacturing industry
		Artisan industries
		Mining industries
		Quarrying and harvest of ice and peat
1900		Manufacturing industry, mining and quarrying industry etc.
		Artisan industries
		Other industries
1910	Rural	Manufacturing industry, mining and quarrying industry
		Artisan industries
		Other smaller industries: works and communications
	Urban	Manufacturing industry, mining and quarrying industry
		Artisan industries
		Other smaller industries: works, communications and others
		Other smaller industries: textile
1920	Rural	Manufacturing industry
		Artisan industries
		Mining and quarry industry, peat harvest etc.
		Construction work
	Urban	Manufacturing industry: factory owners etc.
		Manufacturing industry: clerks etc.
		Manufacturing industry: laborers
		Construction workers

Table A.3: Service sector variables

Census year	Rural/ urban	Category
1891		Trade and banking
		Hotels and restaurants
		Transportation: trains and land-carriage
1900		Trade, banking and transportation (excluding sea transport)
1910	Rural	Trade, banking and transportation
		Trade: sales assistant
	Urban	Trade: merchants, wholesalers
		Trade: sales assistant
		Trade, banking and transportation: others
1920	Rural	Trade activity
		Transportation: carriers, chauffeurs etc. (excluding sea transport)
		Train, post and telegraph etc.
	Urban	Trade: Merchants, wholesalers
		Trade: clerks
		Trade: sales assistant, messengers
		Banking, insurance, brokers, etc.
		Hotels and cafes
		Transportation: carriers, chauffeurs etc. (excluding sea transport)
		Train, post and telegraph etc.

A.4 Individual-level data

The individual-level census records from 1865, 1900 and 1910 can be obtained from `http://www.nappdata.org`. More information on variable usage and linkage is given below in Appendix B.

A.5 Occupational classification

The occupational categories used in the baseline analysis are shown in Table A.4. Percentages refer to the share of the male population aged 20-50 in 1910.

In the section "Did upward occupational mobility cause a hollowing out of the skill distribution?", a more fine-grained classification of the manual occupations is used, based on the SEIUS classification (as implemented by NAPP). The cutoffs were chosen on the basis of the number of individuals in each occupation, to create categories as similar in size as possible.

The lowest-skilled category (SEI 9 or lower) predominantly contains occupations classified as *manual, unskilled* in the baseline specification. The next category, SEI 10-15, contains *manual, unskilled* occupations, but also some *manual, skilled* occupations. *Farmers* are also classified in this category. Occupations in the next two categories, SEI 16-20 and SEI 21-25, predominantly constitute *manual, skilled* occupations in the baseline analysis. The highest-skill category, SEI 26+, also has a substantial share of *manual, skilled* occupations. In addition, nearly all white-collar occupations are placed in this category.

By way of illustration, the largest manual occupation groups are shown with SEIUS rankings and categories in Table A.5.

Table A.4: Occupational classifications, and share of total population (men age 20-50, 1910)

Category	Share of population
White collar	
HISCO: 1100-3100, 3250-6400, 7110, 7600-13300, 14120-16300, 17120-22190, 23160, 31010-36020, 37020-45120, 45220-49030, 51020-51030, 51050-51090, 58500, 59200, 59950, 63220, 77630, 89500, 94920	
Largest categories:	
Dealer, merchant etc. (wholesale and retail trade)	2.2%
Salesmen, wholesale or retail trade	1.0%
Office clerks, specialization unknown	0.8%
Teachers (primary)	0.7%
Ship's navigating officers and ship's mates	0.7%
Other occupational categories	8.1%
Total:	*13.6%*
Manual skilled	
HISCO: 3210-3240, 6500, 7500, 16400, 23110-23150, 23170-24100, 36040-36090, 45190, 49090, 58100-58220, 58420-58430, 62800, 64970-77620, 77640-89200, 89400, 89620-94290, 94930-96900, 97130, 97150-97300, 97440, 98120-98440, 98510-98730, 99200, 99450	
Largest categories:	
Carpenters	3.1%
Seamen	2.3%
Boot and shoe makers and repairers	1.6%
Sawyers and other titled wood/sawmill operatives	1.6%
Paper mill machine operators and paper makers	1.4%
Other occupational categories	22.0%
Total:	*32.1%*
Manual unskilled	
HISCO: 7210, 13990, 51040, 52020-57040, 58300, 59100, 59940, 59990, 61115, 61330, 62110-62740, 62920-63140, 63230-64960, 89300, 97120, 97140, 97410-97430, 97490, 98490, 98900-99150, 99300-99440	
Largest categories:	
Farm workers, specialization unknown	6.7%
Fishermen	6.2%
Lumbermen, loggers and kindred workers	2.5%
Husbandmen or cottars	1.9%
Day laborers (e.g., journalier)	1.8%
Other occupational categories	8.0%
Total:	*27.2%*
Farmer	
HISCO: 61110, 61220-61320, 61400	
Largest categories:	
General farmers and farmers not further specified	18.4%
Farmer and fisherman	4.5%
Other occupation categories	0.4%
Total:	*23.2%*
Occupation missing	
Total:	*3.8%*

Table A.5: Occupational classifications, examples (manual occupations only)

Category (HISCO title)	SEI score	Share of pop.
Highest-skilled (SEI 26 or higher)		
Delivery men and drivers of goods	32	1.0%
Mason not further specifiedor combined	27	1.3%
Mechanics	27	1.6%
High-skilled (SEI 21-25)		
Stone carvers or cutters and stone yard workers	25	1.6%
Tailors and dressmakers	23	1.3%
Bakers	22	1.2%
Medium-skilled (SEI 16-20)		
Carpenters	19	5.3%
Boot and shoe makers and repairers	18	2.8%
Sawyers and other titled wood/sawmill operatives	18	2.6%
Papermill machine operators and paper makers	18	2.4%
Ship's engine men	17	1.7%
Painters, not further specified	16	1.4%
Blacksmiths	16	1.5%
Seamen	16	3.9%
Low-skilled (SEI 10-15)		
Drivers, nec	15	1.7%
Husbandrymen or cotters	14	3.2%
Cotters and fisherman	14	1.5%
Ship and boat loaders and dock workers	11	1.1%
Miners	10	1.6%
Fishermen	10	10.5%
Lowest-skilled (SEI 9 or lower)		
Laborers not further specified	8	1.5%
Other skilled railway workers	8	1.4%
Navvies, excavators and diggers, not further specified	8	0.8%
Day laborers (e.g., journalier)	8	3.1%
Road builders, workers and labourers	8	0.9%
Servants not further specified	7	1.3%
Farm workers, specialization unknown	6	11.4%
Lumbermen, loggers and kindred workers	4	4.2%
Porters	4	1.0%

A.6 Summary statistics

Table A.6: Summary statistics for municipality analyses

	Mean	Std. dev.
Labor force size	1828.6	1222.09
Employment share in manufacturing	9.20	5.99
Employment share in services	2.62	2.07
Employment share in primary sector	39.1	8.72
Number of hydropower plants	0.07	0.32
Indicator of coast	0.61	0.49
Area of land	654.25	913.2
Emigration share (lagged)	6.08	5.4

Table A.7: Summary statistics for municipality level analyses

	Hydropower municipalities				Non-hydropower municipalities			
	1891		1920		1891		1920	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Labor force size (pop. aged 15+)	2165.3	(1284.56)	3037.99	(2039.07)	1605.58	(900.86)	1857.8	(1332.76)
Employment share in manufacturing	10.95	(6.00)	16.15	(8.23)	7.48	(3.97)	8.83	(5.38)
Employment share in services	3.91	(5.34)	8.76	(7.07)	1.24	(2.62)	3.87	(3.63)
Employment share in primary sector	39.14	(7.97)	30.35	(9.87)	43.11	(7.85)	39.95	(9.05)

Municipalities are separated into two groups: municipalities with hydropower production sometime during 1891-1920 (hydropower municipalities) and municipalities without such production in the same period (non-hydropower municipalities).

Table A.8: Summary statistics for upward mobility analyses, linked worker sample

	Mean	Std. dev.	N
Indicator of upward mobility for farmers	0.05	0.22	33001
Indicator of upward mobility for unskilled manual workers	0.13	0.33	30923
Indicator of upward mobility for skilled manual workers	0.05	0.23	16268
Number of hydropower plants	0.09	0.34	86730
Age	34.22	8.92	86730
Age squared	1250.85	622.54	86730
Indicator of being married	0.62	0.48	86432
Number of children	1.91	2.28	86730
Indicator of not being resident in municipality of birth	0.22	0.41	86730
Indicator of coast	0.62	0.49	86730
Area of land	677.86	805.83	86730
Emigration share (lagged)	4.33	3.22	86730
Steamship stop	0.62	0.49	86730
Ship route stop	0.2	0.4	86730
Railwaystation before 1880	0.14	0.35	86730
Number of railwaystations constructed 1880-1890	0.13	0.68	86730
Coach stop	0.05	0.22	86730

Table A.9: Summary statistics for upward mobility analyses, linked father-son sample

	Mean	Std. dev.	N
Indicator of upward mobility for farmers	0.23	0.42	32864
Indicator of upward mobility for unskilled manual workers	0.27	0.44	10588
Indicator of upward mobility for skilled manual workers	0.08	0.28	5213
Number of hydropower plants	0.1	0.38	50999
Age, son 1900	16.79	5.4	50999
Age squared, son 1900	311.01	201.84	50999
Indicator of son being married	0.02	0.14	50834
Sons number of children	0.02	0.19	50999
Indicator of son being born in municipality of residence	0.09	0.28	50999
Indicator of coast	0.62	0.49	50999
Area of land	674.78	803.83	50999
Emigration share (lagged)	4.27	3.21	50999
Steamship stop	0.62	0.49	50999
Ship route stop	0.2	0.4	50999
Railwaystation before 1880	0.14	0.34	50999
Number of railwaystation constructed 1880-1890	0.08	0.49	50999
Coach stop	0.05	0.21	50999

A.7 Incomes by occupational category

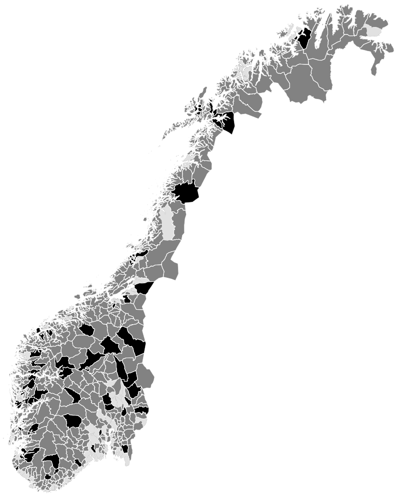
The income information in Footnote 14 of the main paper are calculated from data on income for men aged 30-60 in “Indtægts- og formuesforhold efter skatteligningen 1911 i forbindelse med Folketællingen 1910, Norges Officielle Statistik VI no. 24”, publ. 1915. A general review of this documentation is given in Modalsli (2017), page 14 as well as in Appendix A2 to that paper (figure of general trends).

Figure A.1: City municipalities and municipalities with hydropower production by census year

(a) City municipalities, excluded from sample

(b) Hydropower production in 1900

(c) Hydropower production in 1910



(d) Hydropower production in 1920

B Supplementary information on record linkage

The linked 1900-1910 sample (as well as the 1865-1900 sample used for "historical mobility") was constructed on the basis of an algorithm developed and used by Modalsli (2017). The following exposition is based on the information in that paper, as well as its online appendix.

B.1 Data

Data were obtained from individual-level data sets of the population as recorded in the Norwegian censuses of 1865, 1900 and 1910.

From the census files, the following variables were extracted:

- First name
- Last name
- Name of place of residence
- Information on family relationship of those who reside together
- Birth year
- Municipality of birth

Then, individuals are linked across censuses by personal information: name, birth time and birth place. Time-varying characteristics such as occupation, spouse or other family members are not used for linkage as these are likely to be correlated with the outcome of interest. As fixed surnames were not mandated by law in Norway until 1925, there was still some flexibility in how individuals reported their identity to the authorities during this period. Spelling was somewhat flexible, and individuals could go by inherited surnames, patronymics (the name of their father plus the suffix "-sen"), or surnames based on the farms they grew up on. Over time, patronymics and farm names became fixed as time-invariant surnames that were inherited from fathers to sons.

The census files were obtained from the North Atlantic Population Project (www.nappdata.org). Names were converted to lower case; Norwegian characters were converted to "a" in all censuses (because of a limitation on how the characters were stored in the NAPP database at the time of extraction); special characters were removed and some common substitutions of spelling variants were substituted (such as "ch" for "k"). Patronymics were constructed by adding "sen" to the father's first name; the patronymic for the first names "Ola" and "Ole" was changed to the most common variant "Olsen".

B.2 Matching algorithm: Calculating differences in identifying information

Matches are in principle constructed by comparing all possible pairs from two years; however, this is impractical in practice because of the large number of potential combinations. To improve running time and improve flexibility in formulating match rules, all distances between match elements (e.g. names) were pre-calculated. For each piece of identifying information (as listed above) and year, a file with all unique occurrences was constructed. Then, all occurrences in year A were compared to all occurrences in year B for all variables. The following paragraphs describe how match scores are assigned; this description is partially reproduced from Modalsli (2017, Online Appendix B).

Strings (names)

The Levenshtein distance between any two strings is calculated using a command included in the `strgroup` package for Stata (written by Julian Reif, University of Chicago). The Levenshtein algorithm counts the minimum number of letter removals, additions or swaps needed to go from one string to another. The distance between the strings is divided by the length of the shortest string to get the final score. Only matches with name scores smaller than 0.3 are considered.

Scores are denoted D_F (first names), D_{L-CC} (last names), D_{L-PC} (patronymic in first period, last name in second period), D_{L-LC} (location name in first period, last name in second period), D_{L-CP} and D_{L-CL}.

Birth years

The score is the absolute value of the birth year in the two sources, and is considered if the difference is five years or less. The score is denoted D_Y.

Municipality of birth

Municipalities are aggregated to avoid mismatches due to border changes and mergers.

The score is set to 0 if the municipality cluster matches; 1 if the cluster is different but the county matches; 2 if both periods have missing birth municipality, and 3 if one of the periods has a missing birth municipality. The score is denoted D_M.

Aggregating match scores

Given the above qualifications, all matches between the compared censuses are considered. First, the two lists are merged by potentially similar first names ($D_F < .3$), then the scores for other matches are added. The last name score is constructed as $D_L = \min(D_{L-CC}, D_{L-PC}, D_{L-LC}, D_{L-CP}, D_{L-CL})$. Matches that are not considered (birth times too different or $D_L > .3$) are removed from the data set.

These scores are then combined to create an aggregate score. To balance the impact of name changes with differences in other characteristics, name differences are multiplied by 8.

$$D = 8 \cdot D_F + 8 \cdot D_L + D_Y + D_M \tag{A1}$$

The score D states the distance (difference) between two observations — one observation from each time period. Clearly, we want to pick the pairs of observations with low differences. However, we also have to evaluate the degree of *uniqueness* of each pair. For each observation i from time t, rank the candidates from period $t-1$ in descending order by score. Each $t-1$ candidate j will now have a difference score $D_{i,j}$. The uniqueness parameter R_i is then the difference between the (i, j) combination score $D_{i,j}$ and the score of the next best option (i, j'), $D_{i,j'}$. A higher value of R_i means the match is clearly better than other candidate matches. A similar uniqueness score R_j can be calculated from the viewpoint of the $t-1$ data set.

For a candidate to be accepted, restrictions are placed on the difference score and the uniqueness of each pair of observations. As the matching procedure is computationally intensive, a limited set of combinations is considered. Two different approaches with respect to uniqueness are tried; one where the limit of R increases with D (that is, more uniqueness is required if the match score is relatively poorer) and one where the limit of R is the same regardless of the requirement for D. In both cases, the match procedure is run iteratively; after each round, all accepted matches are removed, and the metrics are re-calculated.

The first round consists of all perfect matches: those where name, birthplace and birth time match perfectly ($D_{i,j} = 0$) and there are no other potential candidates for a match (that is, no candidate pairs where the composite scores are below the consideration thresholds described above).

From the second round onward, the allowable difference is increased in increments of 0.5. The allowable non-uniqueness is set to 0.5 for the second round and then increased by 0.25 in each iteration. Thus, the second round has the requirement $D_{i,j} \leq 0.5, R_i \geq 0.5, R_j \geq 0.5$, the third round $D_{i,j} \leq 1.0, R_i \geq 0.75, R_j \geq 0.75$ and so on. Visual inspection of the results show that the number of potential erroneous matches starts to appear around the sixth or seventh iteration. For this reason, the match procedure is stopped after round 5, the final requirement being $D_{i,j} \leq 2.0, R_i \geq 1.25, R_j \geq 1.25$.

B.3 Evaluating the matching algorithm

The details of the matching algorithm do not affect the mobility estimates. As shown in the online Appendix to Modalsli (2017), the Altham statistic (a commonly used mobility metric based on the full matrix of father-son occupations) hardly differs across data sets constructed with different ranges of parameters. Changing the matching parameters (accepting more or less matches) only changes the baseline estimate of 24.1 for the 1865-1900 period within a narrow range (from 4% below to 1% above).

As is common in matched samples of this type, some selection in matching cannot be completely ruled out. As mentioned in the main text, the main selectivity problem is with respect to farmers — there are higher success rates in matching individuals from this occupation group. This is not a major concern in this paper, as farmers are removed from the analysis of mobility by means of the linked data. Second, individuals from larger municipalities are harder to match (less unique identifiers); the baseline analysis here only encompasses rural areas (where most municipalities are relatively small).

C Supplementary information on estimation

C.1 Standard errors

Throughout the paper we compute heteroscedasticity robust standard errors clustered on the municipality level. This is motivated from potential correlations between components in outcomes within clusters. In the municipality analyses the number of fixed effects are high, as we include fixed effects for each municipality and census year. In addition, the number of observations within clusters in the balanced panel is low ($t = 4$). This causes the cluster robust variance matrix to become highly singular and nonsymmetric when conducting IV estimations.

This challenge to inference is described in detail in Cameron and Miller (2015), pages 330-331, where also a solution is proposed. The solution entails computing the Arellano (1987) cluster robust variance matrix. Properties are also described in Wooldridge (2010), page 275. In practice, this is equivalent to estimate the within estimator standard errors. To obtain reliable inference, we compute such standard errors in the municipality IV-regressions using the Stata command xtivreg2.

D Supplementary analyses and sensitivity tests

D.1 Reduced form results

Table D.1: Reduced form results

Sample:	Municipality sample								Linked samples	
									Workers	Father-sons
Dependent variables:	Ln(labor force)		Percentage of workers in						Upward mobility for unskilled	
			Manufacturing		Services		Primary sector			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: unskilled manual workers from the linked worker sample										
Hydropower potential 1900	0.020	0.013	2.878***	1.732***	0.249**	-0.156	-2.216***	-0.954**		
	(0.029)	(0.019)	(0.544)	(0.652)	(0.112)	(0.097)	(0.491)	(0.370)		
Hydropower potential 1910	0.006	-0.002	2.109***	0.964**	0.123	-0.282*	-1.588**	-0.326	0.015***	0.023
	(0.031)	(0.019)	(0.673)	(0.377)	(0.095)	(0.154)	(0.688)	(0.405)	(0.006)	(0.017)
Hydropower potential 1920	0.009	0.001	2.276***	1.133**	0.021	-0.385	-2.245***	-0.981*		
	(0.039)	(0.034)	(0.657)	(0.441)	(0.134)	(0.235)	(0.761)	(0.560)		
County fixed effects	Y	N	Y	N	Y	N	Y	N	Y	Y
Municipality fixed effects	N	Y	N	Y	N	Y	N	Y	N	N
Adjusted R-Squared	0.32	0.96	0.29	0.74	0.36	0.67	0.37	0.76	0.03	0.06
N	1820	1820	1820	1820	1820	1820	1820	1820	30824	10542

Data from Norwegian censuses of 1891, 1900, 1910 and 1920. For the linked samples only the middle censuses are available. Columns (1)-(8) display the reduced form results of the municipality regressions. Columns (9)-(10) display the results of the linked samples of unskilled manual workers and fathers respectively. Dependent variables: potential labor force and sector sizes, and upward occupational mobility for workers and across generations. Variables of interest: hydropower potential per thousand (anchored to 1900 home municipality in linked samples). Estimator: OLS.

All specifications control for geographical size of municipality (km^2), indicators of coast, historical infrastructure variables and lagged emigration share. In columns (1)-(8) the regressions also control for year fixed effects. In columns (9)-(10) the regressions include 1900 worker (son) characteristics: age, age squared, indicator of being married, number of children, and indicator of not being resident in municipality of birth. Robust standard errors clustered on municipality are in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

D.2 Alternative regression specifications and robustness checks

Table D.2: Hydropower production and general population size

	ln(population size)		
	OLS	FE	FE + IV
	(1)	(2)	(3)
Hydropower production	0.40***	0.14***	-0.11
	(0.07)	(0.03)	(0.23)
Municipality FE	N	Y	Y
First-stage F-statistic	-	-	10.84
Adjusted R-squared	0.33	0.97	-
N	1820	1820	1820

Data: Norwegian censuses from 1891, 1900, 1910 and 1920.

Dependent variables: natural logarithm of population size. Regressions control for year fixed effects, county fixed effects, geographical size of municipality (km^2), indicators of coast, historical infrastructure and lagged emigration share. Instruments are hydropower potential interacted with decade indicators. Robust standard errors clustered on municipality are in parentheses. The formulation of the cluster-robust covariance matrix for the IV-estimate follows Arellano (1987). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table D.3: Hydropower production, labor force size and industry composition. Sample with urban municipalities included

	ln(Labor force size)			Percentage of workers in manufacturing		
	OLS	FE	FE + IV	OLS	FE	FE + IV
	(1)	(2)	(3)	(4)	(5)	(6)
Mean (std. dev.)	7.49	(0.77)		10.66	(7.14)	
Hydropower	0.44***	0.14***	0.01	7.25***	2.38***	6.04**
	(0.08)	(0.03)	(0.22)	(0.95)	(0.62)	(2.73)
Municipality FE	N	Y	Y	N	Y	Y
First-stage F-statistic	-	-	6.23	-	-	6.23
Adjusted R-squared	0.37	0.97	-	0.37	0.83	-
N	2140	2140	2140	2140	2140	2140
	Percentage of workers in services			Percentage of workers in primary sector		
	OLS	FE	FE + IV	OLS	FE	FE + IV
	(7)	(8)	(9)	(10)	(11)	(12)
Mean (std. dev.)	3.38	(3.25)		36.16	(11.68)	
Hydropower	1.24***	0.37	-4.42**	-8.63***	-3.19***	-3.17
	(0.37)	(0.23)	(1.87)	(1.20)	(0.69)	(3.27)
Municipality FE	N	Y	Y	N	Y	Y
First-stage F-statistic	-	-	6.23	-	-	6.23
Adjusted R-squared	0.37	0.85	-	0.44	0.88	-
N	2140	2140	2140	2140	2140	2140

Data: Norwegian censuses from 1891, 1900, 1910 and 1920.

Very small urban municipalities (below 8 km^2) are merged with their adjacent neighbors. Dependent variables: natural logarithm of the labor force size (inhabitants 15 years and older) in columns (1)-(3), percentage worker shares in manufacturing, services and primary sectors in columns (4)-(12). Data on sectoral affiliation are available for persons aged 15 and older. Regressions control for year fixed effects, county fixed effects, geographical size of municipality (km^2), indicator of coast, infrastructure and lagged emigration share. Instruments are hydropower potential interacted with decade indicators. Robust standard errors clustered on municipality are in parentheses. The formulation of the cluster-robust covariance matrix for the IV-estimates follows Arellano (1987). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table D.4: Hydropower effect on the level of sectoral employment. Dependent variables are standardized

	Number of workers in manufacturing			Number of workers in services			Number of workers in primary sector		
	OLS	FE	FE + IV	OLS	FE	FE + IV	OLS	FE	FE + IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mean (std. dev.)	195.29	(266.06)		56.86	(99.25)		673.34	(383.27)	
Hydropower plants	1.16***	0.53***	0.79*	0.74***	0.53**	-0.64	0.25*	0.02	-0.59***
	(0.19)	(0.18)	(0.44)	(0.18)	(0.21)	(0.64)	(0.13)	(0.04)	(0.22)
Municipality FE	N	Y	Y	N	Y	Y	N	Y	Y
First stage F-statistics	-	-	10.84	-	-	10.84	-	-	10.84
Adjusted R-squared	0.32	0.74	-	0.31	0.54	-	0.28	0.96	-
N	1820	1820	1820	1820	1820	1820	1820	1820	1820

Data: Norwegian censuses from 1891, 1900, 1910 and 1920.

Dependent variables: workers in manufacturing, services and primary sectors in columns, respectively. In the regressions variables are standardized to have a mean of zero and standard deviation of unity. Data on sectoral affiliation are only available for persons aged 15 and older. Regressions control for year fixed effects, county fixed effects, geographical size of municipality (km^2), historical infrastructure, indicator of coast and lagged emigration share. Instruments are hydropower potential interacted with decade indicators.

Robust standard errors clustered on municipality are in parentheses. The formulation of the cluster-robust covariance matrix for the IV-estimates follows Arellano (1987). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table D.5: Hydropower production and changes in the traditional manufacturing and mining

	Percentage of workers in manufacturing, narrowly defined			Workers in manufacturing, narrowly defined		
	OLS	FE	FE + IV	OLS	FE	FE + IV
	(1)	(2)	(3)	(4)	(5)	(6)
Mean (std. dev.)	2.67	(4.35)		66.13	(160.38)	
Hydropower	6.19***	1.89***	2.86*	208.93***	118.09***	140.99*
	(1.03)	(0.55)	(1.54)	(33.69)	(34.06)	(75.54)
Municipality FE	N	Y	Y	N	Y	Y
First-stage F-statistic	-	-	10.84	-	-	10.84
Adjusted R-squared	0.31	0.82	-	0.29	0.73	-
N	1820	1820	1820	1820	1820	1820

Data: Norwegian censuses from 1891, 1900, 1910 and 1920.

Dependent variables: percentage and level of worker shares in manufacturing, narrowly defined as traditional manufacturing and mining. Data on sectoral affiliation is available for persons aged 15 and older. Regressions control for year fixed effects, county fixed effects, geographical size of municipality (km^2), indicator of coast, infrastructure and lagged emigration share. Instruments are hydropower potential interacted with decade indicators.

Robust standard errors clustered on municipality are in parentheses. The formulation of the cluster-robust covariance matrix for the IV-estimates follows Arellano (1987). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table D.6: 1920 census excluded

	Manufacturing		Services	Primary
	Broadly defined	Narrowly defined		sector
	FE + IV	FE + IV	FE + IV	FE + IV
	(1)	(2)	(3)	(4)
Panel A: Percentage of workers				
Hydropower	6.37	3.57*	-2.73	-1.12
	(4.14)	(2.07)	(1.68)	(3.51)
First-stage F-statistic	9.84	9.84	9.84	9.84
N	1365	1365	1365	1365
Panel B: Number of workers				
Hydropower	220.34	130.32**	-35.23	-118.01
	(136.75)	(64.25)	(35.55)	(101.74)
First-stage F-statistic	9.84	9.84	9.84	9.84
N	1365	1365	1365	1365

Data: Norwegian censuses from 1891, 1900 and 1910.

Dependent variables: Panel A: percentage worker shares in manufacturing, services and primary sectors, while Panel B includes the level of the same variables. See Table D.5 for definition of the different manufacturing variables. Data on sectoral affiliation is available for persons aged 15 and older. Regressions control for year fixed effects, municipality fixed effects and lagged emigration share. Estimator: 2SLS. Instruments are hydropower potential interacted with decade indicators.

Robust standard errors clustered on municipality are in parentheses. The formulation of the cluster-robust covariance matrix for the IV-estimates follows Arellano (1987). Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table D.7: Hydropower production and services sector changes, 1900-1920

	Workers in services				Workers in profession work			
	OLS	FE	FE + IV	FE + IV Omit 1920	OLS	FE	FE + IV	FE + IV Omit 1920
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Percentage of dependent variables								
Mean (std. dev.)	2.95	(2.19)		-	1.38	(0.55)		-
Hydropower	1.04***	0.42	-2.25	-1.88	0.26***	0.15**	0.19	0.61*
	(0.26)	(0.30)	(1.81)	(1.71)	(0.06)	(0.06)	(0.29)	(0.35)
First-stage F-statistic	-	-	6.58	7.81	-	-	6.58	7.81
Adjusted R-squared	0.35	0.71	-	-	0.22	0.66	-	-
N	1365	1365	1365	910	1365	1365	1365	910
Panel A: Level of dependent variables								
Mean (std. dev.)	66.09	(111.63)		-	28.26	(35.82)		-
Hydropower	71.15***	47.81**	-65.95	-24.58	24.14***	13.29**	-9.19	0.96
	(16.81)	(22.13)	(70.87)	(27.17)	(5.23)	(5.23)	(16.83)	(8.24)
First-stage F-statistic	-	-	6.58	7.81	-	-	6.58	7.81
Adjusted R-squared	0.32	0.62	-	-	0.26	0.74	-	-
N	1365	1365	1365	910	1365	1365	1365	910

Data: Norwegian censuses from 1900, 1910 and 1920.

Dependent variables: workers in services in columns (1)-(4) and workers in profession work in columns (5)-(8). In Panel A the variables are defined as percentages and in Panel B as level. Data on sectoral affiliation are available for people aged 15 and older. Regressions control for year fixed effects, county fixed effects, geographical size of municipality (km^2), indicator of coast, infrastructure and lagged emigration share. Instruments are hydropower potential interacted with decade indicators.

Robust standard errors clustered on municipality are in parentheses. The formulation of the cluster-robust covariance matrix for the IV-estimates follows Arellano (1987). Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table D.8: IV results without municipality fixed effects and excluding the 1891 and 1920 censuses

	ln(labor force size)	Percentage of workers			Number of workers		
		Manu.	Service	Prim.	Manu.	Service	Prim.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Hydropower	0.09	26.05***	1.73	-19.14***	512.63**	18.65	-365.89*
	(0.31)	(8.60)	(1.16)	(6.57)	(229.20)	(34.43)	(198.19)
First-stage F-statistic	9.46	9.46	9.46	9.46	9.46	9.46	9.46
N	910	910	910	910	910	910	910

Data: Norwegian censuses from 1891, 1900, 1910 and 1920.

Dependent variables: natural logarithm of the labor force size (inhabitants 15 years and older) in column (1). Percentage worker shares in manufacturing, services and primary sectors in columns (2)-(4), and the level of the same variables in columns (5)-(7). Data on sectoral affiliation are available for persons aged 15 and older and present at the census count. Regressions control for year fixed effects, county fixed effects, geographical size of municipality (km^2), indicator of coast, infrastructure and lagged emigration share. Instruments are hydropower potential interacted with decade indicators.

Robust standard errors clustered on municipality are in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table D.9: IV results without municipality fixed effects

	ln(labor force size)	Percentage of workers			Number of workers		
		Manu.	Service	Prim.	Manu.	Service	Prim.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Hydropower	0.08	20.94***	0.70	-18.49***	505.81***	-7.24	-402.84**
	(0.31)	(3.87)	(0.90)	(4.59)	(125.25)	(47.12)	(182.26)
First-stage F-statistic	10.23	10.23	10.23	10.23	10.23	10.23	10.23
N	1820	1820	1820	1820	1820	1820	1820

Data: Norwegian censuses from 1891, 1900, 1910 and 1920.

Dependent variables: natural logarithm of the labor force size (inhabitants 15 years and older) in column (1). Percentage worker shares in manufacturing, services and primary sectors in columns (2)-(4), and the level of the same variables in columns (5)-(7). Data on sectoral affiliation are available for persons aged 15 and older and present at the census count. Regressions control for year fixed effects, county fixed effects, geographical size of municipality (km^2), indicator of coast, infrastructure and lagged emigration share. Instruments are hydropower potential interacted with decade indicators.

Robust standard errors clustered on municipality are in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table D.10: Relationship between hydropower production and upward mobility for different occupational groups

		Up from farmer		Up from unskilled		Up from skilled	
		OLS	IV	OLS	IV	OLS	IV
		(1)	(2)	(3)	(4)	(5)	(6)
Mean	Panel A	0.05		0.13		0.05	
(std. dev.)		(0.22)		(0.33)		(0.23)	
	Panel B	0.23		0.27		0.08	
		(0.42)		(0.44)		(0.28)	
Panel A: Linked worker sample							
Hydropower		0.01	-0.01	0.05***	0.14**	0.00	-0.02
		(0.01)	(0.03)	(0.02)	(0.06)	(0.01)	(0.02)
First stage F-value		-	14.73	-	17.05	-	34.70
Adjusted R-squared		0.02	0.02	0.04	0.03	0.01	0.01
N		32904	32904	30824	30824	16193	16193
Panel B: Sample of linked father-son pairs							
Hydropower		0.09***	0.10	0.11***	0.22	0.01	-0.04
		(0.03)	(0.09)	(0.04)	(0.17)	(0.01)	(0.03)
First stage F-value		-	12.23	-	12.21	-	10.88
Adjusted R-squared		0.04	0.04	0.06	0.06	0.01	0.01
N		32771	32771	10542	10542	5198	5198

Data from Norwegian censuses of 1900 and 1910. Panel A displays results from the linked worker sample, while Panel B shows results from the linked father-son sample.

Dependent variables: In columns (1)-(2) the dependent variable is an indicator of change in profession from farmer to skilled and white collar between 1900 and 1910. In columns (3)-(4) it is an indicator of change in profession from unskilled to skilled or white collar between 1900 and 1910, while in columns (5)-(6) it is an indicator of change in profession from skilled to white collar between 1900 and 1910. In the regressions we control for the following characteristics of workers (sons) in 1900: age, age squared, indicator of being married, number of children, and indicator of not being resident in municipality of birth. All regressions include indicators of coast, area of land, infrastructure variables, emigration share and county fixed effects.

Robust standard errors clustered on municipality are in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table D.11: Hydropower adoption and change in worker occupation shares, manual sample

	Lowest-skilled	Low-skilled	Medium-skilled	High-skilled	Highest-skilled
	(1)	(2)	(3)	(4)	(5)
Summary statistics 1900					
Mean	40.50	28.84	23.02	3.30	4.34
(std. dev.)	(21.95)	(26.92)	(15.33)	(3.55)	(4.55)
Summary statistics change					
Mean	-7.64	4.75	0.29	0.60	2.00
(std. dev.)	(11.77)	(10.38)	(9.24)	(3.37)	(4.97)
Hydropower	2.56	-5.67***	3.90**	-2.22*	1.42
	(1.83)	(1.77)	(1.79)	(1.16)	(1.24)
Adjusted R-squared	0.19	0.16	0.09	0.02	0.06
N	452	452	452	452	452

Data: The Norwegian censuses of 1900 and 1910 are used to create a linked sample of workers belonging to detailed occupational categories. Estimator: OLS.

Dependent variables: change in detailed occupation shares between 1900 and 1910, in percent. The five occupation classes are derived using the SEIUS measure. The measure ranks occupations using U.S. data on income and education from 1950. The classes have the following cutoffs: 9, 15, 20 and 25. The mean and standard deviation for 1900 are provided in the top panel. The variable of interest is hydropower status in 1910. Municipalities that received this status earlier are omitted. In the regressions we include an indicator of coast, area of land, share of emigrants in the decade preceding 1900, historical infrastructure variables and county fixed effects. Robust standard errors are in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table D.12: Hydropower adoption and the likelihood of upward mobility for manual workers belonging to different skill classes, OCSCORUS measure

	Lowest-skilled (1)	Low-skilled (2)	Medium-skilled (3)	High-skilled (4)
Panel A: sample of linked workers				
Hydropower production	-0.00	0.09***	0.10**	0.01
	(0.02)	(0.03)	(0.04)	(0.01)
Adjusted R-squared	0.03	0.04	0.07	0.02
N	11473	8545	11053	11680
Panel B: sample of linked fathers and sons				
Hydropower production	0.13*	0.16***	0.13*	-0.01
	(0.08)	(0.03)	(0.07)	(0.02)
Adjusted R-squared	0.09	0.20	0.14	0.03
N	881	7367	2315	3697

Data from Norwegian censuses of 1900 and 1910. Panel A displays results from the linked worker sample, while Panel B shows results from the linked father-son sample.

Dependent variables: indicators for upward mobility of manual workers belonging to four different skill classes. Five skill classes are derived using the OCSCORUS measure. The measure ranks occupations using U.S. data on income from 1950. The classes are based on the following cutoffs: 9, 15, 20 and 25.

In the regressions we control for age, age squared, indicator of being married, number of children, and an indicator of not being resident in municipality of birth. All regressions include an indicator of coast, area of land, emigrant share, historical infrastructure variables and county fixed effects. Robust standard errors clustered on municipality are in parentheses. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

D.3 Robustness of aggregate results using synthetic control methods

To test the robustness of the results using a different estimation approach, we proceed with a synthetic control method with multiple treatment municipalities (Cavallo et al., 2013).[1] We focus on the municipalities that first adopted hydropower technology, just before 1900. Unfortunately, we have a rather limited time series for each municipality. We add data from the 1865 census to obtain a longer pretrend. The following categories from 1865 are included to expand the trend:

Table D.13: Sector variables from the 1865 census

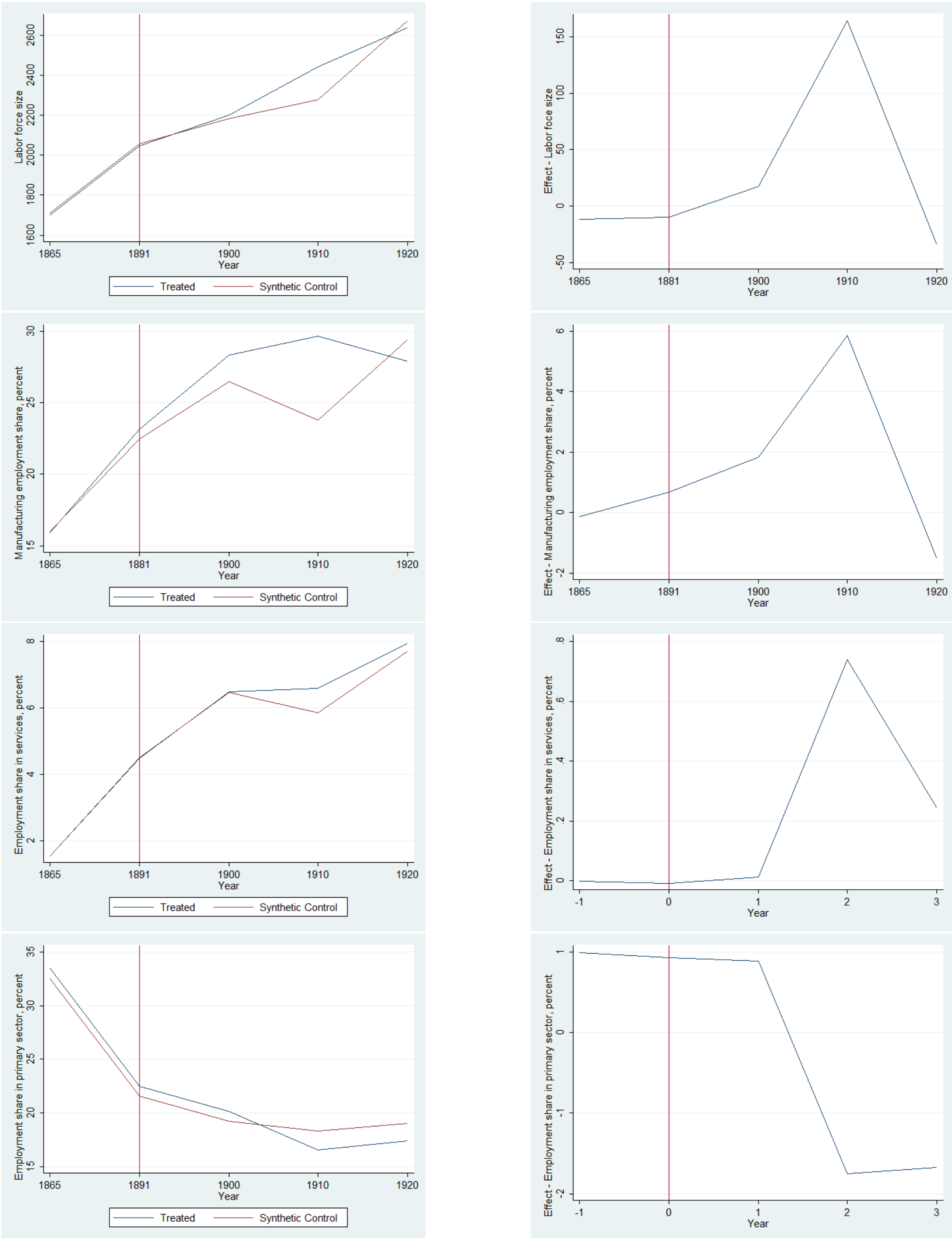
Sector	Rural/ urban	Category
Primary	Rural	Farming and animal husbandry, forestry, fisheries: main persons
		Farming and animal husbandry, forestry, fisheries: servants
	Urban	Farming and animal husbandry, forestry, fisheries: main persons
		Farming and animal husbandry, forestry, fisheries: servants
Manufacturing	Rural	Mining and manufacturing industry: main persons
	Urban	Mining and manufacturing industry: main persons
Services	Rural	Trade: Main persons
		Transport (excluding sea transport), post and telegraph: main persons
	Urban	Trade: merchants, shipowners: main persons
		Trade: sales assistants: main persons
		Trade: workers: main persons
		Trade: liquor and ale merchants, peddlers: main persons
		Trade: sales assistants and workers selling liquor and ale: main persons
		Transport (excluding sea transport), post and telegraph: main persons

The new data enable us to match on the level of the dependent variable in two periods, 1865 and 1891. We exclude municipalities that receive treatment in 1910 and 1920, and effectively match hydropower municipalities with municipalities that do not adopt hydropower technology in this period. The matching procedure is as follows. First, the program focuses on the pretrend of the treated municipalities. It matches the dependent variable by weighing selected non-treated municipalities to replicate the exact levels. The same weight matrix is used to create a counterfactual trend post treatment. The identification assumption is that matching on the level of the observables will also reflect the data-generating process that stems from the unobservables. In this case, because of the limited scope of the data, the method must be regarded as suggestive rather than conclusive.

The results are displayed in Figure D.2. On the left hand side, we have the average trends for the 3 treated municipalities and their controls; to the right, we have the average effects. From the top two figures, which display the result for labor force size, we see that the effect seems to last for two periods before it abates. The same can said for the second and third

[1] We use the synth_runner package for Stata.

Figure D.2: Effect of hydropower technology adoption on labor force size and structural transformation with synthetic control method



row of graphs showing the results for employment shares in manufacturing and services, respectively. However, the effect is stronger in the first period for manufacturing and it also lingers in the third period for services. The primary sector result, in the last row of graphs, shows a small decline in this sector. However, the pretrend is poorly matched. Summing up, the results are quite similar to what we find with other estimation methods. Nonetheless, we are not fulfilling the data requirements for the use of this method, and the results must be interpreted accordingly.

References

Arellano, Manuel (1987). Computing robust standard errors for within-groups estimators. *Oxford Bulletin of Economics and Statistics 49*, 431–434.

Cameron, A. Colin and Douglas L. Miller (2015). A practitioner's guide to cluster-robust inference. *The Journal of Human Resources 50*(2), 317–372.

Cavallo, Eduardo, Sebastian Galiani, Ilan Noy, and Juan Pantano (2013). Catastrophic natural disasters and economic growth. *Review of Economics and Statistics 95*(5), 1549–1561.

Modalsli, Jørgen (2017). Intergenerational mobility in Norway, 1865-2011. *Scandinavian Journal of Economics 119*(1), 34–71.

Norwegian Water Resources and Energy Directorate (1946). *Utbygd vannkraft i Norge.* Oslo: Norges vassdrags- og elektrisitetsforening, Den hydrografiske avdeling.

Wooldridge, Jeffrey M. (2010). *Econometric analysis of cross section and panel data.* Cambridge: MIT Press.