The death of distance in hiring

Maria Balgova, IZA*

This version: January 14, 2024
Most recent version available here

Abstract

I estimate the impact of online job boards on the geography of labour markets. I find that the US cities with earlier access to online recruitment experienced an increase in migration flows in and out of the city accompanied by an increase in wages. To understand the underlying mechanism, I collect a novel data set on firm recruitment across space. I show that firms hire outside of their local labour market to find workers of a specific skill rather than low-wage, unemployed labour. As a result, the reduction in the effective distance between local labour markets increased sorting across cities in jobs with high return to match quality. This mechanism contributed to the divergence in outcomes between US cities.

*Email: balgova@iza.org I would like to thank Gabriel Ahlfeldt, Patrick Kline, Yanos Zylberberg, Tyler Ransom, Simon Jaeger, Abi Adams-Prassl, Simon Trenkle, Marc Witte, Ingo Isphording, Gokay Demir, Jakob Wegmann, the participants of the 12th European Meeting of the Urban Economics Association, the seminar audience at the University of Maastricht, and the audience of Essex/RHUL/Bristol Search and Matching Webinar for helpful comments. I am grateful to Nigar Valiyeva for excellent research assistance, and Kory Kroft, Devin Pope, Milena Djourelova, Ruben Durante and Gregory J. Martin for generously sharing their data on Craigslist popularity and expansion.

1 Introduction

In theory, the internet is the perfect solution to search frictions in the labour market. The ability to transmit large amounts of information cheaply, to communicate across a large distance, and to search through hundreds of job postings and candidate profiles should make it possible for employers to hire better, faster and more cheaply than in the analogue world. Indeed, there is some evidence that the introduction of the internet helped job seekers find employment more quickly (Stevenson, 2009, Kuhn and Mansour, 2014, Gürtzgen, Diegmann, Pohlan, and van den Berg, 2021), and improved efficiency in matching in the local labour market (Bhuller, Kostol, Ferraro, and Vigtel, 2022).

However, the internet doesn't just help to search more efficiently *within* a labour market – it also "kills distance", making it easier for workers and firms to meet *across* markets. This second effect, so far overlooked in the literature, is potentially as important as the improvements to local matching. In the majority of countries, labour markets are predominantly local (Moretti, 2011, Manning and Petrongolo, 2017), and so are the technologies counteracting search frictions: social networks and referrals, the assistance of a caseworker in a local employment agency, or simply putting up a vacancy posting in the shop window. In contrast, matching between cities and towns predominantly relies on the individual efforts of workers and firms, and as a result search frictions are larger between than within labour markets (Schmutz and Sidibe, 2018, Porcher, 2021, Balgova, 2022).

In this paper, I study the impact of online recruitment on the matching of workers and firms across labour markets. Ideally, I would want to examine the impact of an exogenous change in search technology on firms' recruitment across space and the resulting mobility and wages of workers. However, this strategy has several practical challenges. The first lies in the nature of the internet itself: it's a technology that radically changed many aspects of life in developed economies, making it difficult to identify the specific effect of online recruitment. The second challenge is that there is no existing data set describing firm recruitment across space and over time. In contrast to the surveys collecting information about how workers search for jobs, we know virtually nothing about why firms recruit outside of their local labour market and for which types of vacancies.[1] As a result, even if it was possible to estimate the impact of online recruitment on workers' mobility and wages, it would be difficult to make sense of why were some workers affected more than others and to identify the underlying causal mechanism.

I address these challenges in three steps. First, I exploit the staggered introduction of a

[1] The existing literature (Stevenson, 2009, Kuhn and Mansour, 2014, Gürtzgen et al., 2021) analysed the shift to online job search and whether it led to higher job-finding rates for workers, but the change in employers' behaviour is only implied by the fact that there are now online job vacancies to apply to.

particular online job board, Craigslist, across US cities. Craigslist acts as an online version of classified ads in a newspaper, making job openings freely viewable for anyone in the country. To isolate the changes in the labour market from the broader impact of Craigslist (and the internet in general) on society,[2] I focus on a narrow window following the website's unexpected rise in popularity in 2004-2006 which corresponded to a quasi-exogenous shift in search technology. The fact that Craigslist websites are city-specific and that Craigslist didn't enter all cities at the same time creates variation in treatment that makes it possible to estimate the impact of online recruitment on workers' mobility and wages. In the second step, I collect a novel data set on firms' recruitment across space. I exploit the placement of help-wanted ads across US newspapers in 1990: newspapers are similar to online job boards in that they allow firms to recruit outside of their local labour market, but, additionally, firms had to make a conscious choice about which newspaper (labour market) to advertise their vacancy in. As a result, this data uniquely captures which jobs, and why, were advertised non-locally. In the third and final step, I link the two parts by showing that Craigslist's impact at the occupation level follows the expected pattern of non-local recruitment after an efficiency-improving technology shock.

I study the impact of Craigslist on two key labour market outcomes at the city level: geographic mobility and wages.[3] To allow for causal interpretation, I estimate a dynamic difference-in-differences (DiD) model of Craigslist entry, comparing treated cities to an observably similar sample of not-yet-treated and never-treated cities. This design is possible because Craigslist websites are city-specific: once posted, Craigslist ads can be seen by anyone with an internet connection, but the website allows employers to advertise local vacancies only.[4] I complement this main identification strategy with several alternative approaches. To address the issue of the endogeneity of Craigslist entry, I re-estimate the dynamic DiD regression using an alternative control group consisting of cities entered by Craigslist's unsuccessful rival Kijiji.com. To demonstrate that the impact of Craigslist entry is driven by job vacancies rather than other types of classified ads (such as housing), I make use of data

[2]There are several studies documenting the impact of Craigslist on crime, health and political outcomes. Kroft and Pope (2014) show it helps the housing rental market clear faster; Djourelova, Durante, and Martin (2022), Seamans and Zhu (2014) document its negative effect on newspaper profits and political coverage; and Brenčič (2016) shows it has hit negatively the competing online job boards. There is also a large number of papers studying the effect of Craigslist's personal ads and erotic services sections on crime rates and health outcomes such as HIV; see Cunningham, DeAngelo, and Tripp (2023) for a summary.

[3]The migration statistics come from the tax return data from the Inland Revenue Service, and the 2005-2007 waves of the American Community Survey. The data on wages comes from the Occupational Employment and Wage Statistics from the Bureau of Labor Statistics.

[4]Craigslist entry thus corresponds to the treatment of the local employers: the availability of a new technology allows the employers in a given city to advertise their vacancies cheaply across the entire country. Local workers, on the other hand, are treated only in that they can now search local vacancies online; the entry of Craigslist into their city doesn't change the job postings they can view in the rest of the country.

on the number of different posts from an earlier study by Kroft and Pope (2014) and estimate a continuous dynamic DiD model where the main explanatory variable is the number of job posts per city.

I find that the availability of Craigslist significantly increased the migration of workers both in and out of the affected cities. On average, gross inflows increased by 6.3% and gross outflows by 3.8%, leading to a small net increase in city-level population of about 0.07%. However, rather than driving relocation away from control to treated cities, Craigslist led to an increase in migration churn *between* treated cities – as reflected in the simultaneous increase in inflows and outflows. Furthermore, I find that this increase was driven by job posts, rather than advertising in the housing market or the general improvement in online connectivity between cities. In particular, 100 Craigslist job posts results in 2 additional workers moving into the city every year.

This increase in geographic mobility was accompanied by an increase in wages. The average worker earned $0.12 per hour, or 0.8%, more each year following Craigslist entry. The increase was larger (1.2%) at the top decile of the wage distribution. As a consequence, Craigslist entry marked an end to the convergence in average pay between cities.

One possible explanation for the simultaneous increase in migration and wages is that Craigslist allowed for greater – and more efficient – matching and sorting across space. However, this explanation hinges on the assumption that firms use online recruitment to find specific talent or better matches outside of their local labour market, rather than to access larger pools of unemployed at lower wages, or to hire more efficiently locally. To distinguish between these different channels, I analyse firms' revealed preference over recruitment across space as observed in the placement of newspaper help-wanted ads. I use a combination of machine learning and keyword search to collect a random sample of job postings from 216 digitally archived newspapers. My final data set contains more than 300,000 job postings from all occupation groups, covering a wide range of US cities and towns. I find that the propensity to recruit across space varies significantly between occupations, falling as low as 6% for Food Preparation, and climbing as high as 17% for Architecture and Engineering. Comparing labour market conditions in the location of the job to those in the location of the newspaper, I find that non-local vacancies tend to be posted in labour markets that pay marginally more, have lower unemployment rates, and have a significantly larger labour force. Moreover, non-local recruitment is strongly driven by the firms' desire to advertise in markets with a relative abundance of the occupation they're searching for. Overall, these results suggest that firms recruit non-locally to find specific talent rather than to access a large pool of cheap or unemployed labour, and that this behaviour varies significantly across occupations.

In the last part of the paper, I show that Craiglist's impact on migration and wages corresponds to the pattern of firms recruiting non-locally to improve their match quality. First, I examine the changes in migration and wages by occupation and demonstrate that, within cities, occupations with the highest migration churn also experience the largest wage increases. Importantly, this relationship only holds for Craigslist cities; in control cities, changes in wages and mobility over time are negatively correlated, as one would expect from shifts in labour supply rather than labour demand. Second, I show that the impact of Craigslist was more pronounced in occupations that had high non-local recruitment in 1990 and so should react more strongly to a fall in the price of non-local recruitment. Similarly, we would expect the growth in wages to be concentrated at the top of the wage distribution, in jobs where the quality of the match matters the most. I document this pattern holds both across and within occupations. In other words, Craigslist's impact is concentrated in the top end of the distribution within occupations with a high return to the quality of the match, in line with the hypothesis that online recruitment significantly reduced spatial search costs and made it easier for workers to find better opportunities across space.

This paper contributes to several streams of literature. I directly build on the research showing that access to the internet has a mixed-to-positive impact on individual-level job-finding rates (Stevenson, 2009, Denzer, Schank, and Upward, 2021, Kuhn and Mansour, 2014, Kuhn and Skuterud, 2004, Gürtzgen et al., 2021) and matching efficiency and search costs at the market level (Kroft and Pope, 2014, Bhuller et al., 2022). Similarly to Bhuller et al. (2022), Denzer et al. (2021), Gürtzgen et al. (2021), I also exploit the staggered roll-out of an internet-related technology, but, in contrast to the existing literature, I focus on its impact on search and matching between local labour markets. Methodologically, this paper is most closely related to Kroft and Pope (2014) who use the variation in the popularity of Craigslist in 2005-2007 to estimate its impact on local unemployment and housing vacancies. Their key finding that Craigslist didn't significantly lower unemployment rates in the cities it entered can be rationalised by my result that the technology increases mobility *between* markets. To the best of my knowledge, the only other paper examining the impact of online recruitment on geographic mobility is a study of broadband roll-out in Norway by Bhuller et al. (2022). They look at geographic mobility as a part of robustness checks, finding a small increase in commuting distance but no changes in moves between commuting zones. These different results are likely attributable to the differences in economic geography between the US and Norway: much of the latter's territory is sparsely populated, and its within-country mobility is about two-thirds of that in the US.[5]

[5]In 2019, the migration rate between Norway's seven regions was 1.9% compared to the 3.2% between-county mobility rate in the US. Data sources: Statistics Norway and US Census Bureau.

More broadly, this paper deepens our understanding of the role of spatial search frictions in internal mobility and regional divergence in the US. Schmutz and Sidibe (2018), Ransom (2022), Balgova (2022), Porcher (2021), Fujiwara, Morales, and Porcher (2021), Wilson (2021) document how search and information frictions caused by geography influence workers' decision to move across labour markets, reducing their mobility. My findings contribute to this literature by showing how online recruitment removes some of these frictions. At the same time, however, I find that the resulting growth in mobility and wages increases, rather than decreases, regional differences. These patterns are in line with the longer-term trends in regional divergence between US regions, first described by Berry and Glaeser (2005), Ganong and Shoag (2017). Starting in the 1980s, the convergence that saw poorer regions in the US catch up reversed, with high-income and low-income cities now growing at the same rates. Giannone (2022) shows that this divergence is almost entirely driven by the growing return to skill in urban areas, and Dauth, Findeisen, Moretti, and Suedekum (2022) find that the quality of the match has been increasing in large cities in Germany. This paper complements this literature by describing the technological changes that made the process of matching and moving into high-return jobs easier for high-skilled workers.

Methodologically, this paper contributes to the growing literature using job vacancies as a source of data on the labour market. The overwhelming majority of these papers make use of online vacancy data, which is more easily available as well as easier to analyse (Hershbein and Kahn, 2018, Marinescu and Wolthoff, 2020, Adams-Prassl, Balgova, Waters, and Qian, 2020, Clemens, Kahn, and Meer, 2021). There is only a handful of papers utilising the text of newspaper job postings. Notably, Atalay, Phongthiengtham, Sotelo, and Tannenbaum (2018, 2020) conduct a detailed text analysis of help-wanted ads from three major newspapers to map the evolution of task content of jobs.[6] Anastasopoulos, Borjas, Cook, and Lachanski (2021) make similar use of postings from *Miami Herald* to measure the labour demand response to the Mariel Boatlift. Compared to these papers, I rely on a larger sample of newspapers and exploit previously unused data on the spatial dimension of recruitment.

The rest of this paper proceeds as follows. In section 2, I describe the data used in my analysis, including the new data set on recruitment across space. Section 3 provides background information about Craigslist and its expansion and outlines the paper's identification strategy. I present my main empirical results in section 4. In section 5, I summarise the patterns of non-local recruitment before the introduction of the internet, and in section 6, I use these patterns to test predictions about the mechanism driving my main results. Section 7 concludes.

[6] This dataset is used in several papers studying changes in the task and skill content of jobs, such as Deming and Noray (2020), Cortes, Jaimovich, and Siu (2023).

2 Data

In this paper, I study the impact of online recruitment in the years 1999-2008 on two labour market outcomes: within-country migration and wages. I focus on these outcomes at the level of metropolitan statistical areas (MSAs) and MSA-occupation level. I use the terms “MSA” and “city” interchangeably throughout the paper. I combine the data on these outcome variables with data on Craigslist entry and popularity and city-level characteristics. I further complement it by constructing a novel dataset on non-local recruitment across US newspapers in 1990. In this section, I describe the origin, definition and construction of each variable. This information is also summarised in Table A2.

2.1 Craigslist

I measure CL’s expansion using 2 variables: the year the website entered a particular city, and the popularity of city-specific websites. Both of these variables come directly from the website itself. The timing of CL’s expansion across US cities until June 2006 is documented on Craigslist’s About webpage[7]. To map out the following wave of expansion in November 2006 (no cities were added in 2007), I use the list of CL’s websites as archived at the end of 2006 on the internet Archive Wayback Machine, cross-checking it with similar data set constructed by Djourelova et al. (2022). The measure of intensive treatment, CL’s city-specific popularity, is the number of job posts per 1,000 inhabitants posted in April 2007, as collected by scraping the CL websites in Kroft and Pope (2014). I conduct webscraping using the internet Archive Wayback Machine to describe the occupational distribution of the posted vacancies. For both measures, I match the city of the website to its corresponding MSA following Kroft and Pope (2014).

2.2 Outcome variables

The data on the main outcome of interest, city-level number of movers, comes from the tax statistics of the Inland Revenue Service. The annual county-to-county mobility flows are constructed from year-to-year changes in the address of residence on individual tax returns.[8] This data captures 95-98% of all tax filing population and provides an exhaustive coverage

[7] `https://www.craigslist.org/about/expansion`

[8] The year of migration is taken from the tax filing year, not the year in which the tax income was earned. The data includes individuals who filed their tax return by the September deadline, e.g. the data for 2003 moves includes all individuals who filed their tax return by September 2003 and moved since they filed their 2002 taxes (i.e. any time between January 2002 and September 2003).

of the entire US territory.[9] I aggregate the county-level inflows and outflows to the city level using the US Census historical delineation files for MSAs in 2003. I construct two versions of the city-level flows: bilateral city-to-city flows, and total city-level inflows and outflows.

Unfortunately, the IRS data doesn't include any individual-level characteristics of the tax filers. As a result, for migration flows at the occupation level, I draw on the American Community Survey, an annual nationally representative survey of more than 3 million individuals (1% of the US population). Starting in 2005, the ACS records the respondents' city of residence in the previous year, along with the current city of residence. This allows me to extract information about migration inflows and outflows at the city-occupation level. While the sample of movers is too small to construct bilateral city migration flows, it is possible to calculate the number of total inflows and outflows at the city level for major occupation groups.

Wage information is taken from Occupational Employment and Wage Statistics, an annual survey of employers conducted by the US Bureau of Labor Statistics. This data set calculates occupation-specific wages at the metropolitan level. I focus on the mean, top and bottom deciles of hourly nominal wages. While OEWS is the most comprehensive data set on city-specific wages for the US, there are three potential issues with the data. First, wage estimates are not always available for all years and all city-occupation cells, primarily due to data confidentiality and data quality reasons. These gaps in coverage will not bias the estimation as long as the observations are missing at random. More broadly, OEWS surveys 396 out of the 922 MSAs, mostly those in the upper half of the population distribution. However, as I explain in Section 3, this is not an issue given my empirical design.[10] Second, the definition of MSA used in OEWS changed in 2005. Since this switch brought the survey in line with the 2000 Census MSA delineation, I use the 2005 MSA definitions as the baseline and manually construct a crosswalk to the 1999-2004 MSAs. Third, the data collection methodology of semi-annual overlapping panels of establishments means that any sharp changes in wages appear in the data more gradually. While this doesn't preclude the use of the data set for time-series analysis, it means that, in my case, any estimates of changes to wages based on a sudden expansion of Craigslist are likely to be biased to 0.[11]

[9] Not every US resident is obliged to file a tax return. The elderly and the poor are particularly under-represented, as are potentially the very wealthy individuals with complex tax arrangements who need an extension to the September filing deadline.

[10] Since most CL cities are also relatively large, I balance the sample of controls to focus on large cities, As a result, the missing smaller cities in OEWS wouldn't be used in the analysis anyway.

[11] Another challenge with using OWES for comparing occupation wages over time are the changes to occupation classification. However, OEWS was using the 2000 SOC throughout the time period of study, 1999-2008.

2.3 Internet penetration and location characteristics

I measure internet availability, an important factor in CL expansion, using data on the number of high-speed internet service providers from the Federal Communications Commission. The FCC creates this statistic from compulsory semi-annual reporting by internet service providers (so-called Form 477)[12], and it is available at ZIP code level. To aggregate it to the number of internet service providers at the MSA level, I calculate a weighted average for each city, using ZIP code population from the 2000 Census as weights.

Data on MSA characteristics comes from 3 different sources. I use OEWS for employment shares and the US Bureau of Economic Analysis for city-level GDP growth rates. I draw on the 2000 Census for the remaining variables (population, land area and density, demographic characteristics).

For county-level characteristics, I draw on annual averages from the Quarterly Census of Employment and Wages, a combination of administrative and survey data capturing labour market statistics for more than 95 percent of U.S. jobs. I use data for 1990 to correspond to my data set on non-local recruitment from the same year.

2.4 Non-local recruitment in 1990

I measure firms' recruitment across space from a novel source: help-wanted ads in US newspapers in 1990. I focus on hiring via newspapers because firms' decision over which newspaper to place their job posting is one of the few directly observable outcomes of their spatial recruitment strategy.[13] I collect my data from newspapers in the year 1990 for two reasons. First, it allows me to avoid any overlap with the arrival of the new online recruitment tools.[14] Second, I am limited by the digitisation process of commercial newspaper archives, which predominantly focus on historical newspapers and have limited availability in the 1990s.[15]

A shortcoming of this approach is that, even in 1990, not all vacancies were advertised in a newspaper: firms also relied on informal networks, recruitment and employment agencies,

[12]The service providers are required to report on whether they offer internet access at speeds exceeding 200 kbps to at least one customer.

[13]Survey data on recruitment is scant in general. The most recent dataset, the Employment Opportunities Pilot Projects, is from 1980. To the best of my knowledge, there is no existing survey dataset on the geographic dimension of recruitment.

[14]The four main online recruitment websites in the US in 2000s, Craigslist, CareerBuilder, Yahoo Hotjobs and Monster.com were founded between 1994 and 1996.

[15]There appear to be two reasons for this. First, commercial digital newspaper archives are primarily marketed at customers conducting genealogy research, so newspapers from the 19th century and older are more valuable than newspapers from 30 years ago. Second, the older newspaper issues are no longer behind paywalls by their publishers. For example, a Library of Congress project digitising US newspapers only runs to 1963.

trade journals, or simply on posting a piece of paper in their shop window. Furthermore, employers likely selected into the different recruitment channels non-randomly so newspapers not only fail to capture all vacancies, they are probably not fully representative of US jobs either. Despite this, the lack of survey data on vacancies has made newspaper help-wanted ads the main gauge of labour demand in the US for much of the postwar period. Until the introduction of JOLTS by the Bureau of Labor Statistics in 1998, the leading – and sometimes only – measure of open job postings was the Conference Board Help Wanted Index, which captured the number of vacancies posted in 51 metropolitan newspapers across the US.[16] My approach thus builds on and expands the standard measure of recruitment in that period.

The source of the help-wanted ads is the digital newspaper archive *Newspapers.com*. It is one the largest archives of digitised newspapers, offering more than 20,000 titles from the 18th century until the early 1990s. I randomly select four days in each month of 1990,[17] using the website to search for all newspapers pages on the given day that contain the phrase "help wanted".[18] This search results in 19,500 newspaper pages. I use a machine learning algorithm to exclude pages that do not contain readable vacancies, leaving 7,850 usable newspaper pages from 216 different newspapers. This corresponds to about 14% of all daily US newspapers. It excludes the four newspapers with nationwide circulation (The New York Times, The Wall Street Journal, The Washington Post, USA Today), but it does cover a range of different newspapers by size, including half of the 50 large city-level newspapers included in the Help Wanted Index.

In the final step, I run a customised keyword search algorithm to extract, from each help-wanted ad, information on the vacancy's occupation and job location.[19] Further details of the page scraping, data cleaning and text analysis can be found in the Appendix.

Each help-wanted ad contains three pieces of information: the location of the job (employer), the location of the newspaper, and the job title (occupation). When the location of the job is missing, I assume the vacancy is local. The location of the newspaper is recorded as the location as its headquarters; while newspaper readership often extends beyond the local town or county, newspaper circulation statistics show that 85% of the readership of the

[16]The index was used as a reliable and representative measure of labour demand in research on a wide range of topics, such as the trends in structural unemployment, productivity, and to understand labour demand over the business cycle (Abraham and Wachter, 1987). It was made redundant by the massive shift of vacancy posting online and was discontinued in 2008.

[17]I exclude public holidays but include Sundays, as many newspapers publish on Sundays, and these editions often carry the bulk of the week's help-wanted ads.

[18]Other phrases, such as "employment" and "jobs" rarely returned any new relevant pages, and contained a lot of false positives in the form of newspaper articles about the economy.

[19]Note that I do not identify the start and end of each advertisement. Instead, I analyse the whole page in one go, re-constructing each ad from the position and order of the recovered keywords.

median newspaper lies within its home county (Djourelova et al., 2022) . To assign an occupation to each posting, I construct a database of job title keywords from the Alphabetical Index of Industries and Occupations, an exhaustive list of occupations and their related job titles developed by the US Census Bureau. It contains almost 31,000 job titles with their corresponding SOC codes, so each job title can be linked to the occupation hierarchy. The list of geographical keywords comes from a gazetteer of the US created by GeoNames[20] and allows me to identify more than 9,500 towns and cities across the US together with their GPS coordinates. These linkages uniquely identify 89.5% locations and 95.4% job titles, and make it possible to merge the vacancy data with any other information on occupations and locations, as well as calculate the distance (in km) between the location of the job and the recruitment location of the vacancy.

The main features of the dataset are described in Table A3. The final sample is fairly representative of US geography (see also Figure A3). It consists of vacancies from 216 newspapers that cover 44 (out of 50) states, 93 (out of 922) MSAs and 174 (out of 3142) counties. The vacancies themselves are even more diverse, originating from 47 states, 242 MSAs and 864 counties. In terms of occupations, the data includes 93 out of 96 occupation groups at the 3-digit SOC level. In total, the sample consists of 318,183 vacancies.

3 Empirical strategy

I use the expansion of a particular online job website, Craigslist (CL), to understand the impact of online recruitment on the spatial dimension of the labour market. In this section, I introduce the website and outline the empirical strategies I use to identify and estimate its impact on geographic mobility and wages.

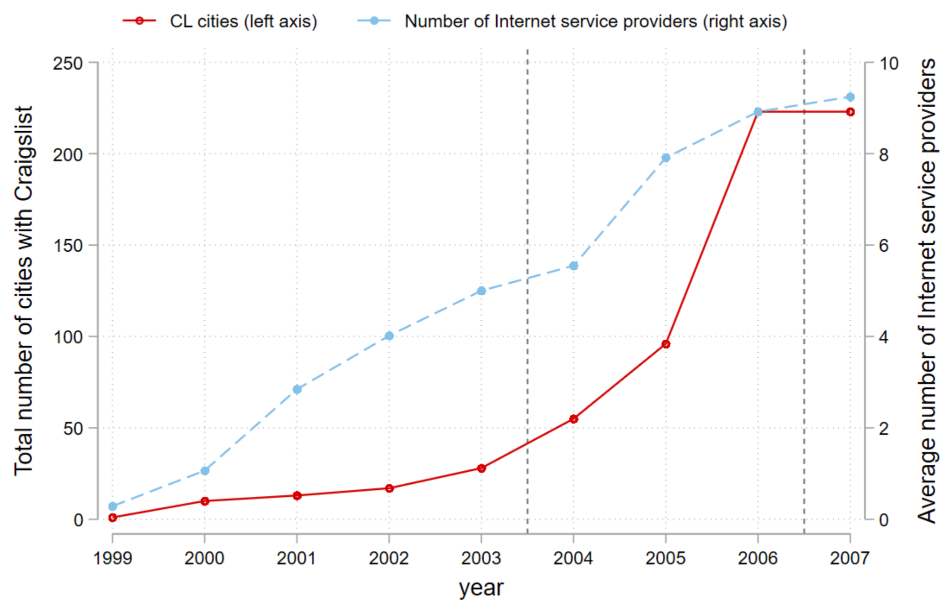
3.1 Craigslist: introduction and expansion

Craigslist is a classified ads website. Just like a local newspaper, it provides space to sell and buy cars and baby clothes, advertise real estate, and look for anything from a babysitter to a romantic partner to a job. Employers wishing to fill a vacancy can advertise under "jobs" (or "gigs" for ad-hoc, short-term work), an online version of a help-wanted ads section. The website also offers a "resumes" section that allows individuals looking for work to upload their CVs.

However, CL is not a typical online recruitment platform. Unlike other online major recruitment websites such as Monster.com, CareerBuilder or Indeed, it doesn't help employers

[20] https://www.geonames.org/

Figure 1: Craigslist expansion and internet availability



Notes: The solid red line (left-hand side axis) plots the expansion of Craigslist until 2007 across US cities. The dashed blue line (right-hand side axis) captures the average number of internet service providers available at the city level. The number of ISP is measures at the ZIP-code level and aggregated by taking a population-weighted average within a city. The vertical lines at 2004 and 2006 mark the start and end of the treatment period considered in this paper. This figure replicates Djourelova et al. (2022) at the city level.

screen applicants, use AI to enhance the quality of job posts or offer an online infrastructure for interested job-seekers to apply directly on the platform. Instead, CL still functions as an online version of a local newspaper classified ads section: it helps to distribute information and leaves the rest of the matching process to the workers and employers. In fact, the website's design has changed relatively little since its inception in 1995. The second major difference is that Craigslist operates a dedicated website for each geographic location. Users are asked to post on their local CL website, and posting on multiple city-specific sites is explicitly prohibited.[21] As a result, even though CL job posts are accessible to job-seekers in any household with an internet connection, only those employers located in a city with a dedicated CL website are able to use it to advertise their vacancies.

These features of CL are a key part of my identification strategy. First, while CL entry makes transmission of information across space considerably faster and cheaper, the lack of

[21] Early on, CL was relying on a system of manual flagging of posts that violated its terms of use. In later years, duplicate postings would be automatically blocked. Source: `https://www.craigslist.org/about/help/faq`

other search-and-match-enhancing capabilities allows me to separately identify the effect of the internet as an information transmitter rather than a direct improvement in matching online. Second, the localised nature of CL websites creates significant geographic heterogeneity in treatment, making it possible to compare treated and control cities (and jobs).

The easy navigation of the website, together with the fact that most posts on CL were free[22], resulted in CL becoming the leading classified-ads website in the US and one of the most visited websites in general during the period studied in this paper. In 2006, CL ranked 47th in terms of unique number of monthly visitors, and 7th in terms of monthly page views – double that of Amazon.com.[23] Kroft and Pope (2014) estimate that in 2007, CL held a two-thirds market share on online job posts from among the 4 major online recruitment websites (CL, Monster, CareerBuilder, and Yahoo/Hotjobs) which, in turn, captured the majority of all online vacancies. CL achieved this level of success in a relatively narrow window between 2004 and 2007. The website was founded by Craig Newmark in 1995 and served only the San Francisco Bay Area until 2000 when it started its slow expansion across other major US cities (see Figure 1). This expansion sped up in 2005, when CL entered 41 MSAs, and was accompanied by a sharp growth in the number of users and monthly visits that wasn't mirrored in that of its main competitors (Kroft and Pope, 2014, Djourelova et al., 2022).[24]

I exploit this sharp and unexpected growth in CL use and availability to identify the impact of online recruitment on labour market outcomes. For this strategy to work, CL's expansion shouldn't be driven by expected growth in the local labour markets it is planning to enter. Importantly, anecdotal evidence as well as several studies (Kroft and Pope, 2014, Djourelova et al., 2022) have shown that CL's expansion and popularity were primarily driven by the idiosyncratic opinions of its CEO and owner rather than strategic profit-maximisation or predictions about economic growth. In a 2004 interview with the San Francisco Chronicle, Craig Newmark (the owner) explained that the entry into a new city is determined by demand from potential local users and Jim Buckmaster's (the CEO) perception and intuition about the city.[25] He identifies just two relevant city characteristics that determine CL's entry – the

[22]Until 2008, CL only charged for job postings in its 11 biggest markets (San Francisco Bay Area, Boston, Chicago, Los Angeles, New York, Orange County, Portland, Sacramento, San Diego, Seattle, and Washington DC), plus a fee for brokered apartment rental listings in New York. Source: `https://web.archive.org/web/20080228034906/https://www.craigslist.org/about/help/posting_fees`. Cajner and Ratner (2016) show that the change in CL's pricing policy at the end of 2012 significantly reduced the number of online job posts, and substantially drove the divergence between JOLTS and the Help-Wanted Online Index.

[23]Source: `https://web.archive.org/web/20170909054652/https://www.forbes.com/2006/12/08/newspaper-classifield-online-tech_cx-lh_1211craigslist.html`

[24]Kroft and Pope (2014) use Corzen.com proprietary data to track the number of posts in a subset of CL websites alongside the main competitors (Monster, Careerbuilder, Yahoo) over time. Djourelova et al. (2022) draw on data from Comscore, a nationally representative survey of web browsing behaviour, to compare CL's traffic to that of Monster and eBay.

[25]Source: `https://www.sfgate.com/business/ontherecord/article/`

city's population size and broadband availability.

I present the descriptive statistics of CL and non-CL cities in the first three columns of Table 1. CL's city entry strategy is reflected in the characteristics of CL cities: they are significantly larger, both in terms of population and land area, than cities where CL didn't enter by 2007. CL cities are also more densely populated, more ethnically and racially diverse, younger, and with a somewhat higher employment share and wages. Interestingly, however, CL and non-CL cities do not differ in their migration in- and out-flows, which hover at around 2% annually.

Nevertheless, systematic differences between CL and non-CL cities aren't necessarily evidence of CL strategically entering high-wage, high-employment, or high-growth cities. I test this hypothesis for the years 2004-2007 directly in Table A4. Each column of the table summarises the results of a probit regression of CL entry on city characteristics before 2004. The first two columns confirm Craig Newmark's statement that CL is significantly more likely to enter large cities with a large number of internet service providers, a reliable proxy for the availability and use of the internet (Djourelova et al., 2022). In column (3), I regress CL entry on realised employment and GDP growth over the 4 years prior to CL entry (2000-2003), and in column (4), I regress it on the predicted growth in these variables for the years 2004-2007.[26] Neither realised nor predicted city growth explains CL entry. Column (5), which simultaneously controls for all variables, shows that if anything, higher predicted GDP growth decreases the probability that CL will enter the city over the given time period. Overall, even though population size and the growth in internet availability are strong predictors of CL entry, all the variables combined explain only 15.5% of the variation in CL entry, supporting the anecdotal evidence that CL entry was primarily driven by the enthusiasm of local supporters and the idiosyncratic preferences of CL management.

3.2 Main empirical strategy: Staggered rollout of Craigslist

My main identification strategy for estimating the impact of CL entry is based on the staggered rollout of CL across US cities between 2004 and 2006. As explained in the previous section, CL entry in US cities varied across time, and this entry wasn't associated with higher past or predicted economic conditions at the city level. As a result, I identify the treatment effect on wages and migration by comparing city-level outcomes between treated and control cities, taking into account the staggered timing of CL entry. I estimate the following dynamic

CRAIGSLIST-On-the-record-Craig-Newmark-2733312.php

[26] I run a linear dynamic forecast of city-level employment and GDP growth using the lagged dependent variable and the following city characteristics from the 2000 US Census, based on Glaeser and Shapiro (2001): education attainment, population density, the share of workers working in the metropolitan city, the age of city buildings, the number of cars per household, the length of the average commute.

Table 1: Descriptive statistics

	CL cities		Non-CL cities		
	(1)	(2)	(3)	(4)	(5)
	All	2004-2006 entry	All	Kijiji cities	Other cities
Panel A: Mobility and demographics					
Gross inflows	17675.22	9360.34	2074.96	10806.59	1580.97
	(27204.21)	(9199.45)	(6056.41)	(21445.90)	(2972.92)
Gross outflows	17769.97	9152.59	2045.01	9468.78	1625.01
	(31286.14)	(7938.97)	(5226.04)	(15273.66)	(3563.51)
Inflow share, %	2.24	2.23	2.01	2.69	1.98
	(1.10)	(1.09)	(1.14)	(1.11)	(1.13)
Outflow share, %	2.25	2.27	2.14	2.53	2.12
	(0.96)	(0.98)	(0.95)	(0.98)	(0.95)
Population (1,000s)	906.88	435.57	86.30	367.74	70.38
	(1833.31)	(390.44)	(166.30)	(553.50)	(87.13)
Land area (km-sq)	3024.88	2656.18	1352.29	2382.52	1294.01
	(2989.56)	(2821.76)	(1980.57)	(4430.41)	(1731.84)
Population density	2840.54	2284.48	922.84	2136.36	854.19
	(2652.78)	(1049.88)	(849.86)	(1531.56)	(738.71)
Population share, white (%)	79.81	81.18	83.72	83.19	83.75
	(11.93)	(11.11)	(15.18)	(11.78)	(15.36)
Population share, black (%)	10.49	10.12	8.97	7.83	9.03
	(10.39)	(10.66)	(13.75)	(9.54)	(13.95)
Population share, hispanic (%)	10.24	9.71	7.48	8.54	7.42
	(15.66)	(16.17)	(13.55)	(11.53)	(13.66)
Median age	34.35	34.27	36.27	35.77	36.30
	(3.66)	(3.82)	(3.73)	(3.84)	(3.73)
Share under 18 (%)	25.38	25.37	25.48	25.34	25.48
	(3.15)	(3.31)	(3.00)	(3.12)	(2.99)
Share over 75 (%)	5.80	5.85	6.70	5.92	6.75
	(1.62)	(1.61)	(1.77)	(1.92)	(1.76)
N	223	195	691	37	654
Panel B: Labor market					
Employment share (%)	45.29	45.17	41.30	43.34	40.38
	(11.58)	(11.50)	(14.65)	(20.02)	(11.56)
Hourly wage, mean	14.66	14.27	13.87	14.27	13.70
	(1.75)	(1.43)	(1.74)	(1.45)	(1.84)
Hourly wage, bottom 10%	7.80	7.63	7.65	7.94	7.52
	(1.15)	(1.10)	(0.88)	(0.82)	(0.88)
Hourly wage, top 10%	23.33	22.74	21.99	22.47	21.78
	(3.62)	(3.49)	(2.31)	(1.90)	(2.45)
N	193	166	87	27	60

Note: The table summarises city characteristics in 2000. Column (1): all MSAs that Craigslist entered before 2008. Column (2): MSAs that Craigslist entered between 2004 and 2006. Column (3): all MSAs without a Craigslist website by 2007. Column (4): MSAs without a Craigslist website but with a Kijiji website in 2007. Column (5): MSAs without either Craigslist or Kijiji website by 2008. Standard errors in parentheses.

two-way fixed-effects regression:

$$Y_{ct} = \alpha_c + \beta_t + \sum_{j=-6}^{j=-2} \delta_j CL_c + \sum_{j=0}^{j=3} \boldsymbol{\delta_j} CL_c + \theta_t X_c + \epsilon_{ct} \tag{1}$$

where Y_{ct} is the outcome for city c in year t, α_c and β_t are city- and year-fixed effects, and δ_j are the parameters of interest, estimating the difference between treated and control cities 6 years before and up to 3 years after the treatment. $\theta_t X_c$ is the interaction of year-FE and baseline city-level internet availability (in levels and growth), allowing for differential growth rates of cities with different connectivity in 2001. The key identifying assumption is that, in the absence of treatment, wages and mobility in CL cities would have evolved in the same way as in non-CL cities (parallel trends assumption). While this assumption cannot be tested empirically, coefficients δ_j allow me to test whether the treated and control cities evolved similarly before the treatment.

I construct the control group so that the two groups are comparable in pre-treatment *level* characteristics. Using Mahalanobis distance matching, I select a subsample of never-treated cities that are similar to treated cities along a range of characteristics in the year 2000 (population size, land area, population density, racial and ethnic composition, age, employment share). I present the balance test for the matched sample in Table A5. The first 3 columns summarise the baseline characteristics of treated and control cities and their difference. The matching procedure does a good job of pairing cities that are similar in terms of migration rates, demographics, and employment share. The only remaining statistically significant differences are in population size and density: treated cities in the balanced sample are on average twice as large as the control cities, with a correspondingly higher population density. As a result, workers in the treated cities are paid on average 50c an hour more. This difference in city size stems from CL's focus on large cities: by the end of 2006, CL had entered 164 out of the 200 most populous cities in the US (see Figure A2). As a consequence, restricting control and treated cities to match on population size would lead to a very small control group.

To rule out the possibility that the estimated treatment effects are driven by city size rather than CL entry, I perform two robustness checks. First, I impose common support in population size on the matched control and treated cities. Second, I only estimate the dynamic diff-in-diff on (all) CL cities, identifying the treatment effect solely from the staggered timing of CL entry. In addition, I address the question of the appropriate control group addressed directly in an alternative empirical strategy (see next section).

Another threat to identification is anticipation effects. However, I argue that these are less relevant in this particular setting. CL entry into a city was unannounced (Cunningham et al., 2023), and in general, CL didn't rely on advertisement to attract users and boost its viewership, making it difficult for employers to increase wages (or workers to bargain for higher pay) in anticipation of CL entry. Anticipation effects are even less plausible for migration flows in and out of the treated cities, since my main hypothesis is that CL websites provided the mechanism that made moving easier, and movers had little to gain from moving in anticipation of CL entry.

I focus on CL entry in the years 2004-2006 for two reasons. First, it offers a cleaner definition of "treatment": Kroft and Pope (2014) demonstrate that CL popularity in terms of pageviews and users only started to grow substantially in 2005. As a result, cities where CL entered before this date had access to the technology, but a much smaller user base likely resulted in a weaker effective treatment. This argument is closely related to the second motivation for focusing on entries starting in 2004: treatment effect heterogeneity. The treatment effect of CL entry for the earlier years might be different not only because of differences in general CL popularity but also because the CL entry in its first few years was heavily concentrated among the few largest MSAs in the country.[27] Including these cities in the treatment group would lead the standard dynamic two-way fixed-effects specification to produce biased estimates of the treatment effects.

Nevertheless, it is still possible that the dynamic treatment effects of CL entry are heterogeneous, for example is the effect of CL entry on wages in year j after treatment is smaller for cities CL entered later. Sun and Abraham (2021) show that if the dynamic treatment effects are not homogenous, the TWFE estimator of δ_j will likely not correspond to an unbiased estimate of the average treatment on the treated and it may be based on "forbidden comparisons" of the newly treated cities with already-treated cities. To test whether my baseline estimates from regression (1) are robust to heterogeneity in treatment effects, I re-estimate (1) using an alternative estimator by Callaway and Sant'Anna (2021) which accommodates heterogeneous dynamic treatment effects and imposes the control group to only consists of never-treated (non-CL) cities.

3.3 Alternative empirical strategies

Externally determined control group The chief potential weakness of the empirical strategy outlined above is that CL picked cities to enter based on predicted strength growth in the local economy. Even though I have shown, in Table A4, that the past nor forecast employment and GDP growth can explain CL entry, it still might be the case that CL

[27] Cl entered all 10 largest MSAs by 2003. Following its first location in San Francisco Bay Area, it entered the 3 largest MSAs (New York, Los Angeles and Chicago) in the first expansion wave in 2000.

management was able to cherrypick the cities that would experience faster wage growth and mobility in the future. This is particularly concerning given that there is no publicly available data explaining how exactly Craig Newmark and Jim Buckmaster selected which cities to operate in. The ideal set of control cities would consist of near-identical MSAs that CL should have entered but didn't. Crucially, they should not only be similar to CL cities in terms of their baseline characteristics, but also in their market potential and profitability.

In this section, I present an externally identified set of control cities that plausibly fit the characteristics of an ideal control group. I then re-estimate the dynamic diff-in-diff specification (regression (1)) using this externally determined control, complementing my main empirical strategy.

This alternative identification strategy originates in a short-lived involvement of eBay, the online marketplace, with Craigslist in 2004-2008. In the early 2000s, eBay developed an interest in the online classified business; the company saw it as a natural next step in its growing portfolio of various trade-related online services. When in 2003 a disgruntled Craigslist stake-holder decided to sell his 28.4% share in the company, eBay seized the opportunity and became a minority shareholder. eBay's objective was to eventually purchase the entire company, but following strong opposition from the other two owners (Craig Newmark and Jim Buckmaster, the CEO), eBay decided to use its minority stake to at least "learn the "secret sauce" of Craigslist's success, presumably so that eBay could spread that sauce all over its own competing classifieds site.".[28] Over the following 3 years, eBay's representatives on CL's board of directions passed on non-public financial statements, capacity projections, and other private Craigslist documents on to its internal team tasked with launching Kijiji.com, eBay's version of classified ads website.

Kijiji launched at the end of June 2007 across 220 US cities: in 183 out of 224 CL cities, plus another 37 cities that CL hadn't entered yet. These non-CL Kijiji cities form a near-ideal control group for CL cities, for three reasons. First, eBay decided to launch Kijiji after it became clear they wouldn't be able to acquire Craigslist, and used private Craigslist data to decide which cities to enter. Given the strong for-profit motivation of eBay, especially in contrast to the "business as a community service"[29] philosophy behind Craigslist, Kijiji cities can be interpreted as the cities CL should have entered to maximise profits and growth. Second, eBay's development of Kijiji didn't affect CL's operations. CL directors were unaware of the Kijiji launch until less than a month before the website went live, and they didn't know that some part of Kijiji strategy was based on internal CL

[28] This is a quote from a ruling from the Supreme Court of Delaware on eBay v. Craigslist. The statement is based on internal eBay communication that was presented in the trial. Source: https://caselaw.findlaw.com/court/de-supreme-court/1558886.html#footnote_33

[29] Source: https://caselaw.findlaw.com/court/de-supreme-court/1558886.html#footnote_33

documents until eBay sued CL over the breach of fiduciary duty in 2008 and their activity was revealed in the legal discovery. Third, even though Kijiji officially launched in June 2007, it didn't start to fully compete with CL until 2008. This means that the Kijiji launch didn't significantly alter CL operations and userbase until after the time period studied in this paper.[30]

I present the characteristics of non-CL Kijiji cities in column (4) of Table 1. Two main patterns stand out. First, non-CL Kijiji cities are substantially different from other non-CL cities (column (5)): they are more than 5 times larger in terms of their population, have higher population density, are somewhat younger, and have larger employment shares and wages. Second, all of these attributes make them relatively much more comparable to CL cities. I make this point explicitly in columns (4)-(6) of Table A5, where I compare CL and non-CL cities within the Kijiji group. While non-CL Kijiji cities are still somewhat smaller than CL cities, this difference is substantially smaller than for the matched treatment and control cities, and it is significant only at 10%. The two groups of cities are statistically indistinguishable in all other characteristics except for migration shares, where non-CL Kijiji cities experience somewhat greater in- and out-flows than CL cities. The same pattern is visible when we compare the control group (non-CL Kijiji cities) to all CL cities (presented in column (1)): the population size difference disappears, and the two groups are near-identical in their demographics and labour market characteristics.

Craigslist popularity The second alternative identification strategy is based on Kroft and Pope (2014) who exploit the variation in CL's popularity across cities rather than the timing of its entry. The advantage of their approach is twofold. First, it identifies the impact of CL within the cities that CL entered, circumventing the problem of endogeneity of CL entry. Kroft and Pope (2014) argue that the popularity of the website was driven by chance much more than its entry (which was, after all, decided by the company's CEO). Second, given that the impact of any technology depends on the extent of its usage, focusing on the intensive margin is probably closer to the ideal experiment in which some cities randomly use CL and others don't. In other words, the estimation using staggered CL entry identifies an intention-to-treat effect, while the estimation using CL popularity identifies a treatment effect.

I combine the two approaches and estimate dynamic difference-in-differences using CL popularity as a continuous treatment. I compare cities where CL became more popular to cities where it was less popular or it never entered, using the dynamic specification to control for the staggered CL entry. I measure CL popularity using the scraped number of

[30] In fact, Kijiji never fully took off in the US and the website was shut down in 2010.

job posts across city websites in April 2007, normalised by city population, from Kroft and Pope (2014). As in my main empirical strategy, I restrict my attention to cities that are matched on baseline covariates. I run the following version of regression (1):

$$Y_{ct} = \alpha_c + \beta_t + \sum_{j=-6}^{j=-2} \delta_j Posts_c + \sum_{j=0}^{j=3} \boldsymbol{\delta_j} Posts_c + \epsilon_{ct} \tag{2}$$

where, as before, δ_j is the coefficient of interest. Unlike in the binary main specification, the interpretation of these coefficients is somewhat more complicated. Callaway, Goodman-Bacon, and Sant'Anna (2021) show that δ_j can be interpreted as a weighted average of causal response at a particular treatment dose, i.e. the change in outcomes due to a change in treatment dose.[31] However, in order to interpret δ_j as the average causal response of CL popularity on labour market outcomes, we need to be willing to make several assumptions about the nature of the treatment.

First, I need to assume so-called strong parallel trends: high-CL-popularity cities would develop similarly to low-CL-popularity cities had CL been less popular there. In other words, not only do non-CL cities need to be a good counterfactual for CL cities, but low-CL-popularity cities must be a valid counterfactual to high-CL-popularity cities. This assumption is satisfied if there is no selection into the intensity of treatment: CL popularity mustn't be driven by expected wage growth or mobility at the city level. While Kroft and Pope (2014) show that, unlike CL entry, the number of posts isn't correlated with population size or internet availability, and that, anecdotally, CL management was surprised by where CL became popular, CL still might have been used more enthusiastically in cities where its returns were higher – for example in tight labour markets which would have seen fast wage growth anyway. To address this issue of selection bias, I instrument the number of job posts with the number of normalised personal posts (community, services, for sale) in the same city. This IV will be valid if CL popularity in one type of posts spills over into the popularity of other types of posts, but, at the same time, CL popularity in personal posts mustn't be related to local labour market conditions.[32]

The validity of the first assumption depends on the validity of the instrumental variable. While it's not possible to prove the exclusion restriction holds, it is supported by the anecdotal evidence that the website's popularity was difficult for CL management to predict, as

[31] The alternative interpretation is that δ_j corresponds to the average treatment effect on the treated, i.e. the average *level* difference between treatment and no treatment. However, as the authors show, TWFE estimates of δ_j are weighted dose-specific ATT effects, using weights that are not always sensible or non-negative.

[32] In the language of Callaway et al. (2021), a valid instrument will remove the selection bias term from the estimates of average causal response.

described in Kroft and Pope (2014). Regarding the instrument's relevance, I document the first-stage relationship between the number of personal posts and the number of job posts in Figure A4. Panel (a) of the figure shows that there was significant variation in CL popularity. About half of the CL cities only recorded 1 job post per 1,000 inhabitants in April 2007, but the top quartile of cities has double the average number of posts. In panel (b), I show that CL popularity in job postings was closely related to CL popularity in personal posts: in other words, cities where CL was a popular platform for personal ads were also the cities with a high number of job postings, making it a relevant instrument for job post popularity.

The second necessary assumption for the correct interpretation of δ_j concerns the homogeneity of the treatment effect. The dynamic diff-in-diff specification in equation (2) aggregates individual dose-specific causal responses. This aggregation can be interpreted as an average causal response to CL popularity when the treatment effect is homogeneous across (i) the timing of the treatment, and (ii) the treatment intensity. In other words, I assume that a city with CL entry in 2004 would have reacted to its treatment in the same way had CL entered in 2006; and that the causal response function of CL popularity is linear, so that low CL popularity and high CL popularity result in the same relative causal response. The linearity assumption is imposed by the functional form of regression (2). Probably the main reason why assumption (i) might not be satisfied is if CL popularity depends on how long the website has been available in the given city. If CL popularity grows over time, the older CL cities will be treated more intensely than the ones where CL entered relatively recently (note that I cannot capture the changes in CL popularity over time because the measure of CL posts is only available for a single month at the end of the studied time period[33]). Two pieces of evidence make this relatively unlikely. First, the period of treatment under study, 2004-2006, is chosen precisely because it was a time of rapid CL expansion and popularity growth across the US; CL websites reached relatively high levels of popularity within a few months of entering the market. Second, I show in panel (c) of Figure A4 that average CL popularity varied only weakly across the three years of entry in my sample. While CL entered its most popular cities in 2004, only 5 cities from this entry cohort experienced higher popularity than the most popular CL city of 2005. Moreover, mean popularity is statistically the same for the 2005 and 2006 entries, suggesting that the intensity of treatment wasn't significantly lower for cities where CL entered late.[34]

[33] The number of posts is measured in April 2007 for all MSAs that CL entered by June 2006. This means that this sample excludes the 61 MSAs CL entered in November 2006.

[34] This finding is at odds with the patterns reported in Kroft and Pope (2014), who find that the time of entry is one of the few predictors of CL popularity in a given city. This discrepancy is entirely driven by the differences in samples. While Kroft and Pope (2014) look at all CL entries since 1995, I restrict my attention to entries in or after 2004. The fact that the first cities CL entered were significantly different from the average US city – and we would thus expect CL popularity to differ - is one of the main reasons why I

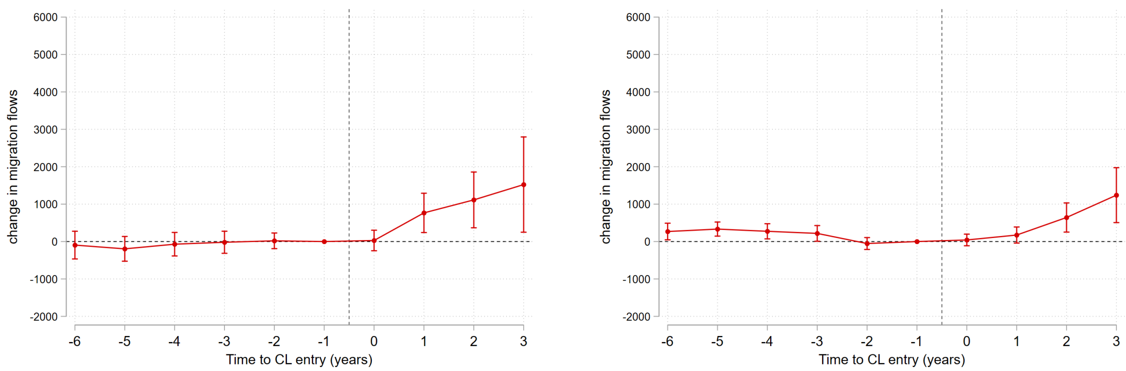
4 The impact of Craigslist on mobility and wages

This section summarises the main empirical findings of the paper. I show that Craigslist's entry into different US cities caused an increase in migration inflows and outflows from these cities, and raised wages, especially at the top of the distribution. The complementary identification strategies and robustness checks provide evidence that this treatment effect is mainly driven by Craigslist's use as an online recruitment tool.

4.1 Impact on geographic mobility

Figure 2: Impact of CL entry on aggregate migration flows

(a) Inflows, baseline estimates (b) Outflows, baseline estimates



Notes: Coefficients from dynamic TWFE difference-in-differences regression (equation (1)). The regressions include city and time fixed effects and baseline levels and growth of Internet service providers interacted with time FE. The model is estimated on a sample of cities balanced on baseline covariates. Vertical bars represent confidence intervals at 95% level of statistical significance.

I present the baseline results of the impact of Craigslist on city-level aggregate migration in Figure 2. The dependent variables are the annual totals of inflows and outflows of individuals from other US cities. They show the results of a TWFE dynamic difference-in-differences regression for Craigslist entries between 2004 and 2006, controlling for city and time fixed effects and allowing mobility to vary across cities with different internet availability. The plots show that CL entry significantly increases both the number of inflows and outflows from the city. By year 2 after CL entry, the treated cities experience on average 1113 more individuals moving into the city, and 643 more individuals moving out of the city compared to cities without CL; this effect increases over time. Given the average number of in- and out-movers per year in a CL city (17,701 and 16,761, respectively, in the year before

restrict my sample this way.

treatment), this CL-induced increase corresponds to a 6.3% increase in inflows and a 3.8% increase in outflows at the city level.

I explore these patterns further by running a series of robustness checks and alternative identification strategies. In Figure A6, I re-estimate the baseline specification using an alternative estimator by Callaway and Sant'Anna (2021), showing that the results are robust to treatment effect heterogeneity. In Table 2, I present the difference-in-difference estimator for a range of alternative specifications[35], starting by replicating my preferred specification in column (1). In column (2), I allow for city-specific linear trends in mobility (instead of allowing for different trends based on internet availability), and in column (6), I show the results hold when I include CL's first entries between 2000 and 2003. Because these early cities are bigger, the size of the effects in terms of the number of in- and out-migrants is also larger. In column (5), I show that my results carry through when mobility flows are expressed in logs rather than levels.[36]

One of the potential caveats of my analysis is the fact that CL focused on entering large cities (CL cities are significantly larger than non-CL cities even in the matched sample; see Table A5). I address this point in several ways. First, in column (4), I impose common support in population size to the cities in the matched sample. This removes some of the largest CL cities from the treatment group, reducing the overall size of the treatment effects, but the results remain similar to the baseline estimates in column (1) and are statistically significant. Second, I restrict my sample to CL cities only, so that the treatment effect is identified from the comparison to not-yet-treated cities. I present these results in column (3), again finding an increase in inflows and outflows from the affected cities. Third, I turn to Kijiji cities as an externally determined control group. These results are presented in column (7). I find that inflows and outflows from CL cities increase compared to non-CL Kijiji cities, although these differences are not statistically significant due to the large variance in outcomes of Kijiji cities.

The estimates presented so far are measuring the intention-to-treat treatment effect of Craigslist entry. However, as I explain in Section 3.3, the actual usage of CL varied significantly across the cities in entered, ranging between 0.2 job posts per 1,000 inhabitants per month in the cities where CL was the least popular to 9 job posts in Santa Barbara, its most

[35]The parameter presented in Table 2 can be interpreted as an average treatment effect of CL entry on city mobility. Recent literature has shown that this estimate corresponds to a *meaningful* ATT only under a set of relatively restrictive assumptions on the homogeneity of the treatment effect across treatment cohorts and time. As a result, I also present the dynamic diff-in-diff estimates (corresponding to versions of regression specifications (1) in Table A6 in the Appendix. While the size of the estimated effect varies as expected, the qualitative results of an increase in migration inflows and outflows hold.

[36]I study changes in the number of movers rather than the changes in mobility growth because geographic mobility was stable during the period in question; see also Figure A5.

Table 2: Impact of Craigslist entry on aggregate migration flows, robustness checks

	Balanced sample				All data		Kijiji cities
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Levels	Levels	Levels	Levels	Log	Levels	Levels
Panel A: Gross Inflows							
Treatment effect	960.0***	659.8**	1388.9**	663.1***	0.0310***	941.2***	484.7
	(290.8)	(298.2)	(558.6)	(214.0)	(0.0106)	(291.9)	(633.7)
Panel B: Gross Outflows							
Treatment effect	313.4**	260.6*	200.6	235.2**	0.0374***	1251.1***	200.3
	(142.1)	(135.8)	(227.4)	(110.9)	(0.00839)	(343.4)	(711.7)
Panel C: Net Inflows							
Treatment effect	646.6**	399.2	1188.3**	427.9*	0.154	-309.9	284.3
	(279.1)	(363.7)	(550.7)	(221.1)	(0.0968)	(347.9)	(775.1)
N	2200	2200	1670	1980	3577	9160	2110
MSA FE	YES	YES	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES	YES	YES
Linear MSA trend		YES					
Covariates × time	YES			YES		YES	YES
CL cities only			YES				
City size common support				YES			
All CL entries						YES	

Note: This table summarises the results of a TWFE regression of CL entry on city-level migration inflows and outflows. The presented coefficient estimates the annual change in migration flows at city level following CL entry, averaged across different years after entry. Column (1) corresponds to my baseline preferred specification, using a matched sample of treated and control cities and including an interaction between period FE and baseline internet service availability. Column (2) uses city-specific linear growth trends instead. Column (3) re-runs the baseline specification from column (1) but only on CL-cities. Column (4) re-runs the baseline specification further imposing common support over population size of treated and control cities. Columns (5) and (6) are estimated on the entire sample of cities; column (5) uses log, rather than levels, of migration flows, and column (6) includes cities that were treated in the years 2000-2003. Column (7) re-runs the baseline specification on Kijiji cities only. Standard errors are in parentheses, clustered at city level throughout. $*p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table 3: The impact of the number of Craigslist job posts on migration flows and wages

	OLS		IV	
	(1)	(2)	(3)	(4)
Panel A: aggregate outflows				
Job posts per 1,000 inhabitants	297.8**	175.6*	432.3***	355.1**
	(120.8)	(103.9)	(164.7)	(178.7)
Panel B: aggregate inflows				
Job posts per 1,000 inhabitants	372.5**	165.8	391.6**	141.0
	(149.8)	(136.9)	(176.9)	(188.3)
Panel C: mean hourly wage				
Job posts per 1,000 inhabitants	0.0991***	0.0960***	0.115***	0.116***
	(0.0178)	(0.0201)	(0.0224)	(0.0254)
Panel D: bottom 10% hourly wage				
Job posts per 1,000 inhabitants	0.0455***	0.0465***	0.0552***	0.0574***
	(0.00882)	(0.00912)	(0.0110)	(0.0115)
Panel E: top 10% hourly wage				
Job posts per 1,000 inhabitants	0.296***	0.283***	0.332***	0.327***
	(0.0453)	(0.0473)	(0.0591)	(0.0655)
N	2354	1590	2356	1590
1st-stage F	239.67	371.62	239.67	371.62
MSA FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES
Covariates time		YES		YES
Matched sample		YES		YES

Notes: The results of a difference-in-differences regression of aggregate migration and wages on the number of CL posts per 1,000 inhabitants, set to 0 for non-CL cities. Each coefficient corresponds to a different diff-in-diff regression. All specifications include city and time fixed effects; columns (2) and (4) also control for time FE interacted with baseline Internet availability, and are estimated on a balanced matched sample. The IV columns use the normalised number of personal posts as an instrument for the number of job posts. Standard errors are in parentheses, clustered at city level throughout. $*p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

popular city (Kroft and Pope, 2014). Furthermore, Craigslist was just an online job board: it also carried sales, housing and personal ads. As the next step, I thus estimate the effect of the actual number of CL job posts (normalised per 1,000 inhabitants) on migration inflows and outflows. Because job posting might be endogenous to local labour market conditions, I instrument it using a normalised measure of personal posts put on CL within the city at the same time. The resulting estimates complement the CL entry estimates by measuring the average causal response of city-to-city mobility to a unit increase in the number of CL *job* posts.

The coefficients are presented in panels (A) and (B) in Table 3. I estimate the average causal response to be about 400 individuals, i.e. one extra job post per 1,000 inhabitants per month increases the number of migrants in and out of the city by about 400 annually. This magnitude is in line with the treatment effects of CL entry estimated in Table 2 (313 for outflows and 960 for inflows in the baseline specification). With the average city recording 1.18 job posts per month, these estimates translate into effect size of 2%, i.e. for every 100 CL job posts per year, 2 additional workers will move into the city on average .

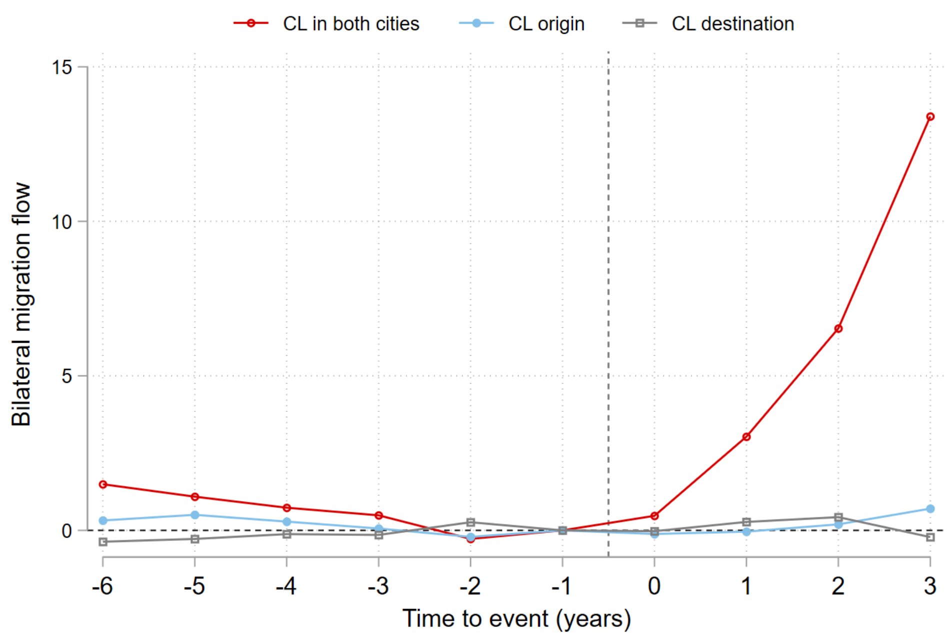
Overall, these results consistently show that Craigslist increased both inflows and outflows of workers from other US cities. The scale of the net effect varies somewhat across specifications, but it is positive and weakly significant across most specifications; in my preferred specification in column (1), the average city population increases by about 0.07% annually. Such an increase, while large as a reaction to a website entry, doesn't correspond to a substantial change in city size. Instead, Craigslist increased migration *churn*, helping more people to move in but also to leave.

I explore this result further by analysing the impact of CL entry on bilateral city-to-city migration flows. A disadvantage of studying aggregate flows is that the total flow in or out of city doesn't just depend on the city characteristics, but also on the conditions of all other potential destination cities. Analysing bilateral flows circumvents this issue by holding the origin and destination city fixed. I assign each city-to-city flow to one of four categories, depending on whether it was the origin city, the destination city, or both that were treated by CL entry. I then estimate the standard dynamic difference-in-differences specification (1), using bilateral flows where neither the origin nor the destination city were treated as control.[37] I plot the results in Figure 3. They show that the increase in mobility due to CL entry is entirely driven by an increase in the flows between CL cities.[38] This

[37]In line with my preferred specification, I run this analysis on a sample of cities matched on baseline covariates. A bilateral flow is included in this balanced sample if both cities are included. The regression controls for time, origin and destination fixed effects, and includes time FE interacted with baseline internet availability in the destination and origin city. In the "both-treated" category, I drop the year when one of the cities was treated first.

[38]Not all bilateral migration flows are non-zero. While this is not an issue in my estimation per se since I

Figure 3: Impact of CL entry on city-to-city (bilateral) migration flows



Notes: Event study of bilateral migration flows (in levels) on CL entry. The control group for all 3 treatment groups are bilateral flows that were never treated (i.e. both destination and origin cities didn't have CL in this time period). The destination and origin cities for all bilateral flows are balanced on baseline covariates. The regression specification includes origin and destination FE, time FE, and baseline Internet access interacted with time dummies, in line with the preferred specification for aggregate migration flows. For confidence intervals, see the Appendix.

result rationalises why CL entry increases both inflows into the city and outflows from the city. Higher inflows are the result of a direct effect of online recruitment becoming available. The higher migration outflows are the consequence of the direct (inflow) effect in other CL cities since workers in CL cities rely on CL in other cities to help them find jobs. In other words, CL entry into one city increases the usage of CL in other cities, too, contributing to

measure migration flows in levels (rather than logs), it does raise the question of whether a linear regression is the best fit for the underlying data-generating process. I explore this question in Figures A7 and A8 in the Appendix. In Figure A7, I show that 2.3% of all possible city-to-city flows are greater than 0. While this share is much higher among CL-to-CL flows than bilateral flows where only one or neither city is treated, it is also almost constant over time: only 0.09% of flows change from zero to non-zero (or back) year on year. This highly constant nature of bilateral flows, combined with the flow fixed effects included in the baseline dynamic diff-in-diff, means that a switching from and to 0 flows is not an important part of my findings. However, the relatively large overall share of zero flows does suggest that a Poisson regression might be more appropriate. I present the estimates in panels (b), (d) and (f) of Figure A8. In the other three panels of this Figure, I present the estimates of a linear dynamic diff-in-diff, restricting the sample to positive flows only. While the standard errors are relatively large for both specifications, they also estimate a positive impact of CL entry on city-to-city migration flows, especially for flows where both the destination and origin city are treated.

migration churn within CL cities.

4.2 Impact on wages

I repeat the dynamic difference-in-differences analysis of Craigslist entry for city-level wages. I focus on three moments: mean, top decile and bottom decile of nominal hourly wages, which allows me to analyse whether CL entry – and the subsequent increase in geographic mobility – had an impact on local wage distribution.

The raw data, plotted in the three left-hand panels of Figure 4 shows that wages in CL and non-CL cities were on different paths before CL expansion. Specifically, wages in non-CL cities were growing faster in the early 2000s, converging to wages in the (often larger) to-be-CL cities. CL entry put an end to this convergence, reversing the trend especially, in the upper half of the wage distribution.

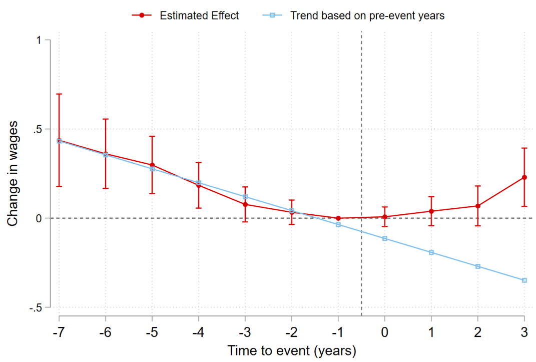
As the next step, I estimate my preferred specification regression, in which I impose the control and treatment cities to be comparable at the baseline, and allow for different time fixed effects for cities with different baseline internet availability. I also allow for a linear trend in wages of CL cities to take into account the convergence pattern in the pre-treatment period. The results are plotted in the right-hand side panels of Figure 4. The confirm the patterns in the raw data: Craigslist caused a relative increase in wages in the cities it entered, especially at the mean and top decile of the distribution. The annual increase was 0.9% for the bottom decile and 1.2% for the top decile of wages.

In Tables A7 and 4, I summarise the results of my robustness checks and alternative identification strategies. I employ different de-trending (city-specific linear trends, column (2)), impose common support in city size (column (4)), expand the sample to include early CL entry before the year 2004 (column (6)), and re-estimate the regression in logs (column (5)). I also identify the treatment effect off of the staggered timing of CL entry only (column (3)) and use eBay's Kijiji cities to construct an alternative, externally-determined control group (column (7)). Throughout, these results confirm that CL entry had a significant impact on city wages. Depending on the specification, the treatment effect on average and bottom decile wages sometimes turns insignificant or weakly negative, highlighting the fact that the CL impact was the largest for the highest-paid jobs. As an additional robustness check, I re-estimate the baseline regression using an alternative dynamic DiD estimator by Callaway and Sant'Anna (2021). I present the results in Figure A9, showing that the wage results are robust to allowing for treatment heterogeneity.

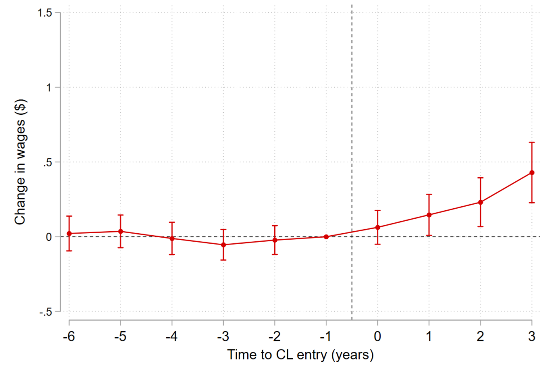
Finally, in panels C, D and E of Table 3, I examine how city-level wages reacted to the *intensity* of treatment as measured by the number of job posts per 1,000 inhabitants in a

Figure 4: Impact of CL entry on city-level hourly wages

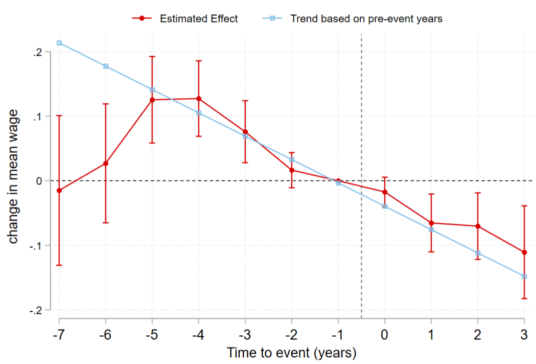
(a) Mean wage, all cities



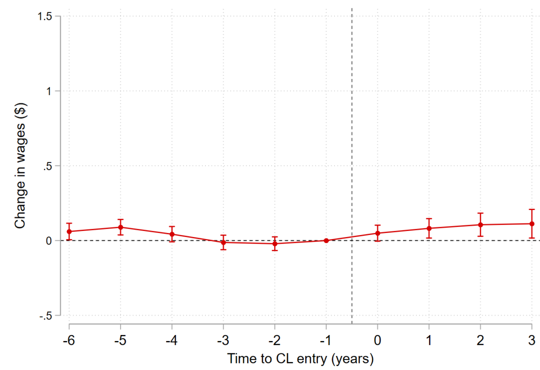
(b) Mean wage, preferred specification



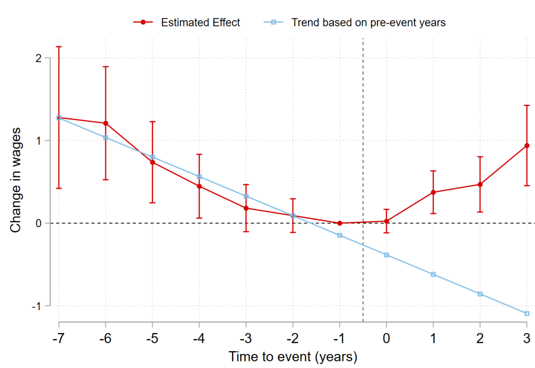
(c) Bottom 10%, all cities



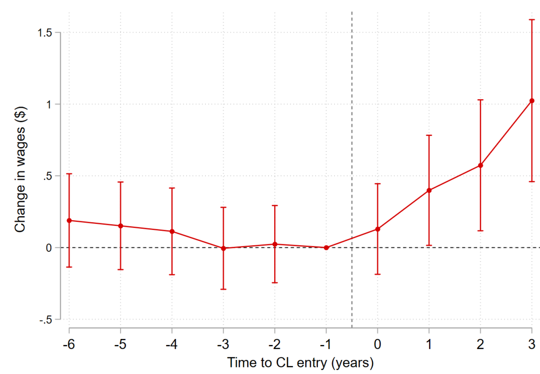
(d) Bottom 10%, preferred specification



(e) Top 10%, all cities



(f) Top 10%, preferred specification



Note: Wages are city-level averages of log hourly wage. "All cities" corresponds to raw event study with MSA and time FE on all cities. The blue line corresponds to linear trend extrapolated from the pre-treatment observations. "Preferred specification" corresponds to event study with city-specific linear trend. This corresponds to diff-in-diff in column (5) in Table 3.

given city. The estimated average causal response is positive: one additional job post per 1,000 inhabitants increases the average wage by about $0.1 per hour (0.7%), the bottom

10th percentile wage by \$0.05 per hour (0.8%), and the wage at the top 10th percentile by \$0.3 (1.1%). These results suggest that the wage impact of CL entry was driven by the cities where CL became popular and widely used, supporting the hypothesis that CL entry increased wages by facilitating a different search and matching in the labour market.

Overall, the results of this section suggest that the availability of a new tool for recruitment and job search caused greater geographic churn between the treated cities, and led to a significant increase in local wages, especially at the top of the wage distribution. In the rest of the paper, I explore the hypothesis that these two results are causally related and explore the underlying mechanisms.

Table 4: Impact of Craigslist entry on average hourly wages, robustness checks

	Balanced sample				All data		Kijiji cities
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Levels	Levels	Levels	Levels	Log	Levels	Levels
Panel A: Mean hourly wage							
Treatment effect at t+0	0.0630*	0.0482	-0.0595*	0.0790**	0.00222	0.0393	0.0256
	(0.0372)	(0.0395)	(0.0321)	(0.0393)	(0.00181)	(0.0328)	(0.0392)
Treatment effect at t+1	0.147**	0.110**	-0.0628	0.169**	0.00540*	0.105*	0.105
	(0.0656)	(0.0501)	(0.0429)	(0.0669)	(0.00280)	(0.0535)	(0.0718)
Treatment effect at t+2	0.231**	0.151**	-0.0659	0.247**	0.00752*	0.156**	0.186*
	(0.0964)	(0.0647)	(0.0444)	(0.0982)	(0.00394)	(0.0772)	(0.110)
Treatment effect at t+3	0.430***	0.236**	0.0194	0.473***	0.0114**	0.298**	0.341**
	(0.152)	(0.0926)	(0.0699)	(0.161)	(0.00539)	(0.120)	(0.163)
Panel B: Bottom 10% hourly wage							
Treatment effect at t+0	0.0492***	0.0633***	0.0200	0.0582***	-0.00125	-0.00506	-0.0261
	(0.0142)	(0.0211)	(0.0148)	(0.0151)	(0.00192)	(0.0126)	(0.0159)
Treatment effect at t+1	0.0817***	0.116***	0.0387*	0.0971***	-0.00561	-0.0286	-0.0740**
	(0.0279)	(0.0267)	(0.0230)	(0.0297)	(0.00384)	(0.0254)	(0.0330)
Treatment effect at t+2	0.106**	0.169***	0.0498**	0.128***	-0.00553	-0.0343	-0.0978**
	(0.0415)	(0.0345)	(0.0234)	(0.0435)	(0.00511)	(0.0343)	(0.0470)
Treatment effect at t+3	0.112*	0.221***	0.0382	0.157**	-0.0121*	-0.0642	-0.137**
	(0.0637)	(0.0494)	(0.0301)	(0.0692)	(0.00682)	(0.0489)	(0.0644)
Panel C: Top 10% hourly wage							
Treatment effect at t+0	0.129	0.0717	-0.189**	0.166*	0.00284	0.0723	0.0972
	(0.0887)	(0.109)	(0.0731)	(0.0921)	(0.00281)	(0.0860)	(0.111)
Treatment effect at t+1	0.399**	0.254*	-0.162	0.446***	0.0150***	0.417**	0.556**
	(0.164)	(0.138)	(0.101)	(0.162)	(0.00474)	(0.164)	(0.224)
Treatment effect at t+2	0.573**	0.262	-0.229**	0.618**	0.0187***	0.555**	0.845**
	(0.243)	(0.178)	(0.102)	(0.239)	(0.00655)	(0.232)	(0.331)
Treatment effect at t+3	1.024**	0.382	-0.106	1.100***	0.0267***	0.939**	1.277***
	(0.395)	(0.255)	(0.175)	(0.410)	(0.00905)	(0.364)	(0.483)
N	2129	2129	1620	1919	3046	3046	1892
MSA FE	YES	YES	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES	YES	YES
Linear treatment trend	YES			YES	YES	YES	YES
Linear MSA trend		YES					
Covariates × time	YES			YES		YES	YES
CL cities only			YES				
City size common support				YES			
All CL entries						YES	

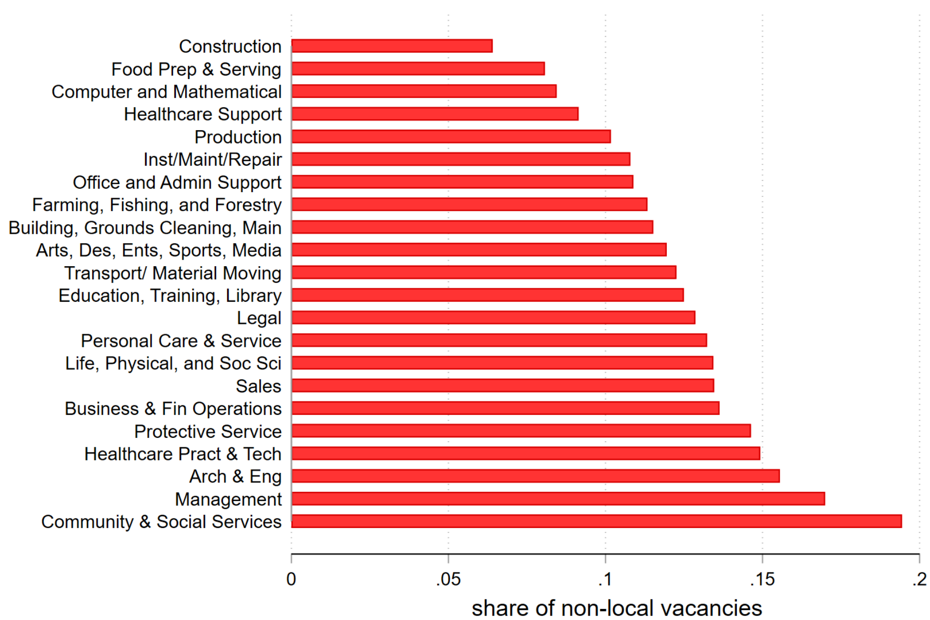
Note: This table summarises the results of a dynamic diff-in-diff version of the robustness checks in Table A7. See the note for more details. Standard errors are in parentheses, clustered at city level throughout. $*p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

5 Understanding firm recruitment

One possible explanation for the simultaneous increase in migration and wages in cities following Craigslist entry is that it was driven by greater sorting and matching across space. If the sudden increase in availability of online recruitment allowed firms to hire more cheaply from other cities, we would expect more city-to-city moves and higher city-level wages, especially at the top end of the wage distribution – in line with the estimates from the previous section. A key assumption underlying this explanation is that firms use online recruitment to find specific talent or better matches outside of their local labour market. Whilst this may sound obvious, it is not necessarily true: the internet may instead be used to access larger pools of unemployed to hire faster and at lower wages. Understanding why firms recruit across space is thus crucial for evaluating the hypothesis that Craigslist improved spatial sorting between cities.

In this section, I analyse the novel data on recruitment across space to explain why firms use Craigslist to advertise their vacancies. I document the patterns in non-local recruitment and show that non-local hiring is predominantly driven by the search for specific skills, rather than to hire cheaply or more quickly.

Figure 5: Share of non-local recruitment in 1990, by occupation



Note: Each bar represents the share of newspaper help-wanted ads advertised in a newspaper not located in the same town as the job (employer).

Table 5: Direction of non-local recruitment

	(1) Non-local hiring	(2) Occup. share in the hiring market
Local occupation employment share	-1.393***	-0.0864**
	(0.327)	(0.0304)
Constant	0.0341***	0.00320***
	(0.000891)	(0.0000670)
Observations	1243	1550
R^2	0.020	0.078

Note: Regressions analysing the placement of help-wanted ads across different cities. Column (1): regression of the share of non-local hiring at occupation level (dependent variable) on the employment share of the given occupation in the location of the job. Column (2): regression of the employment share of an occupation in the location of recruitment (dependent variable) on the employment share of the given occupation in the location of the job for non-local vacancies. Standard errors in parentheses, clustered at occupation level. $*p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Patterns in recruitment across space The newspaper vacancy data shows that while the majority of job postings in 1990 were local, a non-negligible share was not. 11% of vacancies were posted in a newspaper located in a different town, and about a half of these non-local vacancies were plausibly advertised in another local labour market: 5% (of all vacancies) were posted in another county, and about 2% were advertised outside of the commuting zone (a map of the average recruitment distance for each newspaper in the sample is plotted in Figure A3).[39]

Non-local recruitment varied significantly across occupations. As I show in Figure 5, the most local occupation was construction, with less than 6% of vacancies advertised outside of its local town. In the most widely recruited occupations (architects and engineers, managers, and social service workers), the non-local shares were up to three times as large. The heterogeneity across occupations is the first piece of evidence of the drivers of cross-regional recruitment: it is mostly higher-wage, higher-education occupations that are likely to be advertised in a distant newspaper.

Drivers of recruitment across space To explore the drivers of non-local recruitment, I compare, within each help-wanted ad, the labour market characteristics of the job location and newspaper location. I calculate a percentage difference in the local average wages and their growth, the size of the labour force, the competitiveness of the labour market

[39]Interestingly, these figures are comparable to the annual migration rates in 1990 across counties (5-6%), metropolitan statistical areas and state boundaries (both around 3%) (Molloy, Smith, and Wozniak, 2011).

(as measured by the number of establishments) and the percentage point difference in local unemployment rates. If firms recruit non-locally primarily to tap into a large pool of cheap labour, they would post their help-wanted ads in markets with lower pay, higher unemployment rate, and fewer competing establishments. If they are using non-local recruitment to find specific talent, we should instead find that they advertise in higher-pay markets with a large labour force.

I plot the labour market comparisons in panel (a) of Figure 6. The main determinant of where firms post their vacancies is the size of the labour market: recruitment markets are about 3 percentage points larger, both in terms of the labour force and the number of establishments than the labour markets where the job is located. However, firms don't seem to be drawn to large pools of unemployed: they tend to advertise in labour markets with weakly lower unemployment rates than those in their local labour market. Pay doesn't seem to matter much either: firms advertise in markets with somewhat higher pay (but lower wage growth). Put together, these patterns suggest that firms recruit non-locally to access larger markets rather than cheap abundant labour.

However, these patterns don't hold across all occupations. In panel (b) of Figure 6, I focus on the differences in average pay, and in panel (c) on the differences in unemployment rates, calculated separately for each occupation.[40] Both figures show significant variation across jobs. For some occupations, such as Building and Grounds Maintenance, and Food Preparation and Serving, employers do target higher-unemployment, lower-wage regions when recruiting cross-regionally. In other occupations, e.g. Social Services and Business and Financial Operations, the placement of non-local vacancies is the opposite, in low-unemployment-high-pay areas. Comparison with Figure 5 shows that it is the occupations least likely to recruit non-locally which focus on markets with lower wages and higher unemployment. This explains the overall pattern in panel (a): the majority of non-local recruitment happens in large labour markets, regardless of their pay.

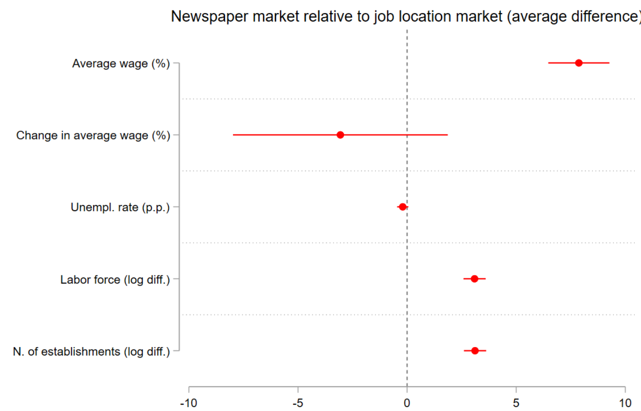
If most non-local recruitment focuses on finding workers with specific skills, we should also see that firms direct their job posts to places where such skills are more abundant. I test this by estimating the relationship between non-local recruitment and the employment share of the vacancy's occupation in the firm's local labour market. Firms that are hiring for an occupation that is relatively abundant in their local labour market (as measured by the occupation's employment share) should be less likely to recruit non-locally. The results in column (1) of Table 5 support this hypothesis: the higher the occupation's share

[40] Note that the data on pay and unemployment rates is still at the county (rather than county-occupation) level. These two panels thus analyse how the pattern of posting varies by occupation, but it doesn't take into account the within-county differences in labour market characteristics.

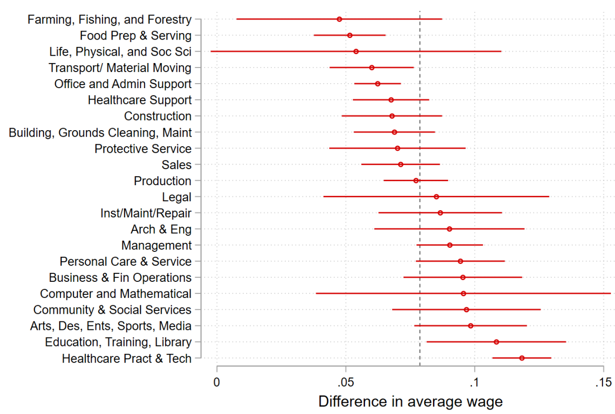
of employment in the local labour market, the less likely is a vacancy posted elsewhere. In column (2) of the table, I go one step further and estimate the relationship between the occupation employment share in the local market and the occupation employment share in the recruitment market, conditional on hiring non-locally. I find that this relationship is also negative: the lower the availability of a particular occupation in the local labour market, the higher it is in the market where the vacancy is posted. In other words, firms post their vacancies in labour markets that specialise in the occupation they are recruiting for.

Figure 6: Differences in labour market conditions between the location of the job and the location of recruitment, for non-local vacancies

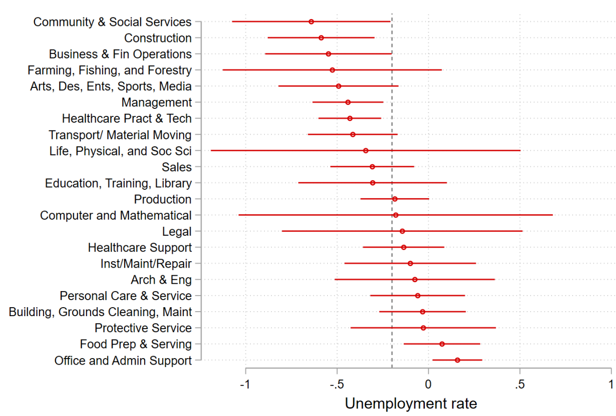
(a) Average differences across occupations



(b) Difference in county-level average pay, by occupation



(c) Difference in county-level unemployment rate, by occupation



Note: Unconditional differences in labour market characteristics at the county of the job location and the county of the newspaper location for no-local vacancies. Panel (a): the difference in average hourly wage, annual growth in hourly wage, the size of the labour force, the unemployment rate, and the number of establishments. Panel (b): the difference in average hourly wage by major occupation groups. Panel (c): the difference in unemployment rate by major occupation groups. The line corresponds to 95% confidence interval.

6 Impact of CL on labour market outcomes: heterogeneity by occupation

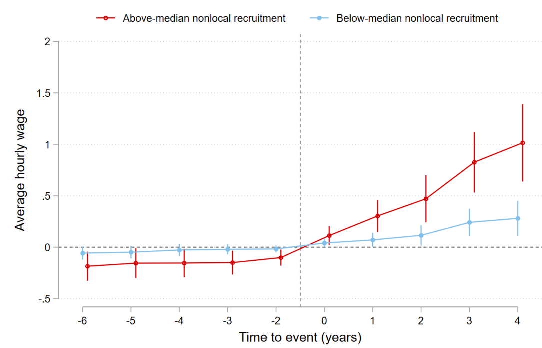
The previous section of this paper has shown that when employers recruit non-locally it is to find workers of specific skill rather than to access large pools of cheap labour. As a consequence, if Craigslist effectively lowered the costs of recruiting non-locally, we should expect an increase in migration churn and wages in the local markets it has entered – exactly in line with the treatment effects estimated in Section 4. However, this hypothesis generates several other testable predictions, which I examine in this section. First, I confirm that the increases in geographic mobility and wages are causally related: the markets with the largest increase in migration churn are also the markets where wages grew the most. Second, I show that the impact of CL entry is greatest in occupations where the value of the match matters most, i.e. in occupations that were recruiting non-locally before the introduction of the internet. Put together, these findings provide evidence on *why* Craigslist entry changed the geography of local labour markets.

Occupation-specific impact on wages I start by showing that the baseline results of CL entry on wages also hold within cities and across occupations. I estimate the baseline regression (1) using city-occupation-specific wages as the dependent variable, progressively adding city- and occupation- fixed effects, and controlling for city- and occupation-specific linear trends in wages. The results are summarised in Table A8. Column (1) contains time- and occupation-fixed effects only, estimating the impact of CL entry on wages across different cities within an occupation. This specification demonstrates that the increase in average city wages isn't driven by a compositional change in workers at the city level: the positive impact of CL entry holds within each occupation. In the other four columns of the table, I also control for city fixed effects. This allows me to compare the impact of CL entry across occupations within a single city, effectively controlling for any general increases in wages at the city level. In column (3), I allow for occupation-specific linear growth in wages, and in column (4), I also add city-specific linear wage growth. Finally, in column (5), I turn to one of the alternative identification strategies in which I use Kijiji cities as an externally determined control group. The estimated effects of CL entry are consistently positive and significant, especially 1-2 years after entry.

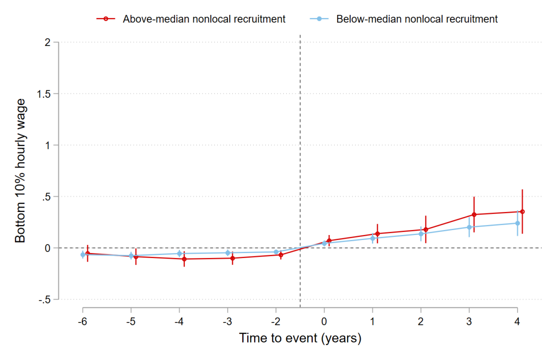
As the next step, I explore the differences in treatment effects between occupations. I split occupations into 2 groups, low and high non-local recruitment, depending on whether its pre-internet share of non-local recruitment in newspapers was below or above the median. There are two ways to interpret this categorisation: jobs that are recruiting cross-regionally at the

Figure 7: Impact of CL entry on wages, by non-local hiring in 1990

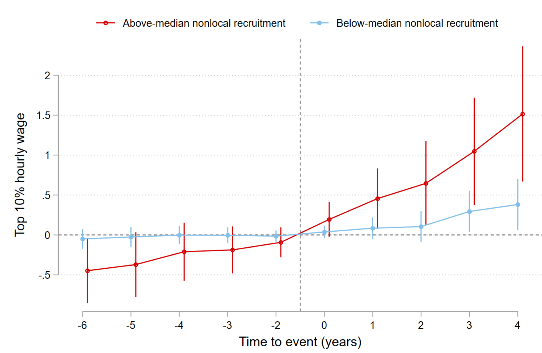
(a) Average hourly wage



(b) Bottom 10% hourly wage



(c) Top 10% hourly wage



Notes: The results of an event study regression for city-occupation-specific wages. The regressions include city-, time- and occupation- fixed effects, as well as occupation- and treatment-group- specific linear trends. The sample consists of treated and control cities matched on baseline covariates. Non-local hiring is measured at occupation level from the dataset of newspaper help-wanted ads in 1990. Standard errors are clustered at occupation and city level. $*p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

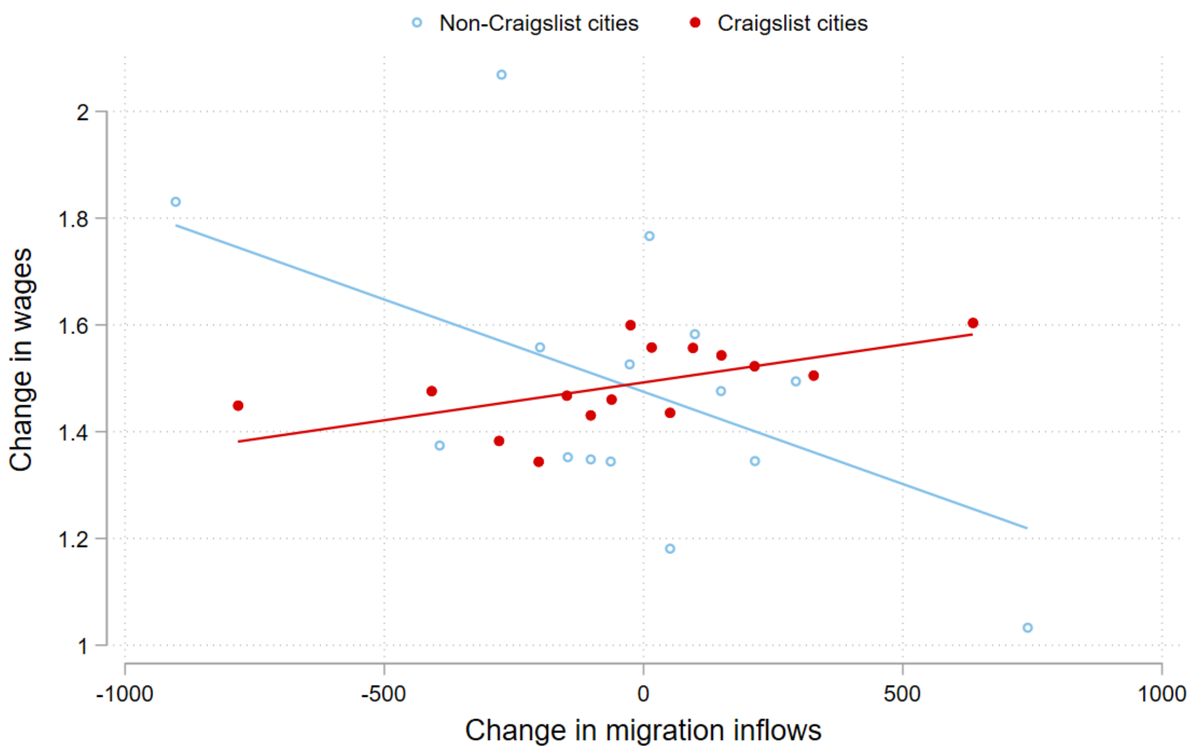
baseline might be more prone to switching to new, cheaper technology (such as Craigslist); or we can use non-local recruitment share as a proxy for the relative importance of match value. Either way, we would expect the occupations with high non-local recruitment before the introduction of the internet to react more strongly to Craigslist entry. To test this hypothesis, I re-run the baseline dynamic diff-in-diff regression separately for high- and low-non-local recruitment occupations. I plot the results in Figure 7. The top panel, focusing on the average wage, shows that while Craigslist entry increases the average wage in all occupations, this increase is significantly more pronounced in occupations with relatively high non-local recruitment share. By year 4 after treatment, wages in high non-local recruitment jobs have increased about three times more than wages in low non-local recruitment occupations. The other two panels of Figure 7 demonstrate that this average difference between the two groups is driven almost entirely by wages at the 90th percentile of city-occupation wages. Panel (b), which replicates the analysis for the bottom decile of the wage distribution within each occupation and city, finds virtually no difference – wages in both groups increase by the same amount. On the other hand, at the top decile (panel (c)), CL entry causes wages in the high non-local recruitment occupations to increase by much more than the top wages in occupations that historically recruited relatively more locally. This result – that not only is the impact of CL greater in occupations that recruit more non-locally in general but that this impact is concentrated at the top of the wage distribution – further supports the hypothesis that online recruitment technologies, such as Craigslist, facilitated greater spatial sorting and better matching across space.

Occupation-specific impact on geographic mobility Ideally, I would like to replicate my analysis of CL impact on migration inflows and outflows at the city-occupation level. Unfortunately, this is not possible due to the lack of migration data at this level of detail. The IRS data I use for my baseline city-level analysis doesn't include any demographic information about the movers, and the alternative – survey data from the American Community Survey – only started to record city-level moves in 2005, after Craigslist entry in a large number of cities. As a result, there is no "before" data for occupation-specific moves that would allow me to replicate my baseline dynamic diff-in-diff specification.[41]

In the absence of detailed time series data, I focus on analysing the relationship between the changes in migration flows and wages *after* CL entry. I use ACS data to calculate the

[41] Technically since CL entered some cities in 2006, I could estimate the baseline diff-in-diff regressions for the subgroup of these cities. While feasible in theory, there are 2 problems with this approach. First, because ACS is a survey of 1% US population, it doesn't capture movers of each occupation into each city every year; in other words, the panel is unbalanced, and data for both years is available for only about 1/6 of all the possible city-occupation cells. The second problem is that the data starts in 2005, giving us only 1 year before treatment.

Figure 8: Relationship between changes in wages and gross migration inflows at occupation level within Craigslist and non-Craigslist cities



Notes: Binscatter of changes in city-occupation-specific wages and city-occupation-specific migration inflows. The changes are calculated between years 2005-2007. The specification includes city- and occupation- fixed effects. Red line and filled circles: Craigslist cities. Blue line and empty circles: non-Craigslist cities.

number of workers of a particular occupation moving into a given city annually between 2005 and 2007, during the treatment period studied in my baseline analysis. I calculate the change in inflows over these three years within each city-occupation cell and compare it to an analogous change in wages over the same time period.[42]This comparison allows me to test whether the occupations that experienced the largest migration inflows are also the occupations that saw the largest increase in wages. Put differently, if it is indeed greater spatial sorting that caused the positive impact of CL on wages, then we should see a positive relationship between these two outcomes.

In Figure 8, I present a binned scatterplot of post-CL entry changes in migration flows and wages at city-occupation level. The graph shows that, in line with the prediction, the relationship between the two labour market outcomes is positive for CL cities: the occupations that experienced the largest increase in migration inflows are also the occupations that saw the largest increase in wages. Importantly, this pattern does not exist within control (non-CL) cities: high inflow of workers within an occupation led to relatively *smaller* increase in wages.

I test this relationship formally by estimating the following regression of wages on migration:

$$\Delta Wage_{co} = \alpha + \beta \Delta Migration_{co} + \gamma \Delta Migration_{co} CL_c + \lambda_o + \kappa_c + \epsilon_{co} \tag{3}$$

where Δ corresponds to the 2007-2005 difference in the given variable, CL is a binary treatment dummy treatment, and λ_o and κ_c are occupation- and city-fixed effects, respectively. The coefficient of interest is γ which allows the relationship between the changes in migration and the changes in wages to vary between treated and control cities. I estimate it to be positive and statistically significant (0.0004812, with a p-value of 0.001), while β, the correlation in control cities is significantly negative (-0.00034, p-value of 0.007). (I plot the estimated relationship between migration and wages, and report the full regression results in Figure A10.) These regression results confirm the patterns shown in the scatterplot above. They show that the impacts of CL on migration and wages are likely causally related. Furthermore, the negative relationship in control cities is consistent with the standard model of labour supply and demand, in which higher labour supply from particular workers causes their wages to relatively decline. Viewed through this lens, the positive relationship for treated cities suggests that the changes in wages observed there were driven by an increase in labour demand, rather than supply – driven, for example, by a more efficient spatial matching technology.

[42]The wage data is the same as in the previous section, taken from the OEWS dataset.

7 Discussion and conclusion

In this paper, I show that the internet didn't just make it easier to search, it also substantially reduced the effective distance between geographically separate labour markets. The introduction of one of the first – and, at one time, the most popular – online job boards in the US led to a significant increase in migration flows in and out of the affected cities. Instead of increasing city population, Craigslist increased migration churn, helping workers sort between different locations. As a result, pay went up, especially at the top of the wage distribution.

I uncover these patterns using a combination of different identification strategies and a novel data set on recruitment. I primarily estimate the impact of Craigslist from the variation in its expansion across time and space, but I complement my analysis with a comparison of Craigslist cities to those entered by an unsuccessful rival online job board. I also make use of the data on Craigslist's actual use to strengthen my claim that the impact of Craigslist on the labour market was predominantly due to its help-wanted ads rather than housing or personal classifieds. To understand the mechanism behind Craigslist's impact, I collect the first data set describing how firms recruit across space. I show that the motivation for cross-regional hiring is to find workers of specific skill rather than to access a large pool of unemployed labour. As a result, when the introduction of online recruitment made non-local hiring cheaper, it was the occupations that historically hired this way the most that experienced the largest increase in migration and wages.

However, the results of this paper come with several caveats. First, Craigslist offers relatively rudimentary online recruitment tools compared to what is on offer today: it allows employers to advertise vacancies online, but it doesn't help to match them with potential hires, nor does it allow employers and candidates to directly communicate or manage the early stages of the recruitment process. As such, the estimated effect of Craigslist might be different from the impact of more sophisticated, AI-powered online recruitment tools.[43] Second, my detailed heterogeneity analysis is just a proxy for directly observing the search and recruitment methods of the moving workers and the changes in the wages of movers. Third, online recruitment is likely to have far-reaching general equilibrium effects, not just within the labour market but across the wider economy. For example, Steffen Altmann and Sebald (2023) document significant negative spillovers of job-search advice in Denmark, and

[43] In particular, the rise of hiring online has been accompanied by an anecdotal increase in the number of applications to each job, and per each job seeker. This simultaneous increase might lead to significant congestion effects in the market, negating much of the improvement thanks to the higher efficiency of the matching function. There is also a large emerging literature on the consequences of algorithm-based selection in recruitment, see for example Hoffman, Kahn, and Li (2017).

Atasoy (2013) shows how the expansion of broadband availability across the US increases firm size. My identification strategy was specifically designed to exclude such general equilibrium effects.

Overall, while the results of this paper show the significant ways in which technology can have a positive impact on the labour market, they also contain some reasons for caution. First, even though the movers in my analysis most likely benefited from relocation and higher pay, the overall wage and regional inequality increased. Policymakers who value regional convergence per se may need to think about ways to harness the efficiency gains of online recruitment while minimising its negative impacts on inequality.

Second, it is unclear why this technology, offering a near-costless advertisement across much larger labour markets, hasn't been adopted universally. Carnevale, Jayasundera, and Repnikov (2014) estimate that about 70% of all vacancies in the US are posted online, which means that about a third of all job openings are not; low-wage, low-skill vacancies are particularly under-represented. This gap in coverage has potentially serious consequences for the equality of opportunity and geographic mobility, but the question of why certain types of employers avoid online recruitment remains so far unexplored in the literature. Given the growth in working from home at the opposite end of the job spectrum and the opportunity it offers to workers to free themselves from the tyranny of geography, the importance of this and related questions about the interaction between the spatial frictions and technology will only increase.

References

Katharine G. Abraham and Michael Wachter. Help-wanted advertising, job vacancies, and unemployment. *Brookings Papers on Economic Activity*, 1987(1):207–248, 1987.

Abigail Adams-Prassl, Maria Balgova, Tim Waters, and Matthias Qian. Flexible Work Arrangements in Low Wage Jobs: Evidence from Job Vacancy Data. IZA Discussion Papers 13691, Institute of Labor Economics (IZA), 2020.

L. Jason Anastasopoulos, George J. Borjas, Gavin G. Cook, and Michael Lachanski. Job vacancies and immigration: Evidence from the mariel supply shock. *Journal of Human Capital*, 15(1):1–33, 2021.

Enghin Atalay, Phai Phongthiengtham, Sebastian Sotelo, and Daniel Tannenbaum. New technologies and the labor market. *Journal of Monetary Economics*, 97:48 – 67, 2018.

Enghin Atalay, Phai Phongthiengtham, Sebastian Sotelo, and Daniel Tannenbaum. The evolution of work in the united states. *American Economic Journal: Applied Economics*, 12(2):1–34, April 2020.

Hilal Atasoy. The effects of broadband internet expansion on labor market outcomes. *ILR Review*, 66(2):315–345, 2013.

Maria Balgova. Moving to look or looking to move? The role of job search in migration decisions. Working paper, 2022.

Christopher R. Berry and Edward L. Glaeser. The Divergence of Human Capital Levels Across Cities. NBER Working Papers 11617, National Bureau of Economic Research, Inc, September 2005.

Manudeep Bhuller, Andreas Kostol, Domenico Ferraro, and Trond Vigtel. The internet, search frictions and aggregate unemployment. Working paper, 2022.

Vera Brenčič. The impact of craigslist's entry on competing employment websites. *IZA Journal of Labor Economics*, 5(1):1–15, 2016.

Tomaz Cajner and David D. Ratner. A cautionary note on the help wanted online data. Working paper, Washington: Board of Governors of the Federal Reserve System, 2016.

Brantly Callaway and Pedro H.C. Sant'Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230, 2021. Themed Issue: Treatment Effect 1.

Brantly Callaway, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. Difference-in-differences with a continuous treatment, 2021.

Anthony Carnevale, Tamara Jayasundera, and Dmitri Repnikov. Understanding Job Ads Data. Technical report, McCourt School of Public Policy, Georgetown University, 2014.

Jeffrey Clemens, Lisa B. Kahn, and Jonathan Meer. Dropouts Need Not Apply? The Minimum Wage and Skill Upgrading. *Journal of Labor Economics*, 39(S1):107–149, 2021.

Guido Matias Cortes, Nir Jaimovich, and Henry E. Siu. The "End of Men" and Rise of Women in the High-Skilled Labor Market. NBER Working Papers 24274, National Bureau of Economic Research, Inc, February 2018.

Guido Matias Cortes, Nir Jaimovich, and Henry E. Siu. The growing importance of social tasks in high-paying occupations. 58(5):1429–1451, 2023.

Scott Cunningham, Gregory DeAngelo, and John Tripp. Did craigslist's erotic services reduce female homicide and rape? 2023.

Wolfgang Dauth, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum. Matching in Cities. *Journal of the European Economic Association*, 20(4):1478–1521, 02 2022.

David J Deming and Kadeem Noray. Earnings Dynamics, Changing Job Skills, and STEM Careers*. *The Quarterly Journal of Economics*, 135(4):1965–2005, 06 2020.

David J Deming and Kadeem L Noray. Stem careers and the changing skill requirements of work. Working Paper 25065, National Bureau of Economic Research, September 2018.

Manuel Denzer, Thorsten Schank, and Richard Upward. Does the internet increase the job finding rate? evidence from a period of expansion in internet use. *Information Economics and Policy*, 55:100900, 2021.

Milena Djourelova, Ruben Durante, and Greg Martin. The impact of online competition on local newspapers: Evidence from the introduction of craigslist. Working paper, 2022.

Thomas Fujiwara, Eduardo Morales, and Charly Porcher. A Revealed-Preference Approach to Measuring Information Frictions in Migration. Working paper, 2021.

Peter Ganong and Daniel Shoag. Why has regional income convergence in the u.s. declined? *Journal of Urban Economics*, 102:76–90, 2017.

Elisa Giannone. Skill-biased technical change and regional convergence. Working paper, 2022.

Edward L. Glaeser and Jesse M. Shapiro. City Growth and the 2000 Census: Which Places Grew, and Why. Technical report, The Brookings Institution Survey Series, Washington DC, May 2001.

Nicole Gürtzgen, André Diegmann, Laura Pohlan, and Gerard J. van den Berg. Do digital information technologies help unemployed job seekers find a job? evidence from the broadband internet expansion in germany. *European Economic Review*, 132:103657, 2021.

Brad Hershbein and Lisa B. Kahn. Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings. *American Economic Review*, 108(7):1737–72, July 2018.

Mitchell Hoffman, Lisa B Kahn, and Danielle Li. Discretion in Hiring*. *The Quarterly Journal of Economics*, 133(2):765–800, 10 2017.

Kory Kroft and Devin G. Pope. Does Online Search Crowd Out Traditional Search and Improve Matching Efficiency? Evidence from Craigslist. *Journal of Labor Economics*, 32 (2):259–303, 2014.

Peter Kuhn and Hani Mansour. Is Internet Job Search Still Ineffective? *Economic Journal*, 124(581):1213–1233, December 2014.

Peter Kuhn and Mikal Skuterud. Internet job search and unemployment durations. *American Economic Review*, 94(1):218–232, March 2004.

Alan Manning and Barbara Petrongolo. How Local Are Labor Markets? Evidence from a Spatial Job Search Model. *American Economic Review*, 107(10):2877–2907, October 2017.

Ioana Marinescu and Ronald Wolthoff. Opening the Black Box of the Matching Function: The Power of Words. *Journal of Labor Economics*, 38(2):535–568, 2020.

Raven Molloy, Christopher L. Smith, and Abigail Wozniak. Internal migration in the united states. *Journal of Economic Perspectives*, 25(3):173–96, September 2011.

Enrico Moretti. Local Labor Markets. In O. Ashenfelter and D. Card, editors, *Handbook of Labor Economics*, volume 4 of *Handbook of Labor Economics*, chapter 14, pages 1237–1313. Elsevier, 2011.

Charly Porcher. Migration with Costly Information. Working paper, 2021.

Tyler Ransom. Labor market frictions and moving costs of the employed and unemployed. *Journal of Human Resources*, 57(S):S137–S166, 2022.

Benoît Schmutz and Modibo Sidibe. Frictional Labour Mobility. *The Review of Economic Studies*, 86(4):1779–1826, 09 2018.

Robert Seamans and Feng Zhu. Responses to entry in multi-sided markets: The impact of craigslist on local newspapers. *Management Science*, 60(2):476–493, 2014.

Robert Mahlstedt Steffen Altmann, Anita Glenny and Alexander Sebald. The Direct and Indirect Effects of Online Job Search Advice. Working paper, 2023.

Betsey Stevenson. The Internet and Job Search. In *Studies of Labor Market Intermediation*, NBER Chapters, pages 67–86. National Bureau of Economic Research, Inc, September 2009.

Liyang Sun and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199, 2021. Themed Issue: Treatment Effect 1.

Riley Wilson. Moving to jobs: The role of information in migration decisions. *Journal of Labor Economics*, 39(4):1083–1128, 2021.

A Appendix: Details on the newspaper help-wanted ad data

A.1 Help-wanted ads as a source of labour data

Newspaper help-wanted ads have long been used as a gauge of labour demand. This was primarily driven by the Conference Board Help Wanted Index, which captured the number of vacancies posted in 51 metropolitan newspapers across the US. The HWI index was the leading measure of labour demand for most of the postwar period, from the 1960s until some point in the 2000s when it was made redundant by the massive shift of vacancy posting online. The index was used as a reliable and representative measure of labour demand in research on a wide range of topics, such as the trends in structural unemployment, productivity, and to understand labour demand over the business cycle (Abraham and Wachter, 1987).

However, the HWI does not provide any information on the content of the job vacancies, such as the occupation, sector, task content, or remuneration. While there has been a remarkable growth in studies that make use of vacancy text posted online[44], only 2 studies[45] so far have analysed the text of newspaper vacancies.

The pioneers of textual analysis of newspaper vacancies are Atalay et al. (2020). They extend the vacancy text analysis to pre-internet era, which poses a few new challenges on its own: newspaper text usually does not conveniently separate out individual vacancies and job titles, and may contain other types of text or ads. They clean and process more than 7 million help-wanted ads from three major metropolitan newspapers (*Boston Globe, New York Times, Wall Street Journal*) from the years 1950-2000 to study the evolution of task content between and within occupations. Working with newspaper vacancies allows them to generate data on the task composition of US jobs over a time span not available in other datasets. This new data suggests that most of the switch to nonroutine tasks in the US happened within, rather than between, occupations. Their follow-up paper uses the same newspaper vacancy text to measure the adoption of new technologies and their role in the observed shift towards nonroutine tasks (Atalay et al. (2018)).

Anastasopoulos et al. (2021) use help-wanted ads to evaluate the impact of immigration on labour demand. Revisiting the case of the Mariel Boatlift in 1980, they use the city-specific HWI to run a difference in differences of the immigration shock on newspaper vacancy posting

[44]to examine things such as the evolving task content of occupations, gendered language, and alternative work arrangements [CITE]

[45]In addition, Cortes, Jaimovich, and Siu (2018) have used the data compiled by Atalay et al. (2020) to study the increasing demand for female-oriented skills across occupations in the US. Deming and Noray (2018) use this data to understand changes in tasks related to STEM skills.

in Miami. They also analyse the content of help-wanted ads in the *Miami Herald* to pin down which type of occupations were most hit by the immigration shock. To do this, they first run a machine learning algorithm on the text of the whole newspaper (for a random sample of issues around the Mariel Boatlift) to "separate out" individual vacancies. They then run another algorithm, commissioned by the US Department of Labor, to classify the vacancy text into standard SOC occupation codes. As a result, the researchers are able to show that the relative drop in vacancy posting visible in the HWI data was primarily driven by a decrease in help-wanted ads for blue collar occupations, which were most directly affected by the arrival of Marielitos.

In this paper, I take a third approach. Instead of trying to capture all vacancies posted at a given point in time (or space), and instead of trying to extract all possible information from a smaller newspaper sample, I extract a limited amount of information from a relatively large number of newspapers. This allows me to build a large sample of specific recruitment data without the need for manual transcription or coding. I explain this process in greater detail below.

A.2 Data collection

The source of the help-wanted ads used in this paper is the digital newspaper archive *Newspapers.com*. It is one the largest archives of digitised newspapers, offering more than 20,000 titles from the 18th century until today, mostly from the US. The 600 million+ newspaper pages in its archive are published as scanned images with searchable text. The major advantage of *Newspapers.com* compared to other archives of digitised historical newspapers is that it covers the early 1990s; many archives focus on earlier periods. The Chronicling America project by the Library of Congress, for example, only runs to 1963.

I randomly pick four days in each month of 1990[46]. I use the *Newspapers.com* website to search for all newspapers pages on the given day that contain the phrase "help wanted"[47]. This search results in URL links to 19,500 newspaper pages.

I then ran a Python script which downloads each newspaper page as an image, reads the text in the image (using a Python optimal character recognition (OCR) module) and saves it as a separate text file. The name of each text file stores the newspaper name and the date it was published.

[46] I exclude public holidays but include Sundays, as many newspapers publish on Sundays, and these editions often carry the bulk of the week's help-wanted ads.

[47] I have experimented with other phrases, such as "employment" and "jobs", but these rarely returned any new relevant pages, and contained a lot of false positives in the form of newspaper articles about the economy.

Before I can proceed to analysing the posted vacancies, there are two issues that need to be resolved. First, while most of the downloaded pages contain (at least a part of) the help-wanted section, the keyword search also resulted in a number of irrelevant pages. These are mostly articles about the national or local economy, or pages that contain the newspaper's advertisement for publishing in classified ads (but not the ads themselves). These pages need to be removed from the sample. The second issue are pages that do contain vacancies, but their scans are not of sufficient quality to be transcribed well by the OCR algorithm. The text files of such pages contain a mix of nonalphanumeric characters, gibberish words and real words; the original content cannot be reconstructed.

I remove the irrelevant and gibberish pages using a two-step machine learning algorithm which sorts through the 19,500 newspaper pages in my sample. I first manually classify about 50 pages, paragraph by paragraph[48], into one of three groups: vacancy text, irrelevant text, and gibberish. I then use this data to train two machine learning models: the first one separates out gibberish text from readable text, and the second one identifies pages that contain text (rather than classified ads). I end up with 7,850 newspaper pages containing help-wanted ads only.

A.2.1 Details of the text classification

I use Python's sklearn library for algorithm and parameter selection for both parts of the machine learning model. In training the algorithm, I reserve 30% of the data for testing. Since the classification is performed on the level of paragraphs, I need to establish a cutoff point for determining whether a page is gibberish, text, or help-wanted ads. I do this by comparing the share of content as classified by the model with the same numbers from the manually coded training dataset.

The first machine learning model classifies paragraphs into those which are gibberish and those which are not. I don't do any cleaning or pre-processing of the text at this stage, because procedures such as lemmatising and remove nonalphanumeric characters would get rid of the identifying features of gibberish text. The model makes use of both words and characters as features, in one- and bi-grams. The model uses the bag-of-words method; the relevance of each character is calculated according to its term frequency – inverse term frequency (TFIDF) value, while the relevance of each word is based on its absolute frequency. I use Support Vector Classification (SVC) algorithm with a linear kernel. I classify a page as gibberish if more than 30% of the paragraphs it contains are classified as such.

In the second step, I take the pages classified as real text (not gibberish) by the first model and train a second model to sort them into text and vacancies. For the second

[48]A paragraph here means a section of text divided by a blank line.

machine learning model, I pre-process the text by removing common stopwords (such as "and", "of", or "with"), lemmatising the words, and removing any special characters. The features of the model are constructed from one- and bi-grams of words, again using TF-IDF transformation to verctorise them. I run a logistic regression with Stochastic Gradient Descent learning (SGD Classifier). I classify a page as contains help-wanted ads if less than 10% of its paragraph are classified as other text. These pages will form the sample I will analyse.

There are several ways to assess how well the two machine learning models identify gibberish and non-vacancy pages. The four main statistics are accuracy, precision, recall, and the F1 score. *Precision* measures the probability of not making a type I error. *Recall* captures the probability of not making a type II error. *Accuracy* is a combination of the two, measuring the share of cases that were correctly classified. *F1* is another combination of precision and recall, one that highlights false negatives and false positives. In the case of imbalanced classes, such as here, F1 gives a more accurate assessment about model performance than the more commonly used accuracy score.

$$precision = \frac{\text{true positives}}{\text{true positives} + \text{false positives}} = \frac{\text{true positives}}{\text{total classified as positive}}$$

$$recall = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}} = \frac{\text{true positives}}{\text{total actually positive}}$$

$$accuracy = \frac{\text{true positives} + \text{true negatives}}{\text{total cases}} = \frac{\text{truly classified}}{\text{total cases}}$$

$$F1 = 2 * \frac{precision * recall}{precision + recall}$$

The classification reports for the two machine learning models are presented in Table A1. Both models achieve a relatively high accuracy of 0.84 and 0.88, However, this number masks some differences in the models' performance across the different categories. In the first model, which classifies text into readable text and gibberish, the gibberish category is significantly smaller than that of readable text. This imbalance makes it relatively more difficult for the model to correctly classify gibberish paragraphs: F1 score for the gibberish category is 0.67, compared to 0.9 for the readable text. This difference is primarily driven by recall: 37% of gibberish paragraphs are incorrectly classified as readable text, in contrast to only 8% of readable text being incorrectly classified as gibberish. While this makes it more likely that some unreadable pages are included in my final sample, it reduces the probability of throwing away good pages.

Table A1: Classification report for newspaper page classification models

	model 1			model 2		
	readable text	gibberish	total	other text	vacancies	total
precision	0.88	0.71	0.84	0.43	0.99	0.93
recall	0.92	0.62	0.84	0.9	0.87	0.88
F1	0.9	0.67	0.84	0.58	0.93	0.89
support	894	300	1194	82	770	852
accuracy			0.84			0.88

The second model, which sorts vacancies from other types of text (primarily newspaper articles), has a similar problem with unbalanced sample: help-wanted ads are much more frequent than other types of text. As a result, the model performs very well in classifying vacancies (F1 score is 0.93), but does reatively poorly on other text (F1 score of 0.58). Unlike in the first model, where the main problem was the weak recall of the smaller group, here the problem lies predominantly in the precision of classifying other text. The precision score is 0.43, which means that more than half of paragraphs which are classified as other text are in fact help-wanted ads (false positives). On the other hand, the model performs very well in correctly classifying most of help-wanted ads, and only rarely misclassifies other text as vacancies. This means that the final sample of help-wanted ads has a low share of other text, at the cost of excluding some valid help-wanted ad pages.

A.2.2 Text analysis

Having collected and cleaned the data, I have the text of 7,850 pages of newspaper help-wanted ads to analyse. Ideally, I would like to extract from each published ad the occupation and location of the job. Combined with the information on the location of the newspaper, this will allow me to calculate recruitment distance for each posted vacancy.

The main difficulty with this procedure is identifying the start and the end of each help-wanted ad. The text of each page in my sample is continuous, with page breaks that may or may not correspond to individual vacancies. Alternatively, we may want to rely on some of the patterns in the ads themselves: a vacancy usually starts with the job title advertised, sometimes printed in capital letters, and often ends either in the employer's address, phone number, or phrases such as "send your CV to" or "equal opportunity employer". However, given the large number of sampled newspapers, these patterns are not regular enough for the procedure to work. Furthermore, it would likely result in discarding very short, local vacancies that consist only of a sparse job description and a phone number, which would bias my sample towards longer-distance postings[49].

For these reasons, I do not attempt to identify the boundaries of each ad and instead analyse the whole newspaper page. I perform a simple keyword search for job titles and locations, and match them according to their position in the text. The key idea is that, within a help-wanted ad, the job title is almost always given before the job location. A job-location pair is thus identified as job keyword and location keyword that satisfy the following three criteria: (i) the job keyword occurs in the text before the location keyword; (ii) the keywords are "consecutive", i.e. there is not another keyword between them in the text; (ii) they lie within a 80 words of each other[50].

A stylised example of this job-location pair extraction is given in Figure A1. The top panel gives an example of three help-wanted ads, with the job and location keywords highlighted. The bottom panel illustrates how this text is read by my algorithm, and how jobs and locations are paired. The first ad represents a clear case of a job title followed by location, which are matched into a job-location pair. The second ad contains a job title, but provides no location information, so no job-location pair can be formed. The third ad does come with location, but includes three different job titles. Even though all of them are located in the now-depopulated town of Franklin, Alaska, the algorithm only picks up one job-location pair, that of the last job title, since the first two violate condition (ii).

The algorithm allows me to reconstruct job-location pairs of a vacancy without the need to identify ad boundaries, but it also has its shortcomings. The first issue is illustrated by the third help-wanted ad in the example in Figure A1: an ad with multiple job titles only matches up the job title most immediately preceding the location keyword. If we assume job titles without a location are local vacancies, multiple-job vacancies such as this one will lead to overestimation of local job posting and underestimation of recruitment distances[51]. The second problem arises when the algorithm picks up a job title from a location-less ad and matches it with a job-less location keyword from the next ad, creating a false positive match. This can happen especially on newspaper pages which combine help-wanted ads with

[49]The other two existing papers analysing the text of newspaper help-wanted ads take different approaches. Atalay et al. (2020) only analyse three newspapers, which allows them to clearly define ad patterns and define vacancy boundaries this way. Anastasopoulos et al. (2021) identify individual postings manually, as a part of the transcription process of the newspaper pages. More details on the methods used by these papers and on the differnces and similarities with this paper can be found in Appendix **??**

[50]This cutoff point was selected based on the average length of the manually analysed vacancies.

[51]A solution to this problem is to explicitly recognise that job titles within a few words of each other belong to the same vacancy, and hence to the same location keyword. I plan to add this to the next version of the algorithm.

Figure A1: A stylised example of extraction of job-location pairs

(a) Example of help-wanted ads

```
Part time sales clerk, nights and weekends, no phone calls please.
Apply at Keynes Café, Springfield.

RECEPTIONIST, part time days. Call 338-6000. 8-5, ask for John or
Joan.

CATERING position open for combination of waiter/bartender. Also
looking for chefs. Will train, apply in person. Johnson's Catering
Depart., 331 Washington Drive, Franklin, AK.
```

(b) Corresponding text analysis

keyword position	job keyword	location keyword	job-location pair
3-4	sales clerk		sales clerk, Springfield
16		Springfield	
17	receptionist		
35	waiter		
36	bartender		
40	chef		chef, Franklin, AK
52-53		Franklin, AK	

other classifieds, and is much more difficult to remedy by tweaking the existing approach. The third issue arises if the help-wanted ad is too long. When the job keyword and location keyword are more than 80 words apart, the algorithm doesn't recognise them as belonging to the same posting, and the job is classified as a local posting. However, my choice of the ad length cutoff (80 words) has been selected to minimise this possibility.

Ultimately, however, the quality of the text analysis depends on the datasets of job titles and location keywords. The list of job keywords would ideally contain all the words and expressions used as job title in regular speech, not just the job titles used by statistical agencies or in occupation hierarchies. For location keywords, we need a list of all relevant geographic entities in the US: towns, cities, counties and states.

I construct the database of job title keywords from the Alphabetical Index of Industries and Occupations for the 2000 Census. This is an exhaustive list of occupations and their related job titles, developed by the US Census Bureau to aide census takers in the 2000 Census. It contains almost 31,000 job titles with their corresponding SOC codes, so each job keyword can be easily linked to the occupation hierarchy. I process and extend this list to capture all the possible variations of job titles. First, I remove some expressions and characters which are unlikely to used in a help-wanted ad, such as removing the abbreviation for "not specified", "\n.s.", from "adjudicator \n.s.". In some cases, this includes striping away some information from the original job title. For example, "health therapist, less than associate degree" contains useful information for the census taker on the education required, but this exact expression is unlikely to be found in a newspaper ad, so I shorten it to "health therapist". Second, I generate all possible permutations of multi-part job titles, i.e. job titles with words separate by commas, the word "or", or backslashes. For example, "victims advocate clerk/specialist" becomes two separate entries: "victims advocate clerk" and "victims advocate specialist". The original job title"apprentice, air conditioning, auto" becomes five separate keywords: "apprentice", "apprentice air conditioning", "air conditioning apprentice", "apprentice auto air conditioning", and "auto air conditioning apprentice". This procedure generates many redundant (unlikely) job keywords, but it also ensures the algorithm captures all its possible relevant versions without having to consider each of the 31,000 job titles individually. By carrying over the SOC code of the original job title, I create a dataset of more than 47,000 job keywords and their SOC codes.

The list of geographical keywords comes from a gazetteer of the US created by GeoNames[52]. The gazetteer contains the names of more than 2.2 million geographical features; each name is accompanied by its coordinates (latitude and longitude), feature class (such as lake, city, region, road, mountain or building), population (if applicable), any alternative

[52]https://www.geonames.org/

names, and which state and administrative divisions it belongs to. I restrict my attention to geographical names that correspond to 6 specific feature classes: *state, county, state capital, county capital, political capital*, and *other towns and cities* with population over 10,000 inhabitants. This generates a list of more than 9,500 locations. Similarly to the list of job keywords, I extended the original gazetteer to include entries commonly found in help-wanted ads. I added the abbreviation "Co." as an alternative to "county" to each county name; and for each location other than state, I created a duplicate followed by the state name or abbreviation. For example, "Drew County" in Arkansas will now be included in the database six times: as "Drew County", "Drew Co.", "Drew County, Arkansa", "Drew County, AR", "Drew Co., Arkansas" and "Drew Co., AR". The final dataset contains about 37,000 location keywords which can be linked to their coordinates.

There are several types of false positives in both keyword searches that need to be coded and excluded. Some job titles have other meanings, or are used in conjunction with other forms to form expressions unrelated to the job market: "shepherd" is an occupation, but "German shepherd" is not. To fix this, I remove a set of job titles with other commonly used meanings, such as "brother" an "advocate". In the case of location keywords, the main issue are geographic names used in other types of names. For example, the US abounds in streets, avenues, plazas, etc. named after Pennsylvania. This presents a problem: since these names usually come before the actual name of city or town, the algorithm would pick them up as the actual location, resulting in a false match. I thus specifically exclude any location keywords which are followed by any "false key terms", such as street, court, church, or university[53]. Similarly to job keywords, I also removed location keywords that have other meaning (such as "Justice", a town in Illinois).

In theory, we should be able to link all job keywords to their SOC codes, and all location keywords to their geo coordinates. In practice, the linkage is not possible for about 10.5% of location keywords and 5.5% of job keywords. The problem are non-unique keywords. There are several job keywords which, if not specified further, cannot be linked to a unique SOC code. For example, there are 277 keywords containing the word "driver". "Fire engine driver" falls under SOC 33-2000 (fire fighting and prevention workers), "farm truck driver" is SOC 53-3000 (motor vehicle operators), and "log truck driver" is classified as SOC 45-4000 (forest, conservation, and logging workers). Any ad which only gives "driver" as its job title thus cannot be associated with a unique SOC code. I keep these job keywords in the text analysis, but they are treated as missing data when SOC codes are needed. For geographic places, a well-known example is the name "Springfield", which refers to inhabited places in

[53] The full list is can be found in Appendix **??**.

34 US states[54]. If such a non-unique name is not followed by a county or state name, I assume it lies in the state of the newspaper. If no such place exists (there is no Springfield in Rhode Island), it cannot be linked to geo coordinates, and it is effectively unidentified for the purposes of this analysis. I do not include it in the list of found locations or matched into job-location pairs. The decision to omit non-identified locations while keeping ambiguous job titles makes my estimated recruitment distances as accurate as possible while not discarding any posted jobs.

Having constructed the datasets of job and location keywords, I run a keyword search and job-location pair matching as described at the beginning of this section and Figure A1. This analysis has three outputs: a list of found locations, a list of found jobs, and a list of identified job-location pairs. For each search result I also know the name of the newspaper it was published in. If we assume that jobs without a known location are local (i.e. posted in the same location as the newspaper), the list of jobs can be added to the list of job-location pairs to form a complete list of all vacancies, their locations, and their place of publication. In total, the text analysis algorithm found 153,353 locations, 318,138 jobs, and identified 87,142 job-location pairs.

A.2.3 Adding geographical information

Each help-wanted ad in my dataset comes with two locations: the location of the job, and the location of the newspaper[55]. Because both locations can be linked to their geo coordinates, I can calculate the main variable of interest, recruitment distance, for each job. Recruitment distance measures how far from the job location is the job advertised. I calculate it as the orthodromic ("as the crow flies") distance in kilometers, based on the geo coordinates and using the the haversine formula. I assign recruitment distance of 0 to all jobs without an identifiable location.

Of course, as-the-crow-flies recruitment distance may over- or under-estimate the true proximity of the job to where it is advertised. Furthermore, newspapers differ in their geographic circulation: if we use newspaper location as a proxy for the location of potential employees, this will add another source of upward or downward bias.

To address this, I merge in for each location a set of additional geographical information. The gazetteer provides the county and state of each location, which allows me to analyse recruitment distance as a function of administrative boundaries (what share of vacancies are advertised within their county or state?). Using a county-MSA crosswalk, I can also

[54] https://www.usgs.gov/faqs

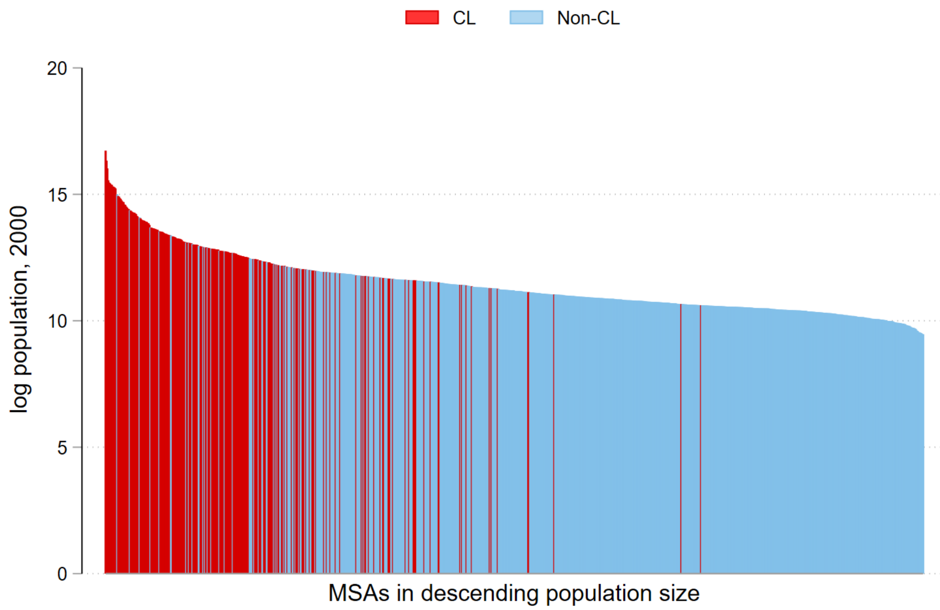
[55] Locations of newspapers are taken from the search result pages on *Newspapers.com*, with any missing data completed manually.

determine whether a location is urban or non-urban, and which of the 284 US metropolitan statistical areas it belongs to. Finally, I merge in population counts from the 1990 Census for each location, county, and MSA (if applicable).

Overall, each advertised job in my dataset comes with data on the position, size, and relative distance of its location and of where it is advertised.

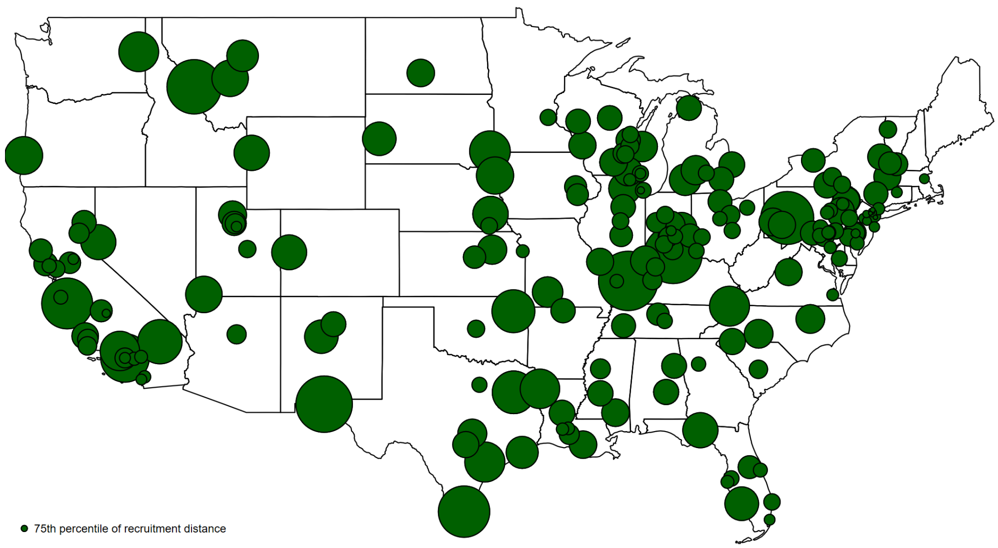
B Appendix figures and tables

Figure A2: CL and non-CL cities in descending order by population size, 2000



Notes: A bar graph of all MSAs in the USA, ordered by their (log) population size in 2000. Red bars correspond to cities with Craigslist by the end of 2006. Blue bars are cities without Craigslist by the end of 2006.

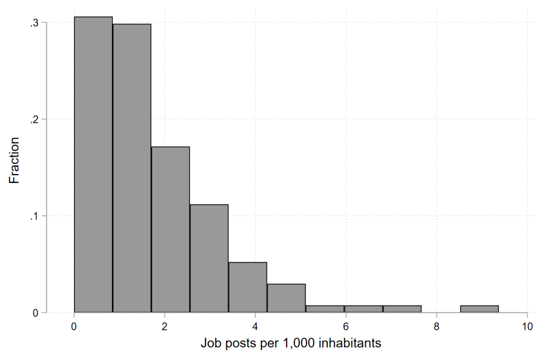
Figure A3: Local labour markets in the newspaper data



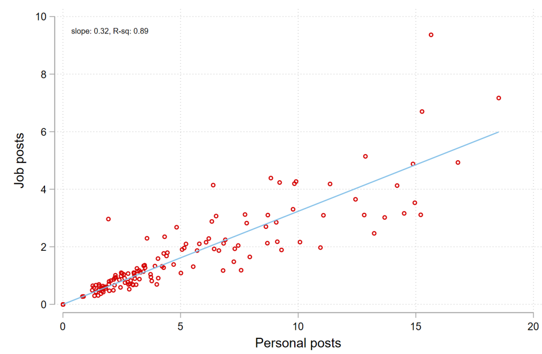
Notes: Each circle on the map represents a newspaper in my dataset of help-wanted ads. The size of the circle corresponds the geographic radius of help-wanted ads advertised in a given newspaper, measured at the 75th percentile of recruitment distance.

Figure A4: Craigslist popularity

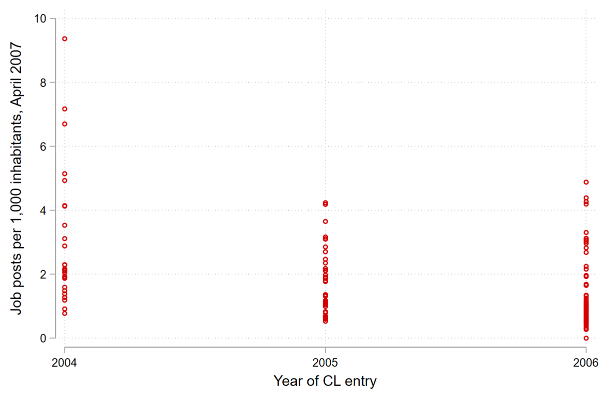
(a) Distribution of the number of job posts across cities



(b) First-stage relationship between job posts and personal posts



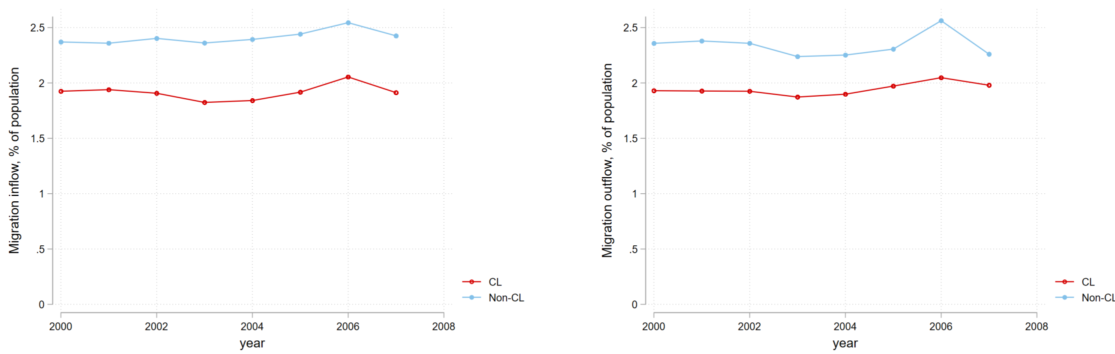
(c) CL popularity across cities, by year of entry



Notes: Descriptive statistics of CL popularity, as measured by the number of job posts per 1,000 inhabitants in April 2007. Data from Kroft and Pope (2014). Panel (a) plots the distribution of the number of job posts across cities where CL entered after 2003. Panel (c) further breaks this distribution down by the year of CL entry. Panel (b) shows the first-stage relationship between the normalised number of job posts and the normalised number of personal posts per city.

Figure A5: Trends in aggregate migration flows

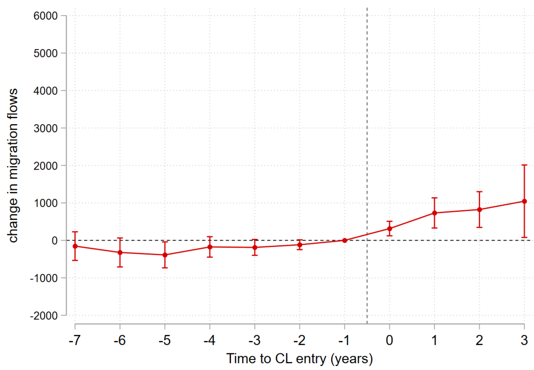
(a) Inflows (b) Outflows



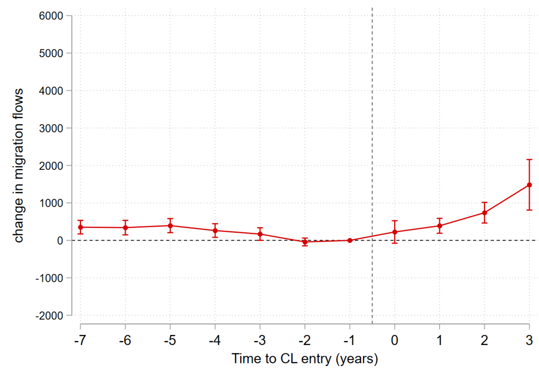
Notes: Total city-level inflows (panel (a)) and outflows (panel (b)) for years 2000-2008. Red line: Craigslist cities (status as of December 2006). Blue line: cities without Craigslist.

Figure A6: Impact of CL entry on aggregate migration flows, robustness checks

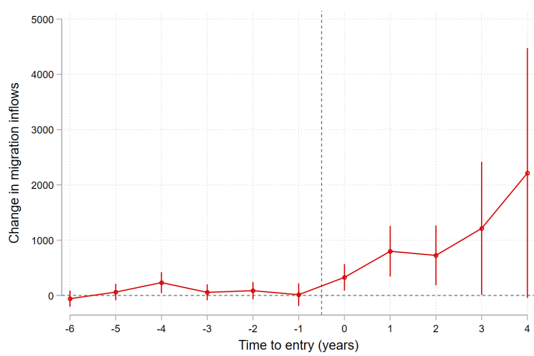
(a) Inflows, raw data



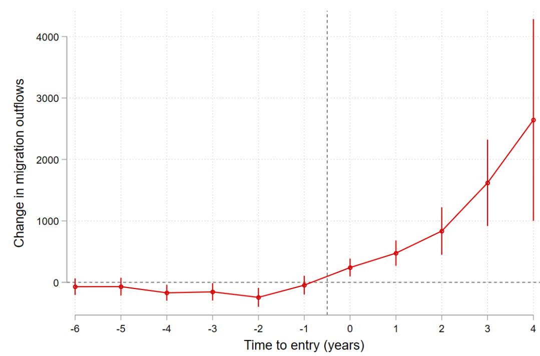
(b) Outflows, raw data



(c) Inflows, Callaway and Sant'Anna (2021) estimator



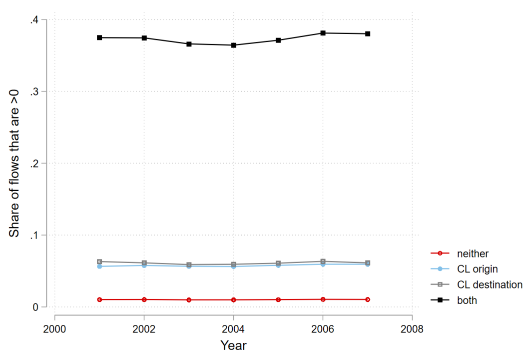
(d) Outflows, Callaway and Sant'Anna (2021) estimator



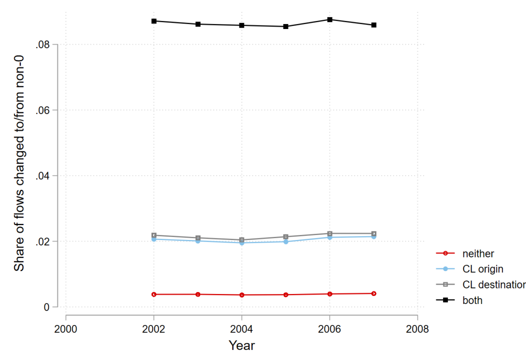
Notes: Coefficients from dynamic difference-in-differences regression (equation (1)). Panels (a) and (b): TWFE estimator on raw data (all cities, city and year fixed effects, no controls). Panels (c) and (d): baseline specification estimated using an alternative estimator by Callaway and Sant'Anna (2021), including city and year fixed effects and baseline levels and growth of Internet service providers interacted with time FE. The model is estimated on a sample of cities balanced on baseline covariates and the control group are never-treated cities. Vertical bars represent confidence intervals at 95% level of statistical significance.

Figure A7: Positive bilateral flows

(a) Positive bilateral flows



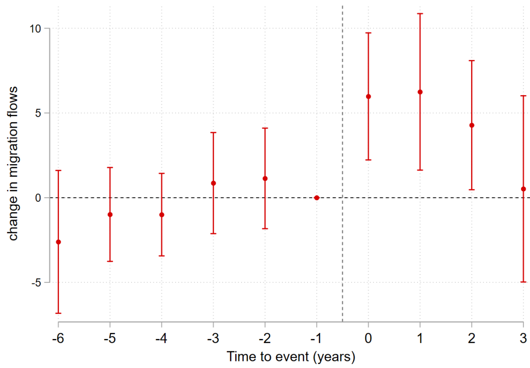
(b) Changes to/from positive flows



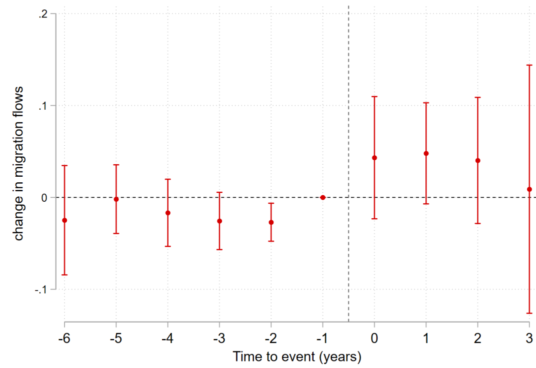
Notes: Panel (a): share of annual city-to-city flows between 2000-2008 (pooled data) that are non-zero. Panel (b): share of city-to-city flows that change between zero and non-zero.

Figure A8: Impact of CL entry on bilateral migration flows

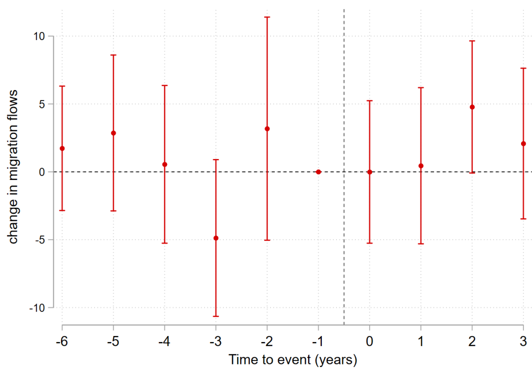
(a) Both cities CL, flows>0



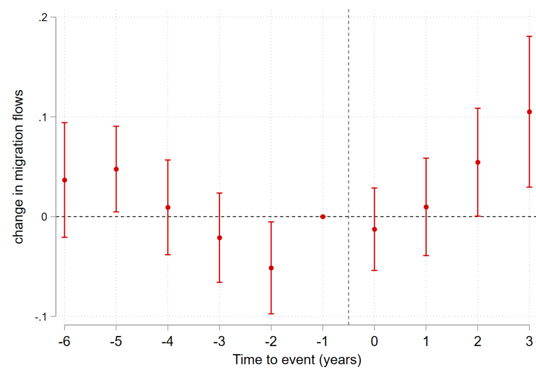
(b) Both cities CL, Poisson



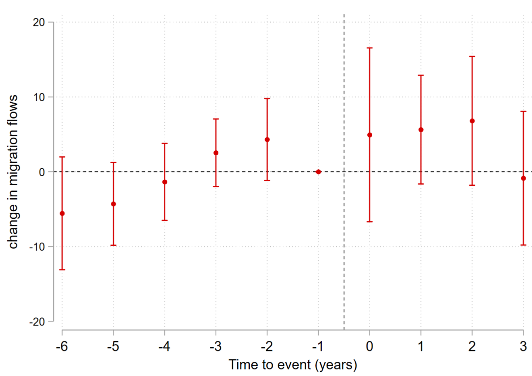
(c) CL origin, flows>0



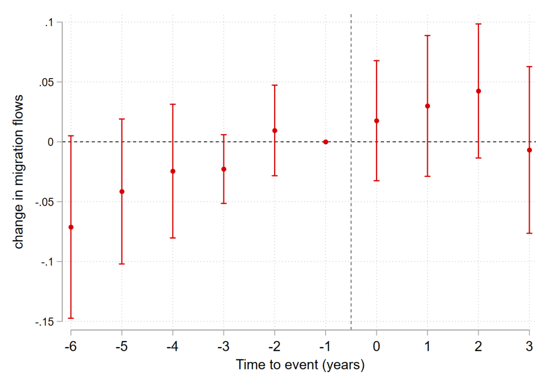
(d) CL origin, Poisson



(e) CL destination, flows>0



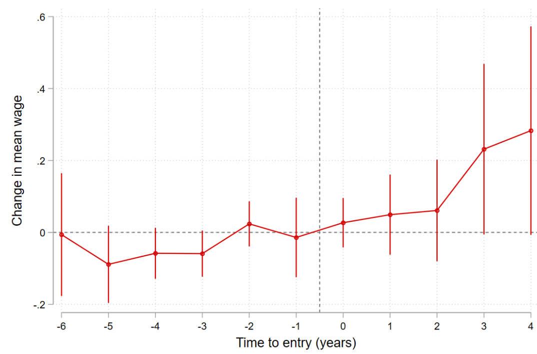
(f) CL destination, Poisson



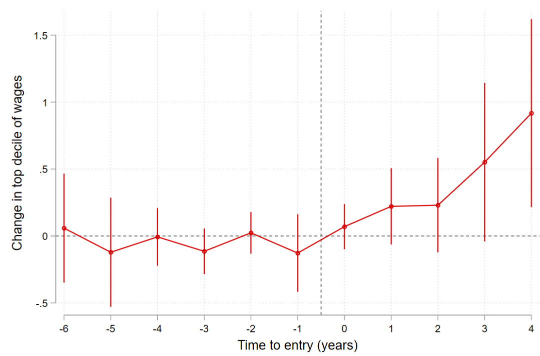
Notes: The left-hand side panels contain only positive (non-0) bilateral flows. The right-hand side contains the full sample and is estimated using a Poisson regression. All regressions contain city and time FE and baseline internet availability interacted with time dummies. The sample used is migration flows between cities such that both origin and destination are drawn from the matched sample.

Figure A9: Impact of CL entry on city-level hourly wages, robustness check

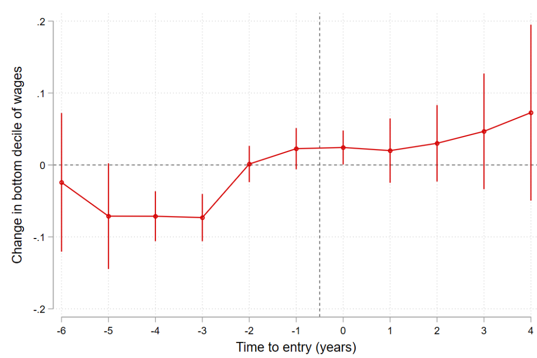
(a) Mean wage



(b) Top 10%

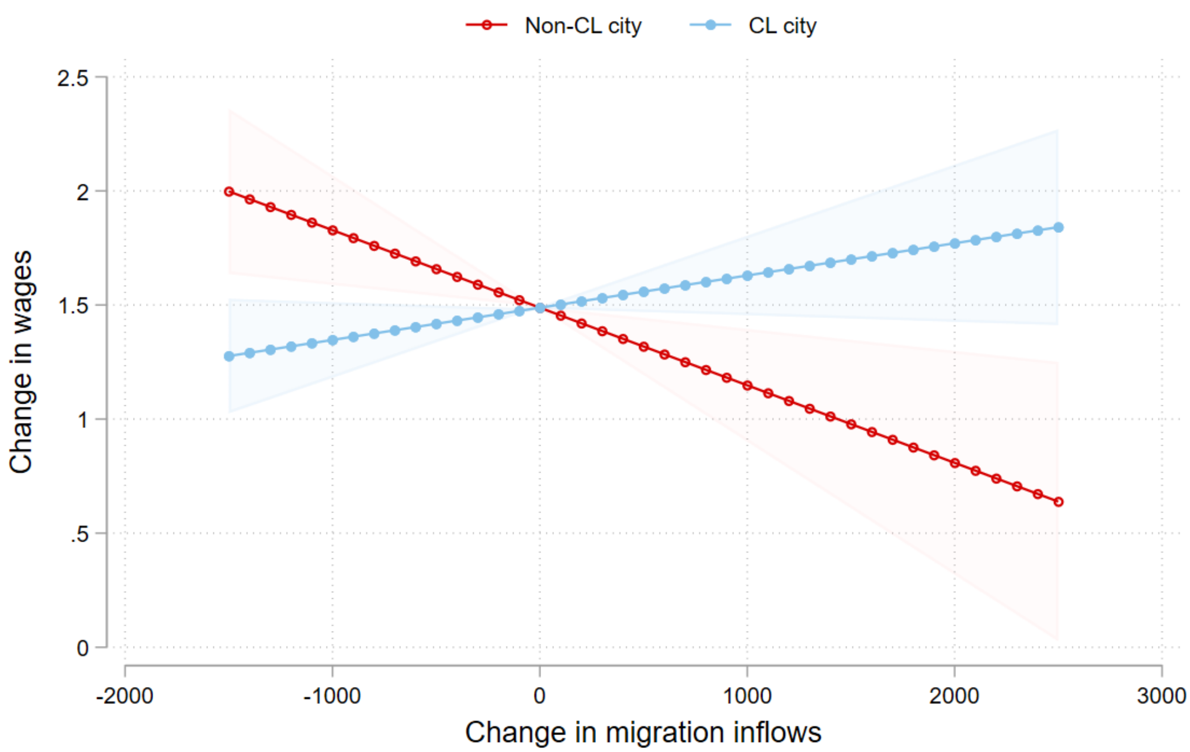


(c) Bottom 10%



Notes: Coefficients from dynamic difference-in-differences regression (equation (1)) estimated using an alternative estimator by Callaway and Sant'Anna (2021), including city and year fixed effects. The model is estimated on a sample of cities balanced on baseline covariates and the control group are never-treated cities. Vertical bars represent confidence intervals at 95% level of statistical significance.

Figure A10: Estimated relationship between changes in wages and gross migration inflows at occupation level within Craigslist and non-Craigslist cities



Notes: Predicted margins from a regression of changes in city-occupation-specific wages on city-occupation-specific migration inflows. The changes are calculated between years 2005-2007. The specification includes city- and occupation- fixed effects. The slope for control cities is -0.00034 (0.0001243), and the difference between the slope for control and treated cities is 0.0004812 (0.0001482). The regression's R-squared is 0.3762, and the sample size is 1,624. The red line and filled circles in the plot: Craigslist cities. Blue line and empty circles: non-Craigslist cities. Standard errors are clustered at city level. The shaded areas correspond to confidence intervals at 95% statistical significance.

Table A2: Variable definitions and data sources

Variable	Description	Source
	City-level variables	
City-level migration	Number of individuals who changed city of residence between two fiscal years	SOI Tax Statistics, Inland Revenue Service
City × occupation migration	Number of individuals who changed their city of residence between two calendar years	American Community Survey
Wage	Mean, 10th percentile and 90th percentile of city or city × occupation hourly wage, annual statistic	Occupational Employment and Wage Statistics
Broadband availability	The number of high-speed Internet service providers within a city	Form 477 data, Federal Communications Commission
Employment share	Total employment (excl. solo-employment) as a share of total population	Occupational Employment and Wage Statistics
GDP growth	Annual growth in real Gross Domestic Product at city level	U.S. Bureau of Economic Analysis
Population	Annual estimate of the number of inhabitants	US Census Bureau, Population Division
Land area	Area of the city in squared miles	2000 Census
Population density	Average number of inhabitants per squared mile	2000 Census
Population share, by race/ethnicity	% of city population that is white/black/Hispanic	2000 Census
Median age		2000 Census
Population share under 18	% of city population under 18 years of age	2000 Census
Population share over 75	% of city population over 75 years of age	2000 Census
	County-level variables	
Average wage	Average weekly wage based on the 12-monthly employment levels and total annual wage levels	Quarterly Census of Employment and Wages
Number of establishments	Annual average of quarterly establishment counts for a given year	Quarterly Census of Employment and Wages
Labour force	Number of individuals in the labour force	Local Area Unemployment Statistics
Unemployment rate	County-level unemployment rate (%)	Local Area Unemployment Statistics

Table A3: Description of the newspaper vacancy dataset

variable	# of unique values	
	in the dataset	in the population
newspapers	216	
locations	200	9548
counties	174	3142
MSAs	93	922
states	44	50
vacancies	318 183	
occupations	93	96
locations	2699	9548
counties	864	3142
MSAs	242	922
states	47	50

Notes: Descriptive statistics of the data on vacancy posting from help-wanted newspaper ads in 1990. The two panels summarise the number of unique values in the dataset and in the population, at the newspaper and vacancy level.

Table A4: Predicting Craigslist entry

	Craigslist entry				
	(1)	(2)	(3)	(4)	(5)
Population, log, 2002	2.300***				1.576***
	(0.171)				(0.307)
Population growth, 2002-03	9.378				5.369
	(11.35)				(27.74)
N. of Internet service providers, 2002		0.560***			-0.146
		(0.0563)			(0.118)
Growth in ISP, 2002-03		0.275***			0.363*
		(0.0860)			(0.194)
GDP growth, 2000-03			-0.0314		0.0253
			(0.0614)		(0.0688)
Employment growth, 2000-03			0.105		0.0583
			(0.118)		(0.183)
Predicted GDP growth, 2004-07				-0.248	-0.611*
				(0.224)	(0.331)
Predicted employment growth, 2004-07				0.116	0.185
				(0.215)	(0.358)
Constant	-28.39***	-3.942***	0.882***	1.546**	-16.98***
	(2.028)	(0.297)	(0.225)	(0.627)	(3.529)
Observations	886	888	233	233	233
Pseudo R^2	0.476	0.145	0.003	0.005	0.155

Note: Predicting CL entry between 2004 and 2006 using population levels and growth up to 2 years before the entry (column (1)), the levels and growth in the availability of high-speed internet connection (column (2)), city-level growth in employment and real GDP (column (3)), and predicted growth of employment and GDP over the years 2004-2007 (column (4)). Column (5) uses all the predicting variables simultaneously. Standard errors are in parentheses. $*p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table A5: Balance test for the matched sample

	All cities			Kijiji cities		
	Craigslist cites	Non-CL cities	Diff.	Craigslist cites	Non-CL cities	Diff.
Inflow share, %	2.13	2.24	0.11	2.14	2.67	0.51**
	(0.99)	(1.15)	(0.16)	(0.97)	(1.21)	(0.20)
Outflow share, %	2.04	2.19	0.15	1.99	2.32	0.37**
	(0.94)	(1.26)	(0.16)	(0.87)	(0.91)	(0.16)
Population (1,000s)	478.53	215.70	-262.83***	553.58	454.41	-162.48*
	(400.30)	(148.69)	(56.25)	(411.52)	(684.34)	(85.23)
Population density	2356.71	1879.16	-477.55***	2414.52	2263.27	-244.32
	(1041.96)	(1234.81)	(172.01)	(1011.57)	(1575.37)	(213.53)
Population share, white (%)	80.28	81.95	1.66	79.25	82.12	3.46
	(11.16)	(13.07)	(1.84)	(11.33)	(12.68)	(2.10)
Share under 18 (%)	25.66	25.34	-0.32	25.90	25.67	-0.46
	(2.95)	(2.20)	(0.44)	(2.92)	(2.37)	(0.57)
Share over 75 (%)	5.81	6.03	0.22	5.80	5.96	-0.00
	(1.44)	(1.36)	(0.22)	(1.45)	(1.63)	(0.32)
Employment share (%)	44.12	40.93	-3.19	44.34	41.79	-2.55
	(11.43)	(14.54)	(1.93)	(10.96)	(19.55)	(2.69)
Hourly wage, mean	14.28	13.75	-0.54**	14.43	14.27	-0.16
	(1.43)	(1.39)	(0.22)	(1.41)	(1.45)	(0.30)
Hourly wage, bottom 10%	7.63	7.60	-0.03	7.68	7.94	0.26
	(1.10)	(0.75)	(0.16)	(1.19)	(0.82)	(0.24)
Hourly wage, top 10%	22.77	21.83	-0.94*	22.95	22.47	-0.49
	(3.49)	(2.04)	(0.51)	(3.65)	(1.90)	(0.72)
Observations	167	53	220	134	27	183

Note: The comparison of city-level baseline characteristics (year 2000) for the matched sample of treated and control cities. The first three columns present the characteristics of CL and non-CL cities and calculate their difference. The last three columns repeat the exercise for the subsample of Kijiji cities. Standard errors are in parentheses. $*p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table A6: Impact of Craigslist entry on aggregate migration flows, event study parameters for robustness checks

	Balanced sample				All data		Kijiji cities
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Levels	Levels	Levels	Levels	Log	Levels	Levels
Panel A: Gross Inflows							
Treatment effect at t+0	28.92	-97.37	316.7	129.8	0.0214**	150.4	-238.1
	(140.0)	(153.3)	(376.3)	(140.9)	(0.00860)	(194.1)	(364.7)
Treatment effect at t+1	766.5***	479.5*	1757.2**	671.6***	0.0387***	536.6*	55.18
	(267.8)	(268.6)	(859.2)	(223.9)	(0.0109)	(283.3)	(574.9)
Treatment effect at t+2	1113.7***	502.2	2899.1**	560.7**	0.0279**	658.7**	378.6
	(380.3)	(320.5)	(1290.2)	(231.9)	(0.0121)	(323.5)	(740.2)
Treatment effect at t+3	1523.4**	112.6	3639.8**	517.5	0.0194	1877.3**	1390.1
	(649.8)	(633.8)	(1620.2)	(384.7)	(0.0161)	(836.6)	(1114.7)
Panel B: Gross Outflows							
Treatment effect at t+0	42.98	245.0**	161.8	82.14	0.0179**	148.3	-847.9
	(79.22)	(98.25)	(169.5)	(77.70)	(0.00795)	(185.5)	(905.1)
Treatment effect at t+1	174.4	538.2***	453.4	194.3**	0.0175**	497.1**	-246.1
	(109.6)	(168.6)	(393.4)	(88.02)	(0.00802)	(209.0)	(821.6)
Treatment effect at t+2	643.1***	1135.2***	1131.4	478.6***	0.0328***	1324.7***	1009.7
	(198.9)	(247.6)	(712.0)	(144.1)	(0.00924)	(372.2)	(1077.4)
Treatment effect at t+3	1240.5***	1835.1***	1842.0*	798.8***	0.0621***	2850.5***	2405.7*
	(374.5)	(383.2)	(991.3)	(254.3)	(0.0129)	(678.3)	(1396.7)
Panel C: Net Inflows							
Treatment effect at t+0	-14.05	-342.3*	154.9	47.66	0.0769	2.099	609.8
	(146.8)	(196.6)	(356.0)	(137.6)	(0.0995)	(209.0)	(805.3)
Treatment effect at t+1	592.1*	-58.69	1303.8	477.3**	0.132	39.46	301.3
	(301.2)	(362.9)	(794.3)	(238.2)	(0.128)	(298.4)	(633.0)
Treatment effect at t+2	470.6	-633.0	1767.7	82.13	0.138	-666.0*	-631.0
	(379.7)	(412.1)	(1145.4)	(246.1)	(0.118)	(396.6)	(963.2)
Treatment effect at t+3	282.9	-1722.5**	1797.8	-281.3	0.0489	-973.2	-1015.6
	(673.7)	(738.4)	(1473.0)	(460.2)	(0.154)	(950.5)	(1396.2)
N	2200	2200	1670	1980	3577	9160	2110
MSA FE	YES	YES	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES	YES	YES
Linear MSA trend		YES					
Covariates × time	YES			YES		YES	YES
CL cities only			YES				
City size common support				YES			
All CL entries						YES	

Note: This table summarises the results of a dynamic diff-in-diff version of the robustness checks in Table 2. Column (1) corresponds to my baseline preferred specification. Column (2) uses city-specific linear growth trends. Column (3) re-runs the baseline specification from column (1) but only on CL-cities. Column (4) re-runs the baseline specification further imposing common support over population size of treated and control cities. Columns (5) and (6) are estimated on the entire sample of cities; column (5) uses log migration flows, and column (6) includes cities that were treated in the years 2000-2003. Column (7) re-runs the baseline specification on Kijiji cities only. Standard errors are in parentheses, clustered at city level throughout. $*p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.010$

Table A7: Impact of Craigslist entry on average hourly wages, robustness checks

	Balanced sample				All data		Kijiji cities
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Levels	Levels	Levels	Levels	Log	Levels	Levels
Panel A: Mean hourly wage							
Treatment effect	0.125***	0.0770**	0.182***	0.134***	0.00352	0.0214	-0.0239
	(0.0473)	(0.0345)	(0.0492)	(0.0500)	(0.00220)	(0.0503)	(0.0638)
Average wage	15.960	15.960	15.960	15.960	2.763	15.960	15.960
Panel B: Bottom 10% hourly wage							
Treatment effect	0.0664***	0.0820***	0.0527**	0.0764***	-0.00302	-0.0213	-0.0508*
	(0.0219)	(0.0187)	(0.0229)	(0.0239)	(0.00317)	(0.0219)	(0.0261)
Average wage	6.886	6.886	6.886	6.886	1.926	6.886	6.886
Panel C: Top 10% hourly wage							
Treatment effect	0.348***	0.220**	0.545***	0.364***	0.0109***	0.145	0.0849
	(0.118)	(0.0948)	(0.133)	(0.122)	(0.00409)	(0.176)	(0.236)
Average wage	28.236	28.236	28.236	28.236	3.330	28.236	28.236
Observations	2129	2129	1620	1919	3046	3046	1892
MSA FE	YES	YES	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES	YES	YES
Linear treatment trend	YES			YES	YES	YES	YES
Linear MSA trend		YES					
Covariates × time	YES			YES		YES	YES
CL cities only			YES				
City size common support				YES			
All CL entries						YES	

Note: This table summarises the results of a TWFE regression of CL entry on city-level hourly nominal wages. The presented coefficient estimates the annual change in average wage at city level following CL entry, averaged across different years after entry. Column (1) corresponds to my baseline preferred specification, using a matched sample of treated and control cities and including linear treatment trend and an interaction between period FE and baseline internet service availability. Column (2) uses city-specific linear growth trends instead. Column (3) re-runs the baseline specification from column (1) but only on CL-cities. Column (4) re-runs the baseline specification further imposing common support over population size of treated and control cities. Columns (5) and (6) are estimated on the entire sample of cities; column (5) uses log, rather than levels, of wages, and column (6) includes cities that were treated in the years 2000-2003. Column (7) re-runs the baseline specification on Kijiji cities only. Standard errors are in parentheses, clustered at city level throughout. $*p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table A8: The impact of CL entry on occupation-city-specific wages

	Balanced sample				Kijiji cities
	(1)	(2)	(3)	(4)	(5)
Treatment effect at t+0	0.0496	-0.0332	-0.0661*	0.0328	0.0264
	(0.0761)	(0.0455)	(0.0338)	(0.0343)	(0.0345)
Treatment effect at t+1	0.381***	0.0259	-0.0219	0.0989	0.103*
	(0.0706)	(0.0678)	(0.0503)	(0.0614)	(0.0614)
Treatment effect at t+2	0.678***	0.0870	0.0248	0.133	0.156*
	(0.105)	(0.0911)	(0.0671)	(0.0910)	(0.0921)
Treatment effect at t+3	1.363***	0.312**	0.252***	0.249**	0.347***
	(0.158)	(0.124)	(0.0874)	(0.126)	(0.131)
Treatment effect at t+4	1.880***	0.393**	0.330***	0.236	0.398**
	(0.245)	(0.179)	(0.119)	(0.175)	(0.184)
N	34353	34353	34353	34353	29167
Time FE	YES	YES	YES	YES	YES
MSA FE		YES	YES	YES	YES
SOC FE	YES	YES	YES	YES	YES
Covariates time FE	YES	YES	YES		
Linear MSA trend				YES	
Linear SOC trend			YES	YES	

Note: Dynamic difference-in-differences regressions of city×occupation wages on Craigslist entry. All specifications include time and occupation fixed effects; specifications in columns (1) - (4) are estimated on a matched sample of treated and control cities. Column (1) also includes the standard interaction between time fixed effects and internet availability (in levels and growth) at city level at the baseline. Column (2) adds city fixed effects. Column (3) further adds a linear occupation-specific trend. Column (4) uses a city-specific linear trend instead of the interaction between time fixed effects and internet availability. Column (5) includes city- and occupation- fixed effects and is estimated on the sample of Kijiji cities. Standard errors in parentheses, clustered at occupation and city level. $*p < 0.10$, ** $p < 0.05$, *** $p < 0.010$