Job Vacancies and Immigration: Evidence from the Mariel Supply Shock

L. Jason Anastasopoulos, George J. Borjas, Gavin G. Cook, and Michael Lachanski

November 2020

Job Vacancies and Immigration: Evidence from the Mariel Supply Shock

L. Jason Anastasopoulos, George J. Borjas, Gavin G. Cook, and Michael Lachanski

Abstract

Beginning in 1951, the Conference Board constructed a monthly job vacancy index by counting the number of help-wanted ads published in local newspapers in 51 metropolitan areas. We use this Help-Wanted Index (HWI) to document how immigration changes the number of job vacancies in the affected labor markets. Our analysis revisits the Mariel episode. The data reveal a sizable drop in Miami's HWI relative to comparable cities in the first 4 or 5 years after Mariel, followed by recovery afterwards. We also examine the text of the help-wanted ads published in a number of newspapers and document a significant post-Mariel decline in the relative number of low-skill vacancies advertised in the *Miami Herald*. The post-Mariel trends are consistent with the observed relation between immigration and the HWI across all metropolitan areas in the 1970-2000 period: the spatial correlation suggests that more immigration reduces the number of job vacancies. We explore some of the macroeconomic implications and show that Miami's Beveridge curve shifted inwards by the mid-1980s, suggesting a more efficient local labor market, in contrast to the outward nationwide shift coincident with the 1981-1982 recession.

Job Vacancies and Immigration: Evidence from the Mariel Supply Shock

L. Jason Anastasopoulos, George J. Borjas, Gavin G. Cook, and Michael Lachanski*

I. Introduction

How do labor supply shocks affect labor market outcomes? This is perhaps the central question in the economics of immigration, in terms of both its economic content and policy implications. Most important questions in labor economics relate to how labor markets adjust to supply and demand shocks. An immigration-induced increase in labor supply lets us observe how firms and workers react and adjust to the changed environment. Similarly, the debate over immigration policy often revolves around how immigration changes the size of the economic pie in the receiving country, and, particularly, with how that pie is split.

The centrality of the question inspired a voluminous amount of empirical research over the past four decades. This literature typically examines how wages change in those labor markets targeted by immigrants. Sometimes the markets are defined by geographic boundaries (Altonji and Card, 1991); sometimes the markets are defined by skill group (Borjas, 2003). Regardless of how markets are defined, the basic identification strategy is the same. Immigrants tend to target some markets more than others. The impact of immigration is then measured by contrasting the evolution of wages in the markets hit by immigration with the evolution in the markets that immigrants avoided.

Due to a host of technical issues (e.g., immigrants do not target markets randomly; firms and workers diffuse the impact of local supply shocks by moving elsewhere; and the available data often yields small samples for many cities and skill groups), the existing literature amply demonstrates the difficulty of measuring the impact of immigration on wages (Blau and Mackie, 2016). Even more problematic, it turns out that the evidence often depends on *researcher choices* about how to frame the empirical analysis.

The recent debate over the wage impact of the Mariel boatlift provides a classic example of how those choices influence the answer. Card (1990) reported that the wage of the average

---

* Anastasopoulos: Department of Public Administration and Policy, Department of Political Science, and Department of Statistics, University of Georgia; Borjas: Harvard Kennedy School; Cook: Department of Sociology, Princeton University; Lachanski: Population Studies Center, University of Pennsylvania. We are grateful to Jonathan Arnold and Christian Kincaid for superb research assistance, to the Peterson Foundation for financial assistance, and to two anonymous referees for constructive comments.

worker in Miami was barely affected by the 8 percent increase in supply that Mariel represented. Borjas (2017) showed that if one focused on the low-skill worker most likely to be affected by Mariel (as represented by the average non-Hispanic man who is 25-59 years old and is a high school dropout), the evidence suggested that the wage fell by at least 15 percent. In contrast, Peri and Yasenov (2019) used an alternative definition of a low-skill worker (i.e., the average non-Cuban man or woman who is 19 to 65 years old and has not graduated from high school), and found that the wage remained unchanged after the supply shock.[1]

We contribute to the literature by reporting findings from a "new" data set, a data set that economists have used often since the 1960s, but that has not been exploited in the immigration context. Beginning in 1951, the Conference Board constructed an index of job vacancies in local labor markets by counting the number of help-wanted classified ads in newspapers in 51 metropolitan areas. Although the rise of online advertising obviously reduced the usefulness of this index beginning sometime around 2000, the Conference Board's Help-Wanted Index (HWI) provides a historical series of the ebbs and flows of labor demand in local labor markets for the last half of the 20th century.

The HWI has been used to study such diverse phenomena as the trend of wages and productivity in the stagnant 1970s (Medoff, 1983); the sectoral shifts explanation of structural unemployment (Abraham and Katz, 1986); the relation between job vacancies and the unemployment rate (Cohen and Solow, 1967; Burch and Fabricant, 1968; and Abraham, 1987); the role of job search in a real business cycle framework (Andolfatto, 1996), and the cyclicality of job vacancies (Shimer, 2005). An important theme runs through these studies: The HWI provides valuable information about labor demand and is highly correlated with various measures of labor market conditions.

The analysis of the HWI has one very valuable feature in the immigration context. It greatly reduces the number of degrees of freedom available to a researcher interested in estimating the labor market impact of immigration. The index was created concurrently with the supply shocks by independent organizations for a purpose totally unrelated to the immigration

[1] Notwithstanding the ongoing confusion about the wage impact of the Mariel boatlift, Monras (2021) documents post-Mariel internal migration flows in and out of the Miami labor market that *reduced* the relative supply of low-skill workers in the city during the 1980s.

question that is at the core of this paper. The historical trends in the HWI in the cities that received many or few immigrants are set in stone.

Even more important, the analysis of the HWI allows us to extend the study of the labor market impact of immigration into areas that have been ignored in the literature, including an examination of the link between immigration and job vacancies and the documentation of some of the macroeconomic consequences of supply shocks.[2] Our analysis begins by revisiting the Mariel episode. The data clearly reveal a marked decrease in Miami's HWI relative to various control groups in the first 4 or 5 years after Mariel, followed by a full recovery afterwards. Our empirical analysis also looks beyond Miami and estimates the generic spatial correlation that dominates the literature, correlating changes in the HWI between 1970 and 2000 with immigration across the 51 metropolitan areas for which the index is available. The spatial correlation also indicates that immigration is typically associated with reduced employer effort to find workers by placing help-wanted ads in the local newspaper.

One weakness of the HWI is that it does not provide any information on the occupational characteristics of vacancies. We overcome this limitation (in the context of the Mariel supply shock) by directly inspecting a large sample of text drawn from help-wanted ads published in the *Miami Herald* and other newspapers between 1978 and 1984. This type of analysis helps identify the types of vacancies that "vanished" from the Miami labor market in the early 1980s. It turns out that the drop in the number of job vacancies in Miami was particularly severe in the low-skill sector. The fraction of vacancies for blue collar positions fell precipitously, and surviving vacancies were for jobs that tended to hire more educated workers. The joint study of the quantity and textual content of help-wanted ads provides complementary evidence strongly suggesting that the labor market fortunes of low-skill natives deteriorated in the aftermath of a

[2] Few studies investigate the quantitative relationship between immigration and job vacancies either in the United States or abroad. Withers and Pope (1985) find no statistically significant relationship between immigration and the Australian job-finding rate, while Pholphirul (2012) reports that immigration reduces short-term job vacancies in Thai manufacturing. Job vacancies are often present as a latent variable in both theoretical (Ortega, 2000) and empirical analyses (Davila and Saenz, 1990) of immigration's impact on labor market outcomes. Chassamboulli and Peri (2015) use the national HWI to calculate a long-run job-finding rate, which they then fix to calibrate a labor market matching model for evaluating various immigration policies. Albert (2021) uses data from the CPS Basic Monthly files to examine immigrant-native differences in the probability that an unemployed worker finds a job. The CPS data suggest that unemployed low-skill immigrants are more likely to exit unemployment than unemployed low-skill natives.

large low-skill supply shock, even when focusing on non-wage components of labor market opportunities.

The job vacancy data also lets us begin an exploration of the macroeconomic implications of the Mariel supply shock, which coincided with the onset of the 1981-1982 recession. The Beveridge curve, the downward-sloping locus relating job vacancies and unemployment, was shifting out in comparable cities during the recession. In Miami, however, after a short-lived outward shift in the first few months after the Mariel supply shock, the curve shifted inwards substantially. The inward shift might have been caused by the changed demographics of the Miami labor market, with the search behavior of the Mariel refugees differing from that of native workers (e.g., the refugees may have searched more aggressively, quickly filling up vacancies).[3] By the late 1980s, however, the relative position of Miami's Beveridge curve had moved back to where it was prior to the supply shock.

The thrust of our evidence indicates that a large supply shock composed mainly of low-skill workers has demonstrable short-run microeconomic and macroeconomic consequences, particularly affecting the low-skill sector. The literature already documents that such supply shocks have both adverse wage effects on native workers (Monras, 2020) and adverse employment effects (Smith, 2012). Our analysis suggests that the supply shock also reduces the number of job vacancies available to low-skill jobseekers (making it harder for out-of-work natives to find jobs), even though it also makes the labor market more efficient (allowing employers to fill vacancies faster).

II. The Conference Board Help-Wanted Index

Beginning in 1951, the Conference Board contacted 51 newspapers, each corresponding to a major metropolitan area, and asked each newspaper for the number of classified ads published in the previous month. The Conference Board adjusted the count for seasonal trends and day-of-the-week frequency to create a monthly index for each metropolitan area. The

[3] Albert (2021) provides a rare study of differences in search behavior between immigrants and natives. His evidence (based on post-1994 data) contradicts the conjecture that immigrants search more intensively. The merged data from the CPS and the American Time Use Survey suggests that unemployed low-skill immigrants spend less time in job search activities than their native counterparts.

metropolitan area indices were then aggregated to create an index for each geographic region, and for the nation as a whole.[4]

Figure 1 illustrates the correlation between the national HWI and the unemployment rate, showing a strong inverse relation between the two variables throughout much of the period.[5] Despite this strong link, several well-known biases in the index influence the interpretation of observed trends.

The first arises from the fact that the number of job vacancies *per ad* is procyclical. During booms, a single ad might advertise explicitly for two or more job openings. But the algorithm used by the Conference Board to construct the HWI counts this as only one advertised job (Preston, 1977). A related problem arises with ads placed by private employment agencies, which often contained several job postings. Some newspapers placed all agency advertising in a section specifically demarcated for labor market intermediaries (and this section may not have been included in the Conference Board counts), while other newspapers made no distinction between agency advertising and ads placed by individuals or firms (Walsh, Johnson, and Sugarman, 1975). Further, the HWI only counts ads placed in the official help-wanted section of the newspaper. Some high-skill jobs, especially those in finance, insurance and real estate (FIRE), were not advertised in the help-wanted section at all; they instead appeared in dedicated FIRE sections.[6]

Finally, there are differences in the market power of the newspapers surveyed by the Conference Board. In some cities, as in Miami (where the sampled paper was the *Miami Herald*), the paper used to construct the index was the key source of job classifieds in the area. In other locations, as in Minneapolis (where the sampled newspaper was the *Minneapolis Star Tribune*), there were other newspapers (the *St. Paul Dispatch-Pioneer*) that also contained many help-wanted ads (Courtney, 1991).

---

[4] Online Appendix A lists the 51 newspapers sampled by the Conference Board. Apart from the removal of the *Newark Evening News* (and the Newark metropolitan area) in 1971, and a swap of the *Dallas Times Herald News* for the *Dallas Morning News* in the early 1990s, the newspapers and cities surveyed did not change after 1970.

[5] Friedman (1985, p. 63) examines trends in the HWI for the Phoenix labor market and concludes that the HWI is "a viable indicator of future employment activity in the Phoenix area, just as it is nationally."

[6] Cohen and Solow (1967, p. 108) noted some of these methodological issues, writing that "we know nothing, for example, about the number of jobs offered per advertisement." They also make the point that "the index can not be decomposed by occupation," making it difficult to determine if the index is a better metric of labor market conditions for some skill groups than for others.

Because of these methodological issues, we focus exclusively on the panel of HWI changes rather than intercity differences in the level of the HWI. Specifically, we rescale the HWI so that the level of the index in each city equals 1 at some point in the pre-treatment period. This rescaling is inconsequential, except when illustrating the trends graphically. Our estimate of the impact of a supply shock on job vacancies uses a difference-in-differences estimator, so that the level of the index washes out in the calculation.[7]

Although the explosion of online job postings reduced the relevance of newspaper-based indices, the HWI remains useful for historical research.[8] For the time period in our analysis, the HWI is the gold standard for job vacancy data. We obtained the entire HWI time series for the 51 metropolitan areas directly from the Conference Board. Our analysis exploits the historical data in a novel context. How does the tightness or slackness of a local labor market respond to immigration-induced supply shocks?

III. Job Vacancies and Mariel

The historical details of the Mariel episode are well known. On April 20, 1980, Fidel Castro declared that Cuban nationals wishing to move to the United States could leave freely from the port of Mariel. About 125,000 Cubans left before an agreement between the Carter administration and the Castro regime ended the flow in October 1980. Borjas (2017, Table 1) documents that the Mariel supply shock, which increased the size of the workforce in Miami by 8.4 percent, was composed of relatively low-skill workers: nearly 60 percent of the adult refugees lacked a high school diploma and only 7.4 percent had a college degree. As a result, the supply shock disproportionately increased the size of the low-skill workforce. The number of workers in Miami without a high school diploma increased by 18.4 percent, but the number of college graduates rose by only 3.4 percent.

---

[7] Medoff (1983) makes a similar point, arguing that although the absolute level of the index in any given region depends on the sample of newspapers surveyed by the Conference Board, cross-region differences in the percent change in the index can capture differences in the rate of growth of job vacancies. Abraham (1987) provides the most comprehensive discussion of the benefits and problems with the HWI and shows that the index (or some normalization of it) is correlated with the true number of job vacancies.

[8] Autor (2001, p. 27) noted that the HWI was "flat throughout the 1990s economic boom" and cited the migration of "vacancy listings… from newspapers to the Internet" as a possible explanation. In response to the declining relevance of newspaper help-wanted sections, the Conference Board ceased the public release of the HWI in July 2008 and stopped data collection in October 2010.

The labor market impact of the Mariel boatlift was first studied in Card's (1990) classic paper. Card's analysis of the Miami labor market, when compared to conditions in other cities that served as a control group, indicated that nothing much happened to Miami despite the large number of new workers. The average wage in Miami did not fall relative to that in the cities that formed the control group.

Following the Borjas (2017) reappraisal, which specifically examined the wage trends of workers who lacked a high school diploma, the debate over the labor market impact of the *Marielitos* intensified in the past few years. It turns out that the wage of this specific low skill group fell substantially relative to control cities. A flurry of subsequent papers argues that the evidence may be sensitive to the definition of a "low skill worker", to changes in the racial composition of the sample, and to the sampling error resulting from the small number of observations in the Miami metropolitan area (see Borjas and Monras, 2017; Peri and Yasenov, 2019; Clemens and Hunt, 2019; Borjas, 2019; and Monras, 2021). Our analysis of the behavior of the HWI around the time of Mariel is impervious to these potential problems.

As noted earlier, part of the cross-city variation in the level of the HWI arises because of the idiosyncratic way in which classified ads were counted by different newspapers. We address this issue by rescaling the HWI so that the level of the index in each city equals 1 in the 1977-1979 period. The top panel of Figure 2 begins the empirical analysis by illustrating the 1975-1989 trend in the HWI in Miami, in the South Atlantic region, and in the entire nation. We choose the 1975-1989 period because Cuban immigration to the United States was relatively low and stable in the pre-treatment period (hovering around 6 thousand annually between 1975 and 1978), and it was again relatively low and stable in the post-Mariel period between 1981 and 1989 (hovering around 10 thousand).[9]

The trends in the *raw* HWI data are visually striking. Because Mariel coincided with the onset of a recession, the number of job vacancies declined everywhere between 1980 and 1982. The index for Miami, however, declined much more, reaching a nadir at the end of 1982 before beginning a recovery through the 1980s. By 1989, the value of the HWI for Miami was again

[9] We also conducted an analysis using a longer pre-treatment period (going back to 1970), and the measured impact of the Mariel supply shock resembles what is reported below. The 1975-1989 period is preferable because the Miami labor market in the early 1970s was still adjusting to the large number of Cubans (over 220,000) who migrated during the 1966-1970 Freedom Flights.

similar to the national index (although it was still lower than the index for the South Atlantic region).

Of course, Miami's distinctive trend should be contrasted with what happened in "comparable" cities rather than regional or national aggregates.[10] We use the synthetic control method (Abadie, Diamond, and Hainmueller, 2010) to construct an appropriate set of comparable cities. We calculate the synthetic control by using a large number of control variables, all calculated from the 1970 and 1980 IPUMS decennial census files. It is important to emphasize that the 1980 census provides information on economic and demographic conditions in the potential control cities as of April 1, 1980—*prior* to the Mariel supply shock.

The control variables are: the education distribution of workers in the city in 1980 and the percent change in the number of workers in each education group between 1970 and 1980; the industrial distribution of workers in the city in 1980 and the percent growth in the number of workers employed in each of the industries between 1970 and 1980; the fraction of immigrants in the workforce in 1980 and the percent growth in the number of immigrants between 1970 and 1980; the fraction of workers who are male, the percent growth in the number of male workers between 1970 and 1980, and the percent growth in the number of female workers; the fraction of workers who are black (in 1980); and the fraction who are Hispanic.[11] The bottom panel of Figure 3 illustrates the trend in the HWI for the synthetic control, and it confirms the uniqueness of Miami's post-Mariel experience.

A simple exercise further illustrates the striking difference between the Miami experience and that of each of the 50 other metropolitan areas in the Conference Board sample. In particular, for each city we calculated the difference in the log HWI between 1985 and the year prior to the Mariel shock. The average log point change across the 51 cities was -0.062 (with a standard deviation of 0.313). The observed decline in Miami, however, was -0.741 log points, larger than the drop observed in any other city. The city with the second largest drop was Milwaukee, where the HWI dropped by -0.643 log points.

---

[10] We use the three-digit *metarea* variable (which defines a metropolitan statistical area) in the IPUMS files of the decennial censuses to gather information about the cities for which the HWI is available. The only exception is for Gary, Indiana (which is officially defined as a "metropolitan division"), where we use the four-digit IPUMS code.

[11] The cities in the synthetic control are Charlotte (with a weight of 0.014), Jacksonville (0.023), Los Angeles (0.338), Memphis (0.141), New Orleans (0.053), and San Antonio (0.432).

Similarly, Miami's HWI exhibited a remarkably strong recovery in the second half of the 1980s. The average growth in the HWI between 1985 and 1989 across all cities was 0.158 log points (with a standard deviation of 0.250). Miami's recovery was the second largest in the sample (0.676 log points), with only Gary, Indiana exhibiting a larger growth (0.702 log points). In short, the distinctive U-shaped trend documented in Figure 3 of the number of job vacancies in Miami during the 1980s was unique, with a remarkable drop in in the first half of the decade (immediately after Mariel) followed by a remarkable recovery in the second half.

To measure the impact of the supply shock on the HWI, we estimate a generic difference-in-differences regression model where the unit of analysis is a city-year-month cell:

$$\log H_{rtm} = \theta_r + \theta_{tm} + \beta(\text{Miami} \times \text{Post-Mariel}) + \varepsilon, \qquad (1)$$

where H_{rtm} is the HWI in city r, year t, and month m; θ_r is a vector of city fixed effects; θ_{tm} is a vector of *interacted* year-month fixed effects (i.e., a fixed effect for every year-month pairing); "Miami" is a dummy variable identifying the Miami-Hialeah metropolitan area; and "Post-Mariel" indicates if cell (r, t, m) was observed after June 1980. The regression uses monthly data from January 1975 through December 1989. Note that the inclusion of the θ_{tm} fixed effects net out the impact of any transitory shock that affected all local labor markets equally (including economy-wide cyclical fluctuations).

The top panel of Table 1 reports the estimated coefficients in the vector β (and robust standard errors) for various specifications of the regression model. Because the graphical analysis shows that the impact of Mariel varies over time, the post-Mariel variable in equation (1) is a vector of fixed effects indicating whether the observation refers to the intervals June 1980-1982, 1983-1984, 1985-1986, or 1987-1989. We first estimate a regression that pools the data across all 51 metropolitan areas. The regression clearly shows a sizable relative decline in Miami's vacancy index (of nearly 50 percent) by the mid-1980s. The second column compares Miami to the synthetic control and shows a decline of over 20 percent by 1981-1982 and of over 40 percent by 1985. Regardless of the specification, the regressions show a full recovery in Miami's HWI by the late 1980s.[12]

[12] Trends in the HWI may also reflect changes in the market share of the paper sampled by the Conference Board in the metropolitan area (perhaps due to entry and exit of competing newspapers). We used circulation data

Modern work on search theory and job vacancies (Pissarides, 1985; and Mortensen and Pissarides, 1994) relies on the fundamental construct of a matching function $M(U, V)$ that gives the number of new hires, M, as a function of the number of unemployed workers, U, and the number of job vacancies, V. If the matching function has constant returns to scale, the job-finding rate (i.e., the probability M/U that an unemployed worker finds a job) can be written as a function of the ratio V/U. The job-finding rate is frequently interpreted as a measure of labor market tightness. Following Shimer (2005), we proxy the job-finding rate by the ratio of the HWI to the unemployment rate at the city-year-month level.[13]

The BLS began to report the monthly unemployment rate for the metropolitan areas used in our analysis in 1976.[14] The bottom panel of Figure 3 illustrates the trends in the job-finding rate in Miami and in the synthetic control for the 1976-1989 period.[15] The figure shows a rapid decline in Miami's job-finding rate soon after Mariel and a full recovery by the late 1980s. The corresponding regressions in the bottom panel of Table 1 show that Miami's job-finding rate dropped by 40 to 50 percent relative to the synthetic control by the mid-1980s (before recovering by the end of the decade).

The analysis of the job-finding rate raises the interesting question of how the unemployment rate in Miami responded to the Mariel supply shock. After all, our proxy for the job-finding rate might decline even if there was no change in the tightness of the labor market (as measured by the unemployment rate) if the HWI does not measure job vacancies accurately. For instance, there is a possibility that the number of help-wanted ads published in the *Miami Herald* may not truly reflect labor market conditions because the demographic shift in the workforce

---

from Gentzkow, Shapiro, and Sinkinson (2011) to determine if changes in market share could be driving the different trends observed in the number of help-wanted ads published in the *Miami Herald* and in the newspapers in the synthetic control. The market share of the *Miami Herald* increased from 78.5 to 87.3 percent between 1978 and 1984, which was roughly similar to the increase in the market share of the newspapers in the synthetic control (62.8 to 70.9 percent).

[13] It is a proxy because not all job vacancies are advertised in newspapers. We make the standard assumption that the HWI captures variation in the true level of vacancies plus mean-preserving noise.

[14] The data are available in the monthly *Employment and Earnings* series published by the BLS. The 1976-1989 unemployment series is not complete for 29 of the 51 cities sampled by the Conference Board. Our analysis of the job-finding rate uses data for 48 metropolitan areas. We excluded cities where the unemployment rate is missing for more than 2 consecutive months and linearly interpolated the missing unemployment rates.

[15] The job-finding rate for each city is rescaled to equal 1 in the 1977-1979 period. The control cities (weights) forming the synthetic control are Jacksonville (0.106), Los Angeles (0.348), Memphis (0.101), New Orleans (0.056), Providence (0.014), and San Antonio (0.376).

encouraged employers to pursue alternative ways of finding workers and filling vacancies. It turns out, however, that the decline in the job-finding rate cannot be solely attributed to trends in the HWI, as Miami's unemployment rate (relative to the control cities) also rose during the early 1980s.

The top panel of Figure 4 shows the close link between the HWI and the unemployment rate in Miami. The correlation coefficient between the two series is -0.86 before 1980 and -0.80 after 1980. As with the nationwide trends illustrated in Figure 1, the local HWI in Miami closely tracks the economic conditions captured by the local unemployment rate. In addition, as the bottom panel of Figure 4 shows, Miami's unemployment rate in the post-Mariel years rose relative to the synthetic control.[16] Miami's unemployment rate was roughly similar to that of the control cities in 1979 (5.9 percent in Miami and 5.6 percent in the synthetic control). By 1983, Miami's unemployment rate had risen to 9.8 percent, while the unemployment rate in the synthetic control rose to 7.9 percent.[17] In short, the observed decline in the HWI in post-Mariel Miami coincides with the deterioration in local labor market conditions.

The trend in the unemployment rate underlines the fact that the Mariel boatlift occurred shortly before the onset of the 1981-1982 recession.[18] The HWI has an obvious cyclical trend, and the regression specification in equation (1) nets out cyclical shocks that affect all localities equally. However, the short-term cyclical fluctuations may vary across metropolitan areas (perhaps *because* of immigration-induced supply shocks). It is easy to illustrate what happens if we net out city-specific transitory fluctuations from the vacancy index by applying the Hodrick-Prescott (HP) filter to our data.

In a panel data context, this widely used filter decomposes the HWI time series for each metropolitan area into a long-run trend and a short-term cyclical component. Note that because the HP filter is applied individually to the time series in each metropolitan area, it ignores the possibility that the intensity of the cyclical fluctuations in a particular city (e.g., Miami) might themselves have been affected by supply shocks. Netting out city-specific transitory fluctuations,

---

[16] The control cities (weights) forming the synthetic control in the unemployment analysis are Jacksonville (0.234), Los Angeles (0.294), Omaha (0.033), Providence (0.055), and San Antonio (0.376), and San Diego (0.039).

[17] The increase is statistically significant: Miami's relative unemployment rate increased by 1.67 percentage points (with a standard error of 0.27).

[18] It also coincided with the short recession between January and July 1980.

therefore, might remove some of the effect of supply shocks on the locality and produces a lower-bound estimate of the impact of the shock.

The two panels of Figure 5 show the behavior of the adjusted HWI and of the adjusted job finding rate in Miami relative to the synthetic control.[19] The exercise obviously removes a lot of the cyclical variation in the HWI for the group of cities that make up the synthetic control. It is also obvious, however, that the post-1980 Miami experience was very different from what was observed elsewhere. The long-run trend in Miami's HWI was obviously dislodged after Mariel. As the regression reported in the last column of Table 1 indicates, the permanent level of the HWI in Miami declined by about 30 percent relative to the synthetic control by the mid-1980s.[20]

An intuitive interpretation of the evidence is that a large supply shock composed of newly arrived immigrants makes it easier for firms to fill existing vacancies, leaving few available jobs for the pre-existing workforce. Over time, of course, firms would be expected to adjust to the supply shock by increasing their capital stock, leading to a corresponding increase in labor demand and a re-equilibration of the labor market. However, as noted by Hamermesh and Pfann (1996, p. 1264), "businesses change their demand for inputs more slowly than the shocks to input demand warrant." The adjustment costs faced by firms as they adjust their inputs to the changed circumstances suggest that the number of job vacancies may not increase sufficiently fast after the supply shock, resulting in a relatively low number of job vacancies for some time.

An alternative (and mechanical) explanation of the evidence is that because of the large and sudden increase in the size of the Cuban community in the post-Mariel period, Miami's employers began to shift their advertising to other venues, such as Spanish-language outlets.[21] In fact, in 1977 the *Miami Herald* began to publish *El Nuevo Herald*, a sister newspaper in Spanish.

---

[19] The long-term trend for the synthetic control in Figure 5 is calculated using a two-step approach. We first remove the impact of the short-run cyclical fluctuation from the HWI time series. We then use the adjusted HWI data to find the synthetic control and generate the trends summarized in the figure. The synthetic control method, therefore, compares Miami with cities that had similar permanent trends prior to the supply shock.

[20] In addition to the Mariel boatlift, Miami received other large waves of Cuban immigrants between 1960 and 2000. Anastasopoulos et al (2019) show that the short-run response of the HWI in Miami to these other shocks was similar to what we find in the Mariel episode: A relative decline in the index of job vacancies within two or three years after the immigrant influx.

[21] As noted earlier, the trend in the HWI may not be a good proxy for the trend in job vacancies when there is a substantial change in workforce composition and an associated shift in the distribution of alternative methods of job search.

The Conference Board did not enumerate help-wanted ads in *El Nuevo Herald*, perhaps leading to a decline in the HWI in post-Mariel Miami.

To determine the validity of the substitution hypothesis, we counted the total number of ads published by *El Nuevo Herald* in each December between 1978 and 1984.[22] Figure 6 illustrates the trend in Miami's HWI and in the number of ads published in *El Nuevo Herald*. The HWI dropped by 62.7 percent (from 126 to 47) between December 1979 and December 1982. The number of ads published in *El Nuevo Herald* dropped by 70.1 percent in the same period (from a monthly total of 2,510 to 750 ads). The data indicate, therefore, that the number of job vacancies advertised directly to Miami's Spanish-speaking community fell by *more* than the number of vacancies in the aggregate Miami labor market.[23]

Although the objective of this paper is to provide empirical evidence of the strong link between job vacancies and immigration, it is worth noting that there are several theories in macroeconomics that can explain such a link. These theories, however, do not necessarily agree on what happens to job vacancies after a labor supply shock.

Shimer's (2007, Proposition 3, p. 1082) "mismatch" theory predicts that the number of job vacancies per worker will decline in equilibrium if there is an exogenous increase in the number of workers. The logic is straightforward: an economy with more workers yields more matches with existing job vacancies (i.e., increasing employment, reducing vacancies). Firms will also create more jobs if there are more workers in the economy (since the returns to creating a job have gone up), but the mismatch between workers and jobs induces an increase in the number of jobs that is too small to offset the first-order effect of additional workers filling existing vacancies. Since firms create relatively few jobs in response to an increase in the stock of workers, the permanent unemployment rate also increases in response to the supply shock.

---

[22] Specifically, we accessed issues of *El Nuevo Herald* archived at the Library of Congress in microfilm and manually counted the number of ads that appeared between December 1 and December 30 of each year between 1978 and 1984. As with the calculation of the HWI, a help-wanted ad in *El Nuevo Herald* that advertised several job openings is counted as a single vacancy.

[23] One hypothesis that is consistent with the decline in both *El Nuevo Herald* and the HWI is that the supply shock led firms to utilize pre-existing Cuban or Spanish-speaking social networks to fill jobs instead of the help-wanted section of newspapers. Pholphirul (2012) provides suggestive evidence that firms which have used immigrants to fill job openings in the past are more likely to do so in the future. Note that from the perspective of an unemployed non-Hispanic worker, the reduction in the number of posted job ads has the same implications regardless of whether it arises from faster vacancy filling or from a substitution from published job advertisements to unobserved social networks.

In contrast, the canonical Diamond-Mortenson-Pissarides (henceforth DMP) model is more ambiguous about how vacancies respond to one-time supply shocks (Pissarides, 2000). In the DMP model, which uses the matching function introduced earlier, exogenous parameters fix the ratio of vacancies to unemployment. Let market tightness be defined as V/U. We can normalize both the number of vacancies and the number of unemployed workers by the size of the labor force, so that market tightness is given by the ratio of the vacancy rate to the unemployment rate v/u. The DMP model predicts that, for fixed values of parameters, v/u equals a constant θ.

If immigrants enter the model as unemployed workers (i.e., initially increasing the unemployment rate) and have no other impact on the model's parameters, the vacancy rate would then be predicted to jump upwards in equilibrium. The intuition is that the returns to creating a job vacancy have increased, so that firms post vacancies until those returns are driven back down to zero. As the new workers obtain jobs, unemployment decreases and so should the number of vacancies outstanding, with the vacancy rate returning to its original level.

The evidence of a strong negative relation between the job-finding rate (as proxied by the ratio of the HWI to the unemployment rate) and supply shocks is consistent with the DMP model only if the Mariel shock changed one of the other parameters of the model (Cahuc, Carcillo, and Zylbergberg, 2014, p. 598). These other parameters include the reservation wage of unemployed workers, the discount rate, the cost of posting a vacancy, or the output generated by a worker-firm pair. For instance, if the reservation wage of unemployed workers fell because the new refugees are more willing to take on the available jobs, the DMP model predicts that the equilibrium job-finding rate would *rise*. In contrast, if the new immigrants reduced the output generated by a particular match, then the job-finding rate would *fall*. In short, a reconciliation of the evidence with the DMP framework requires a detailed examination of how supply shocks shift the "fundamentals" of that model.[24]

IV. The Textual Content of Help-Wanted Ads

[24] Attempts to distinguish between the mismatch and canonical models of labor market frictions have focused on their quantitative predictions about variable co-movements in equilibrium (Shimer, 2007; Sahin et al, 2014). Our reduced-form results suggest that a structural exploration of the response of these models to supply shocks might yield qualitatively divergent (and more easily testable) predictions.

Although the previous section documents a short-term decline in the relative *number* of job vacancies in post-Mariel Miami, the trends in the HWI do not provide information on *which* types of job vacancies vanished. Because of its skill composition, the Mariel supply shock should have particularly affected the low-skill labor market. To determine whether these are indeed the vacancies that were most affected by the increased supply of workers, we need to examine the text of the help-wanted ads. Modern machine learning and natural language processing techniques allow us to extract the relevant content of help-wanted classifieds from the published text of a newspaper. As the pioneering work of Atalay, Phongthiengtham, Sotelo, and Tannenbaum (2020) shows, the analysis of the text of help-wanted ads can provide important insights into how labor markets evolve.

A critical obstacle limits the scope of such a study: digitized files of the text of help-wanted classifieds exist for only a small number of newspapers, and it is an arduous and costly undertaking to expand this small set.[25] We constructed a sample of help-wanted ads from the *Miami Herald* by selecting a random sample of issues published between 1978 and 1984 (essentially one issue from every month in that period).[26] For each of these issues, we (manually) produced digital images of the help-wanted pages from the microfilm archive maintained at the Library of Congress. Using a crowd-sourced, natural language processing pipeline, each ad was then transcribed into text and compiled into a single database, creating a random sample of 95,263 ads. Online Appendix B describes the process of creating this sample of ads from the *Miami Herald* (as well as two other newspapers discussed below).

The next step involves classifying each ad in terms of the type of job being advertised. This classification is straightforward as the Department of Labor commissioned a private firm to build an algorithm, the O*NET-SOC AutoCoder, that classifies job titles and other job information into standard occupational codes. Specifically, we "fed" the entire text of each of the

---

[25] For example, Atalay et al (2020) use only the ads published in the *Boston Globe*, the *New York Times*, and the *Wall Street Journal* to document the changing occupational structure of the labor market. A private company (ProQuest) archives the digital version of these newspapers and makes the data accessible to university researchers. Unfortunately, ProQuest does not include the *Miami Herald* in its archive. The private firm that archives the *Herald* digitally (NewsBank) does not generally make their data accessible to researchers at educational institutions at a non-prohibitive price.

[26] We first picked a random date 2 to 3 years after Mariel: December 20, 1982. Once that date was determined, we worked backwards to find a similarly situated date before Mariel: December 18th, 1978. From these dates, the pre-Mariel and post-Mariel samples were constructed by selecting issues published every 29 days from the initial pre- and post-Mariel dates between January 1978 to December 1984.

help-wanted ads in our sample to the algorithm, which returned the corresponding SOC-2010 occupation code for that vacancy. The algorithm is very reliable when conducting this type of matching; it "guarantee(s) 85% accuracy for codes assigned to job ads (titles plus descriptions)" (Wilson, 2019).[27]

Once each help-wanted ad is assigned an occupation code, it is easy to create a number of variables that describe the skills required by the vacancy. We use several alternative characterizations of the occupation's skill level: 1. Whether the vacancy advertises a blue-collar (or service) job opening; 2. The mean years of educational attainment of workers employed in the occupation; 3. The percent of workers in the occupation without a high school diploma; and 4. The "occupational wage," defined as the (adjusted) log hourly wage of workers in the occupation.[28] All of these variables were created using the sample of workers aged 25-64 in the 1980 Census.

The various panels of Figure 7 illustrate the trends for each of these skill measures. To minimize month-to-month sampling noise, we aggregate the *Miami Herald* data to the annual level. Consider initially the trend in the relative frequency of ads for blue-collar vacancies. Prior to Mariel, about 38 percent of vacancies advertised in the *Miami Herald* were for blue-collar positions. By 1982, only about 30 percent of the ads were for blue-collar jobs. Similarly, the mean years of education of workers typically employed in the advertised job rose from 12.2 years before Mariel to 12.5 years by 1982. The average occupational wage in the advertised vacancy rose by 3 to 4 percentage points. Figure 7 unambiguously reveals that the skill composition of the help-wanted ads appearing in the *Miami Herald* changed in the post-Mariel period, with a noticeable decline in the relative number of vacancies for low-skill workers.

The observed trends in Miami should be contrasted with trends in other cities to determine whether the evolution of the occupational distribution of Miami's job vacancies was unique. We obtained access to the sample of help-wanted ads Atalay et al (2020) created for the

---

[27] Using a standard concordance table, we converted the SOC-2010 code into the 1990 Census occupation code (which is the code we use when merging all the data sets used in this section).

[28] A blue-collar or service job is in the precision production, craft, and repair occupations (with a code between 503 and 699), or the operators, fabricators, and laborers occupations (with a code between 703 and 889), or the service occupations (with a 1990 occupation code between 405 and 469). The occupational wage is constructed by averaging (within each occupation) the residuals from a regression of a worker's log hourly wage on a vector of age fixed effects, and on indicators for race and gender.

New York Times and the *Boston Globe*. These data reported the number of ads published in each occupation-year-month-day cell for each newspaper.[29] We then collected comparable data for two other newspapers, whose digital archives are maintained by ProQuest. These two other newspapers, the *Minneapolis Star-Tribune* and the *St. Louis Post-Dispatch*, were chosen because they sometimes appeared as "control cities" when we estimated synthetic control models of the 1975-1989 trend in the HWI using the small subset of newspapers that we could potentially use in the analysis.[30]

The ProQuest files for these newspapers include their *entire* published content. Using the machine-learning methods developed by Atalay et al (2020), we isolated the help-wanted ads in each issue published between 1978 and 1984. This procedure created a sample of 253,833 ads for Minneapolis and 148,520 ads for St. Louis. We then predicted the ad's occupation code by running each ad through the O*NET-SOC AutoCoder. In sum, the control group consists of help-wanted classifieds published in the *Boston Globe*, the *Minneapolis Star-Tribune*, the *New York Times*, and the *St. Louis Post-Dispatch*.

Figure 7 also shows the trend in the skill characteristics of ads published in the control group. It is visually obvious that the Miami trend is noticeably different from the trend in the control. At the same time that the fraction of blue-collar ads dropped rapidly in Miami between 1980 and 1982, the respective fraction rose in the control cities. Similarly, while the average education or the occupational wage of job vacancies advertised in the *Miami Herald* was rising after Mariel, the respective statistic in the control cities was stable.

---

[29] The Atalay et al (2020) data file consists of 1,211,616 ads in the *New York Times*, and 1,060,920 ads in the *Boston Globe*. The *Boston Globe* data, however, do not contain any ads for calendar year 1984. The data file also contains comparable information for ads published in the *Wall Street Journal*, but we do not use that newspaper in our analysis. The ads published in the *Wall Street Journal* are typically targeted to a very different sector of the labor market than the ads published in other newspapers.

[30] Putting aside the *Boston Globe* and the *New York Times*, we purchased the entirety of ProQuest newspaper holdings that satisfied the following restrictions: (a) the newspaper was in the HWI sample; (b) the newspaper had already been digitized and Optical Character Recognition (OCR) software had been applied so that we could, in theory, work directly with the text; and (c) the digital files existed for the years 1978 through 1984. The few newspapers that satisfied these restrictions at the time of purchase were the *Detroit Free Press*, the *Louisville Courier-Journal*, the *Minneapolis Star-Tribune,* the *Philadelphia Inquirer*, the *Pittsburgh Post-Gazette*, and the *St. Louis Post-Dispatch*. We estimated regression models similar to those reported in the previous section using this small subset of potential control cities, and the Minneapolis and St. Louis papers were the ones that typically showed up as part of the synthetic control.

Table 2 reports the relevant coefficients from the difference-in-differences regression in equation (1) that estimates the impact of the Mariel supply shock on the skill composition of advertised vacancies. The unit of analysis in the regression is a newspaper-year-month cell, and the regression includes newspaper and year-month fixed effects. We estimated both unweighted and weighted regressions (where the weight is the number of ads in the cell). The table reports the interaction coefficients between the Miami indicator and the post-Mariel time period.

The regressions consistently show a significant post-Mariel effect in the skill composition of vacancies in Miami, with much of the effect disappearing by 1984. The relative number of blue-collar vacancies in Miami declined by 6.7 percentage points by 1982; the fraction of workers in the (typical) advertised job who lacked a high school diploma dropped by 2.8 percentage points; the educational attainment of the average occupation being advertised increased by 0.2 years; and the occupational wage rose by 3 percent. In short, a regression analysis that examines the *textual* content of advertised vacancies confirms a substantial drop in the relative availability of low-skill jobs in post-Mariel Miami.[31]

V. Beyond Miami

Up to this point, our analysis has focused on documenting the response of job vacancies in Miami to the Mariel supply shock.[32] Immigration also affected many other cities, with the annual number of immigrants entering the United States increasing from about 250,000 in the 1950s to over 1 million legal and undocumented immigrants by 2000.

A large literature exploits the geographic distribution of immigrants to estimate the labor market impact of immigration. This literature essentially correlates some economic outcome in the city—typically the average wage of some group of workers—with the number of immigrants in that city. Although some studies report a weak spatial correlation (suggesting that native outcomes may not be strongly affected by immigration), the correlation is contaminated by the endogenous settlement of immigrants in high-wage cities. This sorting makes it difficult to detect

---

[31] The O*NET-SOC AutoCoder also reports a "match score," which measures the reliability of the occupation allocation. The data we obtained for New York and Boston, however, do not contain that score. Nevertheless, we estimated regressions in the *Miami Herald* data using a simple difference specification of equation (1) and weighed the observations by the average match score of the month-year cell. The results are very similar to those that ignore the match score altogether.

[32] The working paper version of this article (Anastasopoulos et al, 2019) shows that job vacancies in Miami also declined after other supply shocks of Cuban immigrants in the early 1960s and mid-1990s.

the potential adverse effect of supply shocks on the wage of competing workers. Moreover, the widely used shift-share instrument (which is a nonlinear transformation of lagged immigration in the locality) may not truly solve the endogeneity problem and may understate the adverse wage impact of supply shocks (Jaeger, Ruist, and Stuhler, 2018).

Given the dominance of this methodological approach (and the rarity of "experimental" Mariel-like supply shocks), it is of interest to estimate the analogous spatial correlation between the HWI and immigration. Ideally, we would have monthly or yearly data on the number of immigrants arriving in each of the cities surveyed by the Conference Board over a multi-decade period. Unfortunately, the CPS did not begin to collect immigration data until 1994.[33]

We instead use the decennial censuses to calculate the immigration-induced supply shock in each of the 51 cities. We then correlate these supply shocks with the decadal change in the city's HWI. The regression model is given by:

$$\log \frac{H_{r,\tau}}{H_{r,\tau-1}} = \theta_{\tau} + \gamma \frac{Immigration\ into\ city\ r}{Baseline\ number\ of\ natives\ in\ city\ r} + \eta\,, \qquad (2)$$

where $H_{r,\tau}$ gives the HWI for city r in census year τ (τ = 1970,…,2000). The HWI index for census year τ is defined as the average HWI observed in the three-year interval around τ. For example, the average HWI for Rochester in census year 1980 is the average HWI reported monthly for Rochester between 1979 and 1981. The two variables used to measure the supply shock (the number of immigrants, and the number of natives) give population counts for persons aged 18-64 in a particular city. There seems to be some confusion as to how the ratio measuring the supply shock in equation (2) should be defined (Borjas, 2014; Borjas and Monras, 2017; and Card and Peri, 2016). We will use alternative definitions of the supply shock to demonstrate the robustness of our estimated correlation.

The same endogeneity problem that plagues estimates of the link between wages and immigration also plagues the regression in equation (2). Immigrants are more likely to settle in cities where there are jobs to be had and employers are actively searching for workers. This

[33] The monthly data from the Job Openings and Labor Turnover Survey (JOLTS) since 2000 could, in principle, be used to estimate the spatial correlations in the post-2000 period. The JOLTS vacancy rate statistics, however, are only available for only 18 metropolitan areas, are model-based, and are still in an experimental stage.

endogeneity builds in a positive correlation between the change in the HWI and the number of immigrants settling in that city during the decade.

Table 3 reports estimates of the coefficient γ using a number of alternative specifications. Consider initially the regression reported in the first row. The supply shock is defined as the ratio of the number of immigrants who settled in the city between census year τ-1 and τ, or $M(\tau, \tau\text{-}1)$, to the number of natives residing in the city at time τ-1, or $N(\tau\text{-}1)$.[34] This specification ignores the fact that there may have been a supply response as natives moved in or out of immigrant-receiving cities. The estimated spatial correlation is negative and significant, suggesting that a 10 percent increase in supply reduces the HWI by about 10 percent.

We address the endogeneity issue by using the shift-share instrument (as constructed by Jaeger, Ruist, and Stuhler, 2018).[35] In rough terms, the instrument uses the geographic settlement of earlier immigrant waves belonging to particular national origin groups to allocate new immigration across the cities. The validity of the instrument obviously hinges on whether the local economic conditions that attracted immigrants to a particular place persist for some time after arrival. The IV estimate of the impact is also negative and significant, and larger in absolute value than the OLS estimate.

The other rows of Table 3 use alternative definitions of the supply shock. In row 2, the supply shock is the ratio $M(\tau, \tau\text{-}1)/N(\tau)$. It differs from the definition in the first row because it does not lag the native baseline (thereby allowing for a potential native supply response). Finally, the supply shock in the last row is defined as the difference $\left(\frac{M_{c,\tau}}{N_{c,\tau}} - \frac{M_{c,\tau-1}}{N_{c,\tau-1}}\right)$.[36] Regardless of how we define the supply shock, Table 3 documents a negative correlation between the change in the HWI and immigration, with an elasticity of about -2. The spatial correlation approach, therefore, confirms the basic lesson of the Mariel episode: the number of job vacancies falls after an immigration-induced supply shock.

---

[34] The number of immigrants who migrated to the United States in the interval (τ−1, τ) and settled in city *r* is obtained from the decennial census in year τ.

[35] We are grateful to David Jaeger, Joakim Ruist, and Jan Stuhler for sharing their data. We use the variant of their instrument that uses the settlement of immigrants as of 1960 to predict the geographic sorting of new arrivals in subsequent census years.

[36] If there were only two cross-sections, the regression relating the first difference in the HWI and the first difference in the ratio *M*/*N* is numerically equivalent to a panel regression where the level of the HWI in city *r* and census year τ is regressed on sets of city and year fixed effects and on the ratio *M*/*N*.

VI. Macro Implications: The Beveridge Curve

The matching function (with constant returns), $M(U, V)$, is a key building block in the canonical derivation of the Beveridge curve, the downward-sloping steady-state relationship between job vacancies and unemployment (Elsby, Michaels, and Ratner, 2015). The DMP framework implies that as long as a supply shock is composed of persons who enter the labor market as unemployed workers, the vacancy-unemployment locus should immediately shift outwards (see Blanchard and Diamond, 1989, p. 17). Unemployment increases instantaneously. Firms then respond by increasing the number of vacancies. The larger number of unemployed workers and of vacancies increases the number of matches, helping to reduce both vacancies and unemployment. The Beveridge curve would then tend to shift back to its original position. In short, the simplest model of the Beveridge curve would predict an immediate outward shift in the vacancy-unemployment locus in post-Mariel Miami, followed by a gradual movement back to its original position.[37]

The top panel of Figure 8 illustrates the data scatter forming the national Beveridge curve (using the national HWI and the national unemployment rate) in the two 5-year periods before and after Mariel. The national Beveridge curve shifted out after 1980, concurrent with the deep 1981-1982 recession. The middle panel shows the Beveridge curve for the synthetic control, calculated by taking a weighted average of the unemployment and vacancy rates across metropolitan areas (using the weights implied by the synthetic control analysis of the permanent trend in the job-finding rate. The Beveridge curve for the synthetic control also shifted out. Finally, the bottom panel shows the data for Miami, which generates exactly the opposite pattern, an *inward* shift in the Beveridge curve. The common interpretation of such an inward shift is that the Mariel supply shock made Miami's labor market *more* efficient in the sense that there is less unemployment for a given number of vacancies.[38]

---

[37] Although most discussions of the Beveridge curve focus on national labor markets, Courtney (1991) and Valletta (2005) use the local labor market as the unit of analysis for estimating Beveridge curve regressions, with the vacancy-unemployment relationship differing across metropolitan areas or regions. Bonthuis, Jarvis, and Vanhala (2016) use a similar approach in their examination of Beveridge curve shifts in the European Union. Finally, Anderson and Burgess (2000) estimate the matching function using the local HWI and state-level unemployment data.

[38] Warren (1982) examines the link between immigration and the Beveridge curve in the Australian context. He finds little evidence that changes in Australian immigration policy have shifted the Beveridge curve.

We can estimate the shift in Miami's Beveridge curve in the first half of the 1980s relative to either the national data or the synthetic control. The regressions that estimate these shifts are:

$$\text{National:} \quad \log H_{rtm} = \theta_r + \theta_{tm} - \underset{(0.013)}{0.082}\, u_{rtm} - \underset{(0.028)}{0.273}\,(\text{Miami} \times \text{Post-Mariel}), \qquad (3a)$$

$$\text{Synthetic:} \quad \log H_{rtm} = \theta_r + \theta_{tm} - \underset{(0.014)}{0.117}\, u_{rtm} - \underset{(0.029)}{0.345}\,(\text{Miami} \times \text{Post-Mariel}), \qquad (3b)$$

where u_{rtm} gives the unemployment rate in city $\underline{r}$, year t, and month m; θ_r denotes a vector of city fixed effects; θ_{tm} denotes a vector of year-month fixed effects (i.e., a fixed effect for each year-month combination); the "post-Mariel" period goes from June 1980 through December 1984; and both regressions have 216 observations; The Mariel supply shock shifted down Miami's Beveridge curve by about 30 percent relative to how the curve was behaving in other cities at the time.

The inward shift in Miami's Beveridge curve is still observed when Miami is compared to all other cities in the HWI sample. In particular, we stacked the data across all metropolitan areas over the 1976-1984 period (yielding 5,184 observations) and estimated the model:

$$\text{All cities:} \quad \log H_{rtm} = \theta_r + \theta_{tm} - \underset{(0.002)}{0.062}\, u_{rtm} - \underset{(0.029)}{0.179}\,(\text{Miami} \times \text{Post-Mariel}). \qquad (3c)$$

The pooled data again document an inward shift in Miami's Beveridge curve of almost 20 percent relative to all other cities.

It turns out, however, that measuring the average shift across two 5-year periods hides detectable short-term swings in the relative placement of Miami's Beveridge curve. Following Valletta (2005), we examine the post-Mariel changes in the relative position of that curve by estimating a model where the interactions between Miami and the post-Mariel period are instead done year by year. Table 4 reports the value and standard errors of the (annual) interaction coefficients, while Figure 9 illustrates the year-to-year variation. Regardless of the specification, the vacancy-unemployment locus in Miami seems to have shifted out immediately after Mariel. If we use the national data, for example, the intercept of Miami's curve shifted out by about 0.12

log points (with a standard error of 0.029). This outward shift, however, did not last long. By 1982, the relative intercept of Miami's locus was strongly negative in all specifications, reaching a nadir in 1985. Interestingly, all specifications show a "return to normalcy" in Miami's vacancy-unemployment locus by the end of the decade; the relative Beveridge curve for Miami in 1989 was back at its starting pre-Mariel position.

Although the Mariel evidence shows the immediate outward shift in the vacancy-unemployment locus implied by the simplest search-theoretic approach, that model cannot account for its eventual inward shift.[39] However, heterogeneity in search behavior between immigrants and natives might shift the Beveridge curve because the matching function of the new arrivals may differ from that of pre-existing workers (Barnichon and Figura, 2015). The supply shock could, for example, lead to more efficient matching if the immigrants search more aggressively than natives or if Spanish-language social networks became more effective at matching employers and workers post-Mariel.

This assumed search behavior by unemployed immigrants ensures a more rapid filling of vacancies after the supply shock. Interestingly, the data suggest that the resulting "efficiency gain" in Miami's labor market vanished after a decade. The eventual return of Miami's Beveridge curve to its pre-Mariel position could perhaps be reflecting the process of immigrant assimilation, with the search behavior of immigrants becoming more like that of natives over time.

VII. Summary

This paper addresses a central economic question in the economic analysis of immigration: How do supply shocks affect the labor market in receiving countries? We contribute to the literature by exploiting a data set that has a long history in economics but has not been employed in the immigration context, the Conference Board Help-Wanted Index (HWI). Beginning in 1951, the Conference Board constructed an index of job vacancies in local labor markets by counting the number of help-wanted classified ads in newspapers in 51

[39] Shimer (2007) proposes an alternative micro foundation for the Beveridge curve, based on a mismatch model where the firms with vacancies are not located in the same markets as the unemployed workers looking for jobs. The mismatch framework predicts that a supply shock produces an inward shift in the Beveridge curve (see Anastasapoulos et al, 2019, Appendix C).

metropolitan areas. It is well known that the HWI provides valuable information about trends in local labor demand and is highly correlated with various measures of labor market conditions.

We use the HWI to document how immigration affects job vacancies in labor markets affected by supply shocks. We exploit both the natural experiment produced by the random and sudden Mariel labor supply shock in the city of Miami, as well as estimate spatial correlations that measure how local trends in job vacancies are related to immigration-induced supply shifts. Our findings include:

1. The labor market in Miami responded strongly to the Mariel supply shock in 1980. The HWI dropped relative to the trend observed in control groups between 1980 and 1985. Miami's HWI recovered fully by 1990.
2. We examined the text of a large sample of help-wanted ads published in the *Miami Herald* and other newspapers between 1978 and 1984. The textual analysis documented that the drop in the number of job vacancies in Miami was particularly severe for low-skill jobs.
3. There is a negative cross-city correlation between the change in the HWI and the number of immigrants entering the local labor market. The measured spatial correlation is negative and significant despite the obvious endogeneity bias created by the non-random settlement of immigrants in cities where there are job openings.
4. At the same time that the Beveridge curve was shifting out in other cities in 1982-1985, Miami's Beveridge curve shifted in. The inward shift might have been caused by reduced mismatch between workers and available job opportunities or by the changed demographics of the Miami labor market, with the Mariel refugees perhaps searching more aggressively and more quickly filling up vacancies. Miami's Beveridge curve returned to its pre-Mariel (relative) position by the late 1980s.

In sum, our evidence consistently indicates that immigration-induced supply shocks are typically followed by a short-run drop in the number of advertised job vacancies, perhaps because the existing positions were quickly filled by the immigrants and firms did not expand fast enough to take advantage of the increase in labor supply. The labor market, however, tends to recover after a few years.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association* 105 (June): 493-505.

Abraham, Katharine G. 1987. "Help-Wanted Advertising, Job Vacancies, and Unemployment," *Brookings Papers on Economic Activity* (1): 207-243.

Abraham, Katharine G. and Lawrence F. Katz. 1986. "Cyclical Unemployment: Sectoral Shifts or Aggregate Disturbances?" *J.P.E.* 94 (June): 507-522

Albert, Christoph. 2021. "The Labor Market Impact of Immigration: Job Creation vs. Job Competition," *American Economic Journal: Macroeconomics* 13.

Altonji, Joseph G., and David Card. 1991. "The Effects of Immigration on The Labor Market Outcomes of Less-Skilled Natives." In *Immigration, trade, and the labor market,* edited by John M. Abowd and Richard B. Freeman, 201–234. Chicago: University of Chicago Press.

Anastasopoulos, L. Jason, George J. Borjas, Gavin G. Cook, and Michael Lachanski. 2019. "Job Vacancies and Immigration: Evidence from Pre- and Post-Mariel Miami," Working Paper No. 24580, NBER, Cambridge, MA.

Anderson, Patricia M. and Simon M. Burgess. 2000. "Empirical Matching Functions: Estimation and Interpretation Using State-Level Data," *Review of Economics and Statistics* 82 (February): 93-102.

Andolfatto, David. 1996. "Business Cycles and Labor-Market Search," *A.E.R.* 86 (March): 112-132.

Atalay, Enghin, Phai Phongthiengtham, Sebastian Sotelo, and Daniel Tannenbaum. 2020. "The Evolution of Work in the United States," *American Economic Journal: Applied Economics* 12 (April): 1-34

Autor, David H. 2001. "Wiring the Labor Market," *Journal of Economics Perspectives* 15 (Winter): 25-40.

Barnichon, Regis and Andrew Figura. 2015. "Labor Market Heterogeneity and the Aggregate Matching Function," *American Economic Journal: Macroeconomics* 7 (October): 222-249.

Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller, and Akitaka Matsuo. 2018. "quanteda: An R Package for the Quantitative Analysis of Textual Data," *Journal of Open Source Software* 3 (October): 774.

Blau, Francine D. and Christopher Mackie, eds. 2016. *The Economic and Fiscal Consequences of Immigration*. Washington, DC: National Academies Press.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation," *Journal of Machine Learning Research* 3 (March): 993-1022.

Bonthuis, Boele, Valerie Jarvis, and Juuso Vanhala. 2016. "Shifts in Euro Area Beveridge Curves and Their Determinants," *IZA Journal of Labor Policy*, 5.

Borjas, George J. 2003. "The Labor Demand Curve *Is* Downward Sloping: Reexamining the Impact of Immigration on the Labor Market,' *Q.J.E.* 118 (November): 1335-1374.

Borjas, George J. 2017. "The Wage Impact of the *Marielitos*: A Reappraisal," *Industrial and Labor Relations Review* 70 (October): 1077-1110.

Borjas, George J. 2019. "The Wage Impact of the *Marielitos*: The Role of Race," *Industrial and Labor Relations Review* 72 (August): 858-870.

Borjas, George J. and Joan Monras. 2017. "The Labor Market Consequences of Refugee Supply Shocks," *Economic Policy* 32: 361-413.

Burch, Susan W. and Ruth A. Fabricant. 1968. "A Further Comment on the Behavior of Help-Wanted Advertising," *Review of Economics and Statistics* 50 (May): 278-281.

Cahuc, Pierre, Stéphane Carcillo, and André Zylberberg. 2014. *Labor Economics*, Second Edition. Cambridge, MA: MIT Press.

Card, David. 1990. "The Impact of the Mariel Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review* 43 (January): 245-257.

Chassamboulli, Andri and Giovanni Peri. 2015. "The Labor Market Effects of Reducing Undocumented Immigrants," *Review of Economic Dynamics* 18 (October): 792-821.

Clemens, Michael A., and Jennifer Hunt. 2019. "The Labor Market Effects of Refugee Waves: Reconciling Conflicting Results," *Industrial and Labor Relations Review* 72 (August): 818-857.

Cohen, Malcolm S. and Robert M. Solow. 1967. "The Behavior of Help-Wanted Advertising", *Review of Economics and Statistics*, 49 (February): 108-110.

Courtney, Hugh. 1991. "The Beveridge Curve and Okun's Law: A Re-Examination of Fundamental Macroeconomic Relationships in the US," Ph.D., dissertation, MIT, Cambridge, MA.

Davila, Alberto and Rogelio Saenz. 1990. "The Effect of Maquiladora Employment on the Monthly Flow of Mexican Undocumented Immigration to the U.S., 1978–1982," *International Migration Review*, 24 (February): 96–107.

Elsby, Michael W., Ryan Michaels, and David Ratner. 2015. "The Beveridge Curve: A Survey," *Journal of Economic Literature* 53 (September): 571-630.

Friedman, Susan Krug. 1982. "The Help-Wanted Advertising Index: A Leading Indicator of the Local Economy," *Business Economics* 17 (May): 61-64.

Gentzkow, Matthew, Jesse M. Shapiro, and Michael Sinkinson. 2011. "The Effect of Newspaper Entry and Exit on Electoral Politics," *A.E.R.* 101 (December): 2980-3018.

Jaeger, David A., Joakim Ruist, and Jan Stuhler. 2018. "Shift-Share Instruments and the Impact of Immigration," Working Paper No. 24285, NBER, Cambridge, MA.

Kroft, Kory and Devin G. Pope, 2014. "Does Online Search Crowd Out Traditional Search and Improve Matching Efficiency? Evidence from Craigslist," *Journal of Labor Economics* 32 (April): 259-314.

Medoff, James L. 1983. "U.S. Labor Markets: Imbalance, Wage Growth, and Productivity in the 1970s," *Brookings Papers on Economic Activity* 1: 87-120.

Monras, Joan. 2020. "Immigration and Wage Dynamics: Evidence from the Mexican Peso Crisis," *J.P.E.* 128 (August): 3017-3089.

Monras, Joan. 2021. "Local Adjustment to Immigrant-driven Labor Supply Shocks," Journal of Human Capital, forthcoming.

Ortega, Javier. 2000. "Pareto-Improving Immigration in an Economy with Equilibrium Unemployment," *Economic Journal* 110 (January): 92–112.

Peri, Giovanni and Vasil Yasenov. 2019. "The Labor Market Effects of a Refugee Wave: Synthetic Control Method Meets the Mariel Boatlift," *Journal of Human Resources* 54 (Spring): 267-309.

Pfann, Gerard A. and Daniel S. Hamermesh. 1996. "Adjustment Costs in Factor Demand," *Journal of Economic Literature* 34 (September): 1264-1292.

Pholphirul, Piriya. 2013. "Immigration, Job Vacancies, and Employment Dynamics: Evidence from Thai Manufacturers," *Journal of Asian Economics* 24 (February): 1-16.

Pissarides, Christopher A. 1985. "Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages," *American Economic Review* 75 (September): 676-690.

Pissarides, Christopher A. 2000. *Equilibrium Unemployment Theory*, Cambridge, MA: MIT Press.

Preston, Noreen L. 1977. *The Help-Wanted Index: Technical Description and Behavioral Trends*. Report no. 716. New York: Conference Board.

Şahin, Ayşegül, Joseph Song, Giorgio Topa, and Giovanni L. Violante. (2014). "Mismatch Unemployment," *A.E.R.* 104 (November): 3529-3564.

Shimer, Robert. 2005. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review* 95 (March): 25-49.

Shimer, Robert. 2007. "Mismatch," *American Economic Review* 97 (September): 1074-1101.

Smith, Christopher L. 2012. "The Impact of Low-Skilled Immigration on the Youth Labor Market," *Journal of Labor Economics* 30 (January): 55-89.

Valletta, Robert G. 2005. "Why Has the U.S. Beveridge Curve Shifted Back? New Evidence Using Regional Data," Working Paper #2005-25, Federal Reserve Bank of San Francisco.

Walsh, John, Miriam Johnson, and Marged Sugarman. 1975. *Help-Wanted: Case Studies of Classified Ads*. Salt Lake City: Olympus.

Warren, Donald. 1982. "Immigration and the Natural Rate of Unemployment in Australia," *Journal of Macroeconomics* 4 (Fall): 449-457.

Wilson, Rob. 2019. "Frequently Asked Questions: What is the O*NET-SOC AutoCoder?" R.M. Wilson Consulting, Inc. www.onetsocautocoder.com/plus/onetmatch?action=guide, accessed on August 9, 2019.

Withers, Glenn and David Pope. 1985. "Immigration and Unemployment," *Economic Record* 61 (June): 554-564.

Figure 1. The national Help-Wanted Index (HWI) and the unemployment rate

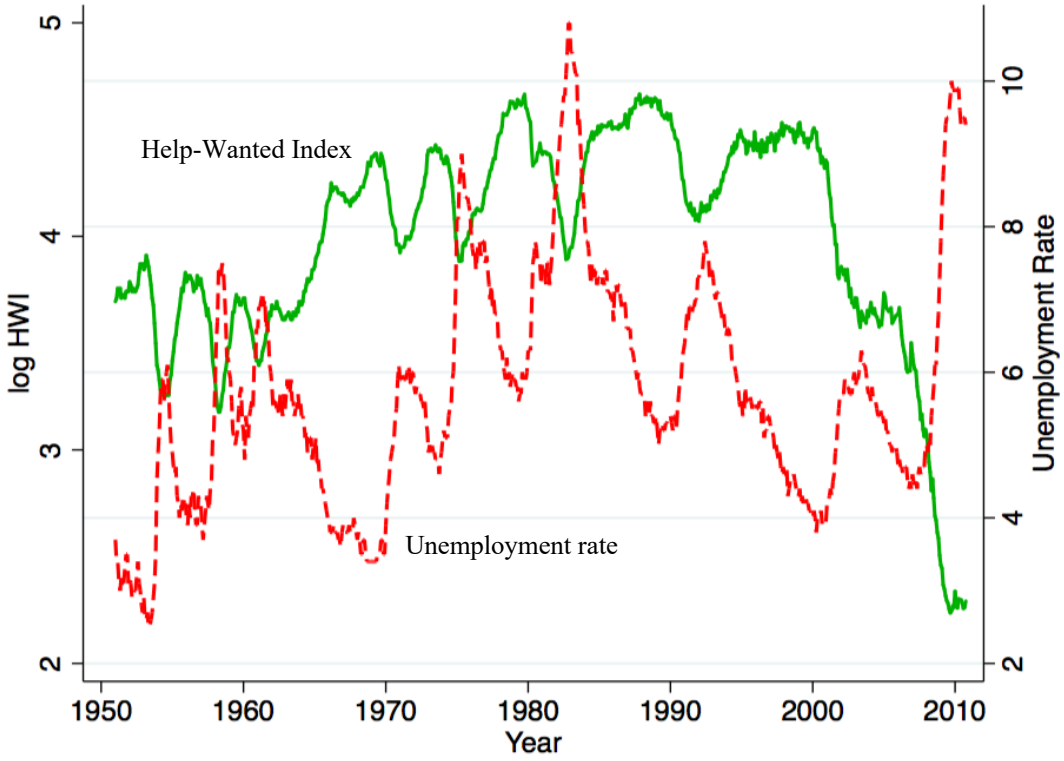
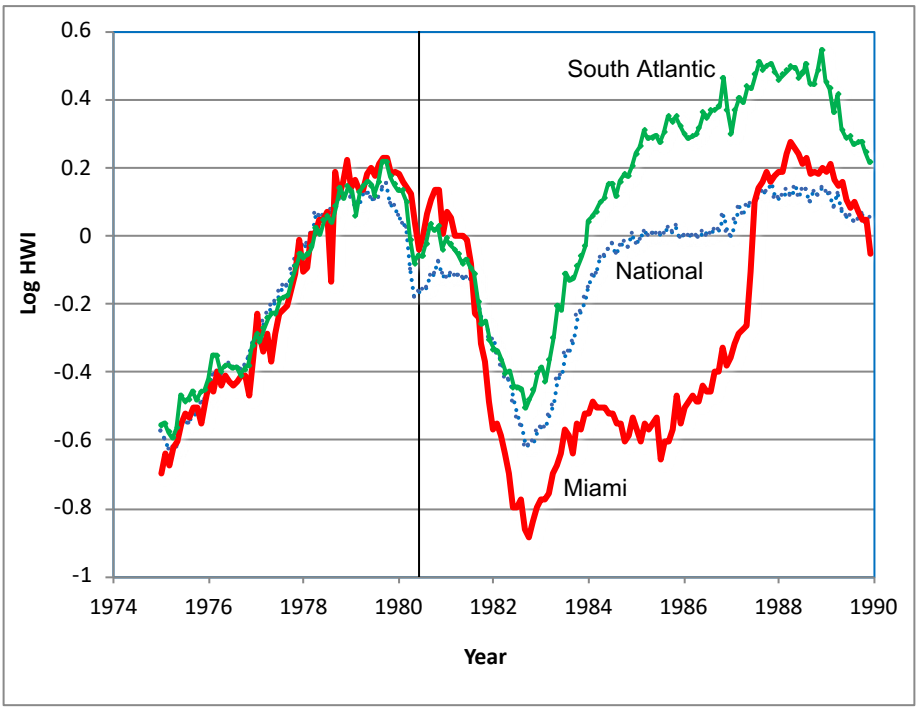


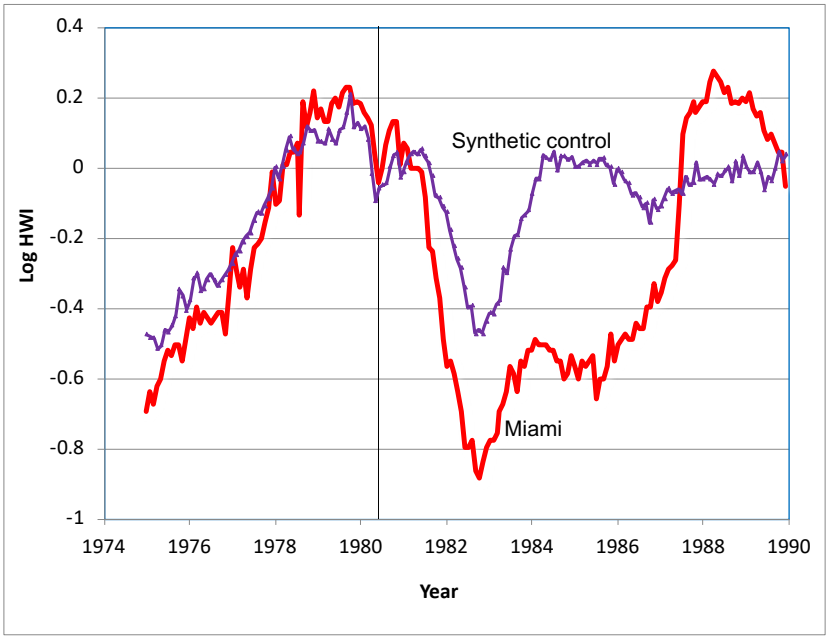
Figure 2. The Help-Wanted Index in Miami, 1975-1989



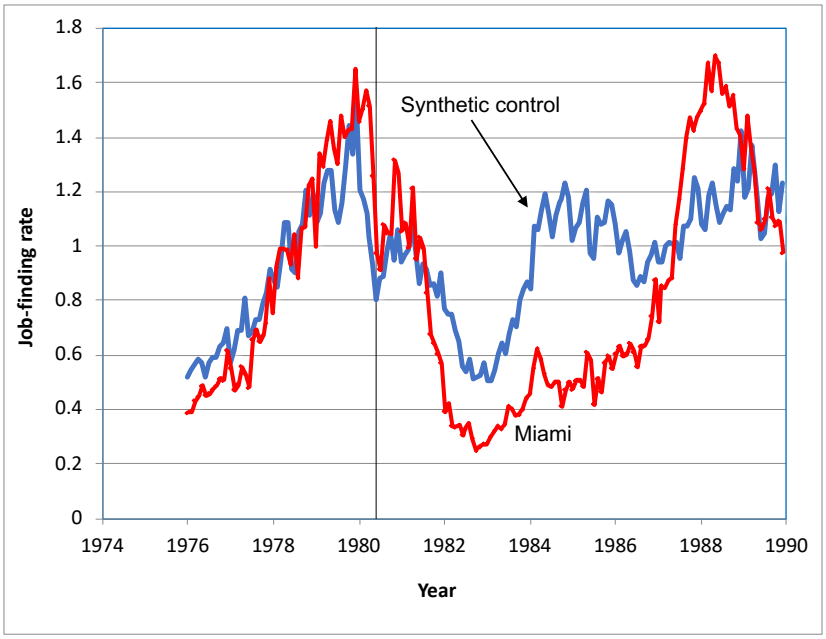
Notes: The HWI for each city/region is rescaled to equal 1 in 1977-1979. The treatment line is drawn as of June 1980.

Figure 3. Job vacancies in Miami relative to the synthetic control, 1975-1989

A. The help-wanted index



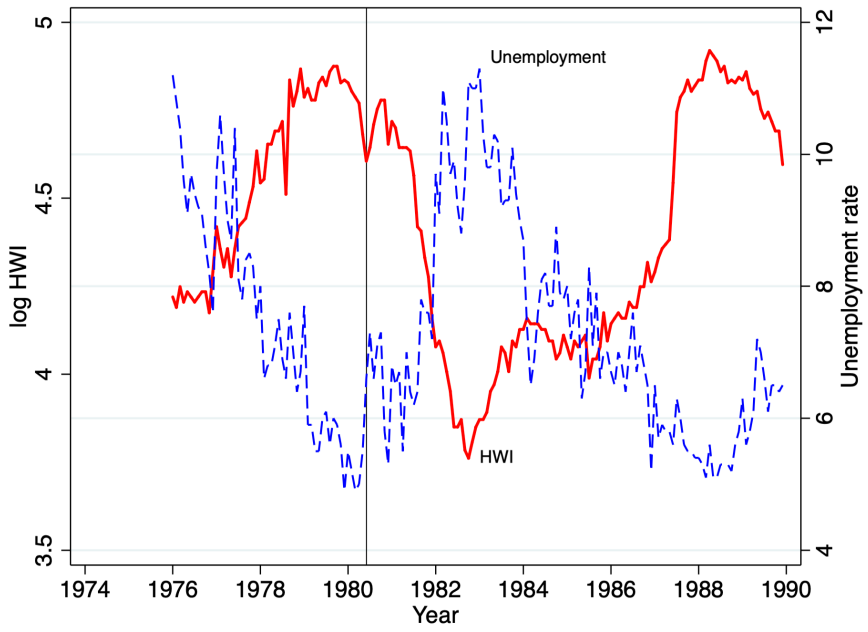
B. The job-finding rate



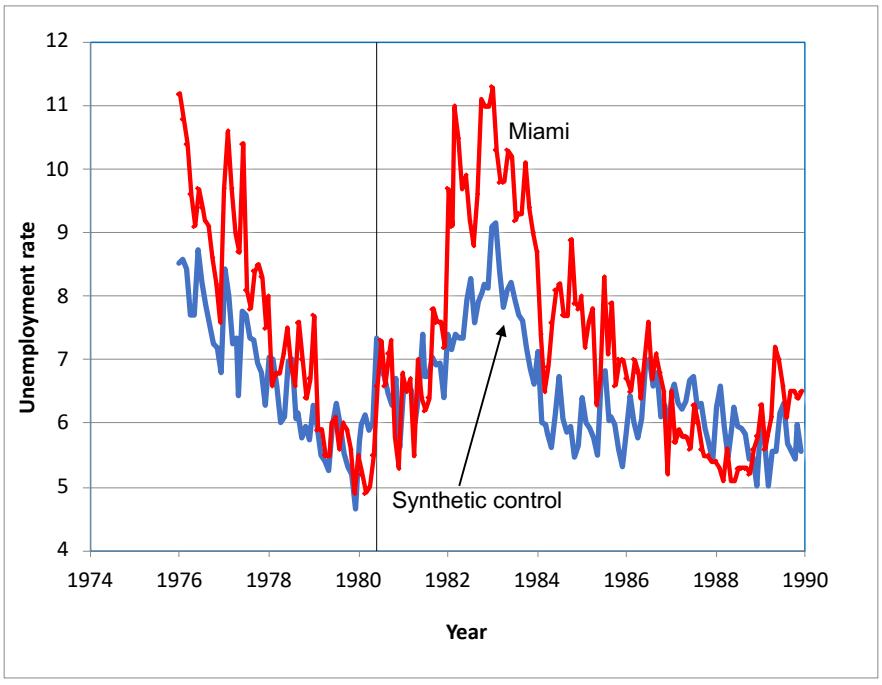
Notes: The HWI for each city is rescaled to equal 1 in 1977-1979. The job-finding rate is defined as the ratio of the HWI to the local unemployment rate and is also rescaled to equal 1 in 1977-1979. The treatment line is drawn as of June 1980.

Figure 4. Unemployment and job vacancies, 1975-1989

A. The HWI and the unemployment rate in Miami



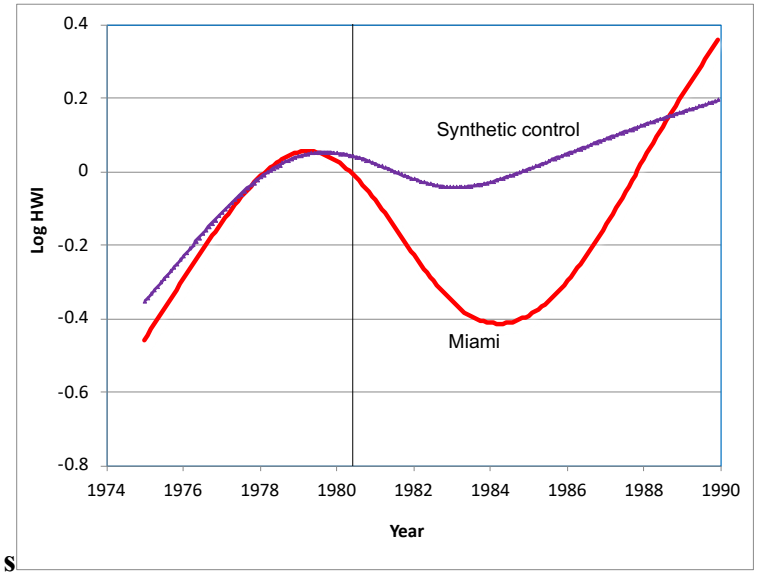
B. The unemployment rate in Miami relative to the synthetic control



Notes: The treatment line is drawn as of June 1980.

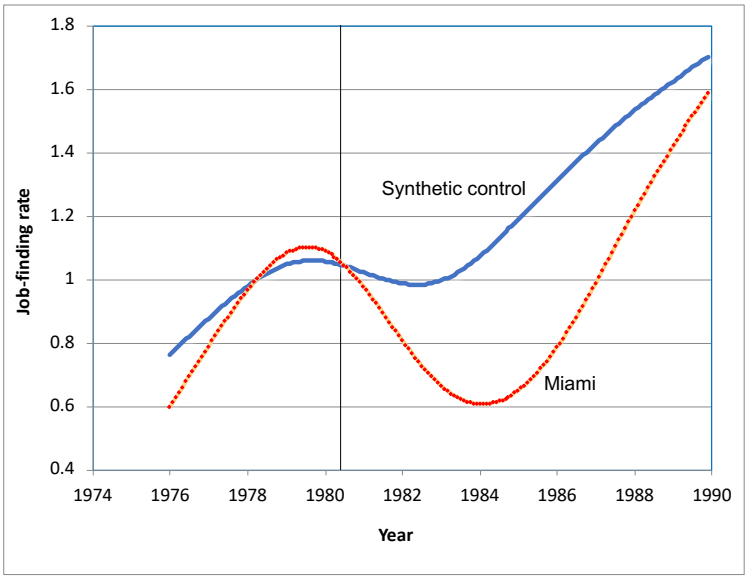
Figure 5. Removing local transitory fluctuations from the HWI using the HP filter

A. The help-wanted index



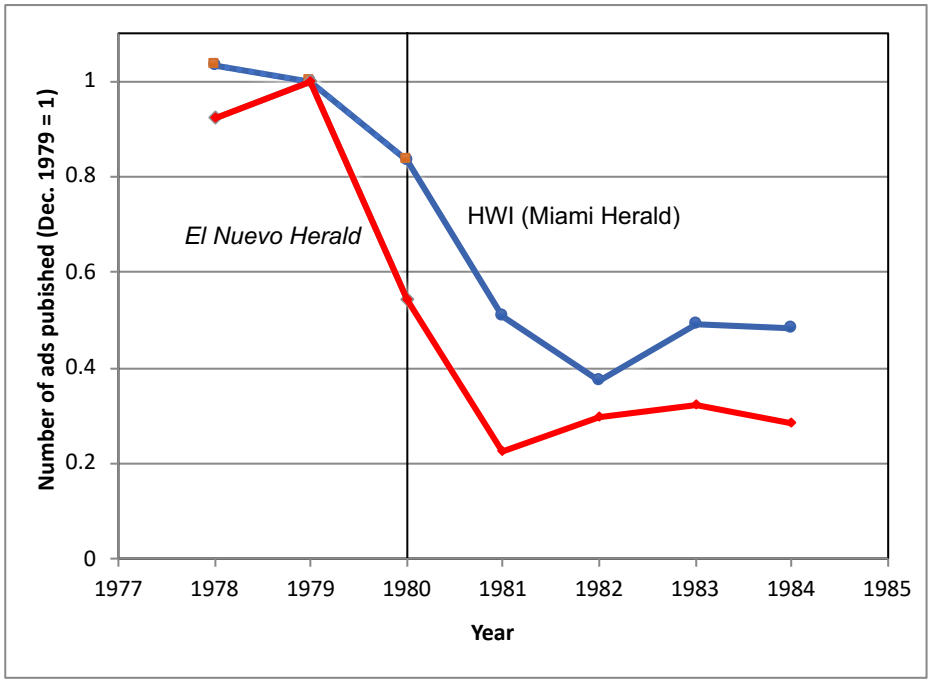
S

A. The job-finding rate



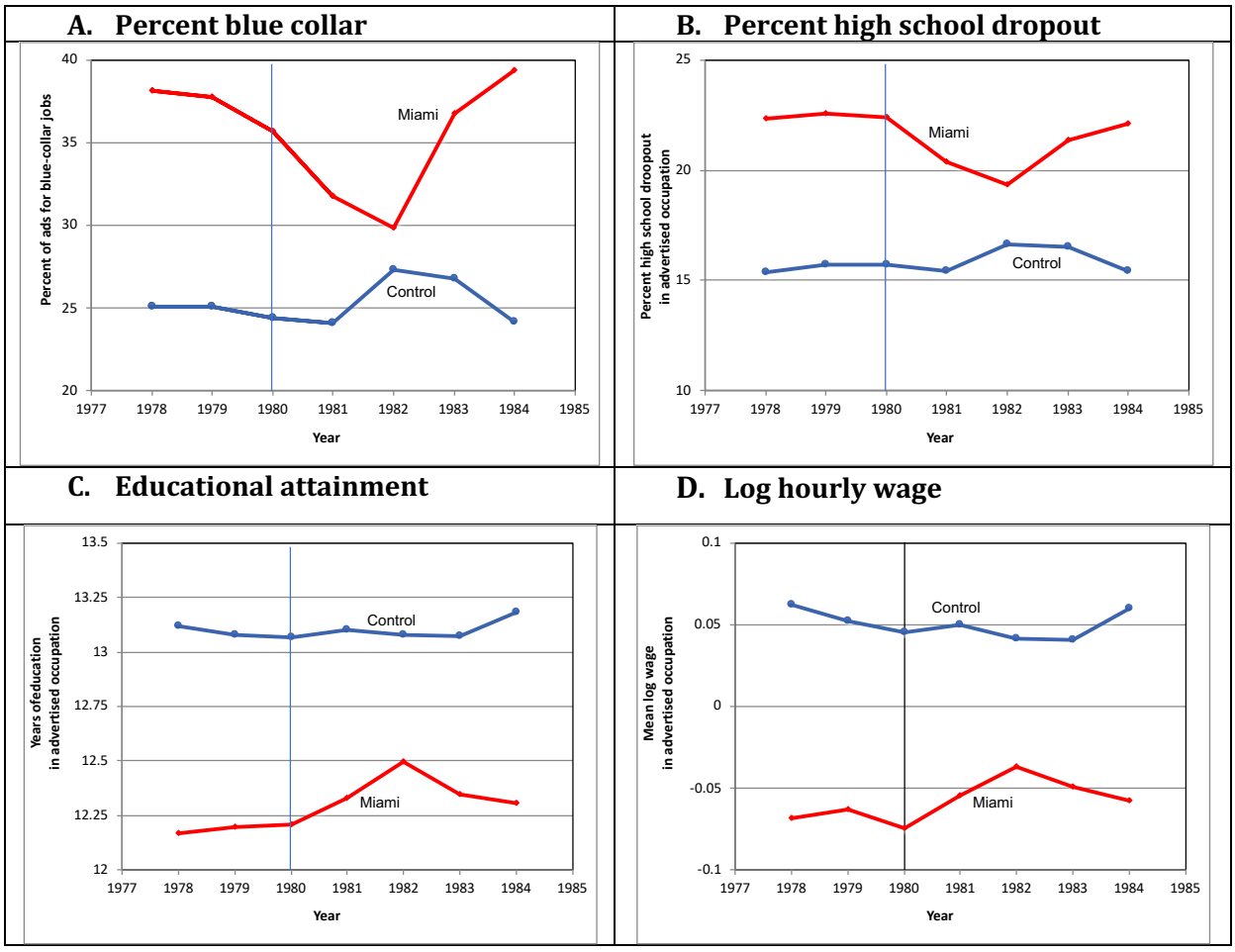
Notes: The adjusted HWI and job-finding rates give the long-term trend predicted by the HP filter with a smoothing parameter of 129,600, using all monthly observations between 1975 and 1989. The HWI and the job-finding rate for each city are each rescaled to equal 1 in 1977-1979.

Figure 6. Number of help-wanted ads published in the *Miami Herald* and *El Nuevo Herald*, 1978-1984



Notes: The number of ads published in the *Miami Herald* is given by the HWI in Miami in December of each year. We manually counted the number of ads published in *El Nuevo Herald* in each December between 1978 and 1984. Each of the trends is rescaled to equal one as of December 1979.

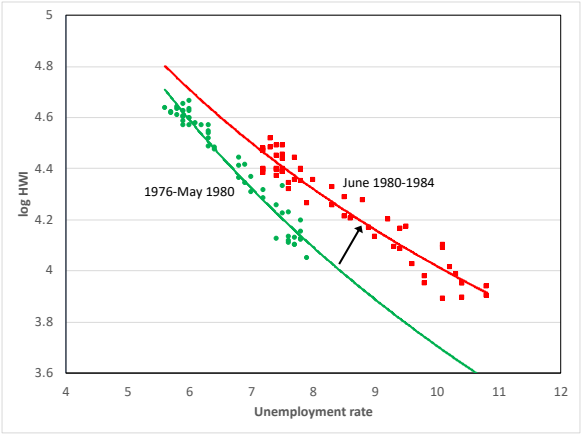
Figure 7. Trends in the skill composition of advertised vacancies, 1978-1984



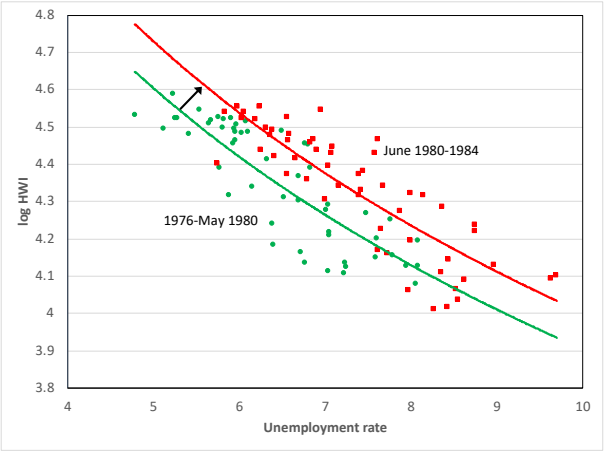
Notes: Panel A gives the fraction of ads advertising for a blue-collar position; panel B gives the fraction of workers who are high school dropouts in the average advertised occupation; panel C gives the average years of education in advertised occupations; and panel D gives the average (adjusted) mean log hourly wage in advertised occupations. The size of the sample in the *Miami Herald* is 95,263 ads; the size of the sample in the control newspapers (the *Boston Globe*, the *Minneapolis Star-Tribune*, the *New York Times*, and the *St. Louis Post-Dispatch*) is 2,674,889 ads.

Figure 8. The Beveridge Curve and Mariel, 1976-1984

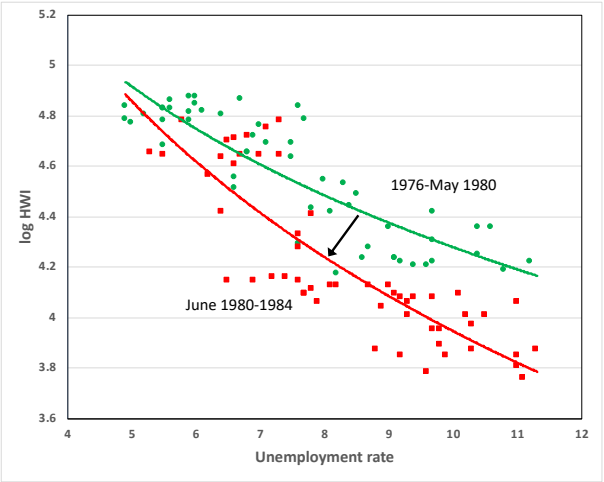
A. The national labor market



B. The synthetic control

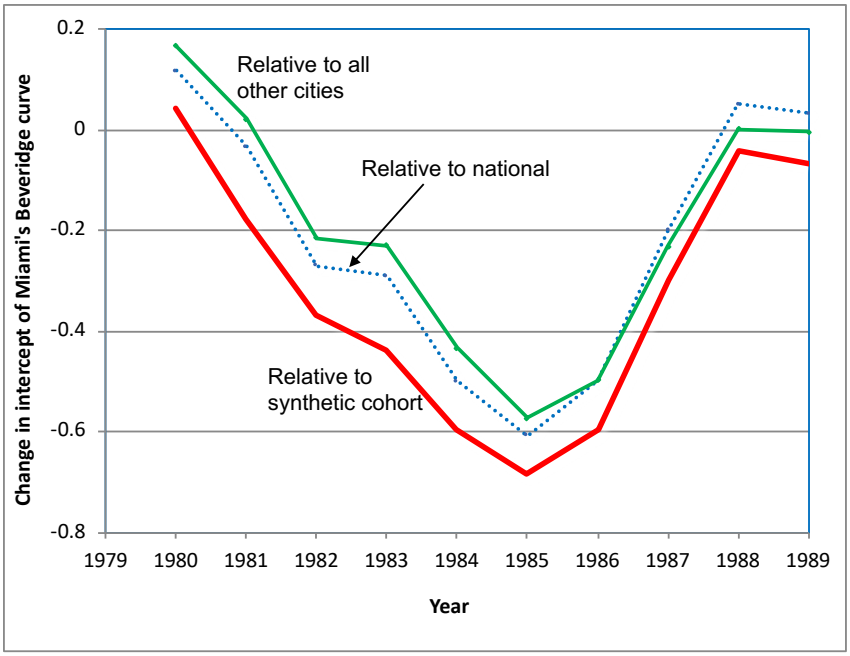


C. Miami



Notes: The figures show the data scatter and the logarithmic trend lines relating the raw HWI and the unemployment rate in a city-year-month cell in the pre-Mariel (January 1976-May 1980) and post-Mariel (June 1980-December 1984) periods. The Beveridge curve for the synthetic control in Panel B is constructed by taking a weighted average of the vacancy and the unemployment rate across metropolitan areas for each year-month cell, weighted by the synthetic control weights obtained in the analysis of the permanent trend in the job-finding rate.

Figure 9. Year-by-year shifts in the relative intercept of Miami's Beveridge curve (Baseline is January 1976-May 1980)



Notes: The figure illustrates the value of the interaction coefficient in a Beveridge curve regression model between a Miami indicator variable and a post-Mariel year (with the 1980 year being denoted by the June 1980-December 1980 period). The dependent variable is the log of the HWI index in an area. The regressions include the area's unemployment rate, area fixed effects, and year-month fixed effects.

Table 1. Difference-in-differences impact of Mariel supply shock on job vacancies

		Synthetic control	
Dependent variable:	All cities	Unadjusted	HP filter
A. Log HWI			
June 1980-1982	0.029	-0.237	-0.172
	(0.037)	(0.036)	(0.014)
1983-1984	-0.296	-0.436	-0.342
	(0.028)	(0.023)	(0.007)
1985-1986	-0.493	-0.448	-0.308
	(0.019)	(0.029)	(0.011)
1987-1989	-0.041	0.138	-0.000
	(0.028)	(0.029)	(0.021)
B. Job-finding rate			
June 1980-1982	0.044	-0.190	-0.151
	(0.048)	(0.043)	(0.019)
1983-1984	-0.289	-0.479	-0.427
	(0.045)	(0.043)	(0.016)
1985-1986	-0.491	-0.462	-0.491
	(0.035)	(0.039)	(0.010)
1987-1989	-0.194	0.121	-0.243
	(0.047)	(0.046)	(0.018)

Notes: Robust standard errors are reported in parentheses. The data consist of monthly observations for each city between 1975 and 1989. All regressions include vectors of city and year-month fixed effects. The table reports the interaction coefficients between a dummy variable indicating if the metropolitan area is Miami and the timing of the post-Mariel period (the baseline period goes from January 1975 through May 1980). The job-finding rate is the HWI divided by the unemployment rate in the city-year-month cell. The HWI (job-finding rate) regressions that use the all-city sample have 9,180 (8,064) observations, and the respective regressions that use the synthetic control have 360 (336) observations.

Table 2. Difference-in-differences impact of Mariel supply shock on skill composition of help-wanted ads, 1978-1984

Regressions	Percent blue collar	Percent dropout	Years of education	Mean log wage
A. Unweighted				
June 1980-1982	-6.705	-2.787	0.231	0.029
	(1.994)	(0.952)	(0.068)	(0.011)
1983-1984	0.507	-0.761	0.069	0.013
	(2.707)	(1.030)	(0.082)	(0.012)
B. Weighted				
June 1980-1982	-6.583	-3.010	0.251	0.032
	(1.895)	(0.817)	(0.070)	(0.011)
1983-1984	-2.971	-2.059	0.141	0.021
	(2.693)	(1.054)	(0.090)	(0.012)

Notes: Robust standard errors are reported in parentheses. The data consist of monthly observations for each city (i.e., newspaper) between 1978 and 1984. All regressions include vectors of city and year-month fixed effects. The table reports the interaction coefficients between a dummy variable indicating if the metropolitan area is Miami and the timing of the post-Mariel period (the baseline period goes from January 1978 through May 1980). The regressions have 403 observations.

Table 3. Supply shocks and the HWI, 1960-2000
(Dependent variable = Decadal change in city's log HWI)

Measure of supply shock:	OLS	IV
1. $M(\tau, \tau\text{-}1)/N(\tau\text{-}1)$	-1.027	-1.809
	(0.399)	(0.606)
2. $M(\tau, \tau\text{-}1)/N(\tau)$	-1.291	-1.707
	(0.402)	(0.553)
3. $M(\tau)/N(\tau) - M(\tau\text{-}1)/N(\tau\text{-}1)$	-0.907	-2.108
	(0.464)	(0.747)

Notes: Robust standard errors are reported in parentheses. The variable $M(\tau)$ and $N(\tau)$ give the number of immigrants and natives (in the relevant city-year cell) enumerated in census year τ, and $M(\tau\text{-}1, \tau)$ gives the number of immigrants who arrived between the two census years. All regressions have 198 observations. The instrument is the predicted size of the immigrant flow settling in a particular city based on the geographic settlement of earlier waves of immigrants belonging to the same national origin group (as constructed by Jaeger, Ruist, and Stuhler, 2018). All regressions are weighted by the size of the city's adult population at the time of the census.

Table 4. Relative intercept of Miami's Beveridge curve, 1980-1989
(Baseline = Jan. 1976-May 1980)

Year	National	All cities	Synthetic cohort
June-December 1980	0.117	0.167	0.042
	(0.029)	(0.035)	(0.027)
1981	-0.033	0.021	-0.178
	(0.039)	(0.038)	(0.041)
1982	-0.270	-0.214	-0.369
	(0.017)	(0.025)	(0.023)
1983	-0.290	-0.230	-0.438
	(0.017)	(0.020)	(0.015)
1984	-0.501	-0.433	-0.595
	(0.021)	(0.023)	(0.016)
1985	-0.608	-0.574	-0.685
	(0.016)	(0.019)	(0.017)
1986	-0.499	-0.498	-0.597
	(0.017)	(0.020)	(0.024)
1987	-0.200	-0.232	-0.300
	(0.058)	(0.058)	(0.063)
1988	0.050	0.001	-0.042
	(0.015)	(0.017)	(0.018)
1989	0.032	-0.004	-0.069
	(0.019)	(0.021)	(0.022)

Notes: Robust standard errors are reported in parentheses. The data consist of monthly observations for each city between 1976 and 1989. The dependent variable is the log HWI for the city-year-month cell. The regressors include the cell's unemployment rate, a vector of city fixed effects, and a vector of year-month fixed effects. The "national" regression compares Miami to the national HWI and unemployment data; the "synthetic cohort" regression compares Miami to the synthetic cohort implied by the analysis of the permanent trend in the job-finding rate; the "all cities" regression pools the city-year-month cells across all 48 cities in the sample. The table reports the interaction between a dummy variable indicating if the metropolitan area is Miami and the year in the post-Mariel period. The regressions using the national or the synthetic cohort data have 336 observations; the regression using the all-city sample has 8,064 observations.

Appendix A. Newspapers sampled by the Conference Board

City	Paper Used for HWI Since at Least 1970
Albany	*The Times Union*
Allentown	*Allentown Morning Call*
Atlanta	*Atlanta Constitution* (became *Atlanta Journal Constitution* in 1982)
Baltimore	*Baltimore Sun*
Birmingham	*Birmingham News*
Boston	*Boston Globe*
Charlotte	*Charlotte Observer*
Chicago	*Chicago Tribune*
Cincinnati	*Cincinnati Enquirer*
Cleveland	*Cleveland Plain Dealer*
Columbus	*Columbus Dispatch*
Dallas	*Dallas Times Herald* until 1991, then *Dallas Morning News*
Dayton	*Dayton Daily News*
Denver	*Denver Rocky Mountain News*
Detroit	*The Detroit News*
Gary	*Gary Post-Tribune*
Hartford	*Hartford Courant*
Houston	*Houston Chronicle*
Indianapolis	*Indianapolis Star*
Jacksonville	*Florida Times-Union*
Kansas City	*Kansas City Star*
Knoxville	*Knoxville News-Sentinel*
Los Angeles	*Los Angeles Times*
Louisville	*Louisville Courier-Journal*
Memphis	*Memphis Commercial Appeal*
Miami	*Miami Herald*
Milwaukee	*Milwaukee Sentinel*
Minneapolis	*Minneapolis Star Tribune*
Nashville	*Nashville Tennessean*
New Orleans	*The Times-Picayune*
New York	*New York Times*
Newark	*Newark Evening News*
Oklahoma City	*The Daily Oklahoman**
Omaha	*Omaha World-Herald*
Philadelphia	*Philadelphia Inquirer*
Phoenix	*Phoenix Arizona Republic*
Pittsburgh	*Pittsburgh Post-Gazette*
Providence	*Providence Journal*
Richmond	*Richmond Times-Dispatch*
Rochester	*Rochester Times-Union*
Sacramento	*Sacramento Bee*
Salt Lake City	*Salt Lake Tribune*
San Antonio	*San Antonio Express-News*
San Bernardino	*San Bernardino Sun*
San Diego	*San Diego Union*
San Francisco	*San Francisco Examiner*
Seattle	*Seattle Post-Intelligencer*
St. Louis	*St. Louis Post-Dispatch*
Syracuse	*Syracuse Herald Journal*
Toledo	*Toledo Blade*
Tulsa	*Tulsa World*
Washington D.C.	*Washington Post*

*We have been unable to confirm that the surveyed paper in Oklahoma City was the *Daily Oklahoman*.

Appendix B. Construction of Samples of Help-Wanted Classifieds

A. The *Miami Herald*

As noted in the text, we selected a random sample of issues published between 1978 and 1984 (essentially one issue from every month). For each issue, we (manually) produced digital images of the help-wanted pages primarily from the microfilm archive maintained at the Library of Congress. In the rare cases where the help-wanted pages from the Library of Congress were either not available or illegible for the selected dates, we used the Princeton University Library's *Miami Herald* archive. For all but one day in our original sampling plan, a human coder judged this procedure sufficient to obtain the entire day's help-wanted advertisements.[40] These digital images were then parsed into many smaller files and the task of transcribing each of these files was crowdsourced to Amazon Mechanical Turk workers. For instance, if a PDF page of help-wanted ads contained 9 columns of ads, this single page might be broken down into a total of 3-page chunks each containing 3 columns of ads, as in Figure B1 below.

Each of these page chunks was stored on *Google Drive* with the filename formatted in the following manner: "YYYY-MM-DD-##.pdf". Using the link sharing capabilities of *Google Drive*, a single link was generated for each of the files to enable us to share the file with Amazon Mechanical Turk workers as needed. After constructing links for every page chunk that we created, a Human Intelligence Task or HIT, was created on Amazon's Mechanical Turk for each page chunk. As part of the HIT, Mechanical Turk workers were asked to transcribe each ad in a format that would allow us to parse the ad using natural language processing (NLP) tools available in the *Python* language. Figure B2 provides an example of an HIT posted to Mechanical Turk requesting the transcription of a single page chunk of help-wanted ads.

After accepting the HIT and viewing the PDF file, each Mechanical Turk worker was instructed to create a file in the ".txt" format which had the identical filename as the PDF. For instance, if the PDF they happened to open was "1984-2-10-01.pdf", then the worker would create a file with the name "1984-2-10-01.txt" which would contain the transcribed ads in the format we requested. When transcribing the ads, workers were asked to separate each ad by a new line and to not include numbers, capitalization, line breaks, addresses, dates or days of the week contained in the ads themselves. These additional instructions allowed us to more easily extract and format the ad text using *Python* NLP tools. Finally, after transcribing the help-wanted page chunk according to our instructions, the Mechanical Turk worker was instructed to upload the transcriptions to a *Dropbox* folder that we had created for this purpose. This resulted in a total of 651 text files containing transcriptions from each of the page chunks.

After obtaining these transcribed files, we passed them through a *Python* script that used the filename and the textual information contained within each of the files to add each help-

---

[40] For the one date in which neither the Library of Congress nor Princeton University Library had a legible copy of the *Miami Herald*, we downloaded the prior day's *Miami Herald* from the Princeton University Library. For the dates in our sample in which the paper was available at both the Princeton University Library and the Library of Congress, the *Miami Herald* was missing more often in the Princeton collection. Additionally, the Princeton archive of the *Miami Herald* can only be read and downloaded one page at a time, making it relatively time-consuming to sample the help-wanted ads, which typically spread over many pages and tend to appear near the back of newspaper (and which must be manually clicked through to access). Moreover, the Princeton archive does not have copies of the *Miami Herald* after 1982. Hence, we defaulted to the Library of Congress archive whenever possible. We used the original print run edition of the *Miami Herald* if it was available; otherwise we used the "Final" edition. This procedure allowed us to obtain complete help-wanted sections for all the dates specified in our original sampling plan.

wanted ad to a single row on a CSV sheet which contained information about the month, day, year, text, and original filename of the transcribed ad. This database was subsequently processed further using a combination of keyword searches and manual auditing to remove duplicates and ads that were not help wanted classifieds. The resulting database from this process contains the 95,263 ads that were used in our analyses.

B. Control Newspapers

As noted in the text, we obtained a file from Atalay et al (2020, henceforth APST) that reported the number of ads in each newspaper-occupation-year-month-day cell for the *New York Times* and the *Boston Globe*. The text of help-wanted ads for the other two newspapers in the control group, *The Minneapolis Star-Tribune* (*MST*) and the *St. Louis Post-Dispatch* (*SLPD*), was acquired in a manner that was nearly identical to the method used by APST and utilized much of the *Python* code they made publicly available on their Github page: https://occupationdata.github.io/.

The digital versions of the *SLPD* and *MST* were acquired from *ProQuest* and were delivered to us in Extensible Markup Language (XML) format. The original data consisted of hundreds of thousands of individual XML files, each of which contained a full newspaper page.

The first step for processing these data involved isolating *ProQuest* newspaper pages from between 1978-1984, the period covered by our analyses. This was conducted using a script in the C# language and this process yielded 302,775 total *MST* newspaper pages and 272,817 *SLPD* newspaper pages, all in XML format. As part of the extraction process, the C# code labeled each of these newspaper pages using date and page number metadata from within each *ProQuest* XML file in the following format: "YYYY-MM-DD-###.xml" which represents year, month, day and page number, respectively. This was done to enable us to easily identify and keep track of each newspaper page in this large database. For instance, the file "1978-01-01-001.xml" within a folder containing the *MST* pages is the first (front) page of the *Minneapolis Star-Tribune* for the edition published on January 1, 1978.

The next stage involved identifying and isolating the newspaper pages that contained help-wanted ads. This stage required four additional steps: (1) removal of XML markup for each newspaper page; (2) conversion of the XML page to plain text; (3) unsupervised learning/topic model estimation on cleaned plain text files and; (4) posterior estimation from the topic model to identify pages with help-wanted ads. Removal of the XML markup for each newspaper page and conversion to plain text was accomplished using modified *Python* code from APST (2019) who had written a script to perform a similar task in their work. After conversion of the pages to plain text, each of the newspaper pages was cleaned through stemming, tokenization into unigrams and bigrams, and the removal of stop words using the *quanteda* package in R (Benoit et al, 2018). The resulting text was then converted into a high-dimensional document-feature matrix that was used to estimate a series of topic models on the pages of each newspaper. Topic models are an unsupervised machine learning technique that uses the words in documents to identify clusters of similar documents (Blei, Ng, and Jordan, 2003).

Topics contained within a corpus are identified using the top 10-20 vocabulary words that make up the topic. Labeling of documents with topics is conducted through posterior estimation, where topics are assigned to each document according to the highest probability topic for that document. After an examination of the output of the top terms from each of the topic models, we found that a 10-topic model was able to most clearly distinguish between pages with help wanted ads for each of the newspapers.

Table B1 provides the top 10 terms from the first 6 topics estimated on *MST* newspaper pages in 1978. In this case, Topic 6 is clearly identifying pages with help-wanted ads. Because of the large scale of the topic modeling analysis required, a ten-topic model was estimated for each newspaper-year, so that a total of 14 ten-topic models were estimated, one for each newspaper-year combination.

Posterior estimation identified the pages in each newspaper-year which corresponded to the highest posterior probability for the help-wanted ad topics. Those pages were then extracted from the larger database. This process yielded a total of 12,078 pages from *MST* and 11,079 pages from *SLPD* identified as containing help-wanted ads. Finally, help-wanted ads from each of these pages were extracted and added to a CSV sheet using code provided by APST. This yielded a total of 253,890 *MST* and 148,535 *SLPD* ads which included ad titles, ad text, along with the day month and year of the ad. As illustrated in Panel B of Figure D3, the trend in the number of annual ads sampled in each of the control newspapers tracks the HWI for the respective city. The tracking is quite good for some newspapers (e.g., the *Boston Globe*), and less precise for others (e.g., the *New York Times* or the *St. Louis Post-Dispatch*). Each of the help-wanted ads in the *MST* and *STPD* samples was then fed to the O*NET-SOC AutoCoder to obtain the occupation code of the advertised job.

Figure B1. Example outline of the process of transcribing a single *Miami Herald* page of help-wanted ads

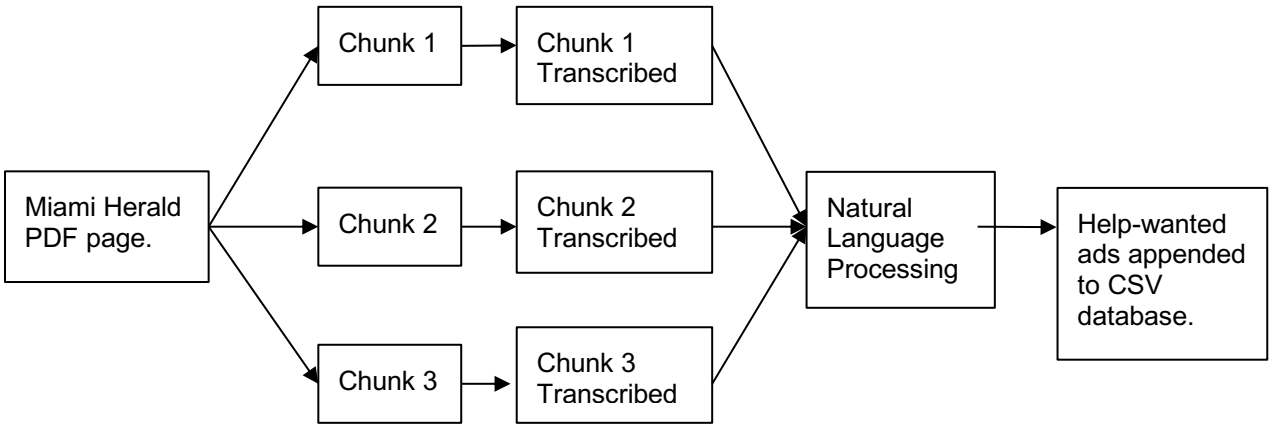
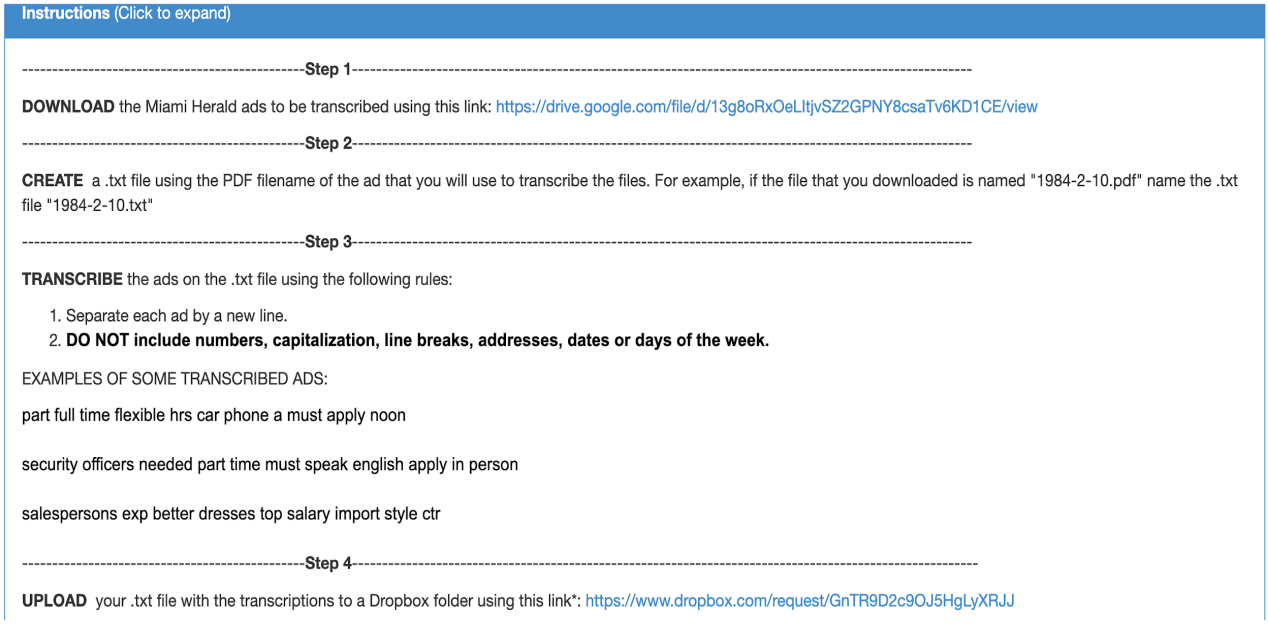


Figure B2. A sample Human Intelligence Task (HIT) posted to Amazon's Mechanical Turk requesting the transcription of *Miami Herald* ads



Instructions (Click to expand)

----------**Step 1**----------

DOWNLOAD the Miami Herald ads to be transcribed using this link: https://drive.google.com/file/d/13g8oRxOeLltjvSZ2GPNY8csaTv6KD1CE/view

----------**Step 2**----------

CREATE a .txt file using the PDF filename of the ad that you will use to transcribe the files. For example, if the file that you downloaded is named "1984-2-10.pdf" name the .txt file "1984-2-10.txt"

----------**Step 3**----------

TRANSCRIBE the ads on the .txt file using the following rules:

1. Separate each ad by a new line.
2. **DO NOT include numbers, capitalization, line breaks, addresses, dates or days of the week.**

EXAMPLES OF SOME TRANSCRIBED ADS:

part full time flexible hrs car phone a must apply noon

security officers needed part time must speak english apply in person

salespersons exp better dresses top salary import style ctr

----------**Step 4**----------

UPLOAD your .txt file with the transcriptions to a Dropbox folder using this link*: https://www.dropbox.com/request/GnTR9D2c9OJ5HgLyXRJJ

Table B1. First 6 topics from a 10-topic model estimated using *Minneapolis Star-Tribune* pages, 1978

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6
home	reg	m	br	of_the	opportun
open	sale	s	s	said	experi
room	t	ii	avail	year	person
br	m	t	ft	in_the	call
new	j	v	apt	state	work
rm	price	n	call	minneapoli	employ
realti	new	j	av	percent	time
bath	w	l	new	minnesota	benefit
lot	r	r	park	new	posit
area	f	i_i	home	to_the	open

Note: Topic 6 is clearly the one that identifies pages containing help-wanted ads.