From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Economics

2024

Essays on the Economics of Newspapers in American History

Le, Quan

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp013t945v14d

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

ESSAYS ON THE ECONOMICS OF NEWSPAPERS IN AMERICAN HISTORY

QUAN C. LE

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE

BY THE DEPARTMENT OF ECONOMICS

ADVISER: KATHERINE HO

MAY 2024

© Copyright by Quan Le, 2024.

All rights reserved.

Abstract:

I present a series of essays related to the economics of newspapers in U.S. history and their impact on American society.

The first chapter establishes the importance of subscriptions to news agencies to reader demand of newspapers in the early 20th century. I show that the value of news agencies exhibits strong network effects due to the sharing of local news across member newspapers within each agency.

The second chapter studies the role of exclusive vertical contracting in competition between news agencies. I show that exclusive territory contracts between the Associated Press (AP) and member newspapers inadvertently created demand for and facilitated the growth of the AP's primary rival, United Press. I do so by estimating a model of newspaper entry and news agency network formation. I find that demand-side network effects established in the previous chapter and supply-side economies of scale form considerable natural barriers to entry for news agencies. In a counterfactual where exclusive contracts were illegal, United Press would have likely exited.

The third chapter, written jointly with Anaïs Galdin, studies the role of daily newspapers in providing information to migrants and facilitating urbanization in the U.S. between 1870 and 1940. We show that access to information about urban health conditions through local newspapers contributed to rural-urban migration during our study period.

Acknowledgments:

First and foremost, I thank Kate Ho for being my adviser from the very first day of graduate school. I am further grateful to Leah Boustan, Nick Buchholz, and Jakub Kastl for their support as my reference writers; to Pierre Bodéré, Bo Honoré, Allan Hsiao, Adam Kapor, Alessandro Lizzeri, and Steve Redding for engaging with my work; and to Melissa Dell, Leander Heldring, and Pablo Querubín for being both coauthors and mentors. I thank my fellow classmates and collaborators Eduard Boehm, Greg Dobbels, Jacob Dorn, Nick Flamang, Rachel Fung, Anaïs Galdin, Jordan Richmond, Adam Rosenberg, Karl Schulze, Carolyn Tsao, So Hye Yoon, Jie Zhou, and many more for being part of my research community; seminar participants at Princeton and elsewhere for their comments and suggestions; and Connor Fogal for excellent research assistance. I am grateful for extensive archival contributions by Valerie Komor and Francesca Pitaro at the Associated Press Corporate Archives. I thank Gustavo Bobonis, Matt Gentzkow, Marco Gonzalez-Navarro, and Jesse Shapiro for their mentorship and support before I started graduate school. Finally, I thank my family and friends for encouraging me to go further than I have ever thought I could.

Dedicated to the free press.

Contents

Abstract iii

Acknowledgments iv

1 The Role of News Agencies in Reader Demand for Historical Newspapers 1

1.1 Introduction[1] 1

1.2 Historical Setting 5

1.2.1 History of the News Agencies 6

1.2.2 News Agency Operation 10

1.3 Data 14

1.3.1 News agency membership and newspaper characteristics data 14

1.4 Descriptive Evidence on the Value of News Agencies 16

1.5 Empirical Model of Demand 18

1.6 Demand Estimation 20

1.6.1 Discussion 20

1.6.2 Estimates 22

1.7 Conclusions 27

2 News Agency Competition and Exclusive Territory Contracts 30

[1]Email: quanle@princeton.edu. This paper was part of a previously circulated manuscript titled *Network Competition and Exclusive Contracts: Evidence from News Agencies.*

2.1 Introduction[2] 30
2.2 Data 38
2.2.1 Newspaper characteristics data 38
2.2.2 Leased wires maps and news agency costs 38
2.2.3 News agency internal operations: assessments, costs, and contracts 39
2.3 Descriptive Evidence 41
2.3.1 Exclusive Contract and News Agency Growth 41
2.3.2 Economies of Scale and Assessments 42
2.4 Empirical model 48
2.4.1 Roadmap 48
2.4.2 Variable profit 49
2.4.3 Newspaper entry and news agency subscriptions 50
2.5 Estimation 53
2.5.1 Demand 53
2.5.2 Supply 54
2.6 Empirical estimates 58
2.6.1 Determinants of Variable Profit 58
2.6.2 Fixed Costs 59
2.7 Model Simulations and Counterfactuals 63
2.7.1 Framework 63
2.7.2 Baseline Simulation and Model Fit 64
2.7.3 Counterfactual Networks: No Exclusive Contracts 65
2.7.4 Counterfactual Networks: Exclusive Territories vs. Exclusive Dealing Contracts 67

[2]Email: quanle@princeton.edu. This paper was part of a previously circulated manuscript titled *Network Competition and Exclusive Contracts: Evidence from News Agencies.*

2.7.5 Counterfactual Newspaper Market Structures and Consumer Welfare 68
2.8 Conclusions 73

3 Urban Migration, Public Health Amenities, and Local Newspapers in 1870–1940 U.S. 81

3.1 Introduction[3] 81
3.2 Data and Descriptives: Sanitation Systems and Urbanization 87
3.2.1 Rural-Urban Migration Flows 87
3.2.2 Public Health Investments and Waterborne Diseases 93
3.3 Newspaper Coverage of Urban Public Health 100
3.3.1 Newspaper Text Corpus 100
3.3.2 Capturing Local News about Urban Public Health 102
3.3.3 Change in coverage with public health infrastructure 104
3.4 Model of Migration Choices and Information Acquisitions about Destination Amenities 107
3.4.1 Theoretical Framework 108
3.4.2 Empirical model 111
3.4.3 Identification assumptions 113
3.5 Empirical estimates 118
3.5.1 Newspaper Presence and Destination Amenities 118
3.5.2 News Coverage of Destination Amenities 124
3.6 Conclusion 127

[3]This paper is co-authored with *Anais Galdin*. We are grateful to Leah Boustan, Janet Currie, Kate Ho, Eduardo Morales and Steve Redding for their continual guidance and support on this project. We also thank Nick Buchholz, Jacob Dorn, Gene Grossman, Jakub Kastl, Ilyana Kuziemko, Alessandro Lizzeri, Chris Neilson and Tianyi Wang for helpful discussions and comments. This research benefited from financial support from the International Economics Section (IES) at Princeton University, whose support is gratefully acknowledged. All errors are our own.

Chapter 1

The Role of News Agencies in Reader Demand for Historical Newspapers

1.1 Introduction[1]

Newspapers in American history underwent a transformative period in the early 20th century when they gained independence from political machines, shifted their focus to informative news provision, and offered less partisan content. Previous work has established a key factor in this shift to be advancements the technology of newspaper production such as cheaper newsprint, which allowed for mass circulation and in turn reduced newspapers' dependency on political parties to cover fixed costs (Gentzkow, Glaeser, and Goldin 2004) and increased revenue from commercial advertisement (Petrova 2011). Significant questions remain about whether there were changes in the value of newspaper content to readers and the production of this content that could have contributed to the increase of newspaper circulation in this period.

[1]Email: quanle@princeton.edu. This paper was part of a previously circulated manuscript titled *Network Competition and Exclusive Contracts: Evidence from News Agencies*.

In this paper, I argue that the development of the telegraphic news business and news agencies was a crucial component of increased newspaper circulation in the early 20th century United States. News agencies were networks of newspapers that formed in order to coordinate the collection and dissemination of the news via the electric telegraph. While telegraphy and the associated commercial telegraph network have been available since the mid-19th century (Wang 2023), the growth of news agencies operating a fully connected leased telegraph network mostly emerged during the early 20th-century. I contend that the expansion of the Associated Press and its primary rival, United Press, during this time was an important factor driving reader demand for newspapers in the US.

In order to motivate the formal analysis, I first present a historical account of these news agencies, detailing their strategy, technical operations, and contracting practices. These components are crucial for understanding the demand for news agencies in this paper and studying the market structure of news agencies in future work. In particular, the historical account shows that the value of news agencies derives from the news they provide to newspapers as opposed to purely as a brand or fact-checking role. This service of news provision can take many forms. First, agencies coordinate the collection of foreign and national news by establishing bureaus and allocating reporter resources, thus eliminating the need for each newspaper to dispatch their own reporters to newsworthy locations. Second, agencies act as a central collection point for local news generated by newspaper members. Upon receiving incoming news, agencies setup telegraphic networks to disseminate this news to individual newspapers. I hypothesize that these roles give rise to *network effects* in the readers' value of news agencies: a larger network can more effectively pool together resources and coordinate the collection of news as well as acquire more local news for dissemination. I further demonstrate that the news agency duopoly in the U.S. arose with an understanding of these network effects. Furthermore, these

agencies employed news sharing technology, strategy, and contracting practices that specifically cultivated these network effects.

Formally, I develop a model of newspaper demand that explicitly takes into account a newspaper's subscription to news agencies and these network effects. To estimate this model, I collect novel data on news agency subscriptions and merge them with a panel of daily newspaper circulation and characteristics from Gentzkow, Shapiro, and Sinkinson (2014). The network value for each news agency is identified by the extent to which its subscribers' local market shares vary with the agency's membership size. I also allow for potential differentiation between news agencies in two ways. First, I allow readers of different observable demographic groups to potentially have heterogeneous preferences for news agencies. Second, readers can prefer newspapers that bundle content from more than one agency. Readers purchase newspapers according to a discrete choice model, where a newspaper's characteristics include the value of its news agency subscriptions.

Demand estimation in newspaper faces an important challenge of subscription prices being endogenous and potentially correlated with the paper's unobserved quality. I follow Gentzkow, Shapiro, and Sinkinson (2014) in calibrating the price coefficients using first order conditions implied by monopolist price setting and a set of observed marginal cost and advertising revenue for newspapers. Another key identification concern in demand estimation is that newspapers that select into different news agencies may have different qualities unrelated to the value of the news agency subscription (such as local reputation or local reporting). I avoid this problem by leveraging the panel structure of the data and directly estimating the component of local quality of a newspaper that does not vary over time. Network effects for news agencies are identified from variation in newspaper market share within newspapers across time rather than across newspapers.

My demand estimates reveal the following patterns regarding network effects. First, demand-side network effects that result from a news agency's membership size are substantial. Readers valued an AP subscription more than 150% as much as a UP subscription in 1908, but this gap decreased over time as UP gained scale. By the end of the study period, the value of AP or UP telegraphic news was approximately one fourth of the median price of newspapers. Second, bundling is highly rewarded. I find that an AP paper bundling UP news would earn 0.56 of the additional network value. These estimates represent a barrier to entry for UP as the AP's larger network was preferred by all readers, even if AP papers have an incentive to bundle UP news. These results hold whether I measure network effects via the number of newspapers in the organization or via the population coverage of the organization.

I further find that UP did not enjoy a significant advantage over the AP among any demographic group. This clashes with public statements made by UP proprietors at the time, who wished for UP to be a progressive organization. Demand estimates suggest that newspapers themselves, on the other hand, are highly horizontally differentiated. These facts are consistent with private communication from the agencies regarding their strategies. In particular, the AP views themselves as producer of hard and objective news. These findings are potentially consistent with the technology of the agencies, a hypothesis that I explore in another chapter of this dissertation.

Related Literature

My paper studies the role of news agencies in reader demand for newspapers. Regarding the news content of newspaper, this literature has focused on production-side quality measure such as the number of journalists / reporters employed by the newspaper (Fan 2013) across categories of news production. There's also a literature studying the role of political information and political slant and bias in newspaper consumption, including Gentzkow and Shapiro (2010, 2008), Chiang and Knight

(2011), and Xiang and Sarvary (2007). I contribute to this literature by highlighting the importance of news agency subscriptions as well as the role of network effects in the value of news agencies to readers.

There is a small literature in economics that studies how newspapers use wire news. Cage et al. (2019) studies the importance of how online French outlets reprint copy from Agence France-Presse and Reuters in 2013, Chiou and Tucker (2017) studies the value of AP news to news aggregators such as Google News, and Djourelova (2023) studies how slanted language used in AP copy can influence public opinion. The paper closest to this is Angelucci et al. (2023) who studies how newspapers bundle news agency content with locally produced content between 1945 and 1960. They find that a decrease in advertising revenue due to the entry of television leads to newspapers cutting back their subscriptions to news agencies.

Roadmap

Section 2 describes the history of news agencies. Section 3 describes the data. Section 4 presents the descriptive evidence that motivates the structural analysis. Section 5 describes the formal empirical model. Sections 6 describes the estimation procedure and estimates. Section 7 concludes.

1.2 Historical Setting

I focus on the history and technology of news agencies as they relate to the value of news agencies to newspapers and contracting practices specifically rather than that of telegraphic news in general. A complete history of news agencies is necessarily outside the scope of this paper. I direct interested reader to the historical appendix for references that relate to the aspects of news agencies that referred to in this paper. Readers interested in the early technology of the electric telegraph and its

effect on news content will find an excellent summary in Wang (2023). John (2015) provides a history of exclusion in American telecommunication networks in general. Silberstein-Loeb (2014, 2012) studies the legal history of the news agencies' exclusive contracts.

1.2.1 History of the News Agencies

Historically, the presence of several competing news agencies is a distinctive characteristic of the United States (UNESCO 1953). The vast majority of countries in the world have a single national news agency, which is often either state-owned or partially subsidized by the state. Among news agencies in the US, the Associated Press (AP) has a unique structure as a non-profit cooperative of newspapers. It had two notable competitors: United Press (UP) was founded by E. W. Scripps which aimed to be a commercial entity challenging the AP, unlike the International News Service (INS) which existed primarily to advance William Randolph Hearst's political views.

The Associated Press (AP) In 1856, the New York Associated Press (NYAP) was established as the earliest precursor to the current Associated Press (AP). Its non-profit cooperative structure can be traced back to its early days when newspapers pooled resources to hire boats in the New York harbor to collect news from arriving vessels (Rosewater 1930). A personal relationship with Abraham Lincoln allowed it exclusive access to the Union's war bulletins which contributed to its rise to prominence. Significant demand for news and additional costs due to coverage of the Civil War and the ensuing Spanish American War led to considerable upheavals in the news agency business. By the end of the 19th century, the Associated Press, now having re-established in Illinois, emerged as the dominant American news agency (Blanchard 1987).

In reaction to the Illinois Supreme Court decision Inter-Ocean Publishing Company. v. Associated Press (1900) which found the AP's ongoing exclusive contracts in restraint of trade as long as it remained an Illinois corporation, the AP decided to "consequently [...] discontinue business and wind up its affairs" in the state of Illinois and reincorporated into a new corporation in New York with an identical name and member list (Associated Press 1901). In 1900, the AP's legal counsels found that the only way to maintain exclusive contracts was to restart the AP as a non-profit under New York's "Membership Corporation Law." The state passed this law in 1895 with the intention to provide for the incorporation of medical and college alumni societies (State of New York 1918). I regard 1900 as the entry point of the modern iteration of the AP in this paper and thus the beginning of the period of study. The AP continues to be the predominant American news agency today. Furthermore, they still maintain their non-profit membership corporation structure.

United Press (UP) The year 1907 saw the inception of United Press (UP) from a merger of three newspaper chains Scripps-McRae, Scripps-Blades, and Publishers' Press Association (Benét 1933; Morris 1957).[2] The family of E. W. Scripps and its associates owned these newspaper chains.

From the outset, UP's primary intent was to serve newspapers who could not get AP service. In a 1912 private letter to UP's General Manager Roy Howard, Scripps claimed that he "wanted to make it possible for any other man to found a newspaper in any city of the Union." This seemingly altruistic desire was also rooted in a practical problem: several of Scripps' newspapers were not welcome in the AP due to existing exclusive agreements.

[2]UP's full name initially was "United Press Association." The organization of interest bears no relation to another short-lived news agency also named "United Press" that went out of business in 1898. UP's morning operation was initially called "United News" until it was brought under the same umbrella.

Scripps matched his verbal commitment to the UP by a financial one. In his letter to Howard, Scripps revealed that he had initially covered "the whole expense of paying for the cables and gathering news from all over the country" and bailed out the Publishers' Press Association's debts with $150,000 (roughly $285k in 1928 or $5m today). A series of letters in 1912 between UP's proprietors deliberating on the organization's valuation reveals that UP is only breaking even because of a "large subsidy" from Scripps.

These letters motivate two important factors in my analysis. First, they reveal that UP needed a large subsidy in order to cover its costs in 1912. This is in line with evidence that UP's overall cost was very high relative to its membership. Drawing from observed expenditures submitted as evidence to AP v. US (1945), I find that UP's expense per member in 1912 was more than twice as much as the AP, while their assessments were roughly similar. Second, it reveals that UP's proprietors had a scrap value: internal valuations of UP stock in 1912 ranged from $75,000 to $600,000 (this upper bound is $1m in 1928 or $19 million today). These two factors lead me to conclude that any potential revenue shortfall the UP might have faced in the counterfactual would need to be offset by additional support from Scripps, and that a revenue shortfall beyond the upper bound of $1m would have led to UP's exit.

The Scripps paper chain was known for his progressive ideals. However, there is scant historical evidence that the United Press news were colored to reflect Scripps' ideals. Private conversations as well as public profiles at the time suggests that the United Press's primary objective was profit-maximization: among nearly a hundred UP promotional documents, I fail to identify a single instance that mentions the organization's political slant. Karl Bickel [who] stated that "the press association business is purely a business" and that UP's purpose is "to make a profit." (Bickel, 1927 "The First Twenty Years", Editor and Publisher, April 30 1927, p11). A 1929 UP internal emphasizes that "the United Press is an impartial news service. It never takes

sides in any controversy" and that it "has but one policy, and that is to present all sides and carry the news." Despite this objective, it seems like UP never significantly turned a profit.

UP eventually gained scale to compete with the AP and was treated as one of the three domestic news agencies by AP v. US (1945). The post-war entry of television and the general decline of newspaper advertising revenues saw UP merged with INS to form United Press International (UPI) in 1958 (Angelucci, Cage, and Sinkinson 2023). Despite this merger, UPI struggled financially and relied heavily on a cross-subsidy arrangement with the far more lucrative United Features Syndicate which owned the comic strip "Peanuts." The decline of UPI in the late 1970s echoed the broader downturn in the American newspaper industry. Ultimately, by 1982, its value had diminished so much that it was sold for a mere dollar. (Cohen and Gordon, 1900).

International News Service (INS) The International News Service (INS) was established in 1910 as an arm of the Hearst newspaper conglomerate. Originally, its primary role was to manage the Sunday comics business, but it soon expanded to serve Hearst-owned newspapers as a full-fledged news agency (Nasaw 2013).

However, rather than being a direct commercial rival to the AP and UP, INS mostly functioned as a medium for Hearst to broadcast his viewpoints (Rosewater 1930). For example, the firm's distinct pro-German stance led to it being banned from Allied cables during the first World War. The substantial deficits INS ran, notably $400,000 in 1919 (or approximately $7.5m today; Schwarzlose 1989), further underscores its non-commercial purpose. Given this history, I have chosen to treat subscriptions to the INS and its growth as independent of the AP and UP which are distinctly commercial entities.

1.2.2 News Agency Operation

News Collection Technology. The collection of local news is the duty of local member newspapers, whose contracts with the AP require that newspapers "furnish for the use of the [AP] the news of [their] territory." In effect, this territory usually spans a 30-mile radius, with exceptions in densely populated regions with multiple publications such as New York City.

One potential concern is that newspapers could free-ride on telegraphic news and scale back their local news production upon obtaining news agency content. Alternatively, newspapers could increase their local news production in order to service the agencies or to complement newly available national news (Angelucci, Cage, and Sinkinson 2023). In this paper, I model the variation in local news collection over time as exogenous to news agency subscription. This is motivated by the fact that telegraphic news only constituted around a quarter of a newspaper's editorial content, and so newspapers still had significant incentives to compete on the quality of their local news (AP v. US 1945).

Agency staff and string correspondents were employed to collect regional and national news. In 1931, the AP and UP had 91 and 51 domestic bureaus respectively. Bureaus also acted as hubs for rewriting national news for local transmission. Given these practices, I model the nearest bureaus as the source of the news for each newspaper market.[3]

News Sharing Technology. Starting in 1900, the vast majority of the AP members received the news via a system of leased wires. Leased wires were dedicated telegraph circuits that connect newspapers to bureau news sources and provided by telegraph companies such as Western Union and Postal Telegraph.

[3]Foreign news are either collected from coastal bureaus through international cables via contracts with international counterpart (e.g. Reuters in the United Kingdom, Havas or later Agence France-Presse, or the Telegraph Agency of the Soviet Union, TASS) or by the agency's own bureaus in foreign countries.

Leased wires were important insofar as they dictate the outgoing news but not incoming news.[4] While the maintenance of the physical lines was done by telegraph companies, the agencies provided their own operators in most instances. In order to obtain the news from an agency, a newspaper needed to integrate the telegraph circuit physically into their editorial headquarters and maintain an operator there to receive the transmitted news. A paper that subscribed to multiple news agencies separate rooms and operators — this is done in order to prevent breaking news from one organization leaking to the other. There could also be additional related editorial costs. These arrangement means that a newspaper's cost upon subscribing to a news agency may depend include additional costs other than assessments.

The sharing of leased wires is the crucial source of economy of scale for news agencies. The primary wire system (as opposed to special events or financial information supplemental wires) consisted of two systems: a trunk system that connected major bureaus and regional circuits. Regional circuits were often contained within state boundaries due to newspapers in the same state having similar interests. The fact that the news agencies did not own the wire but merely leased them via long-term contracts motivates a key modeling decision: the cost of leased wires are an ongoing fixed cost, not a one-time entry cost.[5]

Assessments. The AP charged its member a weekly assessment that depends on three components (Associated Press 1984). First, each member covered its own local costs. Second, members on the same leased wire circuit proportionately shared its

[4]Incoming news are delivered via telephone, commercial telegraph, or carbon copy and constitute a separate expense category.

[5]Not all papers are receive the news via leased wires. In the year 1926, one quarter of the AP membership received the news via a long distance telephone circuit. I do not have reliable data on other years on which AP or UP papers could be telephone clients rather than leased wire clients. Because of this lack of data, I group all papers in the same subscription class.

cost. Third, members paid a share of general expenses at the organization level, which is proportional to the population within 10 miles per the decennial census.[6]

I estimate the assessment formula using the population of the newspaper's headquarter city, the distance to the nearest bureau, the number of newspapers that share the bureau and state circuits, and the region as defined by the nearest news agency headquarter (which closely relates to a Census region). Since the AP formula was first adopted in 1902 and no change was made until 1951 (except for level changes in the general cost due to the Great Depression and wartime budget increases), I maintain this approximation through out the paper (Associated Press 1984).

I do not have direct evidence of UP's assessment formula. An internal letter to UP's board by General Manager Roy Howard states that UP was going to charge a newspaper entrant at cost in a manner similar to the AP before eventually imposing a 15% profit when the paper breaks even. Crucially, Howard did not discuss price discrimination. In Section 4.2.2, I show that the AP formula fits observed UP assessments well. I assume that UP uses the same formula as the AP when calculating UP revenue.

Exclusive Territory Contracts. The AP had exclusive territory contracts since the beginning. These contracts were operationalized via so-called "protest rights," where the rights holder could protest against the admission of a potential competitor.[7] In 1900, protest rights were only extended to 260 papers and varied by their geographic radius (most often 10 to 30 miles), applicability across times of publication (morning vs. evening), and whether there were specific carve-outs for other nearby urban markets. The exclusion radius shrunk quickly geographically to a uniform 10-mile

[6]The population proportional method for calculating this share is rather complicated. First, the AP treats each geographic market as a "city" definition in the Census. To account for differential population across neighboring cities, AP documents allude to a "circle overlap" method that assigns the market a fraction of the neighboring cities' population. For details, please see the appendix.

[7]The protest right could be overcome by an 80% vote by the membership. There were six ever instances of this happening (*AP v. US* 1945).

radius for most members starting in 1915 (Silberstein-Loeb 2012). An internal coup in 1928 extended these rights to all members who were in the organization for more than 5 years and limited them to a 10-mile radius and within the same time of publication.

I make the simplifying assumption that AP exclusivity is automatic to membership and thus exogenous to local market conditions. This is consistent with the empirical probabilities: less than 1% of AP markets had more than one AP paper at any given point in time. Furthermore, an AP board member wrote in 1927 to then AP's general manager Kent Cooper that "It has always been the policy of the Board of Directors whether or not members had a legal right of protest to protect them in their respective fields against the election of a competing member."

I make the further simplying assumption that exclusivity is limited to the city's geographical limits and to the same time of publication. This simplifying assumption keeps the model tractable — otherwise markets have overlapping boundaries and thus entry decisions are interlinked across markets. This is also consistent with a 1928 letter in which AP board members claim that they have rarely enforced exclusion outside of the immediate city boundary. For the demand system, I maintain the same market definition as that used by the AP for the purpose of their contracts, which is a city and time of publication combination.

UP exclusion could have been endogenous. Substantial historical evidence suggests that despite Scripps' proclaimed ideals for an open organization, UP enforced exclusive contracts to all clients (Silberstein-Loeb 2012). In my data, fewer than 2% of UP markets had more than one UP paper at any given point in time.[8] Because of this, I also treat UP's exclusive contracts as exogenous to local market conditions.

[8]I observe several instances of UP signing "asset value" contracts, with a clause akin to liquidated damages clause that forces a second UP newspaper in the same market to recompense the first UP newspaper for their loss of exclusivity. These clauses are a way to implement a barrier to entry without explicitly excluding future firms (Whinston 2008, Aghion and Bolton 1987). However, Asker and Bar-Isaac (2014) shows that exclusionary vertical practices can arise in equilibrium due to a tacit understanding between firms even without an explicit contract.

1.3 Data

1.3.1 News agency membership and newspaper characteristics data

My main dataset on news agency memberships is derived from the Transcript of Record for Associated Press v. United States (1945). In connection to the case, senior executives from AP, UP, and INS were deposed by the government. They were required to submit the names, locations, and time of publication of all their English language daily newspaper clients in the U.S. as of 1941, in addition to the respective dates they first joined.

The data provides information on newspapers that were members of at least one agency in 1941 and their initial subscription date. However, it does not include newspapers who were no longer active in 1941. For newspapers that went through mergers, had short subscription interruptions, or experienced a name or ownership change, their original subscription date is recorded. I trace through these events to categorize different versions of the same newspaper under a single "franchise" designation.[9] A “franchise” subscriber is the relevant unit of subscription from the perspective of the news agency growth: membership is automatically transferred through name and ownership changes and so resulting network effects and economies of scale are preserved. Figure 1.1 shows the membership data.

[9]When two newspapers merge, we see the subscription date as before the merge date. For these, I manually research auxiliary databases of newspaper characteristics and newspaper scan archives to ascertain their subscription status.

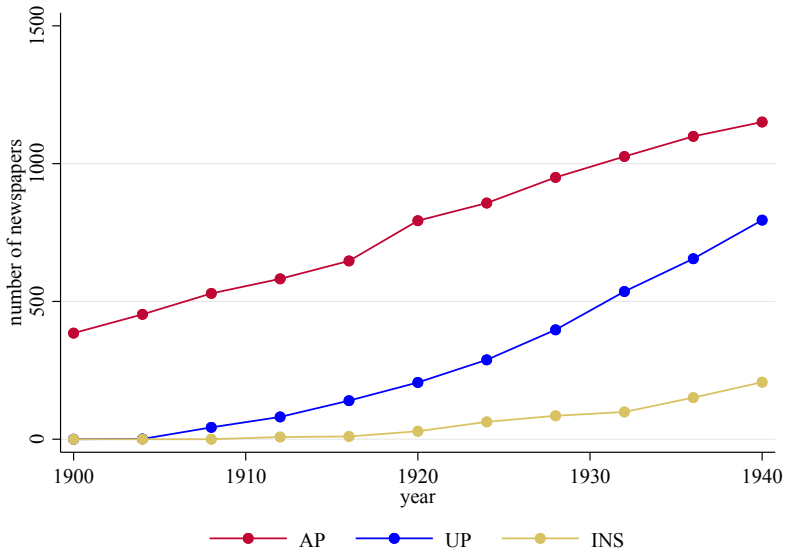


Figure 1.1: News Agency Membership

Note: This figure plots the number of newspapers in the data that subscribes to each news agency over time. The red, blue, and yellow lines denotes membership count for the Associated Press, United Press, and International News Service respectively. The newspaper panel dataset is from Gentzkow, Shapiro, and Sinkinson (2014) and is available for election years. I merge the panel with initial news agency subscription date as reported in *Associated Press v. United States* (1945) and track a newspaper "franchise" across name changes, ownership changes, and mergers.

I merge this data with the United States Newspaper Panel from Gentzkow, Shapiro, and Sinkinson (2014, GSS), which contains newspaper characteristics for election years in the US. This constitutes the bulk of the data that goes into our reader demand system, which includes circulation, subscription prices, and political affiliations. In order to construct market share for each newspaper's home market, I multiply circulation by 75%. This figure comes from GSS (2014) which documents that approximately 75% of an average newspaper's circulation is within its home market. Since subscription prices are unavailable for 1932, I skip this year in the estimation of reader demand.

I incorporate additional data from several other sources. First, I digitize Annual Linage Supplement provided by Editors and Publishers for the years 1932, 1936, and 1940. This provides me information on advertising quantities for a large subset of newspapers. Second, I use the Inland Press Cost and Revenue Study for 1928, also

provided by GSS (2014). This was a non-random sample of newspapers for which I could observe actual marginal costs and advertising revenue. In this chapter, the primary role of the Inland Press data is following GSS (2014) to calibrate the mean marginal costs of newspapers in 1928 and estimate the price coefficient.

1.4 Descriptive Evidence on the Value of News Agencies

The value of a newspaper's subscription to news agencies come from access to foreign, national, regional, and local news. There is a substantial body of work that shows the contribution of telegraph news to newspaper editorial content: Wang (2023) shows that the electric telegraph speeds the news from Washington DC; Silcock and Dell (2023) shows that a substantial fraction of newspaper texts are reprinted, which increases between 1920 and 1945 (with a sharp spike at the start of the second World War). Future versions of the paper will incorporate direct evidence of the value of news agencies in newspaper text using the *American Stories* dataset (Dell et al. 2023).

To demonstrate that news agencies provide additional content for the newspaper, I regress a newspaper's observed number of physical pages on their news agency subscriptions and newspaper fixed-effects:

$$\text{NumPages}_{jmt} = \sum_{w \in \{\text{AP,UP,INS}\}} \kappa_w w_{jmt} + \mathbf{1}\{j \text{ bundles multiple services}\}_{jmt} + \text{Newspaper-FE}_j + \epsilon^{Pages}_{jmt},$$

where the number of pages is either the minimum or maximum number of pages in a typical issue for the paper that year as reported in Editors and Publishers. The right hand side includes w which are indicators for the three news agencies, an indicator for whether the newspaper bundles multiple wire services, and newspaper fixed-effects. ϵ^{Pages}_{jmt} denotes a measurement error.

Table 1.1 presents results. Column 1 presents results for the maximum number of pages observed within the year, and Column 2 presents that for the minimum number of pages. For both specifications, the number of pages that a given newspaper prints increase significantly with news agency subscription. Since the regression controls for newspaper fixed-effects, this increase is due to only within-paper variation.

	(1) Max # of pages	(2) Min # of pages
AP	1.146	0.554
	(0.299)	(0.134)
UP	1.511	0.396
	(0.467)	(0.201)
INS	-0.471	1.577
	(1.651)	(0.738)
Bundling multiple services	2.345	0.531
	(0.972)	(0.435)
Constant	6.077	5.873
	(0.277)	(0.124)
R^2	0.821	0.777
Number of market-years	6,758	

Table 1.1: News Agency Subscription Correlates with Newspaper Length
Note: This table presents the relationship between news agency subscriptions and number of pages in a newspaper. I present two measures in the panel dataset, extracted from Editors and Publishers' reporting a newspaper's the range of number of pages in a typical issue that year. Data on the newspaper length is only available for election years between 1900 and 1912. Both specifications control for newspaper franchise fixed effects.

Subscribing to news agencies significantly increases the page count in newspaper issues. Specifically, an AP subscription leads to an 18% increase relative to the average maximum page count, while UP subscriptions result in a 25% increase (the difference in effect sizes between AP and UP is not statistically significant). The magnitude of this increase in page count is larger for the maximum number of pages compared to the minimum. Furthermore, when newspapers bundle services from multiple agencies,

there's a substantial increase in only the maximum number of pages. This is likely due to to the additional of content available during special events that both agencies cover.

1.5 Empirical Model of Demand

I start by presenting some basic notation. Consider all possible markets $m \in \{1, \ldots, M\}$ spanning finitely many discrete time periods $t \in \{1, \ldots, T\}$. A market is defined as a geography and time-of-publication combination. I refer to the set of incumbents for each market-period combination as $\mathscr{J}_{mt}$ and cap the total number of firms possible in each period at J^{max}. For notational ease, I refer to the vector of market states being considered by the mover in each period as Ω_{mt}: this includes anything relevant to payoffs and known by the mover.

I will refer to a newspaper j's bundle of subscriptions to news agencies w as W, where W is any subset of $\{\text{AP}, \text{UP}, \text{INS}\}$ (including $\emptyset$).[10]

Reader demand My specification for reader demand follows a nested logit specification (Berry 1994). Each reader i in market m at time t chooses to either to buy a newspaper j or the outside option and maximizes the following utility equation across options:

$$u_{ijmt} = -\alpha p_{jmt} + g_{mt}\left(W_{jmt}\right) + X_{mt}\gamma + \overline{\xi}_{jm} + \xi_{jmt} + \left(\zeta_{igmt} + (1-\sigma)\,\epsilon_{ijmt}\right) \quad (1.1)$$

where α is a coefficient on the newspaper's price p_{jmt}, $g_{mt}\left(W_{jmt}\right)$ is the reader's valuation over newspaper j's bundle of news agency subscriptions W_{jmt}, X_{mt} is a vector of market-level shifters of mean utility and γ is a vector of associated coefficients, $\overline{\xi}_{jm}$ is an unobserved vertical quality component of the newspaper that

[10] The letter w is chosen for *wire* services.

persists over time, ξ_{jmt} is an idiosyncratic vertical quality component of the newspaper that gets drawn each period, ζ_{igmt} is an idiosyncratic preference for a product "group," ϵ_{ijmt} is a type-I extreme value error, and σ dictates the correlation of idiosyncratic utilities among the products in the same group. There are two mutually exclusive groups of products: the outside good (not buying a newspaper) or the inside good (buying a newspaper).

The specification for a reader's valuation over a paper's subscription bundles g is:

$$g_{mt}\left(W_{jmt}\right) = \overbrace{\beta_1\left(1-\gamma_{|W|>1}\right)\sum_{w\in W_{jmt}}\log\left(1+|J^w_{mt}|\right)}^{\text{network effects}} + \underbrace{\beta_2 \mathbf{1}\left\{W_{jmt}=\emptyset\right\}}_{\text{preference for no-wire papers}} + \underbrace{\sum_{w,\emptyset}\mathbf{1}\left\{w\in W_{jmt}\right\}\left(X_{mt}\gamma_w\right)}_{\text{demographic-specific preferences}}. \tag{1.2}$$

The first component captures the total network strengths of a newspaper's subscription bundle. This strength is a function of size of the network, J^w_{mt}. β_1 is a scalar that captures overall preference for network strengths. The coefficient $\gamma_{|W|>1}$ captures a "penalty" that is relevant if the newspaper bundling multiple news agencies' content and is analogous to measuring the potential overlap in network values: $\gamma = 0$ implies that there is no penalty and $\gamma = 1/2$ implies that bundling is worthless to the reader. In the baseline specification, I allow an agency w's network strength to be simply the natural logarithm of its membership size $|J^w_{mt}|$ measured by the number of newspapers active in the network. I also report a specification where a network's strength is the covered population.

The second component captures any potential preference for no-wire papers. This allows for potential quality investments that papers without telegraphic wires make in the absence of agency telegraphic news.[11]

[11] In this period, newspapers without a telegraphic news subscription often rewrite news from copies of previously published newspapers to fill space. News only has a "quasi-property rights" for a short duration and published facts cannot be copyrighted.

The third component captures any demographic-specific preference for the different news agencies that are independent of their network effects, such as the agency's brand, style, or slant. X consists of the same market demographic variables as in the main utility equation (1.1). The reported specification has a vector of observable demographics in the Census, including the population share of a town that is Urban, White, works in Manufacturing, or is Foreign-born.

This utility specification yields for a mean utility $\overline{u}_{jmt}$ for each newspaper j.[12] The share of potential readers that buys newspaper j after integrating over idiosyncratic preference shocks, s_{jmt}, follows the familiar nested logit probabilities.[13] The demand curve facing newspaper j is then $q_{jmt} = Pop_{mt}s_{jmt}$.

1.6 Demand Estimation

1.6.1 Discussion

I estimate the demand model using a two-steps routine. In the first step, I guess the price coefficient α. In the second step, I estimate equation (1.1) using two-staged least squares. An instrument is needed for the market share of a newspaper conditional on being an inside good. For each guess of the price coefficient, this routine gives me the demand for each newspaper as a function of prices and product characteristics s_{jmt}. I then use the first order condition implied by the monopolist's price setting problem to back out implied marginal costs MC. I follow the intuition of Gentzkow, Shapiro, and Sinkinson (2014) and iterate on α until the mean marginal cost equates that observed in the Inland Press Cost and Revenue Survey for monopolist papers

[12] $\overline{u}_{jmt} = -\alpha p_{jmt} + g_{mt}\left(W_{jmt}\right) + X_{mt}\gamma + \overline{\xi}_{jm} + \xi_{jmt}$.

[13] The following derivation uses the notation from Miller and Weinberg (2017). Let $\overline{s}_{j|in} = \frac{\exp\left(\frac{\overline{u}_{jmt}}{1-\sigma}\right)}{D_g}$ where $D_{in} = \sum_k \left(\frac{\overline{u}_{kmt}}{1-\sigma}\right)$. The unconditional choice probability for the inside good is $\overline{s}_{in} = \frac{D_{in}^{1-\sigma}}{1+\sum D_{in}^{1-\sigma}}$, and the market share of good j is $s_j = \overline{s}_{j|in}\overline{s}_{in}$.

in towns between the 25th-75th percentile of the population distribution in the year 1928.

Identification of the reader demand equation (1.1) relies on variation in newspaper market shares that correlates with changes in its news agency's membership. An important concern is that papers of different quality could have selected into different news agencies over time. For example, if UP did not gain high-quality papers until later in life, we would have mis-attributed UP's paper strength to its network rather than its local quality. I overcome this selection problem by estimating equation (1.1) with explicit newspaper fixed effects $\overline{\xi}_{jm}$. Since I observe newspapers over time, I interpret this as the component of unobserved local quality of the newspaper that is constant over time.

Another concern is that it could be that newspapers who were and that other features of the newspaper could have changed over time in a way that matched UP's development relative to AP. For example, if papers that join news agencies are more likely to invest in their local quality, we would wrongly attribute the resulting increase in market share to news agency network effects rather than in this local investment. Resolving this concern would require proxying for the extent of national and local news investment using textual data or data on reporters.

Demand estimation requires an instrument for the within-nest market share. The standard instrument in the literature is a shifter of the number of available products inside the nest (Miller and Weinberg 2017). I use a vector of indicators for whether one of the following events happened in a market in that particular period: the entry or exit of a newspaper, the merger of two newspapers, and whether a competitor picked up a news agency subscription.

1.6.2 Estimates

Table (1.2) reports estimates of the reader demand model. I report two different specifications for reader demand. The first specifies a news agency's network strength by the number of newspapers that are its clients. The second specifies a news agency's network strength by the population covered by the news agency. Figure 1.2 demonstrates coefficients for the first specification in units of annual subscription price.

I find that the network strength strongly contributes to a news agency's value to a newspaper reader. The value of subscribing to the AP at the beginning would have been one and a half larger than subscribing to the UP. This mostly converged toward the end of the time frame as UP gained scale. The magnitude is substantial: the mean utility of wire services contributing to a newspaper is about \$1.5 for the reader. Given that median subscription price is about \$7.50, telegraphic news constitute about 20% of the value of a newspaper by the end of the sample. This is consistent with reports in *AP v. US* (1945) noting that about 20% of a newspaper's content is telegraphic news.

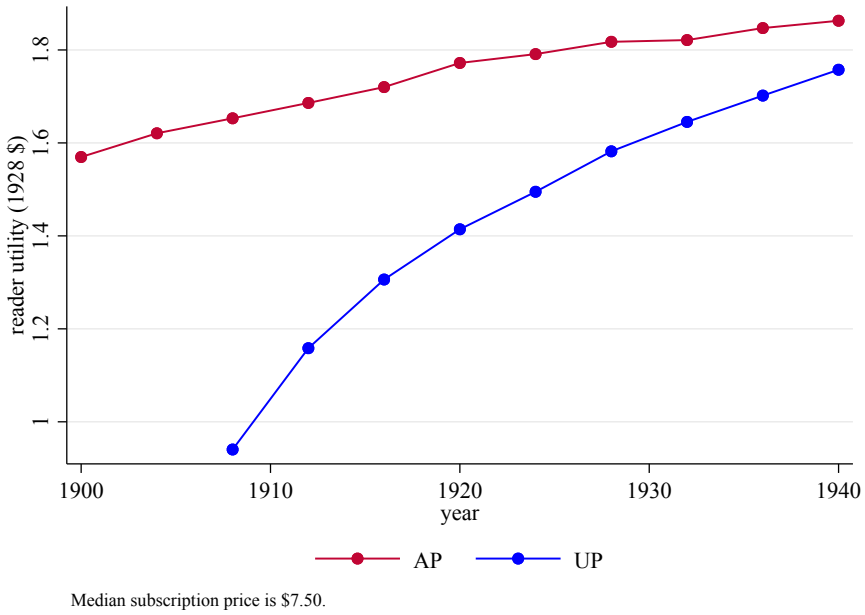
I also find that bundles of news agency exhibit substantial overlap: the utility resulting from subscribing to more than one agency is about 0.7 or 0.8 the sum of utilities from each separate subscription. This significantly lessens the incentive for newspapers to bundle both news agencies' content.

		(1) Membership Size	(2) Population Coverage
Network Effects	Strength (β_1)	0.029	0.028
		(0.012)	(0.011)
	Overlap ($\gamma_{\|W\|-1}$)	0.219	0.292
		(0.139)	(0.104)
	No-Wire (β_2)	-0.024	0.031
		(0.076)	(0.094)
Demographics ($\gamma_\emptyset$)	White	0.513	0.511
		(0.206)	(0.206)
	Manufacturing	-0.278	-0.287
		(0.382)	(0.382)
	Foreign-Born	-0.180	-0.184
		(0.250)	(0.250)
	Urban	0.717	0.718
		(0.144)	(0.144)
Demographics × AP (γ_{AP})	White	0.052	0.053
		(0.015)	(0.015)
	Manufacturing	0.001	0.001
		(0.019)	(0.018)
	Foreign-Born	-0.025	-0.025
		(0.022)	(0.220)
	Urban	-0.300	-0.302
		(0.025)	(0.025)
Demographics × UP (γ_{UP})	White	-0.025	-0.025
		(0.022)	(0.023)
	Manufacturing	0.020	0.021
		(0.382)	(0.022)
	Foreign-Born	-0.001	-0.001
		(0.250)	(0.023)
	Urban	-0.099	-0.100
		(0.144)	(0.025)
	Nesting coefficient (σ)	0.072 (0.07)	0.071 (0.07)
	Price Coefficient (α)	0.10	0.12
	R^2	0.857	0.857
	Number of market-years	14,121	14,121

Table 1.2: Estimates for the Reader Demand Model

Note: This table presents estimates for the demand model described in Section 5.2. All specifications control for city-population-decile fixed effects, franchise fixed effects, and Year × Time-of-Publication fixed effects.

Panel A: Network Effects



Panel B: Demographic-Specific Preferences

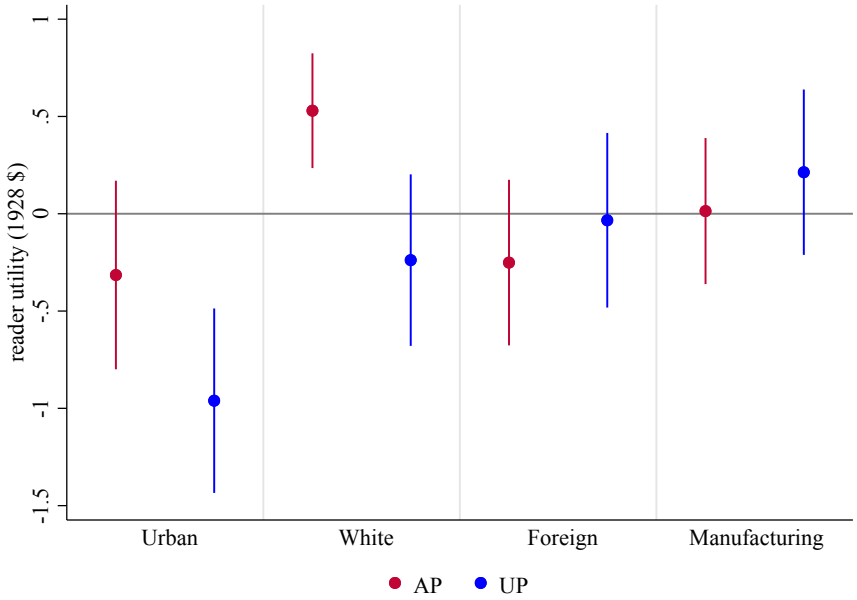


Figure 1.2: Graphical Interpretation of Demand Estimates

Note: This figure depicts a graphical representation of demand estimates in Column 1 of Table 1.2. Panel A depicts the magnitude of the coefficient on network strength (β_1) in terms of 1928 dollars. Each dot is a year in the data and is calculated by $\frac{\beta_1}{\alpha} \times \log(N_w)$ where α is the price coefficient and N_w is the size of the news agency w's membership in each year. Panel B plots the magnitudes of the set of coefficients on demographics γ_{AP} and γ_{UP} respectively in terms of 1928 dollars. Red dots for AP are calculated by $\frac{\gamma_{AP}}{\alpha}$ and blue dots for UP are calculated by $\frac{\gamma_{UP}}{\alpha}$ respectively. Bars are 95% confidence intervals.

I find that observable market demographics exhibit heterogeneous preferences across news agencies, but little that advantages UP over the AP or not having a news agency subscription. Comparing to a baseline newspaper without wire subscriptions, I find that the brand effect of UP corresponds to approximately $1 decrease in reader utility when the market has a 1 s.d. increases in its share-urban measure. I also find that the AP enjoys approximately an $0.50 increase in mean reader utility when the market has a 1 s.d. increase in its share white demographics. Neither foreign or manufacturing share exhibit any significant heterogeneity. I cannot reject the null hypothesis that the nesting coefficient is different from 0, suggesting that newspapers are highly differentiated in unobserved dimensions not captured by news agency choices.

Figure 1.3 displays the distribution of estimated newspaper local qualities. There is significant selection into news agency subscription: papers ever bundling both services are of higher local qualities than those who only ever having one services, which in turn are higher than papers who never had a news agency subscription.

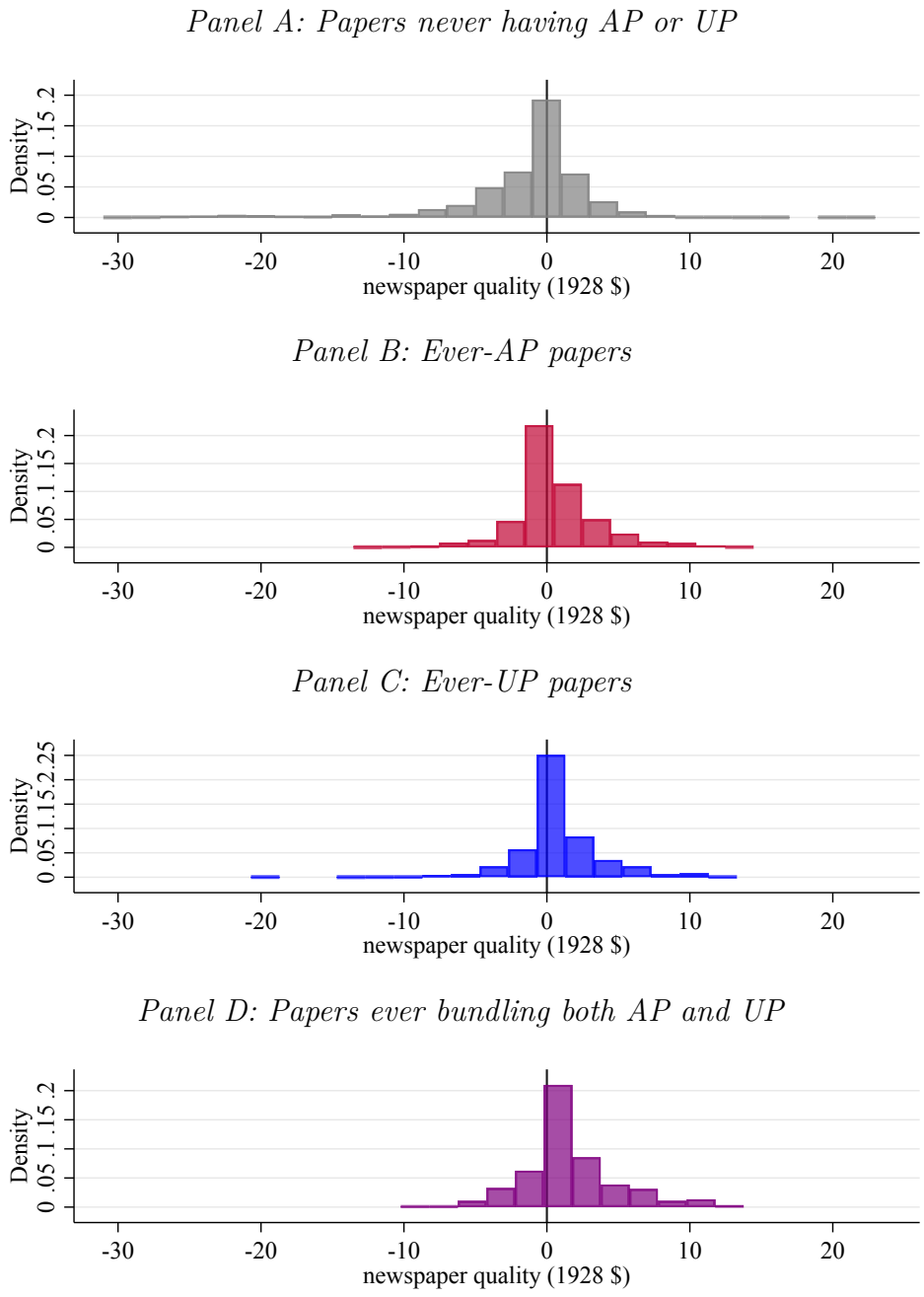


Figure 1.3: Newspaper local qualities (1928 $)

Note: This figure presents the distribution of estimates of newspaper local qualities from Table 1.2, Column 1, residualized on city fixed effects. Estimates are divided by the price coefficient α to be interpreted in units of 1928 dollars. Panel A includes only newspapers that never had either AP or UP service in their lifetime. Panel B, C, and D include papers that had AP, UP, or bundled both services for at least one year in the data. Note that the plotted samples are not mutually exclusive.

1.7 Conclusions

I analyze the role of news agencies toward reader demand for historical American newspapers. Via historical accounts, I show that the value of news agencies to newspapers come from the foreign and national news content they produce as well as acting as an exchange of local news among newspapers. I show via a formal demand analysis that the value of news agencies is substantial and exhibits strong network effects. I further show that news agencies exhibit very little differentiation, unlike newspapers.

Aghion, P. and P. Bolton (1987): "Contracts as a Barrier to Entry," *Am. Econ. Rev.*, 77, 388–401.

Angelucci, C., J. Cage, and M. Sinkinson (2023): "Media Competition and News Diets," *American Economic Journal: Microeconomics (forthcoming)*.

Asker, J. and H. Bar-Isaac (2014): "Raising Retailers' Profits: On Vertical Practices and the Exclusion of Rivals," *Am. Econ. Rev.*, 104, 672–686.

Associated Press (1901): *Annual Report*.

——— (1984): "The Associated Press Assessment Formula History," Tech. rep.

Associated Press v. United States (1945): vol. 326, U.S. Reports.

Benét, S. V. (1933): "The United Press," *Fortune Magazine*, 67–98,100,104.

Berry, S. T. (1994): "Estimating Discrete-Choice Models of Product Differentiation," *Rand J. Econ.*, 25, 242–262.

Blanchard, M. A. (1987): "The Associated Press Antitrust Suit: A Philosophical Clash over Ownership of First Amendment Rights," *Bus. Hist. Rev.*, 61, 43–85.

Cage, J., N. Herve, and M.-L. Viaud (2019): "The Production of Information in an Online World: Is Copy Right?" .

Chiang, C.-F. and B. Knight (2011): "Media Bias and Influence: Evidence from Newspaper Endorsements," *Rev. Econ. Stud.*, 78, 795–820.

Chiou, L. and C. Tucker (2017): "Content aggregation by platforms: The case of the news media," *J. Econ. Manag. Strategy*.

Cohen, R. E. and G. Gordon (1900): *Down to the Wire : UPI's Fight for Survival*, McGraw-Hill Companies.

Dell, M., J. Carlson, T. Bryan, E. Silcock, A. Arora, Z. Shen, L. D'Amico-Wong, Q. Le, P. Querubin, and L. Heldring (2023): "American Stories: A Large-Scale Structured Text Dataset of Historical U.S. Newspapers," .

Djourelova, M. (2023): "Persuasion through Slanted Language: Evidence from the Media Coverage of Immigration," *Am. Econ. Rev.*, 113, 800–835.

Fan, Y. (2013): "Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market," *Am. Econ. Rev.*, 103, 1598–1628.

Gentzkow, M., E. L. Glaeser, and C. Goldin (2004): "The Rise of the Fourth Estate: How Newspapers Became Informative and Why It Mattered," .

——— (2010): "What Drives Media Slant? Evidence From U.S. Daily Newspapers," *Econometrica*, 78, 35–71.

——— (2014): "Competition and Ideological Diversity: Historical Evidence from US Newspapers," *Am. Econ. Rev.*, 104, 3073–3114.

Inter-Ocean Publishing Company. v. Associated Press (1900): 56 N.E. 822,184 Ill. 438.

John, R. R. (2015): *Network Nation: Inventing American Telecommunications*, Harvard University Press.

Miller, N. H. and M. C. Weinberg (2017): "Understanding the Price Effects of the MillerCoors Joint Venture," *Econometrica*, 85, 1763–1791.

Morris, J. A. (1957): *Deadline every minute : the story of the United Press*, Doubleday & Company, Inc.

Nasaw, D. (2013): *The Chief: The Life of William Randolph Hearst*, Houghton Mifflin Harcourt.

Petrova, M. (2011): "Newspapers and Parties: How Advertising Revenues Created an Independent Press," *Am. Polit. Sci. Rev.*, 105, 790–808.

Rosewater, V. (1930): *History of Cooperative News-gathering in the United States*, New York, D, Appleton and Company.

Schwarzlose, R. A. (1989): *The Nation's Newsbrokers: The rush to institution, from 1865 to 1920*, Northwestern University Press.

Silberstein-Loeb, J. (2012): "Exclusivity and Cooperation in the Supply of News: The Example of the Associated Press, 1893–1945," *J. Policy Hist.*, 24, 466–498.

——— (2014): *The International Distribution of News: The Associated Press, Press Association, and Reuters, 1848–1947*, Cambridge University Press.

Silcock, E. and M. Dell (2023): "A Massive Scale Semantic Similarity Dataset of Historical English," .

State of New York (1918): *Annotated Consolidated Laws of the State of New York as Amended to January 1, 1918: Containing Also the Federal and State Constitutions with Notes of Board of Statutory Consolidation, Tables of Laws and Index*, Banks Law Publishing Company.

UNESCO (1953): "News agencies: their structure and operation," Tech. Rep. 42.

Wang, T. (2023): "The Electric Telegraph, News Coverage and Political Participation," .

Whinston, M. D. (2008): "Lectures on Antitrust Economics," *MIT Press Books*, 1.

Xiang, Y. and M. Sarvary (2007): "News Consumption and Media Bias," *Marketing Science*, 26, 611–628.

Chapter 2

News Agency Competition and Exclusive Territory Contracts

2.1 Introduction[1]

The effects of exclusivity on competition remain a matter of considerable debate.[2] A monopolist who controls an essential input can reduce competition by preventing potential rivals from accessing that input. On the other hand, such practices can increase efficiency by preventing excessive entry of rivals and eliminating duplicated fixed costs. Furthermore, excluded rivals may turn to other suppliers and thereby promote competition between suppliers. Thus, exclusive arrangements that are anticompetitive in one segment of an industry can potentially be pro-competitive in another segment. Since U.S. courts currently evaluate exclusive contracts on a case-

[1]Email: quanle@princeton.edu. This paper was part of a previously circulated manuscript titled *Network Competition and Exclusive Contracts: Evidence from News Agencies*.

[2]See Lafontaine and Slade (2008) for a survey of legal attitudes and empirical evidence around this debate and Lee, Whinston, and Yurukoglu (2021) for that of theoretical models and recent structural empirical work on vertical contracting. For the closely related debate on the effects of exclusive franchising on welfare, see Lafontaine and Slade (2014) and Lafontaine and Scott Morton (2010).

by-case basis, there is a growing need for empirical evidence that quantifies these contrasting effects.[3]

I study these effects in the setting of news agencies in the early 20th century United States. News agencies are networks of newspapers that historically exchanged local news via the electric telegraph and coordinated the collection and distribution of national and foreign news.[4] In 1900, the non-profit cooperative Associated Press (AP) was by far the dominant network. Another network called United Press (UP) entered in 1907. Exclusive territory contracts were a key feature of the industry: both firms effectively guaranteed local exclusivity by not selling the news to potential competitors of their existing member newspapers.

When the Supreme Court found the AP's exclusive territory contracts illegal in the landmark antitrust case *Associated Press v. United States* (1945), it had to consider harms from contracts "plainly designed in the interest of preventing competition" against potential benefits from these contracts, such as "foster[ing] [...] the growth of [other] news agencies." The Court's speculation that papers excluded from AP service likely sought telegraphic news from the AP's competitor motivates my research question: would AP v. US (1945) have changed the market structure of news agencies had it been decided during UP's infancy? The history of these contracts further motivates my main counterfactual: state-level policies repeatedly attempted to ban the AP's exclusive territory contracts prior to 1945 but none was enforced until the federal case AP v. US (1945). The answer to this question is also pertinent to

[3]Baseline attitudes toward exclusionary vertical practices have greatly changed over time, in part due to changes in the theoretical and empirical understanding of these contracts. In the early case *United States v. Terminal Railroad Association* (1912), the Supreme Court ruled that a railroad operator was required to let rivals use the only bridge into St. Louis, which the operator owned. In contrast, *Verizon v. Trinko* (2004) held that Verizon was not obliged to share its network with competing local phone companies. Recent matters such as *FTC v. Microsoft and Activision Blizzard* (2023) and *US v. Google* (2023) revolve around similar concerns.

[4]They are also called *wire services* or *news wires*. News agencies are distinct from feature syndicates: news agencies cover hard news while feature syndicates focus on human interest stories, columns, fiction, and comics (Guarneri 2017, Lee 1937). The definition used in this paper also does not cover supplemental news agencies that specialized in one particular news area such as financial news agencies (e.g. Bloomberg) or picture agencies (e.g. Getty Images).

understanding the causes of media competition in the U.S. as well as the impact of exclusive contracts among network industries, in general.[5]

To quantify the effects of exclusive contracts in this industry, I develop a model of news agencies. I use the model to estimate the forces underlying the values and costs of news agencies, endogenize the formation of their networks between 1900 and 1945, and simulate a counterfactual without exclusive contracts. I choose this timeframe specifically to study the dynamics of a nascent network entering and competing with an established monopolist. Furthermore, the setting is amenable to my work because the two firms maintained extensive archives in this era. From these archives, I obtain operational and financial data from both firms as well as then-confidential communications that directly corroborate important components of the model.

The model includes three economic mechanisms relevant to news agencies. First, news agencies exist to pool together newspapers' resources to collect and transmit the news. This function gives rise to *economies of scale*, as larger networks can split the operational costs among more newspapers and so are cheaper for individual members. Second, agencies act as exchanges of local news between newspapers located in different places. This role gives rise to *network effects* on the demand side: larger networks have access to more local news and thus are more valuable to readers. Third, news agencies could potentially be *differentiated* to readers, since some readers could prefer UP's style or brand over that of the AP and newspapers could bundle both services. Economies of scale and network effects are natural barriers to entry, which likely made it harder for rivals to enter and compete with an established monopolist such as the AP. On the other hand, differentiation incentivizes UP's entry.

[5]Network industries represent 9.1% of US GDP (Spulber and Yoo 2014). There has been much recent attention from policymakers and researchers on the effects of exclusion in multi-sided networks such as credit card associations (Rochet and Tirole 2002), Yellow Pages (Rysman 2004), video game consoles (Lee 2013), and hospital networks (Ho and Lee 2019, Ho 2006). Jullien, Pavan, and Rysman (2021) and Evans and Schmalensee (2014) provide comprehensive surveys on this literature. There is less work on exclusion in peer-to-peer communication networks such as social media platforms or news agencies.

In the model, forward-looking newspapers enter and become part of a news agency's network over time. These decisions take into account a stream of discounted future profits, which is an expectation taken over the current and future local market structures as well as the values and costs of news agencies. This dynamic consideration is important since, given exclusive territories, a newspaper subscribing to a news agency can foreclose future rival newspapers who can no longer take up the same subscription.

For each newspaper, the stream of future profits depends on a demand system that is realized period by period. Readers purchase newspapers according to a discrete choice model where a newspaper's characteristics include the value of its news agency subscriptions, which I have estimated in the previous chapter. The newspaper can then sell their readers' attention to advertisers. I allow advertising revenue to vary by the newspaper's readership and market characteristics. I combine the reader demand model, the advertising revenue model, and a two-sided differentiated Bertrand-Nash pricing model to predict a newspaper's variable markup each period. Computing the first-order conditions of the pricing model at realized prices and market shares recovers the newspapers' marginal costs. I use the estimated markups and recovered marginal costs to predict newspapers' expected per-period variable profits at the time of entry. The expectation of these per-period variable profits, less a per-period fixed cost as well as costs related to news agency subscription, constitutes a newspaper's dynamic payoff when it makes entry decisions.

For newspaper and news agency costs, I collect a dataset of assessments on newspapers by agencies, data on annual revenues and expenditures by news agencies, and maps of leased telegraph wires. I use these datasets to estimate the formula that the agencies used to determine their annual assessments on newspapers. These assessments enter as an input to newspaper entry estimation and news agencies'

revenue. The key parameters for entry costs and news agency subscription costs are estimated via preferences revealed by each newspaper's entry decision.

The assessment data shows that telegraph costs are important to the data generating process. A 1% increase in the distance to the nearest news bureau increases payments by 0.25%. Economies of scale matter for both assessment and subscription choices: the entry of a nearby AP paper sharing the same news bureau decreases a paper's AP assessment by 0.28% and increases the probability of it having an AP subscription by 1.8%. For a limited set of large-city markets in which I observe both AP and UP assessments, I find that the two correlate strongly.

Estimates of the supply model show that newspapers internalize the assessments paid to news agencies and the associated economies of scale in their entry and subscription decisions. In 1908, across all possible markets, the average per-capita cost for a newspaper to subscribe to a news agency was $2.26 for UP and $0.50 for the AP, consistent with the original density of the AP's network. These estimates contribute to yet another barrier to entry for UP: members of the sparser network must pay higher costs to obtain the news, even holding constant the network's value.

I take these empirical estimates to a counterfactual simulation in which exclusive territory contracts were already illegal during UP's infancy. Model simulations require searching for new equilibria of newspaper entry and network choice. Solving for the new equilibrium amounts to a fixed-point problem: newspapers join networks based on beliefs over future network effects and economies of scale, which in turn must be internally consistent with the rate at which different networks grow.

I find that exclusive territories granted by the AP to member newspapers contributed to the growth of its rival UP. In the absence of exclusive territory contracts, UP would have 30% lower membership in 1920 and 40% lower membership in 1924. A series of internal letters by UP board members during the organization's infancy revealed that UP was only breaking even due to "a large subsidy" from its primary

proprietor E.W. Scripps.[6] The same letters include a range of valuations of UP. I use the maximum discussed valuation as UP's scrap value, and assume the news agency would exit if it were to lose revenue beyond this amount.[7] In the counterfactual, UP's revenue loss would have exceeded the agency's scrap value within four years. After eight years, the revenue loss would exceed the size of Scripps' newspaper estate, strongly suggesting that UP would have exited in this scenario.

I also study a counterfactual in which the AP enforced an exclusive dealing clause. Exclusive *dealing* contracts would have required AP papers to source telegraphic news only from the AP (unlike exclusive *territory* contracts which required the AP to sell the news to only one newspaper per market). This is a relevant antitrust counterfactual: in 1915, the US Attorney General gave his opinion that the AP had the right to exclusive territory contracts but *not* exclusive dealing. In response, the AP did not enforce its exclusive dealing clauses. Exclusive dealing has complicated theoretical effects in this setting. First, these contracts can hurt UP as AP-incumbents cannot bundle UP news. On the other hand, exclusive dealing potentially weakens the incumbent and incentivizes second entrants who can take up UP service. On net, exclusive dealing would have been detrimental to UP: it loses 15% of newspaper-years compared to the baseline.

In equilibrium, exclusive territories did not achieve their intended effects of limiting competition and making the AP more attractive to local newspapers: I find that AP exclusive territories decreased the number of market-years with AP news. This was due to UP being available as an alternative. If UP was guaranteed to not exist, exclusive territory contracts would have increased the number of market-years with AP news. The AP was not able to preempt the entry of a rival news agency because

[6]Scripps was a newspaper magnate who founded UP in order to "make it possible for any other man to found a newspaper in any city in the Union" in spite of AP exclusion.

[7]Estimating scrap value requires observing many instances of exit (Aguirregabiria et al. 2021). Since there were not many news agencies, I choose to calibrate this parameter using internal industry data. This is similar to Jeon (2022) which uses the sale and demolition prices of container ships and Benkard (2004) which uses publicly reported figures of costs for an aircraft production facility.

its organizational structure as a non-profit cooperative: individual newspapers have incentives to maintain local exclusivity for themselves, even if these contracts lead to the entry of a rival news agency in UP. I find that banning exclusive territory contracts and UP exiting would not have led to higher profits for AP-incumbent newspapers, since these papers would have experienced entry of an AP competitor in their local market. Finally, while exclusive territory contracts harm consumers relative to a free-entry counterfactual, I find that the entry of the upstream network UP significantly dampened such harms relative to a world where exclusive territory contracts were in place and AP was a monopolist.

Related Literature

My paper presents an empirical study of the effect of exclusive contracts on market structure in network industries and emphasizes that a monopolist network's exclusive territory contracts can incentivize rival entry. I am not the first to empirically study questions related to entry incentives due to exclusive contracts: Sinkinson (2020) studies the incentives of handset manufacturer and wireless networks to sign exclusive contracts to soften interbrand price competition; Lee (2013) studies software exclusivity as a tool for entrant video game consoles to gain scale; and Asker (2016) studies whether exclusive distribution leads to cost-based foreclosure of beer brewers in Chicago.

I contribute to the literature on exclusive contracts along two dimensions. First, I explicitly model and estimate the entry probabilities of downstream firms (newspapers) and how entry responds to the costs and values of upstream inputs (news agencies). This model allows me to quantify how changes in upstream values and costs due to exclusives affect downstream entry probabilities, which in turn affects upstream values and costs due to changes in scale. Second, I study the interaction between exclusive territory and exclusive dealing contracts and their effects on entry

incentives of both the downstream and upstream firms. I show that these contracts have complex interactions: exclusive territory protects the downstream monopoly but incentivizes upstream entry, while exclusive dealing protects the upstream firm by preventing the downstream monopoly from bundling suppliers but weakens the downstream monopoly in the process. The setting, while amenable to this analysis, has important simplifying details and limitations. In particular, the setting is in contrast to a general setting where entrant networks can set access prices under costs at the beginning in order to gain a critical mass of users (Cabral, 2011). A more general setting might also have Incumbent networks could also lower access prices in response to the threat of entry (Fudenberg and Tirole 2000).

I also contribute to a literature that studies the causes of media variety in US history by explicitly modeling the vertical nature of the news media market and the market structure of news agencies. The model extends Gentzkow, Shapiro, and Sinkinson (2014)'s analysis of newspapers' ideological choices to allow for news agency bundle choice and dynamic entry decisions.The dynamic extension uses a stochastically alternating-move, finite-horizon model closely related to Igami and Uetake (2019) and is estimated using a full-solution approach as in Rust (1987).

Roadmap

Section 2 describes the data. Section 3 presents the descriptive evidence that motivates the structural analysis. Section 4 describes the formal empirical model. Sections 5 and 6 describe the estimation procedure and estimates. Section 7 presents the model fit and counterfactual simulations. I conclude in Section 8.

2.2 Data

2.2.1 Newspaper characteristics data

I incorporate additional data from several sources other than those described in the previous dissertation chapter.

First, I digitize Annual Linage Supplement provided by Editors and Publishers for the years 1932, 1936, and 1940. This provides me information on advertising quantities for a large subset of newspapers. Second, I use the Inland Press Cost and Revenue Study for 1928, also provided by GSS (2014). This was a non-random sample of newspapers for which I could observe actual marginal costs and advertising revenue.

In this chapter, I use the Inland Press data in following Tiew (2022) to calibrate the mean advertising revenue of newspapers and thus the average discount from their posted advertising rates in the trade journal Editors and Publishers. In order to estimate advertising revenue for years prior to 1932, I collected aggregate data advertising revenue for newspapers over time from Lee (1937). I calculate use the circulation data to measure average per-reader advertising revenue for years before 1932 and calibrate the variation in average per-reader advertising revenue over time accordingly.

2.2.2 Leased wires maps and news agency costs

I collect a number of different leased wire maps for the purpose of the study, whose sources are detailed in the appendix. These maps either come from internal communications or public materials such as promotional circulars. While they differ in specifics, each map shows leased wire connections between cities and whether a node is designated as a bureau.[8]

[8]There is substantial evidence that the reported details of telegraph leased lines are accurate. For the original AP map in 1900, I corroborate their details with the corresponding telegraph bills

Figure 2.1 shows two examples of these maps. In the current version of the paper, I use the maps to identify the relevant nearest bureau for each newspaper as connected via the telegraph lines. Future versions of the paper will estimate an explicit network planning model using observed connections.

2.2.3 News agency internal operations: assessments, costs, and contracts

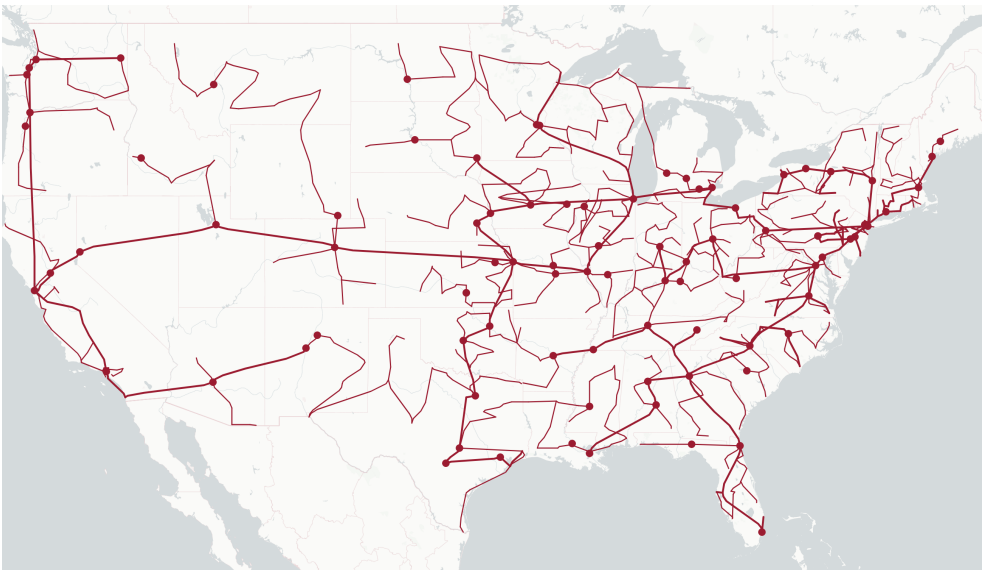
I collect several datasets related to internal news agency operations. The appendix details the sources of this data.

First, I digitize a dataset of payments made by newspapers to the AP in 1926. The data specifies weekly payments and categorizes them as either from leased wire clients or telephone clients. I use this data to estimate the assessment formula for the AP.

Second, I digitize the complete aggregate expenditures from the AP and UP between 1900 and 1945, as submitted to the Supreme Court as part of *AP v. US* (1945). I use this data to evaluate the fit of the assessment model for both the AP and UP as well as for years other than 1926.

recorded in AP's annual internal audits and find them to be consistent. In subsequent years for AP, these records were either part of the annual directory or officially submitted in courts. As for the UP, the information was included in public promotional materials and also used on UP's official stationery to attract new clients. Any discrepancy or misrepresentation would have likely been identified by potential clients.

Panel A: AP 1936



Panel B: UP 1928

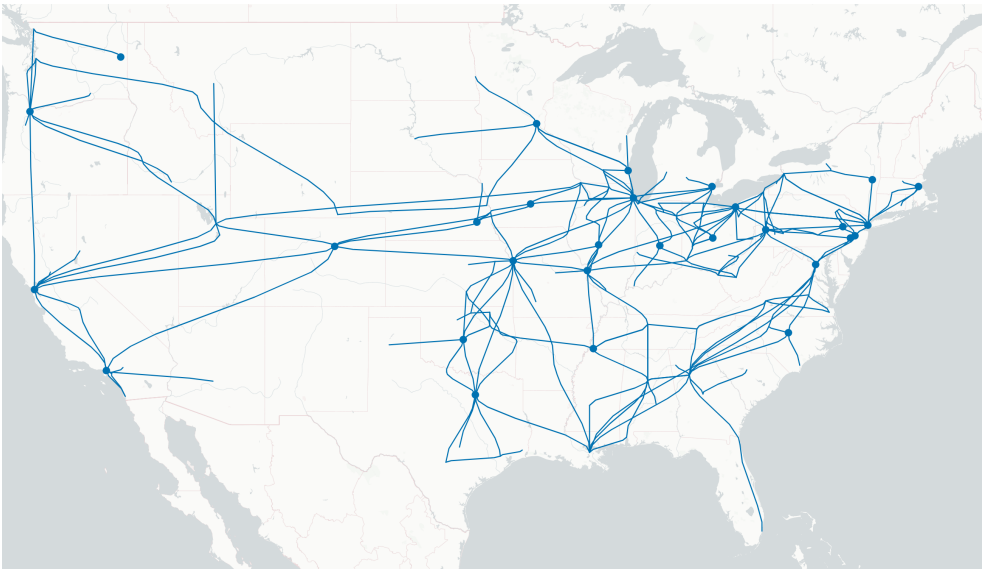


Figure 2.1: Maps of leased wires

Note: These are maps of the leased telegraph line system for each news agency. Lines denote telegraph connections. Dots denote bureaus operated by each news agency at the time. Panel A depicts the Associated Press's network in 1936, including national trunk lines which are bold. This particular AP map was submitted as evidence in *AP v. US* (1945) and printed in Lee (1937). Panel B depicts the United Press's network in 1928. Duplicated connections are possible due to the likely inclusion of both Day and Night wires. This particular map was printed on the back of UP's stationery in 1928 and reprinted in many UP newspapers as part of UP's advertising campaign in 1929.

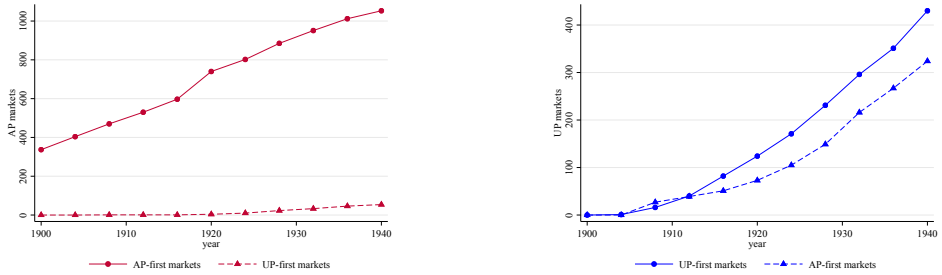
2.3 Descriptive Evidence

2.3.1 Exclusive Contract and News Agency Growth

The key argument of the paper depends on the source of UP's growth. In this section, I show that a substantial share of UP's membership came from markets where the AP already had an incumbent.

In Figure 2.2, for each news agency, I split markets (a city and time-of-publication combination) into two categories: whether a particular news agency was present (i.e. had a member newspaper) before its competitor.

Panel A: News Agencies' Membership by their Markets' Entry Order



Panel B: News Agencies' Population Coverage by their Markets' Entry Order

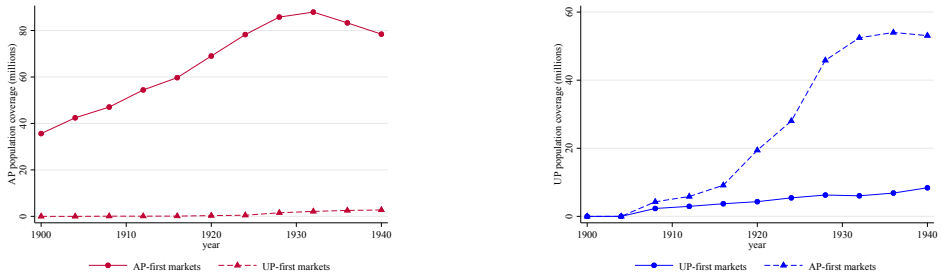


Figure 2.2: Membership and Population Coverage by their Markets' Order of Entry
Note: This figure depicts whether a news agency's size comes from markets where they were first or second relative to the other agency. Panel A shows size in terms of the number of newspapers. Panel B shows size in terms of the total population in markets where the news agency has at least one newspaper subscriber, double-counting Morning and Evening markets. The left column plots the Associated Press's size and the right column plots United Press's size. For each panel, the solid line plots the news agency size from only markets where the agency had presence before the other agency, and the dashed line plots the news agency size from only markets where the agency was had presence after the other agency.

Panel A depicts the distribution of both the AP and UP's membership, measured by the number of subscribing newspapers. Initially, the AP enjoyed a considerable size advantage. Furthermore, the AP was the first entrant in virtually all of its markets. On the other hand, almost half of UP's presence was in markets where it followed an AP paper. This underscores the initial dominance of the AP's installed base and UP's status as a challenger against a dominant monopolist.

For Panel B, I define "population coverage" as referring to the total population of cities where an agency had a newspaper, double-counting across Morning and Evening papers. The figure highlights that virtually all of UP's covered population comes from places where it was second to the AP. This fact is consistent with the fact that the AP already had a substantial presence in most large cities by the time that UP entered. The key consideration in a counterfactual without exclusive contracts is whether UP could have gained this membership and population coverage if new entrants in AP-incumbent markets could have acquired AP service.

2.3.2 Economies of Scale and Assessments

Assessment Model

In this section, I estimate a formula of the AP's assessments using the 1926 data. I perform this exercise using three different specifications. The first and second specifications allow a market's telegraph cost (proxied by its distance to the nearest news bureau) is equally shared across all AP papers using that bureau or across all AP papers in the same state respectively. I also estimate a third specification where costs can vary flexibly with either the number of papers sharing the bureau or state news circuit.[9]

[9]This last specification is motivated by the fact that I do not have access to a precise formula for which fraction of assessments relate to general, state, and local cost. Allowing for a flexible specification avoids committing ex-ante to the way the AP splits costs among papers in the same states or sharing the same bureau.

For each AP member in market m in 1926, I identify its nearest bureau $b(m)$ by tracing along the telegraph maps as observed in 1936 and identifying the nearest city with a bureau by total edge distance. I fit a formula of assessments on the assessment data by the following cross-section regression:

$$\begin{aligned}\log\left(\tau_m^{AP}\right) = \chi_0 &+ \chi_1 \log\left(CityPop_m\right) \\ &+ \chi_2 log\left(1 + \frac{\text{Distance to Bureau}_{m,b(m)}}{N_{b(m)}^{\text{AP same-bureau}}}\right) \\ &+ \chi_3 \mathbf{1}\left\{m \text{ has a bureau}\right\} + \epsilon_m^C. \end{aligned} \tag{2.1}$$

where τ_m^{AP} denotes the total annual payment by a newspaper to the AP. On the right-hand side, the first term denotes the population of the city; the second term includes the as-the-crow-flies distance between the market m and its nearest bureau $b\left(m\right)$, divided by the number of papers sharing that bureau (or the number of papers in the same state); and the third term is an indicator for if the market has a news bureau. The final term ϵ_m^c is an observational error term to allow for potential mismatches in my calculation of the right-hand side variables and the AP's calculation. The most flexible specification separates the distance-related component of cost into three terms: just the distance to the bureau, the natural logarithm of the number of AP papers sharing the same bureau, and the natural logarithm of the number of AP papers in the state.

Table 2.1 describes the results. Columns 1 and 2 report results for specifications where a newspaper's distance to the bureau is shared by other bureau connections vs. the number of AP papers in the state respectively. Column 3 reports results for the flexible specification.

	$\log(\tau^{AP})$		
	(1)	(2)	(3)
$\log(CityPop)$	0.707	0.694	0.726
	(0.020)	(0.021)	(0.020)
$\log\left(1 + \frac{\text{Distance to Bureau}}{\text{Number of AP Papers sharing the Bureau}}\right)$	0.327		
	(0.025)		
$\log\left(1 + \frac{\text{Distance to Bureau}}{\text{Number of AP Papers in the same state}}\right)$		0.301	
		(0.026)	
$\log(1 + \text{Distance to Bureau})$			0.251
			(0.031)
log(Number of AP Papers sharing the Bureau)			-0.278
			(0.041)
log(Number of AP Papers in the same state)			-0.170
			(0.034)
$\mathbf{1}\{\text{Bureau Market}\}$	1.232	1.059	1.354
	(0.084)	(0.082)	(0.154)
Constant	0.679	0.993	1.252
	(0.234)	(0.235)	(0.287)
N	456	456	456
R^2	0.796	0.781	0.805

Table 2.1: Estimates for the AP Assessment Model

Note: This table presents regressions of AP annual assessments for leased wire clients in the year 1926 on observable market characteristics. τ^{AP} denotes AP annual assessments for 1926. *CityPop* is the population of the city in 1926 (linearly interpolated using counts from the decennial census). Distance to Bureau is the distance of the market to the nearest AP news bureau. The number of AP papers sharing the Bureau is defined as the number of AP papers with whom the market shares its nearest bureau with. The indicator $1\{\text{Bureau Market}\}$ denotes whether the market has an AP news bureau.

Across all specifications, a 1% increase of the local market population increases annual assessments between 0.69% and 0.73%. Both telegraph connection costs and the associated economies of scale are economically significant. A 1% increase in the distance of the market to the nearest bureau leads to a 0.25% in costs. Assessments fall as more papers share the same bureau and state news circuit: a 1% increase in the number of newspapers that share a given bureau leads to a 0.28% decrease in annual assessments, and the same figure is 0.17% for a 1% increase in the number of papers in the same state. Since the median AP bureau is serving eight newspapers,

another paper connecting to that same bureau would have decreased those members' assessments by 2.5%.

Panel A of Figure 2.3 shows that the formula fits the AP data well ($R^2 \approx 0.81$). In Panel B, I plot observed assessments for a limited set of cities and newspapers for which I observe both assessments. Regressing the AP's per-capita assessments on that of UP yield a slope of .992 (confidence interval [.624, 1.316]). My inability to reject a slope of 1 motivates me to use the estimated AP formula for calculating UP revenue.

Bureau Sharing Correlates with Subscription Choice

In this section, I show that the determinants of economies of scale correlate with the choice of news agency subscription. In particular, I show that whether a newspaper takes up AP or UP service is highly correlated with the number of other papers of the same service that connects to the same bureau.

I estimate a linear probability model of news agency subscription choice across newspapers as a function of the number of AP and UP papers sharing the bureau:

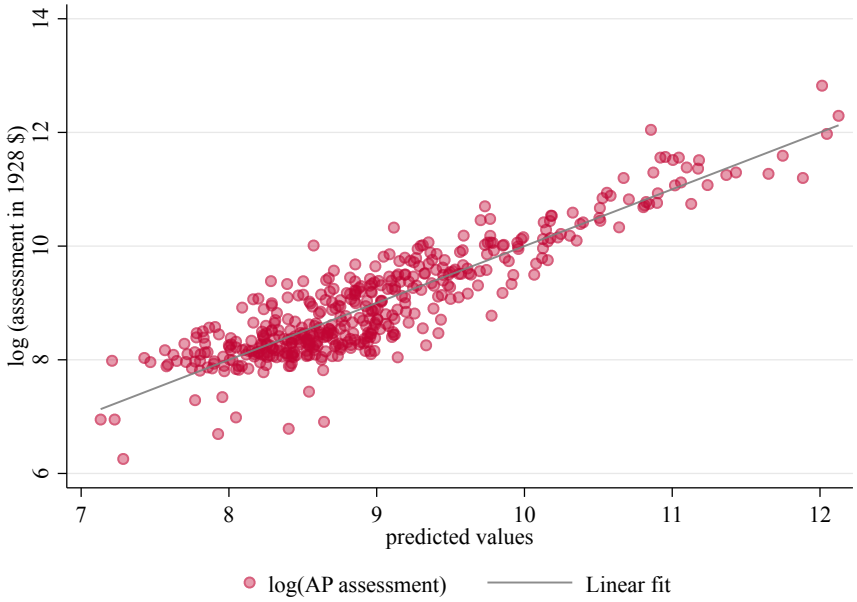
$$w_{jmt} = \kappa_{AP} N_{b(m)}^{\text{AP same-bureau}} + \kappa_{UP} N_{b(m)}^{\text{UP same-bureau}} + \text{Controls}_{jmt} + \epsilon_{jmt}^{w},$$

where two separate specifications for $w \in \{\text{AP}, \text{UP}\}$ denoting a newspaper j's subscription to w in market m in year t, and τ denotes the predicted assessments for either agency. The main regressors are the number of *other* AP papers that share the same bureau (and same for UP). All specifications control for the demographics, newspaper fixed-effects and Year $\times$ Time-of-Publication fixed effects.

Table 2.2 describes results. First, note across Columns 1 and 2 that when an additional other AP newspaper shares the same bureau, the probability of a newspaper subscribing to the AP increases by 1.8% and decreases for UP by 0.7%. Second, an

additional other UP newspaper sharing the bureau increases the probability of joining UP by 1% and decreases that for the AP by 0.7%. Together with the estimated cost formula, these estimates are suggestive that the nearby density of a network influences subscription behavior by reducing telegraph costs born by any particular member.

Panel A: Predicted AP Annual Assessments



Panel B: AP vs. UP Per Capita Assessment

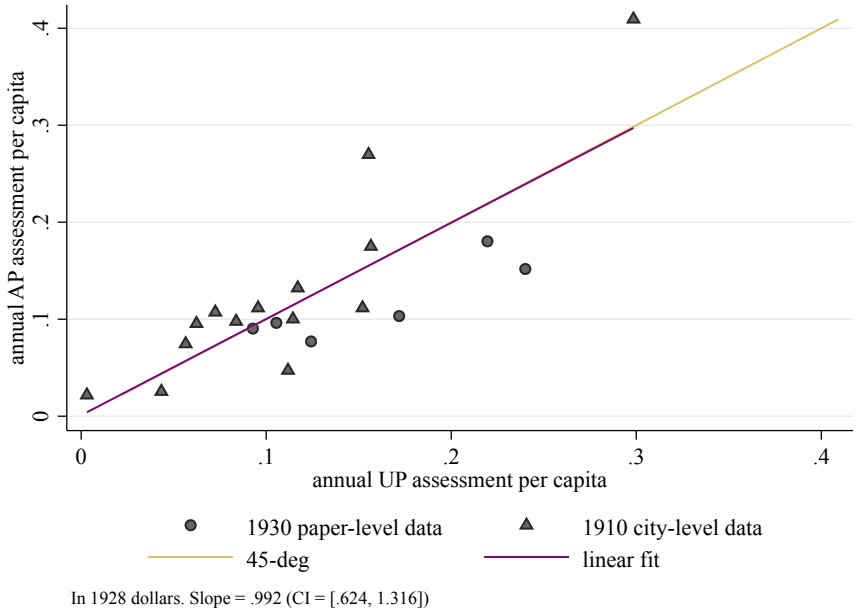


Figure 2.3: Annual Assessments

Note: This figure plots the annual assessments for the Associated Press and United Press normalized to 1928 dollars. Panel A shows the relationship between observed AP annual assessments and predicted values using the estimates from Column 3 of Table 2.1. Each dot represents a newspaper and the line is the linear fit. Panel B shows the relationship between annual AP assessment per capita and annual UP assessment per capita. The circles represent a set of papers bundling both AP and UP services in 1930 for which I observe both assessments, and the triangles represent a set of markets for which I observe both an AP and UP newspaper and their assessments. The yellow line is the 45-degree line and the purple line is the linear fit of AP to UP assessment per capita pooling both datasets.

	(1) AP	(2) UP
Number of other AP papers sharing the bureau	0.018	-0.007
	(0.001)	(0.001)
Number of other UP papers sharing the bureau	-0.007	0.010
	(0.001)	(0.001)
R^2	0.874	0.735
Number of newspaper-years	15,826	15,826

Table 2.2: Bureau-Sharing Predicts Newspaper Subscription to News Agency
Note: This table describes the relationship between a newspaper subscribing to the AP/UP and the number of other AP/UP papers sharing their nearest bureau in a linear probability model. Each observation is a newspaper-year. All specifications control for demographics, newspaper fixed-effects, and Year × Time-of-Publication fixed effects.

2.4 Empirical model

2.4.1 Roadmap

In this section, I develop a model of dynamic newspaper entry and subscription to news agencies that depends on the values and cost structure of news agencies over time. The goal of the model is to capture the key economics forces of economies of scale, network effects, and differentiation in a simple manner. I focus the model on these particular forces as they are most uniquely important to the economics of news agency, thus necessarily abstracting away on important forces that were previously studied in the newspaper literature.

I will first present the a model of static stage game payoffs for newspapers, which nests the reader demand model for newspapers described in the previous chapter, a model of advertising revenue, and a Bertrand-Nash pricing model. I will then present an entry model where entrants and incumbents successively make choices over which news agencies to subscribe to.

I repeat my model notation here, which is identical to that of the previous chapter. Consider all possible markets $m \in \{1, \ldots, M\}$ spanning finitely many discrete time periods $t \in \{1, \ldots, T\}$. A market is defined as a geography and time-of-publication combination. I refer to the set of incumbents for each market-period combination as $\mathscr{J}_{mt}$ and cap the total number of firms possible in each period at J^{max}. For notational ease, I refer to the vector of market states being considered by the mover in each period as Ω_{mt}: this includes anything relevant to payoffs and known by the mover.

I will refer to a newspaper j's bundle of subscriptions to news agencies w as W, where W is any subset of $\{\text{AP}, \text{UP}, \text{INS}\}$ (including $\emptyset$).[10]

2.4.2 Variable profit

Reader demand I omit discussion of the reader demand specification, as it is identical to that of the previous chapter.

Advertising revenue In this section, I specify the per-reader advertising revenue that a newspaper gets given its demand a_{jmt} to be

$$\log a_{jmt} = \lambda_0 + \lambda_1 \log \text{Pop}_{mt} + \lambda_2 \log s_{jmt} + X^{\text{ad}}_{mt}\vec{\lambda}_3 + \iota_{jmt}, \tag{2.2}$$

where λ_0 is a constant, λ_1 is a coefficient on the newspaper's market share s_{jmt}, $\vec{\lambda}_3$ is a vector of coefficients on market-level market characteristics X^{ad}_{mt} that affect advertising, and ι is an i.i.d. error term. In the main specification, X^{ad} is the same vector of observable demographics used for the reader demand model including the population share of a town that is Urban, White, works in Manufacturing, or is Foreign-born.

[10] The letter w is chosen for *wire* services.

Exponentiating both sides of this equation yields a_{jmt} as a function of the vector of prices in the market $\mathbf{p}_{mt}$ via market share s_{jmt}.

Expected variable profits and pricing model For each period, given a vector of prices $\mathbf{p}_{mt}$, an incumbent $j \in \mathscr{J}_{mt}$ earn the following variable profit:

$$v_{jmt}\left(p_{jmt}, \mathbf{p_{-j,mt}}\right) = \left[p_{jmt} - MC_{jmt} + a_{jmt}\left(p_{jmt}, \mathbf{p}^*_{\mathbf{-j,mt}}\right)\right] q_{jmt}\left(p_{jmt}, \mathbf{p}^*_{\mathbf{-j,mt}}\right),$$

where p_j is the firm's own price, $\mathbf{p}_{-j}$ is the vector of other firms' prices, MC_{jmt} is the marginal cost per copy of newspaper, and a_{jmt} per-reader ad revenue function which depends on the reader price vector.

I assume that firms compete according to a Bertrand-Nash oligopoly price setting assumption (Nevo 2001). The equilibrium price vector $\mathbf{p}^*$ will have each firm j's optimal price p^*_{jmt} maximizing its variable profit for that period v_{jmt}, i.e.

$$p^*_{jmt}\left(\mathbf{p}^*_{\mathbf{-j,mt}}\right) = \arg\max_p v_{jmt}\left(p, \mathbf{p}^*_{\mathbf{-j,mt}}\right) \tag{2.3}$$

where $\mathbf{p}^*_{\mathbf{-j,mt}}$ is the vector of other firms' prices and q_{jmt} is j's reader demand curve for that period.

2.4.3 Newspaper entry and news agency subscriptions

Timing. There are two types of periods. I refer to even periods as *Entry* periods where potential entrants can come into the market and odd periods as *Update* periods where incumbents can change their subscriptions. Demand and costs are only realized during *Entry* periods — thus *Update* periods are essentially intermediate stages where incumbents can change their subscriptions and are only included for notation simplicity. I now describe what happens every period in order.

At the very beginning of the game, news agencies publish their schedule of assessments for every market, τ_{mt}^{AP} and τ_{mt}^{UP} . This formula depends on all components described in section 4.2.1. This formula is taken as given for the remainder of the game.

For each period, nature decides the mover. If it is an *Entry* period and $|\mathscr{J}_{mt}| < J^{max}$, a potential entrant newspaper j^{+} is born and becomes this period's mover. This potential entrant can either enter the market with any combination of news agency subscriptions available (including not taking up wire news) or stay out. If it is an *Update* period, nature picks an incumbent $j \in \mathscr{J}_{mt}$ according to a selection rule $\Upsilon(j)$ to be this period's mover, such that $\sum_{j \in \mathscr{J}_{mt}} \Upsilon(j) = 1$. The chosen mover can decide whether to add news agency subscriptions to their current bundle. There is no option to exit the market and I treat exit probabilites for each market as exogenous. The game skips an *Entry* period if there is the maximum number of incumbents in the market. The game also skips an *Update* period if there is no incumbent in the market.

Choices. The mover has access to available news agency subscriptions $W \subseteq \overline{W}_{jmt}$. Recall that W could be any possible subset of $\{\text{AP}, \text{UP}, \text{INS}\}$ including $w = \emptyset$, which denotes an option to enter without telegraphic news. I abuse notation and include in $\overline{W}$ a choice $w = Out$ that is only available to potential entrants. The set of available subscriptions $\overline{W}_{jmt}$ also depends on the exclusion regime in place and the set of incumbents' subscriptions: in my baseline specification, if there is already an incumbent's subscription to a news agency, the agency's availability in future movers' choice sets is removed (except for the incumbent holding the exclusive right).

Payoffs and Parametrization. To simplify the notation in this section, I take the current market structure $\mathbf{W}$ and market states Ω as given in profit and cost functions and omit them when I write the functions.

Choosing a bundle of news agencies W is equivalent to signing a perpetual contract with agencies $w \in W$: the mover agrees to pay the sum of assessments τ_{ms}^{W} for the current period $s = t$ as well as all future periods $s > t$. Thus, the stream of discounted payoffs to mover j from the bundle of subscriptions W is

$$\Pi_{jmt}^{*}(W) = \begin{cases} V_{jmt} + \delta^{s-(t+1)} \sum_{s=t+1}^{T} \mathbb{E}\left[V_{jms}\right] + \eta_{jmt}(W) & W \neq \text{Out} \\ \eta_{jmt}(W) & W = \text{Out} \end{cases}, \tag{2.4}$$

where δ is a discount factor, $\eta_{jmt}(W)$ is a private cost shock specific to choice W, and the per-period overall profit is:

$$V_{jms} = \begin{cases} \mathbb{E}\left[v_{jms}(\mathbf{W})\right] - FC_{jms} - WC_{jms}(W;\tau) & \textit{Entry} \text{ period} \\ 0 & \textit{Update} \text{ period} \end{cases},$$

where v denotes the expected static stage game variable profit per period described in Section 5.2, FC denotes the per-period fixed-cost, and WC denotes the per-period costs related to the newspaper's news agency subscriptions and is a function of the assessment schedules τ. The uncertainty is over possible realizations of the newspaper local quality $\overline{\xi}_{jm}$ and marginal costs MC_{jms} for newspapers that are not yet in the market.

I assume that $\eta_{jmt}(W)$ is a private shock i.i.d. across markets and periods. I follow the intuition of nested logit demand systems and specify two components for η:

$$\eta_{jmt}(W) = \zeta_{jomt}^{Entry} + \left(1 - \sigma^{Entry}\right) \epsilon_{jmt}^{Entry}(W),$$

where $\epsilon_{jmt}^{Entry}(W)$ is distributed EV-1 and specific to each choice W, ζ_{jomt}^{Entry} is an idiosyncratic cost shock that is only specific to choice of an entrant staying out or entering the market. As in Cardell (1991), ζ_{jomt}^{Entry} follows the unique distributions

such that η is also EV-1. σ^{Entry} captures the correlation of cost shocks related only to the choice of news agency as opposed to an entrant's choice to enter the market.

Given this parametrization, for a given mover and realizations of shocks, their unique optimal choice of subscription bundles is

$$W^* = \arg \max_{W \in \overline{W}_{jmt}} \Pi^*_{jmt}(W),$$

and the ex-ante probability of the mover choosing W follows the familiar nested logit choice probability expression.

Beliefs It remains to specify the beliefs that form the expectation $\mathbb{E}[V_{jms}]$. I assume that the mover knows the complete path of fixed costs FC and assessments τ at the time of making their decisions. They also know the path of the demand shifters that change v. Thus, the only uncertainty for an incumbent mover is the realizations of future market structures, which include the mover's and competitor's subscription bundle, marginal costs, and local qualities.

2.5 Estimation

2.5.1 Demand

For reader demand, I use the model estimates obtained in the previous chapter.

For advertising demand, I estimate the model via two-staged least squares for the years 1932, 1936, and 1940. An identification challenge is that papers that have high market share may also have unobservable qualities that make their readers more attractive to advertisers. I instrument for reader market shares using the recovered newspaper local qualities $\overline{\xi}_{jm}$ estimated in the reader demand model. Since reader demand estimates are unavailable for 1932, the specification with this particular

instrument omits observations for that year. I also present the OLS with the 1928 advertising data where results are similar.

I have two measures of total advertising revenues: one implied by the Inland Cost Revenue and Cost Survey data and one implied by multiplying advertising rates posted in *Editor and Publisher* with reported linage[11,12]. I find that the two figures are greatly mismatched: surveyed revenue is approximately 5% that of calculated revenue from advertising quantity and posted prices for small markets (those below the 80th percentile in population), and approximately 10% that for large markets. This is consistent with Tiew (2022) who finds the discount to be 3% for the period from 1932 to 1992 and is likely because advertising rates are often individually negotiated and can involve steep discounts relative to posted rates. I scale the calculated revenue by 5% for small markets and 10% for large markets respectively in order to account for this discount.

2.5.2 Supply

Newspaper entry and news agency subscriptions

To estimate the entry and subscription choice model, I take the demand parameters estimated as given and estimate the supply side using maximum likelihood. The key parameters to be recovered are entry costs and costs specific to news agency subscription bundle choice.

[11] Advertising rates posted in *Editor and Publisher* are often consistent with rates listed in individual newspapers' "rate cards." These cards provide information for the benefit of advertisers which includes the newspaper's rate, potential quantity discounts, deadlines, circulation, and the paper's mechanical specifications.

[12] In this period, a standard unit for advertising buy is an "agate line," which is 1 1/14 inches in height and one column (usually 2 inches) in width.

For each market (a city-time of publication combination), I assume that there are at most $J^{max} = 2$ newspapers active at any given time. I also only estimate the model on entrants and subscription decisions made *after* 1912.[13]

The data is available in four-year intervals. I split each year into two periods: an initial *Update* period where an incumbent could change their wire subscription, and then an *Entry* period where the potential entrant could come in. The data on the supply side consists of market characteristics that vary with entry cost Z^{FC}_{jmt}, market characteristics that vary with wire choice Z^{wire}_{jmt}, and choices W_{jmt} conditional on choice set $\overline{W}_{jmt}$ for each market-period.

I assume that entry cost FC_{jmt} is linear to Z^{FC}_{jmt} and wire cost for each subscribed wire w_{jmt} is linear to Z^{wire}_{jmt}. For each parameter guess, I compute the probability for each choice following the nested logit probability formula (see Section 5.3). The log likelihood of observing the data conditional on parameters is

$$\log L = \sum_m \sum_t \log \Pr \left[W_{jmt} = \arg \max_{W \in \overline{W}_{jmt}} \Pi^*_{jmt} \left(W | \theta^{entry}, \theta^{wire}, \sigma^{entry}, \rho \right) \right]. \quad (2.5)$$

I am not able to estimate the nesting parameter σ^{Entry}. On the reader demand side, the identification of the nesting parameter is from the variation in reader's choice sets. The only variation in choice set that I have in wire subscription choice is from first vs. second entrant (due to exclusion), which may be unreliable toward identifying how subscription costs vary relative to entry costs. Instead, I calibrate σ^{Entry} is so

[13]This sample is motivated by two factors. First, many papers that entered news agencies prior to 1912 constitute UP's installed base (including papers belonging to the Scripps chain). These papers may have a difference entry model than those that came into UP later on. Second, the AP's geographical radius for exclusive territory contracts were not standardized before this period.

that the relative magnitude of the EV-1 shock from wire choice vs. entry choice is approximately the relative magnitude of reported wire cost vs. fixed cost.[14]

I estimate and report costs as scaled by the market population being the per-capita fixed cost $\frac{FC}{Pop}$ and per-capita news-agency related cost $\frac{WC}{Pop}$.

Newspaper-specific specification I assume that $\frac{FC}{Pop}$ consists of two constants: a component that is inversely related to city population, a fixed component, and a component that varies with the distance to the nearest bureau. The last component is meant to capture any investment in news collection that a newspaper would need to make regardless of wire subscription in order to enter. I estimate

$$\frac{FC_{mt}^{AP}}{CityPop_{mt}} = \frac{\theta_0}{CityPop_{mt}} + \theta_1 + \theta_2 \log\left(\frac{Dist_{m,b(m)}}{CityPop_m}\right).$$

News agency-specific fixed cost specification. I flexibly specify the news agency-specific fixed cost to have three components, motivated by the AP assessment formula. The first component is a constant that is inversely related to population. The second component is the general news agency operational cost, which is divided equally among all covered population of a news agency. I normalize this population to be 1 for AP's covered population in the year 1928. The third component is taken directly from the assessment formula.

I include two additional parameters. First, I include a cost multiplier that results from bundling more than one news agencies' content. This is meant to capture any needed investments (e.g. keeping the staff of the two agencies in separate rooms so that they do not accidentally leak breaking news) or savings (e.g. telegraph maintenance cost) from bundling services. Second, I add a cost trend that allows

[14]The intuition is that the relative importance of these two shocks should match the relative importance of FC vs. cost of wire connection, so $(1-\sigma) = \overline{\left(\frac{\tau}{FC}\right)}$. Using reported means for fixed costs in the Inland Press survey and means of the transfer formula, I pick $(1-\sigma) = 0.25$ or $\sigma = 0.75$.

the news agency costs to increase (or decrease) over time according to an exponential process.

Concretely, I estimate

$$\frac{WC_{mt}^{AP}}{CityPop_{mt}} = \frac{\theta_0^{AP}}{CityPop_{mt}} + \frac{\theta_1^{AP}}{PopAP_t} + \theta_2^{AP}\frac{\tau_{mt}}{CityPop_{mt}}$$

and similarly for UP, and the total wire cost is

$$\tau_{mt} = \begin{cases} \theta_{trend}^t \tau_{mt}^{AP} & \text{AP} \\ \theta_{trend}^t \tau_{mt}^{UP} & \text{UP} \\ \theta_{trend}^t \theta_{bundling} \left(\tau_{mt}^{AP} + \tau_{mt}^{UP}\right) & \text{AP and UP} \end{cases}.$$

Mover selection rule. The mover selection rule Υ decides which incumbent is allowed to update their subscription bundle each period. I set the selection rule to be equal probabilities for all incumbents, $\Upsilon(j) = \frac{1}{\mathscr{J}}$. [15]

Beliefs. In this model, a mover must have beliefs regarding future market structures. This includes their own and their competitors' local newspaper qualities, denoted as $\overline{\xi}_{jm}$, and marginal costs, denoted as MC_{jm}. For both variables, I assume that movers expect realizations to match the median estimated values based on the city's population decile. Additionally, movers accurately predict newspaper exit probabilities and assume that any incumbent has an exit probability consistent with the empirical probability for that particular city's population decile and year.

Computational simplifications. I make two simplifications in computing the model.

[15] This also means that in estimation, if I do not see a change in news agency subscriptions for a two-paper market during an Update period, I assume randomly that one of the two incumbents was picked and decided to not change their subscriptions.

First, I do not explicitly recompute the stage-game equilibrium profit v^* for each parameter guess. Instead, I computed v^* on 100,000 grid points of $\left\{\delta_j^{\text{reader}}, \delta_j^{\text{ad}}, MC_j\right\}_{j\in\{1,2\}}$. I then use a cubic fit using all possible combinations of linear, quadratic, and cubic terms of these $3 \times J$ variables. The R^2 of this fit is 0.98 for monopoly markets and 0.95 for duopoly markets. I take the fitted values from this exercise to form equation (2.4). Implementation details of this exercise are in the appendix.

Second, in computation, I impose a shorter horizon for firms than they would have under the full game. This means that firms only take into account $T^{Horizon}$ periods from the period t when it makes its decision (unless $T > T^{Horizon}$). I set $T^{Horizon} = 6$. I calibrate the discount factor δ so that the sum of future values of δ is equivalent to that of $\delta^* = 0.95$ in an infinite game. [16]

2.6 Empirical estimates

2.6.1 Determinants of Variable Profit

Table (2.3) reports estimates of the advertising revenue model respectively. I report both the OLS and the preferred specification, the two-staged least squares (2SLS) that uses newspapers' quality as an instrument for newspaper market share. Consistent with Tiew (2022), I find that observed advertising revenues are far smaller than what would be implied from linage and posted rates, approximately 0.1 for large markets and 0.05 for small markets. I find that per-reader advertising revenue drops sharply as the paper's market share increases. This coefficient is consistent with standard models of media advertising where the marginal eyeball is potentially less valuable to advertisers.[17]

[16] This means finding δ such that $\frac{1}{1-\delta^*} = \frac{1-\delta^{T^{Horizon}}}{1-\delta}$ where $\delta^* = 0.95$. For $T^{Horizon} = 6$, $\delta = 1.483$.

[17] Two sets of theories can imply this result. The marginal reader might makes the readership's demographic more diverse and less niche, making it difficult for advertisers to target their desired demographics (Chandra 2009). The marginal reader could also be more likely to consume other

	(1) OLS	(2) 2SLS
Log Market Share	-0.064	-0.200
	(0.043)	(0.111)
Log City Population	0.142	0.008
	(0.034)	(0.140)
White	0.021	0.410
	(0.315)	(0.324)
Manufacturing	-0.636	0.517
	(0.527)	(0.559)
Foreign-Born	1.486	0.266
	(0.594)	(0.640)
Urban	0.002	0.040
	(0.155)	(0.166)
Big-market discount rate	0.1	
Small-market discount rate	0.05	
First-stage F statistic	-	32.57
N	1,822	855
R^2	0.04	0.22

Table 2.3: Estimates for the Advertising Revenue Model

Note: This table presents the advertising revenue model described in Equation (2.2). The "discount rate" refers to the scaling factor needed to match the Inland Press revenue survey data. The instrument for the 2SLS is the estimated newspaper local quality from the reader demand system. Both specifications control for state fixed effects and Year × Time-of-Publication fixed effects.

2.6.2 Fixed Costs

Table 2.4 report estimates of the entry and news agency subscription choice model.

The magnitude of the newspaper-specific components are similar to previous work.[18] I find that a newspaper's fixed cost increases as it is further away from the nearest bureau, even if it has no telegraphic news subscription. This is plausible since in the absence of telegraph news, it would have had to invest in other technologies

media (such as the radio, magazines, or competing newspapers), making their impressions less valuable to advertisers (Gentzkow, Shapiro, Yang, and Yurukoglu 2022). I am agnostic on the source of this interaction and am holding it fixed in the counterfactual.

[18] I report a $2.28 per-capita for fixed cost. Gentzkow et al. 2014 reports a $9.03 per-reader. The Inland Press survey reports mechanical fixed cost to be $7.73.

(such as hiring its own reporting staff) in order to enter. This is also consistent with per capita fixed costs being lowered when a paper picks up AP or UP.

		(1) Fixed Costs (per capita)
Newspaper-Specific	Inverse population (θ_0)	0.695
		(0.007)
	Constant (θ_1)	2.281
		(0.186)
	log(Distance to bureau / City population) (θ_2)	0.526
		(0.008)
News Agency-Specific	AP inverse population (θ_0^{AP})	-0.235
		(0.008)
	UP inverse population (θ_0^{UP})	-0.004
		(0.009)
	AP general costs (θ_1^{AP})	0.020
		(0.009)
	UP general costs (θ_1^{UP})	0.020
		(0.009)
	AP assessments (θ_2^{AP})	0.463
		(0.363)
	UP assessments (θ_2^{UP})	0.404
		(0.072)
	Cost multiplier from bundling ($\theta_{bundling}$)	1.124
		(0.206)
	Cost trend (θ_{trend})	1.10
		(0.010)
Calibrated parameters	Entry nesting parameter (σ^{entry})	0.75
	LLH	-39,400.94
	Number of market-years	14,121

Table 2.4: Estimates for the Supply Model

Note: This table presents estimates for the entry and news agency subscription choice model described by Equation 2.5. The population-independent constants are a component of cost that does not scale with population size (and thus in a per-capita specification is modeled as $\frac{x}{Pop}$). AP and UP general costs are fixed constants divided by the population covered by the AP and UP in each period (with the 1928 AP population normalized to 1). AP and UP assessments are predicted values estimated from Column 3 of Table 2.1, divided to be per-capita. The cost-trend is an exponential trend that captures secular changes in news agency costs in this period. The nesting parameter dictates the magnitude of the unobserved cost shocks in entry choice relative to news agency choices and is calibrated using ratio of fixed costs measured in the Inland Press data relative to AP assessments in 1926. The model is estimated for election years between 1912 and 1940. Standard errors do not account for uncertainty in the estimation of variable profits.

In order to interpret the news agency-specific costs, I aggregate the estimated components and plot the average across all markets in 2.4. From the revealed preference model, the average per-capita cost of subscribing to the AP in 1928 is \$0.28 (compared to \$0.27 in the assessment data). In 1912, the per-capita cost of news agency subscription is \$0.50 for the AP and \$2.26 for UP respectively. The significant gap and subsequent decrease underscores the importance of economies of scale in this setting: the cost of joining the AP and UP drops over time as the network increases in density.

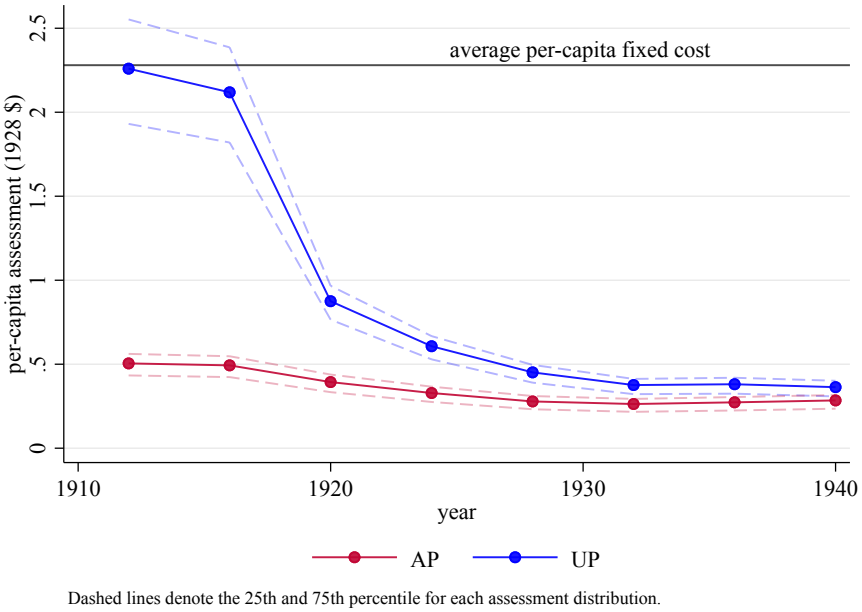


Figure 2.4: Graphical Interpretation of Supply Estimates

Note: This figure presents the implied news agency specific costs as calculated using estimates from Table 2.4. Costs are computed for *all* markets (regardless of observed subscriptions). The figure presents the average cost for a market to connect to the AP vs. UP. over time.

2.7 Model Simulations and Counterfactuals

2.7.1 Framework

To obtain counterfactuals, I fully solve the model and compute new equilibria where the realizations and beliefs regarding economies of scale and network effects could potentially change. In the counterfactual, a small change in membership sizes at the beginning can snowball as firms anticipate that future costs and values will adjust. The equilibrium outcome of the economy is a vector of entry probabilities such that implied network effects and economies of scale matches these entry probabilities. Finding this equilibrium amounts to finding a fixed point.

I solve the model on a set of constructed representative markets. Since local economies of scale are heterogeneous by distance to the nearest news source, I designated different markets into four types: Urban, Suburban, Remote, and Very Remote, corresponding to quartiles of distance to the nearest bureau. I further split each category markets by within-category population quartiles and by time of publication (Morning/Evening). This makes up $4 \times 4 \times 2 = 32$ representative markets. For each representative market, I calculate their exogenous market characteristics such as demographics and population as the average of all cities that fall into that category.[19]

I compute a self-fulfilled expectation equilibrium by solving a fixed point problem for these 32 markets. Formally, I find a set of choice probabilities $\Pr[W_{jmt}]$ such that the following three variables are consistent with each other: entry probabilities for each period, the network values and costs for each period implied by these entry probabilities, and movers' beliefs over them. For each guess of $\Pr\left[W_{jmt}^{(k)}\right]$ in iteration k, I compute the implied network values and costs, update entry probabilities to $\Pr\left[W_{jmt}^{(k+1)}\right]$, and iterate until $sup\left(\left\{\left|Pr\left[W_{jmt}^{(k+1)}\right] - Pr\left[W_{jmt}^{(k+1)}\right]\right|\right\}_t\right) < 0.0001$.

[19]For the set of real-life towns that are most similar to the constructed markets, please see the appendix.

Multiple equilibria are common in the network literature. I compute the fixed point problem by setting starting beliefs to the realized patterns in the data.[20] In order to compute news agency revenue fit, I take the per-capita formula estimated in Table 2.1 as given and sum up the payments to AP and UP across all markets.

2.7.2 Baseline Simulation and Model Fit

Figure 2.4 presents the equilibrium from the model simulation where the model starts from the installed base of 1916. The model approximates patterns in the data well. In particular, the patterns are close to the actual patterns of UP growth. The model overestimates AP's growth. This suggests that there can be costs to joining the AP that are not captured by my model.[21]

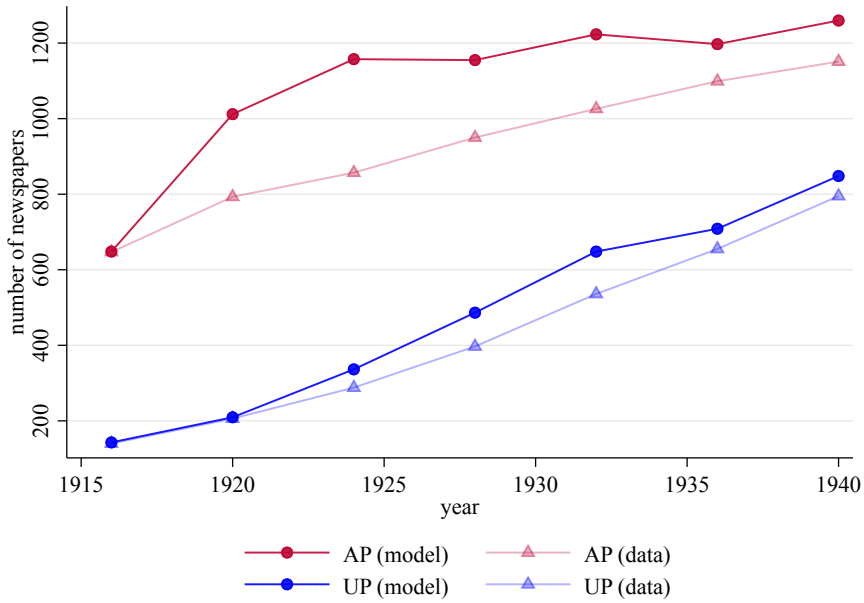


Figure 2.5: Model Fit

Note: This figure presents the model fit from the baseline model simulation described in Sections 8.1.and 8.2 and describes the number of newspapers as implied the model (dark circles) vs. the number of newspapers in the data (light triangles). The starting point for model simulation is the both firms' installed bases in the year 1916.

[20] Future versions of the paper will include an Appendix where I interrogate starting beliefs.

[21] One important candidate is that exclusive contracts beyond the city's immediate radius may have been informally enforced. Future versions of the paper will include robustness where I use the observed exclusive radii in 1900 rather than the city boundaries.

2.7.3 Counterfactual Networks: No Exclusive Contracts

Figure 2.7 presents the counterfactual membership path where there are no exclusive contracts. This is the main counterfactual considered by *AP v. US* (1945): would UP have been able to enter and compete with the AP if there were no exclusive territory contracts?

We observe a marked decline in UP memberships. In the counterfactual without AP exclusive territories, many papers opted to switch away from UP service and joined AP instead. This led to a substantial 30% decrease in UP's membership within the first four years, further deepening to 40% over eight years. Despite this initial loss, UP was still able to gain scale. This is due to its large installed base already established in 1916.

UP's revenue loss in this scenario quickly becomes devastating for the organization. As detailed in Section 2.1, UP was barely breaking even due to subsidies from the Scripps family, which offset its operational expenses. While the precise amount of the subsidy remains unclear, the counterfactual allows us to examine the impact of revenue loss. In this scenario, the cumulative revenue shortfall (illustrated as the shaded area in 2.6) quickly surpasses $1m (1928 dollars) in the first four years. This matches the highest potential scrap value discussed within the UP organization. Within eight years, this deficit nears $8m (1928 dollars), equivalent to the entire worth of E.W. Scripps' newspaper estate. Although I do not have an economic model for the Scripps' subsidizing process, these numbers indicate that no plausible subsidy would have sustained UP's break-even point without AP's exclusive territory agreements.

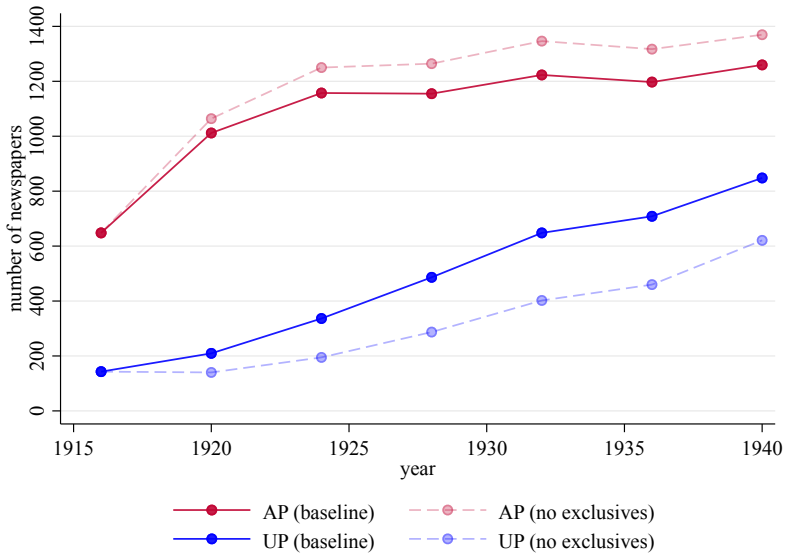


Figure 2.6: Counterfactual with no Exclusive Contracts

Note: This figure presents the counterfactual model simulation where there are no exclusive contracts. The figure represents the path of membership under the baseline and counterfactual.

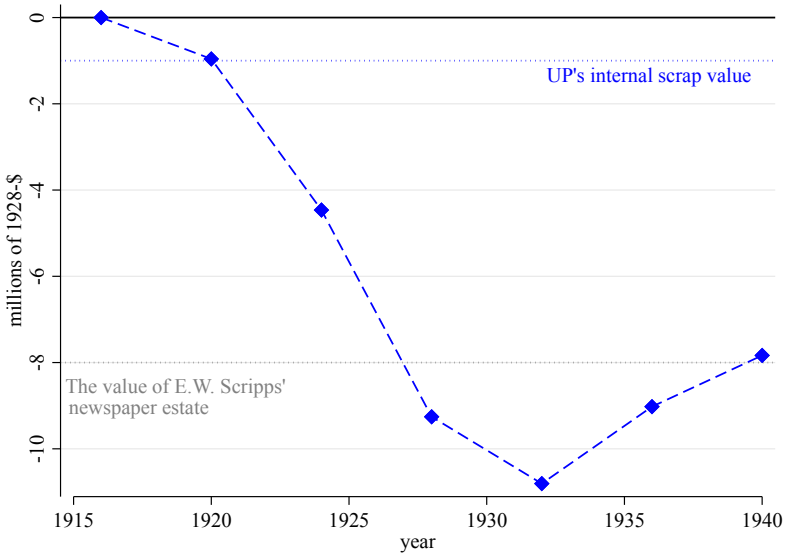


Figure 2.7: Cumulative United Press Revenue Loss

Note: This figure presents the the cumulative change in UP's revenue in a counterfactual without exclusive contracts relative to the baseline simulation in terms of 1928 dollars. The plot includes two horizontal lines that describes plausible scrap values for the UP organization. The top line, at negative \$1m, is the highest value internally discussed as the total valuation of the United Press stocks. The bottom line, at negative \$8m, is the total value of E. W. Scripps's newspaper estate as described in his will in 1921.

2.7.4 Counterfactual Networks: Exclusive Territories vs. Exclusive Dealing Contracts

Figure 2.8 and Table 2.5 presents news agency membership growth under counterfactual contracting regimes. I explore four different scenarios: the baseline with only exclusive territories in place, a counterfactual based on the *AP v. US* (1945) ruling where no exclusive contracts exist, another counterfactual where only exclusive dealing contracts are permitted (excluding exclusive territory contracts), and a final counterfactual where both types of contracts are allowed. These counterfactuals have historical significance: In 1915, the US Attorney General opined that the AP was entitled to exclusive territory contracts, but *not* to exclusive dealing. Internal documents reveal that AP's strategy was to limit UP's expansion by enforcing exclusive dealing clauses, thereby preventing their current members from accessing UP services.

In Table 2.5, I present results in terms of the number of newspaper-years that would have been part of either organizations under different counterfactual. I normalize memberships under the baseline (where there are exclusive territory contracts but no exclusive dealing contracts) to unity. Figure 2.8 shows the growth path.

My analysis indicates that the baseline scenario—with exclusive territories but without exclusive dealing contracts—is the most unfavorable for the dominant AP and the most favorable for the newcomer UP in terms of membership size. In this situation, UP attracts two groups of newspapers: those wanting to combine both AP and UP services, and those excluded from AP. Introducing exclusive dealing contracts from this baseline considerably undermines UP. However, the impact of exclusive dealing contracts on news agency memberships is much less pronounced when exclusive territories are not active, given that UP's reduced value and increased costs in this setting makes it less enticing to potential bundlers.

2.7.5 Counterfactual Newspaper Market Structures and Consumer Welfare

In Table 2.6, I show counterfactual newspaper market structures under counterfactual exclusive contracting regimes. Each figure denotes the fraction of total possible market years in the economy that corresponds to each market structure under various counterfactuals. I present four counterfactuals in addition to the baseline.

First, in the main counterfactual, UP exits in the absence of AP exclusive territories. Here, AP captures an additional 8% of the market-years, roughly half of what UP loses. The overall number of market-years receiving telegraphic news drop by 5%, indicating the reduced incentive for entry without exclusivity. These effects are also present in the second counterfactual where UP is subsidized to remain in the market, though the impacts are less pronounced.

In the third counterfactual, only exclusive dealing contracts are in effect. Here, UP suffers a large drop in market-years due to the loss of newspapers that bundled both services. Surprisingly, these contracts increase the AP's membership. This outcome seems counterintuitive, given that exclusive dealing contracts are designed to protect the upstream AP from competition at the cost of restricting downstream newspapers from bundling services. This unexpected result stems from an increase in AP's quality and a decrease in UP's in the counterfactual. This effect is clearer in the fourth scenario, where both exclusive dealing and exclusive territory contracts are in play. In this scenario, there's a slight uptick in markets featuring two newspapers compared to the baseline (an increase of 0.002% in market-years or 0.6% of the baseline). This minor increase is the downstream effect of exclusive dealing contracts: new newspaper entrants are encouraged to enter as incumbents are weakened by their inability to bundle various services.

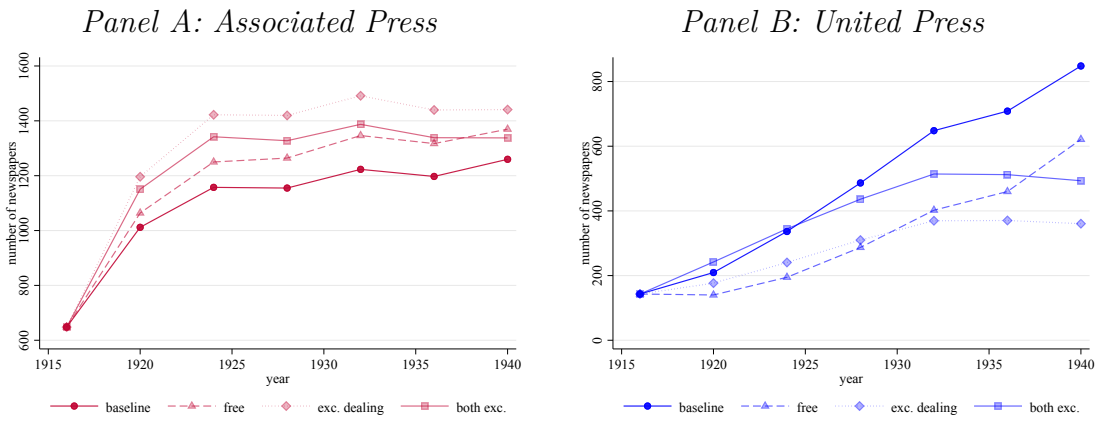


Figure 2.8: News Agency Membership under Counterfactuals

Note: This figure plots the growth rate of the membership of the Associated Press and United Press across counterfactual contracting regimes.

Associated Press		Exclusive Dealing	
		No	Yes
Exclusive Territories	No	1.08	1.17
	Yes	**1**	1.12

United Press		Exclusive Dealing	
		No	Yes
Exclusive Territories	No	0.65	0.63
	Yes	**1**	0.85

Table 2.5: Membership Size Under Exclusive Dealing Counterfactuals

Note: This table presents counterfactual news agency membership sizes (in market-years) under counterfactual exclusive contracting regimes. The top panel shows membership size for the Associated Press and the bottom panel shows that for United Press. For each panel, rows vary by whether exclusive territory contracts are enforced, and columns vary by whether exclusive dealing contracts are enforced. The bolded cell is the baseline observed in the real history and is normalized to 1.

Market structure		(1) Base	(2) UP-exit	(3) Free	(4) Deal	(5) BothExc
Number of newspapers	0	0.17	-0.02	+0.00	-0.02	-0.02
	1	0.56	+0.02	-0.00	+0.02	+0.02
	2	0.27	+0.00	-0.00	+0.00	+0.00
Number of AP papers	0	0.53	-0.09	-0.04	-0.10	-0.07
	1	0.47	+0.08	+0.04	+0.09	+0.07
	2	0.01	+0.01	+0.00	+0.01	-0.00
Number of UP papers	0	0.84	+0.14	+0.06	+0.07	+0.03
	1	0.16	-0.14	-0.06	-0.07	-0.03
Number of bundling papers	0	0.97	+0.03	+0.01	+0.03	+0.03
	1	0.03	-0.03	-0.01	-0.03	-0.03
Number of papers with telegraphic news	0	0.40	+0.03	+0.01	-0.06	-0.07
	1	0.58	-0.03	-0.01	+0.06	+0.07
	2	0.02	+0.00	+0.00	+0.00	-0.00

Table 2.6: Counterfactual Newspaper Market Structures

Note: This table presents newspaper market structures under counterfactual exclusive contracting regimes. Each row corresponds to a market structure, and each column corresponds to an exclusive contracting regime. Column 1 presents the fraction of total market-years that corresponds to the relevant market structure. Columns 2 to 5 present the difference in market-years that corresponds to the same market structure for each counterfactual. Columns 2, 3, 4, and 5 are for the counterfactual regimes where UP exit (UP-exit), with no exclusive contracts but UP stays in the market (Free), only exclusive dealing contracts but UP stays in the market (ExcDealing), and both exclusive territory and exclusive dealing contracts (BothExc).

Finally, to highlight the importance of equilibrium network adjustments, I compare the primary counterfactual to a scenario where UP does not exist even when exclusive territory contracts are in place. This counterfactual underscores the motivation for suppliers to adopt exclusive territory contracts in non-network industries (or when they deem rival entry unlikely). Under this scenario, the number of market-years with AP news decline by 12%. This reflects the reduced entry incentives for newspapers, as they lack protection from potential future competitors.

Market structure		(1) Exclusive	(2) Free Entry
Number of newspapers	0	0.20	-0.04
	1	0.71	-0.15
	2	0.09	+0.18
Number of AP papers	0	0.27	+0.12
	1	0.72	-0.12
	2	0.01	+0.00

Table 2.7: Counterfactual Newspaper Market Structures *without* United Press
Note: This table presents newspaper market structures under a counterfactual where United Press did not exist regardless of contracting regimes. Each row corresponds to a market structure, and each column corresponds to an exclusive contracting regime. Column 1 presents the fraction of total market-years that corresponds to the relevant market structure, when there are AP exclusive territory contracts but UP did not exists. Columns 2 presents the difference in market-years that corresponds to the same market structure for a counterfactual where there is no UP and no exclusive territory contract.

These results highlight the importance of the AP's organizational structure toward the observed market structure in news agencies. In particular, as a non-profit trade association controlled by its member newspapers, it is most likely that AP management was not attempting to maximize membership but rather the profit of its incumbent newspapers. In Figure 2.9, I plot the profit of newspapers who were already members of the AP in 1916 over time under counterfactual regimes relative to the baseline. In the counterfactual where there are no exclusive territory contracts, the threat of an immediate rival entrant with the AP mostly cancels out any gains these incumbents would have had from increased network value and decreased network cost from UP exiting. On the other hand, under a counterfactual where the AP can maintain its exclusive territory contracts *and* operates as a monopolist news agencies, the profit of member newspapers would have increased by approximately 10%. Yet, this scenario never realized due to UP's existence as AP's rival.

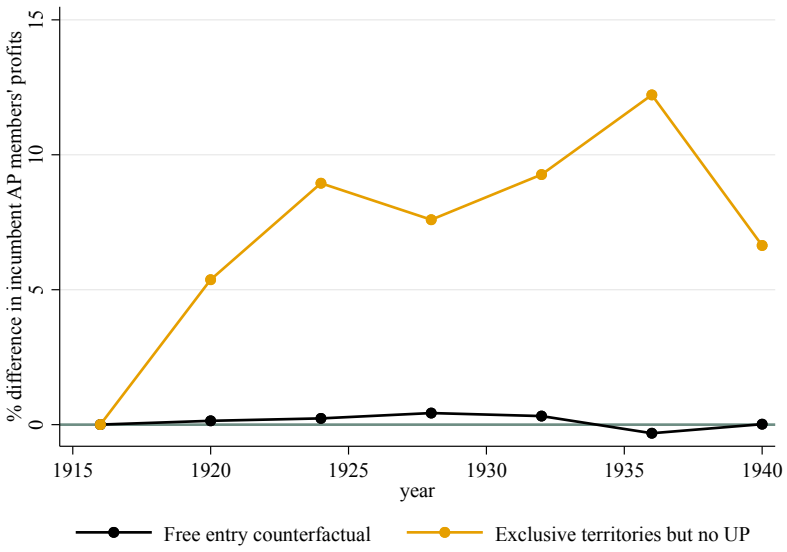


Figure 2.9: Incumbent Newspaper Profit

Note: This figure demonstrates the change in incumbent newspaper profit over time across counterfactual legal regimes relative to the baseline regime. Incumbent newspaper profit is calculated as the average total profit for newspapers who were Associated Press members as of the start of the study period. The green line denotes the baseline regime. The black line denotes a counterfactual where there are no exclusive contracts and United Press exits. The orange line denotes a counterfactual where exclusive territory contracts are in place, there are no exclusive dealing contracts, and United Press regardless exogenously exit at the beginning of the study period.

I further highlight the importance of accounting for equilibrium network competition and their effects on consumer surplus in Figure 2.10. In the counterfactual without exclusive territory contracts, consumer surplus increases by approximately 4%. This gain is the result of increased newspaper competition but is limited by the duplication of news agency fixed cost and the splitting of network effects across the AP and UP. Importantly, the counterfactual where the AP is a monopolist news agency with exclusive territory contracts represents the worst case scenario for consumers, with a decrease in surplus of approximately 12% over time relative to the baseline. Once again, the existence of upstream competition between the AP and UP significantly dampened harms to consumers from exclusive territory contracts.



Figure 2.10: Consumer Surplus

Note: This figure demonstrates the change in consumer surplus over time across counterfactual legal regimes relative to the baseline regime. Consumer surplus is calculated as the average population-weighted reader utility given their utility-maximizing choice across representative markets. The green line denotes the baseline regime. The black line denotes a counterfactual where there are no exclusive contracts and United Press exits. The orange line denotes a counterfactual where exclusive territory contracts are in place, there are no exclusive dealing contracts, and United Press regardless exogenously exit at the beginning of the study period.

2.8 Conclusions

I analyze exclusive vertical contracts in network industries, focusing on early 20th-century U.S. news agencies. I find that while the Associated Press (AP) aimed to establish local newspaper monopolies with exclusive territory contracts, it unintentionally aided the growth of their primary rival, United Press (UP). I develop a structural model of news agencies and estimate it using historical data. I identified strong natural barriers to entry for news agencies due to economies of scale and network effects. Counterfactual simulations show that exclusive territories protected downstream newspaper monopolies but incentivized upstream news agency competition. Exclusive dealing contracts have the opposite effects.

Appendix A: Notation

Use	Symbol	Meaning
Indexing	m,M	markets and the set of markets
	t, T	periods and the last period
	j, J	newspapers and the set of newspapers
	i, I	households and the set of households
	w	news agencies
Reader Demand	u	reader utility
	p	annual newspaper price
	W	bundle of news agency subscriptions
	$g(W)$	value of news agency subscriptions to readers
	X	market-level demographics
	α	price coefficient
	γ	demographics coefficients
	$\overline{\xi}$	time-invariant unobserved newspaper quality
	ξ	time-variant unobserved newspaper quality
	σ	nesting coefficient for newspaper demand
	ϵ	idiosyncratic reader preferences
	β	news agency value coefficients
	s	newspaper market share
Advertising Demand	a	per-reader advertising revenue
	λ	advertising revenue coefficients
	ι	idiosyncratic newspaper advertising value
Pricing	v	per-period newspaper variable profit
	MC	newspaper marginal costs
Entry and Subscription	Υ	mover selection rule for incumbents
	$\overline{W}$	set of news agency subscriptions available to the mover
	τ	annual payments from newspaper to news agencies
	δ	discount factor
	Ω	market states
	η	idiosyncratic choice logit shock
	ρ	nesting coefficient for entry choices

Table 2.8: Model notation

Online Appendix A: United Press Assessments

In this section, I discuss UP's method of calculating assessments on newspapers. My knowledge of the UP's comes primarily from two documents.

The first document includes minutes and exhibits from the "United Press Rate Setting Committee" held in November 1912, hereby referred to as *Minutes*. The committee was appointed by the Scripps newspapers and the Board of UP in order to propose a method for calculating assessments on UP members broadly and Scripps newspapers specifically. At this time, Scripps newspapers contribute over 24% of total receipts of United Press.

The second document is a memo prepared in 1913, titled "System for the Determination of a Proper Basis of Charges for United Press Clients," hereby referred to as *Memo*. This memo was created as a result of the rate setting committee, as the original committee struggled to decide on a cost apportioning method given "the impracticability of adopting any of the [present suggestions]."

The assumption that UP calculates the cost of an additional member in a similar fashion to that of the AP is supported by the following statements:

- The formula for calculating the cost of an additional member depends on a similar method as the AP. Overhead charges (total news collection by the organization and administrative cost) is divided across clients "irrespective of geographical location" . The cost of the trunk line is also divided independently of geographical location as "all clients are [...] dependent upon, and should pay their faire (sic) share of the cost of maintaining a trunk line of communication from Coast to Coast." The cost of the local circuit is divided by members in that circuit: "all wire and operating cost, exclusive of Trunk Line Charges, should be divided equally among the Clients in each zone." (Page 2 of *Memo*)

The assumption that UP does not engage in strategic pricing is supported by the following statements:

- UP did not expect to lose money initially on new members: "[Board member] Rickey asked what percentage of profit the UP should retain, ultimately, as a "goal." Mr. Howard thought 15% and saw no reason why it should not. [...] The profit-making feature [of the UP] was an asset in getting clients." (Page 6 of *Minutes*)

- UP did not price discriminate by local market structure. The organization also used "rate cards" based on the market's exogenous characteristics: "It shall be left to the management of the UP and the management of each individual group to determine a card rate for the United Press service to possible non-Scripps clients in each individual city of that [region]. (Exhibit F of *Minutes*).

- UP was unable to turn a profit at the beginning or even in the long run. Scripps said, as quoted in Morris 1957: "[...] in founding the new UP, the Scripps papers actually paid out several hundred thousand dollars more than they would have paid out had they been members of the AP. However, notwithstanding these comparatively large payments, I felt that the investment, bad as it was financially in one way, was a good one beacuse it secures freedom from the temptation of one huge monopoly." At the time of the merger with INS, a UP executive told *Editor and Publisher* that "Both agencies operated under low profit margins for many years."

Online Appendix B: United Press Scrap Value and Subsidies

In this section, I discuss the evidence surrounding the scrap value of the United Press in 1912 and 1913. Here, I describe a series of letter in Roy Howard's Archive:

First, in a letter dated August 15, 1912 to James G. Scripps (E.W.'s son), Howard asks to discuss a "proper valuation of the United Press property." Howard then wrote to J. C. Harper (general counsel for the E. W. Scripps Company) dated August 9, 1912, asking "to know just what in your opinion this United Press stock is worth." In this letter, Howard mentioned that it would be "absurd to regard the United Press" as worth $700,000.

In a reply to Howard dated August 13, 1912 by Harper, he stated that UP "has no tangible property" and alludes to UP being a going concern purely because of a large subsidy by the Scripps organization:

> If certain Scripps' papers paid to the UP on the basis that non-Scripps papers pay, or on the basis that their competitors pay the AP, how much of UP profit would be left? [...] You know that E.W. Scripps and James G. Scripps both recognize that the payment that certain papers made and are now making to UP is excessive. [...] I submit, therefore, that [...] UP is a subsidized concern.

He countered that even $600,000 would be too high that he "doubt[s] whether [he] would want to pay that figure." Given these discussed figures, I use $600,000 (1912 dollars, approximately $1m in 1928) as the scrap value for UP in 1912.

Since United Press was heavily subsidized, it is plausible that the Scripps family could increase the subsidy to cover the fixed cost of the organization in an event of membership loss. A letter dictated April 26, 1920 by E. W. Scripps to T. L. Sidlo (general counsel for the Scripps media conglomerate starting 1916) states that E. W.'s

newspaper estate is inventoried at $8m and that of his sister Ellen B. Scripps' is at $4m. I use this figure as the maximum plausible additional subsidy by the Scripps family into United Press.

Aguirregabiria, V., A. Collard-Wexler, and S. P. Ryan (2021): "Dynamic games in empirical industrial organization," in *Handbook of Industrial Organization*, ed. by K. Ho, A. Hortaçsu, and A. Lizzeri, Elsevier, vol. 4 of *Handbook of industrial organization*, 225–343.

Asker, J. (2016): "Diagnosing Foreclosure due to Exclusive Dealing," *J. Ind. Econ.*, 64, 375–410.

Benkard, C. L. (2004): "A Dynamic Analysis of the Market for Wide-Bodied Commercial Aircraft," *Rev. Econ. Stud.*, 71, 581–611.

Cabral, L. (2011): "Dynamic Price Competition with Network Effects," *Rev. Econ. Stud.*, 78, 83–111.

Chandra, A. (2009): "Targeted advertising: The role of subscriber characteristics in media markets: Subscriber characteristics and targeted advertising," *J. Ind. Econ.*, 57, 58–84.

Evans, D. and R. Schmalensee (2014): *The Antitrust Analysis of Multisided Platform Businesses*, Oxford University Press.

Fudenberg, D. and J. Tirole (2000): "Pricing a Network Good to Deter Entry," *J. Ind. Econ.*, 48, 373–390.

——— (2014): "Competition and Ideological Diversity: Historical Evidence from US Newspapers," *Am. Econ. Rev.*, 104, 3073–3114.

Gentzkow, M., J. M. Shapiro, F. Yang, and A. Yurukoglu (2022): "Pricing Power in Advertising Markets: Theory and Evidence," .

Guarneri, J. (2017): *Newsprint Metropolis: City Papers and the Making of Modern Americans*, University of Chicago Press.

Ho, K. (2006): "The welfare effects of restricted hospital choice in the US medical care market," *J. Appl. Econ.*, 21, 1039–1079.

Ho, K. and R. S. Lee (2019): "Equilibrium Provider Networks: Bargaining and Exclusion in Health Care Markets," *Am. Econ. Rev.*, 109, 473–522.

Igami, M. and K. Uetake (2019): "Mergers, Innovation, and Entry-Exit Dynamics: Consolidation of the Hard Disk Drive Industry, 1996–2016," *Rev. Econ. Stud.*, 87, 2672–2702.

Jeon, J. (2022): "Learning and investment under demand uncertainty in container shipping," *Rand J. Econ.*, 53, 226–259.

Jullien, B., A. Pavan, and M. Rysman (2021): "Two-sided markets, pricing, and network effects," in *Handbook of Industrial Organization*, ed. by K. Ho, A. Hortaçsu, and A. Lizzeri, Elsevier, vol. 4 of *Handbook of industrial organization*, 485–592.

Lafontaine, F. and F. Scott Morton (2010): "Markets: State Franchise Laws, Dealer Terminations, and the Auto Crisis," *J. Econ. Perspect.*, 24, 233–250.

Lafontaine, F. and M. Slade (2008): "Exclusive contracts and vertical restraints: Empirical evidence and public policy," in *Handbook of Antitrust Economics*, ed. by P. Buccirossi, The MIT Press.

——— (2014): "Franchising and exclusive distribution," *The Oxford Handbook of International Antitrust Economics*, 2, 387.

Lee, A. M. (1937): *The daily newspaper in America: the evolution of a social instrument*.

Lee, R. S. (2013): "Vertical Integration and Exclusivity in Platform and Two-Sided Markets," *Am. Econ. Rev.*, 103, 2960–3000.

Lee, R. S., M. D. Whinston, and A. Yurukoglu (2021): "Structural empirical analysis of contracting in vertical markets," in *Handbook of Industrial Organization*, ed. by K. Ho, A. Hortaçsu, and A. Lizzeri, Elsevier, vol. 4, 673–742.

Nevo, A. (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69, 307–342.

Rochet, J.-C. and J. Tirole (2002): "Cooperation among Competitors: Some Economics of Payment Card Associations," *Rand J. Econ.*, 33, 549–570.

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999–1033.

Rysman, M. (2004): "Competition between Networks: A Study of the Market for Yellow Pages," *Rev. Econ. Stud.*, 71, 483–512.

Sinkinson, M. (2020): "Pricing and Entry Incentives with Exclusive Contracts: Evidence from Smartphones," .

Spulber, D. and C. Yoo (2014): *Antitrust, the internet, and the economics of networks*, Oxford University Press.

Tiew, A. (2022): "Flailing Firms and Joint Operating Agreements: An Application to US Local Daily Print Newspapers from 1932 to 1992," .

Chapter 3

Urban Migration, Public Health Amenities, and Local Newspapers in 1870–1940 U.S.

3.1 Introduction[1]

Potential migrants likely do not have all relevant information about each possible destination. This was particularly true for most of history, when search and information acquisition was relatively costly. This friction suggests a potentially important role for the news media in historical migration: learning about the characteristics of other places from the news may change one's beliefs about those places and subsequently the decision of whether and where to migrate.

We study the effects of local newspapers on urbanization via the provision of information to rural migrants about urban amenities. We focus on the U.S. between

[1]This paper is co-authored with *Anais Galdin*. We are grateful to Leah Boustan, Janet Currie, Kate Ho, Eduardo Morales and Steve Redding for their continual guidance and support on this project. We also thank Nick Buchholz, Jacob Dorn, Gene Grossman, Jakub Kastl, Ilyana Kuziemko, Alessandro Lizzeri, Chris Neilson and Tianyi Wang for helpful discussions and comments. This research benefited from financial support from the International Economics Section (IES) at Princeton University, whose support is gratefully acknowledged. All errors are our own.

1870 and 1940 for two reasons. First, this was a period of rapid urbanization in the U.S., when the percentage of the population living in urban areas increased from 22% to 55%. Second, local newspapers were by far the primary available news media in our setting, as TV had not yet been popularized and radio rarely engaged in news programming. This setting provides us ample variation in exposure to local news and in urban migration.

The urban amenity of interest to our study is the public health conditions in a city, in particular the incidence of waterborne diseases. Waterborne diseases were an important concern for urban life at the time due to their high death rates, severe symptoms, and their ability to disrupt life and cause massive income loss (Anderson et al., 2022; Beach, 2022b). In 1900, diseases such as typhoid, dysentery, and diarrhea were responsible for nearly 200 deaths per 100,000 residents in the U.S., with most deaths happening during mass outbreaks. Our study period also coincides with a wave of investments in public sanitation projects aimed at eradicating these diseases, such as a water filtration plant or sewage system. Sanitation projects mostly eliminated outbreaks, greatly decreased the incidence of disease, and reduced urban mortality. In this paper, we explore the possibility that rural migrants who learned of improved urban conditions due to these projects from local newspapers could have been induced to move there.

We micro-found the migration decision of a individual using a discrete-choice framework inspired by Berry (1994b). To evaluate the importance of different urban characteristics, we explicitly specify the utility of the individual across potential locations to include characteristics such as wages, sanitation, and disease rates. We allow individuals'salience about characteristics to vary depending on whether they have access to a local newspaper. Our empirical implementation includes origin-time and destination-time fixed effects and so uses bilateral variation across origin-destination pairs to identify the importance of characteristics and information access. Specifically,

we identify the role of information in migrants'decision by comparing migration rates across destinations with different characteristics *within* an origin at a given time, after controlling for unobservable destination characteristics that are constant across time.

We apply this framework to historical migration flows in the 1870-1940 U.S., using full-count U.S. censuses from IPUMS (Ruggles et al., 2024a,b). We use record linkage techniques from Abramitzky et al. (2020) to match people over time in the decennial census data. Our data on public health conditions include typhoid death rates from Beach (2022b) and the implementation of water filtration plants as well as other public health intervention such as water chlorination, tuberculosis testing, and sewage systems from Ao (2015) and Anderson et al. (2022). To proxy for the availability of information to rural migrants, we use the United States Newspaper Panel from Gentzkow et al. (2011b) to detect the presence of a newspaper. We directly augment this data with keyword searches on the Chronicling America database from Dell et al. (2023) to measure instances of news coverage of waterborne diseases and changes in public health infrastructure.

First, we present evidence linking the significance of sanitation systems, news coverage of urban health, and rural-urban migration. Utilizing our data on migration flows, we support earlier findings that migration is primarily local (Eckert and Peters, 2018), noting the average mover relocated just 9 miles from their origin. We also find that a significant driver of urbanization during this period was rural-urban migration, rather than direct growth in urban areas. We then demonstrate that a wave of investments in sanitation systems coincided with urbanization between 1870 and 1940. Together, these results indicate that rural migrants'access to information about urban destinations'sanitation systems is a plausible factor in their migration decisions.

We further provide evidence that water filtration meaningfully changed how the news cover an urban destination's quality of life. An informal audit of our news corpus reveals that direct coverage of water filtration systems that reflects improvement in

urban amenity was minimal. On the other hand, these improvements were reflected in news coverage via a large drop in the amount of news coverage of waterborne diseases. In an event study, we find that water filtration systems significantly decreased the number of mentions of typhoid related to an urban destination. We repeat this exercise for a set of non-waterborne diseases and find that, conversely, these diseases received increased coverage after the implementation of water filtration. This suggests that newspapers shifted their coverage away from waterborne diseases after water filtration.

We then implement our primary empirical model, motivated from our discrete choice model of an individual's migration decision. We find that on average, the implementation of a water filtration plant at an urban destination increases the share of residents from any origin moving to that destination by 25%, while the joint presence of a newspaper at an origin and a filtration plant at the destination increases this share by an additional 74%. This result supports our hypothesis that while water filtration and public health reflect an important amenity for rural-urban migrants, migrants with access to newspapers at their local origin are potentially more salient to this amenity change and thus more likely to move there.

We augment our primary results with two additional empirical specifications: instead of using the implementation of a water filtration plant as a proxy for urban amenity, we use first, a direct measure of destination typhoid death rate and second, a measure of the intensity of news coverage about typhoid at an urban destination. Since migration toward an urban destination can also cause increased news coverage, we use the variation in disease news predicted by the implementation of a water filtration system for our preferred specification. We find that on average, a standard deviation increase in typhoid mortality rate leads to a lower in-migration rate for a destination and even lower from origins with newspaper access, although this result is not statistically significant in our preferred specification. On the other hand,

we find that individuals strongly respond to typhoid news coverage: an increase of one standard deviation leads to a 61% lower migration share from an origin with newspapers.

Throughout our empirical results, we study the bilateral variation in migration between origin-destination pairs in response to bilateral variation in destination amenities and origin newspaper presence, after controlling for origin-time and destination-time fixed effects. This helps us overcome potential concerns about unobserved push factors that are common across destinations (or pull factors that are common across origins). However, concerns about pair-specific unobservables remain. For example, if origins that have newspapers and destinations that have water filtration systems are more similar in characteristics, they might experience higher migration rates. We partially alleviate these concerns by including pair-level congruence measures of observable demographics in the Census. Data limitations mean that potentially unobservable congruence among characteristics that are not captured by Census demographics can still bias our results.

Another important concern is that we cannot rule out other urban attributes that may have changed at the same time as public health investments. These include changes in productivity, job opportunities, or leisure amenities at a destination city, to which individuals could be responding rather than to public health conditions directly. This concern limits our ability to make claims about the precise mechanisms through which public health and information interact to increase urbanization. Nonetheless, our results indicate that public health was an important component of urban amenities and that migrants'response to amenity changes depends on access to information through local newspapers.

Related Literature Our paper builds on a literature that found relatively low migration response to local shocks, suggesting the existence of significant barriers to migration (Porcher, 2020 WP). A set of recent papers have found evidence for

information frictions to be a major cause (e.g. Wilson (2022) on migration responses to fracking in the U.S., Porcher (2020 WP) on broadband internet in Brazil). In addition, we show that information frictions hinder the extent to which individuals can migrate in response to changes in amenities. While we do not provide a model of endogenous amenity investment in our setting (e.g. Almagro and Lino (2022 WP) or Diamond (2016)), our results suggest that incorporating information can be a valuable addition for future work in that direction.

Our work also relates to a literature on historical urbanization in the U.S. (see Boustan et al. (2018) for a survey) and that on the importance of sanitation systems and public health (Beach, 2022b; Anderson et al., 2022; Cutler and Miller, 2005). We show that the implementation of a sanitation system in a city likely induced in-migration from rural areas and that this effect is heterogeneous depending on the information environment experienced by potential migrants. Finally, we contribute to a literature that studies the role of historical newspaper content in influencing societal outcomes. In particular, we corroborate the findings of Costa and Kahn (2017) that the historical news media focused more on "negative" news of diseases rather than "positive" news of decreased death rate and improved conditions in the city. We extend this finding by showing that the migration decisions of individuals in fact respond to the absence of disease news, suggesting that individuals can rationally update on the conditions of the city despite the improvement not being directly covered.

Our paper proceeds as followed. Section 2 describes the data. Section 3 describes the historical background for our study. Section 4 describes our theoretical and empirical model and a discussion of our identification strategy. Section 5 describes our results and robustness tests. Section 6 concludes.

3.2 Data and Descriptives: Sanitation Systems and Urbanization

In this paper, we measure the salience of health-related events in local newspapers and capture migration responses to health-related destination amenities across different information environments. We first construct a novel dataset that leverages full-count census data in order to build historical migration flows across U.S. counties and cities. We then pair rural-urban migration flows with data on public health investment and typhoid death rates in cities, as well as a large corpus of digitized local newspaper text from 1870 to 1940.

3.2.1 Rural-Urban Migration Flows

This section first describes how we constructed bilateral migration flows from historical U.S. censuses, before drawing key statistics related to rural-urban migrations around the turn of the 20th century.

Building Migration Flows from Historical U.S. Censuses

In order to create a measure of bilateral migration flows within the U.S., we use historical full-count decennial censuses for years 1870 to 1940 from the Integrated Public Use Microdata Series (IPUMS).[2] Each digitized full-count census provides geocoded individual-level data, including details on education, race, income, occupation, family, health and residence. For the period 1850-1940, the full-count censuses contain over 650 million individual-level records.

[2] At the time of this study, the latest full-count census being published for the U.S. is for the year 1950, as full-count censuses have to wait 72 years before becoming publicly available. The 1960s one should be released in 2032.

Census Linking. In order to build intra-U.S. migration flows, we rely on the Census Linking Project (Abramitzky et al., 2020, CLP)[3] to link individuals between adjacent pair of complete-count census years. Matches between adjacent census years are made based on an individual's first and last name, age and state/country of birth. In order to limit false positives in migration events, our main results exploit the most conservative matching algorithm from the CLP (ABE-exact), which requires a unique and exact match of name and age within a 5-year age band.[4]

Our results are robust to the use of the use of the three alternative automated matching methods from the CLP. Depending on the specific matching algorithm, the linked samples used in our analysis include between 10 and 20% of the total U.S. population compared to the full-count census.[5] The final sample contains 617,483,404 unique matched individuals over the 1870-1940 period.

Bilateral Migration Flows. We measure migration flows from an origin o to a destination d for an intercensal period (10 years) as the share of individuals living in o at the beginning of the period that we find in location d at the end of the period. Our primary outcome for migration flows is thus equivalent to computing the empirical probability that a person moves from an origin location o to a destination d during an intercensal period t, allowing for people to stay put:

$$m_{odt} = \Pr[\text{individual moves from } o \text{ to } d \text{ in period } t]$$

[3] `https://censuslinkingproject.org` provides a series of fully anonymous crosswalks between each pair of censuses.

[4] The approach is described in more detail in Abramitzky et al. (2021).

[5] The sample only includes male individuals due to the nature of the data, as tracking females through adjacent census decades is complicated by changes of family names post marriage. The Census Linking Project thus only provides links for men on the site and our our analysis is consistent with the literature using the linked census data. It moreover does not allow to capture international migration flows.

For instance, if o is Mercer county and d is Trenton city in New Jersey, and t captures the intercensal period 1900-1910, m_{odt} captures the share of Mercer county residents in 1900 found in Trenton in 1910.

As our measure of migration is constructed from moves across linked full-count U.S. censuses, it is limited to intra-U.S. migrants, and the set of origin locations o only captures U.S.-based locations and do not include international migrants. The data, however, captures the birth location of individuals and classifies them into foreign-born, native born, first-generation and native born. We use these individual-specific observables as controls in our regression framework.

Rural-Urban Migrants. We define a rural-urban migrant as an individual who moved from a rural county to an IPUMS location identified as a "city".[6] Figure 3.1 plots the distribution of the population in the full-count U.S. census for different type of geographical units (county, city, urban area). The accompany table to Figure 3.1 details the share of all individuals from the full-count census living in an area that is not identified as a city. This share decreases from around 80% of the population in 1870 to 50% in 1940.

Figure 3.2 compares our newly created measure of urban migration flows to the share of population that is urban in the matched full-count census data. The urban population and the urban migration shares are closely related, both in terms of magnitude and time trend, which supports the adequacy of our urban migration measure.

[6]IPUMS defines a city based on its population: a place is referred to as a city census if it has more than 10,000 contemporary inhabitants in the 1870 and 1880 censuses, and more than 25,000 contemporary inhabitants from 1900 to 1940. Our results are qualitatively similar if we use the less restrictive IPUMS definition of an "urban" county. We prefer the city definition as it allows us to generates a fixed and comparable set of cities across our sample.

Year	Share of Rural Pop.
1870	0.814
1880	0.784
1900	0.671
1910	0.630
1920	0.576
1930	0.525
1940	0.543

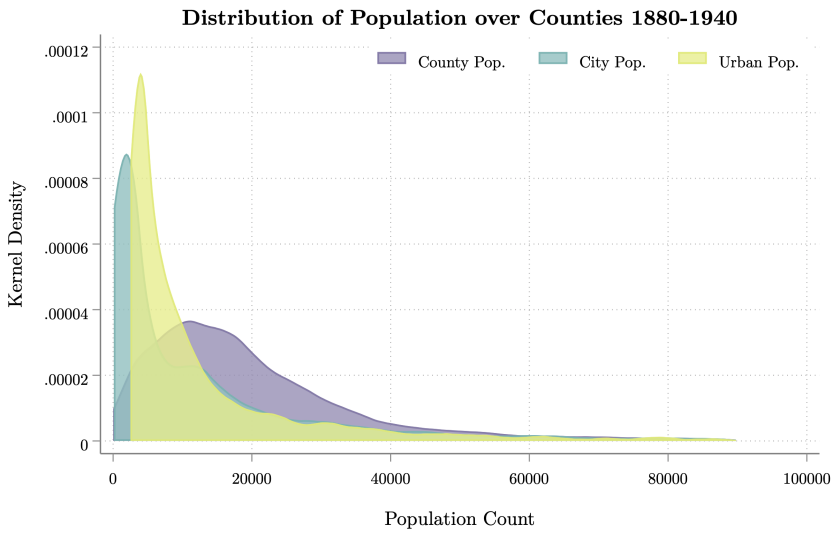


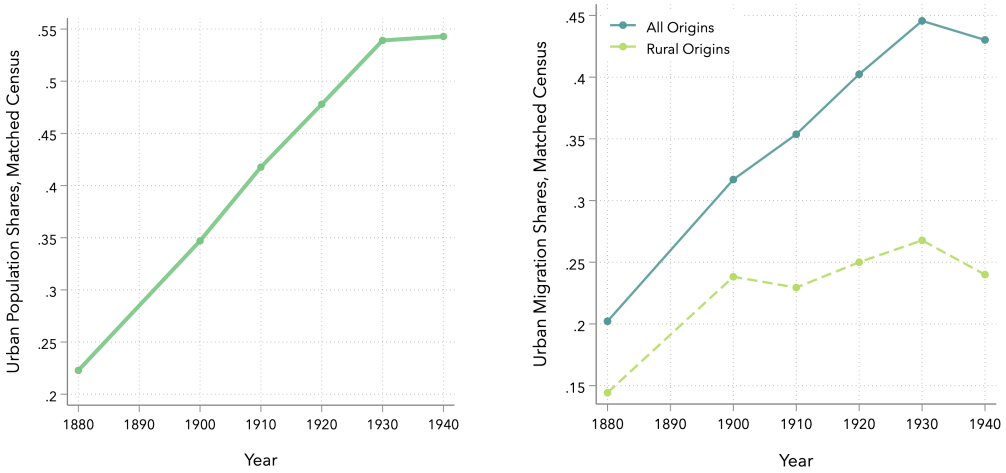
Figure 3.1: **Population by location type in the full-count census data**

Notes: Figure 3.1 is based on the unlinked version of the full-count U.S. census data from IPUMS. A "city" is (mostly) defined as any location with 10,000+ inhabitants. "Urban" areas denote all cities and incorporated places of 2,500+ inhabitants. All areas not classified as urban are designated rural. Individuals living in smaller, non-urban places are assigned to a state, a county, and a "non-identifiable city" code. The left table details the share of IPUMS Population living in a "non-city" area.

Statistics on Migration Flows

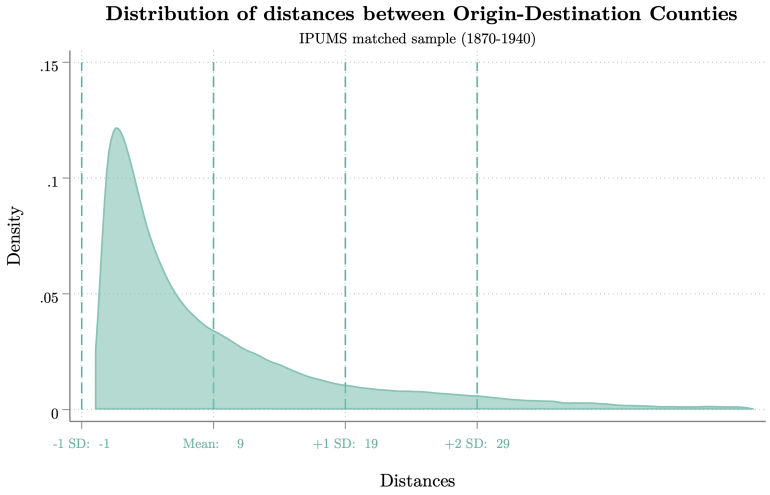
Migration is a local phenomenon. Acquiring information about different potential destinations has historically been costly. It is thus without surprises that migrations in the U.S. primarily remained a local phenomenon between 1870 and 1940 (Figure 3.3). The average mover in our linked census dataset only moved 9 miles away from their origin location between two census years, with a standard deviation of 10 miles. This is consistent with the findings of Eckert and Peters (2018) who finds that American industrialization was also highly local, and suggests that access to information may be an important barrier to migration and integration of labor forces.

Figure 3.2: Urbanization is closely tied to urban migration



Notes: The left figure plots the share of the entire U.S. population living in urban areas, as defined by IPUMS. The right figure plots the share of urban migrants in our data, splitting migrants by type of origin location. Both measures of urban population shares and urban migrations are computed based on the linked version of the full-count U.S. census, using the Census Linking Project. "Urban" areas denote all cities and incorporated places of 2,500+ inhabitants. All areas not classified as urban are designated rural. Before 1950, the urban share only includes residents living in incorporated places. The slight flattening of the trend from 1930 to 1940 may be due to the rise of unemployment in cities during the Great Depression. Looking at urban population shares over longer time series (see e.g. Boustan et al. (2018)) shows that the upward trend continues after 1940 and this flattening is temporary.

Figure 3.3: Migration is mostly a local phenomenon



Notes: Figure 3.3 reports the distribution of distance between origin and destination counties/cities for individuals identified as movers in the matched census data. The average mover in the linked census only moved 9 miles away from their origin location between two census years, with a standard deviation of 10 miles. Intra-U.S. migrations primarily remained a local phenomenon between 1870 and 1940.

Share of Urban Movers. Migration flows to cities accelerated over the turn of the 20th century (Table 3.1). Most individuals who moved before 1920 where going from rural areas to urban areas, or from non-cities to cities. Starting around 1920-1930, as more people live in urban counties, this trend reverses and migration flows become mostly urban-to-urban and city-to-city. An important caveat is that since we do not have international migrants in our data, these figures may be understating the extent of urbanization in this period and international arrivals'contribution to that process.

Table 3.1: Number and Share of Movers to Urban and City Areas, by Year

Year	Urban Areas	Rural to Urban	Urban to Urban	City Areas	Non-City to City	City to City
			Total Number of Movers (Volumes)			
1880	227554	123403	104151	178351	113305	65046
1900	549116	360296	188820	446615	316655	129960
1910	928953	480752	448201	718322	416061	302261
1920	1.420e+06	669103	751136	1.138e+06	639412	499087
1930	1.979e+06	867559	1.111e+06	1.635e+06	872126	762589
1940	1.753e+06	661979	1.091e+06	1.372e+06	636419	735502
			Share of Movers (Among all Movers)			
1880	0.253	0.137	0.116	0.198	0.126	0.0723
1900	0.396	0.260	0.136	0.322	0.228	0.0938
1910	0.437	0.226	0.211	0.338	0.196	0.142
1920	0.512	0.241	0.271	0.411	0.231	0.180
1930	0.572	0.251	0.321	0.473	0.252	0.221
1940	0.521	0.197	0.324	0.407	0.189	0.218
			Share of Movers (Among the whole IPUMS Population)			
1880	0.0868	0.0470	0.0397	0.0680	0.0432	0.0248
1900	0.176	0.116	0.0606	0.143	0.102	0.0417
1910	0.146	0.0756	0.0705	0.113	0.0654	0.0475
1920	0.169	0.0797	0.0894	0.136	0.0761	0.0594
1930	0.186	0.0817	0.105	0.154	0.0821	0.0718
1940	0.140	0.0527	0.0868	0.109	0.0507	0.0585

Notes: The second (middle) part of the table reports the share of all movers that goes from a given type of origin to a given type of destination (e.g.: total number of movers from rural to urban areas in census year t divided by the total number of movers in year t, for all possible origins-destinations). The share of movers going to urban areas increase over the period, and while the early 20th century migrations were mostly from rural to urban areas, migrations become increasingly urban-to-urban starting around 1920-1930. The third (bottom) part of the reports the share of movers as a percentage of the whole U.S. population from census (for instance, total number of movers to cities in year t divided by the total U.S. population in year t).

3.2.2 Public Health Investments and Waterborne Diseases

In this subsection, we present the historical background that motivates our empirical analysis. First, we describe the state of urban public health in the U.S. and investments in public health infrastructure during our sample period. Second, we describe our public health investment and waterborne disease data. Third, we describe the relationship between urbanization and public health investments.

Waterborne Diseases in the 1870-1840

For most of American history, cities were deadly places. Important killers of urban dwellers include waterborne diseases which spread via water contaminated with bacteria or parasites such as typhoid, cholera, and dysentery. In 1900, these diseases combined caused approximately 186 deaths per 100,000 inhabitants in the U.S. (Troesken, 2001).[7]

These diseases were salient for reasons beyond their high death rates. For example, consider typhoid which accounted for roughly 2.5% of total mortality during this period (Cutler and Miller, 2005). Infection typically occurs via drinking water tainted with infected feces. Symptoms include muscle ache, diarrhea, severe fever, delirium, internal hemorrhage, and increase the risks of pneumonia and tuberculosis during the disease. The brutality of the symptoms is intensified by their duration: the disease's active period can last from three to four weeks, but the recovery period can be as long as four months. Reinfections were common and there was no effective treatment or vaccine. Furthermore, epidemics were common (Troesken, 1999). Given the severity, length, and frequency of the disease, typhoid was often a signpost for the filthy environmental conditions of the city.

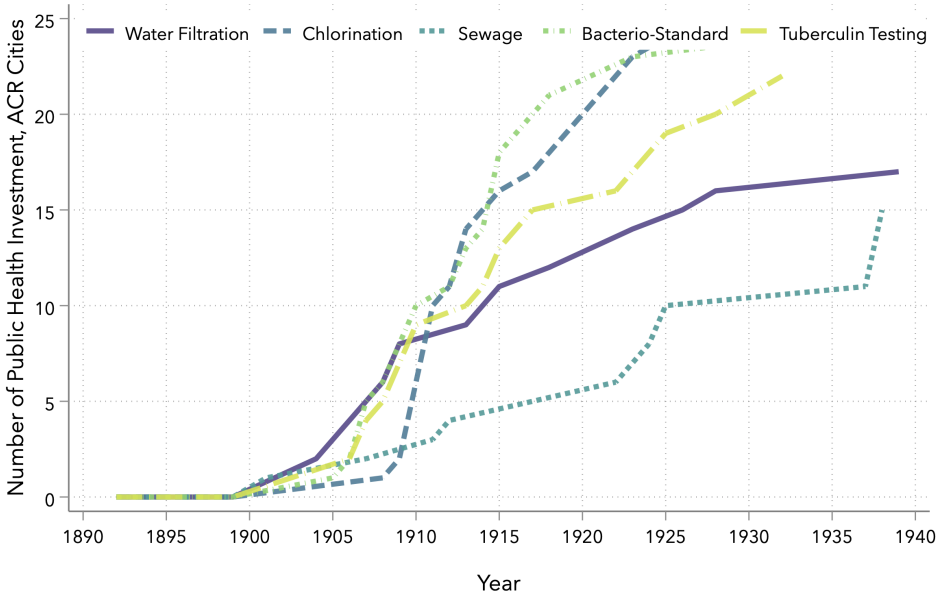
Public Health Developments in the 1870-1840

As of the late 19th century, public health officials understood that cleaner water, e.g. via water filtration or chlorination, could eradicate typhoid and other waterborne diseases (Troesken, 1999). This led to a wave of investments in these structures around the turn of the 20th century in the United States (Figure 3.4). There is an active debate in the academic literature on just how much of the decrease in mortality

[7]For comparison, COVID-19 caused roughly 200 deaths per 100,000 inhabitants in the U.S. between February 2020 to November 2021. (CDC 2021)

in the U.S. during this period was due to these investments, with estimates ranging from 20 to 50%. (Cutler and Miller, 2005; Beach, 2022b; Anderson et al., 2022).

Figure 3.4: Public Health Investment over the beginning of the 20th century



Notes: This figure reports the cumulative sum of U.S. cities adopting public health investment (PHI) systems in the U.S. The subset is limited to the 25 cities studied by Anderson et al. (2022), splitting the data by type of PHI. The figure emphasizes that most of the changes occurred during the first 20 years of the 20th century. Nearly all cities in the study had adopted bacterio-standard, chlorination and tuberculin testing by 1940. The timing of water filtration and sewage systems was more idiosyncratic, with almost half of major U.S. cities not having implemented such systems by 1940.

Several key public health developments, combined with advancements in microbiology and epidemiology, helped significantly reduce the incidence of waterborne diseases like typhoid or cholera in the United States during the 1870-1940 period. The introduction of water filtration systems was one of the most significant step in reducing waterborne diseases. In the late 19th century, cities began implementing sand filtration systems, which helped remove bacteria and other contaminants from the water supply. For example, in 1872, Poughkeepsie, New York, became the first U.S. city to use a sand filtration system for its public water supply. Chlorination also proved to be an effective method for disinfecting water and reducing the spread of

waterborne diseases like typhoid fever and cholera. In 1908, Jersey City, New Jersey, became the first U.S. city to use chlorination for water treatment. The development of piped water systems, such as the Croton Aqueduct in New York City (completed in 1842), played another crucial role in providing clean water to urban populations. These systems helped reduce the reliance on contaminated water sources, such as wells and rivers, which were often breeding grounds for waterborne diseases. Finally, the implementation of sewage treatment systems was another important factor in reducing waterborne diseases. By treating sewage before releasing it into water bodies, cities could reduce the contamination of drinking water sources. In the early 20th century, cities began adopting activated sludge systems and other advanced sewage treatment methods. For a literature review and history of the politics of water infrastructures in the U.S., starting with early waterworks and later sanitation efforts, see Beach (2022a).

Timing of PHI Adoption. The literature on public health argues that the timing of adoption of these water filtration systems in U.S. cities is *exogenous* to the health status of the city, due to political, financial, and legal situations (Anderson et al., 2022; Cutler and Miller, 2005). The average central water filtration system costed one to two times the average municipal budget to implement. The link between waterborne diseases and contaminated water was also not widely acknowledge before the end of the 19th century. For most cities, water filtration systems thus took a long time to plan and yet could suddenly be implemented, so that among all potential candidate cities for public health investment, the timing of installation is exogenous to health outcomes.

Due to these frictions in adoption, it is difficult to generalize a geographic pattern according to which cities adopted water filtration systems and, more generally, clean water supply. Previous work has identified factors that leads to higher probability

adoption including segregation (Beach et al., 2022) and privatization of waterworks (Troesken, 1999). While a city's lack of access to alternative, untainted water sources in principle could drive filtration efforts, but cities also often filtered clean water sources for amenity values (such as taste, Beach (2022a)). The lack of a clear geographic pattern of adoption further motivates us to use destination fixed-effects in our analysis as we cannot provide plausible geographic control variables for why one city would have invested in clean water infrastructure over another.

Data on Public Health Investment and Typhoid Outbreaks

Public Health Investments Data. We use two main sets of data to measure public health investments in the United States over the period 1870-194. The first dataset comes from Anderson et al. (2022), which provides several public health investments measures for 25 U.S. cities. These investments include water filtration, water chlorination, clean water projects such as aqueducts, milk standards (such as tuberculin testing for cows and bacteriological standards), and sewage systems.

As this initial dataset is limited to 25 U.S. cities, we complement the dataset with data related to water filtration systems from Ao (2015). This dataset records the year of installation of these systems for 74 cities located across 30 U.S. states. The creation of the dataset involves compiling data from historical surveys and censuses of filtration plants published by American Water Works Association as well as supplementary textbooks on water filtration practices. The sample of the cities was conditional on having a population of over 30,000 in 1920, having information on water filtration plant installation dates available, and having existed in the 1880 census.

Table 3.2 describes these two data sources across counties as well as the characteristics of those counties, using covariates from ICPSR. By the end of the 1870-1940 period, 59 cities out of the 74 had installed a filtration system. This represents

approximately 1.5% of our county-year observations.

County-level Covariates. We complement the data on public health investment with historical county-level covariates from the Inter-university Consortium for Political and Social Research (ICPSR, Haines and Inter-university Consortium for Political and Social Research (2010)). The covariates include the population share that is white, male, foreign-born, urban, working in the manufacturing sector, literate, and the log manufacturing output per capita.

Table 3.2: **Descriptive Statistics: Public Health Investments**

	Mean	SD	Description
Public Health Investment			
Sewage Treatment/Diversion	0.024	0.152	= 1 if municipality had a sewage treatment plant or diverted sewage away from drinking water supply, = 0 otherwise
Filtration	0.015	0.123	= 1 if municipality had a water filtration plant, = 0 otherwise
Chlorination	0.033	0.179	= 1 if municipality chemically treated water supply, = 0 otherwise
Clean Water Project	0.022	0.146	= 1 if municipality had completed a clean water project, = 0 otherwise
Bacteriological Standard	0.024	0.154	= 1 if municipality set bacteriological standard for milk supply, = 0 otherwise
TB Test	0.023	0.151	= 1 if municipality required tuberculin testing of cows, = 0 otherwise
County-level Covariates			
Diff. in % Foreign, Dest.	-0.267	0.120	Difference in share of County population foreign born between year t and t-4
Diff. in % Manufacturing, Dest.	-0.210	0.204	Difference in share manufacturing in the County between year t and t-4
Diff. in % Male, Dest.	-0.163	0.128	Difference in share of male population in the County between year t and t-4
Diff. in % White, Dest.	-0.084	0.171	Difference in share of white population in the County between year t and t-4
Diff. in % Literacy, Dest.	-0.515	0.124	Difference in share of literate population in the County between year t and t-4
County × Year	4122317		

Typhoid data. Our data on typhoid epidemics come from Brian Beach (Beach, 2022b; Beach et al., 2016). Typhoid rates are measured as deaths per thousand persons. The typhoid rate measures is a consistent yearly series (1880-1930) that combines information from Whipple with information from the mortality statistics. The data is at the city level, for 76 U.S. cities. Figure 3.5 reports the evolution of typhoid rates over time and space.

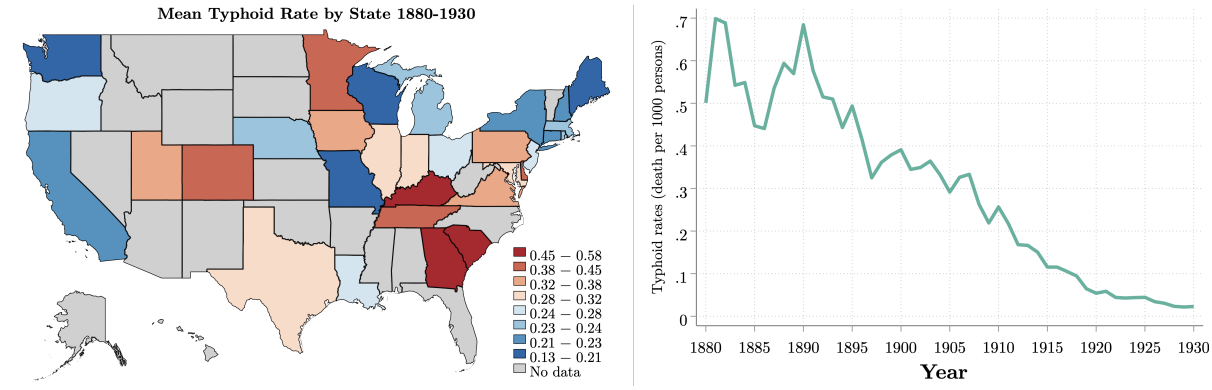


Figure 3.5: **Typhoid death rates over time and space**

Notes: These figures plot the evolution of typhoid death rates over time and space. The map plots the average typhoid rates per state over all years in the sample (1880-1930). The typhoid rates are defined as deaths per 1000 persons. The measure of typhoid rates is a consistent and balanced series (1880-1930) that combines information from Whipple with the mortality statistics. The data were obtained from Brian Beach (Beach, 2022b; Beach et al., 2016).

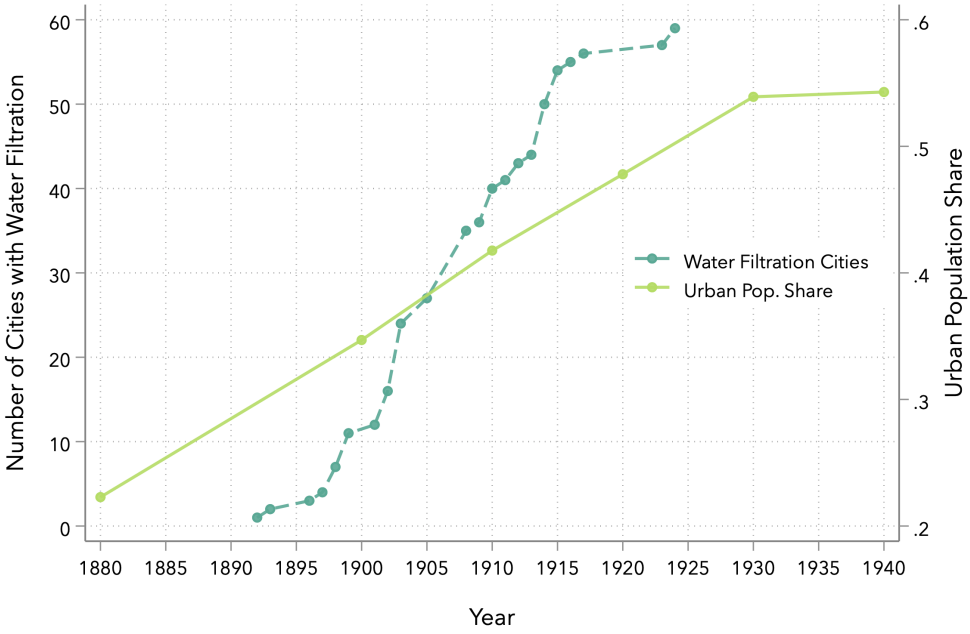
Sanitation Systems and Urbanization

Given the deadliness of the urban conditions, eradicating waterborne diseases was an important goal for cities at the time as they grew. Cities' investment in clean water works closely tracked the rate of urbanization (Beach, 2022b). Municipal ownership of water companies and public investment into filtration systems were viewed as vote-winners by political candidates (Troesken, 1999). Furthermore, industrialization brought in migrants from the rural area which did not have any immunity to typhoid, increased population density, and was considered a contributor to typhoid epidemics (Wohl, 1983). Thus, investments into water infrastructure and urbanization were closely linked at the time.

We show that investment in water filtration systems closely tracked urbanization. In Figure 3.6, we plot the number of cities with water filtration systems from Ao (2015) data against the share of population living in urban counties, which we computed from the U.S. decennial censuses. While the first water filtration systems were only being installed in the 1890s, by 1920, nearly 60 of the cities in our sample had

an operational filtration plant. This closely tracks the linear increase in urbanization in the U.S. from 1900 to 1930, which tapered off between 1930 and 1940.

Figure 3.6: Urbanization & public health coincided



Notes: This figure reports the number of cities with water filtration systems from Ao (2015) data (in dark green) against the share of population living in urban counties (in light green), which we computed from the U.S. census. We find that the timing of urbanization and the implementation of water filtration systems coincided.

3.3 Newspaper Coverage of Urban Public Health

In this section, we document our primary data source of historical local newspapers for the U.S.. We then present descriptive evidence related to newspaper coverage of urban public health.

3.3.1 Newspaper Text Corpus

Our primary data source for the presence of local newspapers is the United States Newspaper Panel from Gentzkow et al. (2011b). The dataset contains detailed characteristics for newspapers for each election year in the US, including readership

size, subscription prices, posted advertising prices, political affiliations, and dates of entry and exit. We use this dataset to construct a measure for the availability of a local daily newspaper for each county in the US during our study period. There were approximately 300 counties in the U.S. with at least one newspaper in 1880 and more than 1,000 of such counties by 1940. Detailed summary statistics of this dataset are available in Gentzkow et al. (2011b). To match the decennial census data level of observation, we aggregate the United States Newspaper Panel data at the decennial level. For each U.S. county and census year, we compute the mean number of newspaper dailies, and the net change in number of newspapers (capturing entry/exit).[8]

Our data on newspaper coverage comes from the Chronicling America database compiled by the Library of Congress (Dell et al., 2023). To construct a measure of newspaper coverage of urban public health, we use a simple keyword search on the Chronicling America database. For each newspaper page, we detect a "hit" when a phrase related to public health appears near the name of one of our sample cities on the page.

We perform this exercise for two sets of public health phrases. First, we look for mentions of specific diseases. Among waterborne diseases, we include "typhoid," "dysentery," and "diarrhea." We also include a set of "placebo" diseases for which we do not expect to be affected by investments into clean water, which include "plague," "influenza," and "tuberculosis." For mentions of public health infrastructures, we look for mentions of a "water filtration" project.

Our measure of newspaper coverage related to public health is the total count of hits for each phrase in any given year divided by the total number of mentions for

[8]We focus on the presence of a newspaper headquarter in a county as our measure for information access. Newspapers could circulate outside of their headquarter county via carriers and mail. Gentzkow et al. (2014) find that in 1924, home-market papers are responsible for 90% of circulation.

any city that year. This gives us the *share* of mentions of a city that also pertains to the public health topic of interest.[9]

3.3.2 Capturing Local News about Urban Public Health

The turn of the 20th century witnessed radical changes in the American press. Newspapers became more independent from political parties and more informative (Petrova, 2011). Increased reach of newspapers also led to an increase in both in-depth investigative coverage as well as sensational journalism (Gentzkow et al., 2004). Our informal audit of the newspaper content suggests that news about public health can fit both categories.

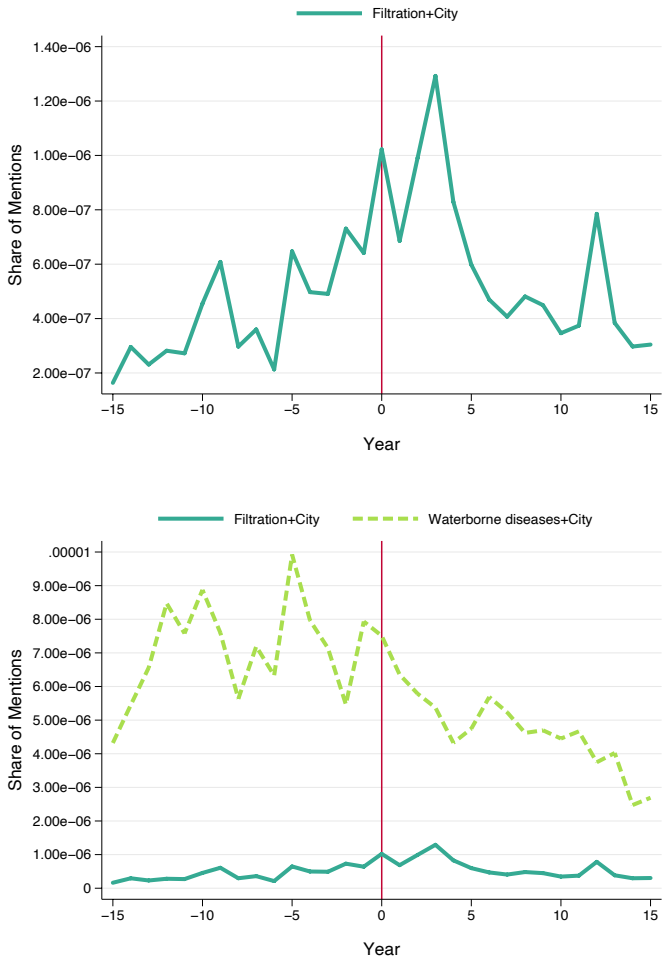
First, we find that the vast majority of news coverage related to water filtration projects revolves around their funding. Since these projects were massive municipal projects at the time, there were often political debates and referenda (e.g. on whether to issue municipal bonds) as well as sensational stories regarding government corruption in implementing these projects. Mentions were not always regarding coverage in the newspaper's own city as there were also discussion of example projects in other cities. Perhaps surprisingly, there was very little coverage celebrating the completion of a project or to advertise the improvement in public health as an amenity.

Second, we find that coverage surrounding diseases at the time broadly fit two types. The first type delivered hard information about local conditions such as case counts, death counts and obituaries. The second type of coverage was more sensational, which included features of urban epidemics and stories of local residents who have a given disease. We note that the majority of disease news was negative and focused on increases in death rates rather than decreases, consistent with Costa and Kahn (2017).

[9]Here, we measure intensity of news coverage related to a place, as salience likely increases with intensity.For example, if there are 1,000 pages across the database that mention 'Milwaukee' in 1900 and there are 5 pages that have 'Milwaukee' and 'typhoid' in proximity of each other, we compute this share to be 0.005.

In Figure 3.7, we report the average share of newspaper articles mentioning either water filtration (top) or waterborne diseases (bottom) in association with a city in our dataset around the time of implementation of a water filtration system. We find that the number of mentions of diseases vastly swamped that of the implementation of a water filtration system. Coverage of water filtration peaks around the time of implementation and shortly after. On the other hand, coverage of waterborne diseases sharply decline after the implementation of filtration systems, consistent with evidence showing that death rates from these diseases also declined after water filtration (Alsan and Goldin, 2019; Anderson et al., 2022; Beach, 2022b). Disaggregating the second series across waterborne diseases. In Section 5, we show these results formally within a difference-in-differences regression alongside that for placebo diseases.

Figure 3.7: Capturing News Coverage



Notes: These graphs report the share of newspaper articles mentioning either water filtration (top) or waterborne diseases (bottom) that are associated to a city in our dataset. The sample is currently limited to the 74 cities in our public health investment dataset, some of these cities not having an implemented water filtration system over the period. In both cases, we compute the number of keyword hits having both a city name and either filtration or typhoid phrases appearing next to each other in an article. Each graph is centered around the year of implementation of a water filtration system in the city. In the top graph, news related to these sanitation events peak around the time of the public health investment. In the bottom graph, we superpose the share of mentions to waterborne diseases (including typhoid, dysentry, and diarrhea) for these cities. The timing of water filtration investments correspond with a decrease in reports of diseases outbreaks in the newspapers. These negative health shocks are also more likely to peak in the news than positive news about health investment.

3.3.3 Change in coverage with public health infrastructure

In this section, we present evidence that news coverage of diseases changed at the same time of the implementation of a water filtration system. We do so by regressing the

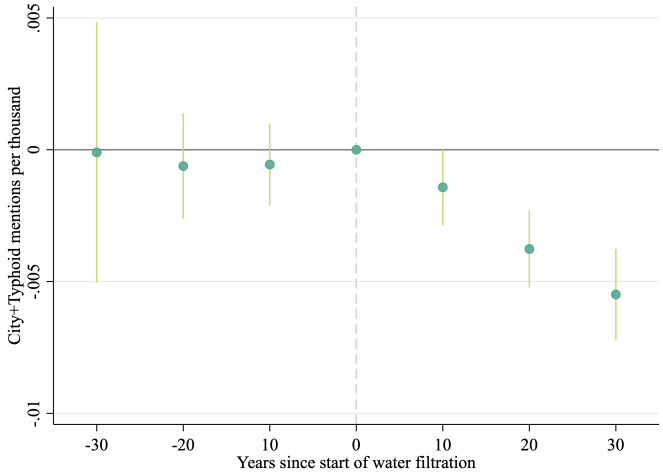
normalized number of hits for "typhoid" linked to a city name per thousand articles (averaged over a decade) on time (in decades) elapsed since the implementation of a water filtration plant. Formally, we perform the following regression

$$Hits_{dt} = \sum_{t=-3,\ldots,0,\ldots 3} \phi_t Filtration_{dt} + \kappa_d + \kappa_t + \eta_{dt},$$

where t denotes decades since the filtration system was put into place and κ denotes city and decade fixed effects.

In figure 3.8, we plot ϕ_t for the left-hand-side variable of hits of "typhoid" linked to the city name . We find that mentions of typhoid decrease after the city implemented a water filtration system, consistent with the disease being gradually eradicated and leading to less news coverage overall (F statistic $> 1,000$). The effect is gradual and pronounced, consistent with works showing a similar trend effect in the eradication of the disease itself. (Beach, 2022b; Alsan and Goldin, 2019).

Figure 3.8: Effect of water filtration system on typhoid news coverage

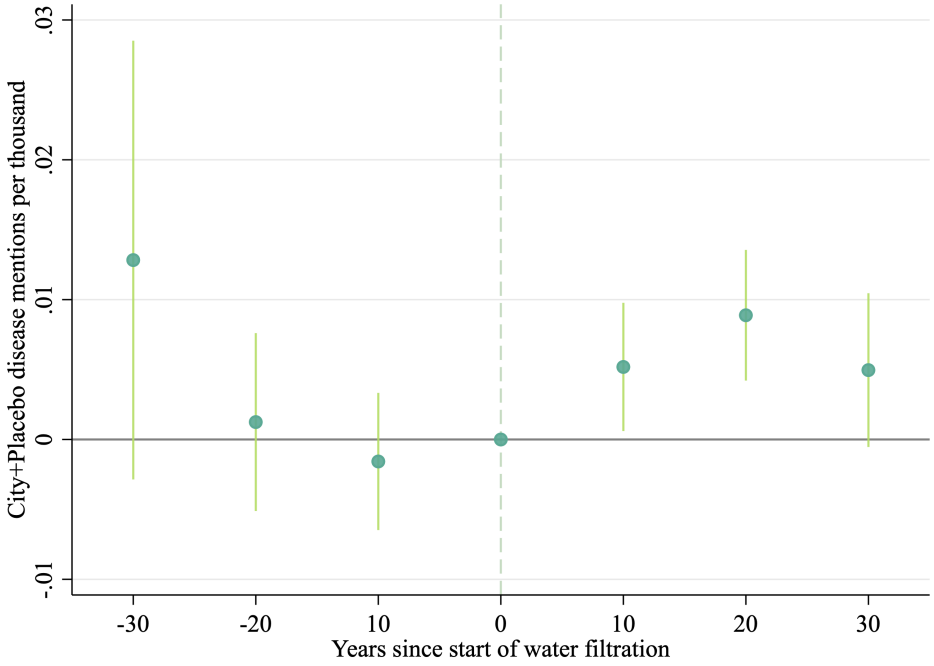


Notes: This figure presents the changes in newspaper coverage of typhoid in response to the implementation of a water filtration system, estimated via an event study specification after controlling for place and time fixed effects. The y-axis denotes the number of hits for a destination city and "typhoid" on a newspaper page divided by the total number of hits for a destination city during that time period. The x-axis denotes years since the implementation of the city's water filtration system, grouped into decades.

Figure 3.8 highlights that the implementation of a water filtration system had direct effects on news coverage of disease. While we cannot rule out that migrants'potential responses could be due to other factors that accompany the implementation of a water filtration system (e.g. job opportunities, leisure amenities, or productivity), here we established that water filtration caused meaningful changes in news coverage of urban public health.

To establish that this was not a coincidence or due to secular improvement in urban life quality or the media production function, we repeat this analysis for a set of diseases likely unaffected by the implementation of a water filtration plant: plague, influenza and tuberculosis. Figure 3.9 reports the resulting event study graphs. The pronounced decrease in coverage of waterborne diseases reported in Figure 3.8 is replaced by an increase in coverage of placebo diseases. This pattern can be consistent with two different hypotheses. First, newspapers, having allocated a fixed amount of space for health and disease coverage, could have shifted coverage to another type of disease. Second, as waterborne diseases became less of a concern, readers demand more news related to other diseases for practical purposes.

Figure 3.9: Effect of water filtration system on placebo disease news coverage



Notes: This figure presents the changes in newspaper coverage of tuberculosis, influenza, and plague in response to the implementation of a water filtration system, estimated via an event study specification after controlling for place and time fixed effects. The y-axis denotes the number of hits for a destination city and tuberculosis, influenza, and plague on a newspaper page divided by the total number of hits for a destination city during that time period. The x-axis denotes years since the implementation of the city's water filtration system, grouped into decades.

3.4 Model of Migration Choices and Information Acquisitions about Destination Amenities

This section presents our model. First, we describe a simple theoretical framework of location choice that emphasizes the role of information. Second, we use the theoretical framework to motivate our empirical exercises. The goal of the model is to make clear the identification assumptions for our empirical exercises, interpret their results and discuss their limitations.

3.4.1 Theoretical Framework

This section builds a discrete choice model of individual location decisions that emphasizes the role of information acquisition and micro-founds our estimating equations.

Set up: Players and Destination-Specific Amenities

Let $t \in \{1, \ldots, T\}$ index periods in time and $\{1, \ldots, D\}$ be the set of geographical divisions. For each period t, a individual i's location at the beginning of the period is referred to as its *origin* $o \in \{1, \ldots, D\}$ and its location at the end of the period is referred to as its *destination* $d \in \{1, \ldots, D\}$. We say that a individual i has migrated during the 10-year intercensal period t if $d \neq o$.

Let H_{dt} denote a vector of characteristics of destination location d at the end of the period t that is relevant to the individual and observable by the researcher. For expositional purposes, this subsection assumes that H_{dt} is a scalar (e.g. the typhoid mortality rate at a destination). The individual may not know H_{dt} and instead have a perception $\tilde{H}_{dt}$ of what H_{dt} may be.

Individual's Perception of Location-specific Amenities

We make several assumptions on the individual's *perception* of the characteristics of different locations.

First, individuals have perfect information on the characteristics of the place they currently live in, so that $\tilde{H}_{ot} = H_{ot}$. Second, there are two information regimes possible for each origin location, captured by the indicator C_{ot} for *access to local information*. $C_{ot} = 1$ indicates that a resident at the origin would have had access to information about potential destinations. Finally, the perceived location characteristic for a destination $d \neq o$ depends on access to information in the origin location

and can be written as

$$\tilde{H}_{dt} = \begin{cases} \overline{H}_{dt} & \text{if } C_{ot} = 0 \\ \rho H_{dt} + (1-\rho)\,\overline{H}_{dt} & \text{if } C_{ot} = 1 \end{cases}$$

where $\overline{H}_{dt}$ is the prior belief about the characteristic of destination d common to individuals across all origins o. $\overline{H}_{dt}$ thus captures a population-weighted average belief about d over all origin locations o. The parameter ρ captures the extent of updating to the real value H_{dt} for individuals living in an origin location with more information (i.e. for individuals living in location such that $C_{ot} = 1$).[10]

Note that this framework nests the strong assumption of perfect updating (e.g. $\rho = 1$ implies that $\tilde{H}_{dt} = H_{dt}$) or completely discarding the information ($\rho = 0$ implies that $\tilde{H}_{dt} = \overline{H}_{dt}$).

Individual's Static Utility from Moving

Suppose a potential migrant i from origin location o makes a one-shot, static decision about whether to move to location d during the 10-years period t.[11] We specify the static, myopic location decision of a individual i over a ten-year periods using the following utility specification for each possible destination d:[12]

$$d_{iot} = \arg \max_{d' \in \{1,\ldots,D\}} u_{iodt}$$

[10] An implicit assumption is that the prior for a given destination is common across origins.

[11] Under this setting, u_{iodt} is not a continuation value. In a dynamic model of migration, locations may have option values that capture the potential future benefits of moving to a location, and influence migration decisions. Our model does not incorporate option values and a dynamic model may yield different insights. We believe these concerns are attenuated by the long 10 years time-frame of the migration decisions in our historical context. The static, one-shot decision of the individuals thus spread a 10-years period. Future research could investigate an extension of this model to a dynamic setting with option values in the spirit of Caliendo et al. (2019).

[12] While a myopic model may not capture all aspects of migration decisions, it can still provide valuable insights into the role of information and other factors in shaping migration patterns over these decennial periods.

$$u_{iodt} = \begin{cases} \tilde{\beta}\left(\tilde{H}_{dt} - \tilde{H}_{ot}\right) + X'_{odt}\alpha + \xi_{dt} - \xi_{ot} + \xi_{odt} + \eta_{iodt} & d \neq o \\ 0 & d = o \end{cases}$$

We are assuming that the individual's outside option is to stay put (i.e. to choose $d = o$) and we normalize the utility from this outside option to zero. As a result of this normalization, preferences are interpreted as relative to the outside option of not moving. $\tilde{\beta}$ denotes the moving individual's preference over the perceived difference in destination and origin characteristic, $\tilde{H}_{dt} - \tilde{H}_{ot}$. X_{odt} is a vector of migration costs between o and d for intercensal period t. ξ_{dt} is a shifter of utility for destination location d at time t, ξ_{ot} captures the disutility from moving away from location o in period t, and ξ_{odt} is an o-d-pair level shifter of utility. η_{iodt} is an idiosyncratic shock to individual i. The individual observes ξ_{dt}, ξ_{ot}, ξ_{odt} and η_{iodt}, while the researcher does not.

We can rewrite the utility specification as following, accounting for the presence of information C_{ot},

$$u_{iodt} = \begin{cases} \underbrace{\tilde{\beta}\left[\rho\frac{\left(H_{dt} - \overline{H}_{dt}\right)}{H_{dt}}\right] C_{ot}H_{dt} + X'_{odt}\alpha + \left(\tilde{\beta}\overline{H}_{dt} + \xi_{dt}\right) - \left(\tilde{\beta}H_{ot} + \xi_{ot}\right) + \xi_{odt}}_{\delta_{odt}} + \eta_{iodt} & d \neq o \\ 0 & d = o \end{cases} \tag{3.1}$$

where δ_{odt} denotes the mean utility of destination d for individuals from origin o in period t.

Assuming that the error term η_{iodt} is identically and independently distributed across destinations and individuals and it follows an extreme value distribution, we can write the predicted share of individuals from o that migrate to d in period t as a

function of the mean utility δ_{odt},

$$s_{odt}\left(\vec{\delta}\right) = \frac{\exp\left(\delta_{odt}\right)}{1+\sum_{d'\neq o}\exp\left(\delta_{od't}\right)} \quad \text{and} \quad s_{oot}\left(\vec{\delta}\right) = \frac{1}{1+\sum_{d'\neq o}\exp\left(\delta_{od't}\right)} \tag{3.2}$$

and obtain the familiar "market share regression" as in Berry (1994b):

$$\log\left(s_{odt}\right) - \log\left(s_{oot}\right) = \delta_{odt}.$$

3.4.2 Empirical model

Our main estimating equation is directly micro-founded by the discrete choice logit model (3.1) and market share equation (3.2),

$$\log\left(s_{odtg}\right) - \log\left(s_{ootg}\right) = \beta C_{ot} \times H_{dt} + \underbrace{\alpha_1 \log\left(\text{dist}_{\text{od}}\right) + \alpha_2 \ln\left(\tau_{\text{odt}}\right)}_{\text{migration costs}} + \alpha_w w_{gdt}$$
$$+ \underbrace{X'_{odt}\alpha + \gamma_{dt} - \gamma_{otg}}_{\text{location-specific cov.}} + \gamma_{\text{occup}} + \varepsilon_{odgt}. \tag{3.3}$$

In this equation, o indexes origins, d indexes destinations, t index time periods, and g index demographic groups (age, race and literacy). The left-hand side captures the log probability of moving from o to d by the end of the intercensal period t, for individuals belonging to group g. C_{ot} captures access to information from local newspapers and is a proxy for information available at the origin. In practice, C_{ot} captures as the fraction of period t during which origin location o has a local daily newspaper. H_{dt} is a destination characteristic of interest. In our setting, it will be related to health investments and either measure (i) the fraction of time during period t that destination d has implemented a water filtration system, or (ii) mortality due to typhoid outbreaks. The interaction term $C_{ot} \times H_{dt}$ represents preference over public health amenities, while the coefficients α_1 and α_2 capture moving costs.

The term X_{odt} is a vector of origin-destination-time specific covariates and γ_{occup} is a work occupation fixed effect. ε_{odtg} is a regression error term. The two γ terms represent preferences over various location characteristics: γ_{dt} and γ_{otg} are destination-time and origin-time-demographic group fixed effects, respectively. The origin-year-demographic group fixed effect ensures that we are comparing two destinations d and d' for a given origin o, while adding a destination-year fixed effect ensures that what is left in the unobserved term is changes in moving costs from o to d over time. The γ_{dt} fixed-effect thus control for any unobserved factors that might vary across destinations and over time, but remain constant within each destination-year combination.

By incorporating both sets of origin-demographic group-year and destination-year fixed effects, we are effectively controlling for unobserved factors that could be correlated with the interaction term $C_{ot} \times H_{dt}$, both at the original location-demographic group level and the destination level, as well as any time-specific shocks that affect all groups or destinations equally. Under this setting, identification comes solely from bilateral variation in the interaction term $C_{ot} \times H_{dt}$. The residual variation in migration after origin-year-group and destination-year fixed effects is an origin-destination-time specific shock. We discuss identification in more details in section 3.4.3 below.

Location-time specific control variables include share white, male, foreign-born, literate and share working in the manufacturing sector. Each control variable is measured as the difference between origin and destination locations.

The parameter of interest is β, which is a product of three distinct objects,

$$\beta = \tilde{\beta}\rho\frac{\left(H_{dt} - \overline{H}_{dt}\right)}{H_{dt}}$$

In our model, $\tilde{\beta}$ represents the direct value of a characteristic on a individual's utility, ρ is the individual's extent of updating of their beliefs given the information

available in C_{ot}, and $\frac{(H_{dt}-\overline{H}_{dt})}{H_{dt}}$ is the extent of correction of the individual's belief given their prior if they fully discard their prior. We cannot separately identify any of these objects without data on the other two. We will instead present a range of possible values for each of these objects given our estimate of β and calibrated values for the remaining parameters.

Our use of fixed effects γ_{dt} and γ_{otg} directly corresponds to objects in a individual's utility that are constant at the destination-time level $\left(\tilde{\beta}\overline{H}_{dt} + \xi_{dt}\right)$ and those at the origin-time level $\left(\tilde{\beta}H_{ot} + \xi_{ot}\right)$. As in Berry (1994b), our regression error term corresponds to the origin-destination-time-occupation level unobservable characteristic ξ_{odtg}.

3.4.3 Identification assumptions

Two-way Fixed Effects. In this setting, the main coefficient of interest is the one on the interaction term $C_{ot} \times H_{dt}$, which represents the effect of information (news coverage at origin) interacted with public health amenities (water filtration or typhoid mortality at destination) on migration flows. The goal is to identify this effect while controlling for potential confounding factors. The two-way fixed effects play a crucial role in this identification strategy:

1. The origin-year fixed effects (γ_{ot}) control for any time-varying factors specific to the origin location that might affect out-migration, such as local economic conditions, population changes, or other shocks. They also absorb the main effect of C_{ot}, the news coverage variable.

2. The destination-year fixed effects (γ_{dt}) control for any time-varying factors specific to the destination location that might affect in-migration, such as local economic conditions, amenities, or other shocks. They also absorb the main effect of H_{dt}, the water filtration or typhoid mortality variable.

By including these fixed effects, the regression effectively controls for all factors that vary across origins and years (but are constant across destinations for a given origin-year) and all factors that vary across destinations and years (but are constant across origins for a given destination-year). This leaves only the bilateral variation in the interaction term $C_{ot} \times H_{dt}$ to identify the effect of interest.

Our primary identification assumption is thus that the co-variation in origin information environment (e.g., the presence of a newspaper) and the destination characteristic of interest (e.g., typhoid rate) is exogenous to the flow of migrants between those two places after controlling for unobservable characteristics specific to each place. The key assumption for this identification strategy to be valid is that there are no other bilateral factors (varying at the origin-destination-year level) that are correlated with both the interaction term and the migration flows, after controlling for the two-way fixed effects. If this assumption holds, then the coefficient on the interaction term can be interpreted as the causal effect of information interacted with public health amenities on migration decisions.

Identification Assumptions and Potential Violations. Any variation in the flow of migrants from o to d that correlates with origin information and destination characteristic other than through preference for that characteristic and belief updating violates this assumption. Formally, we assume that

$$\mathbb{E}\left[\xi_{odgt} | C_{ot} H_{dt}, X_{odt}, \gamma_{dt}, \gamma_{otg}, \gamma_g\right] = 0.$$

Importantly, we do not assume that the origin information environment is exogenous to migration or that the destination characteristic is exogenous to migration. To make this discussion concrete, let us consider C_{ot} as the presence of a local newspaper at the origin and H_{dt} as the presence of a sanitation structure at the destination.

We do not assume that a local newspaper's presence is uncorrelated with migration. In fact, the entry of local newspapers is highly correlated with the predicted profit in that market and thus population growth. We only require that it does not differentially influence migration to places with sanitation structures as opposed to places without, other than through preference for sanitation and updating beliefs about the presence of sanitation structures. To the extent that the presence of local newspapers correlates with unobservable characteristics specific to that origin location in that time period, we can control for those unobservables with the use of origin-time fixed effects.

Similarly, we do not assume that a sanitation structure's presence at the destination is exogenous to in-migration. Sanitation structures are more likely to be implemented in cities that are able to fund large public works projects and in the aftermath of deadly outbreaks, and we are able to control for these factors with the use of destination-time fixed effects. We only require that it does not differentially influence in-migration from places with a local newspaper as opposed to places without, other than through preference for sanitation and updating beliefs about the presence of sanitation structures.

Discussion of Potential Caveats. There are potential caveats to our identification strategy that warrant discussion. First, there is a possibility of omitted variable bias if there are time-varying origin-destination specific factors that are correlated with both the interaction term and the migration flows, even after controlling for the two-way fixed effects. For example, if there are bilateral trade or cultural ties between specific origin-destination pairs that evolve over time and are correlated with both information flows and migration, this could bias our estimates. While we believe such factors are likely to be limited in our historical setting, we cannot entirely rule out this possibility.

Second, there may be concerns about reverse causality, such as the possibility that migration flows might influence the entry of newspapers or the adoption of sanitation structures. While the two-way fixed effects help mitigate this concern by controlling for time-varying factors at the origin and destination level, it is worth acknowledging this possibility. However, given the historical context of the United States in the years 1870-1940, we argue that newspapers were a prime way of obtaining information about potential destinations, and their entry was likely driven by factors such as local economic conditions and population growth rather than migration flows per se.

Third, there is a possibility that sanitation investments at the destination could be made in response to expected future migration flows from population origins, which in turn could be more likely to have newspapers. If this were the case, our estimates could be biased, as the interaction term $C_{ot} \times H_{dt}$ would be correlated with unobserved factors driving both sanitation investments and migration flows. While this concern is partially addressed by the inclusion of destination-time fixed effects, which control for any time-varying factors that affect all origins equally for a given destination-year, we cannot completely rule out the possibility of origin-specific expectations shaping destination investments. On the other hand, political complications related to investments into public health discussed in Section 3.2.2 make it highly unlikely that local government officials would have expended political capital toward these investments on behalf of future residents.

Fourth, it is important to consider whether information about disease could be carried through migrants moving back and forth between an origin and destination. If this were the case, our interaction term $C_{ot} \times H_{dt}$ might not only capture the effect of newspaper-based information but also the effect of information transmitted through personal networks and repeat migration. While we believe newspapers were the primary source of information about distant locations during our study period, we acknowledge that interpersonal information flows could have played a role as well. To

the extent that these interpersonal flows are correlated with newspaper presence at the origin and sanitation investments at the destination, our estimates could be biased. Disentangling these different information channels poses an interesting challenge for future research, which could utilize data on repeat migration or social networks to shed light on the relative importance of various information sources.

Fifth, as we do not observe international migrants, there are potential concerns is that we are either understating the importance of international migrants toward public health or not allowing for international migrants to have different responses to newspapers than native-borns. This concern is potentially important: we document that foreign-born populations are mostly located in urban areas (63.2% of foreign-borns are in urban areas against 36.8% in rural areas). Over the 1870-1940 period, foreign-born individuals only represent 8.3% of the population living in rural areas, against 19.7% of the population living in urban areas. Given that our empirical implementation controls for destination-time fixed effects, we are particularly concerned if international migrants can induce rural-urban migration from one origin over another. Since most rural population is native-born (75.5%), we believe that this particular form of bias is unlikely.

Finally, it is important to consider the external validity of our findings, given the specific historical and geographic context of our study. While we believe our results provide valuable insights into the role of information in shaping migration decisions during a period of rapid urbanization and public health improvements in the United States, the generalizability of our findings to other settings or time periods may be limited.

3.5 Empirical estimates

3.5.1 Newspaper Presence and Destination Amenities

Bilateral Flows Sample. Tables 3.3 and 3.4 report the estimates from our primary estimating equation (**??**), where the destination characteristic of interest H_{dt} stand for the presence of water filtration plants and typhoid outbreaks, respectively. In these tables, we restrict origin locations to rural areas and destination locations to cities. We do not condition on the presence of PHI or a typhoid outbreak at destination. Note that restricting the sample to origin-destination pairs within 50 miles do not change the results (these pairs in fact represent 99% of our sample, see discussion from Section 3.2.1 and Figure).

Regression Framework. In Table 3.4 , H_{dt} is measured as average typhoid rates in destination location d over the decennial period t, in deaths per thousand persons. In Table 3.3 H_{dt} is measured as a percentage; it captures the fraction of the 10-years intercensal period t during which a water filtration plant was present at d. For instance, if a water filtration plant was implemented in location d in 1902, H_{dt} takes a value of zero for all census years up to 1900. For intercensal period 1900-1010, H_{dt} takes a value of 0.8. From 1910 onwards, it takes a value of 1. Similarly, in both tables, C_{ot} captures the fraction of the decennial period t during which origin location o had a local daily newspaper.

The dependent variable $\log(s_{odtg}) - \log(s_{ootg})$ captures bilateral migration flows from the logit model, in logs, subtracting the log share of people choosing the outside option of staying put (i.e. choosing $d = o$) to the share of people migrating to a destination $d \neq o$. s_{odtg} captures the probability that an individual from demographic group g moves from location o to d in intercensal period t. Within the model, this dependent variable is interpreted as the utility of destination d for a migrant from o

in time t of group g relative to his outside option. Since we control for origin-time fixed efects (and so the level of $\log(s_{ootg})$) we can interpret the remaining coefficients as shifters of the log probability of moving from o to d, $\log(s_{odtg})$.

The final sample contains 617,483,404 unique matched individuals over the 1870-1940 period. The share is measured for each demographic group g (age, race and literacy) and census time period t. Age, race and literacy are categorical variables. "Race" is split between white and non white and"age" between '16-40'and '41-65'years old.[13] "Literacy" is binary and is split between "literate" and "illiterate". We consider the ability to read as enough to be literate. We do not condition literacy on the ability to write, as reading is the relevant margin of literacy for our newspaper information channel.

All regressions control for occupation fixed effect. We define five occupation groups that aggregate the original IPUMS occupation categories: "Professionals & Managers", "High-skilled Service", "Low-skilled Service", "Farmers and laborers" and "N/A or unemployed".[14] All regressions also include log distance (in miles) between individuals'origin and destination, the interaction of log distance and public health investment (presence of a water filtration plant), log transportation costs between origin o an destination d from Donaldson and Hornbeck (2016) and the mean wage of demographic group g in destination d.[15] County-level control variables include share male, white, literate, foreign-born and share working in the manufacturing sector,

[13]We drop people below 16 and above 65. The mean age in our data is 27, the median age is 24, and the 75th is 4

[14]See `https://usa.ipums.org/usa/resources/volii/Occupations1950.pdf` for a description of the IPUMS'classification of occupations. The original data includes '3 categories: "Professional, technical and kindred workers", "Farmers and farm managers", "Managers, officials, and proprietors, except farm", "Clerical and kindred workers", "Sales workers", "Craftsmen, foremen, and kindred workers", "Operatives and kindred workers", "Private household workers", "Farm labors and foremen","Laborers, except farm and mine", "Service workers, except private household", "Housekeeping/housewife","N/A, not classified or unemployed".

[15]We use Donaldson and Hornbeck (2016)'s predicted county-to-county transportation cost matrices, in logs. Donaldson and Hornbeck (2016) calculate these costs using a newly constructed geographic information system (GIS) network database, which relies on available railroads,canals, natural waterways, and wagons.

from ICPSR. All controls are bilateral measures and capture the difference between an origin and destination pair, as they are otherwise not separately identified from the fixed effects.

Table 3.3: Regressions of Migration Flows to Cities on Newspaper Presence at Origin and Water Filtration Plants at Destination

	Dependent Var.: Destination Utility			
Fixed Effects:	**Origin-Year-Group**		**Origin-Year-Group & Destination-Year**	
	(1)	(2)	(3)	(4)
Newspaper Presence × Water Filtration	0.2551***	0.2531***	0.1431***	0.1286**
	[0.08973]	[0.0809]	[0.0489]	[0.05236]
Water Filtration	0.9123***	0.7409***		
	[0.07325]	[0.06707]		
Log Distance	-0.6609***	-0.8293***	-0.975***	-0.9838***
	[0.05364]	[0.04094]	[0.01491]	[0.01704]
Mean wage of group g in d	1.7e-07***	1.4e-07***	2.5e-08***	1.5e-08**
	[2.9e-08]	[2.5e-08]	[5.7e-09]	[6.6e-09]
Log Distance × Water Filtration	-0.3192***	-0.3281***	-0.3292***	-0.33***
	[0.0335]	[0.02589]	[0.02034]	[0.02125]
Log Transportation Cost	0.5374***	0.7662***	0.9861***	0.9939***
	[0.1355]	[0.1245]	[0.04301]	[0.05056]
County Covariates, in Diff. $(d - o)$		✓		✓
Occupation FE	✓	✓	✓	✓
Constant	-4.956***	-5.492***	-5.116***	-5.085***
	[.1592]	[.1485]	[.05147]	[.07239]
R2	0.6686	0.7014	0.8069	0.8046
F-Stat	1222	912.9	3110	1567
Dep. Var.: Mean	-4.7	-4.724	-4.7	-4.724
Dep. Var.: Std. Dev.	1.54	1.543	1.54	1.543
N Obs.	9.5e+05	7.6e+05	9.5e+05	7.6e+05

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: All coefficients are standardized beta coefficients. Standard errors are clustered at the origin-group-year and destination-year level. "Water Filtration" captures the implementation of water filtration plants at the destination city. All regressions include the interaction of log distance and public health investment (water filtration plant), log distance, log transportation costs from Donaldson and Hornbeck (2016) and mean wage of deographic group g in destination d. County-level control variables include share male, white, literate, foreign-born and share in the manufacturing sector. All controls are bilateral measures and capture the difference between an origin and destination pair, as they are otherwise not separately identified from the fixed effects.

Regression results: Water Filtration Plants at Destination. The first two columns of Table 3.3 report results from a version of model (**??**) that only includes

origin-year-group fixed effect without destination-year fixed effects. Column (1) does not include county-level covariates, while column (2) includes these covariates, measured in difference between a destination and origin location. All coefficients are standardized beta coefficients. On average, the presence of a water filtration plant increases the share of residents from o moving to d by 25% compared to a location d' without water filtration systems, while the joint presence of a newspaper at the origin location o and the presence of water plants at d increases this probability by an additional 74% relative to origins without newspaper coverage (Column 2). Taken together, for origin-destination pairs where the origin o has news coverage and the destination has a water filtration systems, the share of residents moving across that pair is $(25 + 74) = 99\%$ larger than for pairs without either.

Columns (3) and (4) implement the full model (**??**) including a destination-year fixed effect. This controls for any destination-year specific heterogeneity, so we no longer need to control for the sole presence of water filtration plants at the destination. Under this specification, the fixed effects capture unobserved time-invariant factors specific to each destination and thus control for any unobserved factors that might vary across destinations and over time but remain constant within each destination-year combination. The occupation fixed effects further control for variations across job-specific occupations. We find that on average, the share of residents moving from origin-destination pairs where o has news coverage and d has a water filtration plant is 13% higher than other origin-destination pairs. The results suggest that, while migrants generally respond to better public health conditions in urban destinations, those with access to information about these conditions exhibit significantly stronger responses. The differential migration flows based on news coverage at the origin location provide evidence for the importance of information in shaping migrants'decisions and their ability to accurately assess the benefits of

public health investments at potential destinations.

Regression results: Typhoid Outbreaks at Destination. Similarly, in columns (1) and (2) of Table 3.4, we establish that origin-destination pairs where the origin has a local newspaper and the destination has higher typhoid rates experience a lower share of migrants than origin-destination pairs not sharing these characteristics. The coefficient for interaction term News Coverage × Typhoid rate is negative and statistically significant in columns (1), (3), and (4), suggesting that higher news coverage at the origin interacted with higher typhoid mortality rates at the destination is associated with reduced migration flows.

Given an origin o with news coverage, the share of residents moving from location o to destination d with a one standard deviation higher typhoid mortality rate is 30% lower than for a location d' with a one standard deviation lower typhoid mortality rate (Column 2). However, this effect is not statistically significant when controlling for county-level covariates and origin-year-group fixed effects only. Columns (3) and (4) include destination-year fixed effects, controlling for any destination-year specific heterogeneity. Under this specification, we find that the share of residents moving from origin-destination pairs where o has news coverage and d has a one standard deviation higher typhoid mortality rate is 28% lower than other origin-destination pairs, significant at the 10% level (Column 4).

Positive versus Negative Destination Characteristics. Comparing the results from Table 3.3 and Table 3.4 suggests that newspaper presence at the origin interacts with both positive and negative destination characteristics to influence migration decisions. The magnitude of the interaction effect is larger for typhoid mortality than for water filtration. In the most comprehensive specification (Column 4), a one standard deviation increase in water filtration is associated with a 13%

Table 3.4: Regressions of Migration Flows to Cities on Newspaper Presence at Origin and Typhoid Outbreaks at Destination

	Dependent Var.: Destination Utility			
Fixed Effects:	**Origin-Year-Group**		**Origin-Year-Group & Destination-Year**	
	(1)	(2)	(3)	(4)
Newspaper Presence × Typhoid rate	-0.5806***	-0.3071	-0.3199***	-0.2805**
	[0.192]	[0.1897]	[0.1117]	[0.1375]
Typhoid rate	-0.8369***	-0.1383		
	[0.1695]	[0.2316]		
Log Distance	-0.9487***	-1.02***	-1.153***	-1.136***
	[0.06677]	[0.06379]	[0.02861]	[0.03277]
Mean wage of group g in d	9.6e-08***	3.7e-08**	3.4e-08***	2.8e-08***
	[2.2e-08]	[1.7e-08]	[8.1e-09]	[9.5e-09]
Log Distance × Water Filtration	-0.09105	-0.1126***	-0.2426***	-0.2456***
	[0.05621]	[0.04188]	[0.02454]	[0.02678]
Log Transportation Cost	0.744***	0.769***	1.043***	0.9935***
	[0.1655]	[0.1745]	[0.06335]	[0.07208]
County Covariates, in Diff. $(d-o)$		✓		✓
Occupation FE	✓	✓	✓	✓
Constant	-4.052***	-4.648***	-4.335***	-4.006***
	[0.1912]	[0.2217]	[0.06546]	[0.1187]
R2	0.6708	0.7071	0.8037	0.8032
F-Stat	338.2	318.7	879.6	512.3
Dep. Var.: Mean	-4.421	-4.412	-4.421	-4.412
Dep. Var.: Std. Dev.	1.462	1.472	1.462	1.472
N Obs.	3.8e+05	3.0e+05	3.8e+05	3.0e+05

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: All coefficients are standardized beta coefficients. Standard errors are clustered at the origin-group-year and destination-year level. Typhoid rates are measured as deaths by thousand. All regressions include the interaction of log distance and public health investment (water filtration plant), log distance, log transportation costs from Donaldson and Hornbeck (2016) and mean wage of deographic group g in destination d. County-level control variables include share male, white, literate, foreign-born and share in the manufacturing sector. All controls are bilateral measures and capture the difference between an origin and destination pair, as they are otherwise not separately identified from the fixed effects.

increase in migration flows for origin-destination pairs with news coverage, while a one standard deviation increase in typhoid mortality is associated with a 28% decrease in migration flows.

Finally, the main effect of the destination characteristic (second row of the tables) is significant for water filtration but not for typhoid mortality when controlling for origin-year-group fixed effects only (Column 2). This suggests that the presence of water filtration has a positive impact on migration flows even without considering information, while the effect of typhoid mortality is more dependent on the interaction with newspaper presence.

Section **??** presents additional in-migration regression results based on placebo tests that use past migration flows.

3.5.2 News Coverage of Destination Amenities

Our results in the previous section demonstrate that rural migrants differentially respond to changes in public health conditions of urban destinations depending on whether their rural origin has a local newspaper. It could still be that migrants were responding to other amenities that improved at the same time as public health conditions (e.g. employment opportunities, productivity, or leisure). In this section, we augment the previous results and show evidence that news coverage of public health changed with public health conditions and was likely important in driving migration responses.

Regression framework: Response to Typhoid News. We keep our primary estimating equation (**??**), but change the implementation of H_{dt} to be the degree of coverage of waterborne diseases at an urban destination rather than actual case count. We interpret the results of this regression as evidence that migrants potentially respond to news coverage of waterborne diseases rather than the prevalence of the disease itself.

Using our measure of news coverage constructed in Section 2.3 as a regressor within our primary estimating equation (**??**) presents two primary challenges.

The first challenge is a data limitation: we only observe disease coverage for large urban destinations but not for all destination places (such as potential rural destinations or places not named in the census). We take two approaches to this problem. First, we impute the missing amount of coverage to be 0 for rural areas, while unnamed urban areas get the average coverage measure for cities without typhoid outbreaks/PHI investments during that period and within the same state. We believe that this is justified: the rate of typhoid in rural areas was very low (and likely did not receive any coverage).

The second challenge is related to pair-level reverse causality. News about diseases at a given destination d should correlate with mentions of this destination in a newspaper at the origin o, since migrations from o to d would increase demand for news of d at o. Furthermore, previous migrants from an origin o to a destination d could carry information back to o.

Due to this problem, we predict changes in coverage of typhoid at a destination by using the timing implementation of water filtration systems in each city. Formally, we construct $\hat{H}_{dt}$ as the fitted values from the following regression:

$$H_{dt} = \phi \text{Filtration}_{dt} + \gamma_d + \eta_{dt}$$

where H_{dt} denotes the amount of news coverage of typhoid at a migrant's potential destination and γ_d denotes destination fixed effects. We then estimate the following regression:

$$\log(s_{odtg}) - \log(s_{ootg}) = \beta C_{ot} \times \hat{H}_{dt} + \underbrace{\alpha_1 \log(\text{dist}_{\text{od}}) + \alpha_2 \ln(\tau_{\text{odt}})}_{\text{migration costs}} + \alpha_w w_{gdt} + \underbrace{X'_{odt}\alpha + \gamma_{dt} - \gamma_{otg}}_{\text{location-specific cov.}} + \gamma_{\text{occup}} + \varepsilon_{odgt}. \quad (3.4)$$

where the remaining components are the same as that in our primary estimating equation (**??**).

Regression results: Response to Typhoid News. Table 3.5 describes our results. When we do not use the predicted values for typhoid news and instead use the direct value H_{dt} (Columns 1 and 2), we find that origins with a newspaper have more migrants toward destinations that have a higher incidence of typhoid news. This likely represents the previously discussed bias: news coverage of a destination likely increases as in-migration from origins with newspapers increases.

When we instead predict the amount of coverage of typhoid using the implementation of a water filtration system, we find that migrants from origins with newspapers respond especially negatively to the increase in coverage of typhoid news. In our preferred specification (Column 4), an increase in one standard deviation of typhoid news coverage of a destination leads to a 61% lower migration share from an origin with a newspaper.

We interpret this result as evidence that migrants not only respond to changes in typhoid levels but that news coverage is a plausible channel. This is intuitive given our result in Figure 3.8 showing that news coverage of typhoid decreases sharply given implementation of a filtration system. The fact that migration increase when news coverage of typhoid decreases and news coverage of airborne diseases such as influenza and tuberculosis increases (Figure 3.9) suggests that waterborne diseases were far more detrimental to perception of urban conditions than other. It is possible that waterborne diseases are more salient as a factor of urban conditions due to their concentrated and predictably seasonal outbreaks, and that tuberculosis was already treatable in 1900 with drugs that were widely used (Smith, 1988).

Table 3.5: Predicted Typhoid News

	Actual Coverage		**Predicted Coverage**	
	(1)	(2)	(3)	(4)
Newspaper Presence × Typhoid Coverage (std.)	0.085***	0.11***	-0.50**	-0.61**
	[0.014]	[0.021]	[0.20]	[0.27]
Newspaper Presence × Urban/rural dest.	0.050***	0.039*	0.84***	0.83***
	[0.018]	[0.022]	[0.25]	[0.29]
Log Distance	-0.98***	-0.99***	-0.98***	-0.99***
	[0.0073]	[0.0087]	[0.0078]	[0.0095]
County Covariates, in Diff. $(d - o)$		✓		✓
Origin-Year FE	✓	✓	✓	✓
Destination-Year FE	✓	✓	✓	✓
R2	.756	.758	.5386	.5416
Dep. Var.: Mean	-7.00	-7.02	-7.00	-7.02
N Obs.	2.7e+06	2.0e+06	2.7e+06	2.0e+06

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table describes results for our regression of migration flow on the typhoid news coverage. Columns 1 and 2 describe the regression using the actual amount of coverage measured. Columns 3 and 4 describe the two-stage least squares specification where we predict news coverage using the implementation of a water filtration system at the urban destination, destination fixed effects, and time fixed effects. Each column includes origin-year and destination-year fixed effects, and columns 2 and 4 include controls for demographic congruence between the origin and the destination. We report coefficients for three variables. The first is the interaction term between whether the rural origin has a newspaper with the standardized amount of typhoid coverage of a destination. The second is the interaction term between whether the rural origin has a newspaper and whether the destination is classified as a rural or urban place in the census. The third is the natural logarithm of the distance between the origin and the destination. All standard errors are clustered at the origin-year and destination-year level.

3.6 Conclusion

This paper demonstrates the significant role that daily newspapers in rural areas played in the process of urbanization in the United States between 1870 and 1940. Through the analysis of a new dataset that combines full-count U.S. census information with historical newspaper archives, we provide compelling evidence that the dissemination of information about urban amenities, particularly public health improvements such as water filtration and sewage systems, directly influenced the

migratory decisions of rural individuals. Our findings highlight the critical impact of local newspapers in bridging the informational gap about the benefits of living in cities that were investing in sanitation infrastructure, thereby making these urban centers more attractive to potential migrants. Notably, we observe that rural migrants were significantly more likely to relocate to areas boasting public health investments (PHI) when they had access to a local newspaper that reported on these urban advancements.

The implications of our study extend beyond the historical context of the United States, offering potential insights into the ongoing urbanization trends in regions such as Latin America, Asia, and Africa. These areas are currently experiencing rapid urban growth in mega-cities like SÃ£o Paulo, Jakarta, and Mumbai. One plausible explanation for this phenomenon, as suggested by our research, is the increased availability of information. In today's digital age, where information dissemination is even more widespread and accessible than in the past, our findings underscore the importance of information availability in shaping migration patterns. This is particularly relevant in the context of epidemics or pandemics, where real-time information about public health conditions can significantly influence the reallocation of populations.

As the world continues to urbanize at a rapid rate, the lessons drawn from our study emphasize the need for policymakers to focus on enhancing information dissemination about urban amenities and public health investments. This strategy could be instrumental in promoting urban growth and development, especially in developing regions where the rapid expansion of cities presents both challenges and opportunities. In conclusion, our research not only contributes to the academic discourse on spatial and urban economics but also offers practical insights for contemporary urban planning and policy-making, highlighting the enduring relevance of information in shaping human migratory behaviors.

Abramitzky, R., L. Boustan, K. Eriksson, J. Feigenbaum, and S. PÃ©rez (2021): "Automated Linking of Historical Data," *Journal of Economic Literature*, 59, 865–918.

Abramitzky, R., L. Boustan, K. Eriksson, S. Perez, and M. Rashid (2020): "Census Linking Project: Version 2.0 [dataset]," .

Almagro, M. and T. D. Lino (2022 WP): "Location Sorting and Endogenous Amenities: Evidence from Amsterdam," .

Alsan, M. and C. Goldin (2019): "Watersheds in Child Mortality: The Role of Effective Water and Sewerage Infrastructure, 1880–1920," *J. Polit. Econ.*, 127, 586–638.

Anderson, D. M., K. K. Charles, and D. I. Rees (2022): "Reexamining the Contribution of Public Health Efforts to the Decline in Urban Mortality," *American Economic Journal: Applied Economics*, 14, 126–57.

Ao, C.-K. (2015): "Clean Water and Human Capital Investment," *Working Paper*.

Beach, B. (2022a): "Water and Waste: A History of Reluctant Policymaking in U.S. Cities," Invited contribution to proposed volume: *Managing Water: Lessons from History* for University of Pittsburgh Press' History of the Urban Environment Series (Edited by: Joel Tarr and Martin Melosi).

——— (2022b): "Water infrastructure and health in U.S. cities," *Regional Science and Urban Economics*, 94, 103674, urban Economics and History.

Beach, B., J. Ferrie, M. Saavedra, and W. Troesken (2016): "Typhoid Fever, Water Quality, and Human Capital Formation," *J. Econ. Hist.*, 76, 41–75.

Beach, B., J. Parman, and M. Saavedra (2022): "Segregation and the Initial Provision of Water in the United States," *AEA Papers and Proceedings*, 112, 193–198.

——— (1994b): "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25, 242–262.

Boustan, L., D. Bunten, and O. Hearey (2018): "Urbanization in American economic history, 1800-2000," 2, 75–99.

Caliendo, L., M. Dvorkin, and F. Parro (2019): "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 87, 741–835.

Costa, D. L. and M. E. Kahn (2017): "Death and the media: infectious disease reporting during the health transition," *Economica*, 84, 393–416.

Cutler, D. and G. Miller (2005): "The Role of Public Health Improvements in Health Advances: The Twentieth-Century United States," *Demography*, 42, 1–22.

Dell, M., J. Carlson, T. Bryan, E. Silcock, A. Arora, Z. Shen, L. D'Amico-Wong, Q. Le, P. Querubin, and L. Heldring (2023): "American Stories: A Large-Scale Structured Text Dataset of Historical U.S. Newspapers," .

Diamond, R. (2016): "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, 106, 479–524.

Donaldson, D. and R. Hornbeck (2016): " Railroads and American Economic Growth: A "Market Access" Approach *," *The Quarterly Journal of Economics*, 131, 799–858.

Eckert, F. and M. Peters (2018): "Spatial Structural Change," .

Gentzkow, M., E. L. Glaeser, and C. Goldin (2004): "The Rise of the Fourth Estate: How Newspapers Became Informative and Why It Mattered," .

——— (2011b): "The Effect of Newspaper Entry and Exit on Electoral Politics," *American Economic Review*, 101, 2980–3018.

——— (2014): "Competition and Ideological Diversity: Historical Evidence from US Newspapers," *Am. Econ. Rev.*, 104, 3073–3114.

Haines, M. R. and Inter-university Consortium for Political and Social Research (2010): "Historical, Demographic, Economic, and Social Data: The United States, 1790-2002," .

Petrova, M. (2011): "Newspapers and Parties: How Advertising Revenues Created an Independent Press," *Am. Polit. Sci. Rev.*, 105, 790–808.

Porcher, C. (2020 WP): "Migration with Costly Information," .

Ruggles, S., S. Flood, M. Sobek, D. Backman, A. Chen, G. Cooper, S. Richards, R. Rodgers, and M. Schouweiler (2024a): "IPUMS USA: Version 15.0 [dataset]," `https://doi.org/10.18128/D010.V15.0`.

Ruggles, S., M. A. Nelson, M. Sobek, C. A. Fitch, R. Goeken, J. D. Hacker, E. Roberts, and J. R. Warren (2024b): "IPUMS Ancestry Full Count Data: Version 4.0 [dataset]," `https://doi.org/10.18128/D014.V4.0`.

Smith, F. B. (1988): *The Retreat of Tuberculosis, 1850-1950*, Croom Helm.

Troesken, W. (1999): "Typhoid Rates and the Public Acquisition of Private Waterworks, 1880-1920," *J. Econ. Hist.*, 59, 927–948.

——— (2001): "Race, Disease, and the Provision of Water in American Cities, 1889-1921," *The Journal of Economic History*, 61, 750–776.

Wilson, R. (2022): "Moving to Economic Opportunity," *Journal of Human Resources*, 57, 918–955.

Wohl, A. S. (1983): *Endangered Lives: Public Health in Victorian Britain*, Cambridge, MA: Harvard University Press.