



**HAL**  
open science

# TSE M2 ETE Development: Theory, Public Policy, and Historical Perspectives (2022-2023) Topics in Economic History

Victor Gay

► **To cite this version:**

Victor Gay. TSE M2 ETE Development: Theory, Public Policy, and Historical Perspectives (2022-2023) Topics in Economic History. Doctoral. Development: Theory, Public Policy, and Historical Perspectives, France. 2023, pp.606. hal-04063637

**HAL Id: hal-04063637**

**<https://hal.science/hal-04063637v1>**

Submitted on 9 Apr 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Development: Theory, Public Policy, and Historical Perspectives Topics in Economic History

Victor Gay

University of Toulouse 1 Capitole

Toulouse School of Economics

Institute for Advanced Study in Toulouse

M2 PPD, Semester 2, 2023



Toulouse  
School of  
Economics

# Topic 1

## Economics and the modern economic historian

# Plan of Session

- Does the past have useful economics?
  - A classical approach. [McCloskey \(1976\)](#)
  - From classical to modern approach. [Arrow \(1985\)](#) [Solow \(1985\)](#)
  - A modern approach. [Abramitzky \(2015\)](#) [Cantoni Yuchtman \(2021\)](#)
- Trends in the field.
  - Trends in publishing. [Abramitzky \(2015\)](#) [Cioni et al. \(2019\)](#)
  - Historical econometrics. [Valencia Caicedo \(2021\)](#)
  - Future trends. [Lamoreaux \(2015\)](#)

# A Classical Approach

McCloskey (1976)

- Economic history for itself.
- More and better economic facts.
- Better economic theory.
- Better economic policy.

*“Disinterested intellectual curiosity is the lifeblood of real civilization. [...] There is nothing that more divides civilized from semi-savage man than to be conscious of our forefathers as they really were, and bit by bit to reconstruct the mosaic of the long-forgotten past. To weigh the stars, or to make ships sail in the air or below the sea, is not a more astonishing and ennobling performance on the part of the human race in these latter days, than to know the course of events that had long been forgotten, and the true nature of men and women who were here before us.”*  
Trevelyan (1942)

# Economic History for Itself

*“Disinterested intellectual curiosity is the lifeblood of real civilization. [...] There is nothing that more divides civilized from semi-savage man than to be conscious of our forefathers as they really were, and bit by bit to reconstruct the mosaic of the long-forgotten past. To weigh the stars, or to make ships sail in the air or below the sea, is not a more astonishing and ennobling performance on the part of the human race in these latter days, than to know the course of events that had long been forgotten, and the true nature of men and women who were here before us.”*  
Trevelyan (1942)

*“Every economy and every society has an interesting story to tell.”*  
Mokyr (2003)

⇒ Economic history much alike mathematical economics.

# More and Better Economic Facts

- Inductive or deductive approach: facts are necessary.



# More and Better Economic Facts

- Inductive or deductive approach: facts are necessary.
- History simply provides more facts.
- Whenever institutions are present, statistics are present.
- Even before: archaeology provides facts about the economy.

# More and Better Economic Facts

- More facts about economic growth:
  - Penn World Tables: real GDP PPP 182 countries since 1950.
  - Maddison project database: real GDP since 1000.

# More and Better Economic Facts

- More facts about economic growth:
  - Penn World Tables: real GDP PPP 182 countries since 1950.
  - Maddison project database: real GDP since 1000.
- More facts about populations:
  - IPUMS USA: full count censuses 1850–1950.
  - IPUMS NAPP: full count censuses North Atlantic starting 1703.

# More and Better Economic Facts

- More facts about economic growth:
  - Penn World Tables: real GDP PPP 182 countries since 1950.
  - Maddison project database: real GDP since 1000.
- More facts about populations:
  - IPUMS USA: full count censuses 1850–1950.
  - IPUMS NAPP: full count censuses North Atlantic starting 1703.
- Other facts:
  - Macrohistory: macro-finance indicators, 17 countries since 1870.
  - World Inequality Database: inequality indices since 1875.

# Aside: IPUMS

Ruggles et al. (2015)

- Greatest accomplishment for empirical economics.
- IPUMS: 98 countries, 443 censuses, > 1Bn individuals.
- Historical samples:
  - USA: full count censuses 1850–1950.
  - UK: 1851–1911.
  - Canada: 1852–1911.
  - Norway: 1801–1910.
  - Sweden: 1880–1910.
  - Iceland: 1703–1910.
- Data management:
  - Data preservation and reformatting.
  - Metadata integration: harmonized documentation.
  - Variable integration, e.g., marital status.
  - Dissemination: free, accessible, formats.

# More and Better Economic Facts

- Historical data are routinely richer than modern data.

# More and Better Economic Facts

- Historical data are routinely richer than modern data.
- Full access to companies' records and manufacturing censuses.
- Populations: no anonymity rule.
  - USA: linked samples across censuses (IPUMS).
  - France: linked samples across vital records (TRA).

# More and Better Economic Facts

- Historical data are routinely richer than modern data.
- Full access to companies' records and manufacturing censuses.
- Populations: no anonymity rule.
  - USA: linked samples across censuses (IPUMS).
  - France: linked samples across vital records (TRA).
- Access dramatically improved recently.



# More and Better Economic Facts

- Main limits to more and better facts:
  - Existence of historical data.
  - Inaccuracies and measurement errors.

# More and Better Economic Facts

- Main limits to more and better facts:
  - Existence of historical data.
  - Inaccuracies and measurement errors.
- On existence.
  - True.
  - But be creative and knowledgeable about the data production process.

# More and Better Economic Facts

- Main limits to more and better facts:
  - Existence of historical data.
  - Inaccuracies and measurement errors.
- On existence.
  - True.
  - But be creative and knowledgeable about the data production process.
- On inaccuracies.
  - Less true.
  - Much more cautious approach to data as constructed objects.

*“Much of the problem, I think, arises because of the separation in economics between data producers and data analyzers. By and large, we do not produce our own data and, hence, do not feel responsible for it.” Griliches (1974)*

# More and Better Economic Facts

- History provides better facts through natural experiments.
  - Can we learn from these experiments?
    - People were the same hundred years ago.
    - Structures were different and changed: we can learn a lot from that.
- ⇒ More on that later.

# Better Economic Theory

- Obvious that economic history uses economic theory in Popperian sense (as all applied economics).
- The converse is less obvious but no less true.

# Better Economic Theory

- Obvious that economic history uses economic theory in Popperian sense (as all applied economics).
- The converse is less obvious but no less true.

*In analogous discussions concerning the role of theory in historical research the argument is frequently made (perhaps because it is valid) that the historian will inevitably be guided by some [theory]. [...] Much the same can be said for the use of history by theorists. Even the most scornful ahistorical economist makes some use of history: his own experience, the experience of his generation, or the loose historical generalizations which abound in the folklore of even highly sophisticated societies.*

Cameron (1965)

- Theorists routinely use history (although some get facts wrong):
  - Kaldor stylized facts and neoclassical growth theory.
  - Ricardo and rising land rents.
  - Rostow and the stages of growth.
  - Marx and the impoverishment of the working class.

# Better Economic Theory

- Economic history to test economic theory.



# Better Economic Theory

- Economic history to test economic theory.
- More helpful than contemporaneous economy.
- Enables to test theories in their pure form, as structures sometimes less complex in the past.
- E.g., Koudijs (2015) on information and stock prices volatility.

*“Have a very healthy respect for the study of economic history, because that’s the raw material out of which any of your conjectures or testings will come.” Samuelson (2009)*

*One purpose of history is to broaden our conception of the possible.*

Rockoff (1975)

- Learn from historical crises to inform policy.
- Bernanke is among the best examples.

# Better Economists?

- Smith, Marshall, Schumpeter, Keynes, Friedman. . .
- The great economists were all deeply historical in their thinking.
- Even if not object of study, the long view helps finding important research questions.

# Better Economists?

- Smith, Marshall, Schumpeter, Keynes, Friedman. . .
- The great economists were all deeply historical in their thinking.
- Even if not object of study, the long view helps finding important research questions.
- So read, read, read!

# Does the past have useful economics?

- A classical approach. [McCloskey \(1976\)](#)
- From classical to modern approach. [Arrow \(1985\)](#) [Solow \(1985\)](#)
- A modern approach. [Abramitzky \(2015\)](#) [Cantoni Yuchtman \(2021\)](#)

# From the Classical to the Modern Approach

Arrow (1985) Solow (1985)

- The (caricatured) classical approach might be a naive view of history.
- History is not just a repository of facts for testing economic theory.

*The aim of historical study as such was not simply to serve as a source of data from which to infer and to test social science generalizations.*

# From the Classical to the Modern Approach

Arrow (1985) Solow (1985)

- The (caricatured) classical approach might be a naive view of history.
- History is not just a repository of facts for testing economic theory.

*The aim of historical study as such was not simply to serve as a source of data from which to infer and to test social science generalizations.*

- Economic history would just be applied economics with old data.  
*In this dispensation an economic historian is merely an economist with a high tolerance for dust or a working knowledge of a foreign language.*

# From the Classical to the Modern Approach

- Is economics like physics, true for all time, or are its laws historically conditioned?



# From the Classical to the Modern Approach

- Is economics like physics, true for all time, or are its laws historically conditioned?
- By its general statements, economic theory generates new questions for economic history.

# From the Classical to the Modern Approach

- Is economics like physics, true for all time, or are its laws historically conditioned?
- By its general statements, economic theory generates new questions for economic history.
- But economic theory emphasizes generality at the expense of particularity.
- Economic history helps the development of economic analysis by defining its historical conditioning.

⇒ Economic history is crucial to assess “external validity.”

# Does the past have useful economics?

- A classical approach. [McCloskey \(1976\)](#)
- From classical to modern approach. [Arrow \(1985\)](#) [Solow \(1985\)](#)
- A modern approach. [Abramitzky \(2015\)](#) [Cantoni Yuchtman \(2021\)](#)

# A Modern Approach

Abramitzky (2015)

- Test economic theory.
- Improve economic policy.
- Learn about mechanisms.
- Answer the big questions.

# Better Economic Theory

- Economic history to test economic theory.

# Better Economic Theory

- Economic history to test economic theory.
- Expands possibilities of exogenous variation for identification:

# Better Economic Theory

- Economic history to test economic theory.
- Expands possibilities of exogenous variation for identification:
  - Role of market access for economic development: Redding and Sturm (2008) use the division and reunification of Germany.

# Better Economic Theory

- Economic history to test economic theory.
- Expands possibilities of exogenous variation for identification:
  - Role of market access for economic development: Redding and Sturm (2008) use the division and reunification of Germany.
  - Role of news for stock price fluctuations: Koudijs (2015) uses delay of sailing boats due to adverse weather in 18th century.



# Better Economic Theory

- Economic history to test economic theory.
- Expands possibilities of exogenous variation for identification:
  - Role of market access for economic development: Redding and Sturm (2008) use the division and reunification of Germany.
  - Role of news for stock price fluctuations: Koudijs (2015) uses delay of sailing boats due to adverse weather in 18th century.
  - Impact of supply shocks on technological change: Hanlon (2015) uses drop in cotton from US South after Civil War.

- Economic history to test economic theory.
- Expands possibilities of exogenous variation for identification:
  - Role of market access for economic development: Redding and Sturm (2008) use the division and reunification of Germany.
  - Role of news for stock price fluctuations: Koudijs (2015) uses delay of sailing boats due to adverse weather in 18th century.
  - Impact of supply shocks on technological change: Hanlon (2015) uses drop in cotton from US South after Civil War.
  - Effect of ↓ agricultural labor availability on agricultural development: Hornbeck and Naidu (2014) use Great Mississippi Flood of 1927.

- Economic history to improve economic policy.

# Better Economic Policy

- Economic history to improve economic policy.
- Historical settings allow counterfactual worlds, or repetition of episodes.
- Again, past financial crises deeply inform current policy.

- Economic history to learn about mechanisms.
  - Typically, economists learn about mechanisms indirectly, e.g. by interactions or heterogeneity across subgroups.
  - Deep knowledge of historical setting and direct (qualitative) historical evidence can be much more convincing.
  - Analytic narratives approach using game theory (see Avner Greif's work).

# Answer Big Questions

- Economic history to answer the big questions.

# Answer Big Questions

- Economic history to answer the big questions.
- Only by looking at history can they be answered:

# Answer Big Questions

- Economic history to answer the big questions.
- Only by looking at history can they be answered:
  - Why are some countries rich and other poor?
  - Why did the Industrial Revolution happen in Europe and not elsewhere?
  - What factors caused the demographic transition?
  - How and why did inequality evolve over time?
  - What is the process of cultural change?



# Answer Big Questions

- Economic history to answer the big questions.
- Only by looking at history can they be answered:
  - Why are some countries rich and other poor?
  - Why did the Industrial Revolution happen in Europe and not elsewhere?
  - What factors caused the demographic transition?
  - How and why did inequality evolve over time?
  - What is the process of cultural change?

- These are too big to nail. But...

*Economic history seems to collectively agree to not limit ourselves to answer only those questions with random variation, of which there are not as many.*

# Natural Experiments of History

## Cantoni Yuchtman (2021)

- Economic history is especially popular in mainstream economics for natural experiments.
- Enable to answer important questions that require historical perspective while keeping applied micro approach and credible causal identification.
- Especially popular in economic growth and development.

⇒ Bridge between economic history and all other disciplines.

# Natural Experiments of History

- Taxonomy of three primary motives for examining historical natural experiments:
  - Natural experiments to learn about historical episodes (cat. 1).
  - Natural experiments yielding general economic insights (cat. 2).
  - Natural experiments to understand contemporary development (cat. 3).

# Natural Experiments of History

- Taxonomy of three primary motives for examining historical natural experiments:
  - Natural experiments to learn about historical episodes (cat. 1).
  - Natural experiments yielding general economic insights (cat. 2).
  - Natural experiments to understand contemporary development (cat. 3).
- Not mutually exclusive: most research overlaps categories.
- No one motive is better or worse.

# Natural Experiments of History

- This focus is the result of changes across several fields.

# Natural Experiments of History

- This focus is the result of changes across several fields.
- Labor economics and the credibility revolution.

# Natural Experiments of History

- This focus is the result of changes across several fields.
- Labor economics and the credibility revolution.
  - Push toward exogenous sources of variation for identification of causal effects. Angrist Pischke (2010)
  - Historical variation in public policies, historical human-made or natural events provide better sources of variation for identification (category 2).

# Natural Experiments of History

- This focus is the result of changes across several fields.
- Labor economics and the credibility revolution.
  - Push toward exogenous sources of variation for identification of causal effects. Angrist Pischke (2010)
  - Historical variation in public policies, historical human-made or natural events provide better sources of variation for identification (category 2).
- Economic history beyond the cliometric revolution (category 1).  
⇒ More on the cliometric revolution later



# Natural Experiments of History

- This focus is the result of changes across several fields.
- Labor economics and the credibility revolution.
  - Push toward exogenous sources of variation for identification of causal effects. Angrist Pischke (2010)
  - Historical variation in public policies, historical human-made or natural events provide better sources of variation for identification (category 2).
- Economic history beyond the cliometric revolution (category 1).  
⇒ More on the cliometric revolution later
- Development and the randomization revolution (category 3).

# Natural Experiments of History

- This focus is the result of changes across several fields.
- Labor economics and the credibility revolution.
  - Push toward exogenous sources of variation for identification of causal effects. Angrist Pischke (2010)
  - Historical variation in public policies, historical human-made or natural events provide better sources of variation for identification (category 2).
- Economic history beyond the cliometric revolution (category 1).  
⇒ More on the cliometric revolution later
- Development and the randomization revolution (category 3).

⇒ At the confluence: AJR's *Colonial Origins*

# Learn About Historical Episodes

- Seek historical experiments within pre-specified setting.
- Aim at more compelling causal arguments about process of interest, not testing general economic hypothesis.

# Learn About Historical Episodes

- Seek historical experiments within pre-specified setting.
- Aim at more compelling causal arguments about process of interest, not testing general economic hypothesis.
- Examples:
  - Impact of the invention of the printing press on city growth: quasi-random diffusion of the press from Mainz (Dittmar 2011).

# Learn About Historical Episodes

- Seek historical experiments within pre-specified setting.
- Aim at more compelling causal arguments about process of interest, not testing general economic hypothesis.
- Examples:
  - Impact of the invention of the printing press on city growth: quasi-random diffusion of the press from Mainz (Dittmar 2011).
  - Impact of the Great Migration of Black Americans northward from US South on agricultural technology: quasi-random variation in intensity of Great Mississippi flood of 1927 (Hornbeck and Naidu 2014).

# Learn About Historical Episodes

- Seek historical experiments within pre-specified setting.
- Aim at more compelling causal arguments about process of interest, not testing general economic hypothesis.
- Examples:
  - Impact of the invention of the printing press on city growth: quasi-random diffusion of the press from Mainz (Dittmar 2011).
  - Impact of the Great Migration of Black Americans northward from US South on agricultural technology: quasi-random variation in intensity of Great Mississippi flood of 1927 (Hornbeck and Naidu 2014).
  - Impact of WWI of female labor force participation: variations from cohort-drafting rules (Boehnke and Gay 2020).

# Learn About Historical Episodes: Challenges

- Critique of “looking under the lamppost.”
  - Only study component of episode exhibiting useful identifying variation.
  - Miss many important historical aspects.

# Learn About Historical Episodes: Challenges

- Critique of “looking under the lamppost.”
  - Only study component of episode exhibiting useful identifying variation.
  - Miss many important historical aspects.
- Sometimes, amounts to argue for inexplicability of historical variation of interest.
- But understanding source of variation is valuable (even beyond LATE).



# Yield General Economic Insight

- Seek historical source of variation to estimate a parameter.
- The specific setting is not of interest per se.

# Yield General Economic Insight

- Seek historical source of variation to estimate a parameter.
- The specific setting is not of interest per se.
- Examples:
  - Causal effect of military service on earnings: use Vietnam War draft lottery as source of variation in draft probability (Angrist 1990).

# Yield General Economic Insight

- Seek historical source of variation to estimate a parameter.
- The specific setting is not of interest per se.
- Examples:
  - Causal effect of military service on earnings: use Vietnam War draft lottery as source of variation in draft probability (Angrist 1990).
  - Causal effect of arrival of new information on stock prices: use variation in wind delaying boats between England and Amsterdam as a source of variation in arrival of information (Koudijs 2016).

# Yield General Economic Insight

- Seek historical source of variation to estimate a parameter.
- The specific setting is not of interest per se.
- Examples:
  - Causal effect of military service on earnings: use Vietnam War draft lottery as source of variation in draft probability (Angrist 1990).
  - Causal effect of arrival of new information on stock prices: use variation in wind delaying boats between England and Amsterdam as a source of variation in arrival of information (Koudijs 2016).
  - Causal effect of input relative abundance on technical change: use decline in US South cotton imports to Britain during US Civil War as a source of variation in input availability (Hanlon 2015).

# Yield General Economic Insight: Challenges

- Usually nice “existence” results.
- Sometimes lack of insight on mechanisms.
- Go beyond estimating single parameter.
- Critiques of reduced form à la Heckman (2005) and Deaton (2010).

# Understand Contemporary Development

- Seek historical experiments to identify the causal role of past events (or initial conditions) on the present.
- This is sometimes called the “persistence” literature.

# Understand Contemporary Development

- Seek historical experiments to identify the causal role of past events (or initial conditions) on the present.
- This is sometimes called the “persistence” literature.
- Examples:
  - Long-run effects of colonial institutions: AJR (2001, 2002) use settler’s mortality as IV for quality of historical institutions. See also Dell (2010), Lowes and Montero (2018), Dell and Olken (2019).

# Understand Contemporary Development

- Seek historical experiments to identify the causal role of past events (or initial conditions) on the present.
- This is sometimes called the “persistence” literature.
- Examples:
  - Long-run effects of colonial institutions: AJR (2001, 2002) use settler’s mortality as IV for quality of historical institutions. See also Dell (2010), Lowes and Montero (2018), Dell and Olken (2019).
  - Long-run effects of slave trade: Nunn (2008) uses distance to major slave-receiving ports as IV for intensity of slavery.



# Understand Contemporary Development

- Seek historical experiments to identify the causal role of past events (or initial conditions) on the present.
- This is sometimes called the “persistence” literature.
- Examples:
  - Long-run effects of colonial institutions: AJR (2001, 2002) use settler’s mortality as IV for quality of historical institutions. See also Dell (2010), Lowes and Montero (2018), Dell and Olken (2019).
  - Long-run effects of slave trade: Nunn (2008) uses distance to major slave-receiving ports as IV for intensity of slavery.
  - Long-run effects of disease environment: Alsan (2014) uses suitability for Tze Tze fly as IV for disease environment.

# Understand Contemporary Development: Challenges

- Causally linking past to present is subject to historical contingency:
  - Especially salient when little data around the experiment: challenges to the exclusion restriction.
  - A lot of work is needed along the causal chain.

# Understand Contemporary Development: Challenges

- Causally linking past to present is subject to historical contingency:
  - Especially salient when little data around the experiment: challenges to the exclusion restriction.
  - A lot of work is needed along the causal chain.
- Spatial correlation when persistence is underestimated (Kelly 2019).

# Understand Contemporary Development: Challenges

- Causally linking past to present is subject to historical contingency:
  - Especially salient when little data around the experiment: challenges to the exclusion restriction.
  - A lot of work is needed along the causal chain.
- Spatial correlation when persistence is underestimated (Kelly 2019).
- Reproducibility is challenging, as well as the risk of multiple hypothesis testing.

# Understand Contemporary Development: Challenges

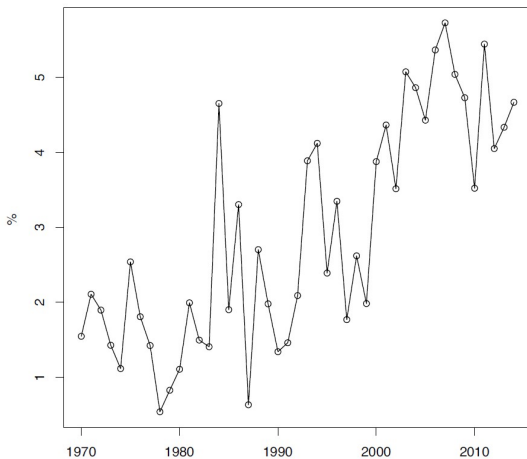
- Causally linking past to present is subject to historical contingency:
  - Especially salient when little data around the experiment: challenges to the exclusion restriction.
  - A lot of work is needed along the causal chain.
- Spatial correlation when persistence is underestimated (Kelly 2019).
- Reproducibility is challenging, as well as the risk of multiple hypothesis testing.
- Bias toward persistence. When and why does history not persist?

# Plan of Session

- Does the past have useful economics?
  - A classical approach. [McCloskey \(1976\)](#)
  - From classical to modern approach. [Arrow \(1985\)](#) [Solow \(1985\)](#)
  - A modern approach. [Abramitzky \(2015\)](#) [Cantoni Yuchtman \(2020\)](#)
- Trends in the field.
  - Trends in publishing. [Abramitzky \(2015\)](#) [Cioni et al. \(2019\)](#)
  - Historical econometrics. [Valencia Caicedo \(2021\)](#)
  - Future trends. [Lamoreaux \(2015\)](#)

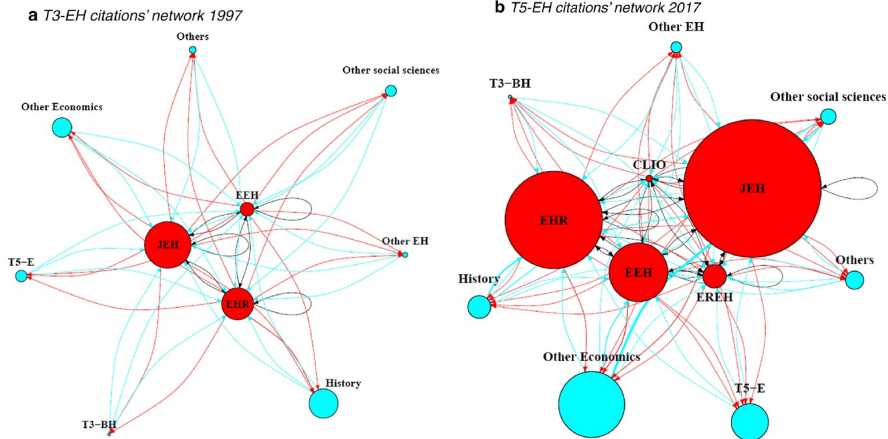
# Trends in Publishing

## Share of EH articles top-5 general interest



Source: Abramitzky (2015), Figure 1

# Trends in Publishing



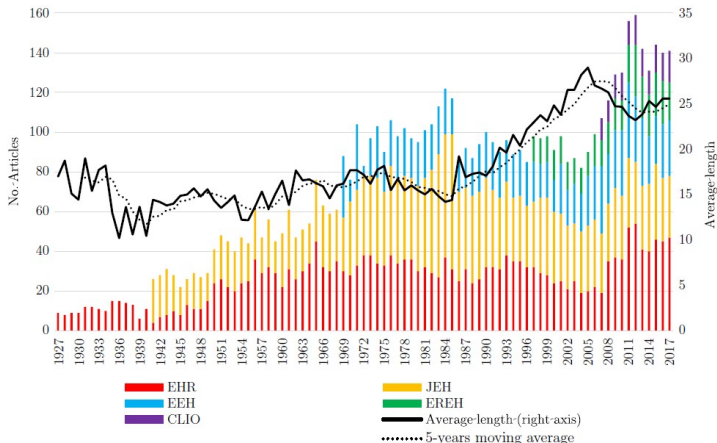
Citation networks

Source: Cioni et al. (2019), Figure 1



# Trends in Publishing

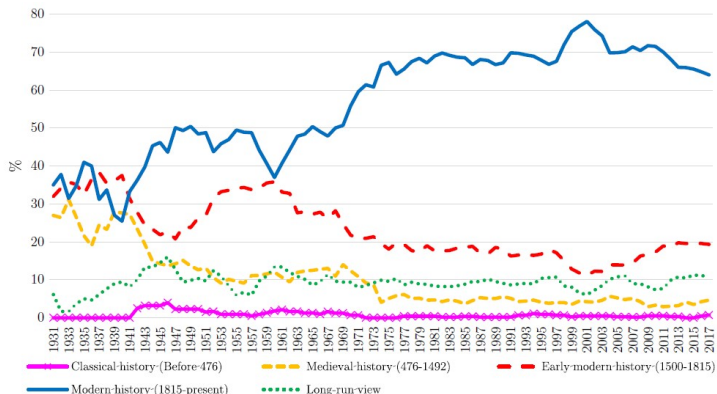
## EH articles and length



Source: Cioni et al. (2019), Figure 2

# Trends in Publishing

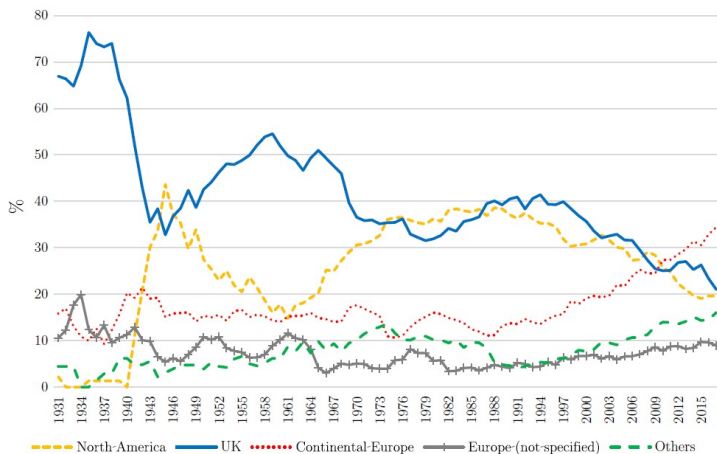
## Share of historical periods



Source: Cioni et al. (2019), Figure 3

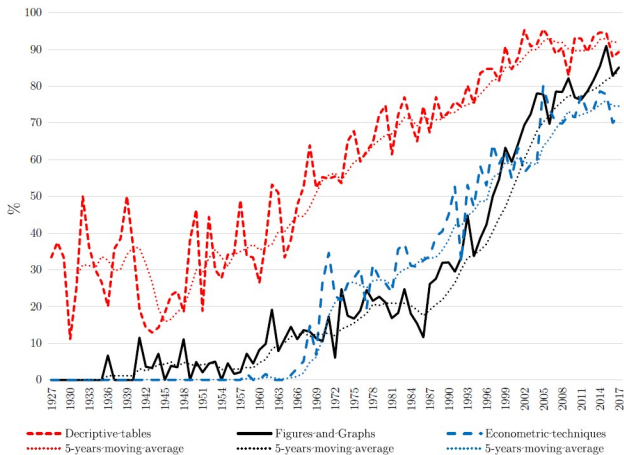
# Trends in Publishing

## Share of continents



Source: Cioni et al. (2019), Figure 5

## Presence of tables, figures, and econometrics



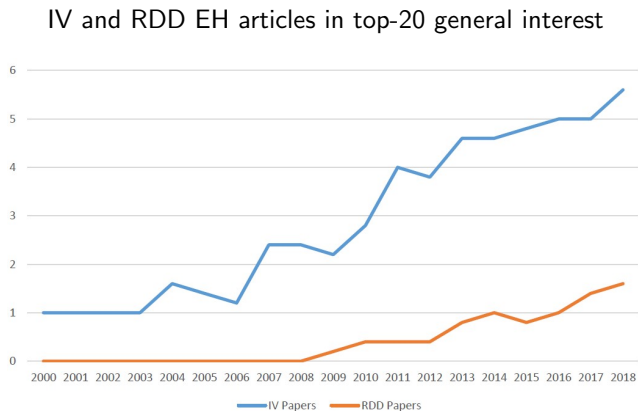
Source: Cioni et al. (2019), Figure 6

## Valencia Caicedo (2021)

- Natural historical experiment literature increasingly centered on identification.
  - Difference-in-differences (parallel counterfactual trends).
  - IV (relevance, excludability, monotonicity).
  - RDD (smooth characteristics along running variable)
- IV and RDD (even DiD) only recover LATE. [Bisin and Moro \(2021\)](#)

- Three samples 2000–2020 from IDEAS/RePEc:
  - EH articles in top-20 general interest journals.
  - Articles in top-5 EH journals.
  - EH articles in top-5 general interest journals.

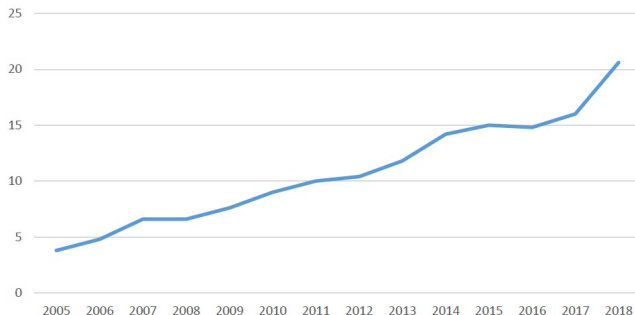
# Historical Econometrics: Trends



Source: Valencia Caicedo (2021), Figure 1

# Historical Econometrics: Trends

Advanced econometrics articles in top-5 EH



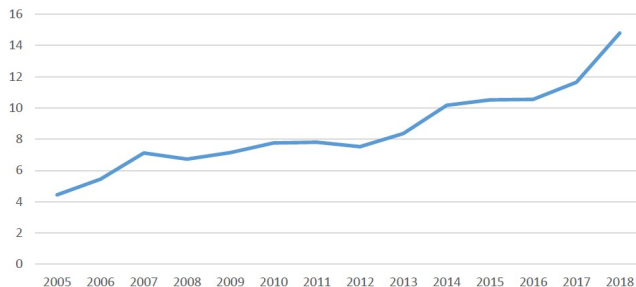
Advanced econometrics: IV, RDD, DiD, panel, matching

Source: Valencia Caicedo (2021), Figure 2A



# Historical Econometrics: Trends

Share of advanced econometrics articles in top-5 EH

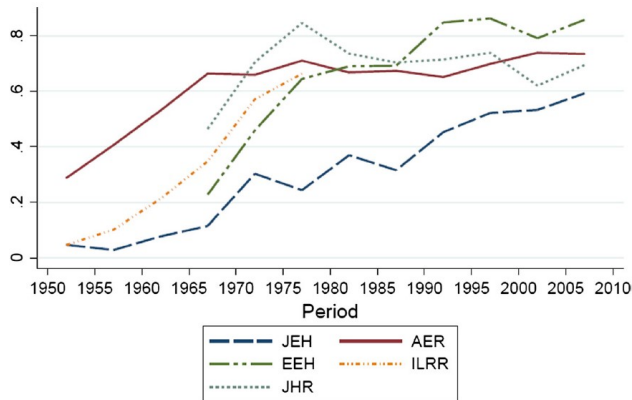


Advanced econometrics: IV, RDD, DiD, panel, matching

Source: Valencia Caicedo (2021), Figure 2B

# Historical Econometrics: Trends

Share of standard econometrics in EH articles



Standard econometrics: regression, logit, probit, MLE, coefficient, standard error

Source: Margo (2018), Figure 1

# First Generation Studies (2001–2011)

- Seminal articles using historical IV: AJR (2001, 2002).
  - Causal impact of institutions on long-run economic development.
  - Endogeneity of institutions to income.
  - IV for historical institutions: settlers mortality.
- Intensive scrutiny of threats to identification.
  - Direct impact of geography through disease environment?
  - Selection of settler mortality data?
  - Better instrument for human capital?

# First Generation Studies (2001–2011)

- Dell (2010)
  - Mining *mita* and development in Peru.
  - Geographical RDD for presence of *mita*.
- Nunn (2008) and Nunn and Wantchekon (2011)
  - Slave trade, development and trust in Africa.
  - IV for historical intensity of slave trade: distance to coast.
- Becker and Woessmann (2009)
  - Protestantism and economic development in late 19th Prussia.
  - IV for prevalence of Protestantism: distance from Wittenberg.
- Dittmar (2011)
  - Printing Press and economic development in Europe.
  - IV for adoption of printing press: distance from Mainz.

# Second Generation Studies (2012–2020)

- Seminal articles using historical RDD: Michalopoulos and Papaioannou (2013, 2014, 2016)
  - Rely on scramble for Africa: haphazard redrawing of African national boundaries in 1884.
  - Random partition of ethnic homelands  $\implies$  random variation of ethnic composition.
  - Explore persistent effects of precolonial ethnic homelands and role of ethnic distribution for conflict.

# Second Generation Studies (2012–2020)

- Alsan (2015)
  - Disease environment and development in Africa.
  - IV for disease environment: suitability for tse-tse fly.
- Adena et al (2015)
  - Media and support for democratic institutions in Nazi Germany.
  - IV for exposure to propaganda: radio signal strength.
- Dell and Querubin (2018)
  - Impact of bombing on communist insurgency.
  - IV for bombing: rounding thresholds in algorithm used to target air strikes.
  - RDD for bombing: military regions with different strategies.

# Second Generation Studies (2012–2020)

- Pascali (2017)
  - Impact of steamship on first globalization (1870–1913).
  - IV for differential reduction in speed: wind patterns.
- Juhasz (2018)
  - Infant industry argument.
  - IV: Napoleonic blockade of UK (1803–1815).
- Becker et al (2020)
  - Forced migration and human capital in post-WWII Poland.
  - Geographical RDD: relocation of expelled Polish.

# Classes of Instruments

- Geography: distance, proximity to rivers, slope of terrain.
- Weather characteristics: wind, rainfall, temperature shocks.
- Land suitability for agriculture.
- Bartik type instruments.
- Forced migrations of ethnic minorities.



# Other RDDs

- Imperial boundaries: Hasburg Empire, Dai Viet. . .
- Battelfronts: Gothic line, Aragon Front. . .
- Pre-colonial boundaries: Kuba Kingdom, matrilineal belt. . .
- Colonial partitions: Dutch cultivation system, rubber concessions. . .
- Assignment thresholds: conscription rules, acreage thresholds . . .

⇒ See Valencia Caicedo (2021) for refinements to RDD.

# Future Trends: More and Better Facts

# Future Trends: More and Better Facts

- “Big data” in economic history:
  - Population data: full count censuses, vital records. . . .
  - Automatic linking techniques.

# Future Trends: More and Better Facts

- “Big data” in economic history:
  - Population data: full count censuses, vital records. . . .
  - Automatic linking techniques.
- Digital humanities:
  - Text as data (see Gentzkow and Shapiro's work).
  - GIS mappings.
  - Ontologies.

⇒ 4D economic history: digitally-driven data design (Mitchener 2015).

# Future Trends: Better Historians

- More interdisciplinarity is necessary (Lamoreaux 2015).

# Future Trends: Better Historians

- More interdisciplinarity is necessary (Lamoreaux 2015).
  - Some applied micro-style economic history lacks . . . history, potentially leading to mistakes.
  - Interesting answers require deep knowledge of the context and data as constructs.
  - In applied micro terms: more history is needed to understand sources of identifying variation, long-run exclusion restrictions, heterogeneity in treatment effects.

# Future Trends: Better Historians

- More interdisciplinarity is necessary (Lamoreaux 2015).
  - Some applied micro-style economic history lacks . . . history, potentially leading to mistakes.
  - Interesting answers require deep knowledge of the context and data as constructs.
  - In applied micro terms: more history is needed to understand sources of identifying variation, long-run exclusion restrictions, heterogeneity in treatment effects.

*Economic history has never been and should never be anything like a closed filed in which practitioners converse mostly with one another. Instead, it stands at a busy intersection of history and the social sciences, where economists, political scientists, sociologists, anthropologists, demographers, and historians come and go.*

Mokyr (2003)

# Topic 2

## Economics History and Labor Economics



# Female labor

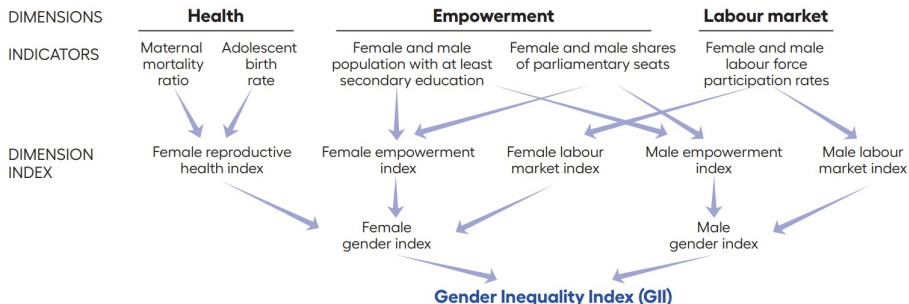
- Female labor and economic development.
- Measuring (female) labor.
- Female labor and wars.

# Female labor and economic development

# Contemporaneous gender inequality worldwide

# Contemporaneous gender inequality worldwide

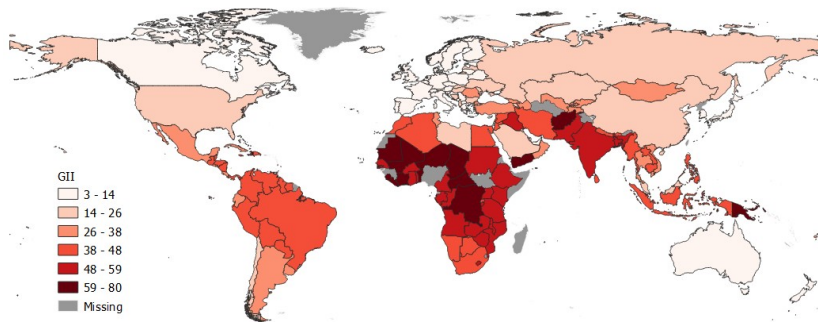
## UN's Gender Inequality Index (GII)



Source: UNDP (2020)

# Contemporaneous gender inequality worldwide

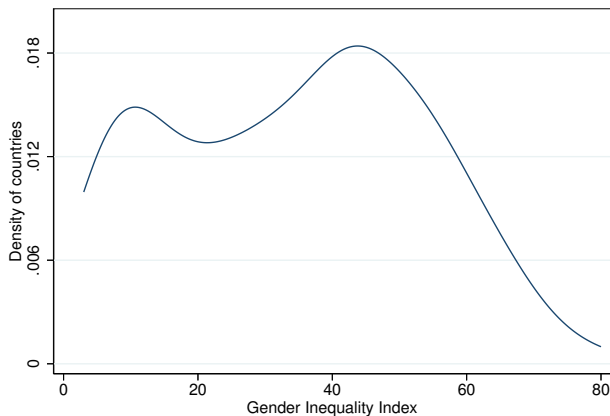
Distribution of Gender Inequality Index (2019)



Source: UNDP (2020)

# Contemporaneous gender inequality worldwide

Distribution of Gender Inequality Index (2019)



Source: UNDP (2020), 162 countries.

# Contemporaneous FLFP worldwide

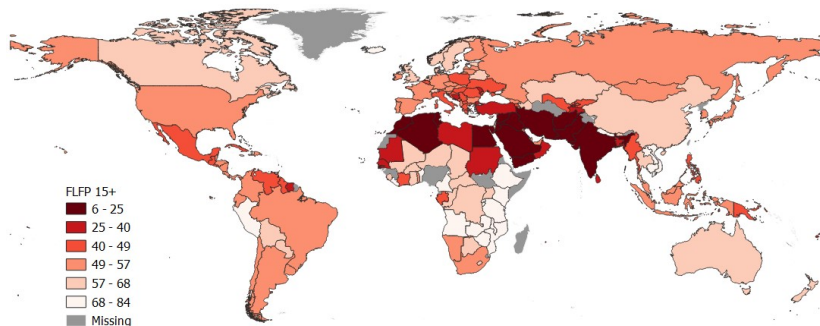
# Contemporaneous FLFP worldwide

- Female labor force participation (FLFP) rate:
  - Share women aged 15 + economically active.
  - Share women aged 30–49 economically active.
- Economically active (ILO definition):
  - Working for pay or profit.
  - Regardless if receive wage.
  - Includes unpaid family farming, family business, self-employed.



# Contemporaneous FLFP worldwide

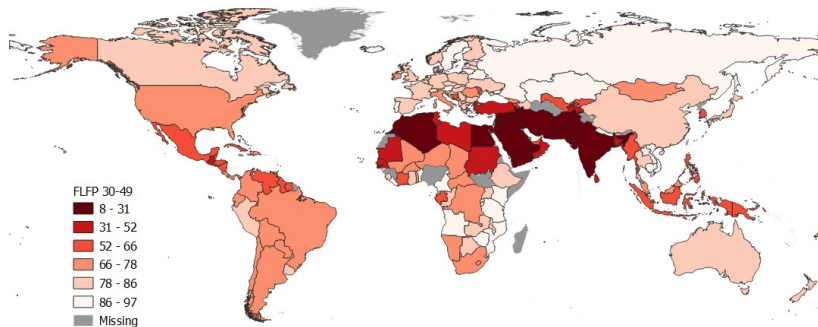
Distribution of FLFP 15+ (2019)



Source: ILO

# Contemporaneous FLFP worldwide

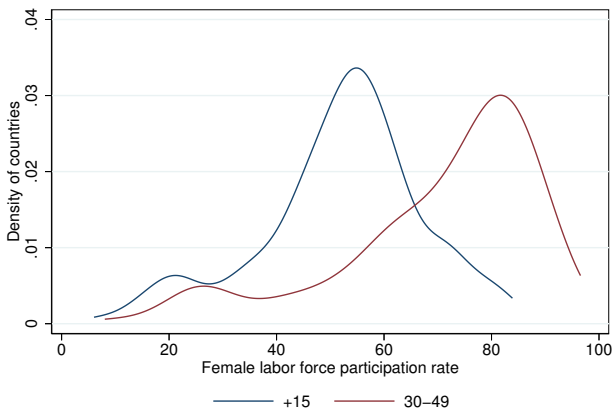
Distribution of FLFP 30–49 (2019)



Source: ILO

# Contemporaneous FLFP worldwide

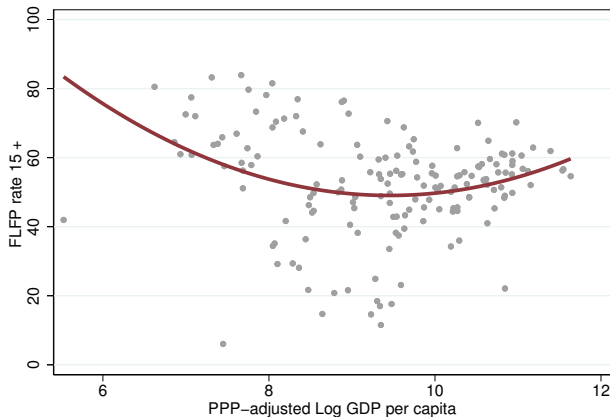
## Distribution of FLFP (2019)



Source: ILO, 189 countries.

# Contemporaneous FLFP and development

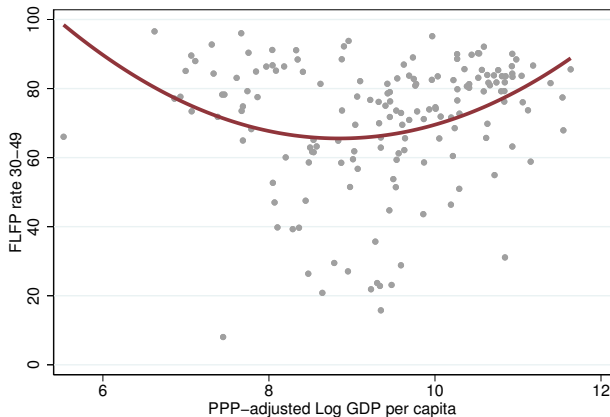
## FLFP and GDP per Capita (2019)



Source: ILO, Penn World Table 10.1, 164 countries.

# Contemporaneous FLFP and development

## FLFP and GDP per Capita (2019)



Source: ILO, Penn World Table 10.1, 164 countries.

# Contemporaneous FLFP and development

- Relationship between FLFP and development?
  - Explore mechanisms behind the U-shape.
  - Use a historical perspective throughout the 20th century.

# Contemporaneous FLFP and development

- Relationship between FLFP and development?
  - Explore mechanisms behind the U-shape.
  - Use a historical perspective throughout the 20th century.
- Why should we care?
  - FLFP instrumental to women's empowerment.
  - Macro perspective: key input in production function.

# Female labor and economic development

- The historical relationship between FLFP and development
  - Stylized facts
  - The structural transformation
  - The 20th century American experience



# Female labor and economic development

- The historical relationship between FLFP and development
  - Stylized facts
  - The structural transformation
  - The 20th century American experience
- Revising old views
  - A steady progress?
  - Measurement issues

# Measuring FLFP in historical perspective

## Olivetti (2014)

- Unbalanced panel 16 countries, 1890–2005.
- Europe, U.K., U.S., Canada.
- FLFP: share women aged 15 + economically active.
- Economically active (ILO definition):
  - Working for pay or profit.
  - Regardless if receive wage.
  - Includes unpaid family farming, family business, self-employed.

# Measuring FLFP in historical perspective

- Pre-WWII data (non-US):
  - International Historical Statistics (IHS) from Mitchell (1998).

“Although imperfect, these data are probably as close as possible to being harmonized in terms of the ILO definition of the employment construct. [. . .] According to Mitchell (1998, 161), the statistics prior to 1968 were unified across different countries and different time periods to adhere to this definition as much as possible.” Olivetti (2014, 170–1)

# Measuring FLFP in historical perspective

- Pre-WWII data (non-US):
  - International Historical Statistics (IHS) from Mitchell (1998).

“Although imperfect, these data are probably as close as possible to being harmonized in terms of the ILO definition of the employment construct. [. . .] According to Mitchell (1998, 161), the statistics prior to 1968 were unified across different countries and different time periods to adhere to this definition as much as possible.” Olivetti (2014, 170–1)
  - IHS data from Bairoch et al. (1968).

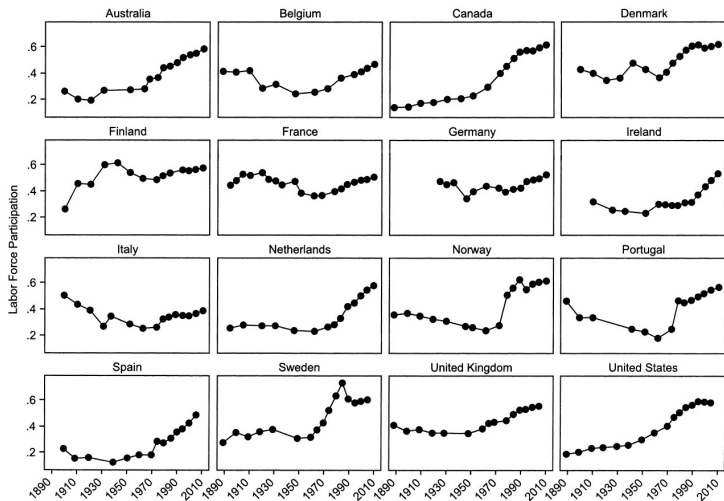
“The immediate source of all statistics is P. Bairoch et al (1968). The original sources are described in detail there. [. . .] [They] ‘tried to as great an extent as possible to unify the statistics in different countries during different periods’, but were unable to achieve anything like perfect comparability.” Mitchell (1998, 161)

# Measuring FLFP in historical perspective

- Pre-WWII data (US):
  - Gainful employment from Goldin (1990, 17).
  - “The 1890 figure is artificially low because it undercounts the paid and unpaid work of married women within the home and on the farm.” Olivetti (2014, 171)
- Post-WWII data: International Labour Organization (ILO).

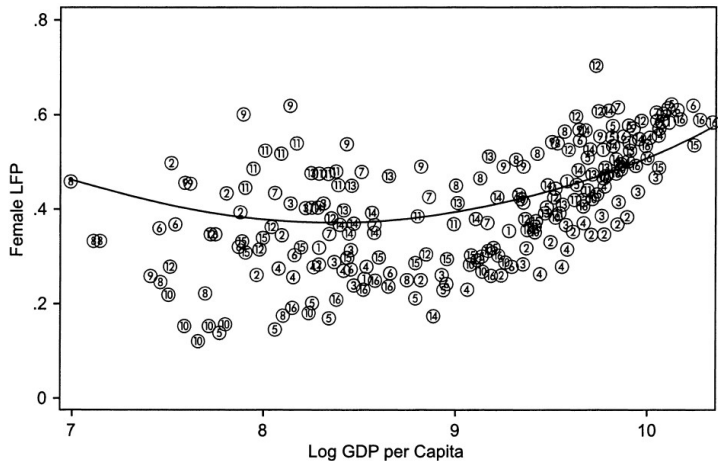
⇒ Challenge in recording labor of married women in farming and family businesses.

# Trends in FLFP, 1890–2005



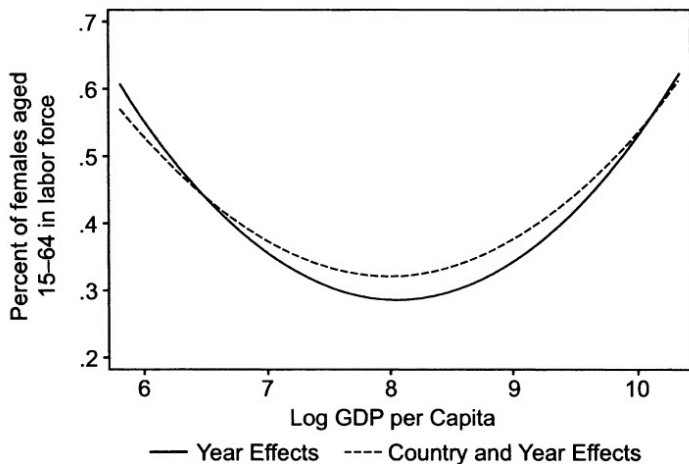
Source: Olivetti (2014, 172).

# FLFP and economic development, 1890–2005



Source: Olivetti (2014, 174).

# FLFP and economic development, 1890–2005



Source: Olivetti (2014, 180).



# The structural transformation

- Structural transformation (Agriculture  $\rightarrow$  Industry  $\rightarrow$  Services):
  - Production of goods relatively intensive in use of “brawn.”
  - Production of services relatively intensive in use of “brain.”
  - Endowments across gender  $\implies$  comparative advantage.

# The structural transformation

- Structural transformation (Agriculture  $\rightarrow$  Industry  $\rightarrow$  Services):
  - Production of goods relatively intensive in use of “brawn.”
  - Production of services relatively intensive in use of “brain.”
  - Endowments across gender  $\implies$  comparative advantage.
- Decreasing portion of the U:
  - Shift from labor to capital intensive agriculture (e.g., plough).
  - Initial manufacturing: use of brawn.
  - Lack of evidence, mainly due to measurement issues.

# The structural transformation

- Structural transformation (Agriculture  $\rightarrow$  Industry  $\rightarrow$  Services):
  - Production of goods relatively intensive in use of “brawn.”
  - Production of services relatively intensive in use of “brain.”
  - Endowments across gender  $\implies$  comparative advantage.
- Decreasing portion of the U:
  - Shift from labor to capital intensive agriculture (e.g., plough).
  - Initial manufacturing: use of brawn.
  - Lack of evidence, mainly due to measurement issues.
- Increasing portion of the U:
  - Expansion of service sector  $\implies$   $\uparrow$  demand for female labor.
  - Marketization of home production  $\implies$   $\uparrow$  supply for female labor.

# The structural transformation

Amplifying factors of increasing part of U [Olivetti and Petrongolo \(2016\)](#)

- Technological change:  $\uparrow$  time for market + elasticity of labor supply.
  - Contraceptive pill. [Goldin and Katz \(2006\)](#) [Bailey \(2006\)](#)
  - Substitutes to maternal lactation. [Albanesi and Olivetti \(2016\)](#)
  - Household production technologies. [Greenwood et al. \(2005\)](#)
  - Availability of child care. [Attanasio et al. \(2008\)](#)

# The structural transformation

Amplifying factors of increasing part of U Olivetti and Petrongolo (2016)

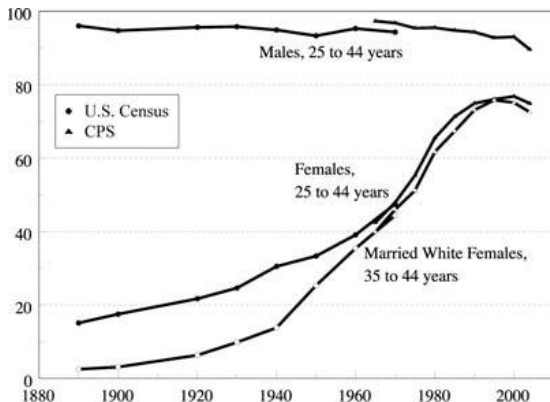
- Technological change:  $\uparrow$  time for market + elasticity of labor supply.
  - Contraceptive pill. Goldin and Katz (2006) Bailey (2006)
  - Substitutes to maternal lactation. Albanesi and Olivetti (2016)
  - Household production technologies. Greenwood et al. (2005)
  - Availability of child care. Attanasio et al. (2008)
- Changes in social attitudes toward female labor.
  - Changes in husbands' attitudes toward working wives. Fernández et al. (2004)
  - Changes in women's beliefs about costs of maternal employment. Fogli and Veldkamp (2011)

# Trends in 20th-century U.S.

Goldin (1991, 2006)

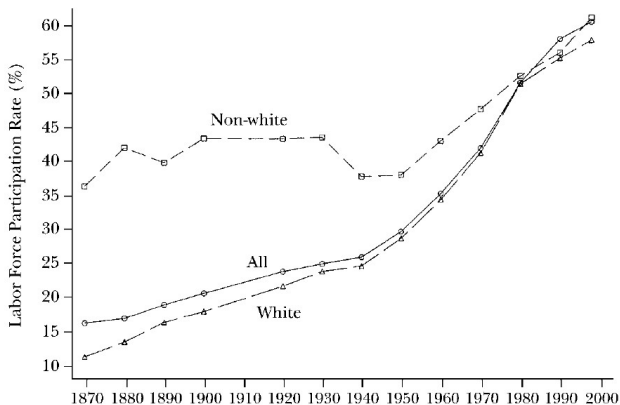
- Phase I (initial): 1890s–1930.
- Phase II (transition): 1930–1950.
- Phase III (roots of revolution): 1950–1970s.
- Phase IV (quiet revolution): post-1970s.

# Trends in 20th-century U.S.



Source: Goldin (2006, 4).

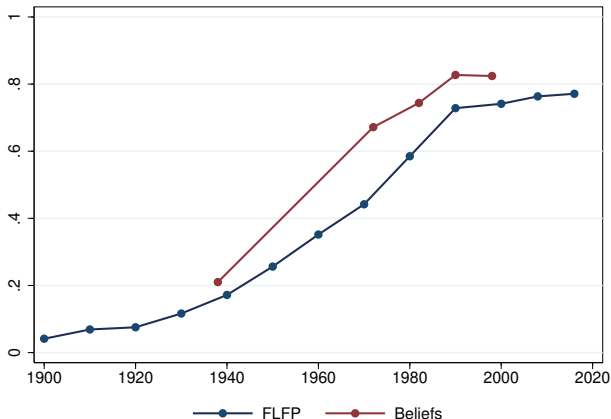
# Trends in 20th-century U.S.



Source: Costa (2000, 104).



# Trends in 20th-century U.S.

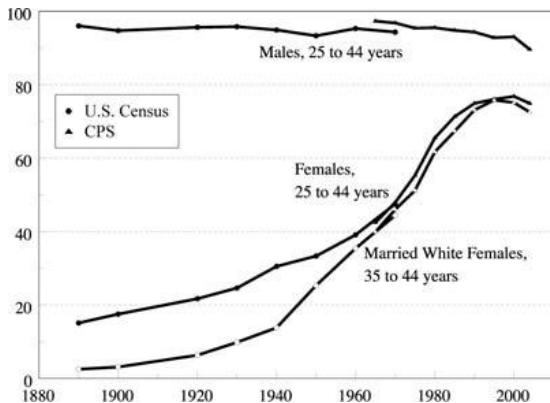


- FLFP sample: white married women aged 30–49 in nonfarm [IPUMS 1900–2016](#)
- Share approves of “A married woman earning money in business or industry if she has a husband capable of supporting her” [Gallup 1938](#) [GSS 1972–1998](#)

# Conceptual Framework

- Changes in female labor supply depends on:
  - Own-wage (compensated) elasticity.
  - Income elasticity (esp. husband income).
- Each phase characterized by different parameters.
- Trends can be explained by:
  - Changes in elasticities of labor supply.
  - Shifts in labor supply.
  - Shifts in labor demand.

# Phase I: 1890s–1930

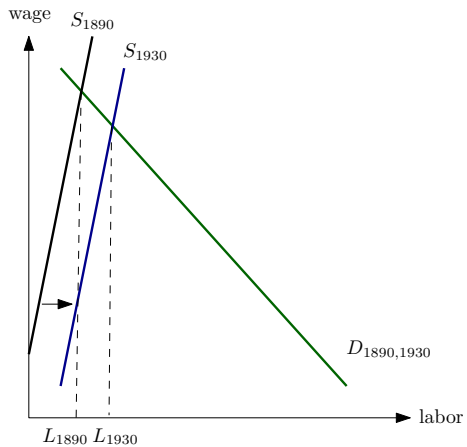


Source: Goldin (2006, 4).

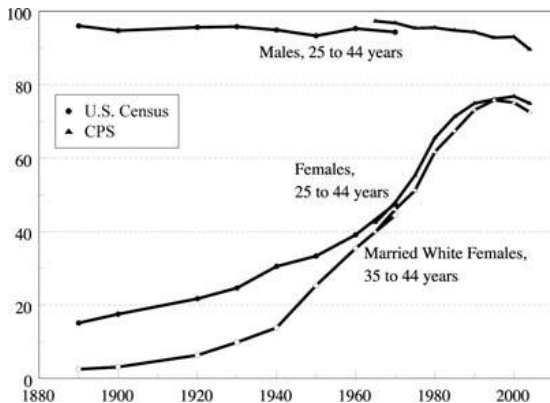
# Phase I: 1890s–1930

- Trends:
  - FLFP married 35–44 increase: 3% to 10%.
  - Female workers: young, unmarried, uneducated, poor.
  - Sectors: manufacturing, domestic, laundresses.
- Driving forces:
  - Strong negative income effect (stigma)  $\implies$  drop out at marriage.
  - Small substitution effect (inflexible schedules).
  - Limited shifts in labor supply ( $\downarrow$  fertility,  $\uparrow$  education)

# Phase I: 1890s–1930



# Phase II: 1930–1950

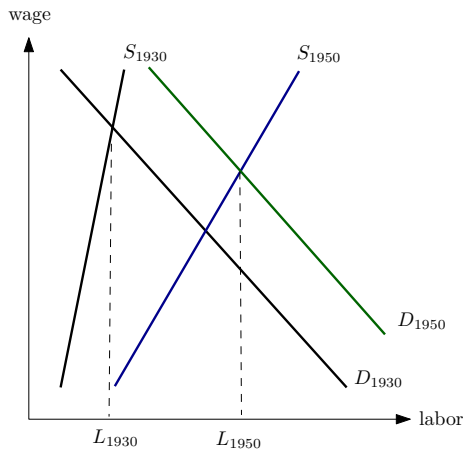


Source: Goldin (2006, 4).

# Phase II: 1930–1950

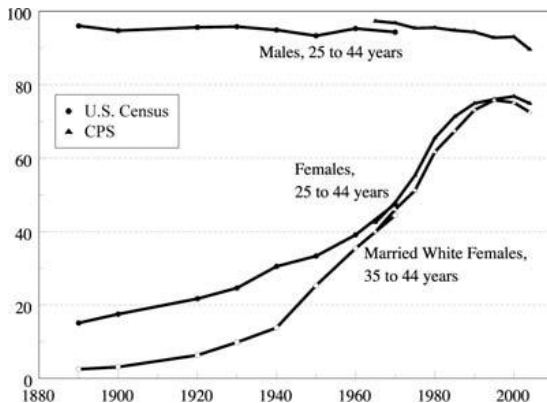
- Trends:
  - FLFP married 35–44 increase: 10% to 25%.
  - Female workers: older, married, more educated.
  - Sectors: more clerical and office workers.
- Driving forces:
  - Weaker income effect (↓ stigma).
  - Increased substitution effect (↑ part-time work, ↑ household technology).
  - Large shifts of labor supply (↑ education) and demand (↑ offices, ↓ marriage bars).
  - Role of WWII?

## Phase II: 1930–1950





# Phase III: 1950–1970

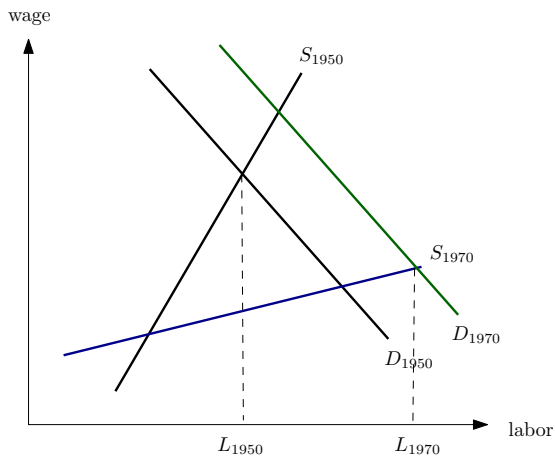


Source: Goldin (2006, 4).

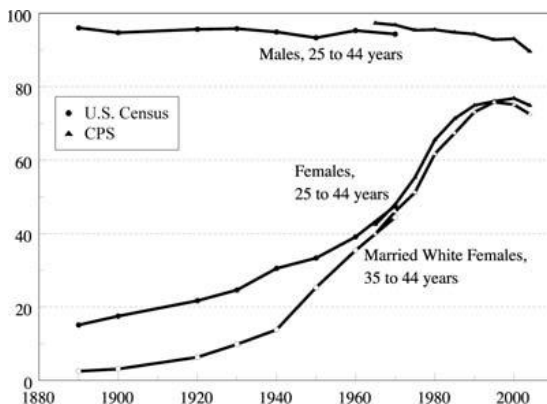
# Phase III: 1950–1970

- Trends:
    - FLFP married 35–44 increase: 25% to 46%.
    - Female workers: generalized.
    - Sectors: services (secretaries, teachers, nurses, social workers. . .).
  - Driving forces:
    - Weaker income effect (↓ stigma).
    - Increased substitution effect (↑ part-time work, ↑ household technology).
    - Large shifts in demand (↑ services).
- ⇒ Still jobs, no careers.

# Phase III: 1950–1970



# The Quiet Revolution – Phase IV: post-1970s



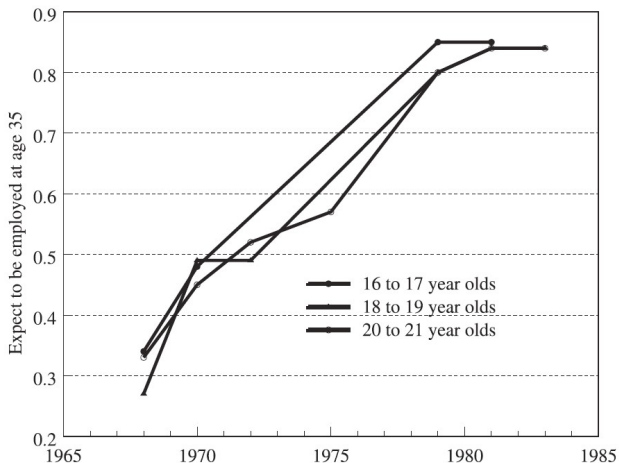
Source: Goldin (2006, 4).

# The Quiet Revolution – Phase IV: post-1970s

- Increases in phases II and III: precondition.
- Increases post-1970 insufficient to characterize revolution.
- Indicators that reflect changes in:
  - Horizons.
  - Identity.
  - Decision making.

# Expanded horizons

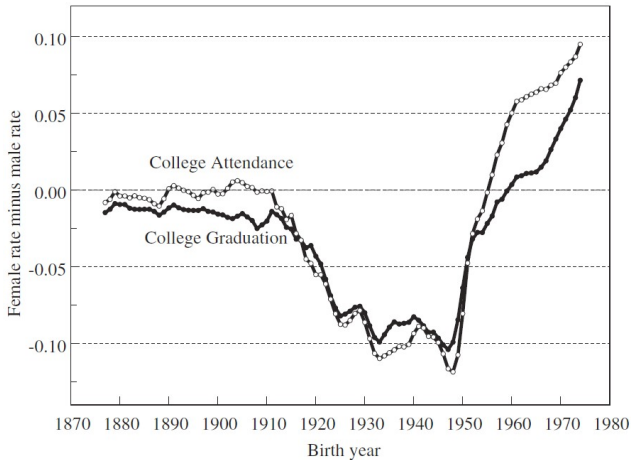
Employment expectations of female youth by age



Source: Goldin (2006, 9) from NLSY

# Expanded horizons

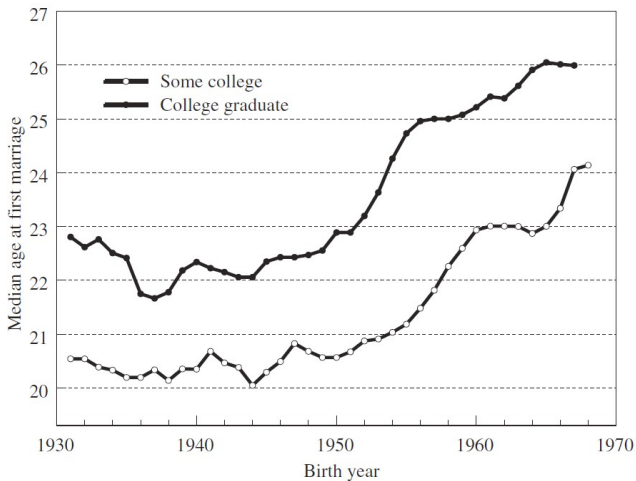
Female minus male college attendance and graduation rates



Source: Goldin (2006, 10)

# Expanded horizons

## Median age at first marriage

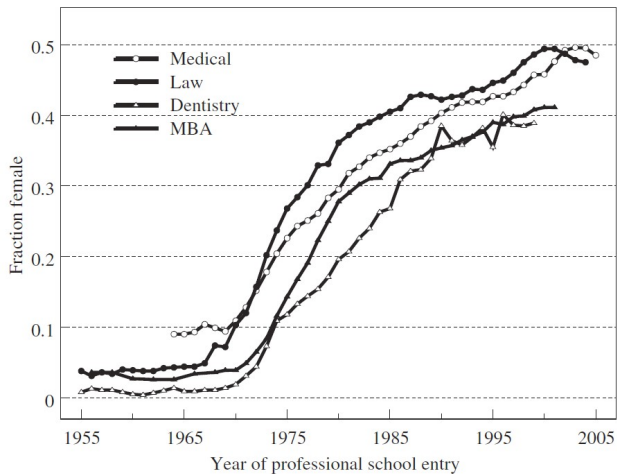


Source: Goldin (2006, 10)



# Expanded horizons

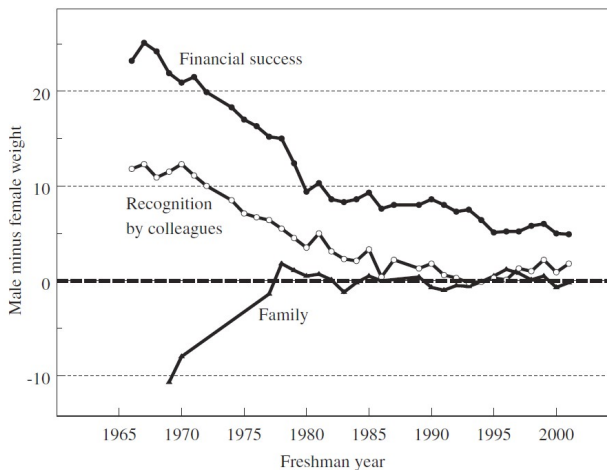
Fraction female among first-year students in professional programs



Source: Goldin (2006, 11)

# Altered identities

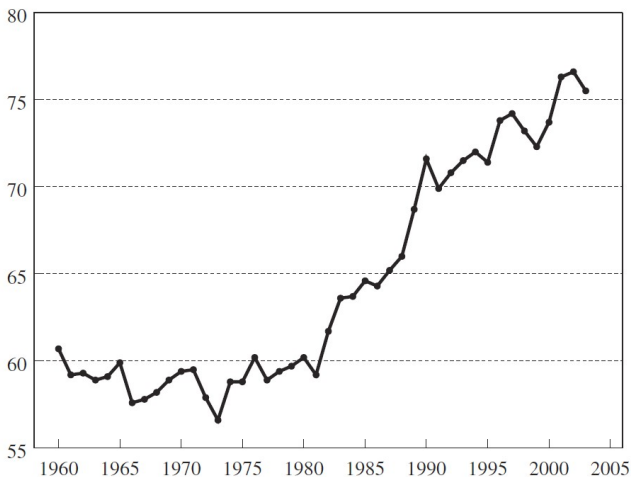
## Personal satisfaction factors for college freshmen



Source: Goldin (2006, 12)

# Relative earnings and occupations

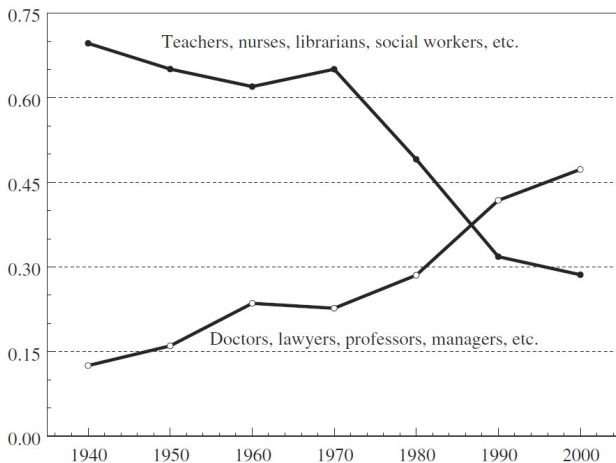
Women's earnings as a percentage of men's earnings



Source: Goldin (2006, 13)

# Relative earnings and occupations

Occupations of college graduate women (30–34)



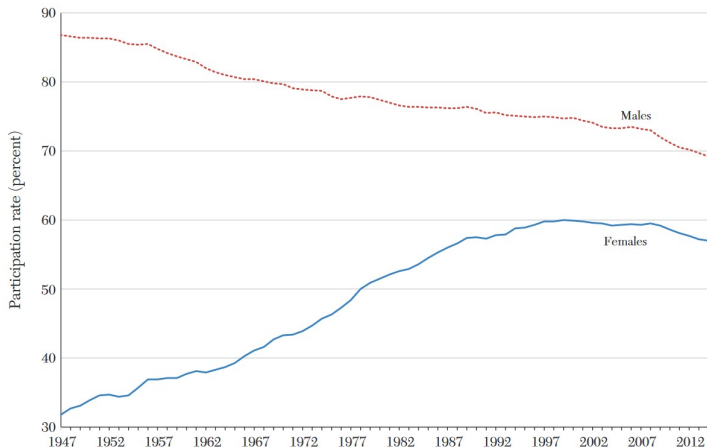
Source: Goldin (2006, 13)

# A steady progress?

- The meta narrative of a steady convergence maybe unwarranted:
  - The “engine of liberation” has stalled in recent years.
  - Historical crises oftentimes have backward consequences.
  - Women are still most affected by the business cycle.

# A stalling engine

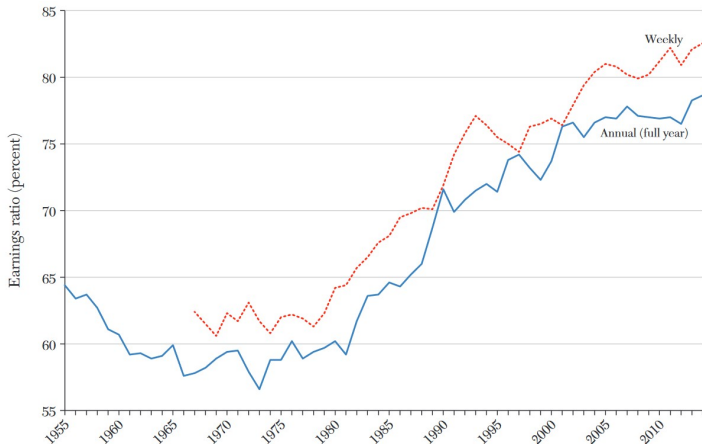
Trends in male and female LFP rates 16+ in the U.S.



Source: Blau and Kahn (2017)

# A stalling engine

Female-to-male earnings ratios of full time workers in the U.S.



Source: Blau and Kahn (2017)

## Measuring (female) labor



# Measurement issues

- Women have always worked.
- The way (gender-biased) statistics capture female labor has changed historically.
- Careful thinking of historical measurement might question standard view.

# Measurement issues: economics vs history

- The heterogeneity of sources reflects methodological and epistemological differences across disciplines.
- In economics (and quantitative history):
  - Preferred sources: nation-wide coverage, all sectors, temporal recurrence.
  - Objective: construct time series and panels for statistical treatment.
  - Main sources for this approach: population censuses.

# Measurement issues: economics vs history

- The heterogeneity of sources reflects methodological and epistemological differences across disciplines.
- In economics (and quantitative history):
  - Preferred sources: nation-wide coverage, all sectors, temporal recurrence.
  - Objective: construct time series and panels for statistical treatment.
  - Main sources for this approach: population censuses.
- Criticism from historians of statistics:
  - Anachronism of long time series.
  - Lack of source criticism, its production context, and representations of its producers.
  - Need to deconstruct census data (Schweitzer 2002, Maruani Meron 2012, *Genèse* 1992).

# Measurement issues: economics vs history

- In history:
  - Preferred sources: micro-analytical scale, centered on specific occupation.
  - Objective: uncover female labor less visible in macro sources.
  - Main sources for this approach focus on specific occupations:
    - Workers: firm archives.
    - Nurses: hospital archives.
    - Domestic workers: private archives of bourgeoisie and nobility.
    - Teachers: school institutions archives.
    - Public servants: public administrations archives.

# Emerging categories

Topalov (1999)

- Categories of active populations emerge and change late 19th-early 20th century as formal markets and waged labor develop.
- Constructed reality (formal, informal) dependent on norms (e.g., occupation's wife).
- Difficult issues to solve for census enumerators:
  - How much "labor" to be "active"? Should it be paid?
  - Account for individual activity or within production unit?
  - Domestic labor and gainful employment vs "housewife."
  - Unoccupied classes (retired, landlords, students, ...).
  - Unemployed and looking for work.

# Emerging categories

- France:
  - Active-inactive distinction: 1861.
  - Notion: “profession, position ou occupation” (1851–1901) then “profession”.
- U.S.:
  - Active-inactive distinction: 1870.
  - Notion: “profession, occupation or trade” (1820–1900), “trade or profession” (1910), then “occupation”.
- U.K.:
  - Active-inactive distinction: 1881.
  - Notion: “rank, profession, or occupation” (1801–1881), “profession or occupation” (1891), then “personal occupation”.

# Measuring female labor in French censuses

- The emergence of the notion of active population:
  - 1836–1856: occupation of head of household attributed to entire household.
  - 1861–1891: active population = heads of households + householders declaring a distinct occupation. Others + domestics = inactive population.
  - 1896: individualization of the notion of gainful employment

# Measuring female labor in French censuses

- Still, notion takes time to stabilize:
  - After 1896: gainful employment (paid domestic vs “housewife”).
  - After 1906: wives of farm owners counted as active.
  - Wives of shopkeepers not all counted as active.
- Changes in counting techniques matter:
  - Starting 1901: centralization in Paris of census counting operations with punched card machines.
  - Starting 1906: cross-tabulations between individual and establishment bulletins.

⇒ Different trends depending on definition.

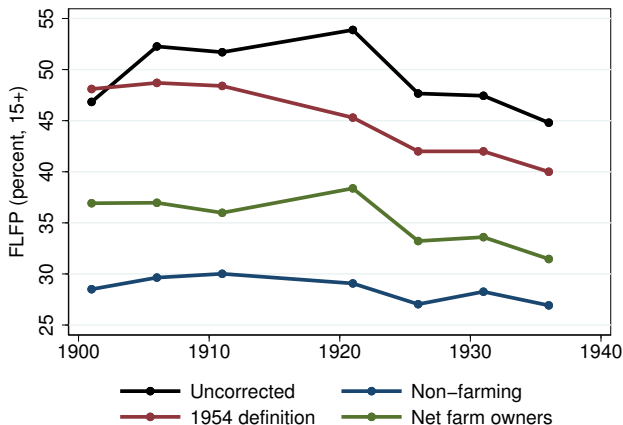


# Measuring female labor in French censuses

- FLFP 15+ uncorrected census data. [Olivetti \(2014, 2016\)](#)
- FLFP 15+ non-farming. [Daric \(1947\)](#) [Maruani and Méron \(2012\)](#)
- FLFP 15+ estimated based on 1954 definition and constant ratio male-female in agriculture by age group in 1954.  
[Marchand and Thello \(1998\)](#) [Costa \(2000\)](#)
- FLFP 15+ net female farm owners. [Beohnke and Gay \(2022\)](#)

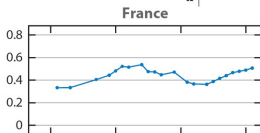
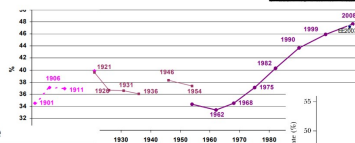
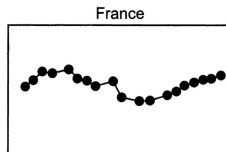
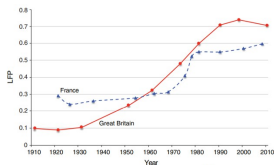
# Measuring female labor in French censuses

Female labor force participation, 15+



# Measuring female labor in French censuses

## Female labor force participation, 15+



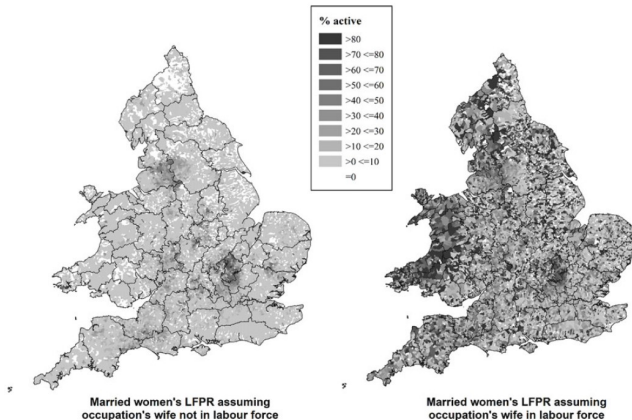
# Measuring female labor in British censuses

You (2020)

- Measure of married women's labor problematic.
- "Occupation's wife" is largest group in British census 1851–1911.
- Compare information in census reports and (100%) Census Enumerators' Books.
- Before 1881: wives of farmers, shopkeepers, inkeepers, butchers assumed to work with husbands in same trade unless explicitly declare other occupation.
- Unclear how census enumerators followed instructions.
- Shows enumerator-level idiosyncrasies (local adherence to norms).

# Measuring female labor in British censuses

## Spatial pattern of married women's LFP, 1881



Source: You (2020, 12)

## Chiswick and Robinson (2021)

- Labor contributions by women to family-operated businesses heavily discounted in official statistics.
- Correcting for farm, merchant, craft family workers, and boarding house keepers dramatically changes FLFP estimates 1860–1920.
- Overall change might be due to shift in unreported work in family businesses to reported work in paid labor market.

# Measuring female labor in US censuses

## Augmented Labor Force Participation Rates, 16+

Formal or Informal Occupation	1860 – All		1920 – All	
	Males	Females	Males	Females
Gainful Occupation	87.6	15.7	89.9	24.0
All Family Workers	5.6	40.9	2.9	25.8
Farm Family Workers	5.1	35.8	2.3	19.9
Merchant Family Workers	0.3	3.0	0.4	3.6
Craft Family Workers	0.1	0.8	0.1	1.5
Boardinghouse Family Workers	0.1	0.8	0.1	0.5
Multi-Job Family Workers	0.1	0.5	0.0	0.3
Total Augmented	93.2	56.6	92.8	49.8
Total Not Working	6.8	43.4	7.2	50.2
Total Adult Population	100. 0	100.0	100.0	100.0
Sample Size	81,657	76,605	357,236	339,181

Source: Chiswick and Robinson (2021)

- Anachronism of long-run time series?
  - Long-run analyses can make unwarranted comparisons over time, assuming the existence of the object (does dichotomy active / inactive make sense?).
  - Be clear about joint measure and construction of object (historicize measure).
  - Comparisons when measure stabilized within short time-frame (though interesting phenomena across changing conditions).



# Female labor and wars

# Why female labor and wars?

- Substantial historical importance (cat. 1).
  - Wars of the 20th century were among most dramatic events.
  - Gender relations dramatically altered by industrial warfare.
  - Critical junctures for social change.

# Why female labor and wars?

- Substantial historical importance (cat. 1).
  - Wars of the 20th century were among most dramatic events.
  - Gender relations dramatically altered by industrial warfare.
  - Critical junctures for social change.
- General economic insights (cat. 2).
  - Test economic theory (sex ratio, marriage and labor market).  
[Abramistky et al \(2011\)](#) [Boehnke and Gay \(2020\)](#)
  - Learn about implications of Revolution of female labor: source of exogenous variation in female labor supply. [Acemoglu et al. \(2004\)](#)
  - Learn about mechanisms of intergenerational transmission: source of exogenous variation in former generation female labor supply.  
[Fernandez, Fogli, and Olivetti \(2004\)](#) [Gay \(2021\)](#)

# Why female labor and wars?

- Substantial historical importance (cat. 1).
  - Wars of the 20th century were among most dramatic events.
  - Gender relations dramatically altered by industrial warfare.
  - Critical junctures for social change.
- General economic insights (cat. 2).
  - Test economic theory (sex ratio, marriage and labor market).  
[Abramistky et al \(2011\)](#) [Boehnke and Gay \(2020\)](#)
  - Learn about implications of Revolution of female labor: source of exogenous variation in female labor supply. [Acemoglu et al. \(2004\)](#)
  - Learn about mechanisms of intergenerational transmission: source of exogenous variation in former generation female labor supply.  
[Fernandez, Fogli, and Olivetti \(2004\)](#) [Gay \(2021\)](#)
- Roots of contemporary gendered structure of labor markets (cat. 3).

# Female labor and wars

- Female labor and WWII in the U.S.
- Female labor and WWI in France.

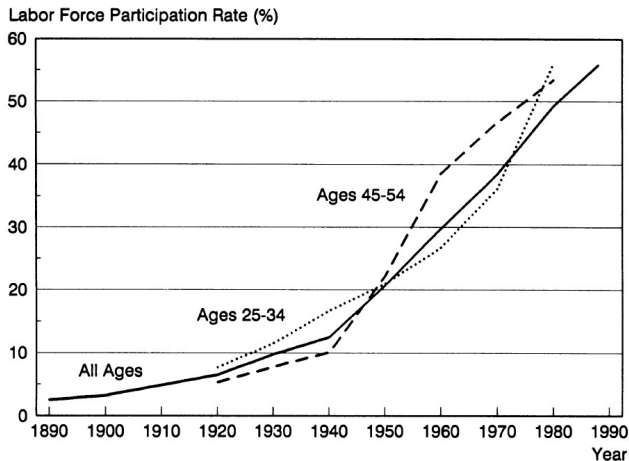
# Female Labor and WWII in the U.S.

# Female Labor and WWII in the U.S.

- Goldin (1991)
- Acemoglu, Autor, and Lyle (2004)
- Goldin and Olivetti (2013)
- Fernandez, Fogli, and Olivetti (2004)
- Rose (2018)

# Female Labor and WWII in the U.S.

## LFP of white married women



Source: Goldin (1991), Figure 1



# WWII: a watershed event?

- Supply side during and retention after:
  - Mobilized husband  $\implies$   $\uparrow$  time +  $\downarrow$  income.
  - Patriotic duty to join war effort.
  - Job training during wartime employment.
  - Changes in husbands' norms.

# WWII: a watershed event?

- Supply side during and retention after:
  - Mobilized husband  $\implies$   $\uparrow$  time +  $\downarrow$  income.
  - Patriotic duty to join war effort.
  - Job training during wartime employment.
  - Changes in husbands' norms.
- Demand side during and retention after:
  - Increased demand and wages for wartime industries.
  - Updated beliefs about women's ability by employers.
  - Erosion of marriage bars.

# WWII: a watershed event?

- Supply side during and retention after:
  - Mobilized husband  $\implies$   $\uparrow$  time +  $\downarrow$  income.
  - Patriotic duty to join war effort.
  - Job training during wartime employment.
  - Changes in husbands' norms.
- Demand side during and retention after:
  - Increased demand and wages for wartime industries.
  - Updated beliefs about women's ability by employers.
  - Erosion of marriage bars.
- But wartimes jobs disappeared after the war (and in industrial sector).

- Retrospective work histories survey (Palmer Survey):
  - 4,350 women.
  - Working full-time for at least a month in 1950.
  - Urban sample across the U.S.
  - Labor status: Jan-1940, Dec-1944, Dec-1949, Jan-1951.
  - Rich information (industry, earnings, education, marital status...)

# Goldin (1991)

- Research questions:
  - Did women who entered the labor force during WWII remained employed in 1950?
  - Did these women left the labor force, and the post-WWII rise came from a post-war movement of women into the labor force?

# Goldin (1991)

- Research questions:
  - Did women who entered the labor force during WWII remained employed in 1950?
  - Did these women left the labor force, and the post-WWII rise came from a post-war movement of women into the labor force?
- Results among white married women aged 35–64 employed in 1950:
  - 52% already employed in 1940.
  - 22% joined labor force during the war.
  - 25% joined the labor force after the war.

# Goldin (1991)

- Research questions:
  - Did women who entered the labor force during WWII remained employed in 1950?
  - Did these women left the labor force, and the post-WWII rise came from a post-war movement of women into the labor force?
- Results among white married women aged 35–64 employed in 1950:
  - 52% already employed in 1940.
  - 22% joined labor force during the war.
  - 25% joined the labor force after the war.
- Results linking with CPS 1944 (with retrospective questions): half of wartime entrants exited labor force by 1951.

⇒ Relatively modest implications based on selected sample.

## LFP of white married women

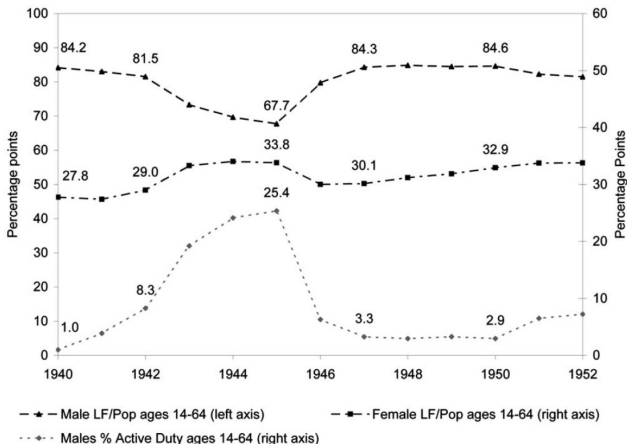
Category	Participation rate (percentage)		
	Age 35–64 in 1950	Age 35–44 in 1950	Age 45–54 in 1950
In labor force in 1950 and in labor force in:			
1940 and 1944	45.5	41.7	48.1
1940 but not 1944	6.6	8.2	5.3
1944 but not 1940	21.8	20.9	22.8
Neither 1940 nor 1944	26.2	29.1	23.7
Number of observations:	1,199	635	451
In labor force in January 1951 and in labor force in:			
1940 and 1944	47.7	44.7	49.8
1940 but not 1944	6.3	7.5	5.4
1944 but not 1940	21.0	19.9	22.3
Neither 1940 nor 1944	25.0	27.7	22.5
Number of observations:	1,067	558	404



- How did increased FLFP affect structure of male and female wages?
- Use WWII as exogenous source of variation in female labor supply.
- Use state-level male mobilization to capture pull force to female labor during the war.

# Female labor and WWII in the U.S.

MLFP, FLFP, military active service personnel, 1940–1952

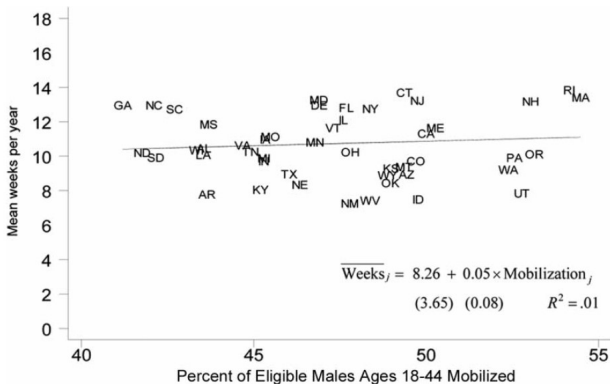


Source: Acemoglu et al. (2004), Figure 2

- Labor data:
  - Decennial censuses 1940–1950 (1% extracts).
  - Sample: women aged 14–64 with earnings information, not in institutional group quarters, not employed in farming.
  - Measure of FLFP: weeks worked.
- Mobilization data:
  - Mandatory draft through 6 waves 1940–1947, men 18–44.
  - Many deferrals for farmers and little nonwhite units.
  - Use state share of enlisted men aged 13–44 in 1940 to proxy for decline in supply of male labor (range from 43 to 53%;  $sd = 3.2$ ).

# Impact of WWII in the U.S. on FLFP?

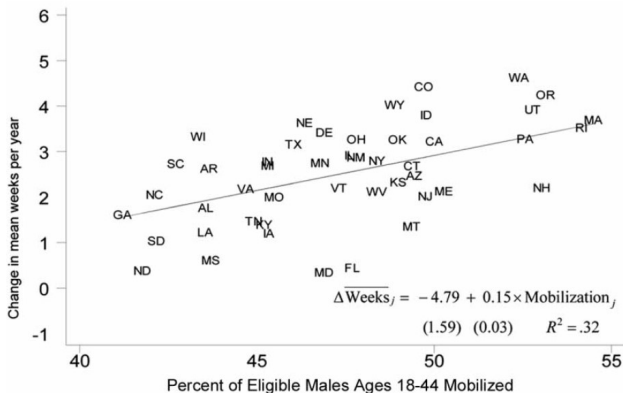
WWII mobilization rates and mean female weeks worked per year, 1940



Source: Acemoglu et al. (2004), Figure 3

# Impact of WWII in the U.S. on FLFP?

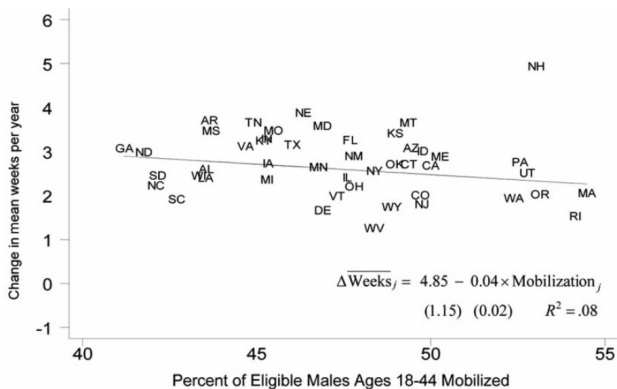
WWII mobilization rates and change in female weeks worked per year, 1940–1950



Source: Acemoglu et al. (2004), Figure 4

# Impact of WWII in the U.S. on FLFP?

WWII mobilization rates and change in female weeks worked per year, 1950–1960



Source: Acemoglu et al. (2004), Figure 5

# Source of variation in mobilization rates

## 1940 state-level determinants of WWII mobilization rates

	MEAN	REGRESSION									
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Share farmers	.15 [.11]	-.15 (.05)	-.16 (.04)	-.17 (.03)	-.17 (.04)	-.23 (.06)	-.26 (.04)	-.22 (.04)	-.17 (.05)	-.16 (.04)	-.17 (.05)
Share nonwhite	.10 [.11]	-.01 (.05)	-.07 (.04)		-.03 (.06)	-.38 (.27)	.04 (.05)	-.03 (.05)	-.03 (.06)	.02 (.06)	-.03 (.06)
Average education	8.89 [.71]	.02 (.01)		.01 (.01)	.01 (.01)	.01 (.01)	.03 (.01)	.01 (.01)	.01 (.01)	.01 (.01)	.01 (.01)
Share aged 13-24	.42 [.03]	.25 (.34)					.73 (.24)				
Share aged 25-34	.31 [.01]	.15 (.48)					.38 (.48)				
Share German	.007 [.006]	-3.19 (.89)						-1.88 (.55)			
Share Italian or Japanese	.010 [.012]	1.70 (.52)							.00 (.42)		
Share married	.50 [.03]	-.10 (.17)								-.22 (.13)	
Share fathers	.47 [.03]	.08 (.13)									.00 (.12)
$R^2$		.78	.57	.58	.58	.39	.68	.67	.58	.61	.58
Southern states		yes	yes	yes	yes	no	yes	yes	yes	yes	yes

Source: Acemoglu et al. (2004), Table 4

# Empirical strategy

$$y_{ist} = \delta_s + \gamma d_{1950} + X'_{ist} \beta_t + \psi d_{1950} m_s + \varepsilon_{ist}$$

- $y_{ist}$ : weeks work by women  $i$  residing in state  $s$  in year  $t$ .
- $\delta_s$ : state of residence FE.
- $X_{ist}$ : covariates (state of birth FE, age, race, state share farmers and nonwhites, average schooling in 1940, interacted with 1950 dummy).
- $m_s$ : state mobilization rate.



## Impact of WWII mobilization rates on female labor supply

	REGRESSION			
	(1)	(2)	(3)	(4)
	A. White Females (N= 530,026)			
Mobilization rate × 1950	11.17 (1.89)	9.85 (2.05)	10.64 (2.65)	8.51 (2.37)
1940 male share farmers × 1950			1.74 (1.08)	1.04 (1.05)
1940 male share nonwhite × 1950			-1.96 (1.15)	-.72 (1.37)
1940 male share average education × 1950				.52 (.16)
$R^2$	.01	.18	.18	.18
Includes marital status, age, state of birth	no	yes	yes	yes

- ↑ 10pp mobilization  $\implies$  ↑ 0.85 week (0.5 sd).

# Baseline results

Impact of WWII mobilization rates on female labor supply

	1940–50		1950–60	
	Weeks Worked (1)	Any Weeks Worked (2)	Weeks Worked (3)	Any Weeks Worked (4)
A. Baseline specification	9.85 (2.05)	.063 (.069)	-7.25 (1.81)	-.057 (.049)
$R^2$	.18	.18	.15	.14
Observations	530,026		615,590	
B. Includes share farmers, share nonwhite, average education	8.51 (2.37)	.174 (.071)	2.15 (1.95)	-.006 (.067)
$R^2$	.18	.18	.15	.14
Observations	530,026		615,590	

- $\uparrow$  10pp mobilization  $\implies$   $\uparrow$  2pp in FLFP (any weeks worked in year).

# Goldin and Olivetti (2013)

- Identify short and long-run impacts of WWII on women's labor supply.
- Focus on young white women aged 16–29 during the war (1941–1945): 25–34 in 1950, 35–44 in 1960.
- Assess differential treatment of the war across groups (age, education, children, marital status).
- Use state-level mobilization rates of men aged 18–44 (ranges from 40 to 54%).
- Data from censuses 1940–1960.
- Sample of white married women with  $\geq 12$  years schooling, not in institutional group quarters, not employed in farming, born and residing in U.S.

# Empirical strategy

$$y_{ist} = \psi_s + \gamma d_\tau + \alpha_t Z_{it} + \beta_t d_\tau X_{s1940} + \delta d_\tau MOB_s + \varepsilon_{ist}$$

- $y_{ist}$ : weeks work or LFP by women  $i$  residing in state  $s$  in year  $t$ .
- $\psi_s$ : state of birth FE.
- $d_\tau$ : year indicator (1950 or 1960).
- $Z_{it}$ : covariates (number of children and age indicators).
- $X_{s1940}$ : state of birth covariates (share farmers, nonwhite, average education).
- $MOB_s$ : state mobilization rate.
- Control group: same characteristics in 1940.

# Baseline results

Impact of WWII mobilization on labor supply of educated white married women

	25–34 years old in 1950		35–44 years old in 1960	
	Weeks worked (1)	LFPR (2)	Weeks worked (3)	LFPR (4)
$\tau \times MOB \times \text{MarrChildWWII}$	23.65 (15.48)	0.90*** (0.3)	42.68** (17.86)	0.68* (0.41)
$\tau \times MOB \times \text{MarrNoChildWWII}$	43.33** (17.14)	0.96*** (0.34)	31.89*** (12.06)	0.64** (0.31)
$\tau \times MOB \times \text{UnmarrWWII}$	37.95* (19.48)	0.79 (0.49)	7.61 (15.11)	0.1 (0.35)
Constant	18.15*** (0.88)	0.39*** (0.02)	13.19*** (1.01)	0.30*** (0.02)
Observations	20,846	20,846	49,009	49,009
$R^2$	0.16	0.14	0.11	0.11

- No effect on less educated and nonwhite women.
- WWII pulled less educated in blue collar occupations, which disappeared after the war.

- Role of husbands' norms for the Revolution of female labor?
- Theoretical mechanism:
  - Growing number of men who grew up with working mother  $\implies$   $\uparrow$  preference toward working wife and/or  $\uparrow$  productivity in household production.
  - Working becomes more valuable trait in the marriage market  $\implies$  women invest in HK and have higher wage draws when adult.

- Role of husbands' norms for the Revolution of female labor?
- Theoretical mechanism:
  - Growing number of men who grew up with working mother  $\implies$   $\uparrow$  preference toward working wife and/or  $\uparrow$  productivity in household production.
  - Working becomes more valuable trait in the marriage market  $\implies$  women invest in HK and have higher wage draws when adult.
- Use WWII in U.S. as source of exogenous variation in labor supply of mothers-in-laws of 1930–1935 cohorts (too young to be directly affected by the war)

- Wartime mobilization: measure from Acemoglu et al (2004).
- Data: 1% decennial censuses 1940–1980.
- Sample: white married women outside farming and group quarters.
- Measure of female labor: weeks worked.



$$w_{ist} = X'_{ist}\beta_1 + \gamma_t \tilde{X}'_{ist}\beta_{2t} + \gamma_t e'_s\beta_{3t} + \alpha_t \gamma_t m_s + d_s + \gamma_t + \varepsilon_{ist}$$

- $w_{ist}$ : weeks worked by woman  $i$  born in state  $s$  at time  $t$ .
- $X'_{ist}$ : individual controls (age, state of residence, husband state of birth indicators).
- $\tilde{X}'_{ist}$ : age indicators interacted with year FE  $\gamma_t$  effects after 1940.
- $d_s$ : state FE.
- $e'_s$ : state characteristics interacted with year indicator after 1940.

# Baseline results

Impact of WWII on labor supply of 42–47 married women

Dependent variable is “Weeks Worked”	
	1940 & 1950
1940 mobilization rate $\times$ year	19.64** (8.71)
1940 share male nonwhite $\times$ year	-2.37 (4.66)
1940 share male farmer $\times$ year	6.12*** (1.79)
1940 male avg years educ $\times$ year	.218 (.483)
Year	-6.80 (6.01)
N. obs.	32,347
Adjusted $R^2$	0.03

- Mothers (1903–1908) of cohort of interest (1930–1935).
- 42–47 in 1950 (keep same group in 1940)

## Impact of WWII on labor supply of married women

Dependent variable is "Weeks Worked"				
Panel A: 25-30				
	(i)	(ii)	(iii)	(iv)
1940 mobilization rate $\times$ 1950	18.11 (11.05)	17.29 (10.99)	21.68 (14.50)	22.59 (14.24)
1940 mobilization rate $\times$ 1960	22.71** (11.30)	19.06* (11.12)	26.68* (15.1)	26.39* (15.15)
Year 1950	-11.23* (6.26)	-11.49* (6.16)	-19.44** (7.51)	-17.06** (8.19)
Year 1960	-5.58 (5.89)	-3.24 (5.72)	-11.73* (7.02)	-14.54* (7.47)
St. of residence & husband's st. of birth		yes	yes	yes
Education				yes
N. obs.	75,748	73,710	50,146	50,146
Adjusted $R^2$	0.01	0.015	0.016	0.027

## Impact of WWII on labor supply of married women

Panel B: 35–40				
	(i)	(ii)	(iii)	(iv)
1940 mobilization rate × 1950	25.25*** (8.36)	23.67*** (7.92)	33.02*** (9.99)	34.49*** (10.28)
1940 mobilization rate × 1960	18.34*** (6.76)	18.17*** (6.29)	22.55*** (7.89)	24.74*** (8.22)
1940 mobilization rate × 1970	14.24* (8.07)	14.78** (7.51)	22.01** (8.74)	25.12*** (8.64)
Year 1950	-2.25 (5.87)	-1.12 (5.63)	-11.88 (7.79)	-20.55*** (7.90)
Year 1960	-3.76 (4.91)	-3.12 (4.75)	-7.30 (6.73)	-16.73** (6.81)
Year 1970	1.79 (6.02)	.99 (5.65)	-6.19 (7.87)	-7.58 (9.02)
St. of residence & husband's st. of birth		yes	yes	yes
Education				yes
N. obs.	112,125	109,864	71,018	71,018
Adjusted $R^2$	0.039	0.045	0.05	0.05

## Impact of WWII on labor supply of married women

Panel C: 45–50				
	(i)	(ii)	(iii)	(iv)
1940 mobilization rate × 1950	12.59 (11.39)	17.92 (11.59)	27.98* (14.79)	26.17* (14.79)
1940 mobilization rate × 1960	11.47 (9.24)	16.22* (9.29)	15.44 (12.54)	15.33 (12.59)
1940 mobilization rate × 1970	3.25 (8.56)	8.99 (8.75)	13.77 (12.31)	15.96 (12.23)
1940 mobilization rate × 1980	17.26* (10.18)	21.72** (10.76)	32.89** (14.78)	33.07** (14.78)
Year 1950	-14.55 (9.59)	-17.96** (9.44)	-27.31** (13.21)	-26.23** (13.08)
Year 1960	.52 (6.53)	-2.09 (6.44)	-7.95 (10.39)	-11.71 (10.81)
Year 1970	9.50 (6.84)	5.07 (6.79)	1.82 (10.27)	-8.88 (10.59)
Year 1980	8.18 (7.46)	5.01 (7.60)	-1.84 (11.39)	-7.01 (11.80)
St. of residence & husband's st. of birth		yes	yes	yes
Education				yes
N. obs.	129,899	126,715	80,261	80,261
Adjusted $R^2$	0.087	0.091	0.098	0.11

## Impact of WWII on labor supply of married women

Dependent variable is "Weeks Worked"				
	1940 & 1950	1950 & 1960	1960 & 1970	1970 & 1980
1940 mobilization rate × year	16.48** (7.87)	-11.66 (8.46)	-2.10 (6.28)	20.65*** (7.07)
1940 share male nonwhite × year	4.56 (4.91)	-9.00* (4.78)	-1.77 (4.54)	-1.81 (2.43)
1940 share male farmer × year	9.84*** (2.03)	-4.39* (2.42)	2.23 (1.86)	2.96 (2.55)
1940 male avg years educ × year	2.12*** (.436)	-.96** (.476)	-.46 (.576)	-.21 (.225)
Year	-22.35** (6.12)	20.93*** (7.62)	10.01 (6.33)	-4.37 (4.69)
N. obs.	15,955	25,127	45,015	45,037
Adjusted $R^2$	0.035	.029	.015	.017

- Assess the impact of WWII female wartime employment directly.
  - Data on female wartime employment at State Economic Areas level (grouping of counties within states).
  - Data on state-level manpower mobilization (same as AAL 2004 and GO 2013), but also induction records and war deaths aggregated at the SEA level, normalized by the 1940 population aged 21 to 54.
  - Data on industrial mobilization with county-level military contracts spending.

- Assess the impact of WWII female wartime employment directly.
  - Data on female wartime employment at State Economic Areas level (grouping of counties within states).
  - Data on state-level manpower mobilization (same as AAL 2004 and GO 2013), but also induction records and war deaths aggregated at the SEA level, normalized by the 1940 population aged 21 to 54.
  - Data on industrial mobilization with county-level military contracts spending.
- Female employment growth 1940–1944 linked to military contracts growth but unrelated to manpower mobilization.



# Female wartime employment and war contracts

Dependent variable: change in female employment (1940–1944), SEA level

	(1)	(2)	(3)
War contracts per capita	0.50*** (0.04)	0.33*** (0.04)	
Inductions			-0.02 (0.03)
War deaths			
1940 share emp. manufacturing		-0.30 (0.59)	
1940 share pop white		1.44*** (0.35)	
1940 median school women 25+		0.03 (0.03)	
1940 female LFP rate		0.05*** (0.01)	
R2	0.29	0.41	0.00
N	407	407	407

- All variable standardized.

# Female wartime employment and war contracts

Dependent variable: change in female employment (1940–1944), state level

	(1)	(2)	(3)	(4)
War contracts per capita	0.27** (0.08)	0.24* (0.10)		
Inductions			0.15 (0.09)	0.01 (0.14)
War deaths				
1940 share emp. manufacturing		1.59 (1.91)		3.16 (2.16)
1940 share pop white		-0.39 (1.09)		-0.28 (1.17)
1940 median school women 25+		0.11 (0.15)		0.20 (0.18)
1940 female LFP rate		-3.21 (2.05)		-3.06 (2.21)
R2	0.17	0.26	0.06	0.15
N	49	49	49	49

- All variable standardized.

$$Y_{ist} = \alpha_s + \beta_0 d_{1950} + \beta_1 WMC_s d_{1950} + \beta_2 X_{ist} + \beta_3 X_{ist} d_{1950} + \varepsilon_{ist}$$

- $Y_{ist}$ : LFP and hours worked by women  $i$  residing in SEA  $s$  in year  $t$  (1940 or 1950).
- $\alpha_s$ : SEA FE.
- $WMC_s$ : female wartime employment growth 1940–1944.
- $X_{ist}$ : covariates (state of birth FE, age, race, SEA share farmers and nonwhites, average schooling in 1940, interacted with 1950 dummy).
- Sample: census 1940 and 1950, women aged 14 to 64 not working in farming.

## Impact of wartime work on FLFP in 1950

	LFP			Hours Worked		
	(1)	(2)	(3)	(4)	(5)	(6)
<b>A. All Women</b>						
WMC female emp. intensity x 1950	-0.002 (0.002)	-0.001 (0.002)	0.003 (0.002)	-0.080 (0.067)	-0.048 (0.067)	0.032 (0.064)
Share emp. farm x 1950			0.075*** (0.020)			1.868** (0.604)
Share pop white x 1950			0.021 (0.012)			1.524*** (0.385)
Med. sch. women x 1950			0.006** (0.002)			0.101* (0.046)
Birth place, age, marital dummies	No	Yes	Yes	No	Yes	Yes
R2	0.03	0.25	0.25	0.02	0.20	0.20
N	1,031,743	1,031,743	1,031,743	1,031,743	1,031,743	1,031,743
<b>B. White Women</b>						
WMC female emp. intensity x 1950	-0.003 (0.002)	-0.002 (0.002)	0.004* (0.002)	-0.165* (0.079)	-0.150 (0.077)	0.045 (0.071)
Share emp. farm x 1950			0.081*** (0.023)			2.068** (0.682)
Share pop white x 1950			-0.042** (0.013)			-2.473*** (0.473)
Med. sch. women x 1950			0.006** (0.002)			0.121* (0.048)
Birth place, age, marital dummies	No	Yes	Yes	No	Yes	Yes
R2	0.03	0.26	0.26	0.02	0.21	0.21
N	919,182	919,182	919,182	919,182	919,182	919,182

- $\uparrow$  1sd wartime employment  $\implies$   $\uparrow$  LFP 0.4pp for white women.

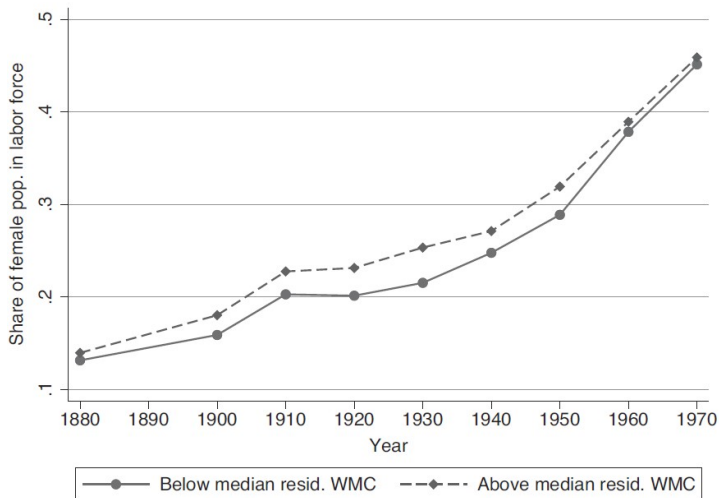
## SEA-level impacts by industry

	All Women		White Women	
	(1)	(2)	(3)	(4)
	LFP	Hours Worked	LFP	Hours Worked
Durable manufacturing	0.003*** (0.001)	0.103*** (0.027)	0.003*** (0.001)	0.117*** (0.030)
Non-durables	-0.002 (0.001)	-0.103* (0.043)	-0.003 (0.001)	-0.118* (0.047)
Transport, telecom, utilities	0.000 (0.000)	0.010 (0.013)	0.000 (0.000)	0.007 (0.014)
Wholesale, retail trade	0.001 (0.001)	0.011 (0.037)	0.001 (0.001)	-0.004 (0.038)
Finance and business	0.001 (0.000)	0.018 (0.014)	0.001 (0.000)	0.018 (0.016)
Personal services	-0.002 (0.001)	-0.087* (0.038)	-0.001 (0.001)	-0.063 (0.035)
Government	0.001* (0.000)	0.024 (0.015)	0.001* (0.000)	0.026 (0.016)

- Driven by durable manufacturing, partly offset by decline in non-durables.

# Baseline results

## Pre-trends in effects of WMC employment on LFP



# No effect of wartime mobilization or deaths

## Impact of inductions and war death on FLFP in 1950

	LFP				Hours Worked			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<b>A. All Women</b>								
Inductions x 1950	0.002 (0.005)	0.001 (0.003)			0.144 (0.111)	0.039 (0.116)		
War deaths x 1950			0.007* (0.003)	0.001 (0.003)			0.282*** (0.081)	0.119 (0.096)
Birth place, age, marital dummies	No	Yes	No	Yes	No	Yes	No	Yes
1940 controls	No	Yes	No	Yes	No	Yes	No	Yes
R2	0.02	0.25	0.03	0.25	0.02	0.20	0.02	0.20
N	671,921	671,921	1,031,743	1,031,743	671,921	671,921	1,031,743	1,031,743
<b>B. White Women</b>								
Inductions x 1950	0.004 (0.006)	0.002 (0.003)			0.250 (0.152)	0.070 (0.134)		
War deaths x 1950			0.005 (0.004)	0.001 (0.004)			0.161 (0.123)	0.151 (0.101)
Birth place, age, marital dummies	No	Yes	No	Yes	No	Yes	No	Yes
1940 controls	No	Yes	No	Yes	No	Yes	No	Yes
R2	0.02	0.26	0.03	0.26	0.02	0.21	0.02	0.21
N	593,076	593,076	919,182	919,182	593,076	593,076	919,182	919,182

# State-level analysis (comparison with literature)

	Weeks Worked		Weeks Worked >0		LFP		Hours Worked	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mobilization x 1950	14.997*** (1.716)	6.104** (2.137)	0.107 (0.064)	0.056 (0.072)	-0.034 (0.045)	-0.022 (0.056)	3.556* (1.633)	10.440*** (2.062)
Share emp. farm x 1950		0.223 (1.094)		0.088** (0.031)		0.116*** (0.026)		5.134*** (0.921)
Share pop white x 1950		2.367* (1.147)		0.011 (0.033)		0.034 (0.022)		2.165** (0.819)
Med. sch. women x 1950		0.406** (0.128)		0.020*** (0.003)		0.013*** (0.002)		-0.059 (0.100)
Birth place, age, marital dummies	No	Yes	No	Yes	No	Yes	No	Yes
R2	0.01	0.19	0.02	0.19	0.01	0.24	0.01	0.19
N	593,944	593,944	593,944	593,944	593,944	593,944	593,944	593,944

- Changing definition of weeks worked: 1940 in full-time equivalent, not in 1950. Inflates part-time workers, with mobilization correlated with 1940 part-time.
- Since mobilization unrelated to female wartime employment, difficult to interpret.



# Female Labor and WWII in the U.S.

- Mixed evidence: what does mobilization capture?
- Using appropriate measurement, little direct effect.
- Indirect implications hold when appropriate measure?

# Female labor and WWI in France

# Female labor and WWI in France

- Boehnke and Gay (2022): effect of WWI during interwar period.
- Gay (2023): effect of WWI in post-WWII period.

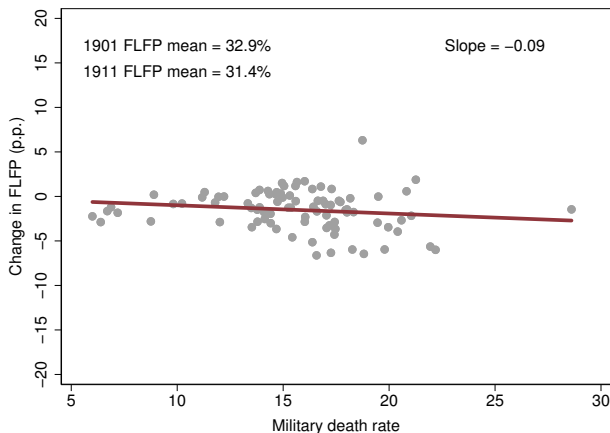
# WWI in France: a shock to the sex ratio



Adult Sex Ratio (1900–2012)

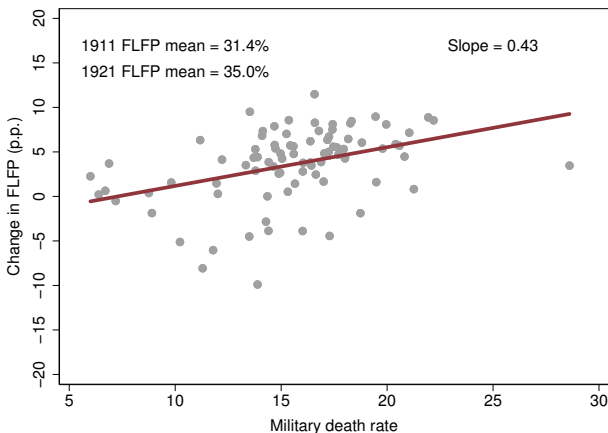
- 1.3 million French fatalities  $\implies$  16% of men aged 15–50 in 1914.

# Impact of WWI in France on FLFP?



WWI Military Death Rates and Changes in FLFP (1901–1911)

# Impact of WWI in France on FLFP?

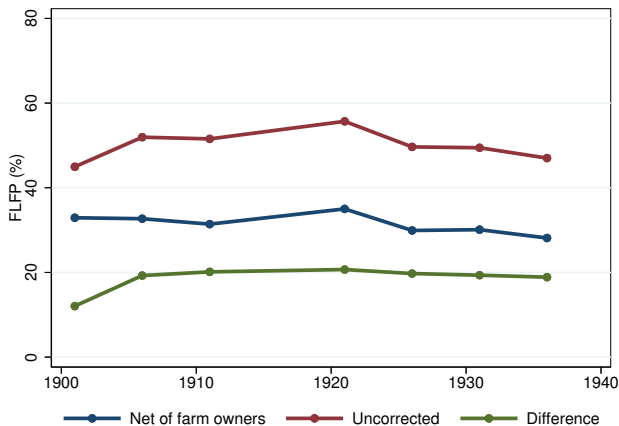


WWI Military Death Rates and Changes in FLFP (1911–1921)

# Measuring FLFP (1901–1936)

- FLFP at département level: censuses 1901–1936.
- FLFP: share of women aged 15+ in labor force.
- Focus on paid employment:
  - Farmer's wives to be classified as labor force participant.
  - Not consistent in 1901. Maruani and Meron (2012)
  - Subtract female farm owners for consistency.

# Measuring FLFP (1901–1936)



Average Female Labor Force Participation Rates (%)



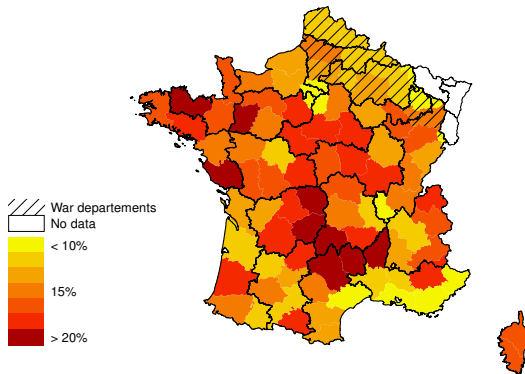
# WWI military fatalities: data

- Original dataset:
  - Individual military records for all 1.3 million deceased French soldiers.
  - Information: département of birth, dates of birth and death.
- Measure of military death rates at the département level:

$$\text{death\_rate}_d = \frac{\text{dead soldiers born in } d}{\underbrace{\text{male pop. aged 15–44 in } d \text{ in 1911}}_{\text{Draft pool in census 1911}}}$$

- Summary statistics:
  - Mean = 15%, 25th percentile  $\simeq$  10%, 75th percentile  $\simeq$  20%.
  - Minimum = 6%, maximum = 29%.

# WWI military fatalities: distribution



Military Death Rates Across 87 Départements

# WWI military fatalities: data generating process

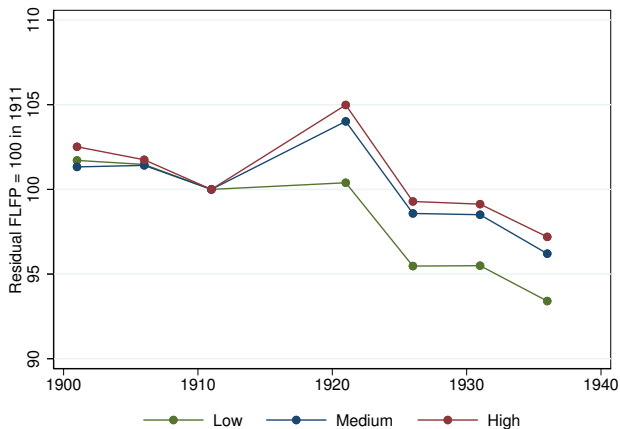
- Until end 1914: territorial organization of military recruitment.
- Policies by the Ministry of War to support the industrial war effort:  
⇒ Up to 15% of soldiers allocated to war industries.
- Pre-war characteristics:
  - Correlated with measures of rurality.
  - Uncorrelated with other pre-war characteristics.
  - Uncorrelated with pre-war levels and trends in FLFP.

# Military fatalities and pre-war characteristics

Dependent variable:	Military death rate		
	(1)	(2)	(3)
FLFP	0.03 [0.03]	-0.06 [0.07]	-0.01 [0.07]
Rural	0.12*** [0.01]	0.11*** [0.03]	0.16*** [0.04]
Born in dép.	0.13*** [0.03]	0.15*** [0.04]	0.11* [0.06]
Other characteristics	No	Yes	Yes
Military region FE	No	No	Yes
Départements	87	87	87
Adj. R <sup>2</sup>	0.739	0.790	0.838

Other characteristics: fertility, wealth, female education, population density, age, height, share in industry, roads and rails per km<sup>2</sup>, share cultivated land, direct taxes, banking deposits, distance to war, share religious schools, turnout, distance to war, share literate.

# Military fatalities and pre-war trends in FLFP



# Baseline results

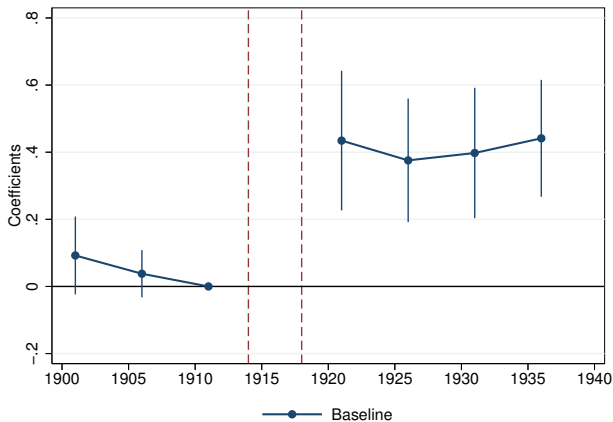
$$\text{FLFP}_{d,t} = \beta \text{ death\_rate}_d \times \text{post}_t + \gamma_d + \delta_t + \varepsilon_{d,t}$$

Dependent variable:	FLFP	
	(1)	(2)
Death rate $\times$ post	0.37*** [0.08]	0.35*** [0.07]
Rural	No	Yes
Born in dép.	No	Yes
Départements	87	87
1911 FLFP mean	31.4	31.4

- 20% vs 10% military death rates  $\implies$  FLFP  $\uparrow$  4 p.p.
- 4 p.p.  $\simeq$  12% of pre-war FLFP mean (0.4 s.d.).

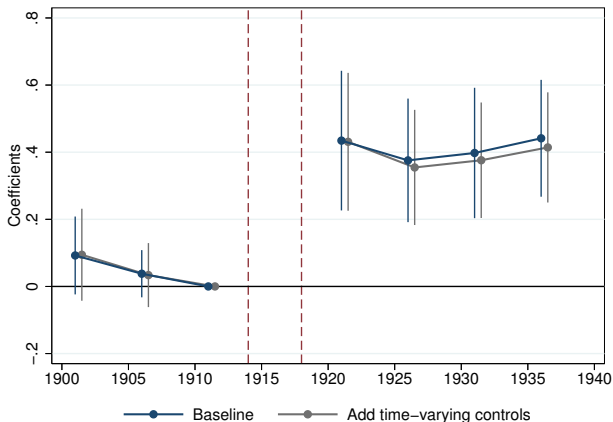
# Baseline results

$$\text{FLFP}_{d,t} = \sum_{t \neq 1911} \beta_t \text{death\_rate}_d \times \text{year}_t + \gamma_d + \delta_t + \varepsilon_{d,t}$$



# Baseline results

$$\text{FLFP}_{d,t} = \sum_{t \neq 1911} \beta_t \text{death\_rate}_d \times \text{year}_t + \gamma_d + \delta_t + \varepsilon_{d,t}$$





# Potential mechanisms

- Supply channel through post-war marriage market.
- Demand channel through substitution across inputs.
- Persistence of female wartime employment.

# Supply factors: a marriage market channel?

- Demographic changes of female marital statuses:
    - Increased share of single women.
    - Increased share of widowed (and divorced) women.
- ⇒ Non-married women structurally supply more labor.

# Supply factors: a marriage market channel?

- Demographic changes of female marital statuses:
  - Increased share of single women.
  - Increased share of widowed (and divorced) women.
- ⇒ Non-married women structurally supply more labor.
- Increased relative labor supply due to negative income shocks:
  - War widows (very low subsidies until 1930s).
  - Young single women (if deceased soldier in family).
  - Married women.
    - Lower quality husbands. [Abramitky et al. \(2011\)](#)
    - Lower bargaining power in household. [Grossbard \(2014\)](#)

# The post-war marriage market

$$Y_{m,a,d,t} = \beta \text{ death\_rate}_d \times \text{post}_t + \gamma_d + \delta_t + \varepsilon_{m,a,d,t}$$

Dependent variable:	Share Widowed		
	20–29	30–39	40–49
Age group:	(1)	(2)	(3)
Death rate $\times$ post	–0.00 [0.01]	0.09*** [0.03]	0.17*** [0.03]
Départements	87	87	87
1911 outcome mean	1.3	4.7	11.5

Controls: rurality, year and département FE.

- No differential pre-war trends.

# The post-war marriage market

$$Y_{m,a,d,t} = \beta \text{ death\_rate}_d \times \text{post}_t + \gamma_d + \delta_t + \varepsilon_{m,a,d,t}$$

Dependent variable:	Share Single		
	20–29	30–39	40–49
Age group:	(1)	(2)	(3)
Death rate $\times$ post	0.27*** [0.07]	0.23*** [0.05]	0.14*** [0.04]
Départements	87	87	87
1911 outcome mean	38.9	14.6	11.2

Controls: rurality, year and département FE.

- No differential pre-war trends.

# FLFP by age

$$\text{FLFP}_{a,d,t} = \beta \text{ death\_rate}_d \times \text{post}_t + \gamma_d + \delta_t + \varepsilon_{a,d,t}$$

Dependent variable:	FLFP		
	20–29	30–39	40–49
Age group:	(1)	(2)	(3)
Death rate $\times$ post	0.27*** [0.09]	0.48*** [0.10]	0.47*** [0.10]
Départements	87	87	87
1911 outcome mean	34.7	30.1	27.7

Controls: rurality, year and département FE.

# FLFP of widows

$$\text{FLFP}_{w,d,t} = \beta \text{ death\_rate}_d \times \text{post}_t + \gamma_d + \delta_t + \varepsilon_{a,d,t}$$

Dependent variable:	FLFP		
Definition:	Active women / all women	Active widows / all women	Active widows / widows
	(1)	(2)	(3)
Death rate $\times$ post	0.26*** [0.07]	0.12*** [0.03]	0.54*** [0.11]
Départements	87	87	87
1911 outcome mean	32.9	5.7	34.5

Controls: rurality, year and département FE.

# Demand factors: a substitution channel?

- Increase in FLFP compatible with (large) increase in demand for female labor?
- Changes in female wages across occupations with different substitutability relative to male labor.
  - Textile manufacturing sector: low substitutability.
  - Domestic services sector: high(er) substitutability.



# The textile manufacturing sector

- Nearly exclusively women.
- Occupations: ironers, seamstresses, milliners.
- No other female occupations available before the war.
- City-level hourly wage rates 1901–1926.
- Enables to fix demand curve across départements with  $\neq$  military death rates.

# The textile manufacturing sector

$$\log w_{o,d,t} = \beta \text{ death\_rate}_d \times \text{post}_t + \gamma_d + \delta_t + \varepsilon_{o,d,t}$$

Dependent variable:	Log Hourly Wage Rate		
	Ironer	Seamstress	Milliner
Occupation:	(1)	(2)	(3)
Death rate $\times$ post	-0.010*** [0.004]	-0.012*** [0.003]	-0.008** [0.003]
Départements	87	87	87
1911 outcome mean (francs)	0.21	0.23	0.25

Controls: rurality, year and département FE.

- No differential pre-war trends.

# The domestic services sector

- Balanced gender composition.
- Occupations: cooks and housekeepers.
- City-level yearly wage rates 1913 and 1921.
- Upper bound for changes in demand curve.

# The domestic services sector

$$\log w_{o,d,t} = \beta \text{death\_rate}_d \times \text{post}_t + \gamma_d + \delta_t + \varepsilon_{o,d,t}$$

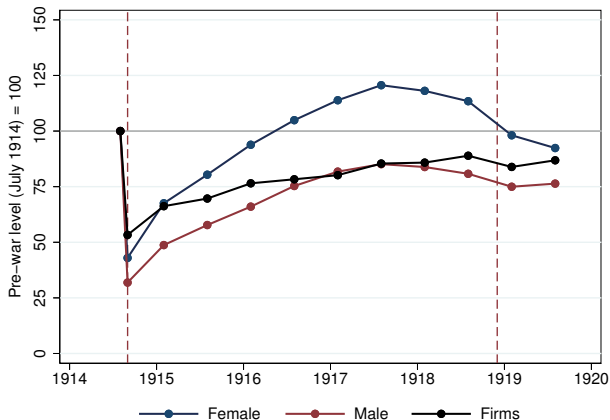
Dependent variable:	Log Hourly Wage Rate	
	Cook	Housekeeper
Occupation:	(1)	(2)
Death rate $\times$ post	-0.006*** [0.002]	-0.002* [0.001]
Départements	87	87
1911 outcome mean (francs)	0.20	0.16

Controls: rurality, year and département FE.

- No differential pre-war trends.

# Female wartime employment

## Wartime Employment in Industrial Sector



# Female wartime employment

- Women who entered labor force could have remained, improved skills, information. . .
- Measure relative intensity of female wartime employment as:

$$\frac{L_{\text{July 1917}} - L_{\text{July 1914}}}{L_{\text{July 1914}}}$$

- Not correlated with military death rates.
- Same result if use July 1918.

# Female wartime employment

Dependent variable:	FLFP		
	(1)	(2)	(3)
Death rate $\times$ post	0.35*** [0.07]		0.35*** [0.07]
Wartime employment $\times$ post		-0.01 [0.01]	-0.00 [0.01]
Départements	87	87	87

Controls: rurality, year and département FE.

# Female labor and WWI in the France

- Limited data to explore mechanisms: need micro-level data  
Gay and Kesztenbaum 2022.
- Long-run consequences and intergenerational transmission.  
Gay (2023)



# Female labor and wars

- Historically unclear how marriage and labor markets interact.
- Relevance of female labor *during vs after* wars?
- Other ambiguous consequences: social and political (family norm, social movements, voting)?
- Heavily dependent upon historical context.

# Topic 3

## Economics history and innovation

# Economic History and Innovation

- Focus on roots of scientific innovation.

# Economic History and Innovation

- Focus on roots of scientific innovation.
- Three articles using same historical natural experiment.
  - What are the relative roles of human and physical capital in production of scientific knowledge? [Waldinger \(2016\)](#)
  - How important are localized peer effects for innovation? [Waldinger \(2012\)](#)
  - What are the roles of field-specific knowledge spillovers for innovation? [Moser, Voena, and Waldinger \(2014\)](#)

# Economic History and Innovation

- Focus on roots of scientific innovation.
- Three articles using same historical natural experiment.
  - What are the relative roles of human and physical capital in production of scientific knowledge? [Waldinger \(2016\)](#)
  - How important are localized peer effects for innovation? [Waldinger \(2012\)](#)
  - What are the roles of field-specific knowledge spillovers for innovation? [Moser, Voena, and Waldinger \(2014\)](#)
- Other studies using same historical natural experiment:
  - Effect of faculty quality on PhD student outcomes in mathematics. [Waldinger \(2010\)](#)
  - Role of social ties within academic networks in facilitating escape of academics from Nazi Germany. [Becker et al. \(2021\)](#)

# Bombs, Brains, and Science

- [Waldinger \(2016\)](#), “Bombs, Brains, and Science: The Role of Human and Physical Capital for the Creation of Scientific Knowledge.”
- General research question: what are the inputs of innovation?

# Bombs, Brains, and Science

- [Waldinger \(2016\)](#), “Bombs, Brains, and Science: The Role of Human and Physical Capital for the Creation of Scientific Knowledge.”
- General research question: what are the inputs of innovation?
- Narrower research question: what are the relative roles of human and physical capital in the production of scientific knowledge?
- Focus on academic scientific knowledge.

# Bombs, Brains, and Science

- Challenging to estimate:
  - Selection of human capital (scientists) into more productive departments.
  - Unobserved determinants of physical capital (labs) productivity.
  - Complex complementarities between physical and human capital
  - Measurement is challenging.



# Bombs, Brains, and Science

- Challenging to estimate:
  - Selection of human capital (scientists) into more productive departments.
  - Unobserved determinants of physical capital (labs) productivity.
  - Complex complementarities between physical and human capital
  - Measurement is challenging.
- Overcome challenges by historical natural experiments:
  - Exogenous variation in human capital: dismissal of (mostly Jewish) scientists in Nazi Germany 1933–1940.
  - Exogenous variation in physical capital: destruction of science facilities during WWII Allied bombings.

# Human Capital Shock

- Dismissal of scientists in Nazi Germany.
  - April 1933: National Socialist party passes the “Law for the Restoration of the Professional Civil Service.”
  - Jewish and “politically unreliable” persons dismissed from civil service positions.
  - Anybody with at least one Jewish grandparent dismissed.
  - Communist party members dismissed.
  - Law extended to Austrian universities in March 1938.

# Human Capital Shock

- Dismissal of scientists in Nazi Germany.
  - April 1933: National Socialist party passes the “Law for the Restoration of the Professional Civil Service.”
  - Jewish and “politically unreliable” persons dismissed from civil service positions.
  - Anybody with at least one Jewish grandparent dismissed.
  - Communist party members dismissed.
  - Law extended to Austrian universities in March 1938.
- More than 1,000 academics dismissed from German universities (up to 60% of faculty).
- Most emigrated and got positions at foreign universities (U.S., U.K., Turkey, Switzerland).

# Human Capital Shock

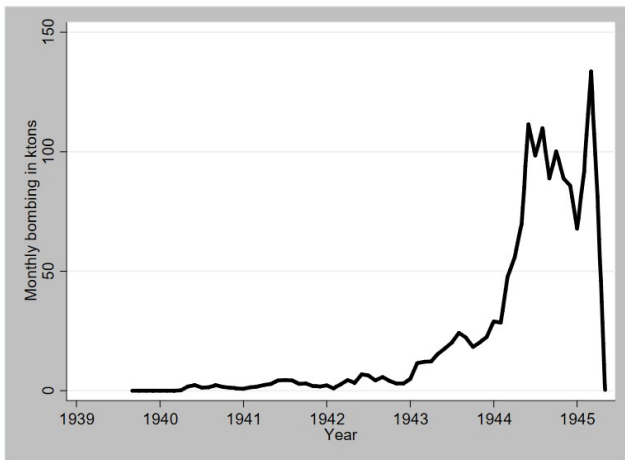
- Physicists:
  - 15% of dismissed.
  - Produced 24% of top journal articles pre-1933.
  - Received 64% of citations to articles pre-1933.
- Chemists:
  - 14% of dismissed.
  - Produced 22% of top journal articles pre-1933.
  - Received 23% of citations to articles pre-1933.
- Mathematicians:
  - 19% of dismissed.
  - Produced 31% of top journal articles pre-1933.
  - Received 61% of citations to articles pre-1933.

# Physical Capital Shock

- Steps in Allied Bombings:
  - From 1939: UK RAF targeted raids on German fleet.
  - From 1940: RAF bombings on other targets (reservoirs, factories, ...)
  - From 1941: RAF bombings on German cities (but small accuracy).
  - From 1942: RAF bombings strongly intensified.
  - From 1943: US AAF bombings on German targets.
  - From 1944: intensification of bombings by US AAF.
- Overall 1.35 million tons of bombs.
- Bombings destroyed 19% of homes in (later) Federal Republic of Germany.
- Scientific departments never targeted, and very inaccurate bombings within cities.

# Physical Capital Shock

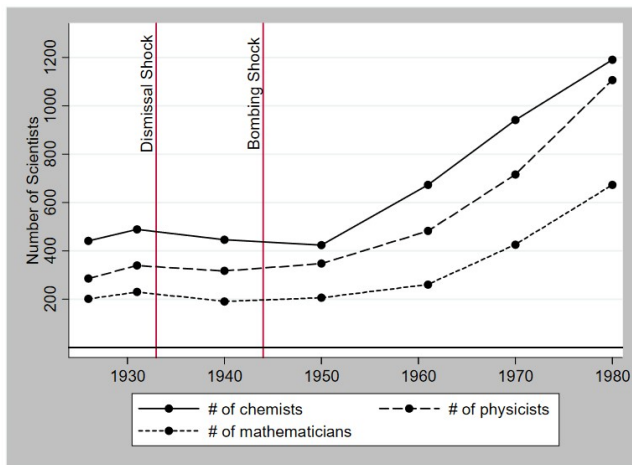
Allied Bombs Dropped over German Territory



Source: Waldinger (2016) Figure A2

- Focuses on physicists, chemists, mathematicians in German and Austrian universities.
- Information from registers of all German university professors (1926, 1931, 1940, 1950, 1961, 1970, 1980).
- Complete faculty roster of science departments 1926–1980.
- Select all 35 German or Austrian universities than existed in 1926 (105 science departments).
- Database of 5,716 scientists, 10,387 person-year observations.

## Number of Scientists in 1926 German Universities



Source: Waldinger (2016) Figure A3



- Scientists' output:
  - Publications and citation-weighted publications for each scientist and year.
  - Most cited mathematician: Johann von Neumann.
- Departments' output:
  - Add individual output within departments.
  - Because not yearly data, aggregate within 5-year window.
  - Standardize total department output in each subject.

# Dismissals, Bombings, and Other Data

- Dismissals:
  - List of Displaced German Scholars (pre-1938 dismissals, no Austria).
  - Other sources  $\implies$  individual data on 90% of all dismissals

# Dismissals, Bombings, and Other Data

- Dismissals:
  - List of Displaced German Scholars (pre-1938 dismissals, no Austria).
  - Other sources  $\implies$  individual data on 90% of all dismissals
- Bombings:
  - University archives of funding requests to fund reconstruction.
  - Damage of each building in percent (in best case) or verbal description (worse case).

# Dismissals, Bombings, and Other Data

- Dismissals:
  - List of Displaced German Scholars (pre-1938 dismissals, no Austria).
  - Other sources  $\implies$  individual data on 90% of all dismissals
- Bombings:
  - University archives of funding requests to fund reconstruction.
  - Damage of each building in percent (in best case) or verbal description (worse case).
- Control variables: university age, share of firms in armament-related industries in 1933, fraction of Jews in city in 1933, distance to iron curtain.

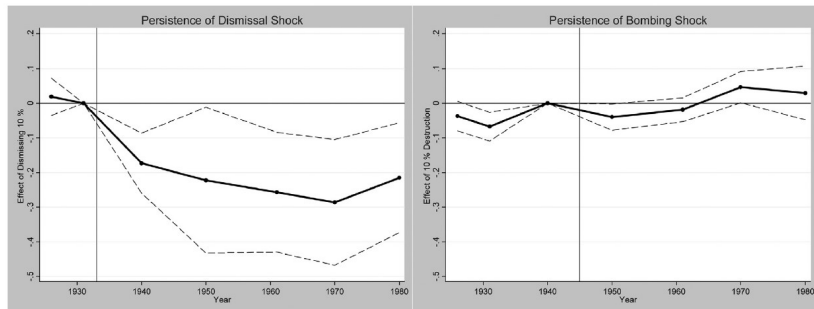
# Empirical Strategy

$$\begin{aligned}\text{Output}_{dt} &= \beta_1 + \sum_{t \neq 1931} \beta_{2t} \text{HCSHock}_{1933-40,d} \times \text{Year}_t \\ &+ \sum_{t \neq 1940} \beta_{3t} \text{PCSHock}_{1942-45,d} \times \text{Year}_t \\ &+ \delta_d + \gamma_t + \beta_4 X_{dt} + \varepsilon_{dt}\end{aligned}$$

- $\text{Output}_{dt}$ : department  $d$  research output in year  $t$  (publications or citation-weighted publications).
- $\text{HCSHock}_{1933-40,d}$ : dismissal of scientists in department  $d$  1933–40.
- $\text{PCSHock}_{1942-45,d}$ : Allied bombings of department  $d$  1942–45.
- $\delta_d$ : 105 department FE.
- $\gamma_t$ : year FE.

# Main Results

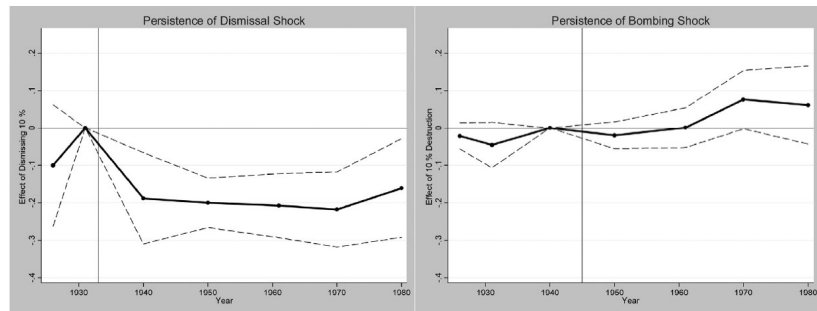
## Persistence of 10% Shocks: Publications



Source: Waldinger (2016), Figure 1. 10% human capital shock is 1.015 scientists. Coefficients from specification  $\text{subject} \times \text{year}$  FE,  $\text{occupation zone} \times \text{Post 1945}$ , and  $\text{share city destruction} \times \text{year}$  FE.

# Main Results

## Persistence of 10% Shocks: Citation-Weighted Publications



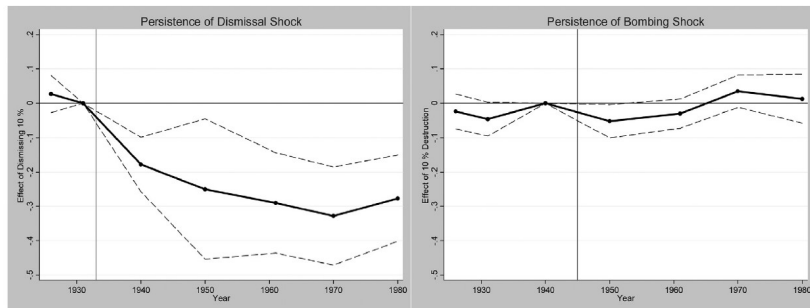
Source: Waldinger (2016), Figure 2. 10% human capital shock is 1.015 scientists. Coefficients from specification  $\text{subject} \times \text{year}$  FE,  $\text{occupation zone} \times \text{Post 1945}$ , and  $\text{share city destruction} \times \text{year}$  FE.

# Robustness

- Subject  $\times$  year FE.
- City-level destruction  $\times$  year FE.
- Occupation zone  $\times$  post 1945 indicator.
- Further controls: Federal state indicator, university age and age squared, number of departments within 50 km, share of firms in armament  $\times$  year FE, fraction of Jews in city  $\times$  post 1945, distance to iron curtain.
- Alternative samples: drop Austrian, drop East Germany, Swiss as control.
- Alternative measurements of shocks.
- Little complementarities (triple interaction).

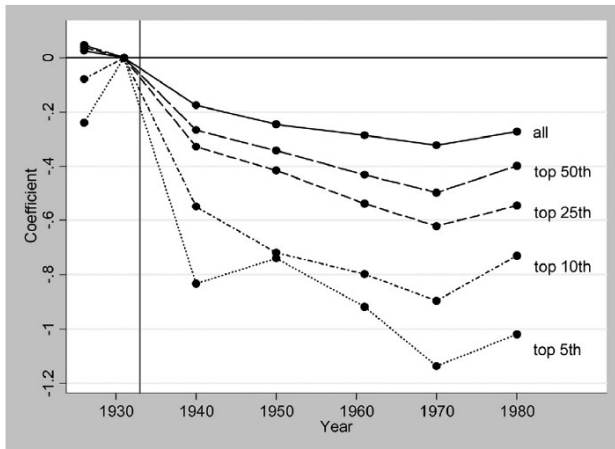


## Persistence of 10% Shocks: Publications with All Controls



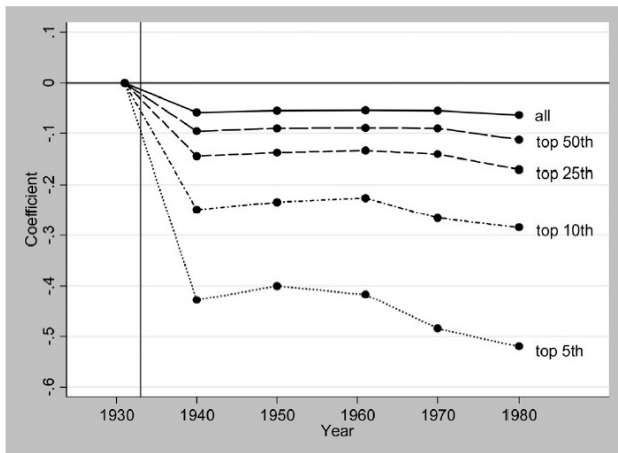
Source: Waldinger (2016), Figure 3. 10% human capital shock is 1.015 scientists. Coefficients from specification with all controls

## Persistence of High-Quality Dismissals



Source: Waldinger (2016), Figure 4.

## Dismissal and the Quality of Hires



Source: Waldinger (2016), Figure 5.

- [Waldinger \(2012\)](#), “Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany.”
- General research question: how important are localized peer effects for innovation?

- [Waldinger \(2012\)](#), “Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany.”
- General research question: how important are localized peer effects for innovation?
- Narrower research question: focus on academic scientific knowledge.

- Challenging to estimate:
  - Positive sorting of scientists in same university  $\implies$  positive correlation of productivity within university not caused by peer effects.
  - Unobservables affecting productivity of scientist and their peers.
  - Measurement challenges.

- Challenging to estimate:
  - Positive sorting of scientists in same university  $\implies$  positive correlation of productivity within university not caused by peer effects.
  - Unobservables affecting productivity of scientist and their peers.
  - Measurement challenges.
- Overcome challenges by historical natural experiment:
  - Exogenous variation in peer group of researchers in Germany: dismissal of scientists by Nazi government in 1933.

# Dismissal of Scientists

- Dismissal of scientists in Nazi Germany.
  - April 1933: National Socialist party passes the “Law for the Restoration of the Professional Civil Service.”
  - Jewish and “politically unreliable” persons dismissed from civil service positions.



# Dismissal of Scientists

- Dismissal of scientists in Nazi Germany.
  - April 1933: National Socialist party passes the “Law for the Restoration of the Professional Civil Service.”
  - Jewish and “politically unreliable” persons dismissed from civil service positions.
- More than 1,000 academics dismissed from German universities (15% of all 7,266 university researchers).
- Most emigrated right away and got positions at foreign universities (U.S., U.K., Turkey, Switzerland).
- Most dismissals in 1933–1934.

# Dismissal of Scientists

## Quality of Dismissed Scholars

	Physics				Chemistry				Mathematics			
	All	Stayers	Dismissed 1933–1934		All	Stayers	Dismissed 1933–1934		All	Stayers	Dismissed 1933–1934	
			No.	% Loss			No.	% Loss			No.	% Loss
Researchers (beginning of 1933)	287	248	39	13.6	466	405	61	13.1	224	183	41	18.3
Researchers (beginning of 1933)	287	248	39	13.6	466	405	61	13.1	224	183	41	18.3
No. of chaired professors	109	97	12	11.0	156	136	20	12.8	117	99	18	15.4
Average age (1933)	49.5	50.2	45.1	—	50.4	50.5	49.7	—	48.7	50	43	—
No. of Nobel Laureates	15	9	6	40.0	14	11	3	21.4	—	—	—	—
Publications 1925–1932												
Average publications	0.47	0.43	0.71	20.5	1.69	1.59	2.31	17.9	0.33	0.27	0.56	31.1
Average publications (citation weighted)	5.10	3.53	14.79	39.4	17.25	16.07	25.05	19.0	1.45	0.93	3.71	46.8
% co-authored	32.0	32.1	31.4	—	75.2	74.8	76.9	—	16.9	15.1	20.3	—
% co-authored with faculty	11.1	10.3	14.5	—	11.8	12.3	9.7	—	9.9	9.7	10.2	—
(with dismissed)	(3.1)	(2.0)	(8.1)	—	(1.9)	(1.9)	(2.0)	—	(4.6)	(3.8)	(6.1)	—
% co-authored with faculty (same uni)	3.7	2.9	7.4	—	4.3	4.4	4.1	—	2.6	1.8	4.3	—
(with dismissed)	(1.5)	(0.5)	(5.9)	—	(0.9)	(0.9)	(1.1)	—	(1.7)	(1.2)	(2.7)	—
Publications 1935–1938												
Average publications		0.35	0.32		1.24	0.55			0.20	0.15		
Average publications (citation weighted)		2.53	11.12		13.61	5.28			0.80	0.57		
% co-authored		43.0	50.0		81.0	69.6			14.9	28.0		
% co-authored with faculty		6.9	7.0		3.9	2.0			6.0	4.0		
(with dismissed)		(0.6)	(4.0)		(0.4)	(2.0)			(0.0)	(4.0)		
% co-authored with faculty (same uni)		2.6	—		0.9	—			0	—		
(with dismissed)		(0.0)	—		(0.1)	—			(0.0)	—		

Source: Waldinger (2012), Table 3.

- Focuses on scientists (physicists, chemists, mathematicians) at German universities.
- Information from registers of all German university scientists (annual 1925–1938).
- Select rank of at least lecturer.
- Information on specialization of all scientists.

- Collect publications of each researcher in academic journals.
- For each researcher, two yearly productivity measures:
  - Sum of publications in top journals in a given year.
  - Citation-weighted publications: sum of citations to all articles published in a given year in any journal included in Web of Science in first 50 years after publication.

# Measuring Peer Effects

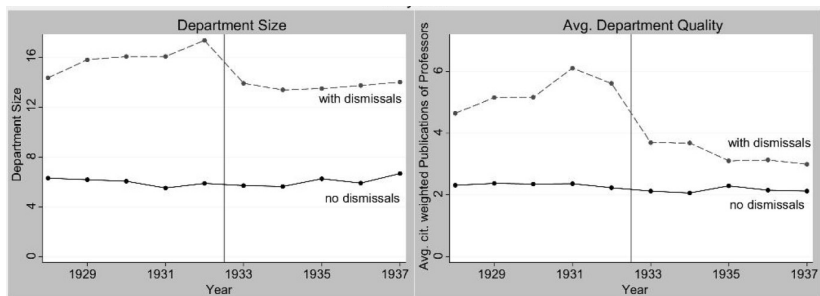
- Different types of peer effects:
  - Joint research projects involving joint publication.
  - Discussions and collaborations without co-publishing.
  - Peer pressure affecting productivity.
  - Peer funding and networks affecting productivity.
- Empirical approach differentiates two dimensions of peer effects:
  - Quality of peer effects measured by mean average productivity of peers (researchers of same specialty in same department).
  - Number of peer interactions measured by number of peers at department.
- Use one-year lags to account for publication lags.

$$\begin{aligned} \# \text{Publications}_{iut} &= \beta_1 + \beta_2 (\text{Peer Quality})_{ut} + \beta_3 (\# \text{of Peers})_{ut} \\ &+ \beta_4 \text{Age FE}_{iut} + \beta_5 \text{Year FE}_t + \beta_6 \text{Ind. FE}_i + \varepsilon_{iut} \end{aligned}$$

- Scientist  $i$  at university  $u$  in year  $t$ .
- Instrument peer group size and quality by dismissals.

# Empirical Strategy

## Effects of Dismissals on Department Size and Peer Quality (Physics)



Source: Waldinger (2012), Figure 1.

# Empirical Strategy

- Two first-stage regressions.
- One to instrument peer quality, with “Dismissal-Induced Fall in Peer quality”: difference average peer quality before 1933 – average peer quality before 1933 among stayers.
- First first-stage regression:

$$\begin{aligned} \text{Peer Quality}_{ut} &= \gamma_1 + \gamma_2 (\text{Dismissal-Induced } \downarrow \text{ in Peer Quality})_{ut} \\ &+ \gamma_3 (\# \text{Dismissed})_{ut} \\ &+ \gamma_4 \text{Age FE}_{iut} + \gamma_5 \text{Year FE}_t + \gamma_6 \text{Ind. FE}_i + \varepsilon_{iut} \end{aligned}$$



- Second first-stage regression:

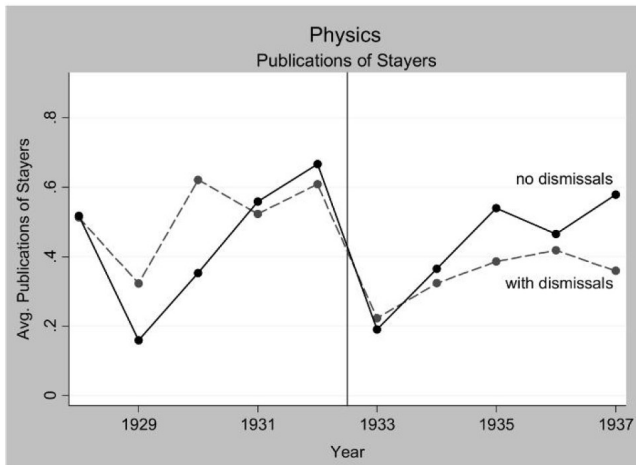
$$\begin{aligned}\# \text{of Peers}_{ut} &= \delta_1 + \delta_2 (\text{Dismissal-Induced } \downarrow \text{ in Peer Quality})_{ut} \\ &+ \delta_3 (\# \text{Dismissed})_{ut} \\ &+ \delta_4 \text{Age FE}_{iut} + \delta_5 \text{Year FE}_t + \delta_6 \text{Ind. FE}_i + \varepsilon_{iut}\end{aligned}$$

- Reduced form:

$$\begin{aligned}\#Publications_{iut} &= \theta_1 + \theta_2 (\text{Dismissal-Induced } \downarrow \text{ in Peer Quality})_{ut} \\ &+ \theta_3 (\#Dismissed)_{ut} \\ &+ \theta_4 \text{Age FE}_{iut} + \theta_5 \text{Year FE}_t + \theta_6 \text{Ind. FE}_i + \varepsilon_{iut}\end{aligned}$$

- Compares changes in publications from the pre- to the post-dismissal period for researchers in affected departments to the change between the two periods for unaffected researchers.

## Effect of Dismissal on Stayers



Source: Waldinger (2012), Figure 2.

## Reduced Form

	(1)	(2)	(3)	(4)	(5)	(6)
	Physics		Chemistry		Mathematics	
Dependent variable	Publications	Cit. weighted publications	Publications	Cit. weighted publications	Publications	Cit. weighted publications
Dismissal induced fall in peer quality	0.029 (0.015)	0.312 (0.235)	0.012 (0.015)	0.383 (0.303)	0.022 (0.031)	-0.464 (0.337)
Number dismissed	-0.021 (0.017)	-0.017 (0.302)	-0.018 (0.009)*	-0.130 (0.222)	-0.018 (0.015)	-0.016 (0.167)
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2261	2261	3584	3584	1538	1538
No. of researchers	258	258	413	413	183	183
R-squared	0.39	0.25	0.67	0.54	0.32	0.20

Source: Waldinger (2012), Table 4.

## First Stages

Dependent variable	Physics		Chemistry		Mathematics	
	Peer quality	Department size	Peer quality	Department size	Peer quality	Department size
Dismissal induced fall in peer quality	-0.644** (0.099)	-0.147 (0.130)	-1.114** (0.196)	0.011 (0.110)	-1.355** (0.149)	-0.228 (0.174)
Number dismissed	0.017 (0.098)	-0.570** (0.117)	-0.047 (0.162)	-0.998** (0.091)	0.160** (0.053)	-0.470** (0.062)
Age dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2261	2261	3584	3584	1538	1538
No. of researchers	258	258	413	413	183	183
$R^2$	0.59	0.90	0.66	0.91	0.70	0.81
$F$ —Test on instruments	81.9	103.10	18.3	64.3	47.8	66.2
Cragg–Donald EV statistic		12.8		89.8		46.7

Source: Waldinger (2012), Table 5.

## First Stages

	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
Physics				
Dependent variable:	Publi- cations	Publi- cations	Cit. weigt. Pubs.	Cit. weigt. Pubs.
Peer quality	0.004 (0.005)	-0.054 (0.035)	-0.048 (0.075)	-0.488 (0.496)
Department size	-0.007 (0.004)	0.035 (0.034)	-0.177** (0.062)	0.016 (0.553)
Age dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
Observations	2261	2261	2261	2261
No. of researchers	258	258	258	258
$R^2$	0.39		0.25	
Cragg-Donald EV Stat.		12.79		12.79

Source: Waldinger (2012), Table 6.

- Drop 1933 and 1934.
- Independent from age of dismissal scientists.
- Independent from baseline size or quality of department.
- Robust to including university FE and university-specific time trends.
- Independent from dismissal at sub-field level or top scientists.

# Interpretation

- No evidence of localized peer effects in science.
- Potential explanations:
  - No long-run data.
  - Strong integration of pre-WWII German scientific community.
  - Science very specialized.
  - Mature fields at the time.



# German Jewish Emigrés and US Invention

- Moser, Voena, and Waldinger (2014) “German Jewish Emigrés and US Invention.”
- Strong migration of German Jewish to the US fleeing from Nazi Germany:
  - 133,000 German Jewish émigrés by 1944.
  - 2,400 academics among them.
- Did they generate spillovers to US innovation?
- Focus on chemists, which innovation is well measured by patents.

# Emigré Chemistry Professors

- Collect information on all 535 chemistry professors and postdoctoral fellows at German and Austrian Universities in 1931–1933.
- 93 chemistry professors were dismissed between 1933 and 1941.
- Among these, 26 emigrated to the US and remained professionally active.

# US Patents of German Chemists

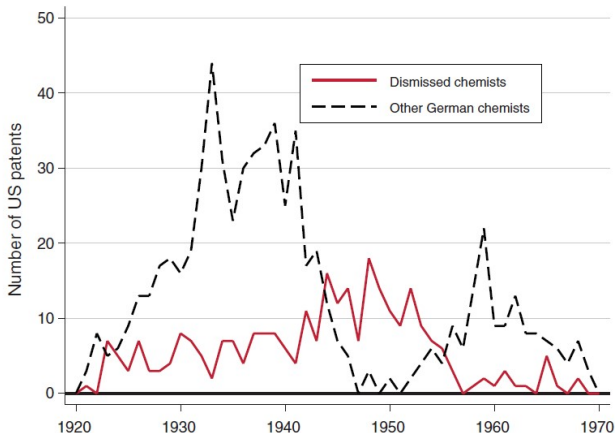
- Collect information on all US patents issued to all 535 German chemistry professors between 1920 and 1970:
  - Description of invention and USPTO class.
  - Date of patent.
  - Location of patentee.
  - Employment history of patentee to ensure match.

# US Patents of German Chemists

- Collect information on all US patents issued to all 535 German chemistry professors between 1920 and 1970:
  - Description of invention and USPTO class.
  - Date of patent.
  - Location of patentee.
  - Employment history of patentee to ensure match.
- Total patents 1920–1970 by 535 German chemists:
  - 946 patents.
  - 282 patents by 43 dismissed German chemists.
  - 157 patents by 13 German Jewish émigrés to the US.

# US Patents of German Chemists

US Patents per Year by German Chemists



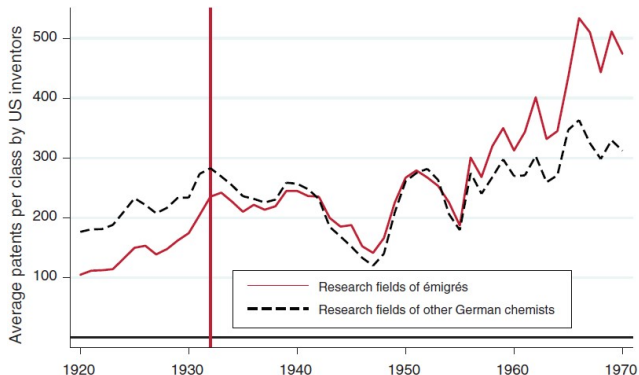
Source: Moser, Voena, and Waldinger (2014), Figure 1.

# US Patents of US Inventors

- Collect all US patents in 166 classes with patents by German chemists 1920–1970.
- Keep only patents by inventors born in the U.S.
- Also collect information on 964k US inventors of 1.3M patents in research fields of German chemists 1920–1970.

# Effect of Emigrés on Domestic Invention

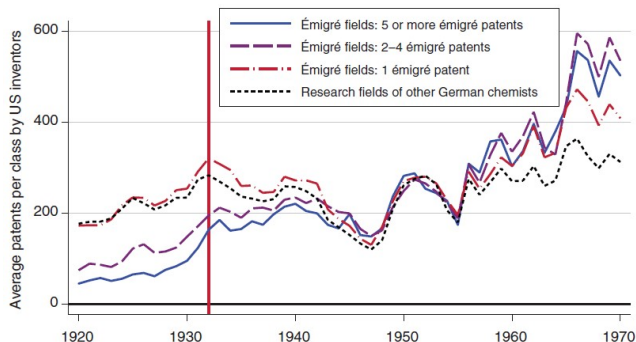
US Patents per Class and Year by Domestic US Inventors



Source: Moser, Voena, and Waldinger (2014), Figure 2A.

# Effect of Emigrés on Domestic Invention

## US Patents per Class and Year by Domestic US Inventors



Source: Moser, Voena, and Waldinger (2014), Figure 2B.



# Effect of Emigrés on Domestic Invention

$$\begin{aligned} \text{Patents by US inventors}_{c,t} &= \alpha_0 + \beta \text{émigré class}_c \cdot \text{post}_t \\ &+ \gamma' \mathbf{X}_{c,t} + \delta_t + f_c + \varepsilon_{c,t} \end{aligned}$$

- Technology class  $c$  and year  $t$  between 1920 and 1970.
- émigré class $_c$ : indicator for technology class  $c$  includes at least one patent by German Jewish émigré (number of émigré patents in alternative specifications).
- post $_t$ : indicator for post 1933 in Germany and post 1938 in Austria.
- Control group: classes with patents by other German chemists but not émigrés.
- $\mathbf{X}$ : controls for variation in patenting at level of research field and year.

# Effect of Emigrés on Domestic Invention

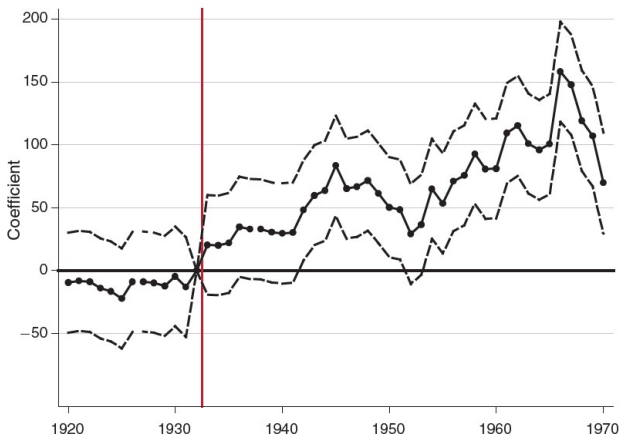
## OLS Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Émigré class × post</i>	105.222*** (22.203)	91.712*** (19.212)	84.803*** (18.950)	75.439*** (19.326)				
<i>Number émigré patents × post</i>					5.848* (3.058)	4.992* (2.561)	4.527** (2.182)	3.991** (1.956)
Number foreign patents	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Quadratic class age	No	No	Yes	Yes	No	No	Yes	Yes
Patent pools	No	No	No	Yes	No	No	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,466	8,466	8,466	8,466	8,466	8,466	8,466	8,466
$R^2$	0.783	0.845	0.849	0.851	0.779	0.842	0.846	0.848

Source: Moser, Voena, and Waldinger (2014), Table 2. Given mean 241 patents per class and year in classes with patents by other German chemists, arrival of émigré chemist increased US patenting by 31%.

# Effect of Emigrés on Domestic Invention

Year-Specific OLS Estimates



Source: Moser, Voena, and Waldinger (2014), Figure 3.

# Effect of Emigrés on Domestic Invention

- Sources of endogeneity of OLS estimates:
  - Dismissed inventors may enter fields becoming more productive.
  - Historical evidence: negative selection of Jewish dismissed inventors, and work in less productive fields.

# Effect of Emigrés on Domestic Invention

- Sources of endogeneity of OLS estimates:
  - Dismissed inventors may enter fields becoming more productive.
  - Historical evidence: negative selection of Jewish dismissed inventors, and work in less productive fields.
- Use pre-1933 patents of dismissed chemists to instrument for 1920–1970 patents of émigrés to the US.
- Exploits fact that research decisions by German Jewish chemists prior dismissal independent from expectations of research in U.S. after 1933.
- First stage:

$$\begin{aligned} \text{Emigré class}_c \cdot \text{post}_t &= \zeta_0 + \phi \text{pre-1933 dismissed class}_c \cdot \text{post}_t \\ &+ \theta' \mathbf{X}_{c,t} + \lambda_t + \mu_c + \nu_{c,t} \end{aligned}$$

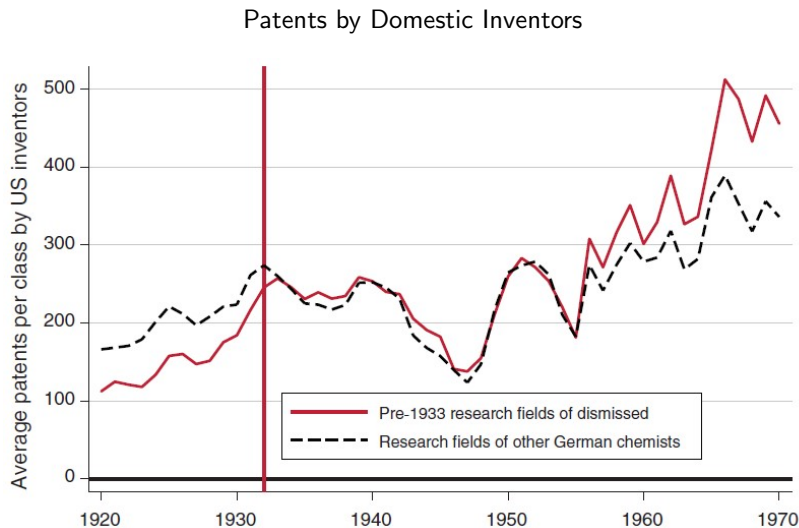
# Effect of Emigrés on Domestic Invention

## First-Stage and Reduced Form

	First stage							
	<i>Émigré class</i> × <i>post</i>		Number <i>émigré patents</i> × <i>post</i>		Reduced form			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Dismissed class</i> × <i>post</i>	0.370*** (0.081)	0.339*** (0.079)			80.821*** (23.155)	57.752*** (19.436)		
<i>Number dismissed patents</i> × <i>post</i>			1.384*** (0.442)	1.303*** (0.435)			35.595*** (6.547)	22.330*** (6.339)
Number foreign patents	No	Yes	No	Yes	No	Yes	No	Yes
Quadratic class age	No	Yes	No	Yes	No	Yes	No	Yes
Patent pools	No	Yes	No	Yes	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,466	8,466	8,466	8,466	8,466	8,466	8,466	8,466
$R^2$	0.801	0.809	0.770	0.773	0.779	0.849	0.782	0.849
$F$ -statistic	20.80	18.25	9.79	8.99				

Source: Moser, Voena, and Waldinger (2014), Table 3.

# Effect of Emigrés on Domestic Invention



Source: Moser, Voena, and Waldinger (2014), Figure 4.

# Effect of Emigrés on Domestic Invention

## Instrumental Variable Regression

	(1)	(2)	(3)	(4)
<i>Émigré class × post</i>	218.707*** (60.614)	170.136*** (57.992)		
<i>Number émigré patents × post</i>			25.717*** (8.750)	17.137** (6.909)
Number foreign patents	No	Yes	No	Yes
Quadratic class age	No	Yes	No	Yes
Patent pools	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Class fixed effects	Yes	Yes	Yes	Yes
Observations	8,466	8,466	8,466	8,466

Source: Moser, Voena, and Waldinger (2014), Table 4. Given mean 241 patents per class and year in classes with patents by other German chemists, arrival of émigré chemist increased US patenting by 71%. LATE estimates increase in patenting by US inventors for classes in which émigrés to the US patented because dismissed chemists had patented in same class before 1933.

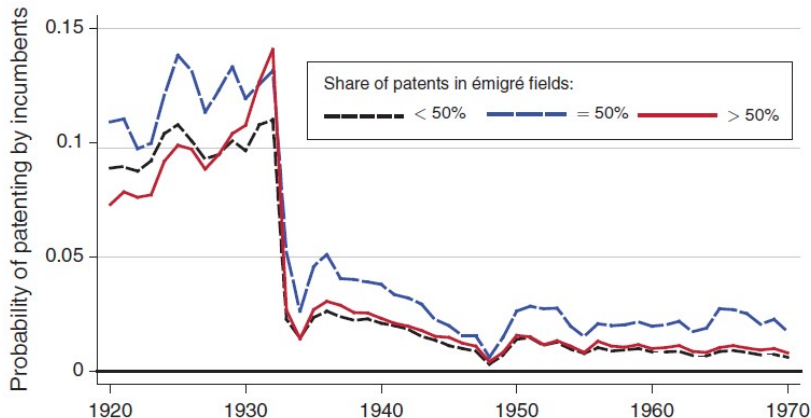


# Mechanisms: Effect on Incumbent US Inventors

- Examine changes in patenting for 210k US inventors who had patented before 1933 in a research field of German chemists.
- Estimate differential effects of émigrés on incumbent inventors depending on share of incumbent's patents in research fields of émigrés.

# Mechanisms: Effect on Incumbent US Inventors

Patenting by Incumbent Innovators in Fields of Emigrés vs Other German Chemists



Source: Moser, Voena, and Waldinger (2014), Figure 5.

# Mechanisms: Effect on Incumbent US Inventors

$$\text{Patenting}_{i,t} = \alpha + \beta \text{share of patents in émigré classes}_i \cdot \text{post}_t \\ + \gamma' \mathbf{Z}_{i,t} + \delta_t + f_i + \varepsilon_{i,t}$$

- $\text{Patenting}_{i,t}$ : indicator for incumbent US inventor  $i$  patented at least one invention in year  $t$ .
- $\mathbf{Z}_{i,t}$ : controls for variation in productivity over life cycle of inventor.
- Incumbent inventor FE and year FE.
- Instrument share of inventor's overall patents in research fields of émigrés with share of inventor's pre-1933 patents in fields with pre-1933 patents of dismissed chemists.

# Mechanisms: Effect on Incumbent US Inventors

## First Stage and Reduced Form

	First stage		Reduced form	
	(1)	(2)	(3)	(4)
<i>Share of pre-1933 patents in dismissed classes <math>\times</math> post</i>	0.403*** (0.086)	0.402*** (0.086)	-0.0003 (0.001)	-0.009*** (0.002)
Quadratic time to first patent	No	Yes	No	Yes
Quadratic time since first patent	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Inventor fixed effects	Yes	Yes	Yes	Yes
Observations	10,730,910	10,730,910	10,730,910	10,730,910
F-statistic	21.83	21.65		
R <sup>2</sup>	0.434	0.434	0.045	0.147

Source: Moser, Voena, and Waldinger (2014), Table 7.

# Mechanisms: Effect on Incumbent US Inventors

## OLS and IV Estimates

	OLS (linear probability)		IV	
	(1)	(2)	(3)	(4)
<i>Share of patents in émigré classes</i> × <i>post</i>	0.002** (0.001)	-0.007*** (0.002)	-0.001 (0.002)	-0.022*** (0.006)
Quadratic time to first patent	No	Yes	No	Yes
Quadratic time since first patent	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Inventor fixed effects	Yes	Yes	Yes	Yes
Observations	10,730,910	10,730,910	10,730,910	10,730,910
$R^2$	0.045	0.147	—	—

Source: Moser, Voena, and Waldinger (2014), Table 6. Incumbent inventors who had a 10% larger share of their patents in émigré classes became 22% less likely to patent an invention after 1932.

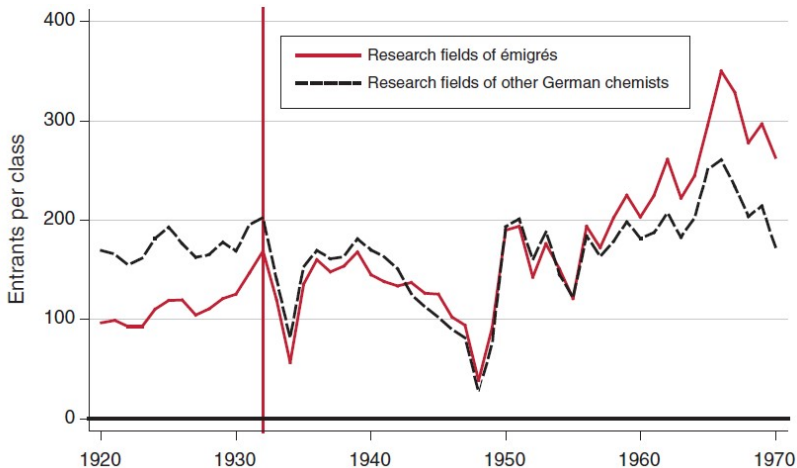
⇒ Knowledge spillovers from émigrés to incumbent inventors unlikely to be driving force behind increase in US patenting after 1933.

# Mechanisms: Entry into Research Fields of Emigrés

- Arrival of émigrés might have encouraged US scientists to switch into fields of émigrés, or attract a new group of US scientists.
- Measure researcher's year of entry in new field by first patent in émigré class.
- Distinguish new inventors from entry by former inventors in other fields.

# Mechanisms: Entry into Research Fields of Emigrés

## Entry of US Patentees into Research Fields of Emigrés



Source: Moser, Voena, and Waldinger (2014), Figure 7.

# Mechanisms: Entry into Research Fields of Emigrés

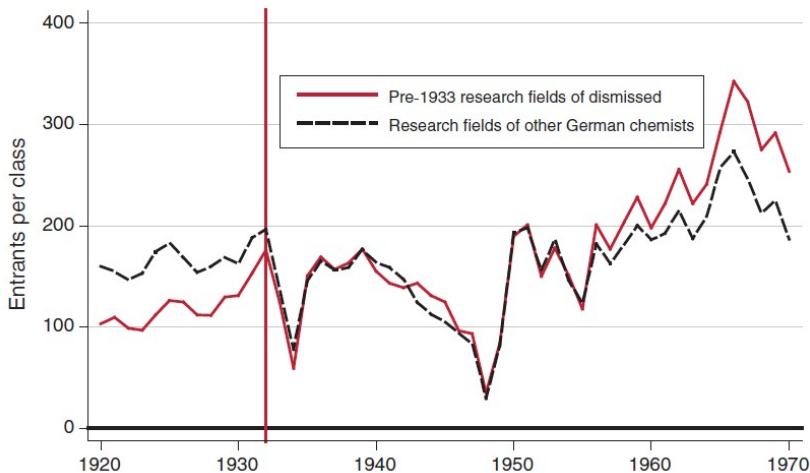
$$\text{Entry}_{c,t} = \alpha + \beta \text{émigré class}_c \cdot \text{post}_t + \gamma' \mathbf{X}_{c,t} + \delta_t + f_c + \varepsilon_{c,t}$$

- $\text{Entry}_{c,t}$ : new researchers per class and year (measured by researcher first patent in class  $c$ ).
- $\text{émigré class}_c$ : indicator if class  $c$  includes at least one patent by an émigré.
- Class FE and year FE.
- Instrument fields of émigrés by pre-1933 research fields of dismissed chemists.



# Mechanisms: Entry into Research Fields of Emigrés

## Entry of US Patentees into Research Fields



Source: Moser, Voena, and Waldinger (2014), Figure 8.

# Mechanisms: Entry into Research Fields of Emigrés

## OLS and IV Estimates

	OLS				Instrumental Variables			
	Entrants into field		Entrants into patenting		Entrants into field		Entrants into patenting	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Émigré class</i> <i>× post</i>	73.799*** (15.674)	58.181*** (14.715)	53.434*** (12.522)	43.967*** (12.261)	162.287*** (44.195)	142.119*** (45.982)	116.707*** (34.565)	109.466*** (37.863)
Number foreign patents	No	Yes	No	Yes	No	Yes	No	Yes
Quadratic class age	No	Yes	No	Yes	No	Yes	No	Yes
Patent pools	No	Yes	No	Yes	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,466	8,466	8,466	8,466	8,466	8,466	8,466	8,466
R <sup>2</sup>	0.781	0.835	0.763	0.805	0.767	0.824	0.750	0.792

Source: Moser, Voena, and Waldinger (2014), Table 9. Given mean 166 entrants per class and year in classes with patents by other German chemists, arrival of émigré chemist increased entrants 35%.

- Evidence that collaborators of émigrés switched into research fields of émigrés after 1933.
- Amplification through network of collaborations including co-inventors of these induced co-inventors of Jewish émigrés.

- Evidence that collaborators of émigrés switched into research fields of émigrés after 1933.
- Amplification through network of collaborations including co-inventors of these induced co-inventors of Jewish émigrés.
- Further including junior émigrés not professor at time of dismissal suggests broader effects.

# Topic 4

## Economics history and urban economics

# Why turn to history in urban economics?

Hanlon and Hebllich (2020)

- Existence of cities are the result of slow-moving processes.
- Urban infrastructures characterized by path dependence.
- Expand set of experiences of rare events to inform policy.
- Historical natural experiments for identification.
  - Exogenous variation.
  - Fewer interdependencies.
- Economic development characterized by long processes.

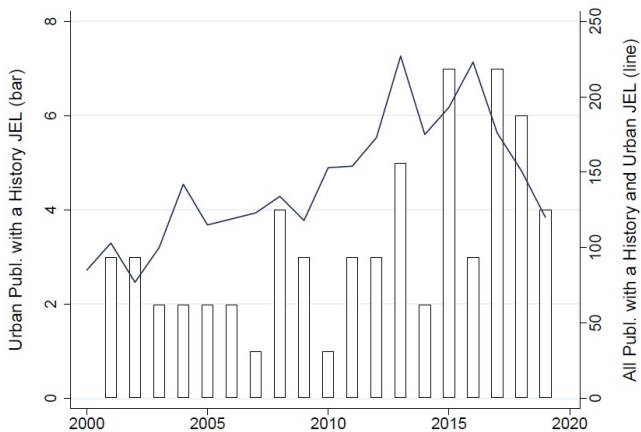
# Plan of session

- Innovations in historical research.
- Origins of cities development.
- Spatial distribution of economic activity within cities.
- Cities and growth

# Innovations in historical urban economics



## Publications in historical urban economics



Source: Hanlon and Heblich (2020), Figure 1a

# Historical statistics

- Increased digitization of archival data sources with spatial dimension:
  - Censuses
  - Birth and death records
  - Patent records
  - Immigration lists
  - Tax records
- See [Giuliano and Matranga \(2020\)](#)

# Historical statistics

- Increased digitization of archival data sources with spatial dimension:
  - Censuses
  - Birth and death records
  - Patent records
  - Immigration lists
  - Tax records
- See [Giuliano and Matranga \(2020\)](#)
- Archival material digitization technologies:
  - Optical character recognition (OCR) technologies (though low quality images, heterogeneous typesetting and fonts, hand-written documents).
  - Crowdsourcing to citizen historians (e.g., zooniverse platform). See special issue of *Historical Methods* (2019), 52(3)
  - Machine learning for digitizing maps. [Combes et al. \(2020\)](#)

# Historical statistics: Challenges

- Standardized frame of reference (HGIS). [Gay \(2021\)](#)
- Matching spatial units.
  - Non-overlapping levels of aggregation: modifiable area unit problem (MAUP).
  - Boundary changes over time. [Eckert et al. \(2020\)](#)

# Historical statistics: Challenges

- Standardized frame of reference (HGIS). [Gay \(2021\)](#)
- Matching spatial units.
  - Non-overlapping levels of aggregation: modifiable area unit problem (MAUP).
  - Boundary changes over time. [Eckert et al. \(2020\)](#)
- Data source with in-text spatial information (e.g., patents): computational linguistics. [Perlmann \(2015\)](#)

# Historical statistics: Challenges

- Standardized frame of reference (HGIS). [Gay \(2021\)](#)
- Matching spatial units.
  - Non-overlapping levels of aggregation: modifiable area unit problem (MAUP).
  - Boundary changes over time. [Eckert et al. \(2020\)](#)
- Data source with in-text spatial information (e.g., patents): computational linguistics. [Perlmann \(2015\)](#)
- Linking individuals over time.  
[Abramitzky et al. \(2020\)](#) [Bailey et al. \(2020\)](#)

# Historical maps

- Increasingly available, rich information (boundaries, buildings, neighborhood, land use, transportation infrastructures. . . ).
- Key for structures of cities and neighborhood.  
[Hornbeck and Keniston \(2017\)](#)
- Challenges in digital and automatic processing.  
[Combes et al. \(2020\)](#)
- Often lack precision, yet key for geographical RDDs.  
[Sequeira et al \(2020\)](#)

# Time invariant factors

- To compensate for lack of historical information, use time invariant factors that explain urban patterns:
  - Geological conditions.
  - Climate conditions.
  - Terrain and soil characteristics
- Interact characteristics with time FE to assess when time invariant factor gained importances. [Nunn and Qian \(2011\)](#)



# Time invariant factors

- To compensate for lack of historical information, use time invariant factors that explain urban patterns:
  - Geological conditions.
  - Climate conditions.
  - Terrain and soil characteristics
- Interact characteristics with time FE to assess when time invariant factor gained importances. [Nunn and Qian \(2011\)](#)
- Genetic markers and linguistic traits. [Spolaore and Wacziarg \(2009\)](#)

# Predicting outcomes

- Use models from outside economics to predict historical conditions.
  - Atmospheric models (e.g., wind patterns for sailing time). [Pascali \(2017\)](#)
  - Engineering rules for in-land transportation cost and time. [Bogart et al \(2019\)](#)
  - Environment models for pollution dissemination. [Heblich et al \(2016\)](#)
- See *Annual Review* article “Quantitative Spatial Economics.” [Redding and Rossi-Hansberg \(2017\)](#)

# Boundary changes over time

- [Eckert et al \(2020\)](#) “A Method to Construct Geographical Crosswalks with an Application to US Counties since 1790.”
    - Historical data at heterogeneous levels of aggregation.
    - Mapping heterogeneous geographies requires cross-walks within and across years.
- ⇒ Develop geographical crosswalks Python routines to build consistent panels.

# Boundary changes over time

- [Eckert et al \(2020\)](#) “A Method to Construct Geographical Crosswalks with an Application to US Counties since 1790.”
    - Historical data at heterogeneous levels of aggregation.
    - Mapping heterogeneous geographies requires cross-walks within and across years.
- ⇒ Develop geographical crosswalks Python routines to build consistent panels.
- Map US county-level data from 1790–2010 to:
    - 1990 and 2010 county boundaries.
    - 1990 commuting zones.
  - Generic code can be applied to construct crosswalks for arbitrary geographies.

# Method to construct geographical crosswalks

- General idea: intersecting maps of the same territory partitioned in different ways.
- Two shapefiles required (in same projection):
  - *Reporting* map for spatial units for which data is reported.
  - *Reference* map for spatial units to which data is to be re-aggregated.
- Intersect maps to generate a new map of smaller mutually exclusive spatial subunits.
- Recombine to form partition in either reporting or reference map, with areas of subunits as weights.

# Method to construct geographical crosswalks

## New York state county boundaries



(a) 1790 COUNTIES

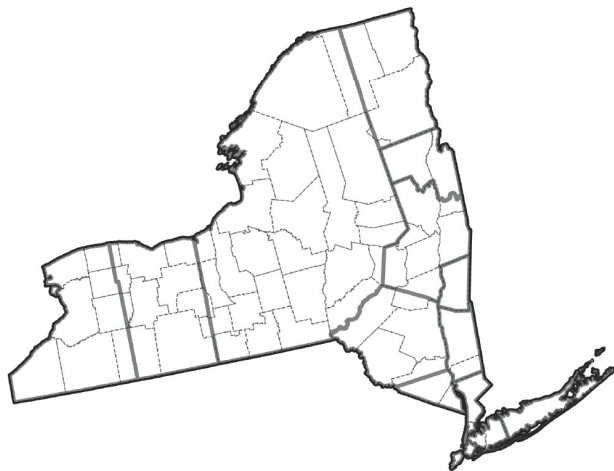


(b) 2010 COUNTIES

Source: Eckert et al (2020), Figure 1

# Method to construct geographical crosswalks

Intersection between New York state county boundaries



Source: Eckert et al (2020), Figure 2

# Method to construct geographical crosswalks

- Compute for each subunit (disjoint) their share of total land area.
- For each region in reporting map, compute share of its subunits in total land area.
- Each spatial unit in the two maps has weights associated.
- Map 1790 population to 2010 counties:
  - Take population counts by county in 1790.
  - Multiply 1790 county population counts by associated weights.
  - This uniformly distribute population across subunits it contains.
  - Sum these counts across subunits contained within each of 2010 county delineations.



# Method to construct geographical crosswalks

- Underlying assumption: variable distributed proportionally to land area size across subunits within each spatial unit of reporting map.
- For plausibility: use most disaggregated maps for which data are reported to construct geographical crosswalks.
- Use for stock variables.
- Different weights required for average outcomes (e.g., average wages).
- To use Python code, only specify reporting and reference map.

# Application: US counties since 1790

- Apply method to construct crosswalk for US counties since 1790.
- Three features responsible to vast changes in counties:
  - Territorial expansion from 1790 (Louisiana, Oregon, Mexico, Alaska) to 1898 (Hawaii).
  - The creation of the Union and integration US territories until 1959.
  - Changing county boundaries due to population changes (generally partitions).

# Application: US counties since 1790

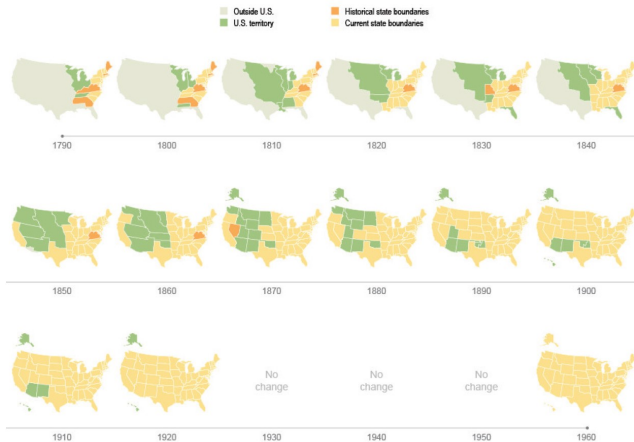
## Territorial history of the United States



Source: Eckert et al (2020), Figure 3

# Application: US counties since 1790

## The territory and the Union States over time



Source: Eckert et al (2020), Figure 4

# Application: US counties since 1790

## Changing county numbers and boundaries

Year	Total Subdivisions	Fraction of 2010 Territory	Fraction of Crosswalk Weights <i>between</i>		
			0.4 and 0.6	0.2 and 0.8	0.01 and 1
1790	292	0.2221	0.018	0.065	0.265
1800	427	0.232	0.021	0.085	0.291
1810	587	0.4882	0.020	0.071	0.278
1820	779	0.4797	0.021	0.075	0.306
1830	1002	0.4955	0.022	0.084	0.259
1840	1285	0.4945	0.021	0.078	0.210
1850	1632	0.8305	0.015	0.059	0.195
1860	2126	0.8377	0.012	0.042	0.157
1870	2334	0.8377	0.010	0.036	0.129
1880	2614	0.9982	0.010	0.028	0.085
1890	2799	0.9982	0.007	0.020	0.062
1900	2848	1	0.006	0.018	0.050
1910	2963	1	0.003	0.012	0.035
1920	3075	1	0.001	0.003	0.017
1930	3110	1	0	0.001	0.009
1940	3108	1	0	0.001	0.008
1950	3111	1	0	0.001	0.007
1960	3134	1	0	0.001	0.007
1970	3141	1	0	0.001	0.005
1980	3137	1	0	0.001	0.003
1990	3141	1	0	0	0.002
2000	3141	1	0	0	0.002
2010	3143	1	0	0	0

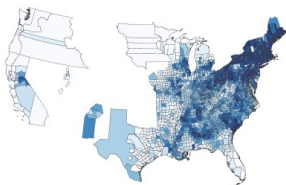
Source: Eckert et al (2020), Table 2

# Application: US counties since 1790

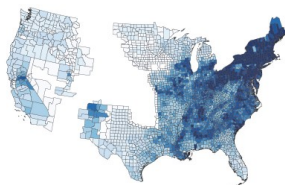
- Historical shapefiles (including territories outside US) from NHGIS. [Manson et al \(2017\)](#)
- Application:
  - For each decade 1790–2000: intersect county shapefile with 2010 county shapefile.
  - For each county-decade, compute share of its area and assign to 2010 list of counties.
  - Repeat for 1990 counties and 1990 commuting zones reference maps.

# Application: US counties since 1790

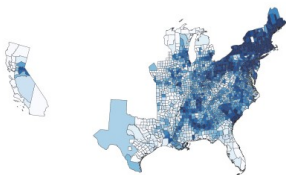
County-level population in 1850 for different county delineations



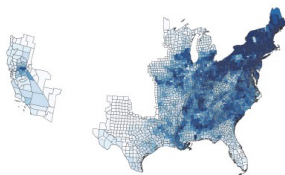
(a) 1850 COUNTIES, 1850 US TERRITORY



(b) 2010 COUNTIES, 1850 US TERRITORY



(c) 1850 COUNTIES, 1850 UNION TERRITORY



(d) 2010 COUNTIES, 1850 UNION TERRITORY

Source: Eckert et al (2020), Figure 5

# Potential challenges

- Need to assume uniform distribution of stock variable across subunits within units.
- Are these units similarly meaningful across time, especially if use unit FE in panel setting?
- Might generate biases if externalities at county level not absorbed by the FE.
- Think about meaning of “units” and changes over time, especially in relation to process of interest.



# Alternative strategy

- Use administrative information on changes without resorting to map.
- Example: Third-Republic Information System [Gay \(2021\)](#)
- Need “building blocks” (e.g., communes) and attribute them to aggregated units.
- Challenge: infra-building-block constituencies.

# Origins of cities development

# Origins of cities development

- Central issues in urban economics:
  - How to explain the distribution of population across space?
  - How to explain the existence and development of cities?
- Part of broader importance of history for long-run development.  
Nunn (2014)

# Origins of cities development

- Central issues in urban economics:
  - How to explain the distribution of population across space?
  - How to explain the existence and development of cities?
- Part of broader importance of history for long-run development.  
Nunn (2014)
- Main answers:
  - Location fundamentals.
  - Productivity effects of agglomeration.
  - Amenities and disamenities.
  - Trade and market access.
  - Education and human capital.

# Location fundamentals

- Role of location fundamentals for city distributions and growth?
- Broad use of historical natural experiments.
  - Bombings during wars, e.g. US bombings of Japan during WWII. [Davis and Weinstein \(2002\)](#)  
⇒ No persistent effects.
  - Economic shocks to cities e.g., shock to cotton imports in Britain during US Civil War. [Hanlon \(2017\)](#)  
⇒ Persistent impact on city size.
  - Temporary relocation of populations e.g., resettlement of expelled Germans after WWII. [Schumann \(2014\)](#)  
⇒ Persistent impact on city size.

⇒ Location fundamentals importance ↑ as differences across locations ↑.

# Location fundamentals

- Importance of location fundamentals is time-dependent on advantages provided by introduction of new technologies:
  - Soil quality and potato introduction in Europe. [Nunn and Qian \(2011\)](#)
  - Coal reserves and the Industrial Revolution in Europe. [Fernihough and O'Rourke \(2014\)](#)
  - Ports naturally accommodating large ships and containerized shipping. [Brooks et al \(2016\)](#)
- Interactions location fundamentals and technological change is active area of research.

# Bombs and break points

- [Davis and Weinstein \(2002\)](#) “Bones, Bombs, and Break Points: The Geography of Economic Activity.”

# Bombs and break points

- Davis and Weinstein (2002) “Bones, Bombs, and Break Points: The Geography of Economic Activity.”
- Two theories of distribution of economic activity across space, and predictions to temporary shock:
  - Agglomeration effects:
    - Importance of knowledge spillovers, labor-market pooling, proximity of supply and demand when trade is costly.
    - Strong role of path dependence.
    - Temporary shocks have permanent consequences.
  - Location fundamentals:
    - Location-specific characteristics might be randomly allocated, but systematically determine later growth.
    - Importance varies depending of substitutes and interactions with technology.
    - Temporary shocks have temporary consequences.



# Bombs and break points

- Test theories through historical natural experiment of Allied bombing of Japanese cities during WWII.
- Data:
  - 303 Japanese cities with population  $> 30,000$  in 1925.
  - Population counts 5-year intervals 1925–1965 (except 1945→1947).
  - Population growth rates 1925–1940, 1940–1947, 1947–1965.

# Bombs and break points

- Test theories through historical natural experiment of Allied bombing of Japanese cities during WWII.
- Data:
  - 303 Japanese cities with population  $> 30,000$  in 1925.
  - Population counts 5-year intervals 1925–1965 (except 1945→1947).
  - Population growth rates 1925–1940, 1940–1947, 1947–1965.
- Measure of intensity of shocks:
  - Number of dead and missing city residents due to war.
  - Number of building destroyed per resident.
- Control for (limited) US and Japanese government expenditures for reconstruction.

# Bombs and break points

- Magnitude of the shocks
  - Allied strategic bombing of 66 Japanese cities.
  - Destruction of 1/2 structures (2.2 million buildings), 2/3 productive capacity.
  - 300,000 Japanese killed.
  - Nuclear bombings:
    - Hiroshima: 2/3 city destroyed, 20% of population (80,000).
    - Nagasaki: 40% city destroyed, 8.5% of population (25,000).

- Variance of the shocks
  - 80% of cities (37% of urban population) untouched by bombings (e.g., Kyoto).
  - Variability in destruction within cities.
  - Other factors: evolving defense of Japanese cities, topography, growth of B-29s in service.

# Bombs and break points

- Size of any city at any point in time:

$$s_{it} = \Omega_i + \varepsilon_{it}$$

- $s_{it}$ : log city  $i$  share of total population at time  $t$ .
- $\Omega_i$ : initial city size.
- $\varepsilon_{it}$ : city-specific shock.

# Bombs and break points

- Size of any city at any point in time:

$$s_{it} = \Omega_i + \varepsilon_{it}$$

- $s_{it}$ : log city  $i$  share of total population at time  $t$ .
  - $\Omega_i$ : initial city size.
  - $\varepsilon_{it}$ : city-specific shock.
- Persistence of shock to population:

$$\varepsilon_{it+1} = \rho\varepsilon_{it} + \nu_{it+1}$$

- $\rho \in [0, 1]$
- $\nu_{it+1}$  is iid error term.

# Bombs and break points

- Evolution of city size:

$$s_{it+1} - s_{it} = (\rho - 1)\nu_{it} + [\nu_{it+1} + \rho(1 - \rho)\varepsilon_{it-1}] \quad (1)$$

# Bombs and break points

- Evolution of city size:

$$s_{it+1} - s_{it} = (\rho - 1)\nu_{it} + [\nu_{it+1} + \rho(1 - \rho)\varepsilon_{it-1}] \quad (1)$$

- Empirical strategy:
  - Estimate (1) by regressing growth rate 1947–1960 on growth rate 1940–1947 (get  $\rho - 1$ ).
  - Instrument growth rate 1940–1947 with buildings destroyed and deaths per capita.



# Bombs and break points

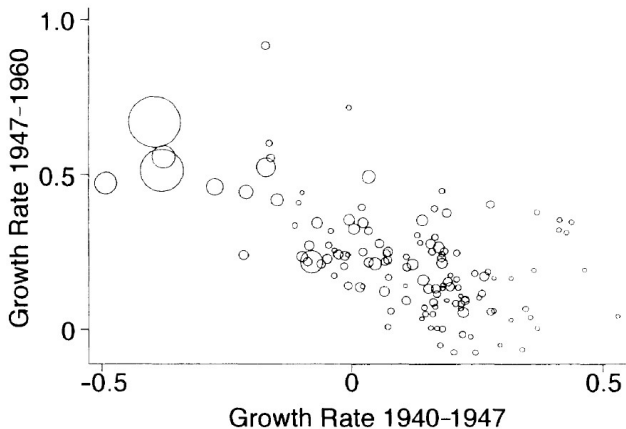
- Evolution of city size:

$$s_{it+1} - s_{it} = (\rho - 1)\nu_{it} + [\nu_{it+1} + \rho(1 - \rho)\varepsilon_{it-1}] \quad (1)$$

- Empirical strategy:
  - Estimate (1) by regressing growth rate 1947–1960 on growth rate 1940–1947 (get  $\rho - 1$ ).
  - Instrument growth rate 1940–1947 with buildings destroyed and deaths per capita.
- Two theories:
  - $\rho - 1 = 0$ : all shocks are permanent, city share is random walk with  $s_{it+1} = s_{it} + \varepsilon_{it+1}$ .
  - $\rho \in [0, 1)$ : city share is stationary, shocks dissipate.
  - $\rho - 1 = -1$ : shock was dissipated by 1960

# Bombs and break points

Effects of bombing on cities > 30,000 inhabitants



Source: Davis and Weinstein (2002), Figure 1. Only contains cities with positive casualty rates

# Bombs and break points

## Instrumental variables equation

Independent variable	Coefficient
Constant	0.213 (0.006)
Deaths per capita	-0.665 (0.506)
Buildings destroyed per capita	-2.335 (0.184)
$R^2$ :	0.409
Number of observations:	303

Source: Davis and Weinstein (2002), Table 2. Dependent variable: rate of growth in city population 1940–1947.

# Bombs and break points

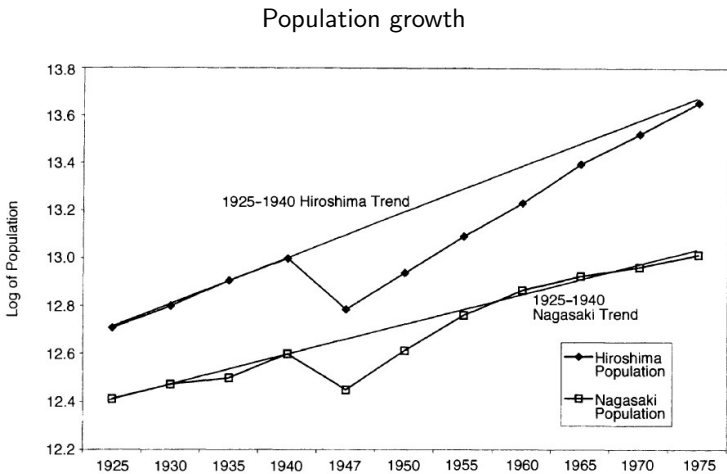
2SLS estimates of impact of bombing on cities

Independent variable	Dependent variable = growth rate of population between 1947 and 1960		Dependent variable = growth rate of population between 1947 and 1965
	(i)	(ii)	(iii)
Growth rate of population between 1940 and 1947	-1.048 (0.097)	-0.759 (0.094)	-1.027 (0.163)
Government reconstruction expenses	1.024 (0.387)	0.628 (0.298)	0.392 (0.514)
Growth rate of population between 1925 and 1940		0.444 (0.054)	0.617 (0.092)
$R^2$ :	0.279	0.566	0.386
Number of observations:	303	303	303

Source: Davis and Weinstein (2002), Table 3. Instruments: deaths per capita and buildings destroyed per capita.

- Objection to interpretation: refugees.
    - Population movement driven by people forced to live elsewhere, but later came back.
    - Argument cannot be made for Hiroshima and Nagasaki.
- ⇒ Full recovery within 20 years.

# Bombs and break points



Source: Davis and Weinstein (2002), Figure 2

# Productivity effects of agglomeration

- City size and clustering of specific industries generate productivity advantages (suppliers proximity, localized knowledge, workers pool).
  - Localized knowledge and supply of capable entrants in tire manufacturing in Akron, Ohio.  
[Buenstorf and Klepper \(2009, 2010\)](#)
  - Inter-industry local externalities (carriage and wagon industry) in regional agglomeration of US auto industry.  
[Cabral et al \(2018\)](#)
  - Origins of industry clusters among automakers in Detroit and semi-conductor manufacturers in Silicon Valley.  
[Klepper \(2010\)](#)
  - Impact of proximity on productivity of European classical composers in late 18th–early 19th century.  
[Borowiecki \(2013\)](#)

⇒ Cross-industry vs within-industry spillovers question.

# Productivity effects of agglomeration

- Implications for place-based policies.
  - Tennessee Valley Authority project 1930s–1960s  $\implies$  long-run impact on manufacturing but not agriculture thanks to agglomeration economies. [Kline and Moretti \(2013\)](#)
  - Subsidies to West German regions to compensate for loss in market access  $\implies$  persistent impact on economic density due to physical infrastructures investment. [Ehrlich and Seidel \(2018\)](#)
  - Transfer of capital of West Germany to Bonn in 1949 and public employment. [Becker et al \(2021\)](#)



# Impact of public employment: Bonn

- [Becker, Heblich, and Sturm \(2021\)](#) “The Impact of Public Employment: Evidence From Bonn.”
- General research question: spillover effects of changes in spatial distribution of public employment on the private sector.

# Impact of public employment: Bonn

- [Becker, Heblich, and Sturm \(2021\)](#) “The Impact of Public Employment: Evidence From Bonn.”
- General research question: spillover effects of changes in spatial distribution of public employment on the private sector.
- Use creation of West German government in Bonn as a source of variation to assess causal impact of public employment on spatial distribution of private sector activity.

# Theoretical framework: preferences and technology

- Mass  $L$  of workers, mobile across cities, endowed with unit labor supplied inelastically with zero disutility.

# Theoretical framework: preferences and technology

- Mass  $L$  of workers, mobile across cities, endowed with unit labor supplied inelastically with zero disutility.
- Consumption in city  $c$  over:
  - Tradable goods:  $C_c^T$
  - Non-tradable goods:  $C_c^N$
  - Housing:  $C_c^H$
  - Amenities:  $B_c$
- Cobb-Douglas utility function:

$$U_c = B_c \left(C_c^T\right)^\mu \left(C_c^N\right)^\theta \left(C_c^H\right)^{1-\mu-\theta}$$

# Theoretical framework: preferences and technology

- Transportation cost  $\tau_{ic} > 1$  for tradables produced in  $i$  consumed in  $c$ .
- Tradable price index CES (Dixit-Stiglitz):

$$P_c^T = \left[ \sum_i n_i^T \left( p_i^T \tau_{ic} \right)^{1-\sigma} \right]^{1/(1-\sigma)}$$

- Tradables  $n_i^T$  produced in  $i$  face same elasticity of demand and charge same equilibrium price to city  $c$ :  $p_{ic}^T = p_i^T \tau_{ic}$ .

# Theoretical framework: preferences and technology

- Transportation cost  $\tau_{ic} > 1$  for tradables produced in  $i$  consumed in  $c$ .
- Tradable price index CES (Dixit-Stiglitz):

$$P_c^T = \left[ \sum_i n_i^T \left( p_i^T \tau_{ic} \right)^{1-\sigma} \right]^{1/(1-\sigma)}$$

- Tradables  $n_i^T$  produced in  $i$  face same elasticity of demand and charge same equilibrium price to city  $c$ :  $p_{ic}^T = p_i^T \tau_{ic}$ .
- Non-tradables can only be consumed in city where produced:

$$P_c^N = \left[ n_c^N \left( p_c^N \right)^{1-\sigma} \right]^{1/(1-\sigma)} = p_c^N \left( n_c^N \right)^{1/(1-\sigma)}$$

# Theoretical framework: preferences and technology

- Both tradables and non-tradables produced by monopolistically competitive firms with same production technology.
- Fixed labor cost  $F$  and constant variable cost.
- Labor  $l$  to produce  $x$  units with productivity of firms in city  $i$   $\varphi_i$ :

$$l = F + \frac{x}{\varphi_i}$$

- Equilibrium supply function (constant mark-up over marginal cost):

$$p_i = \left( \frac{\sigma}{\sigma - 1} \right) \frac{w_i}{\varphi_i}$$

- Housing supply function:

$$P_c^H = (L_c)^\gamma$$

- $\gamma$ : elasticity of house price wrt employment.



# Theoretical framework: public sector

- Public sector produces global public good that affects all agents equally in all cities.
- Utility derived from public good normalized to zero.
- Number of public sector workers is constant.
- Policy decision: distribution of public employment across cities  $L_i^G$ .
- Public sector workers paid  $w_i^G = w_i$ , the private sector rate in city  $i$ .
- Public sector financed through flat income tax  $t$  levied on all workers:

$$t \sum_i (w_i L_i) = \sum_i (w_i L_i^G)$$

# Theoretical framework: public sector

- Public sector spillover to productivity of private sector worker in city  $i$ :

$$\varphi_i = \xi_i \left( L_i^G \right)^\alpha$$

- $\xi_i$ : sources of productivity other than public employment.
- $\alpha$ : elasticity of productivity of private sector to size of public sector.

# Theoretical framework: public sector

- Public sector spillover to productivity of private sector worker in city  $i$ :

$$\varphi_i = \xi_i \left( L_i^G \right)^\alpha$$

- $\xi_i$ : sources of productivity other than public employment.
  - $\alpha$ : elasticity of productivity of private sector to size of public sector.
- Mechanisms of spillovers:
    - Classical knowledge spillovers (e.g., about public policy).
    - Boosts in local demand from public sector, equivalent to productivity shock.

# Theoretical framework: public sector

- Public sector spillover to consumption amenities in city  $i$ :

$$B_c = \psi_i \left( L_c^G \right)^\delta$$

- $\psi_i$ : sources of amenities other than public employment.
- $\delta$ : elasticity of amenities to size of public sector.

# Theoretical framework: public sector

- Public sector spillover to consumption amenities in city  $i$ :

$$B_c = \psi_i \left( L_c^G \right)^\delta$$

- $\psi_i$ : sources of amenities other than public employment.
  - $\delta$ : elasticity of amenities to size of public sector.
- Mechanisms of spillovers:
    - Public sector workers could improve local schools and other public goods.
    - Ability to divert larger share of national public spending to these locations.

# Theoretical framework: spatial equilibrium

- Workers are freely mobile.
- Equal utility across cities implies:

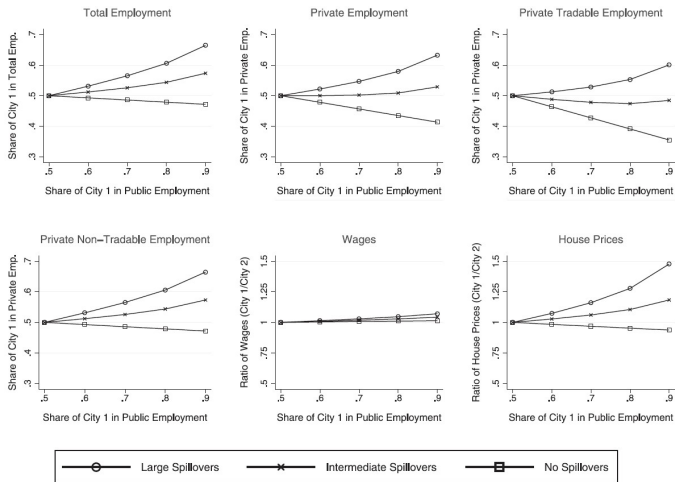
$$\frac{w_c B_c}{(P_c^T)^\mu (P_c^N)^\theta (P_c^H)^{1-\mu-\theta}} = k, \quad \forall c$$

- Higher wages and amenities make a city more attractive.
- Higher prices make a city less attractive.

# Simulations

- Simulate impact of changes in location of public employment on distribution of private sector employment in a two-city version of the model.
- Parameter values:
  - Elasticity of substitution  $\sigma = 5$ .
  - Share of expenditure on housing  $1 - \mu - \theta = 1/3$ .
  - Share of expenditure on tradable and non-tradable  $\mu = \theta = 1/3$ .
  - Share of public employment in total employment  $L_G/L = 12\%$ .
  - Trade cost between two cities  $\tau_{ic} = 1.25$ .
  - Elasticity of house prices wrt employment or population  $\gamma = 0.57$ .
- Compare different spillover values when shifting public employment from city 2 to city 1 (ex-ante symmetric):
  - No spillovers.
  - Intermediate spillovers:  $\alpha = \delta = 0.004$ .
  - Large spillovers:  $\alpha = \delta = 0.008$ .

## Simulating the impact of public employment



Source: Becker et al (2021), Figure 1



- Short run (before workers relocation):
  - Demand from public sector met by reduction in employment in private sector.
  - Housing prices are unchanged (population unchanged).
  - Because constant share spent on non-tradables, change in employment from tradable sector (imports from other cities).
  - Price of tradables  $\uparrow \implies \downarrow$  real wages and utility.
  - Thanks to productivity spillovers, firms can pay higher wages.
  - Thanks to amenities spillovers, utility  $\uparrow$ .

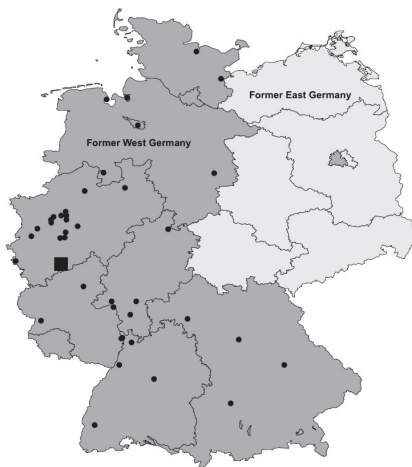
- Short run (before workers relocation):
  - Demand from public sector met by reduction in employment in private sector.
  - Housing prices are unchanged (population unchanged).
  - Because constant share spent on non-tradables, change in employment from tradable sector (imports from other cities).
  - Price of tradables  $\uparrow \implies \downarrow$  real wages and utility.
  - Thanks to productivity spillovers, firms can pay higher wages.
  - Thanks to amenities spillovers, utility  $\uparrow$ .
- Long run (mobility, if spillovers large enough):
  - Inflow of workers from city 2 to city 1.
  - Production in private sector  $\uparrow$ .
  - Inflow until house prices high enough to equalize utility.

# Historical background

- After WWII, Germany divided in East Germany, West Germany, parts to Poland and Soviet Union.
- West Germany founded in 1949, and capital voted as Bonn (other contenders: Frankfurt, Stuttgart, Kassel).
- Fall of Berlin Wall in November 1989.
- Formal reunification on October 1990.
- Relocation of capital to Berlin on June 1991.

- Census data before and after WWII: 1925–1987.
- 40 control cities in West Germany: 20 just below/above Bonn's population in 1939 (Bonn 31st largest West German city in 1939).
- State-level statistics for city-level employment data in 10 sectors.
- Public employment includes public and social security administration, but excludes health, education, and state-owned enterprises.
- Data on consumption amenities, gross value added per worker, wages, housing prices.

## Bonn and the control cities



Source: Becker et al (2021), Figure 2

- Difference-in-differences comparing Bonn to 40 control cities:

$$Y_{ct} = \gamma + T_c + T_c \times D_t + \lambda_t + \varepsilon_{ct}$$

- $Y_{ct}$ : employment in city  $c$  at time  $t$ .
- $T_c$ : indicator for Bonn.
- $D_t$ : indicator for post-1949.
- $\lambda_t$ : time FE.

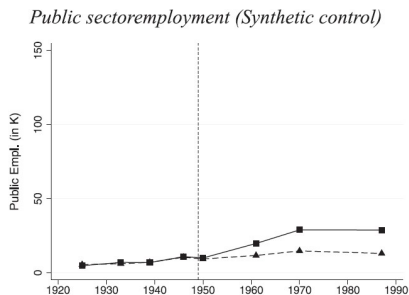
# Empirical strategy

- Difference-in-differences comparing Bonn to 40 control cities:

$$Y_{ct} = \gamma + T_c + T_c \times D_t + \lambda_t + \varepsilon_{ct}$$

- $Y_{ct}$ : employment in city  $c$  at time  $t$ .
  - $T_c$ : indicator for Bonn.
  - $D_t$ : indicator for post-1949.
  - $\lambda_t$ : time FE.
- 
- Issues when one control unit:
    - Similar control group.
    - Clustering at city level underestimates standard errors.
- ⇒ Synthetic control city for Bonn (on pre-1949 levels).  
Abadie and Gardeazable (2003) Abadie et al (2010)

## Public sector employment

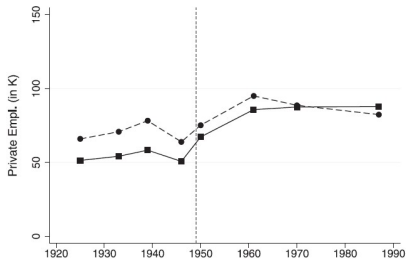


Source: Becker et al (2021), Figure 3. Synthetic: 1987 vs pre = +15,600 workers.

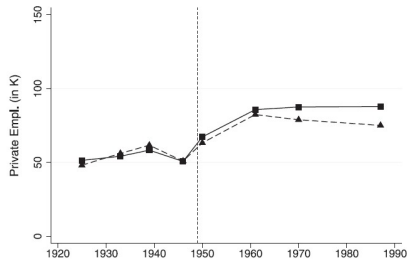


## Private sector employment

Private sector employment (Diff-in-diff)

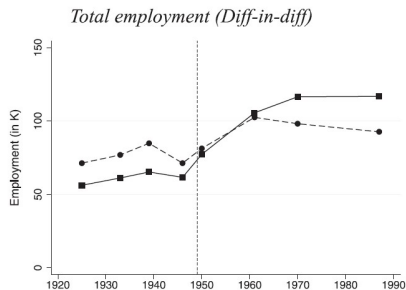


Private sector employment (Synthetic control)



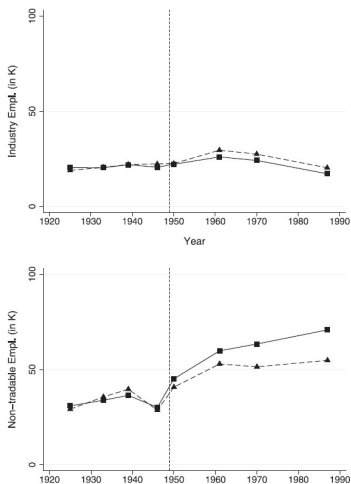
Source: Becker et al (2021), Figure 3. Synthetic: 1987 vs pre = +13,400 workers.  
⇒ Multiplier of public to private employment = 0.86.

## Total employment



Source: Becker et al (2021), Figure 3. Synthetic: 1987 vs pre = +29,000 workers.

## Tradable vs non-tradable private employment (synthetic control)



Source: Becker et al (2021), Figure 5.  $\uparrow 1$  public = +1.05 non-tradable, -0.19 tradable.

# Mechanisms

- Evidence for amenity spillovers in Bonn, proxied by cultural expenditures, theater budgets and visitors, number of artists, doctors, hospital beds.
- No evidence for productivity spillovers in Bonn, proxied by gross value added.
- Evidence for small increase in nominal wages in Bonn.
- Evidence for higher housing prices in Bonn.

# Mechanisms

- Evidence for amenity spillovers in Bonn, proxied by cultural expenditures, theater budgets and visitors, number of artists, doctors, hospital beds.
- No evidence for productivity spillovers in Bonn, proxied by gross value added.
- Evidence for small increase in nominal wages in Bonn.
- Evidence for higher housing prices in Bonn.
- Using moments from empirical analysis, estimate best fits for spillover parameters: productivity spillovers close to zero, amenities spillovers positive and very large.

# Amenities and disamenities

- Role of amenities and disamenities for city growth:
  - Location of baroque opera houses across Germany attracted high human capital workers with positive effects on city growth. [Falck et al \(2012\)](#)
  - Industrial pollution in late 19th century Britain reduced city growth through lower productivity of workers and firms. [Hanlon \(2020\)](#)

- Impact of trade or “market access” on the development of local economies.
  - Local impact of transportation infrastructure improvement on spatial distribution of economic activity in colonial India. [Donaldson \(2018\)](#)
  - Impact of railroads in 19th century US on local development and input factors reallocation.  
[Donaldson and Hornbeck \(2016\)](#) [Hornbeck and Rotemberg \(2019\)](#)
  - Long-run effects of Roman roads network across Europe for inter-regional trade. [Flückiger et al \(2019\)](#) [Dalgaard et al \(2018\)](#)

# Education and human capital

- Educational institutions:
  - Impact of medieval universities using founding of new universities after Papal Schism of 1386. [Cantoni and Yuchtman \(2014\)](#)
  - Founding of US universities on local innovation rates. [Andrews \(2020\)](#)
- Transfers of knowledge or skills across urban areas:
  - Access to the printing press in 15th century. [Dittmar \(2011\)](#)
  - Huguenot (skilled labor) Diaspora to Prussia after expulsion from France in 17th century. [Hornung \(2014\)](#)
- Education and industrialization:
  - Education and industrialization in 19th century Prussia. [Becker et al \(2011\)](#)
  - Knowledge elite (purchasers of Encyclopédie in mid-18th century France) and industrialization. [Squicciarini and Voigtländer \(2015\)](#)



# Spatial distribution of economic activity within cities

# Distribution of economic activity within cities

- Key questions with historical perspective:
  - Equilibrium distribution of residents, workers, and land rents.
  - Urban infrastructures (roads, sewers) and shape and size of cities.
  - Source of contemporaneous residential segregation.

# Distribution of economic activity within cities

- Key questions with historical perspective:
  - Equilibrium distribution of residents, workers, and land rents.
  - Urban infrastructures (roads, sewers) and shape and size of cities.
  - Source of contemporaneous residential segregation.
- Topics:
  - Transport infrastructure.
  - Urban health amenities.
  - Residential amenities.
  - Segregation.
  - Zoning and urban policies.

# Transportation Infrastructure

- Before industrial revolution, city structures constrained by distance (walking). [Voth \(2001\)](#).
- The first transportation revolution: railways and subways. [Bogart \(2014\)](#)
  - Constraint on city structure relaxed.
  - Employment and residence became increasingly segregated. [Heblich et al \(2020\)](#)
  - Led to highly localized production and residential externalities (e.g., Berlin using variation from division and reunification 1930s–2000s). [Ahlfeldt et al \(2015\)](#)

# Transportation Infrastructure

- The second transportation revolution: cars and trucks. Glaeser and Kohlase (2004)
  - First wave: trucks freed firms' location choices  $\implies$  sub-centers.
  - Second wave: urban highways freed residential choices  $\implies$  suburbanization (e.g., 1947 highway plan in U.S. as source of variation). Baum-Snow (2007)

# Transportation Infrastructure

- The second transportation revolution: cars and trucks. Glaeser and Kohlase (2004)
  - First wave: trucks freed firms' location choices  $\implies$  sub-centers.
  - Second wave: urban highways freed residential choices  $\implies$  suburbanization (e.g., 1947 highway plan in U.S. as source of variation). Baum-Snow (2007)
- Strong persistence of urban layout (e.g., coordinating effects of rectangular grids in cities, Manhattan since 1811). O'Grady (2014)

$\implies$  Rapidly expanding area of research!

# Transportation Infrastructure

- Baum-Snow (2007) “Did Highways Cause Suburbanization?”
- General research question: see the title.
- Paradox: 1950–1990, populations in central US cities ↓ 17% while population in metropolitan areas ↑ 72%.
- Contribution of new limited access highways?
- Answer: 1/3 of relative changes in population.

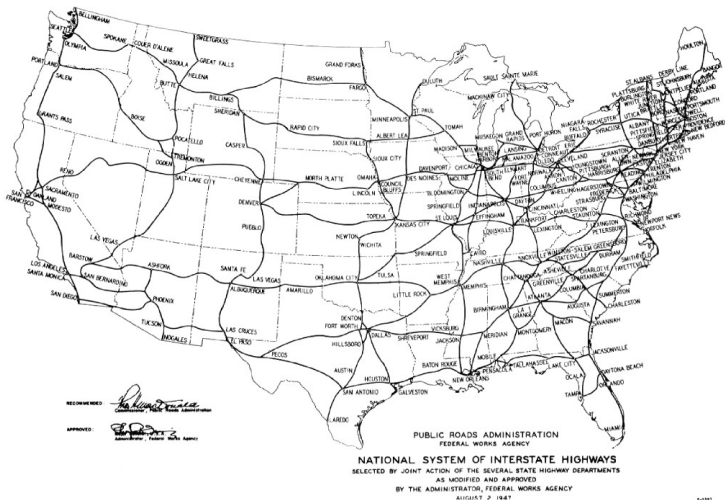
# The Interstate Highway System

- Federal Aid Highway Act of 1933: development of national system of interstate highways.
- In 1947, 37k miles of roadways approved by federal government.
- Funding mostly provided by 1956 Highway Revenue Acts.
- Database of yearly highway development 1950–1990.
- Primary measure: segments of roads connecting central business district of central city to region outside central city.
- 1947 plan not correlated with changes in commuting demand, but trade.
- Segments in the 1947 plan are strong predictors of constructed segments 1950–1990  $\implies$  used planned segments as instrument.



# The Interstate Highway System

The projected system of interstate highways in 1947



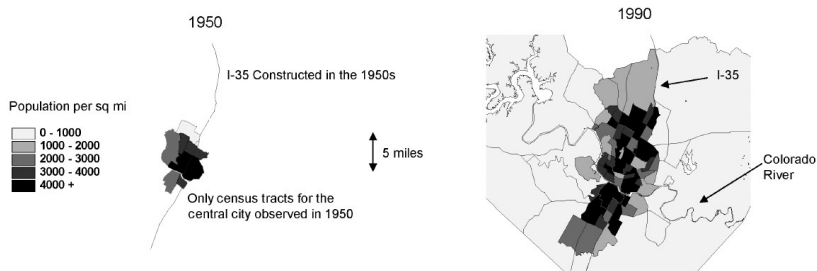
Source: Baum-Snow (2007), figure 1

# Conceptual Framework

- Land use theory: higher commuting speed implies lower population density.
- New highway segment  $\implies$   $\uparrow$  supply of land for given commuting time in section of city.
- Prices of rent  $\downarrow$  throughout metro area.
- Implications, when space is normal good:
  - Population in metro area spread out along new highways.
  - Central city population decreases with number of highways.

# Conceptual Framework

## Development patterns in Austin, TX



Note: Each shaded region is a separate census tract.

Source: Baum-Snow (2007), figure 2

# Empirical Strategy

- Long difference estimates (1950–1990):

$$\Delta \log N_i^c = \delta_0 + \delta_1 \Delta ray_i + \delta_2 r_{ci} + \delta_3 \Delta \tilde{w}_i + \delta_4 \Delta \log N_i^{MSA} + \delta_5 \Delta \tilde{G}_i + \varepsilon_i$$

- $N_i^c$ : central city population MSA  $i$ .
  - $ray_i$ : highway segment from central city to outside region.
  - $r_c$ : central city radius.
  - $N_i^{MSA}$ : MSA population.
  - $\tilde{w}_i$ : mean log annual personal income.
  - $\tilde{G}_i$ : GINI of income distribution of MSA.
- Panel estimates
    - Panel based on census years 1950 to 1990
    - Better account for timing of adjustments and speed of convergence.
    - But lower quality data for in-between years.

## Long-difference IV regressions

Large MSAs in 1950						
	Change in log population in constant geography central cities					
	OLS3	IV1	IV2	IV3	IV4	IV5
Change in number of rays	-.059 (.014)**	-.030 (.022)	-.106 (.032)**	-.123 (.029)**	-.114 (.026)**	-.101 (.046)*
1950 central city radius	.080 (.014)**		.111 (.023)**	.113 (.023)**	.106 (.023)**	.125 (.021)**
Change in simulated log income	.084 (.378)			.048 (.417)	-6.247 (6.174)	-.137 (.480)
Change in log of MSA population	.363 (.082)**			.424 (.094)**	.374 (.079)**	.405 (.108)**
Change in Gini coeff of simulated income					-23.416 (23.266)	
Log 1950 MSA population						-.062 (.062)
Constant	-.640 (.260)*	-.203 (.078)*	-.359 (.076)**	-.588 (.281)*	4.580 (5.091)	-.611 (.265)**
Observations	139	139	139	139	139	139
R-squared	.39	.00	.01	.30	.33	.37

Source: Baum-Snow (2017), table 4

## Panel IV regressions

	Large MSAs in 1950					
	Log central city population					
	1	2	3	4	5	6
Number of rays	-0.111 (0.016)**	-0.142 (0.026)**	-0.140 (0.028)**			
(1990 Rays) × (Fraction of Ray miles completed at $t$ )				-0.097 (0.016)**	-0.089 (0.012)**	-0.086 (0.013)**
Log simulated income		-0.083 (0.117)	-0.061 (0.109)		-0.288 (0.075)**	-0.229 (0.077)**
Log MSA population		0.266 (0.104)*	0.263 (0.105)*		0.294 (0.100)**	0.286 (0.098)**
Simulated Gini coefficient			-0.623 (1.106)			-1.415 (0.847)
MSA Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.20	0.22	0.22	0.14	0.56	0.57

Source: Baum-Snow (2017), table 6

# Urban health amenities

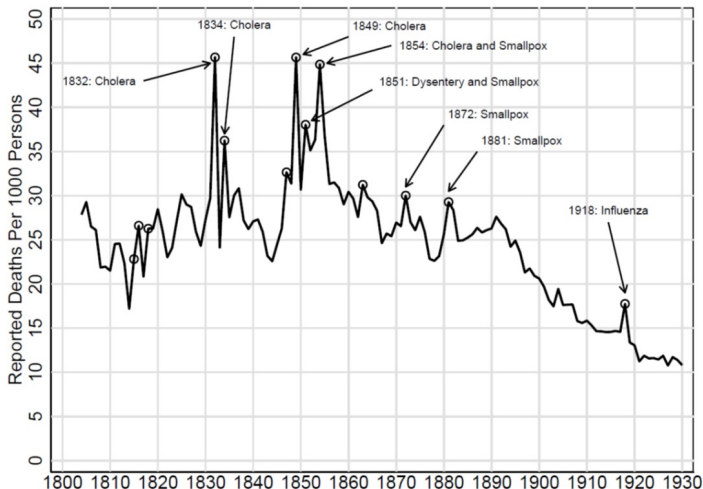
- Water infrastructures.
  - Equally important factor for internal organization of cities (and health distribution): plumbing and sewer systems.
  - Before sewerage infrastructure, endemic cholera epidemics, aggravated by industrialization and increased density.  
[Ferrie and Troesken \(2008\)](#) [Kesztembaum and Rosenthal \(2011\)](#)
  - Review of water infrastructure and urban mortality in the U.S.  
[Beach \(2021\)](#)

# Urban health amenities

- Water infrastructures.
  - Equally important factor for internal organization of cities (and health distribution): plumbing and sewer systems.
  - Before sewerage infrastructure, endemic cholera epidemics, aggravated by industrialization and increased density.  
[Ferrie and Troesken \(2008\)](#) [Kesztembaum and Rosenthal \(2011\)](#)
  - Review of water infrastructure and urban mortality in the U.S.  
[Beach \(2021\)](#)
- Air pollution.
  - Industrialization increased revenues but at dramatic health cost (e.g., coal smoke and population growth in Britain 1851–1911). [Hanlon \(2020\)](#)
  - Interactions between urban pollution and spread of epidemics (e.g., Spanish Influenza in U.S. cities). [Clay and Troesken \(2011\)](#)



## Epidemics in New York City



# Residential amenities

- Factors influencing amenity value of neighborhoods:
  - Public water systems (plumbing and sewers).
  - Pollution.
  - Public safety (e.g., professionalization of police force in early 19th century-London). [Bindler and Hjalmarsson \(2019\)](#)
  - Natural amenities (lakes, mountains). [Lee and Lin \(2018\)](#)
  - Urban infrastructures (roads, public transportation, opera houses, architecture...). [Ahlfeldt and Holman \(2018\)](#)
- Amenities have strong persistence influences on urban segregation and distribution of income within cities.

# Urban segregation

- Very large literature on origins and lasting effects of racial segregation within cities, especially in the U.S.
  - Neighborhood-level segregation doubled 1880–1940. [Logan and Parman \(2017\)](#)
  - Segregation accelerated 1940–1970 with white flight from city centers to suburbs and black migration from rural South to industrial North. [Boustan \(2010\)](#) [Shertzer and Walsh \(2019\)](#)
  - Important implications for neighborhood-level provision of public goods.

# Urban segregation

- Very large literature on origins and lasting effects of racial segregation within cities, especially in the U.S.
  - Neighborhood-level segregation doubled 1880–1940. [Logan and Parman \(2017\)](#)
  - Segregation accelerated 1940–1970 with white flight from city centers to suburbs and black migration from rural South to industrial North. [Boustan \(2010\)](#) [Shertzer and Walsh \(2019\)](#)
  - Important implications for neighborhood-level provision of public goods.
- Role of transportation infrastructures:
  - Development of transportation facilitated / perpetuated segregation (e.g., highways and suburbanization).
  - Highways generated little benefits to (black) housing owners in city centers, but decreased land value (lower rents, increased pollution and noise). [Brinkman and Lin \(2020\)](#)

# Zoning and urban policies

- Most prominent urban policy (in U.S.): zoning. [Fischel \(2004\)](#)
  - First wave early 20th century: protect home owners from industrial development when urban mobility improved.
  - Persistent implications for city structures (e.g., Chicago zoning ordinance of 1923). [Shertzer et al. \(2018\)](#)

# Zoning and urban policies

- Most prominent urban policy (in U.S.): zoning. [Fischel \(2004\)](#)
  - First wave early 20th century: protect home owners from industrial development when urban mobility improved.
  - Persistent implications for city structures (e.g., Chicago zoning ordinance of 1923). [Shertzer et al. \(2018\)](#)
- Relationship between land development and urban policies after fires: Boston 1873 and San Francisco 1906.  
[Hornbeck and Keniston \(2017\)](#) [Siodla \(2015, 2017\)](#)
- Other restrictions: vertical development in U.S. cities (Chicago, New York) and London after WWII bombings.  
[Barr \(2013\)](#) [Ahlfeldt and McMillen \(2018\)](#) [Koster and Dericks \(2020\)](#)

# The Great Boston Fire of 1872

- [Hornbeck and Keniston \(2017\)](#) “Creative Destruction: Barriers to Urban Growth and the Great Boston Fire of 1872.”
- Research question: is urban growth restricted when landowners do not internalize positive spillover effects?
- Use Boston Fire of 1872 as a source of widespread simultaneous reconstruction, encouraging upgrades in nearby buildings.
  - Fire destroyed 776 buildings in downtown Boston in 22 hours (11% of total real estate value of Boston).
  - Reconstruction managed privately (non zoning law or city coordination).
  - Fire in commercial neighborhood: upgrades in economic spaces.

# The Great Boston Fire of 1872

Historical downtown Boston and the burned area





# The Great Boston Fire of 1872

- Plot-level records:
  - Land and building value and occupant's industry and business value at plot level for 1867–1894.
  - Plots in burned and surrounding areas.
  - 44k land plots pooled over 7 years.

# The Great Boston Fire of 1872

- Plot-level records:
  - Land and building value and occupant's industry and business value at plot level for 1867–1894.
  - Plots in burned and surrounding areas.
  - 44k land plots pooled over 7 years.
- Pre-fire values in burned and unburned areas:
  - Burned area was most valuable parts of downtown Boston.
  - Burned areas more valuable prior to the Fire.
  - But no difference in pre-Fire values trends across burned and unburned areas.
  - “Close” vs “distant” sample with 400 meters cutoff.

# The Great Boston Fire of 1872

- Theoretical framework to assess cross-building externalities:
  - Each landowner owns one (homogeneous) plot.
  - Maximizes net present value of rents from plot  $r(q, Q, \omega)$ 
    - $q$ : building quality.
    - $Q$ : average quality of nearby buildings.
    - $\omega$ : city overall productivity.
  - Spillovers 1: quality of nearby buildings  $\partial r / \partial Q > 0$ .
  - Spillovers 2: marginal returns to building quality wrt nearby quality  $\partial^2 r / \partial q \partial Q > 0$ .
- Direct spillover and multiplier effects in burned and unburned areas.

# The Great Boston Fire of 1872

- Cost of replacing building of quality  $q'$ :  $c(q')$ .
- Buildings destroyed with probability  $d$ .
- Total value is Bellman equation:

$$V(q, Q, \omega) = \max \begin{cases} r(q, Q, \omega) + \delta \mathbb{E}[V(q, Q', \omega')] \\ r(q^*, Q, \omega) + \delta \mathbb{E}[V(q^*, Q', \omega')] - c(q^*) \end{cases}$$

- where  $q^*$  maximizes:

$$r(q, Q, \omega) + \delta [(1 - d)V(q, Q', \omega') + dV(0, Q, \omega')] - c(q^*)$$

- $q^*$ : optimal choice of building quality when choosing to replace old building or forced to reconstruct when destroyed.

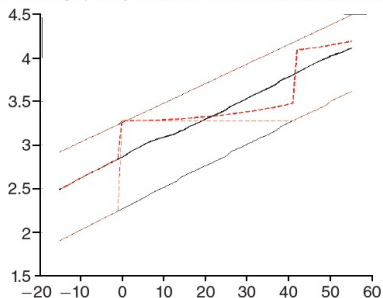
# The Great Boston Fire of 1872

- Absent of destruction, landowners replace building when obsolete ( $\underline{q}$ ) and city productivity has increased enough.
- Total value of owning building is composed of:
  - Land value component:  $V(0, Q, \omega) = V(\underline{q}, Q, \omega)$
  - Building value component:  $V(q, Q, \omega) - V(\underline{q}, Q, \omega)$ .
- Building rents have Cobb-Douglas form:  $r(q, Q, \omega) = q^\alpha (Q^\eta \omega^{1-\eta})^\beta$ .
  - No spillovers:  $\eta = 0$
  - Spillovers:  $\eta > 0$ , with  $\eta = 0.8$  for simulations.

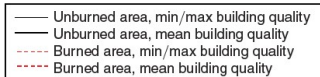
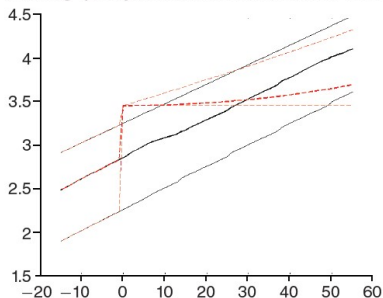
# The Great Boston Fire of 1872

## Simulated changes in building quality and land value after Great Fire

Panel A. Without cross-plot spillovers:  
Building quality in burned and unburned areas



Panel A. With cross-plot spillovers:  
Building quality in burned and unburned areas



Source: Hornbeck and Keniston (2017), figure 3

# The Great Boston Fire of 1872

- Testable predictions:
  - Fire increases plot land values in burned regions and land values converge over time.
  - Fire increases land values and buildings in nearby unburned areas.

# The Great Boston Fire of 1872

- Empirical strategy, plots 1876–1894:

$$Y_{it} = \alpha_t + \rho I_i^{\text{Fire}} + \beta_{1873} I_i^{\text{Fire}} \times I_t^{1873} + \beta_{1882} I_i^{\text{Fire}} \times I_t^{1882} + \beta_{1894} I_i^{\text{Fire}} \times I_t^{1894} + \varepsilon_{it}$$

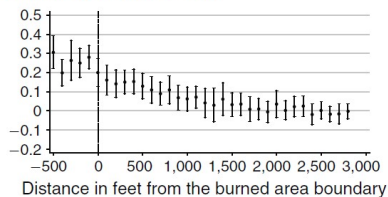
- $I_i^{\text{Fire}}$ : indicator for within burned area.
- $\alpha_t$ : year FE.
- Also include interaction between year FE and pre-Fire values.



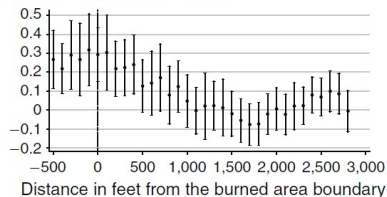
# The Great Boston Fire of 1872

Estimated impacts on land value in 1873 and 1882 by distance to fire

Panel A. Land value in 1873



Panel B. Land value in 1882

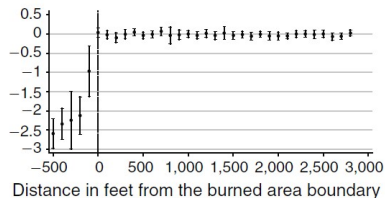


Source: Hornbeck and Keniston (2017), figure 4

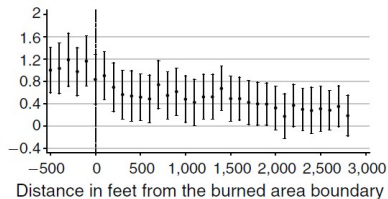
# The Great Boston Fire of 1872

Estimated impacts on building value in 1873 and 1882 by distance to fire

Panel C. Building value in 1873



Panel D. Building value in 1882

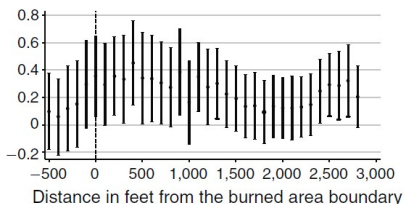


Source: Hornbeck and Keniston (2017), figure 4

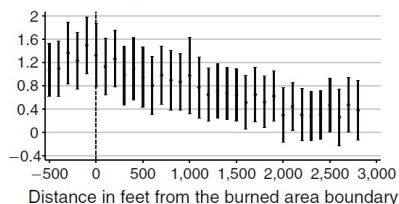
# The Great Boston Fire of 1872

Estimated impacts on building and land value in 1894 by distance to fire

Panel A. Land value in 1894



Panel B. Building value in 1894



Source: Hornbeck and Keniston (2017), figure 5

# Cities and growth

# Urbanization as an indicator of growth

- When GDP estimates nonexistent, urbanization or population density can be proxy for growth:
  - Link between economic and population growth is key prediction of Malthusian model.
  - Malthusian trap: technological progress absorbed by population growth. [Ashraf and Galor \(2011\)](#)

# Urbanization as an indicator of growth

- When GDP estimates nonexistent, urbanization or population density can be proxy for growth:
  - Link between economic and population growth is key prediction of Malthusian model.
  - Malthusian trap: technological progress absorbed by population growth. [Ashraf and Galor \(2011\)](#)
- Alternative: local agricultural productivity. [Bairoch \(1988\)](#) [Dittmar \(2011\)](#)
  - Cost of supplying food to cities key for city size.
  - Technological progress or fall in transportation cost translates in urbanization.

# Contribution of cities to growth

- Cities are central contributors to technological development.
  - Direction of causality between urbanization and productivity growth (e.g., impact of new Swedish towns on local productivity).  
[Carmeno and Enflo \(2019\)](#)
  - Some theoretical framework developed.  
[Desmet and Rossi-Hansberg \(2014\)](#) [Desmet et al \(2018\)](#)
- Trade-off between optimal city size and technological change (e.g., guilds as barrier to entry). [Desmet and Parente \(2014\)](#)

# Topic 5

## Economics history and culture



# Economic History and Culture

- Research questions involving culture often require historical perspective:
  - How can we explain observed differences in culture?
  - Mechanisms through which culture affects behavior?
  - How does culture transmit across generations?
  - When does culture persist over time? When doesn't it?
  - What are the mechanisms of cultural change?
  - How does culture emerge?
- Specific (and standard) empirical issues still apply.

# Economic History and Culture

- Explain contemporaneous distribution of norms about gender roles: pre-industrial division of labor. [Alesina, Giuliano, and Nunn \(2013\)](#)

# Economic History and Culture

- Explain contemporaneous distribution of norms about gender roles: pre-industrial division of labor. [Alesina, Giuliano, and Nunn \(2013\)](#)
- Explain contemporaneous distribution of trust in Africa: the slave trade. [Nunn and Wantchekon \(2011\)](#)

# Economic History and Culture

- Explain contemporaneous distribution of norms about gender roles: pre-industrial division of labor. [Alesina, Giuliano, and Nunn \(2013\)](#)
- Explain contemporaneous distribution of trust in Africa: the slave trade. [Nunn and Wantchekon \(2011\)](#)
- Explain contemporaneous distribution of economic status in Africa: ancestral lifeways. [Michalopoulos, Putterman and Weil \(2019\)](#)

# Economic History and Culture

- Explain contemporaneous distribution of norms about gender roles: pre-industrial division of labor. [Alesina, Giuliano, and Nunn \(2013\)](#)
- Explain contemporaneous distribution of trust in Africa: the slave trade. [Nunn and Wantchekon \(2011\)](#)
- Explain contemporaneous distribution of economic status in Africa: ancestral lifeways. [Michalopoulos, Putterman and Weil \(2019\)](#)
- Mechanisms of persistence of historical shocks: contemporaneous gender norms in Africa and the slave trade. [Teso \(2019\)](#)

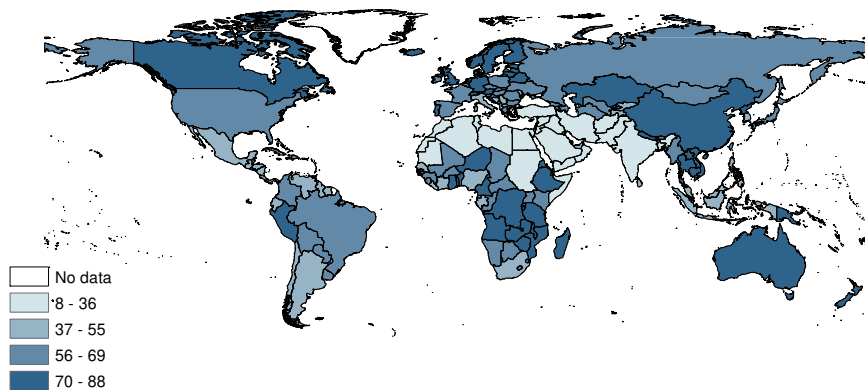
# Economic History and Culture

- Explain contemporaneous distribution of norms about gender roles: pre-industrial division of labor. [Alesina, Giuliano, and Nunn \(2013\)](#)
- Explain contemporaneous distribution of trust in Africa: the slave trade. [Nunn and Wantchekon \(2011\)](#)
- Explain contemporaneous distribution of economic status in Africa: ancestral lifeways. [Michalopoulos, Putterman and Weil \(2019\)](#)
- Mechanisms of persistence of historical shocks: contemporaneous gender norms in Africa and the slave trade. [Teso \(2019\)](#)
- Mechanisms of persistence of historical shocks: contemporaneous and past gender norms in France and WWI. [Gay \(2023\)](#)

# This Session

- Focus on method to identify culture: the epidemiological approach.
- Main example: traditional gender roles and FLFP.
- Leave aside important aspects:
  - Economic theory of preference formation and cultural transmission.  
[Bisin and Verdier \(2011\)](#)
  - Interactions between culture and institutions.  
[Alesina and Giuliano \(2015\)](#)
- Then apply the method to economic history questions

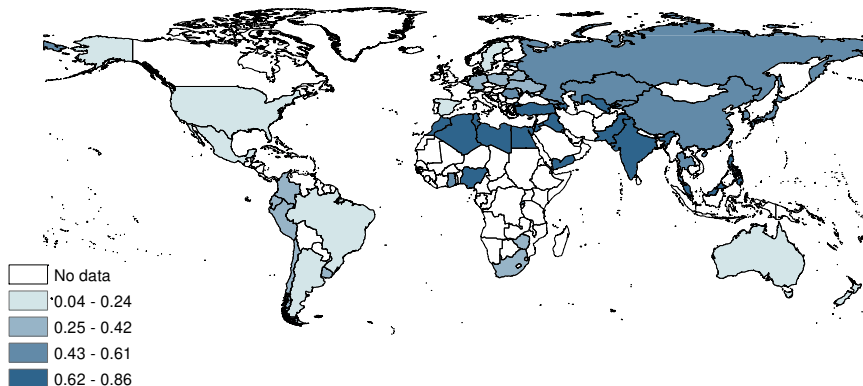
# Distribution of FLFP across Countries



- FLFP 15–65 (2010–2014) across 187 countries [ILO Estimates](#)

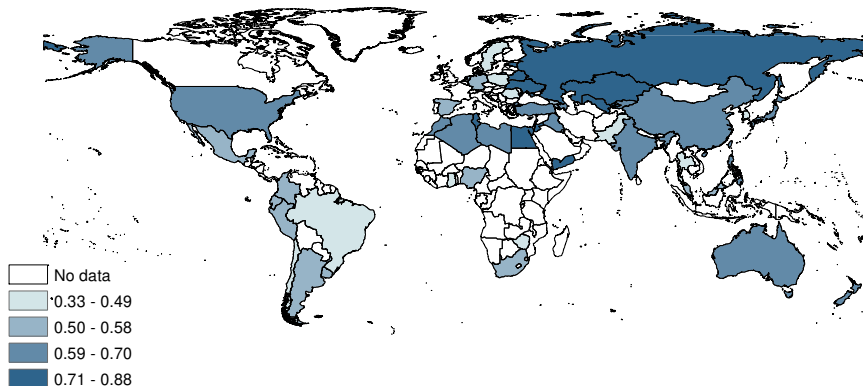


# Distribution of Attitudes across Countries



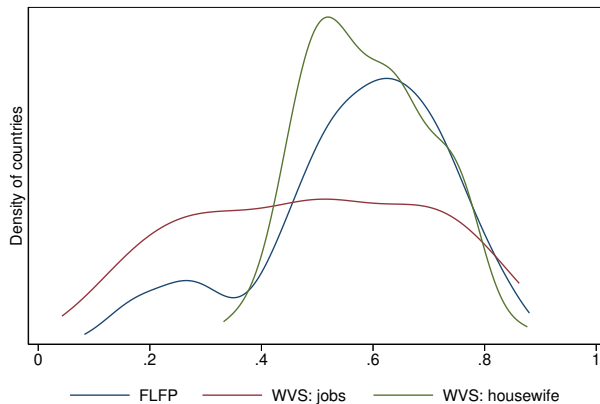
- Share approves of "When jobs are scarce, men should have more right to a job than women" across 58 countries [WVS, Wave 6, 2010–2014](#)

# Distribution of Attitudes across Countries



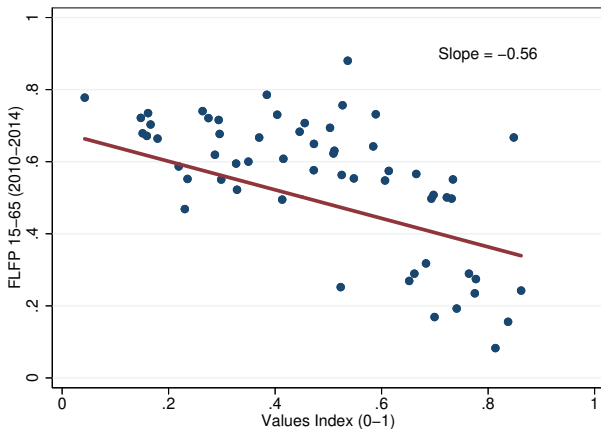
- Share approves of “Being a housewife is just as fulfilling as working for pay” across 58 countries [WVS, Wave 6, 2010–2014](#)

# Distribution of FLFP and Attitudes across Countries



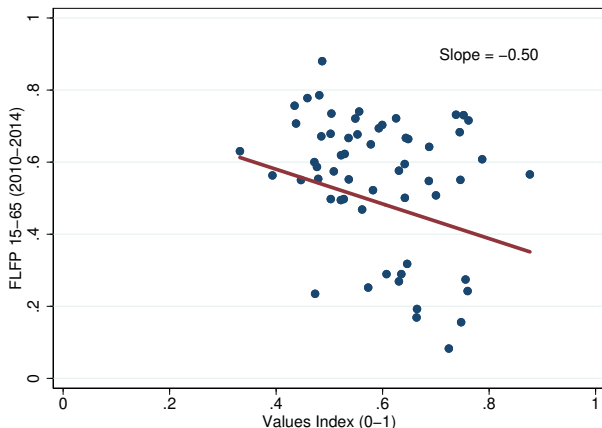
Source: WVS, ILO.

# Correlation between Attitudes and FLFP



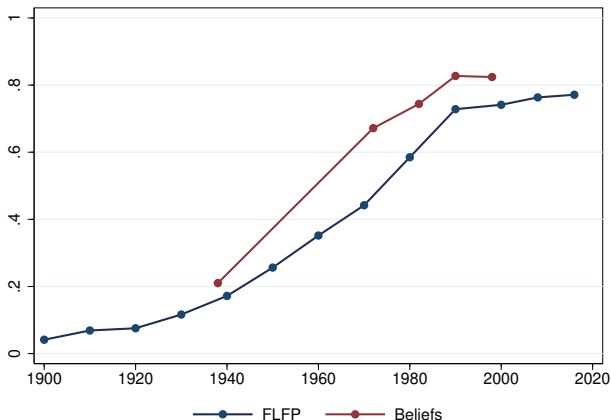
- Values: share approves of “When jobs are scarce, men should have more right to a job than women”

# Correlation between Attitudes and FLFP



- Values: share that approves of “Being a housewife is just as fulfilling as working for pay”

# FLFP and Attitudes Across Time in the U.S.



- FLFP sample: white married women aged 30–49 in nonfarm [IPUMS 1900–2016](#)
- Share approves of “A married woman earning money in business or industry if she has a husband capable of supporting her” [Gallup 1938](#) [GSS 1972–1998](#)

# Correlation between Attitudes and Behaviors

Holds across many dimensions e.g.:

- Gender roles and fertility, distribution of household tasks.
- Family ties and civic participation, geographical mobility.
- Thrift and saving rates.
- Belief in luck and redistribution.
- Trust and civic participation, shirking, financial participation.

# Research Questions

- Are these correlations:
  - Driven by systematic cultural differences?
  - Driven by economic and institutional differences?
- Traditional view:
  - Economics: explain  $\Delta$  in behaviors by  $\Delta$  in prices and income.
  - Stigler and Becker (1977) *De Gustibus Non Est Disputandum*



# Research Questions

- Is there credible empirical evidence that:
  - Culture matters for economic behavior?
  - Cultural change matters for changes in economic behavior?

# Research Questions

- Is there credible empirical evidence that:
  - Culture matters for economic behavior?
  - Cultural change matters for changes in economic behavior?
- If culture matters, many “new” economic questions:
  - Mechanisms through which culture affects behavior?
  - How does culture transmit across generations?
  - When does culture persist over time? When doesn't it?
  - What are the mechanisms of cultural change?
  - How does culture emerge?

# Measuring culture

# Definition of Culture

*Culture is the integrated pattern of human knowledge, belief, and behavior that depends upon the capacity for learning and transmitting knowledge to succeeding generations, and the customary beliefs, social forms, and material traits of a racial, religious group, or social group; and the set of shared attitudes, values, goals, and practices that characterizes an institution or organization.*

*(Webster Dictionary)*

- Broadly, culture is a body of shared knowledge and practices.

# Definition of Culture

- Need practical definition for empirical analysis.
- Identification: more useful to think about *differences in culture*.
- Differences in culture are:
  - Systematic variations in preferences and beliefs.
  - Across time, space, and/or social groups.
- Most common criteria: space. But not always most relevant.
- No attempt to differentiate preferences, expectations, beliefs.  
Bicchieri (2006, 2016)

# Definition of Culture

Alternative approach:

- Culture as a focal factor. [Myerson \(2009\)](#)
- Social structures characterized as a game with multiple equilibria.
- Culture is a mechanism of equilibrium selection and coordination.
- No need for differences in preferences.

# Definition of Culture

Alternative approach:

- Culture as a focal factor. [Myerson \(2009\)](#)
- Social structures characterized as a game with multiple equilibria.
- Culture is a mechanism of equilibrium selection and coordination.
- No need for differences in preferences.
- Related to the emergence of culture?
- Challenging to provide empirical evidence (analytical narratives).

# Cross-Cultural Surveys: Datasets

- World-wide:
  - World Values Survey: 6 waves, 1981–2014, up to 60 countries.
  - International Social Survey: 4 waves, 1985–2016, up to 43 countries.
  - Global Preferences Survey: 76 countries (2012). [Falk et al. \(2018\)](#)
- Regional:
  - European Values Study: 4 waves, 1981–2009, up to 47 countries.
  - Demographic and Health Surveys: 1985–2017, 90 countries.
  - Afro Barometer (1999–2019), Latino Barometer (1995–2017), Asian Barometer (2001–2016).
- Country:
  - United States: General Social Survey (1976–2017).
  - France: DREES, ECAM, CREDOC...



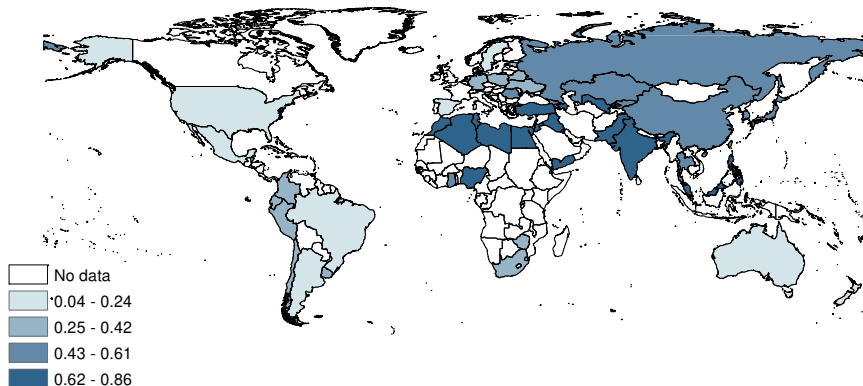
# Cross-Cultural Surveys: Cultural Traits

- Trust.
- Gender roles.
- Time preference.
- Individualism vs collectivism.
- Family ties.
- Morality.
- Work and Poverty.

# Measurement Challenges: Data Availability

- Across countries.
  - WVS: 60 countries in wave 6 (2010–2014).
  - GPS: 76 countries, 90% of world population.
  - Cultures as units matter.
- Across time.
  - WVS: wave 1 (1981–1984), 10 countries.
  - GSS: since 1976, but only for the U.S.
  - Culture is a long-run process, changes over multiple generations.
- Within countries.
  - WVS: wave 6 (2010–2014), average 13 regions per country.
  - GPS: state/province of residence.
  - Usually: no precise residence location, no birth location.

# Data Availability across Countries



- Share approves of “When jobs are scarce, men should have more right to a job than women” across 58 countries [WVS, Wave 6, 2010–2014](#)

# Measurement Challenges: Aggregation Issues

- Question level:
  - “When jobs are scarce, men should have more right to a job than women”
    - Agree: 1→1.
    - Neither: 2→0.5.
    - Disagree: 3→0.
  - “Being a housewife is just as fulfilling as working for pay”
    - Strongly agree: 1→1.
    - Agree: 2→0.66.
    - Disagree: 3→0.33.
    - Strongly disagree: 4→0.
  - Several issues with Likert scales.
- Index level: linear, PCA, ... ?
- Country level: flat or sample weights?

# Measurement Challenges: Fundamental Issues

- What is measured?
  - Expectations, beliefs (1st, 2nd order), preferences, attitudes?
  - GPS: cross-validated with experiments (but rare).
- Comparing cultures
  - Do universal categories of culture exist?
  - Cultural biases (ethnocentrism).

# Potential Solutions

- Past behavior e.g. past FLFP
  - Captures culture + economic and institutional conditions.
  - Data comparability and coverage.
- (Deep) historical measures
  - E.g. Pre-modern division of labor, language structures.
  - Fixed over time.
  - Measurement issues.
  - Little within-country variation.
  - Assumptions on historical processes (persistence, transmission).

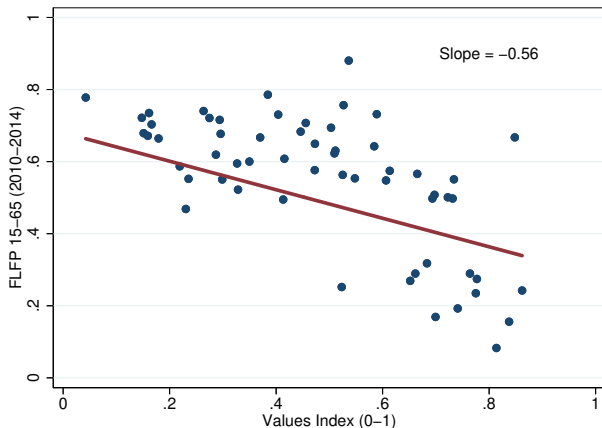
# Potential Solutions

- Criteria for measure:
  - Consistent across time, at least across 3 generations.
  - Consistent across space, especially within a country.
  - Data availability.
  - Represent local preferences and beliefs.
- Some ideas:
  - Textual analysis of books and newspapers. [Giorcelli et al. \(2019\)](#)
  - Legislative behaviors (voting, activities, speeches).

# Culture and behavior



# Cross-Country Approach



- Values: share approves of “When jobs are scarce, men should have more right to a job than women.” [WVS, Wave 6, 2010–2014](#)

# Cross-Country Approach

- Outcome: FLFP 15–65 (2010–2014). ILO
- Cultural variable: approves of “When jobs are scarce, men should have more right to a job than women” WVS
- (Very rough) controls (more later): area, GDP, schooling, fertility, migration, population.

Dependent variable:	FLFP 15–65	
	(1)	(2)
WVS question (0–1)	-0.61*** [0.08]	-0.82*** [0.16]
Controls	No	Yes
Continent FE	No	Yes
Countries	51	51
Adj. R2	0.477	0.502

# Cross-Country Approach

Results lack credibility:

- Reverse causality.
- Endogeneity with economic incentives and institutional structures.
- Few countries available in surveys.
- No data for past attitudes.

# The Epidemiological Approach

- Identify culture by fixing economic and institutional factors.  
Fernández (2011)

# The Epidemiological Approach

- Identify culture by fixing economic and institutional factors.  
Fernández (2011)
- Compare individuals
  - Within same institutional and economic environment.
  - Whose preferences and beliefs might differ systematically.

# The Epidemiological Approach

- Identify culture by fixing economic and institutional factors.  
Fernández (2011)
- Compare individuals
  - Within same institutional and economic environment.
  - Whose preferences and beliefs might differ systematically.
- Compare movers (1st, 2nd, higher order generations).
- (Testable) assumptions:
  - Cultural beliefs persist across generations.
  - Cultural beliefs vary systematically across groups (spatially defined?).
  - Individuals in the same area face identical institutional environment.

# The Epidemiological Approach: Main Issues

- Origin culture might attenuate over time.
- Institutional environment might vary within host environment.
- Movers are a selected sample.

# The Epidemiological Approach: A Simple Model

- Gender roles and married women LFP.
- Married women  $i$  from country  $k$  maximizes:

$$U(c, v_i) = u(c) - \mathbb{1}v_i$$

- $u(\cdot)$ : strictly increasing concave utility function.
- $\mathbb{1}$ : indicator for working.
- $c$ : household consumption.
- $v_i$ : disutility from working.

$$c = w_{hk} + \mathbb{1}w_{fk}$$

- $w_f$ : wife labor income if works.
- $w_h$ : husband's income (always works).



# The Epidemiological Approach: A Simple Model

Sources of heterogeneity:

- Wages:
  - Identical and exogenous within countries.
  - (Potentially) heterogeneous across countries and sex.
- Disutility of work:
  - Drawn from country-specific distribution  $G_k(m_k, \sigma)$ .
  - Varies across women within a country.
  - Varies across countries.
- Differences in culture: differences in mean  $m_k$ .
- Approximation: country  $\simeq$  culture.

# The Epidemiological Approach: A Simple Model

- $L_k$ : FLFP in country  $k$  is given by

$$L_k = \Pr(v_k \leq v_k^*) = G_k(v_k^*)$$

$$\text{with } v_k^* = v^*(w_{hk}, w_{fk}) = u(w_{hk} + w_{fk}) - u(w_{hk})$$

- Suppose  $G \sim \mathcal{N}(m_k, \sigma^2)$ . Then:

$$L_k = G_k(v_k^*) = \Phi\left(\frac{v_k^* - m_k}{\sigma}\right)$$

# The Epidemiological Approach: A Simple Model

- Culture matters through  $m_k$ :  $\uparrow m_k \implies \downarrow L_k$ .

$$\frac{\partial L_k}{\partial m_k} = -\frac{1}{\sigma} \phi \left( \frac{v_k^* - m_k}{\sigma} \right) < 0$$

- Institutional framework matters through wages and thus  $v_k^*$ .

$$\frac{\partial L_k}{\partial w_{fk}} = \frac{1}{\sigma} \phi \left( \frac{v_k^* - m_k}{\sigma} \right) u' (w_{hk} + w_{fk}) > 0$$

$$\frac{\partial L_k}{\partial w_{hk}} = \frac{1}{\sigma} \phi \left( \frac{v_k^* - m_k}{\sigma} \right) [u' (w_{hk} + w_{fk}) - u' (w_{hk})] < 0$$

# The Epidemiological Approach: A Simple Model

- Women from different countries in same environment and same husbands:
  - Same  $w_{hi}$  and same  $w_{fi}$ .
  - Same threshold  $v^*$ .
- But draw from different distributions  $G_k$ .
- (Assumption: culture persists for higher order generations of movers.)
- Proportion of women from (ancestry)  $k$  who work in country  $j$  is:

$$L_{kj} = \Phi \left( \frac{v_j^* - m_k}{\sigma} \right)$$

- Structural estimation approach left aside  
Fernández (2013) Fogli and Veldkamp (2011)

# A Proxy for Culture

Cultural differences  $m_k$  not observable, so need a proxy.

- Measure of attitudes: measurement issues.
- Economic behaviors directly.
  - Reference group behavior e.g ancestry FLFP.
  - Captures culture + economic and institutional conditions.
  - Leverage economic and institutional conditions are fixed (same  $v^*$ ).
  - Left with cultural component.
  - (Relatively) easy to implement.
- (Deep) historical measures:
  - Measurement challenges.
  - Potentially (much more) meaningful for identification.

# A Proxy for Culture

Further issues common to all measures of culture:

- Other (endogenous) sources of heterogeneity beyond culture.
- Proxy might capture inherited social capital.
- Which dimension of culture is captured?

# Application: Gender Roles and FLFP

- Case: FLFP of Immigrants to the United States.
- [Fernández and Fogli \(2009\)](#) [Blau \(2015\)](#)
- Data: ACS 2005–2017 [IPUMS](#)
- Sample selection:
  - Born outside the U.S. from non US parents.
  - Age: 30–49.
  - Regular household, no farms, married spouse present.
  - Born in single identified country. [BPL](#)

⇒ 330k women from 142 countries.

# Cultural Proxy: Country of Birth FLFP

- Data: International Labor Organization [ILO](#)
- Variables:
  - FLFP or ratio FLFP to MLFP.
  - Age bounds: 15+ or 15–65.
- Assignment: year or decade of migration (1956–2017).
- Country coverage: 7 countries missing (635 obs).
- Time coverage:
  - [ILO Estimates](#): 1990–2017. Consistent. Drop 50k obs (15%).
  - ILO Yearly Indicators: 1950–2017. Not consistent.
    - 15+: drop 89k obs (27%), 10k (3%) if interpolated.
    - 15–65: drop 112k obs (34%), 16k (4%) if interpolated.



# Standard Empirical Issues

- 1 Data quality vs data coverage.
- 2 Controls and comparison groups.
- 3 Interpretation: magnitude and precision.
- 4 Clustering.
- 5 Sources of identification.
- 6 Outliers.
- 7 Choice of cultural proxy.
- 8 Choice of outcome.
- 9 Regression model.

# Standard Empirical Issues

These might be helpful:

- [Gentzkow and Shapiro \(2014\) Code and Data for the Social Sciences](#)
- [tips4economists](#)
- [Beamer advice](#)
- In general: document everything!
- Replication will be very important in the future. [Maniadis, Tufano and List \(2017\)](#)
- BTW: register to [NEP](#) mailing list.

# 1. Data Quality vs Data Coverage

- Cultural Proxy: FLFP 15+
  - ILO Estimates (EST). 1990–2017, 50k missing.
  - ILO Yearly Indicators, raw (YI). 1956–2017, 89k missing.
  - ILO Yearly Indicators, interpolated (YII). 1956–2017, 10k missing.
- Outcome: indicator for *In the labor force*.
- Timing of assignment: year of migration.
- Level of aggregation: State.
- No controls.
- Two-way clustering: State + COB.

# 1. Data Quality vs Data Coverage

Dependent variable:	Active		
	(1)	(2)	(3)
COB FLFP 15+	0.48*** [0.05]	0.29*** [0.07]	0.27*** [0.08]
ILO variable	EST	YI	YII
Time coverage	1990–2017	1956–2017	1956–2017
COB coverage	135	132	133
Observations	280,003	240,333	320,044

- Attenuation bias:
  - (1)  $\rightarrow$  (2) =  $-0.19$  ( $-40\%$ )
  - (1)  $\rightarrow$  (3) =  $-0.21$  ( $-44\%$ )
- But regression samples not comparable.

# 1. Data Quality vs Data Coverage

Dependent variable:	Active				
	(1)	(2)	(3)	(4)	(5)
COB FLFP 15+	0.48*** [0.05]	0.29*** [0.07]	0.37*** [0.06]	0.27*** [0.08]	0.41*** [0.06]
ILO variable	EST	YI	YI	YII	YII
Time coverage	1990–2017	1956–2017	1990–2017	1956–2017	1990–2017
COB coverage	135	132	132	133	133
Observations	280,003	240,333	222,292	320,044	272,104

- Different sample:
  - (2)  $\rightarrow$  (3) = +0.08 (+28%)
  - (4)  $\rightarrow$  (5) = +0.14 (+52%)

# 1. Data Quality vs Data Coverage

Dependent variable:	Active				
	(1)	(2)	(3)	(4)	(5)
COB FLFP 15+	0.48*** [0.05]	0.29*** [0.07]	0.37*** [0.06]	0.27*** [0.08]	0.41*** [0.06]
ILO variable	EST	YI	YI	YII	YII
Time coverage	1990–2017	1956–2017	1990–2017	1956–2017	1990–2017
COB coverage	135	132	132	133	133
Observations	280,003	240,333	222,292	320,044	272,104

- Attenuation bias:
  - (1)  $\rightarrow$  (3) =  $-0.11$  ( $-23\%$ )
  - (1)  $\rightarrow$  (5) =  $-0.07$  ( $-15\%$ )
- But regression samples still not comparable.

# 1. Data Quality vs Data Coverage

Dependent variable:	Active		
	(1)	(2)	(3)
COB FLFP 15+	0.48*** [0.05]	0.45*** [0.07]	0.37*** [0.06]
ILO variable	EST	EST	YI/YII
Time coverage	1990–2017	1990–2017	1990–2017
COB coverage	135	132	132
Observations	280,003	222,292	222,292

- Different sample: (1)  $\rightarrow$  (2) =  $-0.03$  ( $-6\%$ )
- Attenuation bias: (2)  $\rightarrow$  (3) =  $-0.08$  ( $-18\%$ )

# 1. Data Quality vs Data Coverage

Dependent variable:	Active			
	(1)	(2)	(3)	(4)
COB FLFP 15+	0.45*** [0.07]	0.38*** [0.07]	0.38*** [0.07]	0.26*** [0.06]
ILO variable	EST	YI/YII	EST	YI/YII
Time coverage	1990–2017	1990–2017	1990–2017	1990–2017
COB coverage	109	109	109	109
Controls	No	No	Yes	Yes
Observations	218,832	218,832	218,832	218,832

- Attenuation bias:
  - No controls: (1)  $\rightarrow$  (2) =  $-0.07$  ( $-16\%$ )
  - Controls: (3)  $\rightarrow$  (4) =  $-0.12$  ( $-32\%$ )



# 1. Data Quality vs Data Coverage

How different are these estimates?

- Ideally, estimate covariance matrix via seemingly unrelated regression.
- Shortcut:
  - Take the estimates difference.
  - Compute s.e. around this difference:  $\sqrt{se_1^2 + se_2^2}$ .
  - Assumes zero covariance for *parameter* estimates.
  - Good approximation.

# 1. Data Quality vs Data Coverage

Dependent variable:	Active			
	(1)	(2)	(3)	(4)
COB FLFP 15+	0.45*** [0.07]	0.38*** [0.07]	0.38*** [0.07]	0.26*** [0.06]
ILO variable	EST	YI/YII	EST	YI/YII
Time coverage	1990–2017	1990–2017	1990–2017	1990–2017
COB coverage	109	109	109	109
Controls	No	No	Yes	Yes
Observations	218,832	218,832	218,832	218,832

- Attenuation bias:
  - No controls: (1)  $\rightarrow$  (2) =  $-0.07$ , s.e.  $\simeq 0.96$ .
  - Controls: (3)  $\rightarrow$  (4) =  $-0.12$ , s.e.  $\simeq 0.95$ .

## 2. Controls and Comparison Groups

- Controls build implicit comparison groups.
- If correlated with  $X$  and  $Y$ , capture mediating channels.
- Comparison groups:
  - Individual and spouse characteristics.
  - COB characteristics.
  - Residence location.
- Cultural Proxy: FLFP 15+, ILO Estimates (EST), 1990–2017.
- Outcome: indicator for *In the labor force*.
- Timing of assignment: year of migration.
- Two-way clustering: State + COB.

## 2. Controls and Comparison Groups

- Survey year: 13 categories.
- Individual characteristics:
  - Cohort: 32 categories (1956–1987).
  - Education: 4 categories (below HS, HS, College, College+). [EDUC](#)
  - English: 3 categories (No English, some, English+). [SPEAKENG](#)
  - Ethnicity: 4 categories (White, Black, Asian, Other). [RACE](#)
  - Year immigration: 62 categories (1956–2017).
  - Number of children (in household): 0–9.
- Spouse characteristics:
  - Same as above.
  - Employment status (indicator).
  - Citizenship status (indicator).
  - Total income in 1999 US\$.

## 2. Controls and Comparison Groups

- COB geography:
  - Area, distance, continent, landlocked. [CEPII](#)
- COB history:
  - Identity of colonizer: 14 categories. [CEPII](#)
  - Genetic distance to US. [Spolaore and Wacziarg \(2018\)](#)
- COB economy:
  - Real GDP PPP 2011 US\$. [PWT](#)
  - Population. [PWT](#)
  - Female years of schooling and female/male ratio. [Barro and Lee \(2013\)](#)
  - Total fertility rate. [UN DESA](#)
  - Net migration rate. [UN DESA](#)

## 2. Controls and Comparison Groups

Dependent variable:	Active				
	(1)	(2)	(3)	(4)	(5)
COB FLFP 15+	0.48*** [0.05]	0.50*** [0.06]	0.40*** [0.07]	0.43*** [0.11]	0.42*** [0.07]
Controls					
Individual	No	No	Yes	No	Yes
Spouse	No	No	Yes	No	Yes
COB	No	No	No	Yes	Yes
Residence level	State	State	State	State	State
Residence FE	51	51	51	51	51
Observations	280,003	268,452	268,452	268,452	268,452

- Different source of variation ( $\simeq 12k$ ): (1)  $\rightarrow$  (2) = +0.02 (+4%).
- Mediating factors: (2)  $\rightarrow$  (5) = -0.08 (-16%). s.e.  $\simeq 0.09$ .

### 3. Interpretation: Magnitude

- How to gauge effect size in a given study? Depends on context.
- Usually most useful: share of variation explained.
- $\uparrow$  1 s.d. treatment is associated with  $\Delta \times$  s.d. of outcome.
- s.d. relative to which sample? Depends on context.
  - Relative to regression sample? Here, individuals.
  - Relative to sample of interest? Here, COB.
  - Tied to effective regression sample? Here, residualized s.d.  
[Mummolo and Peterson \(2018\)](#)
- Still, give actual estimates: meaning depends on context.

### 3. Interpretation: Magnitude

“Cohen's d” (arbitrary rule of thumb):

- $x \leq 0.2$ : small effect.
- $0.2 \leq x \leq 0.8$ : medium effect.
- $x > 0.8$ : large effect.



### 3. Interpretation: Magnitude

Sample:	Individuals				COB		
	Obs.	Mean	S.d.	$\widehat{S.d.}$	Obs.	Mean	S.d.
COB FLFP 15+	268,452	0.45	0.12	0.05	109	0.49	0.14
Active	268,452	0.52	0.50	0.47	109	0.57	0.13

- Initial question: explain  $\Delta$  in economic behavior across cultures.
- Cultures are captured here by COB.
- Relevant statistics are at the COB level.
- Mechanically larger magnitude ( $\times 4$ ).

### 3. Interpretation: Magnitude

- Cultural Proxy: FLFP 15+, ILO Estimates (EST), 1990–2017.
- Controls: individual, spouse, COB.
- Outcome: indicator for *In the labor force*.
- Timing of assignment: year of migration.
- Two-way clustering: State + COB.
- Residence location FE: State.

### 3. Interpretation: Magnitude

Dependent variable:	Active	
	(1)	(2)
COB FLFP 15+	0.50*** [0.06]	0.42*** [0.07]
Controls	No	Yes
Outcome mean	0.57	0.57
Outcome s.d.	0.13	0.13
COB FLFP 15+ mean	0.49	0.49
COB FLFP 15+ s.d.	0.14	0.14
Observations	268,452	268,452

- Estimate

- (1):  $\uparrow$  1pp COB FLFP associated with  $\uparrow$  0.5pp Active.
- (2):  $\uparrow$  1pp COB FLFP associated with  $\uparrow$  0.4pp Active.

### 3. Interpretation: Magnitude

Dependent variable:	Active	
	(1)	(2)
COB FLFP 15+	0.50*** [0.06]	0.42*** [0.07]
Controls	No	Yes
Outcome mean	0.57	0.57
Outcome s.d.	0.13	0.13
COB FLFP 15+ mean	0.49	0.49
COB FLFP 15+ s.d.	0.14	0.14
Observations	268,452	268,452

- Magnitude (mean)
  - (1):  $\uparrow$  10pp COB FLFP associated with  $\uparrow$  9% of mean in Active.
  - (2):  $\uparrow$  10pp COB FLFP associated with  $\uparrow$  7% of mean in Active.

### 3. Interpretation: Magnitude

Dependent variable:	Active	
	(1)	(2)
COB FLFP 15+	0.50*** [0.06]	0.42*** [0.07]
Controls	No	Yes
Outcome mean	0.57	0.57
Outcome s.d.	0.13	0.13
COB FLFP 15+ mean	0.49	0.49
COB FLFP 15+ s.d.	0.14	0.14
Observations	268,452	268,452

- Magnitude (s.d.)
  - (1):  $\uparrow$  1 s.d. COB FLFP associated with  $\uparrow$  0.54 s.d. in Active.
  - (2):  $\uparrow$  1 s.d. COB FLFP associated with  $\uparrow$  0.45 s.d. in Active.
- Cohen's  $d$ : effect size is  $0.5\sigma$  (medium).

### 3. Interpretation: Magnitude

Dependent variable:	Active	
	(1)	(2)
COB FLFP 15+	0.50*** [0.06]	0.42*** [0.07]
Controls	No	Yes
Outcome mean	0.57	0.57
Outcome s.d.	0.13	0.13
COB FLFP 15+ mean	0.49	0.49
COB FLFP 15+ s.d.	0.14	0.14
Observations	268,452	268,452

- Relative magnitudes
  - $\uparrow$  1 s.d. COB FLFP associated with  $\uparrow$  0.5 s.d. in Active.
  - $\downarrow$  1 s.d. children (0.44) associated with  $\uparrow$  0.18 s.d. in Active.

### 3. Interpretation: Precision

Dependent variable:	Active	
	(1)	(2)
COB FLFP 15+	$0.50 \pm 0.11$ [0.06]	$0.42 \pm 0.15$ [0.07]
Controls	No	Yes
Observations	268,452	268,452

- $x \pm y/2$ , where  $y$  is width of 95% confidence interval.
- [Abadie \(2019\) Statistical Non-Significance](#)

## 4. Clustering

- Cultural Proxy: FLFP 15+, ILO Estimates (EST), 1990–2017.
- Controls: individual, spouse, COB.
- Outcome: indicator for *In the labor force*.
- Timing of assignment: year of migration.
- Residence location FE: State.
- Clustering: country, State, COB, both?
- Cameron and Miller (2015) Abadie et al. (2017)



## 4. Clustering

- Regressor correlated across groups  $\implies$  cluster on these groups.
- One-way clustering: inflation factor of standard errors increases with
  - Cluster size.
  - Within-cluster correlation of regressor.
  - Within-cluster correlation of error.
- Two-way clustering (for one coefficient):
  - Estimate s.e. for both one-way clustering separately.
  - Estimate s.e. for one-way clustering on intersection.
  - Compute  $\sqrt{se_1^2 + se_2^2 - se_{1 \cap 2}^2}$ .

## 4. Clustering

Dependent variable:	Active			
	(1)	(2)	(3)	(4)
COB FLFP 15+	0.42*** [0.02]	0.42*** [0.04]	0.42*** [0.07]	0.42*** [0.07]
Cluster level	None	State	COB	S+COB
Controls	Yes	Yes	Yes	Yes
Observations	268,452	268,452	268,452	268,452

51 States of residence, 109 countries of birth.

## 5. Source of Identification

- Cultural Proxy: FLFP 15+, ILO Estimates (EST), 1990–2017.
- Controls: individual, spouse, COB.
- Outcome: indicator for *In the labor force*.
- Timing of assignment: year of migration.
- Two-way clustering: State + COB.
- Residence location FE: State.

## 5. Source of Identification

Dependent variable:	Active	
	(1)	(2)
COB FLFP 15+	0.50*** [0.06]	0.42*** [0.07]
Controls	No	Yes
State FE	Yes	Yes
COB FLFP 15+ variance	0.013	0.013
COB FLFP 15+ residual variance	0.013	0.002
Adj. R2	0.026	0.106
Adj. within R2	0.013	0.036
Observations	268,452	268,452

- Residual variance: variance conditional on controls.
- (2): residual variance = 18% of variance.  
⇒ 18% of original variation left for identification.

## 5. Source of Identification

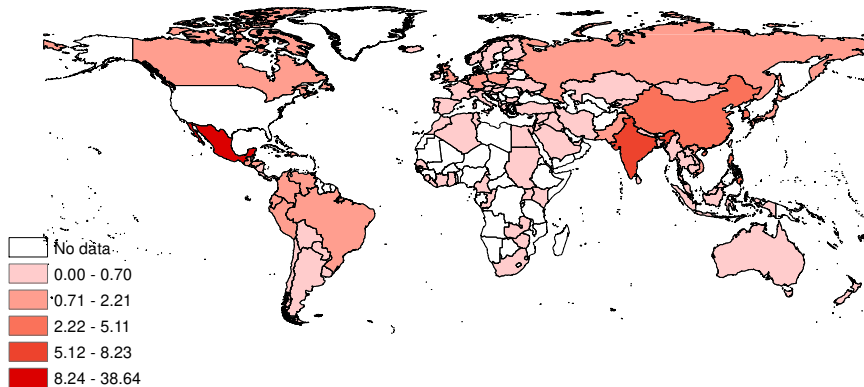
- Implicit regression weight: share of total residual variance.
- [Aronow and Samii \(2016\)](#) How representative are OLS estimates?
- Compute for each cultural group its contribution.
- Share of residual variance of each COB.

## 5. Source of Identification: Effective Sample

Sample (%)		Regression weight (%)			
		No controls		Controls	
MEX	38.64	CHN	23.24	TUR	6.39
IND	8.23	VNM	10.78	CHL	5.49
CHN	5.11	IND	9.73	ISR	5.41
PHL	3.59	MEX	9.56	NPL	4.41
SLV	3.05	PAK	4.29	HUN	4.06
KOR	3.04	CAN	3.75	GHA	3.61
JPN	2.21	IRN	3.02	PER	3.36
GTM	1.95	NPL	2.86	COL	3.07
CAN	1.92	PER	2.25	IND	3.05
VNM	1.77	IRQ	1.96	JPN	2.73
TUR	0.32	TUR	0.92	MEX	1.86
CHL	0.24	CHL	0.16	IND	3.05
ISR	0.34	ISR	0.24	CHN	1.87

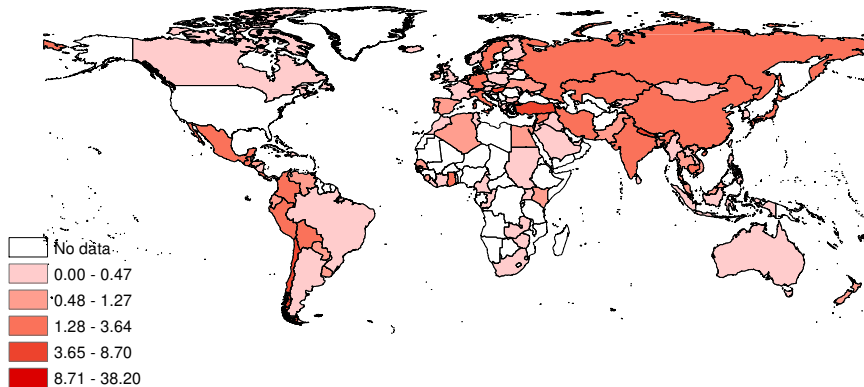
- Effective sample very different from nominal sample.
- Good news: contribution of each COB effectively balanced.

## 5. Source of Identification: Nominal Sample



- Nominal weight (%) across 109 COB.

## 5. Source of Identification: Effective Sample



- Regression weight (%) across 109 COB.



## 6. Outliers

- Cultural Proxy: FLFP 15+, ILO Estimates (EST), 1990–2017.
- Controls: individual, spouse, COB.
- Outcome: indicator for *In the labor force*.
- Timing of assignment: year of migration.
- Two-way clustering: State + COB.
- Residence location FE: State.
- Outliers: sample size, residual variance.

## 6. Outliers: Nominal Sample

Dependent variable:	Active				
	(1)	(2)	(3)	(4)	(5)
COB FLFP 15+	0.42*** [0.07]	0.41*** [0.07]	0.40*** [0.08]	0.37*** [0.08]	0.31*** [0.07]
Controls	Yes	Yes	Yes	Yes	Yes
COB dropped level	None	MEX	IND	CHN	All
Observations	268,452	171,098	245,988	253,365	133,547

- Different source of variation (50% of sample).
- (1)  $\rightarrow$  (5) =  $-0.11$  ( $-26\%$ ). s.e.  $\simeq 0.10$ .
- China outlier: very high FLFP at home (communism).

## 6. Outliers: Effective Sample

Dependent variable:	Active				
	(1)	(2)	(3)	(4)	(5)
COB FLFP 15+	0.42*** [0.07]	0.40*** [0.08]	0.41*** [0.08]	0.38*** [0.08]	0.33*** [0.09]
Controls	Yes	Yes	Yes	Yes	Yes
COB dropped level	None	TUR	CHL	ISR	All
Observations	268,452	267,507	267,776	267,530	265,909

- Different source of variation (99% of sample).
- (1)  $\rightarrow$  (5) =  $-0.09$  ( $-21\%$ ). s.e.  $\simeq 0.12$ .

## 7. Choice of Cultural Proxy

- Cultural Proxy: FLFP 15+ vs 15–65, FLFP vs ratio male to female.
- Data source: ILO Estimates (EST), 1990–2017.
- Controls: individual, spouse, COB.
- Outcome: indicator for *In the labor force*.
- Timing of assignment: year of migration.
- Two-way clustering: State + COB.
- Residence location FE: State.

## 7. Choice of Cultural Proxy

Dependent variable:	Active			
	FLFP		M to F LFP	
COB measure:	15+	15-65	15+	15-65
	(1)	(2)	(3)	(4)
COB measure	0.42*** [0.07]	0.34*** [0.07]	0.37*** [0.06]	0.33*** [0.06]
Outcome mean	0.57	0.57	0.57	0.57
Outcome s.d.	0.13	0.13	0.13	0.13
COB measure mean	0.49	0.55	0.67	0.69
COB measure s.d.	0.14	0.16	0.19	0.21
Observations	268,452	268,452	268,452	268,452

## 7. Choice of Cultural Proxy

- Measurement:
  - 15–65 vs 15+: less bias from mortality.
  - Ratio male to female LFP:
    - Comparable measurement across countries.
    - Captures COB gender roles more directly.
- Results

	FLFP		M to F LFP	
	15+	15–65	15+	15–65
$\Delta$ pp	4.2	3.4	3.7	3.3
$\Delta$ sd	0.45	0.42	0.54	0.53

## 8. Choice of Outcome

- Cultural Proxy: male to female LFP 15–65, ILO EST, 1990–2017.
- Controls: individual, spouse, COB.
- Timing of assignment: year of migration.
- Two-way clustering: State + COB.
- Residence location FE: State.

## 8. Choice of Outcome

- Extensive margin:
  - Indicator for *In the labor force*. LABFORCE
  - Indicator for *Employed*. EMPSTAT
  - Indicator for *Worked for pay last week*. WRKLSTWK
  - Indicator for *Worked for pay last year*. WORKEDYR
- Intensive margin:
  - Number of *Weeks worked last year*. WKSWORK2
  - Number of *Usual hours worked per week*. UHRSWORK
- Within-household division of labor.



## 8. Choice of Outcome: Extensive Margin

Dependent variable:	Active	Employed	Work LW	Work LY
	(1)	(2)	(3)	(4)
COB M/F LFP 15–65	0.33*** [0.06]	0.35*** [0.06]	0.36*** [0.06]	0.38*** [0.05]
Outcome mean	0.57	0.52	0.50	0.59
Outcome s.d.	0.13	0.13	0.13	0.14
COB measure mean	0.69	0.69	0.69	0.69
COB measure s.d.	0.21	0.21	0.21	0.21
Observations	268,452	268,452	257,217	268,452

- $\uparrow$  1 s.d. COB measure associated with  $\uparrow$  0.53–0.58 s.d. in outcome.

## 8. Choice of Outcome: Intensive Margin

Dependent variable:	Weeks		Hours	
	(1)	(2)	(3)	(4)
COB M/F LFP 15–65	18.55*** [2.66]	4.32*** [1.11]	16.29*** [2.27]	4.80*** [0.79]
Sample	All	Working	All	Working
Outcome mean	25.25	42.74	21.17	35.77
Outcome s.d.	6.18	2.32	5.36	1.86
COB measure mean	0.69	0.69	0.69	0.69
COB measure s.d.	0.21	0.21	0.21	0.21
Observations	268,452	145,964	268,452	145,964

- Include zeros:
  - $\uparrow$  1 s.d. COB measure associated with  $\uparrow$  3.9 weeks per year.
  - $\uparrow$  1 s.d. COB measure associated with  $\uparrow$  0.63 s.d. weeks per year.

## 8. Choice of Outcome: Intensive Margin

Dependent variable:	Weeks		Hours	
	(1)	(2)	(3)	(4)
COB M/F LFP 15–65	18.55*** [2.66]	4.32*** [1.11]	16.29*** [2.27]	4.80*** [0.79]
Sample	All	Working	All	Working
Outcome mean	25.25	42.74	21.17	35.77
Outcome s.d.	6.18	2.32	5.36	1.86
COB measure mean	0.69	0.69	0.69	0.69
COB measure s.d.	0.21	0.21	0.21	0.21
Observations	268,452	145,964	268,452	145,964

- Include zeros:
  - $\uparrow$  1 s.d. COB measure associated with  $\uparrow$  3.4 hours per week.
  - $\uparrow$  1 s.d. COB measure associated with  $\uparrow$  0.64 s.d. hours per week..

## 8. Choice of Outcome: Intensive Margin

Dependent variable:	Weeks		Hours	
	(1)	(2)	(3)	(4)
COB M/F LFP 15–65	18.55*** [2.66]	4.32*** [1.11]	16.29*** [2.27]	4.80*** [0.79]
Sample	All	Working	All	Working
Outcome mean	25.25	42.74	21.17	35.77
Outcome s.d.	6.18	2.32	5.36	1.86
COB measure mean	0.69	0.69	0.69	0.69
COB measure s.d.	0.21	0.21	0.21	0.21
Observations	268,452	145,964	268,452	145,964

- Excluding zeros:
  - $\uparrow$  1 s.d. COB measure associated with  $\uparrow$  0.9 week per year.
  - $\uparrow$  1 s.d. COB measure associated with  $\uparrow$  0.39 s.d. week per year.

## 8. Choice of Outcome: Intensive Margin

Dependent variable:	Weeks		Hours	
	(1)	(2)	(3)	(4)
COB M/F LFP 15–65	18.55*** [2.66]	4.32*** [1.11]	16.29*** [2.27]	4.80*** [0.79]
Sample	All	Working	All	Working
Outcome mean	25.25	42.74	21.17	35.77
Outcome s.d.	6.18	2.32	5.36	1.86
COB measure mean	0.69	0.69	0.69	0.69
COB measure s.d.	0.21	0.21	0.21	0.21
Observations	268,452	145,964	268,452	145,964

- Excluding zeros:
  - $\uparrow$  1 s.d. COB measure associated with  $\uparrow$  1.0 hour per week.
  - $\uparrow$  1 s.d. COB measure associated with  $\uparrow$  0.54 s.d. hour per week.

## 8. Choice of Outcome: Household Division of labor

Dependent variable:	Wife - Husband				
	Employed	Weeks		Hours	
	(1)	(2)	(3)	(4)	(5)
COB M/F LFP 15-65	0.35*** [0.06]	18.11*** [2.59]	3.17*** [0.99]	15.53*** [2.40]	3.32*** [1.13]
Sample	All	All	Working	All	Working
Outcome mean	-0.36	-18.96	-4.60	-19.31	-7.73
Outcome s.d.	0.13	5.70	2.33	5.56	2.90
COB measure mean	0.69	0.69	0.69	0.69	0.69
COB measure s.d.	0.21	0.21	0.21	0.21	0.21
Observations	268,452	268,452	137,039	268,452	137,039

- Same magnitudes for within household division of labor.

## 9. Regression Model

- Cultural Proxy: male to female LFP 15–65, ILO EST, 1990–2017.
- No controls.
- Outcome: indicator for *In the labor force*.
- Timing of assignment: year of migration.
- Two-way clustering: State + COB.
- Residence location FE: State.

# 9. Regression Model

- Linear probability model:
  - OLS.
  - `reghdfe` High-dimensional FE, multi-way clustering. [Correia \(2017\)](#)
  - Good approximation for small changes.
  - Nicer interpretation.
- Non-linear probability model.
  - logit and probit.
  - Report estimates evaluated at the mean of the data.
  - Compute standard errors manually.



# 9. Regression Model

Dependent variable:	Active		
	(1)	(2)	(3)
COB M/F LFP 15–65	0.50*** [0.05]	0.51*** [0.06]	0.51*** [0.06]
Model	OLS	Logit	Probit
Observations	268,452	268,452	268,452

# Baseline Specification

- Cultural Proxy: female to male LFP 15–65, ILO EST, 1990–2017.
- Controls: individual, spouse, COB.
- Outcome: indicator for *In the labor force*.
- Timing of assignment: year of migration.
- Two-way clustering: State + COB.
- Residence location FE: State.

# Standard Regression Table

Dependent variable:	Active			
	(1)	(2)	(3)	(4)
COB F/M LFP 15–65	0.31*** [0.05]	0.24*** [0.05]	0.33*** [0.09]	0.33*** [0.06]
Controls				
State and year FE	Yes	Yes	Yes	Yes
Individual and Spouse	No	Yes	No	Yes
COB	No	No	Yes	Yes
COB	109	109	109	109
Observations	268,452	268,452	268,452	268,452
Adj. R2	0.024	0.093	0.049	0.106
Outcome: mean = 0.57, s.d. = 0.13				
COB proxy: mean = 0.69, s.d. = 0.21				

- (4):  $\uparrow$  1pp COB FLFP associated with  $\uparrow$  0.3pp Active.

# Standard Regression Table

Dependent variable:	Active			
	(1)	(2)	(3)	(4)
COB F/M LFP 15–65	0.31*** [0.05]	0.24*** [0.05]	0.33*** [0.09]	0.33*** [0.06]
Controls				
State and year FE	Yes	Yes	Yes	Yes
Individual and Spouse	No	Yes	No	Yes
COB	No	No	Yes	Yes
COB	109	109	109	109
Observations	268,452	268,452	268,452	268,452
Adj. R2	0.024	0.093	0.049	0.106
Outcome: mean = 0.57, s.d. = 0.13				
COB proxy: mean = 0.69, s.d. = 0.21				

- (4) ↑ 10pp COB FLFP associated with ↑ 6% of mean in Active.

# Standard Regression Table

Dependent variable:	Active			
	(1)	(2)	(3)	(4)
COB F/M LFP 15–65	0.31*** [0.05]	0.24*** [0.05]	0.33*** [0.09]	0.33*** [0.06]
Controls				
State and year FE	Yes	Yes	Yes	Yes
Individual and Spouse	No	Yes	No	Yes
COB	No	No	Yes	Yes
COB	109	109	109	109
Observations	268,452	268,452	268,452	268,452
Adj. R2	0.024	0.093	0.049	0.106
Outcome: mean = 0.57, s.d. = 0.13				
COB proxy: mean = 0.69, s.d. = 0.21				

- (4):  $\uparrow$  1 s.d. COB FLFP associated with  $\uparrow$  0.53 s.d. in Active.

# Standard Regression Table

Dependent variable:	Active					
	(1)	(2)	(3)	(4)	(5)	(6)
COB M/F LFP 15–65	0.33*** [0.06]	0.31*** [0.06]	0.31*** [0.06]	0.27*** [0.04]	0.21*** [0.06]	1.00 [0.82]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Controls interacted						
State and year FE	No	No	Yes	Yes	Yes	Yes
Individual and Spouse	No	No	No	Yes	Yes	Yes
COB	No	No	No	No	Yes	Yes
All	No	No	No	No	No	Yes
COB	109	109	109	109	109	33
Observations	268,452	233,199	233,199	233,199	233,199	1,046
Adj. R2	0.106	0.104	0.106	0.179	0.179	0.308

Outcome: mean = 0.59, s.d. = 0.13

COB proxy: mean = 0.69, s.d. = 0.21

- 662 State-year FE; 38,303 individual and spouse FE; 43 COB FE.

# Specific Empirical Issues

- 1 Interpretation of cultural proxy.
- 2 Unobserved human capital.
- 3 Assimilation.
- 4 Timing of assignment.
- 5 Institutional environment.
- 6 Referential group: parents, partners, neighborhood.
- 7 Within cultural heterogeneity.
- 8 External validity.
- 9 Sample selection.

# 1. Interpretation of Cultural Proxy

How does cultural proxy compare to COB overall effect?

- Estimate COB overall effects:
  - At least 50 obs per COB (100 COB, drop 212 obs, 9 COB).
  - Regress outcome on COB FE and individual controls.
  - 99 purged COB FE (relative to MEX).
- Regress 100 COB FE on cultural proxy and COB characteristics.
  - Average time-varying COB characteristics.



# 1. Interpretation of Cultural Proxy

	COB FE	
	Unpurged	Purged
	(1)	(2)
MEX	0.00	0.00
IND	0.07***	0.02*
CHN	0.20***	0.12***
PHL	0.26***	0.14***
SLV	0.16***	0.14***
KOR	-0.07***	-0.15***
JPN	-0.07***	-0.14***
GTM	0.04***	0.03***
CAN	0.18***	0.13***
VNM	0.22***	0.18***

- MEX male to female LFP 15–65 = 0.47 (among lowest).
- MEX % active = 0.44.

# 1. Interpretation of Cultural Proxy

Dependent variable:	COB FE					
	Unpurged			Purged		
	(1)	(2)	(3)	(4)	(5)	(6)
COB M/F LFP 15–65	0.36*** [0.05]		0.34*** [0.06]	0.28*** [0.04]		0.28*** [0.05]
COB controls	No	Yes	Yes	No	Yes	Yes
Observations	100	100	100	100	100	100
Adj. R2	0.432	0.308	0.575	0.369	0.280	0.529

- COB proxy alone captures large part of COB overall effect on FLFP.
- COB controls do not mediate this relationship.

# 1. Interpretation of Cultural Proxy

Does cultural proxy captures role of cultural change?

- Include COB FE.
- Control for years since in U.S. rather than immigration year FE.
- Identification: within COB time variation in cultural proxy.
- Very demanding on data.

# 1. Interpretation of Cultural Proxy

Dependent variable:	Active		
	(1)	(2)	(3)
COB F/M LFP 15–65	0.31*** [0.05]	0.35*** [0.06]	0.25* [0.13]
Controls	No	Yes	Yes
COB FE	No	No	Yes
COB proxy variance	0.0304	0.0304	0.0304
COB proxy residual variance	0.0284	0.0039	0.0005
Observations	268,452	268,452	268,452

- Very demanding on data: little residual variance left for identification.
- 1990–2017:  $\Delta$  COB F/M LFP 15–65  $0.48 \implies 0.66$ .
- $\Delta$  COB F/M LFP 15–2017 associated with  $\uparrow$  0.09 s.d. in active.

## 2. Unobserved Human Capital

- COB proxy could capture unobservable heterogeneity in HK.
- Social capital might vary with COB and correlate with proxy.
- Control for parental education. [Fernández and Fogli \(2009\)](#)
- Control for pre-migration characteristics. [Blau and Kahn \(2015\)](#)
- If unobservables, reflected in market wage.
  - Use Mincer equation.
  - Compute log hourly wage rates.
  - OLS and Heckman selection model.
  - Selection equation: number of children and spouse's characteristics.

## 2. Unobserved Human Capital

Dependent variable: Estimator:	Log hourly wage			
	OLS		Heckit	
	(1)	(2)	(3)	(4)
COB M/F LFP 15–65	0.56 [0.46]	0.20** [0.08]	0.35 [0.40]	0.21*** [0.06]
Controls	No	Yes	No	Yes
Hourly wage mean	23.23	23.23	23.23	23.23
Hourly wage s.d.	7.60	7.60	7.60	7.60
Observations	134,262	134,262	268,455	268,455
Censored observations			134, 193	134, 193

COB measure: mean = 0.69, s.d. = 0.21

- Magnitude in (4):
  - $\uparrow$  1 s.d. COB M/F LFP 15–65 associated with  $\uparrow$  4.41% hourly wage.
  - At mean, hourly wage  $\uparrow$  from \$23.23 to \$24.25.
  - Corresponds to  $\uparrow$  0.13 s.d. in hourly wage.

### 3. Assimilation

- COB culture might dissipate over time.
- [Blau, Kahn, and Papps \(2011\)](#) analyze convergence to native levels.
- See also [Neuman \(2018\)](#)
- Include interaction with *years since in the U.S.*
- Might depend on cultural trait.

### 3. Assimilation

- Cultural Proxy: F/M LFP 15–65, ILO EST, 1990–2017.
- Controls: individual, spouse, COB.
  - No FE for year of immigration.
  - FE for year of birth.
- Outcome: indicator for *In the labor force*.
- Timing of assignment: year of migration.
- Two-way clustering: State + COB.
- Residence location FE: State.



### 3. Assimilation

Dependent variable:	Active			
	(1)	(2)	(3)	(4)
COB M/F LFP 15–65	0.29*** [0.08]		0.35*** [0.06]	0.26*** [0.09]
Years since in U.S.		0.01*** [0.00]	0.01*** [0.00]	0.00 [0.00]
Interaction				0.01* [0.01]
Observations	268,452	268,452	268,452	268,452

Years since in the U.S: mean = 10.21, s.d. = 6.10

- (Some) evidence of assimilation over time.
- Assimilation driven by migrants from progressive countries?
- Explore further (e.g., critical age).

## 4.a. Assignment: Year vs Decade Average

- Cultural Proxy: M to F LFP 15–65, ILO Estimates (EST), 1990–2017.
- Controls: individual, spouse, COB.
- Outcome: indicator for *In the labor force*.
- COB characteristics assignment: year vs decade of migration.
- Two-way clustering: State + COB.
- Residence location FE: State.

## 4.a. Assignment: Year vs Decade Average

Dependent variable:	Active			
	Year		Decade	
COB assignment:	(1)	(2)	(3)	(4)
COB M/F LFP 15–65	0.31*** [0.05]	0.33*** [0.06]	0.31*** [0.05]	0.35*** [0.06]
Controls	No	Yes	No	Yes
COB measure mean	0.69	0.69	0.70	0.70
COB measure s.d.	0.21	0.21	0.21	0.21
Observations	268,452	268,452	268,452	268,452

- Practically same results.
- Year enables more time variation.

## 4.b. Assignment: Year of Migration vs Current

- Cultural Proxy: M to F LFP 15–65, ILO Estimates (EST), 1990–2017.
- Controls: individual, spouse, COB.
- Outcome: indicator for *In the labor force*.
- COB characteristics assignment: year of migration vs current year.
- Two-way clustering: State + COB.
- Residence location FE: State.

## 4.b. Assignment: Year of Migration vs Current

Dependent variable:	Active				
	(1)	(2)	(3)	(4)	(5)
COB M/F LFP 15–65 (at migration)	0.31*** [0.05]	0.33*** [0.06]			0.34*** [0.12]
COB M/F LFP 15–65 (current)			0.32*** [0.08]	0.27*** [0.07]	-0.05 [0.14]
Controls	No	Yes	No	Yes	Yes
Observations	268,452	268,452	268,452	268,452	268,452

COB measure (at migration): mean = 0.69, s.d. = 0.21  
COB measure (current): mean = 0.72, s.d. = 0.20

- Straightforward: past rather than current measure matters.
- Little evidence of cultural links after migration.

## 4.b. Assignment: Migration vs Age 15 vs Current

Dependent variable:	Active			
	(1)	(2)	(3)	(4)
COB M/F LFP 15–65 (age 15)	0.28*** [0.06]			-0.10 [0.09]
COB M/F LFP 15–65 (at migration)		0.39*** [0.06]		0.57*** [0.15]
COB M/F LFP 15–65 (current)			0.29*** [0.07]	-0.16 [0.16]
Controls	Yes	Yes	Yes	Yes
Observations	114,413	114,413	114,413	114,413
COB measure (age 15): mean = 0.66, s.d. = 0.22				
COB measure (at migration): mean = 0.70, s.d. = 0.21				
COB measure (current): mean = 0.73, s.d. = 0.21				

- Preference formation relevant throughout adulthood.

## 5. Institutional Environment

- Approach requires fixing institutional environment.
- What is the relevant level?
  - States: 51 categories.
  - Counties: 475 identifiable counties, 70k obs not identifiable (21%).
  - PUMAs: 3,199 categories but heterogeneous (varies across years).
- Potentially selected migration across local labor markets.

## 5. Institutional Environment: States vs PUMAs

Dependent variable:	Active		
	(1)	(2)	(3)
COB FLFP 15+	0.42*** [0.07]	0.42*** [0.07]	0.41*** [0.07]
Controls	Yes	Yes	Yes
Residence level	U.S.	State	PUMA
Residence FE	1	51	3,189
Observations	268,452	268,452	268,445



## 5. Institutional Environment: States vs Counties

Dependent variable:	Active				
	(1)	(2)	(3)	(4)	(5)
COB FLFP 15+	0.42*** [0.07]	0.43*** [0.08]	0.43*** [0.08]	0.42*** [0.07]	0.42*** [0.08]
Controls	Yes	Yes	Yes	Yes	Yes
Residence level	U.S.	U.S.	State	PUMA	County
Residence FE	1	1	51	3,189	474
Observations	268,452	211,593	211,593	211,589	211,593

## 6. Referential Group

Implicit model of cultural transmission, with a relevant referential.

- Individuals themselves (1st generation immigrants).
- Parents through vertical transmission.
- Spouse through within-marriage transmission / selection.
- Neighborhood through oblique socialization.

## 6.a. Referential Group: Parents

- Directly, 2nd generation immigrants.  
Fernández and Fogli (2009) Blau, Kahn, Liu, and Papps (2013)
- 3rd generation immigrants.  
Costa-Font, Giuliano and Ozcan (2018) (savings in UK)
- Rough approximation: “1.5th generation”, arrived young.

## 6.a. Referential Group: Parents

- Fernández and Fogli (2009)
- U.S. ACS 1970 1% Metro (contains father country of birth).
- Married women 30–40, father born outside U.S.
- Assign FLFP 10+ of father COB in 1950 using ILO YI.
- 7k women over 25 countries of ancestry.
- Outcome: Hours worked in previous week.
- $\uparrow 1$  s.d. FLFP COB  $\implies$  with  $\uparrow 0.82$  hours, or 0.3 s.d. across COA.

## 6.a. Referential Group: Parents

- Fernández and Fogli (2009) Table 2

Dependent variable:	Hours worked			
	(1)	(2)	(3)	(4)
FCOB FLFP 10+ 1950	0.047*** [0.012]	0.041** [0.016]	0.072*** [0.015]	0.045*** [0.016]
Education FE	No	Yes	Yes	Yes
Husband education FE	No	No	Yes	Yes
Husband income	No	No	Yes	Yes
Number of children	No	No	No	Yes
SMSA FE	Yes	Yes	Yes	Yes
COB clusters	25	25	25	25
Observations	6,774	6,774	6,774	6,774
Adj. R2	0.018	0.024	0.053	0.053

## 6.a. Referential Group: Parents

- Blau, Kahn, Liu, and Papps 2013
- U.S. March CPS 1995–2011 (contains parents country of birth).
- Married women 25–49, father or mother born outside U.S.
- Assign hours worked of immigrants in ACS 1970–2000 at age 10.
- 20k women over 69 countries of ancestry.
- Outcome: annual hours worked.
- Positive effect of mother COB hours.
- Larger for mother COB than father's COB women hours.
- Decomposition by family type.

## 6.a. Referential Group: Parents (1.5th Generation)

- Cultural Proxy: M to F LFP 15–65, ILO Estimates (EST), 1990–2017.
- Controls: individual, spouse, COB.
- Outcome: indicator for *In the labor force*.
- COB characteristics assignment: year of migration.
- Two-way clustering: State + COB.
- Residence location FE: State.

## 6.a. Referential Group: Parents (1.5th Generation)

Dependent variable:	Active			
	(1)	(2)	(3)	(4)
COB M/F LFP 15–65	0.30*** [0.06]	-0.02 [0.06]	0.26*** [0.06]	0.27*** [0.06]
Age migrated	All	< 10	< 15	< 20
Outcome mean	0.64	0.71	0.72	0.70
Outcome s.d.	0.12	0.28	0.19	0.14
COB measure mean	0.69	0.64	0.66	0.67
COB measure s.d.	0.21	0.23	0.22	0.21
Observations	411,973	3,375	16,477	71,785



## 6.b. Referential Group: Spouse

- Spouse's culture might matter:
  - Assortative matching.
  - Household bargaining.
- Sample selection, spouse's side:
  - Born outside U.S. from non US parents.
  - No age restriction.
  - Born in single identified country.

## 6.b. Referential Group: Spouse

Compare main results across regression samples.

- Cultural Proxy: wife's M to F LFP 15–65, ILO EST, 1990–2017.
- Controls: individual, spouse, wife's COB.
- Outcome: indicator for *In the labor force*.
- Wife's COB characteristics assignment: wife's year of migration.
- Two-way clustering: State + wife's COB.
- Residence location FE: State.

## 6.b. Referential Group: Spouse

Dependent variable:	Active			
	(1)	(2)	(3)	(4)
COB M/F LFP 15–65	0.31*** [0.05]	0.33*** [0.06]	0.29*** [0.08]	0.36*** [0.06]
Sample	Original	Original	Restricted	Restricted
Controls	No	Yes	No	Yes
Outcome mean	0.52	0.52	0.51	0.51
Outcome s.d.	0.50	0.50	0.50	0.50
COB measure mean	0.59	0.59	0.58	0.58
COB measure s.d.	0.17	0.17	0.18	0.18
Observations	268,452	268,452	172,262	172,262

## 6.b. Referential Group: Spouse

His or her culture?

- Isolate each spouse's culture.
  - Include spouse's COB FE.
  - Include spouse's COB  $\times$  year immigration FE.
- Compare spouses' cultures directly.
  - Include both spouses' cultures.
  - Differentiate spouses from same COB and different COB.

## 6.b. Referential Group: Isolate Spouses' Culture

- Cultural Proxies: spouses' COB M to F LFP 15–65, ILO EST, 1990–2017.
- Controls: individual, spouse, wife's COB or husband's COB.
- Spouse's COB FE or COB  $\times$  year immigration FE.
- Outcome: indicator for wife *In the labor force*.
- Spouse COB characteristics assignment: spouse's year of migration.
- Two-way clustering: State + spouse's COB.
- Residence location FE: State.

## 6.b. Referential Group: Isolate Spouses' Culture

Dependent variable:	Active					
	Wife's COB culture			Husband's COB culture		
	(1)	(2)	(3)	(4)	(5)	(6)
COB M/F LFP 15–65	0.36*** [0.06]	0.24*** [0.05]	0.27*** [0.04]	0.41*** [0.05]	0.25*** [0.04]	0.21*** [0.04]
Controls						
Individual	Yes	Yes	Yes	Yes	Yes	Yes
Spouse COB	Yes	Yes	Yes	Yes	Yes	Yes
Spouse COB FE	No	Yes	No	No	Yes	No
Spouse COB × YI FE	No	No	Yes	No	No	Yes
COB proxy variance	0.0293	0.0293	0.0293	0.0300	0.0300	0.0300
COB proxy residual variance	0.0040	0.0015	0.0015	0.0056	0.0015	0.0018
Observations	149,737	149,737	149,737	149,737	149,737	149,737
Outcome: mean = 0.54, s.d. = 0.14						
COB proxy: mean = 0.69, s.d. = 0.19						

- Share of couples same COB: 92%.
- Share of couples same COB and same year immigration: 48%.

## 6.b. Referential Group: Isolate Spouses' Culture

Dependent variable:	Active					
	Wife's COB culture			Husband's COB culture		
	(1)	(2)	(3)	(4)	(5)	(6)
COB M/F LFP 15–65	0.36*** [0.06]	0.24*** [0.05]	0.27*** [0.04]	0.41*** [0.05]	0.25*** [0.04]	0.21*** [0.04]
Controls						
Individual	Yes	Yes	Yes	Yes	Yes	Yes
Spouse COB	Yes	Yes	Yes	Yes	Yes	Yes
Spouse COB FE	No	Yes	No	No	Yes	No
Spouse COB × YI FE	No	No	Yes	No	No	Yes
COB proxy variance	0.0293	0.0293	0.0293	0.0300	0.0300	0.0300
COB proxy residual variance	0.0040	0.0015	0.0015	0.0056	0.0015	0.0018
Observations	149,737	149,737	149,737	149,737	149,737	149,737
Outcome: mean = 0.54, s.d. = 0.14						
COB proxy: mean = 0.69, s.d. = 0.19						

- (3):  $\uparrow 1$  s.d. in wife's COB proxy  $\implies \uparrow 0.37$  s.d. in outcome.
- (6):  $\uparrow 1$  s.d. in husband's COB proxy  $\implies \uparrow 0.14$  s.d. in outcome.

## 6.b. Referential Group: Compare Spouses' Culture

Compare spouses' cultures directly

- Cultural Proxies: spouses' COB M to F LFP 15–65, ILO EST, 1990–2017.
- Controls: individual, spouse, wife's COB and husband's COB.
- Differentiate if same COB.
- Outcome: indicator for wife *In the labor force*.
- Spouse COB characteristics assignment: spouse's year of migration.
- Three-way clustering: State + spouse's COB.
- Residence location FE: State.



## 6.b. Referential Group: Compare Spouses' Culture

Dependent variable:	Active			
	(1)	(2)	(3)	(4)
Same × COB M/F LFP 15–65	0.37*** [0.06]	0.35*** [0.07]	0.39*** [0.07]	0.40*** [0.07]
Different × wife's COB M/F LFP 15–65	0.36*** [0.06]		0.19*** [0.04]	0.24*** [0.04]
Different × husband's COB M/F LFP 15–65		0.36*** [0.07]	0.22*** [0.06]	0.29*** [0.07]
Different × Interaction				-0.17*** [0.05]
Controls				
Individual	Yes	Yes	Yes	Yes
Wife's COB	Yes	No	Yes	Yes
Husband's COB	No	Yes	Yes	Yes
Observations	169,531	169,531	169,531	169,531

## 6.b. Referential Group: Spouse

Impact of migrant husbands' culture on US wives?

- Born in the U.S. from US parents.
- Age: 30–49.
- Regular household, no farms, married spouse present.
- Husband born from non US parents.
- Husband born in single identified country outside U.S.
- Husband migrated after 1990.

## 6.b. Referential Group: Spouse

Impact of migrant husbands' culture on US wives?

- Cultural Proxy: husband COB M to F LFP 15–65, ILO EST, 1990–2017.
- Controls: individual, spouse, husband's COB.
- Outcome: indicator for wife *In the labor force*.
- Spouse COB characteristics assignment: spouse's year of migration.
- Two-way clustering: State + spouse's COB.
- Residence location FE: State.

## 6.b. Referential Group: Spouse

Dependent variable:	Active	
	(1)	(2)
Husband's COB M/F LFP 15–65	0.16*** [0.03]	0.08** [0.03]
Controls		
Individual	Yes	Yes
Husband's COB	No	Yes
Observations	51,019	51,019
COB measure: mean = 0.68 ; s.d. = 0.22		
Outcome: mean = 0.77 ; s.d. = 0.10		

- (2):  $\uparrow$  s.d. in husband's COB proxy  $\implies$   $\uparrow$  0.18 s.d. in outcome.

## 6.c. Referential Group: Neighborhood

Role of neighborhood through oblique socialization?

- Migrants clustering might preserve and transmit COB culture.
- Mechanism: role models or rewards/punishment.
- Compute COB immigrants density + interact COB cultural.
- Three levels of aggregation (data limitations):
  - States-year: 51 categories.
  - PUMAs-year: 3,199 categories but heterogeneous definition.
  - Counties-year: 475 identifiable counties.
- Two reference groups:
  - General population.
  - Married women aged 30–49.

## 6.c. Referential Group: Neighborhood

Dependent variable:	Active			
	General		Women 30–49	
Reference population:	(1)	(2)	(3)	(4)
COB M/F LFP 15–65	0.33*** [0.06]	0.32*** [0.76]	0.33*** [0.06]	0.32*** [0.06]
COB density		-1.60** [0.76]		-0.93** [0.45]
Interaction		2.92** [1.44]		1.62* [0.83]
Controls	Yes	Yes	Yes	Yes
Aggregation	State	State	State	State
Units	51	51	51	51
Observations	268,451	268,451	268,451	268,451

## 6.c. Referential Group: Neighborhood

Dependent variable:	Active			
	General		Women 30–49	
Reference population:	(1)	(2)	(3)	(4)
COB M/F LFP 15–65	0.32*** [0.05]	0.30*** [0.24]	0.32*** [0.05]	0.30*** [0.05]
COB density		-0.51** [0.24]		-0.35*** [0.12]
Interaction		1.01** [0.44]		0.61*** [0.23]
Controls	Yes	Yes	Yes	Yes
Aggregation	PUMA	PUMA	PUMA	PUMA
Units	3,189	3,189	3,189	3,189
Observations	268,444	268,444	268,443	268,443

## 6.c. Referential Group: Neighborhood

Dependent variable:	Active			
	General		Women 30–49	
Reference population:	(1)	(2)	(3)	(4)
COB M/F LFP 15–65	0.33*** [0.06]	0.31*** [0.48]	0.33*** [0.06]	0.31*** [0.06]
COB density		-0.86* [0.48]		-0.53** [0.23]
Interaction		1.81** [0.90]		1.03** [0.44]
Controls	Yes	Yes	Yes	Yes
Aggregation	County	County	County	County
Units	474	474	474	474
Observations	211,592	211,592	211,592	211,592



## 6. Referential Group: Mechanisms to Explore

- 1st generation immigrants.
- 1.5th generation immigrants (vertical transmission).
- 2nd generation: mothers, fathers (vertical transmission).
- 1st generation: partner (transmission through marriage).
- 2nd generation: parents in-law (transmission through marriage).
- Social spillovers within neighborhoods (oblique transmission).

## 7. Within Culture Heterogeneity

- What about within-group cultural heterogeneity?
- Solution:
  - Cultural proxy by sub-group e.g. COB FLFP 30–25, married, college +.
  - E.g. use [IPUMS-I](#) ( $\sim 300\overline{M}$  observations across  $\sim 100$  countries)
- Enables using within-COB variation  $\implies$  improves identification.
  - Alleviates OVB on origin side.
  - Credible that environment is fixed (rules + networks).
- E.g. [Marcén and Morales \(2018, 2019\)](#) [Marcén et al. \(2018\)](#)
- Estimates increase in magnitude and precision.
- But not possible/relevant depending on question.

## 8. External Validity

- In this context: use effective sample approach.
- Similar results across time (migration waves)?
- Similar results across host environments (countries)?
- Dependent upon culture and cultural trait (assimilation, identity).
- Replicate across time and space.
- Meta-analysis or single analysis [Luttmer and Singhal \(2011\)](#)  
[Rodríguez-Planas and Nollenberger \(2018\)](#)  
[Beblo, Görges, and Markowsky \(2018\)](#) [Hoorn \(2019\)](#)

# 9. Selection

- Are immigrants representative of COB?
  - Compare characteristics.
  - Pre-immigration data. [Blau and Kahn \(2015\)](#)
- Selective in-migration.
  - Self-selection into more dynamic local labor markets.
  - Compare within local labor market.
- Selective out-migration.
  - Depends on context.
  - Here: tied movers, so less likely.
  - Generally: does cultural proxy predicts migration?
  - Unlikely to drive pattern in data.

# Applications

Similar strategy to assess role of culture for economic outcomes (1/2).

- Divorce rates [Furtado, Marcén, and Sevilla \(2013\)](#)
- Self-employment [Marcén \(2014\)](#)
- Living arrangements [Giuliano \(2007\)](#) [Marcén and Morales \(2018\)](#)
- Fertility [Bellido et. al \(2016\)](#) [Stichnoth and Yeter \(2016\)](#)  
[Marcén et. al \(2018\)](#) [Salari \(2018\)](#)
- Son preference [Almond, Edlund, and Milligan \(2013\)](#)
- Savings rates [Costa-Font et al. \(2018\)](#)

# General Applications

Similar strategy to assess role of culture for economic outcomes (2/2).

- Mortgage finance [Rodríguez-Planas \(2018\)](#)
- Home-ownership [Marcén and Morales \(2019\)](#)
- Unemployment rates [Brugger et al. \(2009\)](#)
- Division of labor within family [Frank and Hou \(2015\)](#)
- Science education [Nollenberger et al. \(2016\)](#)  
[Rodríguez-Planas and Nollenberger \(2018\)](#)
- Human capital accumulation [Hoorn \(2019\)](#)
- Health behavior [Rodríguez-Planas and Sanz-de-Galdeano \(2019\)](#)
- Preference for redistribution [Luttmer and Singhal \(2011\)](#)

# Applications in Economic History

- Most useful to study bigger questions with historical perspective:
  - How can we explain observed differences in culture?
  - Mechanisms through which culture affects behavior?
  - How does culture transmit across generations?
  - When does culture persist over time? When doesn't it?
  - What are the mechanisms of cultural change?
  - How does culture emerge?
- Specific (and standard) empirical issues still apply.



# Applications in Economic History

- Explain contemporaneous distribution of norms about gender roles: pre-industrial division of labor. [Alesina, Giuliano, and Nunn \(2013\)](#)

# Applications in Economic History

- Explain contemporaneous distribution of norms about gender roles: pre-industrial division of labor. [Alesina, Giuliano, and Nunn \(2013\)](#)
- Explain contemporaneous distribution of trust in Africa: the slave trade. [Nunn and Wantchekon \(2011\)](#)

# Applications in Economic History

- Explain contemporaneous distribution of norms about gender roles: pre-industrial division of labor. [Alesina, Giuliano, and Nunn \(2013\)](#)
- Explain contemporaneous distribution of trust in Africa: the slave trade. [Nunn and Wantchekon \(2011\)](#)
- Explain contemporaneous distribution of economic status in Africa: ancestral lifeways. [Michalopoulos, Putterman and Weil \(2019\)](#)

# Applications in Economic History

- Explain contemporaneous distribution of norms about gender roles: pre-industrial division of labor. [Alesina, Giuliano, and Nunn \(2013\)](#)
- Explain contemporaneous distribution of trust in Africa: the slave trade. [Nunn and Wantchekon \(2011\)](#)
- Explain contemporaneous distribution of economic status in Africa: ancestral lifeways. [Michalopoulos, Putterman and Weil \(2019\)](#)
- Mechanisms of persistence of historical shocks: contemporaneous gender norms in Africa and the slave trade. [Teso \(2019\)](#)

# Applications in Economic History

- Explain contemporaneous distribution of norms about gender roles: pre-industrial division of labor. [Alesina, Giuliano, and Nunn \(2013\)](#)
- Explain contemporaneous distribution of trust in Africa: the slave trade. [Nunn and Wantchekon \(2011\)](#)
- Explain contemporaneous distribution of economic status in Africa: ancestral lifeways. [Michalopoulos, Putterman and Weil \(2019\)](#)
- Mechanisms of persistence of historical shocks: contemporaneous gender norms in Africa and the slave trade. [Teso \(2019\)](#)
- Mechanisms of persistence of historical shocks: contemporaneous and past gender norms in France and WWI. [Gay \(2023\)](#)

THE  
QUARTERLY JOURNAL  
OF ECONOMICS

---

Vol. 128

May 2013

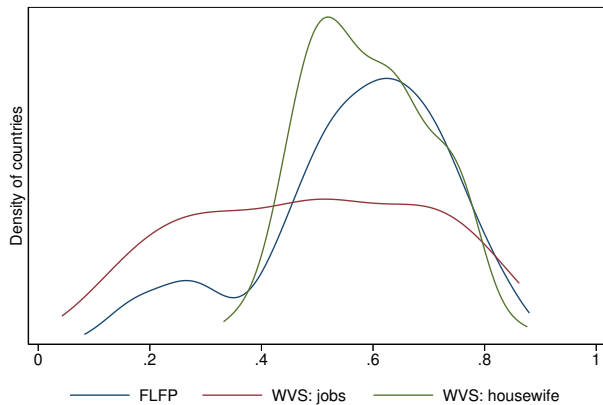
Issue 2

---

ON THE ORIGINS OF GENDER ROLES: WOMEN AND THE  
PLOUGH\*

ALBERTO ALESINA  
PAOLA GIULIANO  
NATHAN NUNN

# Alesina, Giuliano, and Nunn (2013)



Source: WVS, ILO.

# Alesina, Giuliano, and Nunn (2013)

- How can we explain current dispersion in norms about gender roles?



# Alesina, Giuliano, and Nunn (2013)

- How can we explain current dispersion in norms about gender roles?
  - Hypothesis of Boserup (1970):
    - Form of traditional agriculture in pre-industrial societies.
    - Shifting versus plough.
    - Plough: capital intensive, requiring more strength to control animal.
    - Men have comparative advantage in farming.
- ⇒ Specialization along gender lines ⇒ gender norms.

# Alesina, Giuliano, and Nunn (2013)

- How can we explain current dispersion in norms about gender roles?
- Hypothesis of Boserup (1970):
  - Form of traditional agriculture in pre-industrial societies.
  - Shifting versus plough.
  - Plough: capital intensive, requiring more strength to control animal.
  - Men have comparative advantage in farming.  
⇒ Specialization along gender lines ⇒ gender norms.
- Generated norms about gender roles that persisted:
  - Underlying trait reinforced by institutions.
  - Complementarities between cultural beliefs and industrial structure.
  - Cultural beliefs are inherently sticky.

Does traditional plough use have a causal impact on subsequent cultural norms?

# Historical data

- Murdock's *Ethnographic Atlas*.
- Information for 1,265 ethnic groups (before European contact).
- Widely used but source criticism needed.  
[Giuliano and Nunn \(2018\)](#) [Giuliano and Matranga \(2020\)](#)
- Historical plough use:
  - 1 Absent (plough = 0)
  - 2 Existed but not aboriginal (plough = 1)
  - 3 Aboriginal (plough = 1)

# Historical data

- Traditional female participation in agriculture:
  - ① Males only (Female specialization = 1)
  - ② Males appreciably more (Female specialization = 2)
  - ③ Equal participation (Female specialization = 3)
  - ④ Females appreciably more (Female specialization = 4)
  - ⑤ Females only (Female specialization = 5)
- Presence of domesticated animals.
- Density of ethnic groups' settlements.
- Political complexity.

# Geographic data

- Geographic coordinates of centroid of historical group.
- Land suitability for cultivation of crops (FAO).
- Tropical or subtropical characteristics.

# Historical Impact of Traditional Plough Agriculture

## Traditional plough use and female participation in pre-industrial agriculture

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent variable: Traditional participation of females relative to males in the following tasks:						
	Overall agriculture	Land clearance	Soil preparation	Planting	Crop tending	Harvesting	
Mean of dep. var.	3.04	2.83	1.45	2.15	2.86	3.16	3.23
Traditional plough agriculture	-0.883*** (0.225)	-1.136*** (0.240)	-0.434** (0.197)	-1.182*** (0.320)	-1.290*** (0.306)	-1.188*** (0.351)	-0.954*** (0.271)
Ethnographic controls	yes	yes	yes	yes	yes	yes	yes
Observations	660	124	129	124	131	122	131
Adjusted R-squared	0.13	0.19	0.14	0.10	0.09	0.13	0.16
R-squared	0.14	0.23	0.18	0.14	0.13	0.18	0.20

s.d. outcome in (1) is 1.0

# Historical Impact of Traditional Plough Agriculture

## Traditional plough use and female participation outside agriculture

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Dependent variable: Traditional participation of females relative to males in the following tasks:								
	Caring for small animals	Caring for large animals	Milking	Cooking	Fuel gathering	Water fetching	Burden carrying	Handicrafts	Trading
Mean of dep. var.	3.53	1.73	3.25	4.65	3.90	4.64	3.47	2.78	2.47
Traditional plough use	0.14 (0.517)	0.064 (0.254)	0.63 (0.697)	-0.019 (0.108)	-0.638 (0.403)	-0.052 (0.205)	-0.962** (0.378)	-0.157 (0.274)	-0.155 (0.542)
Ethnographic controls	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	88	95	48	173	159	154	135	74	59
Adjusted R-squared	-0.02	-0.02	0.03	0.01	-0.001	0.01	0.12	0.07	-0.01
R-squared	0.05	0.04	0.14	0.04	0.04	0.04	0.16	0.15	0.10

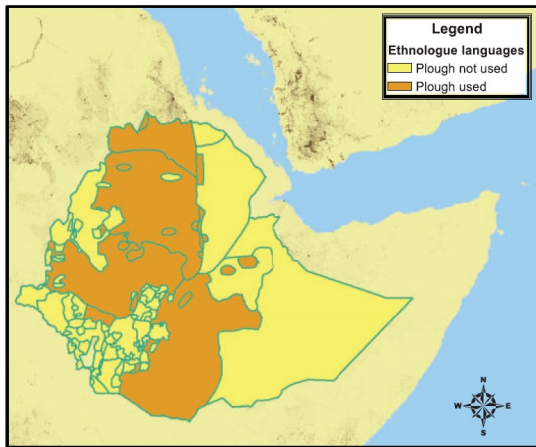


# Long-Term Impact Traditional Plough

- Data:
  - Use *Ethnologue* for location of language groups.
  - Use *Landscan 2000* for world's population by 1 km<sup>2</sup> grid cell.
  - Combine for distribution of language groups today.
  - Match with *Ethnographic Atlas* plough data by grid cell.
  - Aggregate at country or district level.
- Main variable: fraction of location with ancestors traditionally using plough.

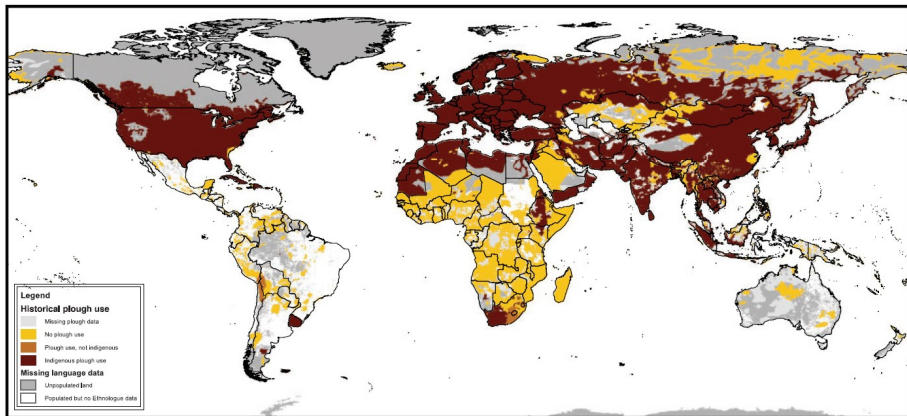
# Long-Term Impact Traditional Plough

Language groups and historical plough use in Ethiopia



# Long-Term Impact Traditional Plough

Traditional plough use among ethnic groups



# Country-Level Estimates

- Contemporaneous outcomes:
  - FLFP in 2000 (World Bank).
  - Proportion of firms owned by women (World Bank).
  - Share women in Parliament (UN).

# Country-Level Estimates

- Contemporaneous outcomes:
  - FLFP in 2000 (World Bank).
  - Proportion of firms owned by women (World Bank).
  - Share women in Parliament (UN).
- OLS regression:

$$y_c = \alpha + \beta \text{Plough}_c + \mathbf{X}_c^H \boldsymbol{\Gamma} + \mathbf{X}_c^C \boldsymbol{\Pi} + \varepsilon_c$$

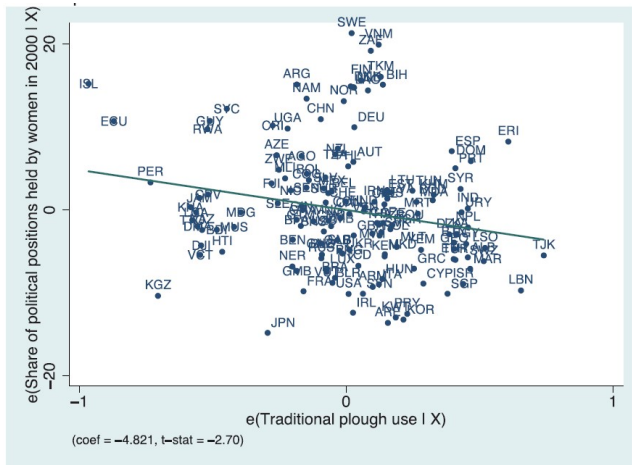
- $\mathbf{X}_c^H$ : historical ethnographic controls (domesticated animals, density of settlements, tropical climate, ...).
- $\mathbf{X}_c^C$ : contemporaneous controls (GDP per capita, flexible).





# Country-Level Estimates

## Partial Correlation: Plough and Politics





# Country-Level Estimates: Magnitudes

- Summary statistics: mean (s.d.)
  - Plough: 0.53 (0.47).
  - FLFP: 51.3 (15.6).
  - Ownership: 35.2 (15.0).
  - Politics: 11.8 (8.9).

# Country-Level Estimates: Magnitudes

- Summary statistics: mean (s.d.)
  - Plough: 0.53 (0.47).
  - FLFP: 51.3 (15.6).
  - Ownership: 35.2 (15.0).
  - Politics: 11.8 (8.9).
- Interpretation in means:
  - $\uparrow$  1 s.d. plough associated with  $\downarrow$  11% FLFP mean.
  - $\uparrow$  1 s.d. plough associated with  $\downarrow$  20% ownership mean.
  - $\uparrow$  1 s.d. plough associated with  $\downarrow$  19% politics mean.

# Country-Level Estimates: Magnitudes

- Summary statistics: mean (s.d.)
  - Plough: 0.53 (0.47).
  - FLFP: 51.3 (15.6).
  - Ownership: 35.2 (15.0).
  - Politics: 11.8 (8.9).
- Interpretation in means:
  - $\uparrow$  1 s.d. plough associated with  $\downarrow$  11% FLFP mean.
  - $\uparrow$  1 s.d. plough associated with  $\downarrow$  20% ownership mean.
  - $\uparrow$  1 s.d. plough associated with  $\downarrow$  19% politics mean.
- Interpretation in s.d.:
  - $\uparrow$  1 s.d. plough associated with  $\downarrow$  0.38 FLFP s.d.
  - $\uparrow$  1 s.d. plough associated with  $\downarrow$  0.48 ownership s.d.
  - $\uparrow$  1 s.d. plough associated with  $\downarrow$  0.25 politics s.d.

# Country-Level Estimates: Magnitudes

- Summary statistics: mean (s.d.)
  - Plough: 0.53 (0.47).
  - FLFP: 51.3 (15.6).
  - Ownership: 35.2 (15.0).
  - Politics: 11.8 (8.9).
- Interpretation in means:
  - $\uparrow$  1 s.d. plough associated with  $\downarrow$  11% FLFP mean.
  - $\uparrow$  1 s.d. plough associated with  $\downarrow$  20% ownership mean.
  - $\uparrow$  1 s.d. plough associated with  $\downarrow$  19% politics mean.
- Interpretation in s.d.:
  - $\uparrow$  1 s.d. plough associated with  $\downarrow$  0.38 FLFP s.d.
  - $\uparrow$  1 s.d. plough associated with  $\downarrow$  0.48 ownership s.d.
  - $\uparrow$  1 s.d. plough associated with  $\downarrow$  0.25 politics s.d.
- Based on R2, explains 6% of FLFP, 11% of ownership, 3% of politics.

# Subnational Estimates 1

- Individual-level outcomes from WVS:
  - Indicator for participation in labor force.
  - Attitudes toward work (“When jobs scarce, men have priority”).
  - Attitudes toward politics (“Men better politicians”).

# Subnational Estimates 1

- Individual-level outcomes from WVS:
  - Indicator for participation in labor force.
  - Attitudes toward work (“When jobs scarce, men have priority”).
  - Attitudes toward politics (“Men better politicians”).
- OLS regression:

$$y_{i,d,c} = \alpha_c + \beta \text{Plough}_d + \mathbf{X}_d^H \boldsymbol{\Pi} + \mathbf{X}_i \boldsymbol{\Phi} + \varepsilon_{i,d,c}$$

- $\alpha_c$ : country FE.
- $\mathbf{X}_d^H$ : historical ethnographic controls in district  $d$ , country  $c$ .
- $\mathbf{X}_i$ : individual controls (age, marital status, gender, education).

# Subnational Estimates 1

## Individual-level OLS estimates using WVS data

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable:					
	Female labor force participation, 1995–2007		When jobs are scarce, 1995–2007		Men better political leaders, 1995–2007	
Mean of dep. var.	0.55	0.55	0.46	0.47	2.62	2.64
Traditional plough use	-0.177*** (0.035)	-0.002 (0.031)	0.193*** (0.033)	0.100* (0.059)	0.224*** (0.069)	0.304*** (0.117)
Individual & district controls	yes	yes	yes	yes	yes	yes
Contemporary country controls	yes	n/a	yes	n/a	yes	n/a
Fixed effects	continent	country	continent	country	continent	country
Number of countries	73	78	74	79	50	55
Number of districts	672	698	674	700	453	479
Observations	43,801	47,587	80,303	87,528	64,215	72,152
Adjusted R-squared	0.17	0.27	0.21	0.28	0.19	0.26
R-squared	0.17	0.27	0.21	0.28	0.19	0.26

# Subnational Estimates 2

- Individual-level outcomes from IPUMS-I:
  - Indicator for participation in labor force.
  - 8 countries with ethnic identify information.



# Subnational Estimates 2

- Individual-level outcomes from IPUMS-I:
  - Indicator for participation in labor force.
  - 8 countries with ethnic identify information.
- OLS regression:

$$y_{i,d,e} = \alpha_d + \beta \text{Plough}_e + \mathbf{X}_e^H \boldsymbol{\pi} + \mathbf{X}_i \boldsymbol{\Phi} + \varepsilon_{i,d,e}$$

- $\alpha_d$ : district FE.
- $\mathbf{X}_e^H$ : historical ethnographic controls in district  $d$ , ethnicity  $e$ .
- $\mathbf{X}_i$ : individual controls (age, marital status, gender, education).

# Subnational Estimates 2

## Individual-level OLS estimates using IPUMS-I data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Dependent variable: Female labor force participation indicator								
	Bolivia 2001	Chile, 2002	Cambodia, 2008	Malaysia, 70, 80, 91, 00	Mongolia, 1989, 2000	Nepal, 2001	Philippines, 1990	Uganda, 1991, 2002	All countries
Mean of dep. var.	0.44	0.40	0.78	0.40	0.38	0.54	0.39	0.56	0.49
Traditional plough use	-0.035*** (0.002)	-0.073*** (0.003)	-0.064** (0.027)	-0.080*** (0.016)	-0.006 (0.013)	-0.100** (0.043)	0.035 (0.023)	-0.079*** (0.020)	-0.040** (0.019)
Individual & ethnicity controls	yes	yes	yes	yes	yes	yes	yes	yes	yes
District fixed effects	9	26	24	15	23	14	77	4	192
Ethnic groups	6	5	11	21	10	16	21	60	150
Observations	173,804	505,114	432,481	319,580	125,349	710,662	1,266,363	1,003,321	4,536,674
Adjusted R-squared	0.07	0.17	0.19	0.10	0.50	0.19	0.13	0.09	0.15
R-squared	0.07	0.17	0.19	0.10	0.50	0.19	0.13	0.09	0.15

## Countr-level 2SLS estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. First stage 2SLS estimates. Dependent variable: Traditional plough use								
Mean of dep. var.	0.53		0.44		0.54		0.51	
Plough-positive environment	0.744*** (0.084)	0.629*** (0.089)	0.861*** (0.078)	0.673*** (0.103)	0.820*** (0.082)	0.685*** (0.104)	0.874*** (0.089)	0.717*** (0.118)
Plough-negative environment	0.119 (0.122)	0.185 (0.133)	0.100 (0.166)	0.115 (0.171)	0.132 (0.130)	0.187 (0.141)	0.129 (0.181)	0.142 (0.188)
Equality of coefficients (p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
<i>F-stat (plough variables)</i>	40.21	25.06	66.80	21.88	51.96	21.88	49.54	18.52

# Country-Level Estimates: 2SLS

## Country-level 2SLS estimates

	Dependent variable (panels B & C):							
	Female labor force participation in 2000		Share of firms with female ownership, 2005–2011		Share of political positions held by women in 2000		Average effect size (AES)	
Plough-negative environment	18.928*** (6.506)	19.571*** (6.329)	6.072 (9.926)	9.134 (10.401)	-2.975 (6.093)	-2.868 (6.258)	0.607 (0.391)	0.653* (0.393)
Equality of coefficients (p-value)	0.00	0.00	0.02	0.02	0.56	0.47	0.00	0.00
<i>F-stat (plough variables)</i>	14.87	12.49	5.41	4.46	3.44	3.40	9.19	7.11
Panel C. Second-stage 2SLS estimates								
Traditional plough use	-21.630*** (5.252)	-25.013*** (7.513)	-17.486*** (5.533)	-22.689*** (7.620)	-6.460*** (2.334)	-9.726*** (3.750)	-0.918*** (0.225)	-1.313*** (0.388)
Hausman test (p-value)	0.02	0.04	0.56	0.40	0.22	0.10	0.33	0.16
Hansen J	0.00	0.00	0.41	0.31	0.72	0.86	0.05	0.06
Historical & contemporary controls	yes	yes	yes	yes	yes	yes	yes	yes
Continent FEs	no	yes	no	yes	no	yes	no	yes
Observations	160	160	122	122	140	140	104	104

# Cultural Transmission

- Separate cultural transmission from correlated effects from institutions.

# Cultural Transmission

- Separate cultural transmission from correlated effects from institutions.
- Country FE with WVS and district FE with IPUMS first step.

# Cultural Transmission

- Separate cultural transmission from correlated effects from institutions.
- Country FE with WVS and district FE with IPUMS first step.
- Epidemiological approach on 2nd generation immigrants within U.S. (CPS) and Europe (ESS).

# Cultural Transmission

- Separate cultural transmission from correlated effects from institutions.
- Country FE with WVS and district FE with IPUMS first step.
- Epidemiological approach on 2nd generation immigrants within U.S. (CPS) and Europe (ESS).
- OLS regression on daughters of immigrants, aged 15–64:

$$y_{i,s,c} = \alpha_s + \beta \text{Plough}_c + \mathbf{X}_c^C \boldsymbol{\Gamma} + \mathbf{X}_d^H \boldsymbol{\Pi} + \mathbf{X}_i \boldsymbol{\Phi} + \varepsilon_{i,s,c}$$

- $\alpha_s$ : state FE.
- $\mathbf{X}_c^H$ : historical ethnographic controls (domesticated animals, density of settlements, tropical climate, ...).
- $\mathbf{X}_c^C$ : contemporaneous controls (GDP per capita, flexible).
- $\mathbf{X}_i$ : individual controls (age, marital status, gender, education).



## Determinants of FLFP for US children of immigrants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Dependent variable: Labor force participation indicator, 1994–2011								
	All women			Married women					
	Woman's ancestry			Woman's ancestry			Husband's ancestry		
	Father's country	Mother's country	Parents same country	Father's country	Mother's country	Parents same country	Father's country	Mother's country	Parents same country
Mean of dep. var.	0.63	0.63	0.60	0.68	0.69	0.69	0.70	0.71	0.70
Traditional plough use	-0.044*** (0.015)	-0.043** (0.018)	-0.062*** (0.020)	-0.094** (0.046)	-0.118*** (0.043)	-0.136** (0.054)	-0.065*** (0.024)	-0.045** (0.022)	-0.058** (0.024)
Observations	57,138	55,341	32,776	10,206	9,508	6,835	35,393	35,158	23,124
Adjusted R-squared	0.23	0.23	0.25	0.10	0.10	0.11	0.08	0.08	0.08
R-squared	0.23	0.23	0.26	0.11	0.11	0.12	0.09	0.08	0.09

- Cultural transmission accounts for 35–50% of total effect.

## Determinants of gender attitudes of European children of immigrants

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variables: "When jobs are scarce..." survey response, 2004–2011					
	Father's country		Mother's country		Same country	
	1–5 scale	Indicator	1–5 scale	Indicator	1–5 scale	Indicator
Mean of dep. var.	2.54	0.32	2.53	0.32	2.62	0.35
Traditional plough use	0.219** (0.091)	0.073** (0.034)	0.214** (0.086)	0.070** (0.033)	0.298*** (0.096)	0.094** (0.038)
Observations	15,545	13,024	15,260	12,788	10,535	8,780
Adjusted R-squared	0.18	0.16	0.17	0.16	0.17	0.16
R-squared	0.18	0.17	0.17	0.16	0.17	0.17

- Cultural transmission accounts for 36–49% of total effect.

*American Economic Review* 101 (December 2011): 3221–3252  
<http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.7.3221>

## The Slave Trade and the Origins of Mistrust in Africa<sup>ii</sup>

By NATHAN NUNN AND LEONARD WANTCHEKON<sup>iii</sup>

# Nunn and Wantchekon (2011)

- How can we explain current dispersion of trust in Africa?

# Nunn and Wantchekon (2011)

- How can we explain current dispersion of trust in Africa?
- Why did Africa's slave trade have detrimental effects on economic development? [Nunn \(2008\)](#)

# Nunn and Wantchekon (2011)

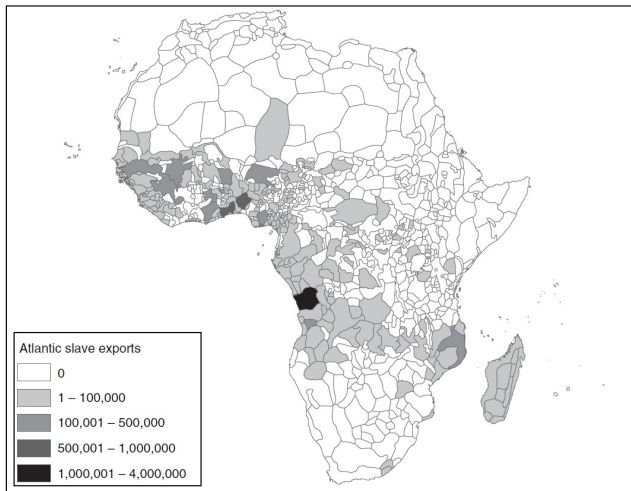
- How can we explain current dispersion of trust in Africa?
- Why did Africa's slave trade have detrimental effects on economic development? [Nunn \(2008\)](#)
- Hypothesis:
  - Slave trade = environment of ubiquitous insecurity.
  - Individuals turned on others to sell each other into slavery (including friends and family) through kidnapping or trickery.
  - A culture of mistrust might have developed because higher returns (Boyd and Richerson, 1985).
  - This cultural trait then persisted (stickiness, complementarities with institutions, self-reinforcement).

- Contemporaneous data on trust:
  - Individual-level data Afrobarometer 2005 across 17 sub-Saharan countries.
  - 21,000 individuals with identified ethnicity.
  - Questions about trust in relatives, neighbors, local government, own and other ethnic groups.

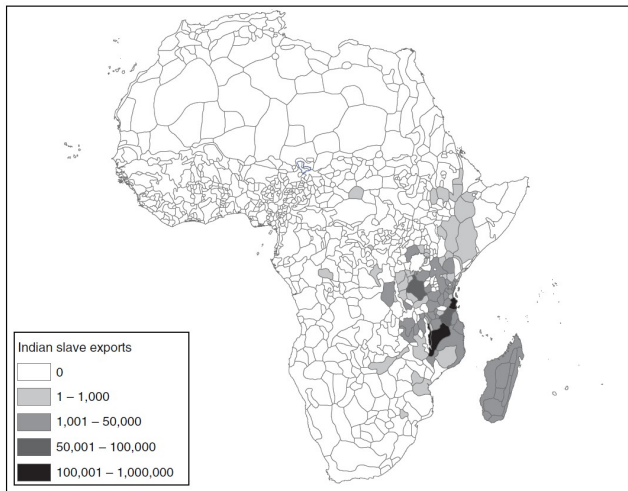
- Contemporaneous data on trust:
  - Individual-level data Afrobarometer 2005 across 17 sub-Saharan countries.
  - 21,000 individuals with identified ethnicity.
  - Questions about trust in relatives, neighbors, local government, own and other ethnic groups.
- Historical data on slave exports:
  - Number of slaves taken from each ethnic group and country from Nunn (2008).
  - Four slave trades 1400–1900.
  - Matching slave ethnicities to Afrobarometer ethnicities through Murdock Atlas.



## Transatlantic slave trade



## Indian Ocean slave trade



# Slave Trade and Contemporaneous Trust

$$\text{trust}_{i,e,d,c} = \alpha_c + \beta \text{slave exports}_e + \mathbf{X}'_{i,e,d,c} \boldsymbol{\Gamma} + \mathbf{X}'_{d,c} \boldsymbol{\Omega} + \varepsilon_{i,e,d,c}$$

- $\text{trust}_{i,e,d,c}$ : trust of individual  $i$ , ethnic group  $e$ , district  $d$ , country  $c$ .
- $\text{slave exports}_e$ : number of slaves taken from ethnic group  $e$  (log normalized by land area).
- $\mathbf{X}'_{i,e,d,c}$ : individual controls (age, gender, urban, living conditions, education, religion, occupation).
- $\mathbf{X}'_{d,c}$ : ethnic controls for district (fractionalization, share same ethnicity).
- Two-way clustering on ethnic group and district.

# Slave Trade and Contemporaneous Trust

OLS estimates of the determinants of the trust of others

	Trust of relatives (1)	Trust of neighbors (2)	Trust of local council (3)	Intra- group trust (4)	Inter- group trust (5)
$\ln(1 + \text{exports/area})$	-0.133*** (0.037)	-0.159*** (0.034)	-0.111*** (0.021)	-0.144*** (0.032)	-0.097*** (0.028)
Individual controls	Yes	Yes	Yes	Yes	Yes
District controls	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	20,062	20,027	19,733	19,952	19,765
Number of ethnicity clusters	185	185	185	185	185
Number of district clusters	1,257	1,257	1,283	1,257	1,255
$R^2$	0.13	0.16	0.20	0.14	0.11

- S.d. of trust and slave measure close to 1.  
⇒ Coefficients close to standardized.

- S.d. of trust and slave measure close to 1.  
⇒ Coefficients close to standardized.
- Alternative: compare explanatory power with other variables.
  - Standard variance decomposition.
  - Slave exports and other covariates explain 5.4% of outcome.
  - Slave exports explain 16–27% of this variation.

- Control for observables.
  - Ethnicity-level determinants of colonial rule: disease environment and precolonial prosperity.
  - Ethnicity-level characteristics of colonial rule: railway lines in 1911, European missionary contact.
  - Same results.

- Control for observables.
  - Ethnicity-level determinants of colonial rule: disease environment and precolonial prosperity.
  - Ethnicity-level characteristics of colonial rule: railway lines in 1911, European missionary contact.
  - Same results.
- Potential bias from unobservables.
  - Altonji et al (2005): how much stronger selection on unobservables must be to explain away the full estimated effect.
  - To explain main effects, selection on unobservables need to be four times greater on average.



- Control for observables.
  - Ethnicity-level determinants of colonial rule: disease environment and precolonial prosperity.
  - Ethnicity-level characteristics of colonial rule: railway lines in 1911, European missionary contact.
  - Same results.
- Potential bias from unobservables.
  - Altonji et al (2005): how much stronger selection on unobservables must be to explain away the full estimated effect.
  - To explain main effects, selection on unobservables need to be four times greater on average.
- IV using historical distance from the coast with falsification tests.

# Slave Trade and Contemporaneous Trust

OLS estimates of the determinants of the trust of others

	Trust of relatives (1)	Trust of neighbors (2)	Trust of local council (3)	Intragroup trust (4)	Intergroup trust (5)
$\ln(1 + \text{exports/area})$	-0.178*** (0.032)	-0.202*** (0.031)	-0.129*** (0.022)	-0.188*** (0.033)	-0.115*** (0.030)
Colonial population density	Yes	Yes	Yes	Yes	Yes
Ethnicity-level colonial controls	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes
District controls	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	16,709	16,679	15,905	16,636	16,473
Number of ethnicity clusters	147	147	146	147	147
Number of district clusters	1,187	1,187	1,194	1,186	1,184
$R^2$	0.13	0.16	0.21	0.16	0.12

# Slave Trade and Contemporaneous Trust

## IV estimates of the effect of the slave trade on trust

	Trust of relatives (1)	Trust of neighbors (2)	Trust of local council (3)	Intragroup trust (4)	Intergroup trust (5)
Second stage: Dependent variable is an individual's trust					
ln (1 + exports/area)	-0.190*** (0.067)	-0.245*** (0.070)	-0.221*** (0.060)	-0.251*** (0.088)	-0.174** (0.080)
Hausman test ( $p$ -value)	0.88	0.53	0.09	0.44	0.41
$R^2$	0.13	0.16	0.20	0.15	0.12
First stage: Dependent variable is ln (1 + exports/area)					
Historical distance of ethnic group from coast	-0.0014*** (0.0003)	-0.0014*** (0.0003)	-0.0014*** (0.0003)	-0.0014*** (0.0003)	-0.0014*** (0.0003)
Colonial population density	Yes	Yes	Yes	Yes	Yes
Ethnicity-level colonial controls	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes
District controls	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	16,709	16,679	15,905	16,636	16,473
Number of clusters	147 / 1,187	147 / 1,187	146 / 1,194	147 / 1,186	147 / 1,184
$F$ -stat of excl. instrument	26.9	26.8	27.4	27.1	27.0
$R^2$	0.81	0.81	0.81	0.81	0.81

# Slave Trade and Contemporaneous Trust

Falsification test: reduced form in Africa and Asia

	Trust of local government council			
	Afrobarometer sample		Asiabarometer sample	
	(1)	(2)	(3)	(4)
Distance from the coast	0.00039*** (0.00009)	0.00031*** (0.00008)	-0.00001 (0.00010)	0.00001 (0.00009)
Country fixed effects	Yes	Yes	Yes	Yes
Individual controls	No	Yes	No	Yes
Number of observations	19,913	19,913	5,409	5,409
Number of clusters	185	185	62	62
$R^2$	0.16	0.18	0.19	0.22

# Mechanisms: External vs Internal Factors

- Slave trade might have weakened institutions, generating mistrust.

# Mechanisms: External vs Internal Factors

- Slave trade might have weakened institutions, generating mistrust.
- Measure of this: mistrust in local government.
  - Control for trust in local government.
  - Use various measures through fixed effects.
  - Half of the overall effect can be explained by effect through quality of local institutions.

# Mechanisms: External vs Internal Factors

- Slave trade might have weakened institutions, generating mistrust.
- Measure of this: mistrust in local government.
  - Control for trust in local government.
  - Use various measures through fixed effects.
  - Half of the overall effect can be explained by effect through quality of local institutions.
- Effects through untrustworthiness of others.
  - Measure of intergroup trust in local area: weighted average slave export intensity of other ethnic groups in same area.
  - No effect from deteriorated intergroup trust.

# Mechanisms: External vs Internal Factors

## Identifying channels of causality

	Intergroup trust				
	Trust of local council		Within town	Within district	Within province
	(1)	(2)	(3)	(4)	(5)
Ethnicity-based slave export measure (baseline measure)	-0.072*** (0.019)	-0.070*** (0.019)	-0.102*** (0.028)	-0.120*** (0.027)	-0.098*** (0.029)
Average slave export measure among other ethnicities in the same location			-0.037 (0.029)	-0.063** (0.030)	-0.091*** (0.035)
Council trustworthiness fixed effects	Yes	Yes	No	No	No
Five public goods fixed effects	No	Yes	No	No	No
Colonial population density	Yes	Yes	Yes	Yes	Yes
Ethnicity-level colonial controls	Yes	Yes	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	12,827	12,203	9,673	12,513	15,999
Number of clusters	146/1,172	145/1,130	147/725	147/737	147/1,127
$R^2$	0.37	0.37	0.12	0.12	0.12



# Mechanisms: External vs Internal Factors

- Effect of slave trade through external environment (institutions, trustworthiness of others) versus individuals' internal norms.

# Mechanisms: External vs Internal Factors

- Effect of slave trade through external environment (institutions, trustworthiness of others) versus individuals' internal norms.
- Location-based measure slave intensity: slaves taken from area of residence.
- When both variables included, identification from movers (45% of sample).

# Mechanisms: External vs Internal Factors

- Effect of slave trade through external environment (institutions, trustworthiness of others) versus individuals' internal norms.
- Location-based measure slave intensity: slaves taken from area of residence.
- When both variables included, identification from movers (45% of sample).
- Ethnicity-based measures decrease by 10–15%, so explains 85–90% of the overall effect.
- It is always twice as large in magnitude than location-based effects.

# Mechanisms: External vs Internal Factors

## Identifying channels of causality

	Trust of relatives (1)	Trust of neighbors (2)	Trust of local council (3)	Intragroup trust (4)	Intergroup trust (5)
Ethnicity-based slave export measure (baseline measure)	-0.155*** (0.029)	-0.182*** (0.029)	-0.100*** (0.023)	-0.169*** (0.033)	-0.090*** (0.030)
Location-based slave export measure	-0.045*** (0.014)	-0.045*** (0.016)	-0.045** (0.018)	-0.043** (0.018)	-0.047** (0.020)
Colonial population density	Yes	Yes	Yes	Yes	Yes
Ethnicity-level colonial controls	Yes	Yes	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	15,999	15,972	15,221	15,931	15,773
Number of clusters	146/269	146/269	145/272	146/269	146/269
$R^2$	0.13	0.16	0.20	0.16	0.12

# THE INFLUENCE OF ANCESTRAL LIFEWAYS ON INDIVIDUAL ECONOMIC OUTCOMES IN SUB-SAHARAN AFRICA

---

**Stelios Michalopoulos**  
Brown University

**Louis Putterman**  
Brown University

**David N. Weil**  
Brown University

# Two Approaches to History

- Place-based approach:
  - Focuses on institutions and geography (Neolithic transition).
  - More developed: easy to map space across time.
- Individual-based approach:
  - Focuses on lineage, cultural transmission, intergenerational mobility.
  - Less developed: little data on lineage.

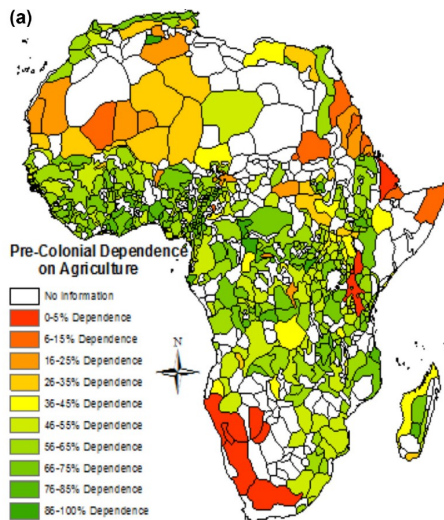
- Does lineage to preindustrial subsistence matter for economic outcomes today in sub-Saharan Africa?
- Compare pastoralist vs agriculture ancestral subsistence lifeways.
- Focus on Africa because easy to match individuals to ethnic groups and ancestral lifeways.
- Distinguish between institution and cultural channels through within location variation.

The introduction of location fixed effects is crucial, since it allows us to absorb characteristics related to the geographic, ecological, and institutional environment of a given region that recent studies have highlighted as important determinants of regional African development. Moreover, it allows us to uncover the importance of portable ethnic-specific traits whose influence is not limited to the ancestral homeland of a given group. This methodology is similar to Nunn and Wantchekon (2011), who investigate the impact of slavery on individual trust among respondents residing outside their ethnic enclaves.



- Ethnicity data:
  - DHS surveys.
  - Information for both ethnicity and enumeration area coordinates.
  - 337k respondents across 21 countries, 492 ethnicity-country groups.
- Matching modern ethnicities to ancestral groups:
  - Murdock's Ethnographic Atlas.
  - Match DHS ethnicities to those in atlas.
  - Average distance moved: residence vs nearest border of ancestral homeland (51% movers).

- Modern outcomes (DHS):
  - Education: categories and years completed.
  - Wealth: quintiles.
- Historical mode of subsistence:
  - Activities in Ethnographic Atlas: gathering, hunting, fishing, animal husbandry, agriculture.
  - Shares of subsistence in 9 bands.
  - In regressions: pastoralism (husbandry) is default activity, focal independent variable on degree of reliance on agriculture (the rest are controls).



# Predictive Power of Ethnicity

$R^2$ from FE regressions	All education	All wealth	Movers education	Movers wealth
Country FE	0.159	0.013	0.147	0.038
Homeland FE	0.291	0.231	0.290	0.295
Ethnicity FE	0.265	0.138	0.257	0.182
Country-ethnicity FE	0.281	0.159	0.282	0.209
Country-homeland FE	0.301	0.248	0.304	0.319
Country-homeland FE and country-ethnicity FE	0.325	0.283	0.337	0.362
Observations	285,255	285,263	154,744	154,747

# The Influence of Ancestral Characteristics

$$y_{i,e,h,c,v} = \beta \text{Agriculture}_e + \delta \mathbf{X}_{i,e,h,c,v} + \alpha_c \cdot \alpha_{h,c} \cdot \alpha_v + \varepsilon_{i,e,h,c,v}$$

- $y_{i,e,h,c,v}$ : outcome for individual  $i$  of ethnicity  $e$  residing in homeland  $h$  in country  $c$  and enumeration area  $v$ .
- $\text{Agriculture}_e$ : subsistence share of agriculture of ancestral group  $e$ .
- $\mathbf{X}_{i,e,h,c,v}$ : controls (age, female, mover status).
- $\alpha_c$ : country FE.
- $\alpha_{h,c}$ : country-homeland FE.
- $\alpha_v$ : enumeration area FE.
- Clustering: ethnicity level.

# The Influence of Ancestral Characteristics

## Benchmark: DHS regressions within ethnic homelands

Variables	(1) Education	(2) Education	(3) Education	(4) Education	(5) Wealth	(6) Wealth	(7) Wealth	(8) Wealth
Agriculture	0.2337*** (0.0432)	0.1498*** (0.0259)	0.1034*** (0.0212)	0.1011*** (0.0191)	0.2233*** (0.044)	0.1797*** (0.0324)	0.0970*** (0.0221)	0.0972*** (0.0219)
Gather/hunt/fish	0.2095*** (0.0547)	0.1026*** (0.0214)	0.0918*** (0.0200)	0.0911*** (0.0185)	0.1574*** (0.0457)	0.0681*** (0.0249)	0.0488*** (0.0152)	0.0524*** (0.0146)
Urban			0.9199*** (0.0357)	0.7079*** (0.0292)			1.6401*** (0.0460)	1.4330*** (0.0468)
Simple controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	No	No	No	Yes	No	No	No
Country-ethnic homeland FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Occupation FE	No	No	No	Yes	No	No	No	Yes
Observations	285,192	285,192	285,192	285,192	285,200	285,200	285,200	285,200
R <sup>2</sup>	0.239	0.350	0.418	0.493	0.052	0.266	0.491	0.518

- Country characteristics only explain 1/3.
- Part explained by migration to urban locations.
- Differences in ancestral lifeways explain 1/3 of individual variation in outcomes due to ethnic identity.

# The Influence of Ancestral Characteristics

Benchmark: DHS regressions within villages

Variables	(1) Education	(2) Education	(3) Education	(4) Wealth	(5) Wealth	(6) Wealth
Agriculture	0.1034*** (0.0212)	0.0731*** (0.0147)	0.0694*** (0.0128)	0.0970*** (0.0221)	0.0389*** (0.0070)	0.0379*** (0.0068)
Gather/hunt/fish	0.0918*** (0.0200)	0.0708*** (0.0155)	0.0681*** (0.0140)	0.0488*** (0.0152)	0.0176** (0.0068)	0.0190*** (0.0065)
Urban	0.9199*** (0.0357)			1.6401*** (0.0460)		
Simple controls	Yes	Yes	Yes	Yes	Yes	Yes
Country-ethnic Homeland FE	Yes	No	No	Yes	No	No
Coordinates FE	No	Yes	Yes	No	Yes	Yes
Occupation FE	No	No	Yes	No	No	Yes
Observations	285,192	285,192	285,192	285,200	285,200	285,200
R <sup>2</sup>	0.418	0.506	0.559	0.491	0.677	0.682

- Shifting from pastoralism to agriculture raises education by 0.28 points, or 0.8 years of education.
- Shifting from pastoralism to agriculture raises wealth by 20% within

# Selection into Migration

- Identification based on movers (54% of sample).
- Differential selection into migration across lifeway groups?
- Assess if ancestral lifeway predicts migration.
- Explicitly assess selective migration.



# Selection into Migration

## Determinants of migration

Variables	(1) Mover	(2) Mover	(3) Moved in life	(4) Moved in life
Agriculture	-0.0465** (0.0212)	-0.0464** (0.0211)	0.0052 (0.0052)	0.0048 (0.0050)
Gather/hunt/fish	-0.1179*** (0.0254)	-0.1177*** (0.0253)	-0.0062 (0.0053)	-0.0061 (0.0051)
Simple controls	Yes	Yes	Yes	Yes
Coordinates FE	Yes	Yes	Yes	Yes
Occupation FE	No	Yes	No	Yes
Observations	285,200	285,200	188,304	188,304
$R^2$	0.768	0.768	0.206	0.210

# Selection into Migration

## Determinants of migration

Variables	(1) Education	(2) Education	(3) Wealth	(4) Wealth
Agriculture	0.2808*** (0.0607)	0.2710*** (0.0543)	0.2431*** (0.0559)	0.2368*** (0.0483)
Gather/hunt/fish	0.2816*** (0.0683)	0.2747*** (0.0638)	0.1826*** (0.0594)	0.1879*** (0.0499)
Mover	0.7459** (0.3505)	0.7805** (0.3280)	0.5330 (0.3417)	0.5994* (0.3029)
Mover × agriculture	-0.0570 (0.0515)	-0.0731 (0.0482)	-0.0251 (0.0441)	-0.0462 (0.0393)
Mover × gather/hunt/fish	-0.1085** (0.0482)	-0.1155*** (0.0441)	-0.0372 (0.0562)	-0.0529 (0.0472)
Simple controls	Yes	Yes	Yes	Yes
Occupation FE	No	Yes	No	Yes
Country FE	Yes	Yes	Yes	Yes
Observations	285,192	285,192	285,200	285,200
$R^2$	0.241	0.38	0.053	0.222

## **THE LONG-TERM EFFECT OF DEMOGRAPHIC SHOCKS ON THE EVOLUTION OF GENDER ROLES: EVIDENCE FROM THE TRANSATLANTIC SLAVE TRADE**

---

**Edoardo Teso**

Harvard University and IQSS

- What are the mechanisms of persistence of historical shocks?
- The case of gender norms in Africa and the slave trade.

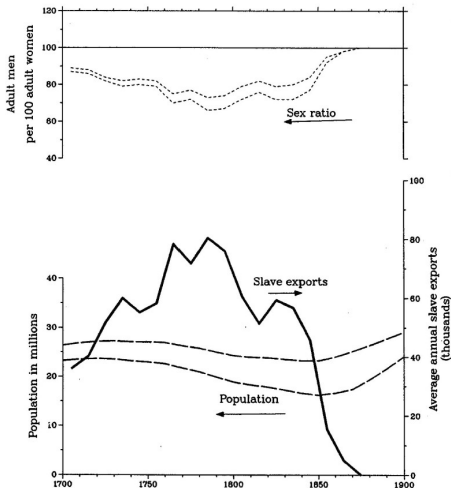
# Teso (2019)

- What are the mechanisms of persistence of historical shocks?
- The case of gender norms in Africa and the slave trade.
- Alternative framing: explain current distribution in FLFP in Africa.

- What are the mechanisms of persistence of historical shocks?
- The case of gender norms in Africa and the slave trade.
- Alternative framing: explain current distribution in FLFP in Africa.
- Hypothesis:
  - Male slaves outnumbered female slaves.
  - This generated historically skewed sex ratios in Africa.
  - Women had to substitute for men in previously male-dominated activities.
  - This changed attitudes and beliefs in the long run, affecting FLFP.
  - Multiple equilibria, intergenerational learning, institutional change.

# Historical Experiment

## Demographic impact of the transatlantic slave trade

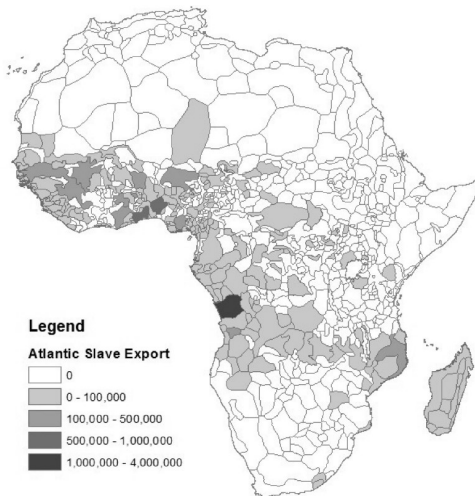


- Contemporaneous data.
  - FLFP data from DHS (61 surveys, 21 countries, 1992–2014).
  - Sample of 662k women aged 15–49.
  - Also information on occupation, fertility, age at first birth, attitudes, household tasks.



- Contemporaneous data.
  - FLFP data from DHS (61 surveys, 21 countries, 1992–2014).
  - Sample of 662k women aged 15–49.
  - Also information on occupation, fertility, age at first birth, attitudes, household tasks.
- Historical data.
  - Slaves taken from Nunn and Wantchekon (2011).
  - Ethnicity data from Murdock's Atlas (1959).
  - Matched with DHS: 583k women and 223k men.

## Ethnic group-level exposure to transatlantic slave trade



$$y_{i,e,c} = \alpha_c + \beta \text{Trades}_e + X'_{i,e,c} \Delta + Z'_e \Omega + \varepsilon_{i,e,c}$$

- $y_{i,e,c}$ : outcome of woman  $i$ , in country  $c$ , ethnic group  $e$ .
- $\text{Trades}_e$ : slaves taken in transatlantic or indian ocean slave trades, normalized by historical area.
- $X'_{i,e,c}$ : individual controls (age, marital status, urban, religion).
- $Z'_e$ : ethnicity controls (disease environment, precolonial urbanization, jurisdictional hierarchies, population density, colonial railroads, missions, agricultural practices, historical warfare).
- Clustering: ethnic group level.

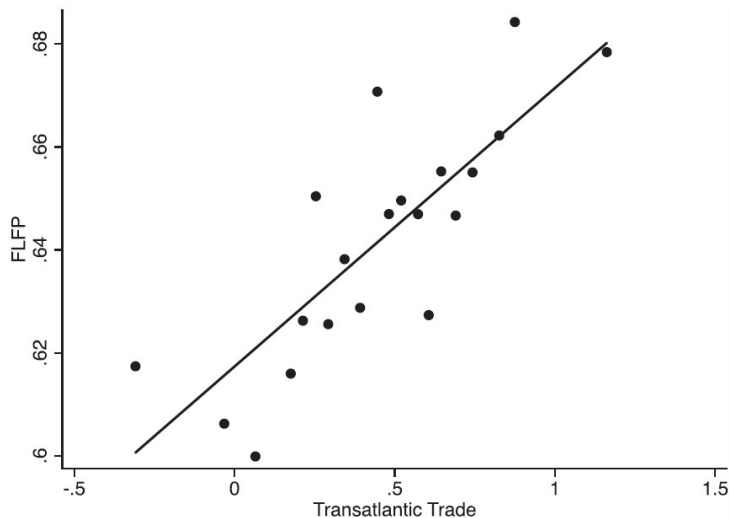
# Long-Run Impact of Slave Trade on FLFP

OLS estimates, the effect of the slave trade on FLFP

	FLFP (1)	FLFP (2)	FLFP (3)	FLFP (4)	FLFP (5)	FLFP (6)
Transatlantic Trade	0.048*** (0.013)	0.054*** (0.011)	0.059*** (0.013)	0.056*** (0.011)	0.073*** (0.012)	0.072*** (0.012)
Indian Ocean Trade	-0.059 (0.140)	-0.120 (0.158)	-0.061 (0.175)	-0.111 (0.174)	-0.146 (0.205)	-0.133 (0.196)
Observations	583,562	563,379	470,183	563,054	386,503	386,317
R-squared	0.16	0.17	0.18	0.18	0.14	0.14
Ethnic Groups	261	243	170	243	241	241
Country-survey FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Historical Controls	No	Yes	Yes	Yes	Yes	Yes
Hist. Part. Agriculture	No	No	Yes	No	No	No
Education	No	No	No	Yes	Yes	Yes
Polygyny	No	No	No	No	Yes	Yes
Transatlantic std. dev.	0.564	0.564	0.570	0.564	0.564	0.564
Indian Ocean std. dev.	0.033	0.031	0.034	0.031	0.031	0.031
Dep. var. mean unaffected	0.588	0.586	0.589	0.586	0.635	0.635

# Long-Run Impact of Slave Trade on FLFP

Residuals plot (Column 2)



## Magnitudes (Column 2)

- $\uparrow 1$  s.d. ancestor exposure to slave trade  $\implies \uparrow 3$ pp FLFP.
- $\uparrow 1$  s.d. ancestor exposure to slave trade  $\implies \uparrow 5\%$  FLFP mean.
- $\uparrow 1$  s.d. ancestor exposure to slave trade  $\implies \uparrow 0.06$  FLFP s.d.

# Long-Run Impact of Slave Trade on FLFP

OLS estimates, the effect of the slave trade on occupational choices

	Agriculture (1)	Clerical (2)	Manual (3)	Domestic (4)	High Ranking (5)
Transatlantic Trade	0.018 (0.016)	-0.000 (0.001)	-0.010 (0.008)	-0.004** (0.002)	0.048*** (0.011)
Observations	549,009	549,009	549,009	549,009	549,009
R-squared	0.23	0.02	0.05	0.07	0.14
Ethnic Groups	243	243	243	243	243
Country-survey FE	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes
Historical Controls	Yes	Yes	Yes	Yes	Yes
Transatlantic std. dev.	0.564	0.564	0.564	0.564	0.564
Dep. var. mean unaffected	0.276	0.011	0.061	0.026	0.224

- High ranking.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\uparrow$  3pp high ranking.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\uparrow$  12% high ranking mean.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\uparrow$  0.06 high ranking s.d.



# Magnitudes

- High ranking.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\uparrow$  3pp high ranking.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\uparrow$  12% high ranking mean.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\uparrow$  0.06 high ranking s.d.
- Domestic.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\downarrow$  0.2pp domestic.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\downarrow$  9% domestic mean.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\downarrow$  0.19 domestic s.d.
- But think of base rates.

# Male Falsification Test

OLS estimates, women's versus men's employment

Sample	Men (1)	Women (2)	Men (3)	Women (4)	Men (5)	Women (6)
Transatlantic Trade	-0.010** (0.005)	0.050*** (0.013)	-0.012** (0.006)	0.054*** (0.010)	-0.008 (0.005)	0.056*** (0.010)
Observations	222,970	548,178	216,419	528,006	216,125	527,687
R-squared	0.31	0.16	0.31	0.17	0.32	0.18
Ethnic Groups	235	261	219	243	219	243
Country-survey FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Historical Controls	No	No	Yes	Yes	Yes	Yes
Education	No	No	No	No	Yes	Yes
Transatlantic std. dev.	0.564	0.567	0.563	0.567	0.563	0.567
Dep. var. mean unaffected	0.831	0.593	0.831	0.591	0.831	0.591

# Magnitudes (Columns 3 and 4)

- Women.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\uparrow$  3pp FLFP.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\uparrow$  5% FLFP mean.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\uparrow$  0.06 FLFP s.d.

# Magnitudes (Columns 3 and 4)

- Women.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\uparrow$  3pp FLFP.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\uparrow$  5% FLFP mean.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\uparrow$  0.06 FLFP s.d.
- Men.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\downarrow$  0.7pp MLFP.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\downarrow$  0.8% MLFP mean.
  - $\uparrow$  1 s.d. ancestor exposure  $\implies$   $\downarrow$  0.02 MLFP s.d.

# The Marriage Market Channel

- Hypothesis from Fernandez et al (2004):
  - Working mothers transmit positive views about female labor to their sons.
  - Women whose husbands' ancestors were more exposed to slave trade are more likely to have working mother, and thus working wife.

# The Marriage Market Channel

- Hypothesis from Fernandez et al (2004):
  - Working mothers transmit positive views about female labor to their sons.
  - Women whose husbands' ancestors were more exposed to slave trade are more likely to have working mother, and thus working wife.
- Use women's ancestry FE to isolate role of husband's origins.

# The Marriage Market Channel

- Hypothesis from Fernandez et al (2004):
  - Working mothers transmit positive views about female labor to their sons.
  - Women whose husbands' ancestors were more exposed to slave trade are more likely to have working mother, and thus working wife.
- Use women's ancestry FE to isolate role of husband's origins.
- To assess relative magnitude, use husband ancestry FE.

# The Marriage Market Channel

- Hypothesis from Fernandez et al (2004):
  - Working mothers transmit positive views about female labor to their sons.
  - Women whose husbands' ancestors were more exposed to slave trade are more likely to have working mother, and thus working wife.
- Use women's ancestry FE to isolate role of husband's origins.
- To assess relative magnitude, use husband ancestry FE.
- Cannot have causal interpretation:
  - Husbands have an impact, holding beliefs constant.
  - Selection into marriage with different types of men.



# The Marriage Market Channel

OLS estimates, the marriage market channel

	FLFP (1)	FLFP (2)	FLFP (3)
Transatlantic Trade	0.071*** (0.015)		0.045*** (0.009)
Transatlantic Trade Husband		0.021** (0.009)	
Observations	109,310	109,294	109,293
R-squared	0.14	0.18	0.17
Ethnic Groups	232	228	232
Country-survey FE	Yes	No	No
Country-survey-woman's ethnicity FE	No	Yes	No
Country-survey-husband's ethnicity FE	No	No	Yes
Individual Controls	Yes	Yes	Yes
Historical Controls	Yes	Yes	No
Transatlantic std. dev.	0.558	0.559	0.558
Dep. var. mean unaffected	0.652	0.657	0.652

# Cultural Transmission

- Separate culture and institutions effects.
- Epidemiological approach using enumeration area FE.
- Find that half is due to cultural transmission.
- Not affected by heterogeneous response of movers relative to non-movers.

## OLS estimates, the cultural transmission channel

	FLFP (1)	FLFP (2)	FLFP (3)	FLFP (4)	FLFP (5)
Transatlantic Trade	0.027*** (0.007)	0.029*** (0.005)	0.029*** (0.005)	0.036*** (0.006)	0.035*** (0.006)
Observations	583,377	563,092	562,766	386,121	385,935
R-squared	0.32	0.32	0.32	0.33	0.33
Ethnic Groups	261	243	243	241	241
EA-survey FE	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes
Historical Controls	No	Yes	Yes	Yes	Yes
Education	No	No	Yes	No	Yes
Polygyny	No	No	No	Yes	Yes
Transatlantic std. dev.	0.564	0.564	0.564	0.564	0.564
Dep. var. mean unaffected	0.588	0.586	0.586	0.635	0.635

## OLS estimates, women's empowerment

	Share HH Decisions (1)	Share Violence (2)	Share Violence (3)	Rights Politics (4)	Rights Politics (5)	Rights General (6)	Rights General (7)
Transatlantic Trade	0.048*** (0.013)	-0.008 (0.011)	-0.005 (0.009)	0.018 (0.031)	-0.057* (0.031)	0.098*** (0.037)	-0.019 (0.057)
Observations	337,994	426,485	163,173	40,394	40,536	24,215	24,389
R-squared	0.26	0.22	0.11	0.07	0.07	0.10	0.09
Ethnic Groups	223	225	189	275	275	261	262
Sample	DHS	DHS	DHS	Afrob.	Afrob.	Afrob.	Afrob.
Gender	Female	Female	Male	Female	Male	Female	Male
Transatlantic std. dev.	0.567	0.570	0.564	0.536	0.536	0.540	0.538
Dep. var. mean unaffected	0.451	0.336	0.198	4.031	3.578	4.081	3.712

**The Intergenerational Transmission of World War I  
on Female Labour**

Victor Gay\*

March 2023

Forthcoming at *The Economic Journal*

# Gay (2023)

- What are the mechanisms of intergenerational transmission?
- What is the process of cultural change?

- What are the mechanisms of intergenerational transmission?
- What is the process of cultural change?
- Use historical shock to FLFP: WWI in France.
- Cultural proxy: WWI military fatalities ( $\implies$  interwar  $\uparrow$  FLFP)
- Outcomes: FLFP and attitudes toward gender roles.
- Hypotheses:
  - Parental transmission [Fernández \(2013\)](#)
  - Transmission through marriage  
[Fernández, Fogli and Olivetti \(2004\)](#)
  - Oblique and horizontal social spillovers  
[Fernández \(2013\)](#) [Fogli and Veldkamp \(2011\)](#)

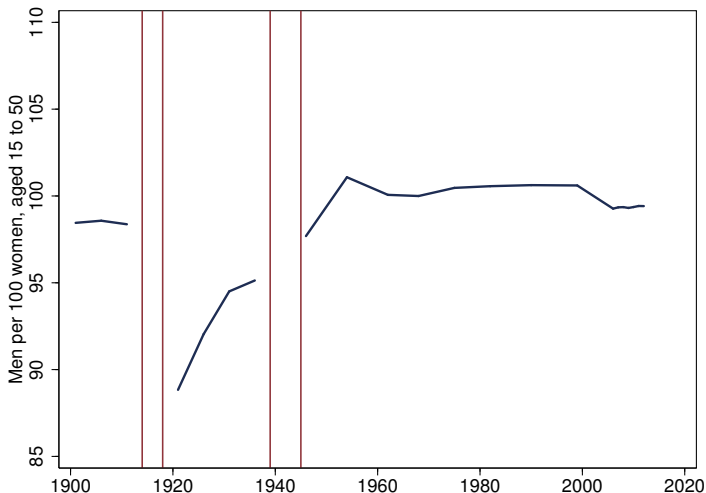
# WWI Military Fatalities: Context

- WWI: 1914–1918, mostly in North-Eastern France.
- Universal conscription system (men aged 20–48):
  - 8 million drafted out of 10 million men aged 15–50 in 1914.
  - 1.3 million died in combats  $\implies$  death rate of 16%.



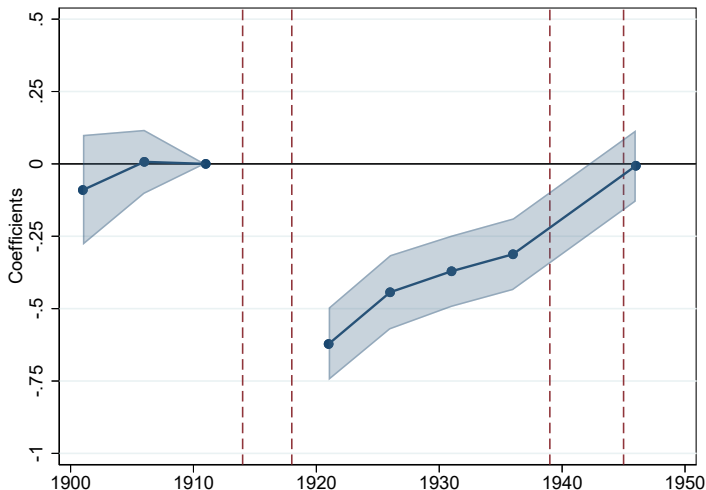
# WWI Military Fatalities: Impact on Sex Ratio

Adult Sex Ratio (1901–2012)



# WWI Military Fatalities: Impact on Sex Ratio

## Difference-in-Differences Estimates



# WWI Military Fatalities: Data

- Original dataset:
  - Individual military records for all 1.3 million deceased French soldiers.
  - Information: département of birth, dates of birth and death.

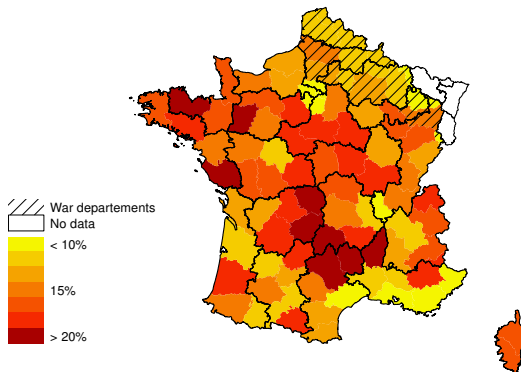
# WWI Military Fatalities: Data

- Original dataset:
  - Individual military records for all 1.3 million deceased French soldiers.
  - Information: département of birth, dates of birth and death.
- Measure of military death rates at the département level:

$$\text{death\_rate}_d = \frac{\text{dead soldiers born in } d}{\underbrace{\text{male pop. aged 15–44 in } d \text{ in 1911}}_{\text{Draft pool in census 1911}}}$$

- Summary statistics:
  - Mean = 15%, 25th percentile  $\simeq$  10%, 75th percentile  $\simeq$  20%.
  - Minimum = 6%, maximum = 29%.

# WWI Military Fatalities: Distribution



Military Death Rates Across 87 Départements

# WWI Military Fatalities: Data Generating Process

- Until end 1914: territorial organization of military recruitment.
- Policies by the Ministry of War to support the industrial war effort:  
⇒ Up to 15% of soldiers allocated to war industries.
- Pre-war characteristics:
  - Correlated with measures of rurality.
  - Uncorrelated with other pre-war characteristics.
  - Uncorrelated with pre-war levels and trends in FLFP.

# Military Fatalities and FLFP: Interwar Period

- Main results. [Boehnke and Gay \(2022\)](#)
  - Data: 7 censuses 1901–1936, département level.
  - Empirical strategy: difference-in-differences.
  - Estimate: 20% vs 10% death rates  $\implies \uparrow$  4pp.
  - Magnitude:  $\sim$  12% of pre-war FLFP mean (0.4 s.d.).

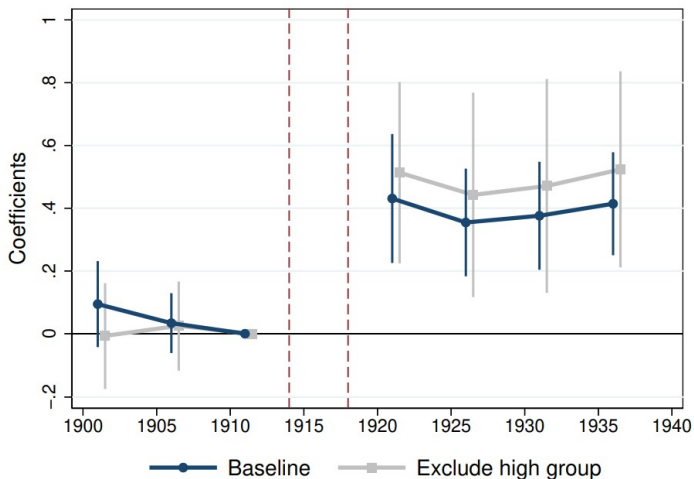
# Military Fatalities and FLFP: Interwar Period

- Main results. [Boehnke and Gay \(2022\)](#)
  - Data: 7 censuses 1901–1936, département level.
  - Empirical strategy: difference-in-differences.
  - Estimate: 20% vs 10% death rates  $\implies \uparrow 4\text{pp}$ .
  - Magnitude:  $\sim 12\%$  of pre-war FLFP mean (0.4 s.d.).
- Mechanisms.
  - Evidence for labor supply channels (war widows, single women).
  - Little evidence for labor demand channels (substitution).
  - Little role of female wartime employment.



# Military Fatalities and FLFP: Interwar Period

Interwar impact of WWI military fatalities on FLFP



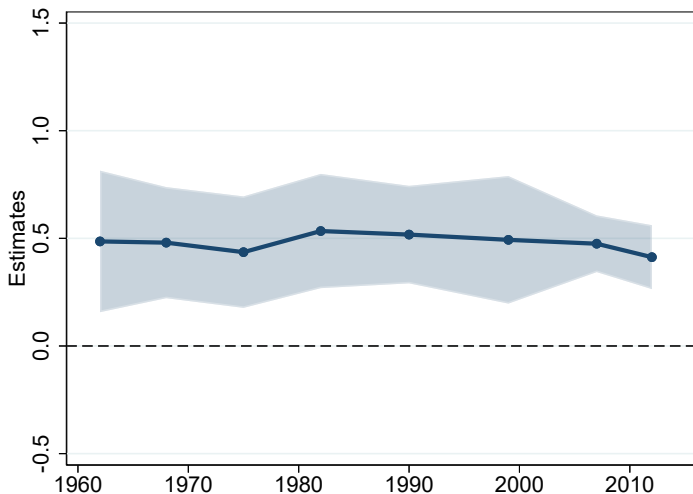
# Military Fatalities and FLFP: Long Run

- Data: 8 micro census 1962–2012 (20–25% samples).
- Sample: **internal migrant** French married women aged 30–49 (3.5 million observations).
- Empirical strategy: repeated cross-sections

$$\text{employed}_{ibrt} = \beta \text{death\_rate}_b + \alpha \mathbf{X}'_{1911,b} + \gamma_i \cdot \delta_r \cdot \mu_{1914,b} + \varepsilon_{ibrt}$$

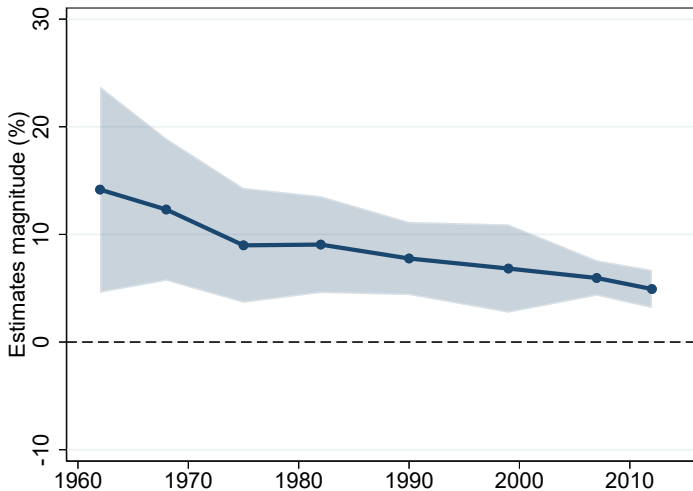
- $i$  = individual,  $b$  = dép. of birth,  $r$  = dép. of residence,  $t$  = census.
  - $\mathbf{X}'_{1911,b}$ : pre-war controls (rurality, determinants of FLFP).
  - $\gamma_i$ : year of birth FE.
  - $\delta_r$ : département of residence FE ( $\sim 90$  dép).
  - $\mu_{1914,b}$ : region FE ( $\sim 4$ – $5$  dép per region).
- Two-way clustering: département of birth and of residence.

# Military Fatalities and FLFP: Long Run



- Born in dép 20% vs 10% death rates  $\implies$  4–6 pp more likely to work.

# Military Fatalities and FLFP: Long Run



# Epidemiological versus Location

- Strategy 1: location-based approach.
  - Sample: non-migrant French married women aged 30–49.
  - Identification: across locations.
  - Capture culture and location effects.

$$\text{employed}_{irt} = \beta \text{death\_rate}_r + \alpha \mathbf{X}'_{1911,r} + \gamma_i \cdot \mu_{1914,r} + \varepsilon_{irt}$$

# Epidemiological versus Location

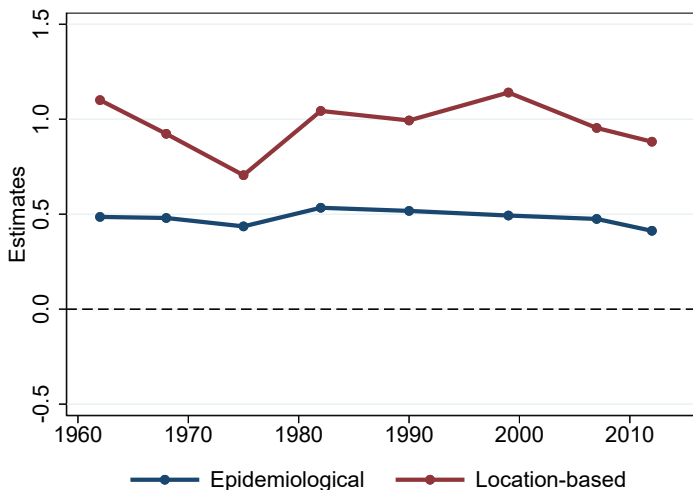
- Strategy 1: location-based approach.
  - Sample: non-migrant French married women aged 30–49.
  - Identification: across locations.
  - Capture culture and location effects.

$$\text{employed}_{irt} = \beta \text{death\_rate}_r + \alpha \mathbf{X}'_{1911,r} + \gamma_i \cdot \boldsymbol{\mu}_{1914,r} + \varepsilon_{irt}$$

- Strategy 2: inverted epidemiological.
  - Sample: migrant French married women aged 30–49.
  - Identification: across locations within origins.
  - Capture location effects independent of culture.

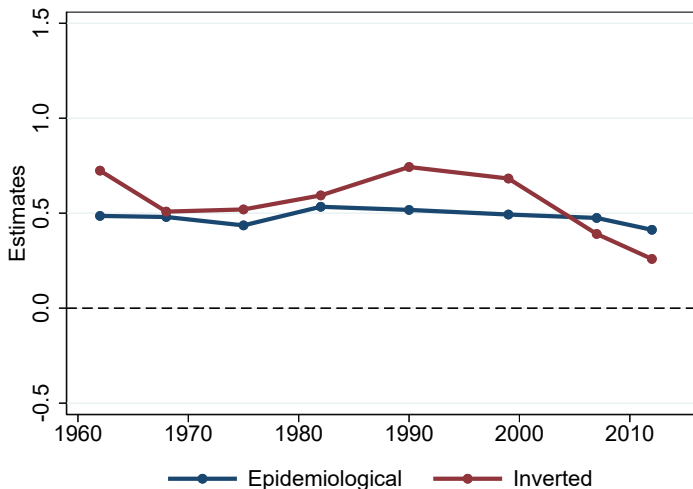
$$\text{employed}_{ibr} = \beta \text{death\_rate}_r + \alpha \mathbf{X}'_{1911,r} + \gamma_i \cdot \boldsymbol{\omega}_b \cdot \boldsymbol{\mu}_{1914,r} + \varepsilon_{ibr}$$

# Epidemiological versus Location



- Cultural transmission  $\simeq$  43–62% of total effect.

# Epidemiological versus Location



- Cultural transmission  $\simeq$  39–60% of total effect.



# Mechanisms of Intergenerational Transmission

Transmission on behavior side:

- Parental transmission: mothers and fathers to daughters.
- Transmission through marriage:
  - Husbands to wives.
  - Mothers in-law to daughters in-law.
- Local social interactions.

# Vertical Transmission: Data

- Mother-to-daughter channel.
- Data: extended labor surveys 2005–2012 (contain parental origins).
- Sample: second-generation internal migrant married women.
  - Second-generation internal migrant = non-migrant with both parents born in another département than daughter.

# Vertical Transmission: Specification

- Specification:

$$\text{employed}_{imfr} = \beta \text{death\_rate}_m + \alpha \mathbf{X}'_{1911,m} + \gamma_i + \delta_r + \mu_{1914,m} + \omega_f + \varepsilon_{imfr}$$

- $i$  = individual,  $m$  = mother,  $f$  = father.
  - $\tilde{\mathbf{X}}'_{1911,m}$ : mother pre-war controls.
  - $\gamma_i$ : cohort FE.
  - $\delta_r$ : dép of residence FE.
  - $\mu_{1914,m}$ : mother region FE.
  - $\omega_f$ : father birth dép FE.
- 
- Two-way clustering: département of residence and mothers' birth.

# Vertical Transmission: Identification

Identify  $\beta$  from variations in the working behavior of second-generation internal migrant women:

- residing in the same département
- born in the same département
- whose fathers were born in the same département
- but whose mothers were born in neighboring départements

# Vertical Transmission: Results

Dependent variable:	Employed		
	(1)	(2)	(3)
Military death rate, mother	1.05*** [0.37]	1.11*** [0.40]	1.24*** [0.35]
Pre-war controls, mother (1911)	Yes	Yes	Yes
Birth-residence dép. FE	Yes	Yes	Yes
Mother birth region FE	No	Yes	Yes
Father birth dép. FE	No	Yes	Yes
Individual controls	No	No	Yes
Household controls	No	No	Yes
Observations	15,095	15,095	15,095

\*\*\* Significant at the 1 percent level.

# Vertical Transmission: Mothers vs Fathers

Dependent variable:	Employed		
	(1)	(2)	(3)
Same $\times$ Military death rate (parents)	1.19*** [0.42]	1.08* [0.60]	1.09* [0.60]
Different $\times$ Military death rate (mother)	0.73* [0.43]	0.86* [0.47]	0.94** [0.45]
Different $\times$ Military death rate (father)	0.46 [0.34]	0.13 [0.60]	0.07 [0.60]
Parents birth region FE	No	Yes	Yes
Other controls	No	No	Yes
Observations	15,095	15,095	15,095

Includes pre-war controls (parents), birth-residence dép. FE, parents birth region FE.

# Vertical Transmission: 2SLS

Dependent variable:	Mother worked	Employed
	FS	SS
Military death rate (mother)	1.94*** [0.64]	
Mother worked		0.49** [0.23]
Pre-war controls, mother (1911)	Yes	Yes
Birth-residence dép. FE	Yes	Yes
Mother birth region FE	Yes	Yes
Father birth dép. FE	Yes	Yes
Observations	15,095	15,095

F-stat: 44.4

# Transmission Through Marriage

- Husband to wife channel.
- Data: censuses 1962–2012.
- Sample: all women aged 30–49 married with internal migrant man.

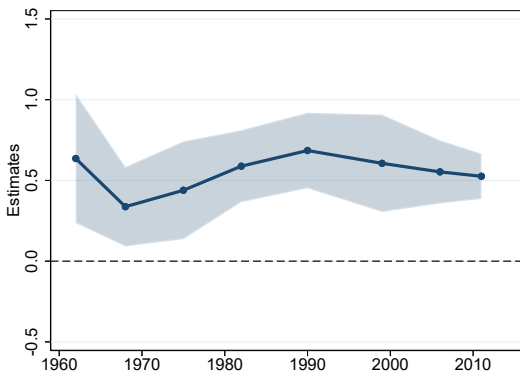
$$\text{employed}_{ihbrt} = \beta \text{death\_rate}_h + \alpha \mathbf{X}'_{1911,h} + \gamma_i \cdot \delta_r \cdot \omega_b + \varepsilon_{ihbrt}$$

- $i$  = individual,  $h$  = husband.
- $\tilde{\mathbf{X}}'_{1911,h}$ : husband pre-war controls.
- $\gamma_i$ : cohort FE.
- $\delta_r$ : dép of residence FE.
- $\omega_b$ : birth dép FE.



# Transmission Through Marriage

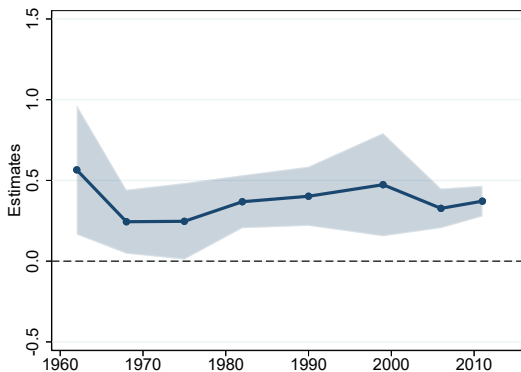
Without wives birth département FE



- Husband born in dép 20% vs 10% death rates  $\implies$  4–7 pp more likely to work.

# Transmission Through Marriage

With wives birth département FE



- Estimates ↓ by 11–41%.

# Transmission Through Marriage

- Assortative matching? A little.
- Working women = valuable trait? Probably.
- Household tasks sharing? Not the case.

# Transmission Through Marriage

- Mother-in-law to daughter channel.
- Data: extended labor surveys 2005–2012 (contain parental origins).
- Sample: second-generation internal migrant married women.
  - Second-generation internal migrant = non-migrant with parents born in another département than daughter.
- Same strategy as before but:
  - Assign key variables to mother-in-law département of birth.
  - Add own parents département of birth FE.

# Transmission Through Marriage

Dependent variable:	Employed		
	(1)	(2)	(3)
Military death rate, mother-in-law	1.28* [0.75]	1.48** [0.67]	1.64** [0.66]
Pre-war controls, mother-in-law (1911)	Yes	Yes	Yes
Birth-residence dép. FE	Yes	Yes	Yes
Mother-in-law birth region FE	No	Yes	Yes
Parents birth dép. FE	No	Yes	Yes
Individual controls	No	No	Yes
Household controls	No	No	Yes
Observations	15,095	15,095	15,095

\*\* Significant at the 5 percent level.

\* Significant at the 10 percent level.

# Transmission Through Local Social Interactions 1

- Cultural assimilation at the local (municipality) level.
- Cultural composition in military death rates:

$$\text{death\_rate}_m = \sum_o \text{sh\_res}_{o,m} \times \text{death\_rate}_o$$

- $\text{sh\_res}_{o,m}$ : share residents in municipality  $m$  born in dép.  $o$ .

# Transmission Through Local Social Interactions 1

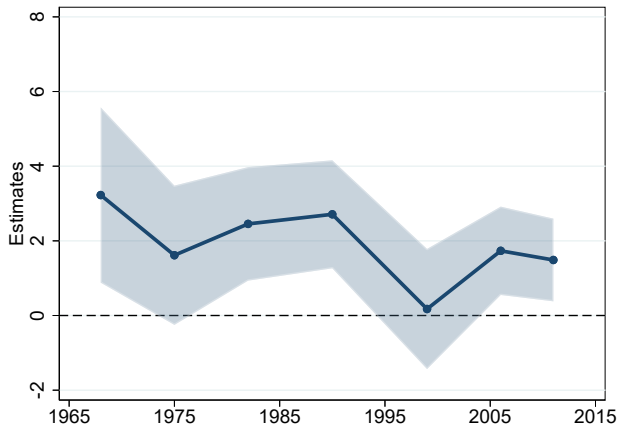
- First-generation internal migrant married women 30–49.
- Censuses 1968–2012.

- Specification:

$$\text{employed}_{ibmlt} = \beta \text{death\_rate}_m + \alpha \mathbf{X}'_{1911,m} + \gamma_i + \delta_l + \omega_b + \varepsilon_{ibmlt}$$

- $\mathbf{X}_{1911,m}$ : pre-war controls.
  - $\delta_l$ : local labor market FE (cantons).
  - $\omega_b$ : birth dép. FE.
  - Values assigned using previous census.
- Across municipality / within canton variation, same dép. of birth.

# Transmission Through Local Social Interactions 1





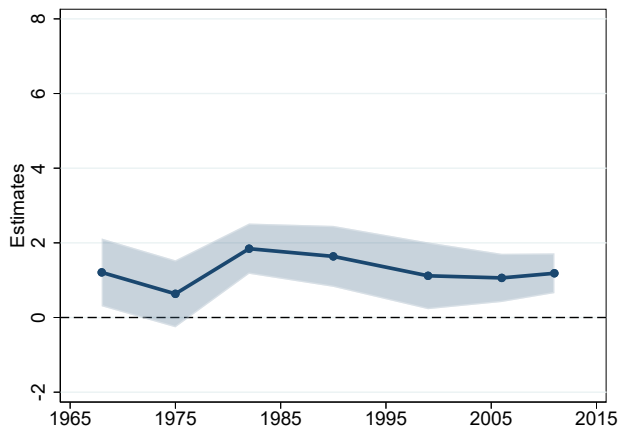
# Transmission Through Local Social Interactions 2

- Cultural preservation at the local level.
- Share of residents with the same origin:  $sh\_res_{b,m}$
- Same sample.
- Specification:

$$\begin{aligned} employed_{ibmlt} &= \beta sh\_res_{b,m} \times death\_rate_b + \alpha sh\_res_{b,m} \\ &+ \gamma_i + \delta_l + \omega_b + \varepsilon_{ibmlt} \end{aligned}$$

- Across municipality / within canton variation, same *dép.* of birth.

# Transmission Through Local Social Interactions 2



# Military Fatalities and Current Attitudes

- Data: GSS 2005.
- Sample: men and women, all ages.
- Higher values indicate disagreement with the statement.

	Statement	Men	Women
1	If a woman earns more than her partner, it is bad for their relationship	0.80 (0.28)	0.76 (0.31)
2	Women should not be able to decide how to spend the money they earned without asking their partners	0.70 (0.35)	0.76 (0.33)
3	In an economic crisis, men should keep their jobs in priority	0.72 (0.35)	0.76 (0.35)
	Cultural values index (four-points scale)	2.22 (0.65)	2.28 (0.64)
	Cultural values index (one-point scale)	0.74 (0.22)	0.76 (0.21)
	Observations	1,345	1,900

# Military Fatalities and Current Attitudes

- Specification

$$\begin{aligned} \text{values}_{ibr} &= \beta_1 \text{death\_rate}_b + \beta_2 \text{female}_i \\ &+ \beta_3 \text{death\_rate}_b \times \text{female}_i \\ &+ \alpha \mathbf{X}'_{1911,b} + \gamma_i + \delta_r + \varepsilon_{ibr} \end{aligned}$$

- Clustering: département.

# Military Fatalities and Current Attitudes

Dependent variable	Gender values index		
	(1)	(2)	(3)
Military death rate	0.95*** [0.32]	0.96*** [0.32]	1.00*** [0.33]
Female	-0.01 [0.03]	-0.01 [0.03]	-0.01 [0.03]
Individual controls	No	Yes	Yes
Parental controls	No	No	Yes
Observations	2,688	2,688	2,688

\*\*\* Significant at the 1 percent level. \*\* Significant at the 5 percent level. \* Significant at the 10 percent level.

- $\uparrow$  1 s.d. in military death rate  $\rightarrow$   $\uparrow$  0.24 s.d. in index.
- No impact on other dimensions (religion, marriage, family).

## A. First-Generation Internal Migrants

Analysis	Year	$\Delta$	Estimate		
			Raw	Standardized	Residual
Overall	1962–2012	Birth dép.	0.48***	0.046***	0.011***
Inverted	1962–2012	Res dép.	0.56**	0.044**	0.010**
Husbands	1962–2012	Husb dép.	0.37***	0.036***	0.006***
Assimilation	1968–2012	Res muni.	1.73***	0.113***	0.006***
Preservation	1968–2012	Residence municip.	1.24***	0.037***	0.009***
Attitudes	2005	Birth dép.	1.04***	0.196***	0.066***

## B. Second-Generation Internal Migrants

Analysis	Year	$\Delta$	Estimate		
			Raw	Standardized	Residual
Mother	2005–2012	Mother dép.	1.10***	0.120***	0.030***
Father	2005–2012	Father dép.	0.51	0.054	0.015
In-law	2005–2012	In-law dép.	1.48**	0.145**	0.029**