Job Vacancies, the Beveridge Curve, and Supply Shocks:
The Frequency and Content of Help-Wanted Ads in Pre- and Post-Mariel Miami

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

August 2019

Job Vacancies, the Beveridge Curve, and Supply Shocks:
The Frequency and Content of Help-Wanted Ads in Pre- and Post-Mariel Miami

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 the 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 marked drop in Miami's HWI relative to many alternative control groups in the first 4 or 5 years after Mariel, followed by recovery afterwards. The Miami evidence is consistent with the observed relation between immigration and the HWI across all metropolitan areas in the 1970-2000 period: these spatial correlations suggest that more immigration reduces the number of job vacancies. We also explore some of the macro implications of the Mariel supply shock and show that Miami's Beveridge curve shifted inwards by the mid-1980s, suggesting a more efficient labor market, in contrast to the outward nationwide shift coincident with the onset of the 1980-1982 recession. Finally, we examine the text of the help-wanted ads published in a number of newspapers and document a statistically and economically significant post-Mariel decline in the relative number of low-skill vacancies advertised in the *Miami Herald*.

Job Vacancies, the Beveridge Curve, and Supply Shocks:
The Frequency and Content of Help-Wanted Ads in Pre- and Post-Mariel Miami

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 political 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 available to 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 (Blau and Mackie, 2016). In the U.S. context, this literature typically uses microdata, such as the decennial censuses or the Current Population Surveys (CPS), to document how wages change in those markets targeted by immigrants. Sometimes the markets are defined by geographic boundaries; sometimes the markets are defined by skill group (Borjas, 2003). But the basic strategy is the same. Immigrants tend to target some markets more than others. We then measure the impact of immigration by contrasting the evolution of wages in the markets hit by immigration with the evolution in the markets that immigrants shunned.

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; 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. 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 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 prime-age, non-Hispanic man who is a high school dropout), the evidence suggested that the wage fell by at least 15 percent. And Peri and Yasenov (2019) reported that the wage remained unchanged if one used an alternative definition for 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).

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 and important 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 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.

We use the HWI to document how immigration affects the number of job vacancies in local labor markets.[1] Our analysis begins by revisiting the Mariel episode. The data clearly show 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. We also look beyond Miami and estimate the generic spatial correlations that dominate the literature, correlating changes in the HWI with immigration across the 51 metropolitan areas for which the index is available. These spatial correlations also indicate that immigration is typically associated with reduced employer effort to find workers by placing help-wanted ads in the local newspaper.

The job vacancy data lets us explore some of the macroeconomic implications of the Mariel supply shock, which coincided with the onset of the 1980-1982 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 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 had lower reservation wages, quickly filling up vacancies). 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.

We conclude our study by examining how the Mariel supply shock changed the *textual content* of help-wanted ads. We examine the text of a large sample of help-wanted ads published in the *Miami Herald* and in other newspapers between 1978 and 1984. This type of analysis helps identify the types of jobs 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 labor market. The fraction of vacancies for blue collar positions fell precipitously, and surviving vacancies were for jobs that tended to hire much more educated workers. In short, low-skill labor markets do seem to respond to low-skill supply shocks.

---

[1] As far as we know, our study is the first to investigate the quantitative relationship between immigration and job vacancies in the U.S. labor market. Withers and Pope (1985) find no statistically significant relationship between immigration and the Australian job-finding rate (see also Warren, 1982), 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.

II. The Conference Board Help-Wanted Index

Unlike many OECD countries, the United States did not maintain any continuous official statistics on job openings until 2000, when the Bureau of Labor Statistics (BLS) introduced the Job Openings and Labor Turnover Survey (JOLTS). Prior to 2000, researchers interested in understanding the trends and determinants of job openings or vacancies relied primarily on the Conference Board Help-Wanted Index (HWI).

Beginning in 1951, the Conference Board contacted 51 newspapers (see Appendix A), each corresponding to a 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 metropolitan area indices were then aggregated to create an index for each geographic region, and for the nation as a whole.[2]

Figure 1 illustrates the very strong correlation between the national HWI and the unemployment rate, showing an inverse relation between the two variables throughout much of the period.[3] However, 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.[4] 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 internal data collection in October 2010.

Despite the obvious link between the HWI and the unemployment rate, there are several biases in the index that 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

---

[2] 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. Zagorsky (1998) combined previous surveys of help-wanted classifieds by the Metropolitan Life Insurance Company with the HWI to create a help-wanted index that dates back to 1923.

[3] 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."

[4] Kroft and Pope (2014) show that the growth of local online ads in Craigslist caused a reduction in a city's HWI.

Board to construct the HWI counts this as only one advertised job (Preston, 1977). Figure 2*a* illustrates the bias through an ad published in the *Miami Herald* on March 2, 1975. This single posting advertised "several openings" for test technicians.

A related problem arises with ads placed by private employment agencies, which often contained several job postings (as in Figure 2*b*). 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, when reporting private employment agency ads to the Conference Board, some newspapers counted a posting by a private employment agency as a single help-wanted ad, but other newspapers did not (Abraham, 1987).[5]

Further, the HWI only counts ads placed in the official help-wanted section of the newspaper. As Figure 2*c* shows, again drawn from the March 2, 1975 issue of the *Miami Herald*, many 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, juxtaposed with ads extolling "excellent land opportunities" or a "3,895 acre operating ranch," and those sections were not included in the computation of the index.

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 job classifieds (Courtney, 1991).

These methodological issues imply that intercity differences in the *level* of the HWI do not provide a good metric for making comparisons in local labor market conditions.[6] The Conference Board index depended partly on how many job openings were posted in a specific ad

---

[5] Cohen and Solow (1967, p. 108) noted these methodological issues, writing that "we know nothing, for example, about the number of jobs offered per advertisement." They also make the point, which we return to below, 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.

[6] As an example of the cross-section variation in the HWI, the 5th percentile value of the index in 1980 is 40.2 and the 95th percentile value was more than four times larger, or 166.9.

and how that ad was counted, on where the ad was physically placed in the newspaper, and on the market power of the sampled newspaper.

We address this problem by rescaling 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]

The HWI data has been used by many widely-cited studies. 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. For example, she compares the HWI with administrative data from Minnesota (one of only two states that collected information on the number of job openings at the time) and concludes: "the Minnesota data… suggest that the normalized help-wanted index is a reasonably good vacancy proxy" (Abraham, 1987, p. 213).

Although the explosion of online job postings reduced the relevance of newspaper-based indices, the HWI remains useful for historical research. 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

Figure 3 shows the number of Cuban immigrants arriving in the United States each year between 1955 and 2010. Most of them settled in the Miami metropolitan area.[8] The Miami labor market was hit by several sudden and short-lived supply shocks during this period. The first was in the early 1960s, beginning soon after the Castro takeover and ending abruptly in 1962 with the

---

[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.

[8] The 1970 census indicates that 49.2 percent of the Cuban immigrants from the 1960-1964 cohort lived in Miami; the 1990 census indicates that 63.5 percent of the 1980-1981 cohort lived in Miami; and the 2000 Census indicates that 63.7 percent of the 1994-1995 cohort lived in Miami.

Cuban missile crisis. The second is the very noticeable spike in 1980 associated with the Mariel boatlift. And the third was in 1994-1995 when a new boatlift of Cuban refugees was rerouted to Guantanamo Bay by the Clinton administration.[9] This detour did not last long, however, and most of the refugees ended up reaching the United States in 1995. The figure also shows that a sizable number of Cubans entered between 1966 and 1970, as part of the "Freedom Flights" that were designed to reunite families separated by the halt in 1962, but this migration was less sudden, not entirely unexpected, and did not end abruptly in a year or two.

This section documents the response of job vacancies in the Miami labor market to the Mariel supply shock, which is, by far, the "cleanest" of the shocks illustrated in Figure 3. It was unexpected, large, and it was preceded and followed by a long period of little Cuban migration.[10]

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 closed the escape valve 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.

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 labor markets that served as a control group, indicated that nothing much happened to Miami despite the large number of refugees. The average wage did not fall, and the unemployment rate remained unchanged relative to what was happening in the cities that formed the control group.

---

[9] Writing before the 2000 census data became available and documented the delayed arrival of the immigrants, Angrist and Krueger (1999, p. 1328) branded this supply shock as "The Mariel Boatlift That Did Not Happen."

[10] Appendix B analyzes the relation between Miami's HWI and the other supply shocks. The short-run response in the Miami labor market, relative to comparable cities, was similar across the various shocks—a decline in the index of job vacancies within two or three years.

The debate over the labor market impact of the *Marielitos* intensified in the past few years as a result of the Borjas (2017) reappraisal, which used data from both the March CPS and the CPS Outgoing Rotation Group (ORG) to specifically examine the wage trends of workers who lacked a high school diploma. It turns out that the wage of this low skill group in Miami fell substantially relative to control cities. A flurry of subsequent papers argues that the evidence is sensitive to the definition of the "low-skill" workforce, to changes in the racial composition of the sample, and to the inevitable sampling error resulting from the small number of workers that the CPS surveyed in the Miami metropolitan area.[11] 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. Figure 4 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).[12]

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 similar to the national index (although it was still lower than the index for the South Atlantic region).[13]

---

[11] The flurry of papers includes Borjas (2019), Borjas and Monras (2017), Peri and Yasenov (2019), and Clemens and Hunt (2019).

[12] We also conducted much of the analysis using a longer pre-treatment period (going back to 1970), and the measured impact of the Mariel supply shock is very similar to what is reported below. The 1975-1989 period may be preferable because the Miami labor market in the early 1970s may still be recovering from the aftermath of the 1966-1970 Freedom Flights.

[13] The relative trend in Miami's HWI resembles the wage trend for high school dropouts reported in Borjas (2017, Figure 2). This similarity similarity between the trends in Miami's low-skill wage and the HWI suggests that the HWI, which presumably measures "average" labor market conditions, may be doing a particularly good job at capturing changing opportunities for low-skill workers. We return to this point below.

Of course, Miami's distinctive trend should be contrasted with what happened in "comparable" cities rather than regional or national aggregates.[14] To illustrate the uniqueness of the Miami experience, we use several alternative controls. We start with the set of metropolitan areas selected by Card (1990) as a control group for the Miami of the early 1980s. This original control group was composed of Atlanta, Los Angeles, Houston, and Tampa-St. Petersburg. Unfortunately, the HWI was never calculated for the Tampa-St. Petersburg metropolitan area, so our "Card control" only has the other three cities. Figure 5 shows the trend in the HWI index for Miami and for the cities in the Card control group.[15] It is evident that Miami's HWI declined dramatically after Mariel relative to what was observed in the Card control.

We next use the control group of the four metropolitan areas that had similar growth rates for low-skill employment in the pre-Mariel period (as reported in Borjas, 2017). These four cities were Gary, Houston, Indianapolis, and Los Angeles. Figure 5 also shows that the post-1980 Miami experience was unusual when compared to this "low-skill control."

Finally, we use the synthetic control method (Abadie, Diamond, and Hainmueller, 2010). The method essentially searches across all 50 potential control cities and derives a weight that combines cities to create a new synthetic city. This synthetic city is the one that best resembled the pre-Mariel Miami labor market along some set of pre-specified conditions.

We calculate the synthetic control by using a large number of control variables, all calculated from the 1970 and 1980 IPUMS census files. 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

---

[14] 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.

[15] The average index for the control group is a weighted average of the index across the cities, where the weight is the city's employment in 1980.

the fraction who are Hispanic.[16] Figure 5 also illustrates the trend in the HWI for the synthetic control, and it again shows the uniqueness of Miami's post-Mariel experience.

To calculate 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 cyclical fluctuations).

The cities r included in the regression are Miami and the cities in a specific control group. For example, if the Miami experience is being compared to that of cities in the low-skill control group, there would be five cities in the data, and each of these cities would be observed 180 times between 1975 and 1989, for a total of 900 observations. The regression comparing Miami to the synthetic control is similar in spirit, but there are only two "cities" in this regression: Miami and the synthetic city, for a total of 360 observations. Because the impact of Mariel might vary 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.

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. If we use the Card control group, the relative HWI in Miami declined by about 20 percent in 1981-1982 and by over 40 percent in 1983-1984. The regression in the second column uses the low-skill control group and also shows sizable declines (of about 20 percent) by the mid-1980s. The third column estimates a regression that pools the data across all 51 cities in our sample and it again shows sizable

---

[16] 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).

declines in the vacancy index by the mid-1980s. Finally, the fourth 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.[17] Regardless of the control group, all regressions indicate a full recovery in Miami's HWI by the late 1980s (in fact, Miami's HWI is sometimes higher by the end of the decade).

Some of the studies that examine trends in the HWI "normalize" the index by dividing it by the size of the workforce (Abraham and Katz, 1986). Define the normalized index for the city-year-month cell as $\hat{H}_{rtm} = H_{rtm}/E_{rtm}$, where E_{rtm} gives non-agricultural employment in city r, year t, month m.[18] Panel B of Table 1 reports the regressions that use the normalized index as the dependent variable. The data again document the very different experience of post-Mariel Miami relative to the control cities. If we use the synthetic control baseline, the normalized index declined by about 40 percent by the mid-1980s, and fully recovered by the late 1980s. Because of the similarity in results using the two alternative measures of the HWI, the remainder of the paper uses the simpler, raw measure of the HWI constructed by the Conference Board.

The obvious 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. One alternative explanation is that because of the large and sudden increase in the size of the Cuban community, Miami's employers began to shift their advertising to Spanish-language outlets. In fact, in 1977 the *Miami Herald* began to publish *El Nuevo Herald*, a sister newspaper in Spanish. The Conference Board did not enumerate help-wanted ads in *El Nuevo Herald*, perhaps mechanically leading to a decline in the HWI in post-Mariel Miami.

---

[17] The regression estimated using the synthetic control differs in a slight (and numerically trivial) way from the other regressions. As noted above, we rescaled the HWI index so that the index for each city equals 1 in the 1976-1978 period. This rescaling is irrelevant in the regressions that use the Card control group, the low-skill control group, or the entire set of 51 metropolitan areas, as the regressions include city fixed effects. The synthetic control regression aggregates the control cities with positive weights into a synthetic city and does not include individual fixed effects for each of those cities. The coefficients in the vector β are almost identical if we skip the step of rescaling the index to equal 1 in the pre-Mariel period.

[18] The BLS employment data at the city-year-month level is available in a file maintained by the ICPSR, "Employment, Employment, Hours, and Earnings in States and Areas of the United States, 1940-1991" (ICPSR 9928).

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.[19] 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. The difference suggests that the Mariel refugees were perhaps more easily substitutable with the pre-existing Cuban workers than with native workers.[20]

As noted earlier, the Mariel boatlift coincided with the onset of the 1980-1982 recession. The HWI has an obvious cyclical trend, and the regression analysis in equation (1) nets out cyclical shocks that affected all localities equally when estimating the impact of the Mariel supply shock. 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,

---

[19] 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.

[20] 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. This argument would not explain why the HWI recovered unless immigrant social networks break down quickly. Unfortunately, our data does not allow us to distinguish between these explanations. 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 the substitution from published job advertisements to unobservable social networks.

therefore, might remove some of the effect of supply shocks on the locality and produces a lower-bound estimate of the impact of a supply shock. If the long-run trend implied by the HP filter is shifted by supply shocks, there would be strong evidence for the hypothesis that such shocks have a sizable and more persistent effect on local labor market conditions.

Figure 7 shows the behavior of the adjusted index in Miami and the synthetic control.[21] The exercise obviously removes a lot of the cyclical variation in the HWI for 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.

IV. Robustness Checks

This section presents a number of extensions to help determine if our empirical finding of a negative association between supply shocks and job vacancies is robust.

A. Randomizing the Control Group

The inference that supply shocks reduce the number of job vacancies may hinge crucially on the counterfactual. One way of determining how much the choice of a control group matters is to bypass the decision of picking control variables to create a synthetic city control, and instead estimate the wage impact for *every* potential control group. We illustrate this approach by estimating the short-run effect of the Mariel supply shock relative to each of the 230,300 four-city control groups that can be constructed from the Conference Board data.

In particular, we use the data from the 1979-1984 period, set the treatment date as of June 1980, and estimate the generic difference-in-differences regression model in equation (1) for each of the 230,300 control groups. The relevant coefficient is the interaction of an indicator variable identifying the Miami metropolitan area and whether the observation refers to a post-

---

[21] The long-term trend for the synthetic control in Figure 7 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.

Mariel period (i.e., between June 1980 and December 1984). Figure 8 shows the density function for all the 230,300 estimated short-run effects (i.e., the distribution of the coefficient β in the regression model).

The mean impact is -0.232, with a standard deviation of 0.106. Only 1.7 percent of the coefficients were zero or positive. The figure also shows the location of the impact implied by the comparable synthetic control regression. Evidently, the synthetic control is "picking" a short-run impact that is large relative to the potential mean.

B. Beyond Miami

Up to this point, our analysis has focused on documenting the response of job vacancies in Miami to the Mariel supply shock. Immigration also affected many other cities in recent decades, 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. It is well known that immigrants cluster in a relatively small number of metropolitan areas, although no other city experienced the sudden and large supply shock that hit Miami in 1980.

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 settling in that city. Although the studies often 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 the potential adverse effect of supply shocks on the wage of competing workers. Moreover, the widely used shift-share instrument (which is essentially 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 collect immigration data until 1994. Moreover, the reliability of the HWI index deteriorated in the late 1990s with the explosion of on-line job

postings. The CPS data, therefore, would only let us analyze a very short time series, making it unlikely that a credible spatial correlation could be estimated.[22]

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 defining 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 spatial correlation.

Note that the same endogeneity problem that plagues estimates of the spatial correlation between wages and immigration 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 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 2 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 moved to the city between census year τ-1 and τ, or $M(\tau, \tau$-1),

---

[22] 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. Unfortunately, the data are not publicly available at the metropolitan area level. In fact, the JOLTS vacancy rate series are only publicly released for four broad census regions (the Northeast, the South, the Midwest, and the West), greatly reducing the possibility that the data can be used to estimate credible spatial correlations.

to the number of natives residing in the city at time τ-1, or $N(\tau\text{-}1)$.[23] This specification ignores the fact that there may have been a supply response as natives either 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 that is commonly employed in the literature. 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. We use the instrument for the total immigrant share for each city calculated by Jaeger, Ruist, and Stuhler (2018).[24] The IV estimate of the impact is also negative and significant, and larger in absolute value than the OLS estimate. The estimated elasticity $\left(\frac{d\log H}{d\log L}\right)$ is about -1.8.

The other rows of Table 2 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 $\left(\frac{M_{c,\tau}}{N_{c,\tau}} - \frac{M_{c,\tau-1}}{N_{c,\tau-1}}\right)$.[25] Regardless of how we define the supply shock, Table 2 documents a negative correlation between the change in the HWI and aggregate 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.

C. The HWI and Low-Skill Labor Markets

---

[23] The number of immigrants who migrated to the United States in the interval $(\tau-1, \tau)$ and settled in city r is obtained from the decennial census in year τ.

[24] We are very grateful to David Jaeger, Joakim Ruist, and Jan Stuhler for generously 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.

[25] 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.

The HWI presumably provides information about aggregate conditions in the local labor market. The Mariel supply shock was composed of relatively low-skill workers. The sensitivity of the HWI to this shock suggests that perhaps the index is doing a particularly good job at capturing labor market trends at the lower end of the skill distribution. As noted earlier, this tendency may be a mechanical consequence of where classified ads appeared in the newspapers sampled by the Conference Board, and how those ads were counted.

We use annual data from the CPS to document that changes in the HWI are, in fact, most strongly correlated with wage and employment trends for the least-educated workers. Consider the following regression model:

$$y_{rst} = \theta_r + \theta_s + \theta_t + \beta_0 \log H_{rt} + \beta_1 (\log H_{rt} \times \theta_s) + \varepsilon, \tag{3}$$

where y_{rst} is a labor market outcome for city r, education group s, and calendar year t; θ_r is a vector of city fixed effects; θ_s is a vector of education fixed effects; θ_t is a vector of calendar year fixed effects; and H_{rt} is the HWI index for city r in year t. By including the city fixed effects, the coefficient vector (β_0, β_1) is essentially estimating the correlation between a within-city change in the HWI and the corresponding change in labor market outcome y, and how that correlation varies across education groups.

We classify workers into four education groups: high school dropouts (who have less than 12 years of education), high school graduates (exactly 12 years), some college (13-15 years), and college graduates (at least 16 years). We examine two labor market outcomes: wages and employment. The average wage or employment propensity for cell (r, s, t) is calculated from residuals to individual-level regressions estimated in the CPS data. We use a simple first-stage model to calculate the age- and gender-adjusted outcome in cell (r, s, t). Specifically, we estimate the following individual-level regression separately in each CPS cross-section:

$$y_{irst} = \alpha_t + X_i \gamma_t + \varepsilon, \tag{4}$$

where y_{irst} is the outcome for worker i in city r with education s at time t; and X_i is a vector of fixed effects giving the person's age and gender.[26] We estimate the regressions using both the March CPS (for survey years 1973-2000), and the ORG (for years 1979-1999).[27] The dependent variable in the wage regressions is either the log weekly wage (in the March CPS) or the log hourly wage rate (in the ORG). The employment regressions use a linear probability model where the dependent variable equals one if the person worked at some point in the previous calendar year (in the March CPS) or during the reference week (in the ORG). The average residual from the regression for cell (r, s, t) gives the age- and gender-adjusted mean outcome in the cell.[28]

Table 3 reports the estimated vector β for alternative specifications of the regression model. Consider initially the regression coefficients in the first row of the top panel of the table, which measure the correlation between the HWI and wage trends in the March CPS. The interactions of the HWI with the education fixed effects indicate that the wage of less-educated workers is more strongly correlated with the HWI than the wage of workers with more education. In fact, the correlation declines monotonically with education. A 10 percent increase in the index is associated with a 1.3 percent increase in the wage of high-school dropouts, a 0.7 percent increase in the wage of high school graduates, a 0.4 percent increase in the wage of workers with some college, and no change in the wage of college graduates. The table reports a similar pattern in the wage regression coefficients using the ORG.

The second row of each panel shows the correlation between employment trends and the HWI. The March CPS shows that the HWI and employment propensities are most strongly correlated for the least-educated workers. A 10 percent increase in the index is associated with a 1.1 percentage point increase in the employment probability of high school dropouts (relative to college graduates), a 0.7 percentage point increase in the probability for high school graduates,

---

[26] The employment regressions use the sample of persons aged 18-64, while the wage regressions use the sample of workers aged 25-59 to minimize the sensitivity of the measured wage to school enrollment or early retirement. The age fixed effects indicate if the person is aged 24 or less, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, or 60-64.

[27] The 1973 March CPS is the first that identifies a relatively large number of metropolitan areas, including Miami.

[28] The calculation of the average residual wage in the cell weighs each individual observation by the product of the person's earnings weight times the number of weeks worked in the year (in the March CPS) or the usual number of hours worked weekly (in the ORG).

and a 0.5 percentage point increase in the probability for workers with some college. The ORG also shows a similar monotonic decline, although the drop is not as steep.

We can combine estimates of the correlation between the low-skill wage and the HWI with estimates of the link between supply shocks and the HWI to back out a (very) rough estimate of the implied wage elasticity—the short-run wage drop resulting from a supply shock. Our discussion of the Mariel episode indicated that an increase in the size of the low-skill workforce of about 18 percent reduced Miami's HWI by about 40 percent, so the elasticity relating the job vacancy index to an increase in supply $\left(\text{or } \frac{d\log H}{d\log L}\right)$ is about –2. Table 3 indicates that the elasticity relating the low-skill wage to the HWI $\left(\text{or } \frac{d\log w}{d\log H}\right)$ is +0.125. The implied wage elasticity is given by the product of these two numbers, or –0.25.

This estimate sits comfortably in the ballpark of what is usually reported in the literature (Blau and Mackie, 2016, Chapter 5), where the spatial correlations are often less than –0.2, while the national-level correlations tend to be larger (between –0.3 and –0.6). The –0.25 estimate is also remarkably close to the short-run elasticity implied by the simplest theoretical model that uses an aggregate (and linear homogeneous) Cobb-Douglas production function to examine the impact of immigration in a competitive market. The wage elasticity implied by that model is the negative of capital's share of income, which typically hovers around 0.3.

V. The Beveridge Curve and Supply Shocks

The BLS began to report the monthly unemployment rate for the metropolitan areas used in our analysis in 1976.[29] These data allow us to examine how the Mariel supply shock affected what is known as the "job-finding rate" in the macro literature as well as document how exogenous supply shocks shift the Beveridge curve.

Building on implications from search theory (Pissarides, 1985; and Mortensen and Pissarides, 1994), the canonical model of the Beveridge curve is closely linked to the matching

---

[29] The pre-1990 unemployment rate 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 and the Beveridge curve 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.

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 or slackness. We proxy the job-finding rate by the ratio of the HWI to the unemployment rate at the city-year-month level.

The top panel of Figure 9 illustrates the trend in the job-finding rate in Miami and in the synthetic control for the 1976-1989 period.[30] The figure shows a rapid decline in Miami's job-finding rate soon after Mariel, suggesting a weaker labor market, with a full recovery by the late 1980s. The bottom panel of the figure uses the HP filter to estimate the long-run trend in our proxy of the job-finding rate after removing the transitory (and city-specific) fluctuations. It is again evident that unemployed workers found it much harder to find jobs in post-Mariel Miami.

Table 4 reports the difference-in-differences regressions measuring the impact of the supply shock on the job-finding rate. Regardless of the specification, the regressions indicate 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).

As noted above, the matching function (with constant returns) 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). This model trivially 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 (Blanchard and Diamond, 1989, p. 17). Vacancies are fixed in the short run, but the number of unemployed workers rose. Over time, firms expand in response to the supply shock and the number of vacancies increase. 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.

---

[30] 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).

The top panel of Figure 10 illustrates the data scatter that forms 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, coincident with the deep 1980-1982 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). As with the national data, the Beveridge curve for the synthetic control also shifted out. Finally, the bottom panel shows the comparable 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—there is less unemployment for a given number of vacancies.[31]

We can estimate the size of 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:

$$\text{National:} \quad \log H_{rtm} = \theta_r + \theta_{tm} - 0.082\, u_{rtm} - 0.273\, (\text{Miami} \times \text{Post-Mariel}), \quad (5a)$$
$$\hspace{3.5cm} (0.013) \hspace{1.2cm} (0.028)$$

$$\text{Synthetic:} \quad \log H_{rtm} = \theta_r + \theta_{tm} - 0.117\, u_{rtm} - 0.345\, (\text{Miami} \times \text{Post-Mariel}), \quad (5b)$$
$$\hspace{3.5cm} (0.014) \hspace{1.2cm} (0.029)$$

where u_{rtm} gives the unemployment rate in cell (r, t, m); 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:

---

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

All cities: $\log H_{rtm} = \theta_r + \theta_{tm} - 0.062\ u_{rtm} - 0.179\ (\text{Miami} \times \text{Post-Mariel}).$ (5c)
              $\quad\qquad\qquad\qquad\qquad\quad (0.002)\qquad (0.029)$

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 "average" shifts 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 5 reports the value and standard errors of the (annual) interaction coefficients, while Figure 11 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.

The finding that a supply shock shifts the Beveridge curve inwards in the medium term is corroborated by an analysis that goes beyond Miami and allows immigration to affect other cities as well. In particular, we use decennial census data to estimate the vacancy-unemployment relationship using a sample of all 51 metropolitan areas in the HWI survey over the entire 1970-2000 period. In particular, consider the regression model:

$$\log H_{r\tau} = \theta_r + \theta_\tau + \delta\ u_{r\tau} + \beta\ p_{r\tau},$$ (6)

where p_{rt} gives the foreign-born share of the population in city r and census year τ. The unemployment rate is calculated in each decennial Census and is defined by the fraction of labor force participants in year τ who are unemployed in the survey week. We estimated the regression model in (6) using decadal census data (analogous to the spatial correlation models estimated in

the previous section), and the IV regression again uses the Jaeger, Stuhler, Ruist (2018) shift-share instrument calculated from the 1960 Census. The estimated regression models are:[32]

$$\text{OLS:} \qquad \log H_{r\tau} = \theta_r + \theta_\tau - 0.178\ u_{r\tau} - 0.007\ p_{r\tau}, \qquad (6a)$$
$$\qquad\qquad\qquad\qquad (0.042) \quad\ \ (0.008)$$

$$\text{IV:} \qquad \log H_{r\tau} = \theta_r + \theta_\tau - 0.147\ u_{r\tau} - 0.036\ p_{r\tau}. \qquad (6b)$$
$$\qquad\qquad\qquad\qquad (0.050) \quad\ \ (0.018)$$

It is evident that immigration produces an inward shift in the Beveridge curve once we adjust for the endogeneity of the geographic sorting of immigrants. If we take the IV coefficient at face value, a supply shock that increases the number of workers by about 8 percent (i.e., roughly the size of the Mariel shock) shifts the Beveridge curve inward by about 30 percent.

Although the Mariel evidence shows the immediate outward shift in the vacancy-unemployment locus implied by the simplest search-theoretic approach to the Beveridge curve, that model cannot account for its eventual inward shift. However, an intuitive extension of the model can easily generate such a shift.[33] The extension allows for the reality that the Mariel supply shock changed the demographics of Miami's workforce. As suggested by the framework in Barnichon and Figura (2015), heterogeneity in search behavior between immigrants and natives might shift the Beveridge curve because the matching function of the new arrivals differs from that of pre-existing workers. Such a difference could arise, for example, if the immigrants have a different reservation wage than natives.

To fix ideas, consider a special case of the Barnichon-Figura model. There are a total of I sectors in a labor market, and the matching function in sector i is the (constant returns) Cobb-Douglas function:

$$M_i = V_i^{1-\sigma}(s_i\ U_i)^\sigma, \qquad (7)$$

---

[32] The unit of analysis is a city-year cell. The regressions have 198 observations and are weighted by the size of the sample used to calculate the cell's unemployment rate.

[33] Shimer (2007) proposes an alternative micro foundation for the Beveridge curve, based on a mismatch model where the firms with vacancies are located in different markets than the unemployed workers looking for jobs. Appendix C shows that the mismatch model also predicts that a supply shock produces an inward shift in the Beveridge curve.

where M_i gives the number of matches in sector i, V_i the number of vacancies, and U_i the number of unemployed workers. In our context, it is useful to think of the labor market as a metropolitan area, with sectors delineated along industrial or neighborhood boundaries.[34] The variable s_i measures the average search efficiency of the unemployed in sector i, which depends on the demographic composition of the pool of unemployed workers in that sector. We scale s_i so that matching efficiency equals 1 for the typical unemployed worker.

There are two types of unemployed workers, natives ($j = 1$) and immigrants ($j = 2$). To easily derive the key insight, we make the simplifying assumption that each type of worker has the same search efficiency s_j regardless of which sector the worker searches in. Suppose further that newly arrived immigrants have a relatively low reservation wage, implying that $s_2 > s_1$ because unemployed immigrants are more likely to accept job offers. The average search efficiency in sector i is defined by:

$$s_i = \left(\frac{U_{i1}}{U_i}\right) s_1 + \left(\frac{U_{i2}}{U_i}\right) s_2, \tag{8}$$

where $U_i = U_{i1} + U_{i2}$. The total number of matches across sectors is $M = \sum_i M_i$, and the aggregate matching function is:

$$M = \mu V^{1-\sigma} U^\sigma, \tag{9}$$

where V and U give the total number of vacancies and unemployed workers, respectively. To a first-order approximation, the parameter μ, which measures the matching efficiency in the labor market, can be written as:

$$\mu = 1 + \sigma \left[\kappa_1 (s_1 - 1) + \kappa_2 (s_2 - 1)\right], \tag{10}$$

---

[34] 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 state-level data.

where $\kappa_j = U_j/U$, the fraction of unemployed workers in group j (Barnichon and Figura, 2015, pp. 228-229). Equation (10) implies that matching efficiency μ will be greater the larger the relative number of immigrants and the more that the search efficiency of immigrants exceeds that of natives.

An increase in μ in the matching function produces an inward shift in the Beveridge curve for the particular labor market (Elsby, Michaels, and Ratner, 2015, pp. 580-582, 584). The (assumed) search behavior of unemployed immigrants ensures a more rapid filling of vacancies after the supply shock.[35] Interestingly, the data indicate that the resulting "efficiency gain" in Miami's labor market vanished after a decade. The eventual return to the pre-Mariel equilibrium might be reflecting the process of immigrant assimilation, with the search efficiency of immigrants becoming more like that of natives over time.

In sum, Miami's vacancy-unemployment locus shifted out immediately after Mariel, consistent with the notion that the refugees initially entered the labor market as unemployed workers. The curve then shifted in (relative to its original position) by the mid-1980s, indicating that the large influx was composed of workers who were very efficient in their search behavior. Finally, the curve ended up where it began by the late 1980s, suggesting that the refugees' advantage in search efficiency disappeared over time.

VI. The Textual Content of Help-Wanted Ads

Although our analysis documents a short-term decline in the *number* of job vacancies in post-Mariel Miami, the trends in the HWI do not provide direct insight into *which* types of job vacancies are vanishing. 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, we need to examine the actual text of the help-wanted ads. In the past, such an analysis would have required a great deal of effort to

---

[35] Increasing returns to scale in the matching function may also explain why the Miami Beveridge curve shifted inwards (but not the outward shift in 1980 nor its eventual return towards the pre-1980 position). Petrongolo and Pissarides (2001) report mixed evidence on estimates of the returns to scale in matching functions.

read and categorize the advertisements. Modern machine learning and natural language processing techniques, however, 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 (2019; henceforth APST) shows, the analysis of the textual content of help-wanted ads can provide insights into how labor markets evolve.

An important obstacle, however, constraints the scope of such a study: digitized files that contain 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. For example, APST 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.

We pursued an alternative approach to construct a large sample of help-wanted ads from the *Miami Herald*. Specifically, we selected a random sample of issues published between 1978 and 1984 (essentially one issue from every month in that period).[36] 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[37], each ad was transcribed into text and then compiled into a single database. This approach created a random sample of 95,263 ads. Appendix D describes the process of creating this sample of ads from the *Miami Herald* (as well as two other newspapers introduced below). Appendix Figure D3 shows that the secular trend in the number of ads in our sample closely tracks the annual trend in the Miami HWI over the period.

The next step involves classifying each ad in terms of the type of job being advertised. This classification is relatively straightforward as the Department of Labor commissioned a private firm to create an algorithm, the O*NET-SOC AutoCoder, that classifies job titles and

---

[36] 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 from January 1978 to December 1984.

[37] The pipeline was developed using *Python*.

other job information into standard occupational codes. This algorithm "is the first affordable, commercially available system developed specifically to assign SOC-O*NET occupational codes to jobs, resumes and UI claims at an accuracy level that exceeds the level achieved by human coders" (Wilson, 2019). Specifically, we "fed" the entire text of each of the help-wanted ads in our sample to the algorithm (using code in the *Python* language), 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).[38]

Once each help-wanted ad is assigned an occupation code, it is easy to create a number of variables that describe the skill composition of that vacancy. To easily summarize the evidence, we use four alternative characterizations of the occupation's skill level: 1. Whether the vacancy advertises a blue-collar (or service) job opening; 2. The mean educational attainment (in years) of workers 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.[39] All of these variables were created using the sample of workers aged 25-64 in the 1980 Census.

The various panels of Figure 12 illustrate the 1978-1984 trends for each of these skill measures. To minimize month-to-month sampling error, 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. Finally, the average occupational wage in the advertised vacancy rose by 3 to 4 percentage points. Figure 12 leaves little doubt 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 low-skill vacancies.

---

[38] 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).

[39] 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.

Of course, 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. As noted earlier, however, the construction of an appropriate control group is difficult because the digital files of the ads published by many newspapers is not yet available (or is often available only at a prohibitive cost).

APST generously gave us access to the samples of help-wanted ads they created for the *New York Times* and the *Boston Globe*. In particular, they provided us with a data file that reported the number of ads published in each occupation-year-month-day cell for each newspaper.[40] We then collected comparable data for two other newspapers from the set that ProQuest made available to the Princeton University Library.[41] 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.

The ProQuest files for these newspapers include their *entire* published content. Using the machine-learning methods developed by APST, 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 obtained 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 12 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

---

[40] The APST data consist 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. APST also shared the comparable data they created for 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.

[41] 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 newspapers that satisfied these restrictions at the time 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*.

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.

Figure 13 shows a more detailed illustration of exactly what happened to the *distribution* of the advertised wage offers in Miami and in the control cities before and after Mariel. Specifically, it documents the change in the cumulative distribution of wage offers. It is evident that the distribution of the occupational wage remained unchanged in the control newspapers. In contrast, Miami's cumulative distribution shifted downwards. In fact, the figure suggests that the post-Mariel distribution stochastically dominates the pre-Mariel distribution. Moreover, the shift was economically important. Before 1980, the top quartile of vacancies offered an occupational wage of at least 0.21. By 1983-1984, 31.1 percent of the surviving vacancies offered an occupational wage in that range.

Table 6 reports the relevant coefficients from the difference-in-differences regression model 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.

Regardless of the specification, 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 6.5 percent by 1982; the fraction of workers in the (typical) advertised job who lacked a high school diploma dropped by 3.0 percent; 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 focuses on the *textual* content of advertised vacancies confirms a substantial drop in the relative availability of low-skill jobs in post-Mariel Miami.[42]

---

[42] 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 score altogether.

The analysis of the HWI reported in earlier sections and the complementary analysis of the textual content of ads in this section document that both the number of vacancies *and* the relative number of low-skill vacancies declined in post-Mariel Miami. We can combine these two separate findings to get a sense of what happened to the actual number of low-skill vacancies. Our simple back-of-the-envelope exercise takes the HWI as the "official" count of help-wanted ads in Miami and in the control cities in each year between 1978 and 1984. We then multiply the HWI in each of the cities times the fraction of ads advertising a blue-collar job and rescale the product in each city to equal 1 in 1979.

Figure 14 illustrates the resulting trends. It is obvious that the number of low-skill vacancies fell substantially in both Miami and the control group after 1980 because of the recession. It is also evident, however, that the low-skill labor market in Miami was the hardest hit, with the number of low-skill vacancies dropping by nearly 70 percent between 1979 and 1982 (as compared to a 50 percent drop in the control group). The end of the recession led to a widespread recovery, but the recovery was not as strong in Miami. The number of low-skill vacancies in that city still lagged behind by about 20 percent in 1984. In short, our statistical text analysis strongly suggests that unemployed low-skill workers in Miami suffered not only from the reduction in advertised job openings but additionally from shifts in the *composition* of advertised job openings away from low-skill work.

VII. Summary

This paper addresses what is perhaps the central economic question in the immigration debate: 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 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 natural experiments created by random and sudden

labor supply shocks 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 boatlift in 1980. The HWI dropped relative to the trend observed in many alternative control groups between 1980 and 1986. Miami's HWI recovered fully by 1990.

2. 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.

3. In the past few decades, immigrant flows to the United States have been disproportionately low-skill. Although the HWI presumably measures "average" local labor market conditions, the persistent negative correlation between immigration and the HWI suggests that the index might be a particularly good barometer for labor market conditions at the bottom end of the skill distribution. Our analysis indeed indicates that the HWI seems to be more strongly correlated with wage and employment trends for the least educated workers.

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 the changed demographics of the Miami labor market, with the Mariel refugees perhaps having lower reservation wages and more quickly filling up vacancies. Miami's Beveridge curve returned to its pre-Mariel (relative) position by the late 1980s.

5. 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.

In sum, our evidence consistently indicates that immigration-induced supply shocks are typically followed by a short-run period of slackness in the local labor market (as measured by the number of advertised job openings). 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 2010): 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?" *Journal of Political Economy*, Vol. 94, No. 3, Part 1.

Anastasopoulos, Jason, George J. Borjas, Gavin G. Cook, and Michael Lachanski. 2018. "Job Vacancies and Immigration: Evidence from Pre- and Post-Mariel Miami," NBER Working Paper No. 24580.

Anderson, Patricia M. and Simon M. Burgess. 2000. "Empirical Matching Functions: Estimation and Interpretation Using State-Level Data," *Review of Economics and Statistics* 82(1): 93-102.

Andolfatto, David. 1996. "Business Cycles and Labor-Market Search," *American Economic Review*, 86(1): 112-132.

Atalay, Enghin, Phai Phongthiengtham, Sebastian Sotelo, and Daniel Tannenbaum. 2019. "The Evolution of Work in the United States," *American Economic Journal: Applied Economics*, forthcoming.

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

Barnichon, Regis and Andrew Figura. 2015. "Labor Market Heterogeneity and the Aggregate Matching Function," *American Economic Journal: Macroeconomics* 7(4): 222-249.

Belasen, Ariel R. and Solomon W. Polachek, 2009. "How Disasters Affect Local Labor Markets: The Effects of Hurricanes in Florida," *Journal of Human Resources* 44: 251-276.

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 (30): 774.

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

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation," *Journal of Machine Learning Research* 3, no. Jan (2003): 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,' *Quarterly Journal of Economics* 118 (November 2003): 1335-1374.

Borjas, George J. 2017. "The Wage Impact of the *Marielitos*: A Reappraisal," *Industrial and Labor Relations Review*, 70(5): 1077-1110.

Borjas, George J. 2019. "The Wage Impact of the *Marielitos*: The Role of Race," *Industrial and Labor Relations Review*, 72(4): 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*, Vol. 50, No. 2.

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

Chassamboulli, Andri and Giovanni Peri. 2015. "The Labor Market Effects of Reducing Undocumented Immigrants," *Review of Economic Dynamics* 18: 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: 818-857.

Cohen, Malcolm S. and Robert M. Solow. 1967. "The Behavior of Help-Wanted Advertising", *Review of Economics and Statistics*, 49(1).

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

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 1990): 96–107.

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

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

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

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: 259-314.

Lazear, Edward, Kathryn Shaw, and Christopher Stanton. 2018. "Who Gets Hired? The Importance of Competition among Applicants," *Journal of Labor Economics* 36(1): 133-181.

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

Ortega, Javier. 2000. "Pareto-Improving Immigration in an Economy with Equilibrium Unemployment," *Economic Journal* 110 (January 2000): 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(2), pp. 267-309.

Petrongolo, Barbara and Christopher A. Pissarides. 2001. "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature* 39(2): 390-431.

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

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

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

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

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

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, 1975.

Warren, Donald. 1982. "Immigration and the Natural Rate of Unemployment in Australia," *Journal of Macroeconomics* 4 (Fall 1982): 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 1985): 554-564.

Zagorsky, Jay L. 1998. "Job Vacancies in the United States: 1923 to 1994," *Review of Economics and Statistics*, 80(2).
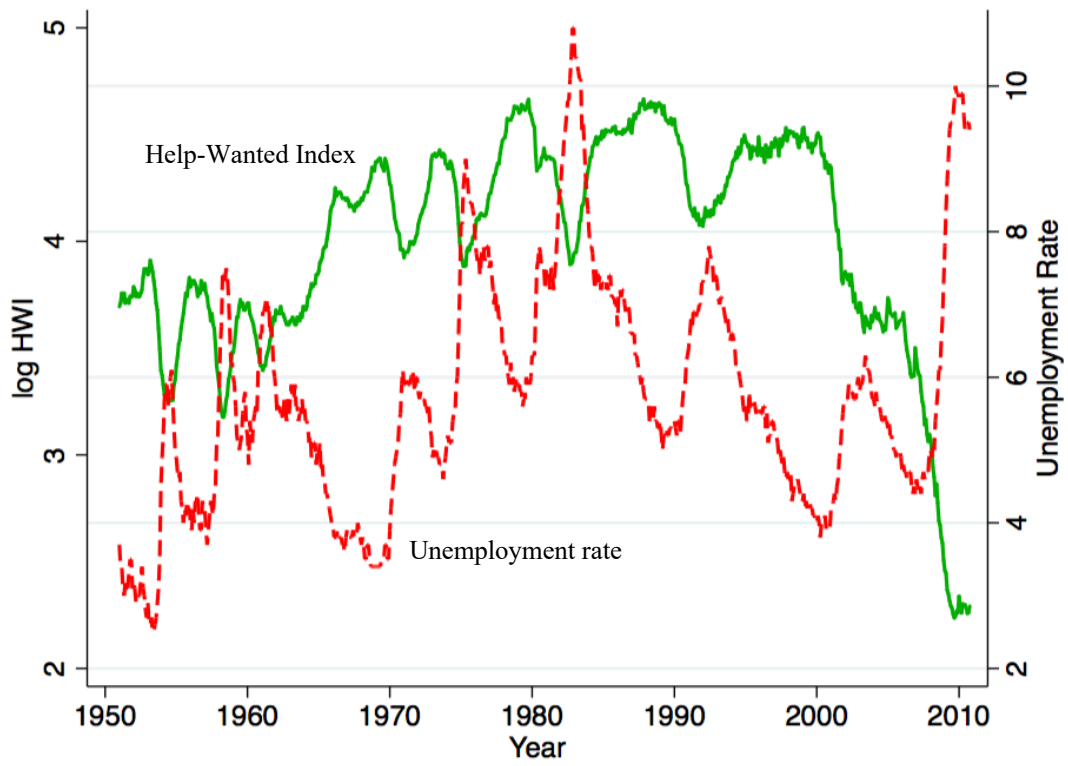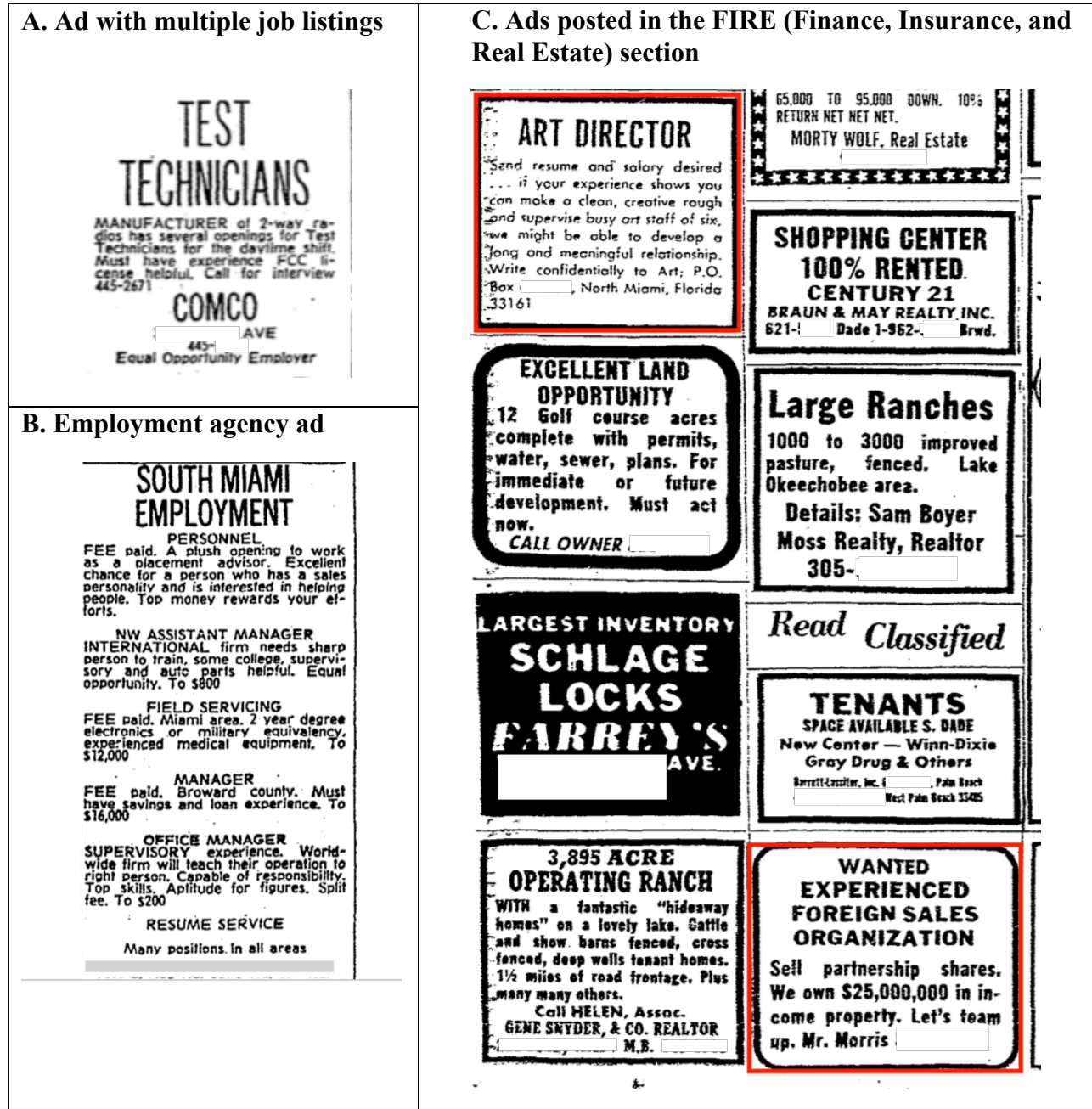
**Figure 1. The national Help-Wanted Index (HWI) and the unemployment rate**

Figure 2. Enumeration biases in the Help-Wanted Index

A. Ad with multiple job listings	C. Ads posted in the FIRE (Finance, Insurance, and Real Estate) section
B. Employment agency ad	

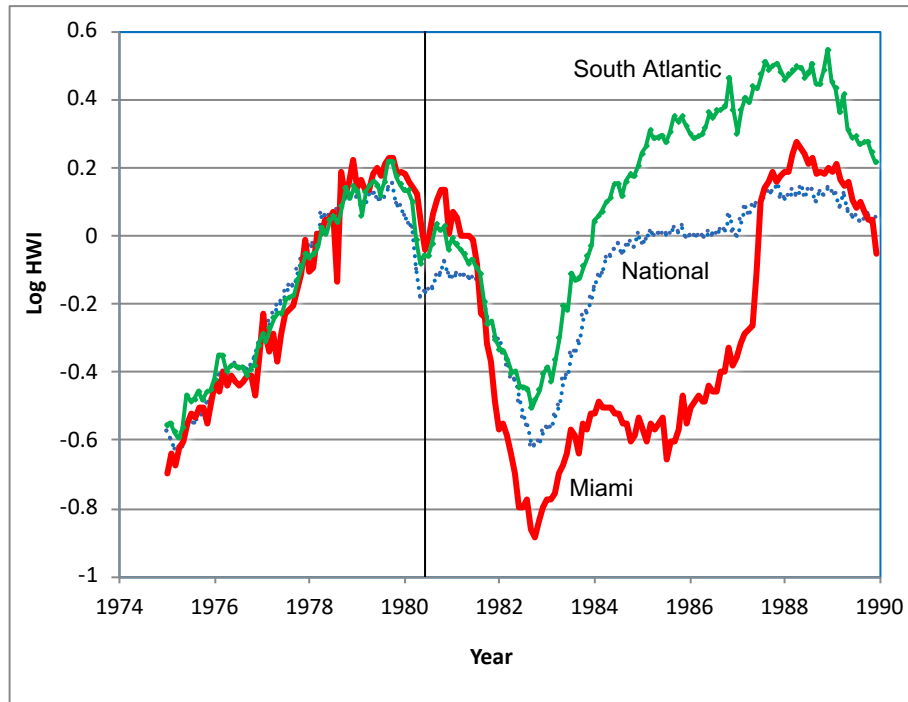Notes: All ads appeared in the March 2, 1975 edition of the *Miami Herald*.

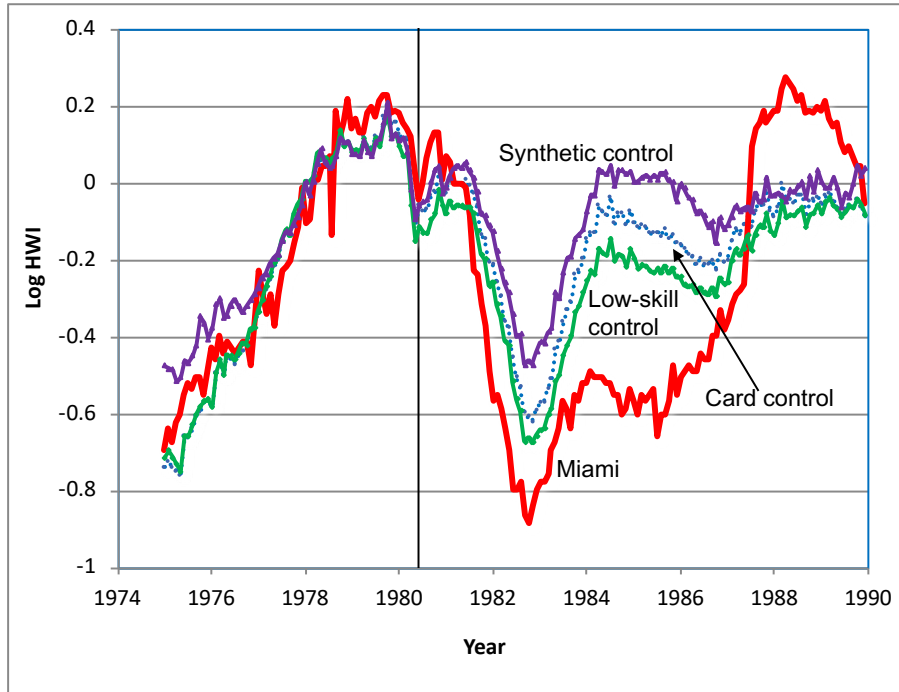**Figure 3. Cuban immigration to the United States, 1955-2010**
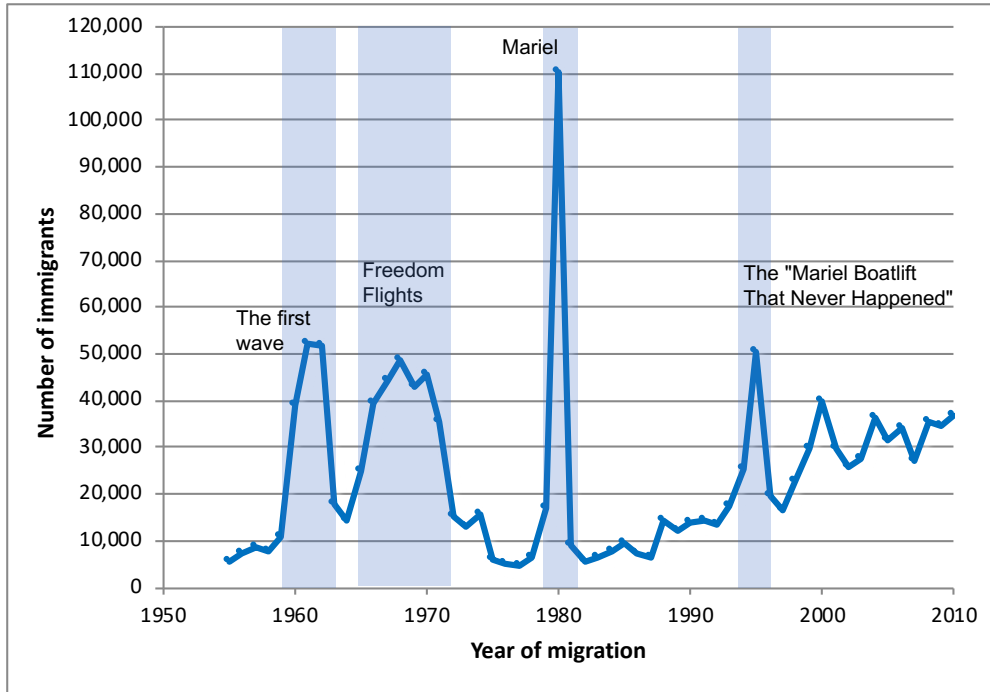
Source: Adapted from Borjas (2017), p. 1080.

**Figure 4. 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 5. Job vacancies in Miami relative to control cities, 1975-1989**

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

**Figure 6. Number of help-wanted ads published
in the *Miami Herald* and *El Nuevo Herald,* 1978-1984**
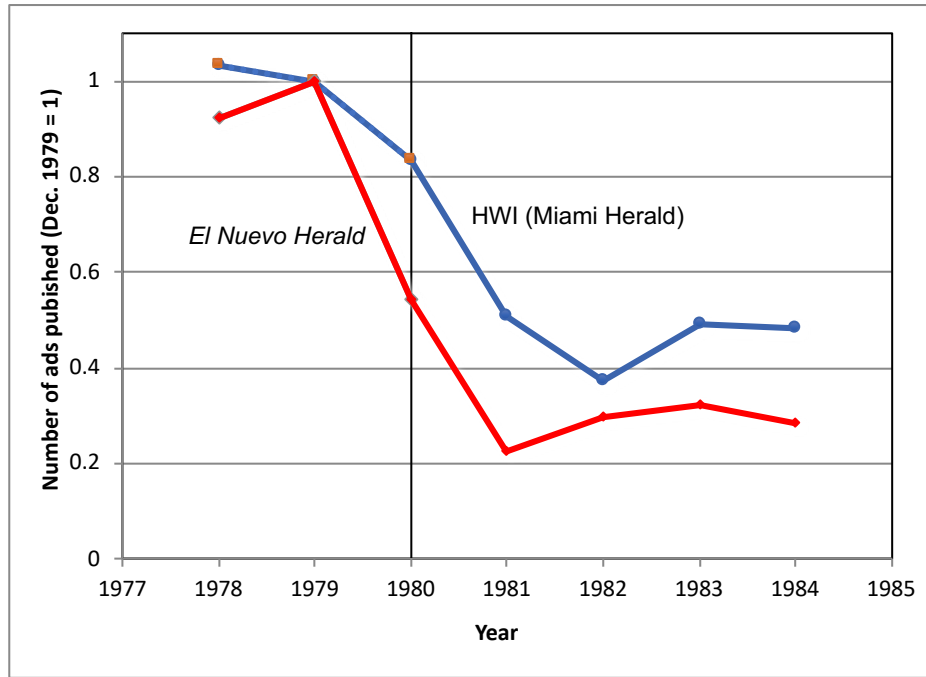
Notes: The number of ads published in the *Miami Herald* is given by the HWI in Miami in December of each year. We manually counted the number of ads published in *El Nuevo Herald* in each December between 1978 and 1984. Each of the trends is rescaled to equal one as of December 1979.

Figure 7. Removing local transitory fluctuations from the HWI using the HP filter
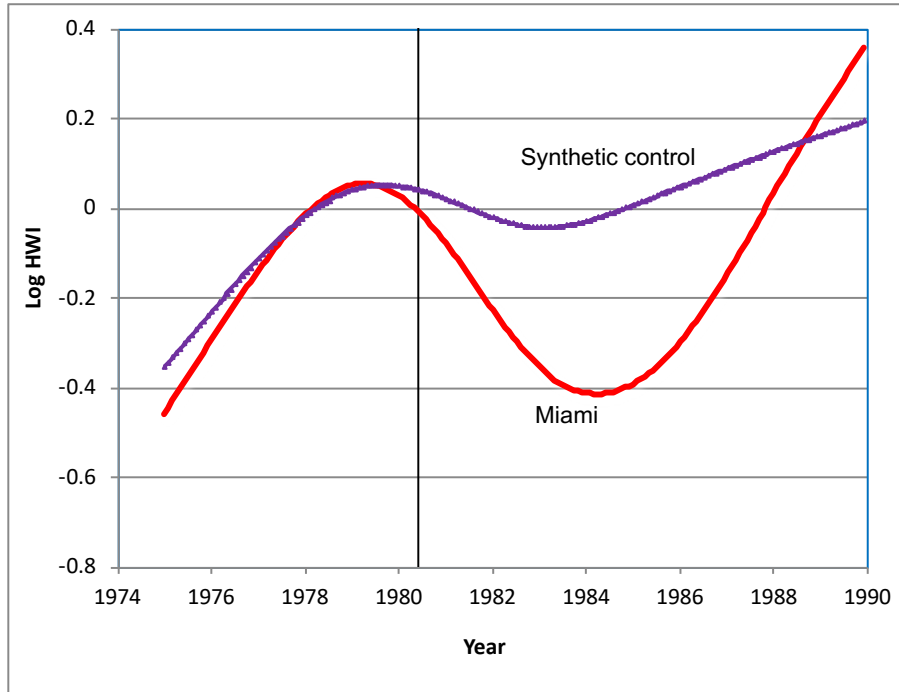
Notes: The HWI in in the figure is the long-term trend predicted by the HP filter with a smoothing parameter of 129,600, using all monthly observations between 1975 and 1989. The HWI for each city is rescaled to equal 1 in 1977-1979.

**Figure 8. Frequency distribution of short-run impact of Mariel
(relative to all potential four-city control groups)**
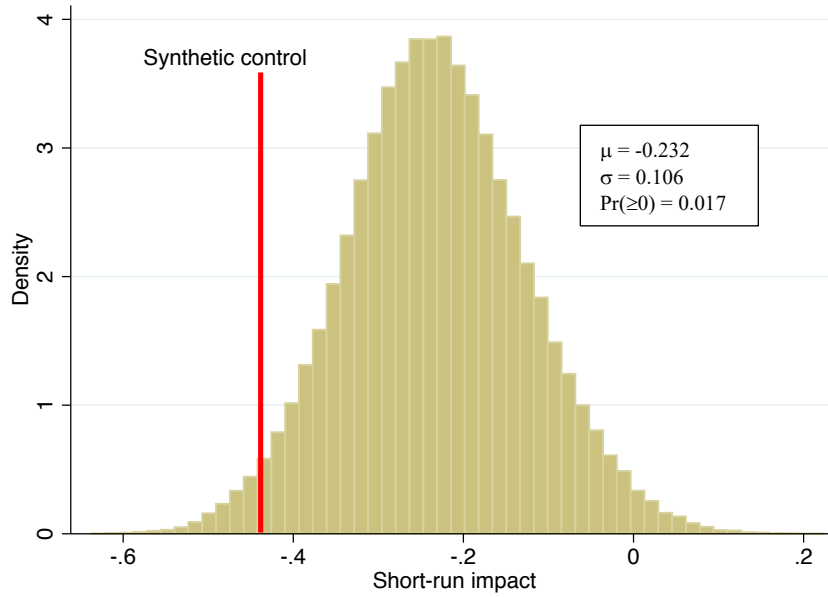
Notes: The sample period used in the analysis is 1979-1984 with a treatment date of June 1980. The figure shows the density function for the relevant coefficient from the difference-in-differences (log) HWI regression that compares Miami to all potential 230,300 four-city control groups. All regressions include city and year-month fixed effects.

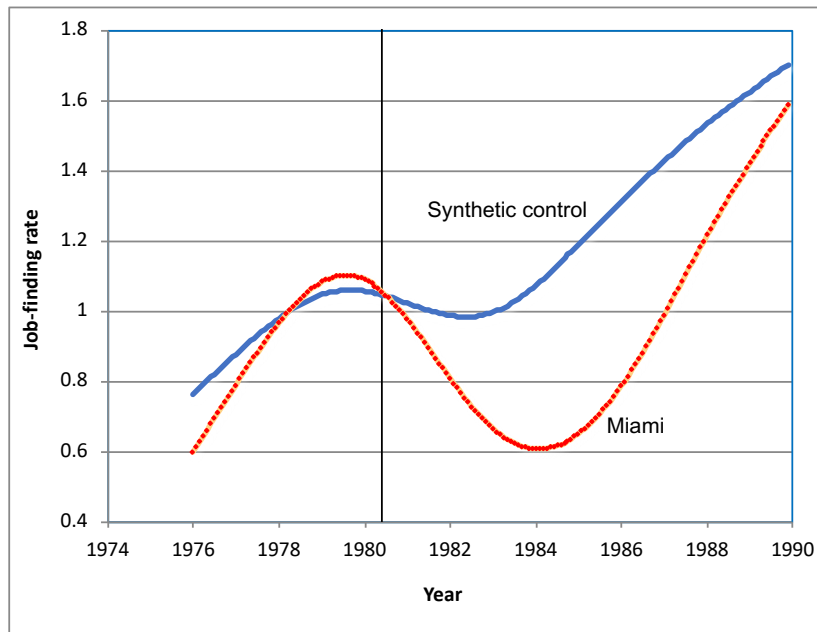**Figure 9. The job-finding rate in Miami relative to the synthetic control**

A. **Unadjusted data**
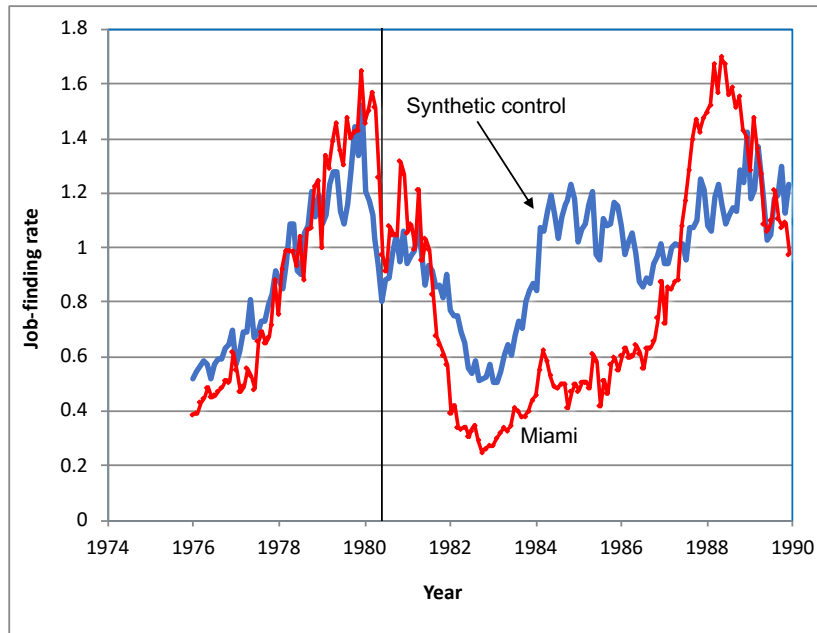
B. **Removing local transitory fluctuations using HP filter**

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

Figure 10. The Beveridge Curve and Mariel, 1976-1984

A. The national labor market
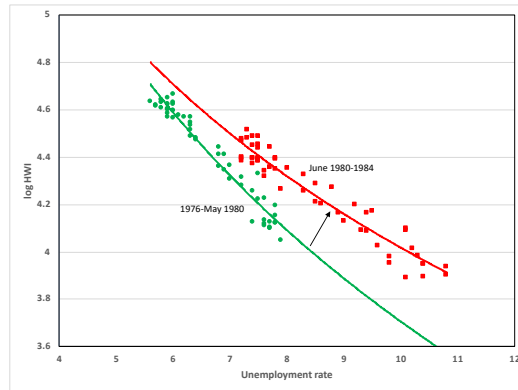
B. The synthetic control
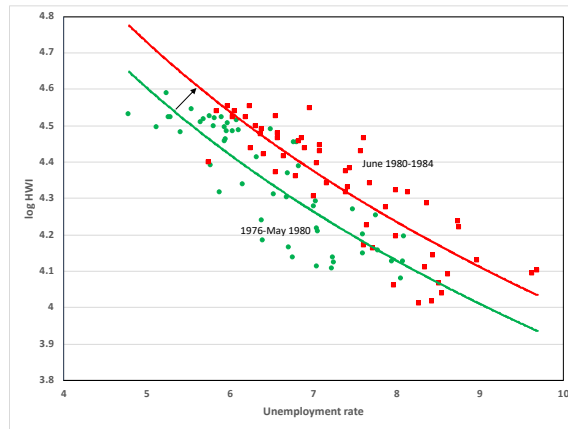
C. Miami
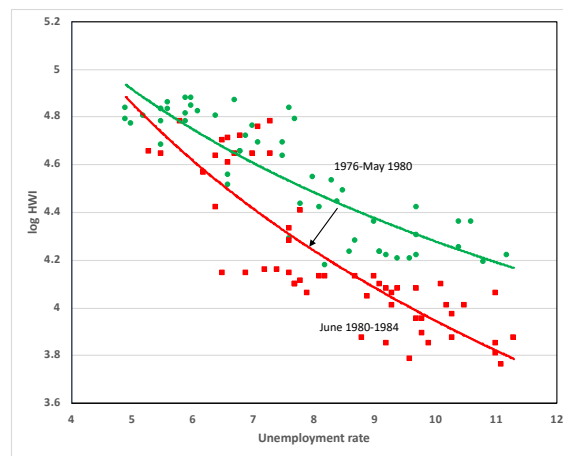
Notes: The figures show the data scatter and the logarithmic trend lines relating the raw HWI and the unemployment rate in a city-year-month cell in the pre-Mariel (January 1976-May 1980) and post-Mariel (June 1980-December 1984) periods. The Beveridge curve for the synthetic control in Panel B is constructed 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 (Table 4, column 2).

**Figure 11. Year-by-year shift in relative intercept of Miami's Beveridge curve
(Baseline is January 1976-May 1980)**
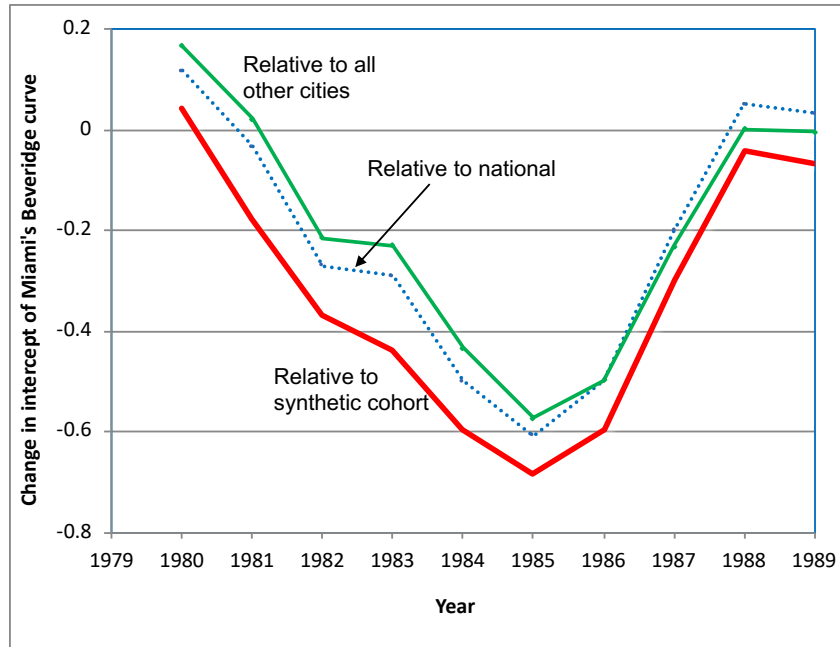
Notes: The figure illustrates the value of the interaction coefficient in a Beveridge curve regression model between a Miami indicator variable and a post-Mariel year (with the 1980 year being denoted by the June 1980-December 1980 period). The dependent variable is the log of the HWI index in an area. The regressions include the area's unemployment rate, area fixed effects, and year-month fixed effects.

Figure 12. Trends in the skill composition of advertised vacancies, 1978-1984

A. Percent blue collar

B. Percent high school dropout

C. Educational attainment

D. Log hourly wage
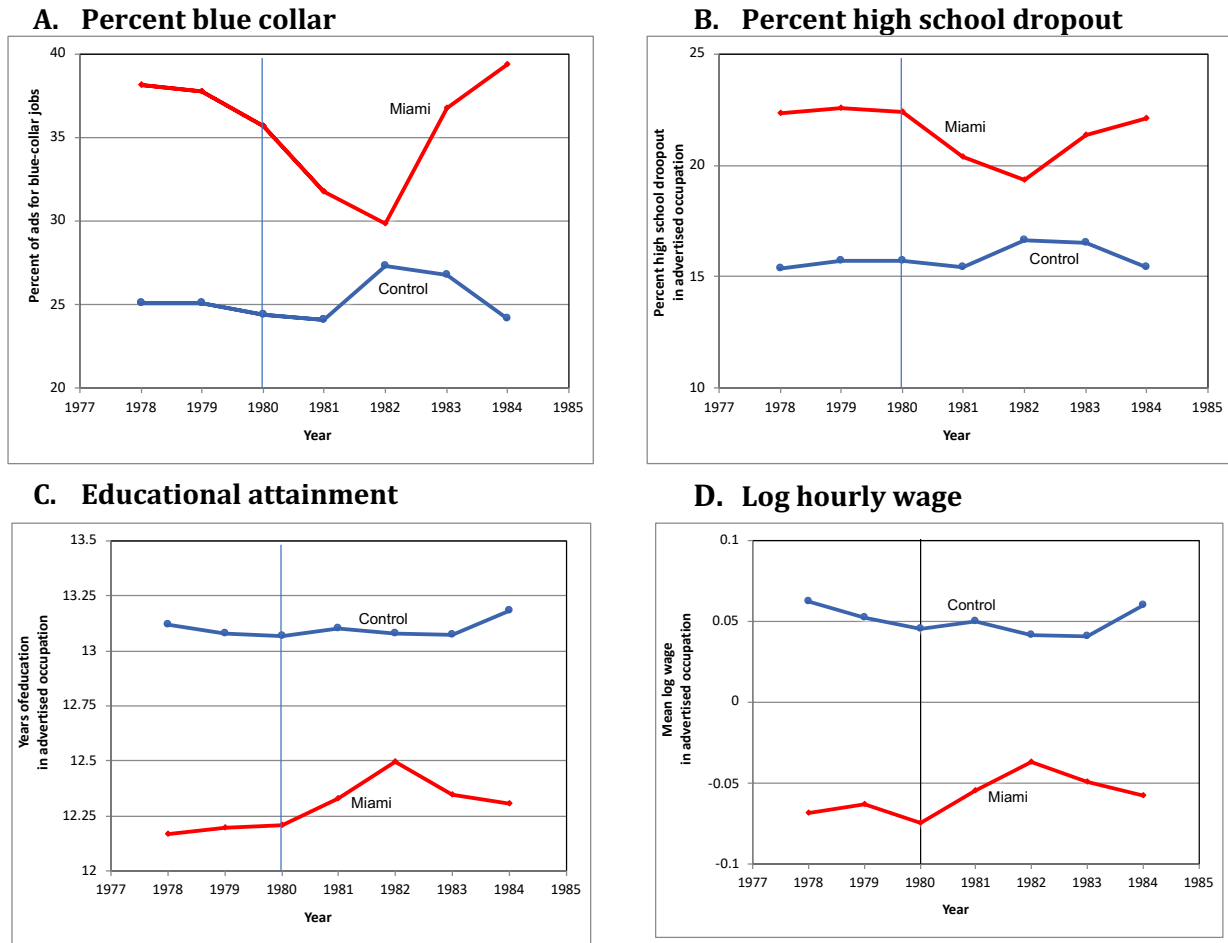
Notes: Panel A gives the fraction of ads advertising for a blue-collar position; panel B gives the fraction of workers who are high school dropouts in the average advertised occupation; panel C gives the average years of education in advertised occupations; and panel D gives the average (adjusted) mean 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 13. Cumulative distributions of occupational wage offers in advertised vacancies

A. Miami Herald
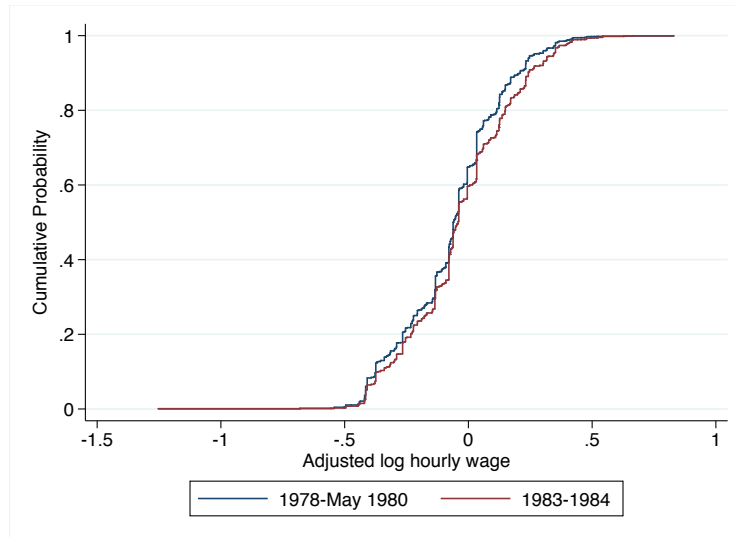
B. Control newspapers
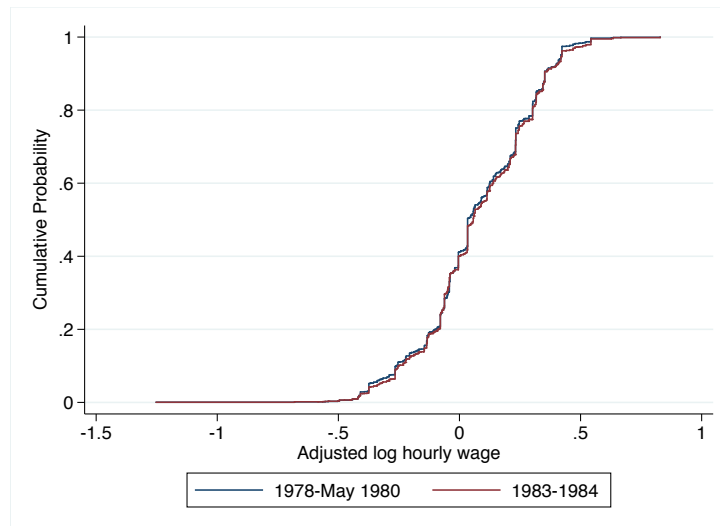
Notes: The figures show the cumulative distribution of the adjusted log hourly wage in the (average) advertised occupation. The occupational wage is calculated from the 1980 census. The size of the sample in the *Miami Herald* is 57,946 ads; the size of the sample in the control newspapers (the *Boston Globe*, the *Minneapolis Star-Tribune*, the *New York Times*, and the *St. Louis Post-Dispatch*) is 1,634,227 ads.

Figure 14. Predicted number of blue-collar vacancies, 1978-1984
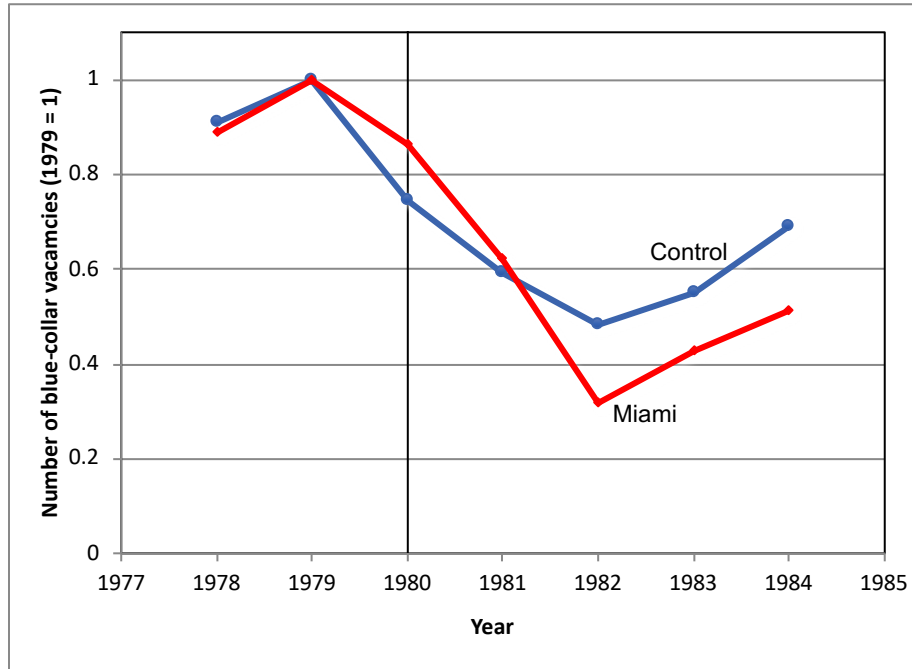
Notes: The predicted number of blue-collar vacancies is given by the product of a city's HWI times the fraction of blue-collar ads calculated from the textual analysis of the city's newspaper. The product is rescaled to take on a value of 1 in 1979 in both Miami and in the control group. The size of 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.

Table 1. Difference-in-differences impact of Mariel supply shock
(Dependent variable = log HWI)

Variable:	Card control	Low-skill control	All cities	Synthetic control	
				Unadjusted	HP filter
A. Log HWI					
June 1980-1982	-0.204	0.006	0.029	-0.237	-0.172
	(0.041)	(0.045)	(0.037)	(0.036)	(0.014)
1983-1984	-0.482	-0.186	-0.296	-0.436	-0.342
	(0.029)	(0.026)	(0.028)	(0.023)	(0.007)
1985-1986	-0.446	-0.225	-0.493	-0.448	-0.308
	(0.042)	(0.025)	(0.019)	(0.029)	(0.011)
1987-1989	0.109	0.129	-0.041	0.138	-0.000
	(0.033)	(0.032)	(0.028)	(0.029)	(0.021)
B. Log normalized HWI					
June 1980-1982	-0.179	-0.059	-0.039	-0.257	-0.198
	(0.034)	(0.044)	(0.037)	(0.033)	(0.014)
1983-1984	-0.449	-0.298	-0.364	-0.410	-0.372
	(0.025)	(0.023)	(0.026)	(0.020)	(0.007)
1985-1986	-0.403	-0.358	-0.544	-0.395	-0.345
	(0.036)	(0.025)	(0.018)	(0.026)	(0.010)
1987-1989	0.130	-0.017	-0.099	0.138	-0.049
	(0.028)	(0.034)	(0.027)	(0.028)	(0.020)

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 normalized HWI is the index divided by employment in the city-year-month cell. The regression that uses the Card control has 720 observations; the regression that uses the low-skill control has 900 observations; the regression that uses the all-city sample has 9,180 observations; and the regressions that use the synthetic control have 360 observations.

Table 2. 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 3. Correlation between labor market conditions and the HWI, by education group

| | | Log HWI interacted with: | | |
	Log HWI	High school dropout	High school graduate	Some college
March CPS (1972-1999)				
1. Log weekly wage	-0.009	0.125	0.069	0.038
	(0.013)	(0.021)	(0.013)	(0.013)
2. Employment propensity	0.002	0.043	0.034	0.023
	(0.004)	(0.006)	(0.004)	(0.004)
CPS-ORG (1979-1999)				
1. Log hourly wage	-0.037	0.109	0.066	0.045
	(0.012)	(0.021)	(0.012)	(0.012)
2. Employment propensity	0.019	0.038	0.034	0.026
	(0.004)	(0.006)	(0.005)	(0.004)

Notes: Robust standard errors in parentheses. The number of observations is 4,604 in the March CPS regressions and 3,657 in the ORG regressions. Both the wage and employment variables are age- and gender-adjusted. The employment variable in the March CPS gives the probability that the person worked at some point during the calendar year prior to the survey, while the employment variable in the CPS-ORG gives the probability that the person worked during the CPS reference week. All regressions are weighted by the number of observations used to calculate the dependent variable.

Table 4. Difference-in-differences impact of Mariel supply shock on the job-finding rate (relative to synthetic control)

Variable:	Unadjusted series	HP filter
June 1980-1982	-0.190	-0.151
	(0.043)	(0.019)
1983-1984	-0.479	-0.427
	(0.043)	(0.016)
1985-1986	-0.462	-0.491
	(0.039)	(0.010)
1987-1989	0.121	-0.243
	(0.046)	(0.018)

Notes: Robust standard errors are reported in parentheses. The data consist of monthly observations for each city between 1976 and 1989. The job-finding rate is the ratio of the HWI to the unemployment rate in each city-year-month cell. It is rescaled to equal 1 in each city in 1977-1979. 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). All regressions have 336 observations.

Table 5. 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.

**Table 6. Difference-in-differences impact of Mariel supply shock
on skill composition of help-wanted ads, 1978-1984**

Variable:	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 by number of ads in cell				
June 1980-1982	-6.503	-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.

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: Job Vacancies and the Other Supply Shocks in Miami

1. The First Wave of Cuban Refugees

Fidel Castro toppled the U.S.-supported government of Fulgencio Batista on January 1, 1959. There was little migration of Cubans to the United States during the Batista years (see Figure 3). The number of immigrants remained low in 1959, but began to increase rapidly soon after that, as the totalitarian nature of the communist regime came to the surface. In 1960, nearly 40,000 Cubans migrated, and over 50,000 Cubans migrated in both 1961 and 1962.[43]

The first wave of Cuban refugees ended abruptly. Most of those refugees left Cuba on a twice-daily commercial flight that Pan American Airways operated between Havana and Miami. On October 16, 1962, President Kennedy was informed that the Soviet Union had placed medium-range ballistic missiles in Cuba, setting in motion the Cuban Missile Crisis. Within a week, the Pan American flights were discontinued, and the Cuban exodus ended.

The flow of refugees resumed in December 1965, after the establishment of the "Freedom Flights" that reunited families separated by the migration hiatus that followed the Missile Crisis. An average of 44,000 Cubans migrated to the United States each year between 1966 and 1969. The flow dwindled down to a small number (about 15,000 a year) by the early 1970s.

The top panel of Table B1 uses the 1960 and 1970 decennial censuses to calculate some basic facts about the size and skill composition of the first wave of Cuban immigrants that arrived between 1960 and 1962. The influx of refugees increased the size of Miami's workforce by 16.6 percent—almost twice as large as the 8.4 percent increase resulting from the Mariel boatlift. The supply shock was remarkably large at the upper end of the education distribution. The first wave of Cuban refugees increased the number of college graduates in Miami's workforce by over 40 percent.

However, the typical immigrant in that early wave experienced a substantial amount of "down-skilling," with employers heavily discounting their credentials. The bottom two rows of the panel show the extent to which the Miami labor market downgraded (at least initially) the educational attainment of the immigrants. The average Cuban who did not have a college diploma ended up working at a job where the average native was a high school dropout. And even the average Cuban college graduate ended up at a job that employed natives who had just one year more than a high school education. In an important sense, therefore, the initial supply shock of Cuban refugees was also disproportionately low-skill.

The top panel of Figure B1 shows the raw trends in the HWI for the 1954-1974 period, comparing Miami with both the South Atlantic region and the nation as a whole, while the bottom panel compares Miami to the synthetic control. The synthetic control in the bottom panel of the figure uses exactly the same specification employed in the Mariel analysis, except that all the control variables are now calculated using the 1950 and 1960 IPUMS census files (the treatment is timed as of January 1960).[44] The figure shows that the HWI for Miami and the

---

[43] As far as we know, we are the first to examine the labor market consequences of this supply shock, most likely because the earliest public-use CPS microdata starts in 1962 and the Miami metropolitan area is not identified until 1973.

[44] The HWI data for some of the metropolitan areas did not begin to be collected until after 1954. We use the subsample of the 44 metropolitan areas that have a complete time series for the period. The cities in the synthetic control are Charlotte (0.31), Jacksonville (0.085), Memphis (0.064), San Bernardino (0.373), Sacramento (0.023), and San Diego (0.144).

synthetic control had similar trends in the Batista years, but that the post-1960 Miami experience is unusual. The index for Miami began to decline (relative to the synthetic control) soon after the Cuban refugees began to arrive in 1960 and had begun to recover by the mid-1960s. The recovery, however, did not take hold. As soon as the Freedom Flights started in 1966, the HWI for Miami again began to diverge. The two indices reached parity again after the end of the Freedom Flights in the early 1970s, but the recovery was short-lived. The deep 1973-1975 recession led to a severe (relative) decline in Miami's HWI.

The first column of Table B2 estimates the regression model in equation (1) using data from the 1954-1974 period. The coefficients indicate that Miami's HWI declined relative to the synthetic control between 1960 and 1963 and had recovered somewhat by the middle of the 1960s. There was continued recovery after the beginning of the Freedom Flights, but the gap had again grown to about 25 percent by the time that program ended in 1970. The situation improved in the early 1970s; by 1972-1974, the HWI in Miami was only 9.8 percent below that of the synthetic control.

2. The "Mariel Boatlift That Did Not Happen"

There was another spike in Cuban immigration in 1995. The number of immigrants rose from about 18,000 in 1993 to over 50,000 in 1995, before quickly falling to below 20,000 in 1996. This spike coincides with the period examined by Angrist and Krueger (1999, p. 1328):

> In the summer of 1994, tens of thousands of Cubans boarded boats destined for Miami in an attempt to emigrate to the United States in a second Mariel Boatlift that promised to be almost as large as the first one…Wishing to avoid the political fallout that accompanied the earlier boatlift, the Clinton Administration interceded and ordered the Navy to divert the would-be immigrants to a base in Guantanamo Bay. *Only a small fraction of the Cuban émigrés ever reached the shores of Miami*. Hence, we call this event, "The Mariel Boatlift That Did Not Happen." [emphasis added]

At the time that Angrist and Krueger wrote about this episode, they could not know what the 2000 Census would eventually uncover: The Mariel boatlift that did not happen indeed happened; it was just delayed by a year. The refugees diverted to Guantanamo made it to the United States after President Clinton reversed course in May 1995 and permitted their entry.

We examine the labor market consequences of this supply shock by focusing on the 1990-1999 period. This supply shock differs in crucial ways from the other waves of Cuban immigrants. First, although it involved a sizable number (with over 75,000 Cubans migrating in 1994-1995), Miami was a much larger city by the mid-1990s. As a result, the relative increase in supply was much smaller—only about 3.9 percent (see the bottom panel of Table B1). Second, the aftermath of the 1994-1995 supply shock was not followed by a hiatus in Cuban immigration. The steady increase in Cuban immigration after the 1994-1995 shock may make it more difficult to detect any recovery.

Figure B3 shows the raw trend in Miami's HWI during the 1990s, contrasting it with the trend in both the South Atlantic region and in the national index. A notable feature of the figure is the huge upward spike in the Miami index in the last half of 1992. The value of the HWI for Miami almost doubled between August and November. This spike coincides exactly with the aftermath of Hurricane Andrew, a Category 5 hurricane that made landfall in Homestead,

Florida, on August 24, 1992. At the time, it was the strongest hurricane to ever make landfall in the United States, causing $45 billion in damage (in 2018 dollars).

The behavior of Miami's HWI in the aftermath of Andrew dramatically shows how a local labor market changes after a major environmental disaster that requires a lot of rebuilding (Belasen and Polachek, 2009). After the quick spike, the HWI began to decline slowly until late 1993. It then remained relatively stable through 1995, at which point the index began a steady relative decline. Note that the raw data do not suggest any type of recovery in the Miami index in the post-1995 period.

We applied the synthetic control method to examine the impact of the 1994-1995 supply shock. We time the treatment as of August 1992, so that the method will create a synthetic city that resembled Miami prior to *both* the hurricane and the supply shock. The set of controls is exactly the same as that used earlier in our analysis of the other shocks, with the only change being that we used the 1980 and 1990 IPUMS census files to construct the control variables.[45]

The bottom panel of Figure B3 shows the trend in Miami's HWI relative to the synthetic control. The comparison reveals the same insights as the raw data. There was a remarkable spike due to Hurricane Andrew with the index becoming relatively stable throughout the 1994 calendar year, but then the index began to decline with the arrival of the refugees from Guantanamo Bay. Notably, Miami's HWI continued to decline throughout the last half of the 1990s (coincident with the steady increase in Cuban immigration).

The second column of Table B2 reports the coefficients from the difference-in-differences regression model.[46] The specification of the model allows for the separate identification of the impact of Hurricane Andrew. In particular, we use the interregnum between Andrew and the supply shock (i.e., the period between January 1994 and June 1995) as the baseline. As the first column of the table shows, Andrew increased the HWI by about 37 percent, and the index fell (relative to the baseline) by 15 to 20 percent between 1995 and 1999. Notably, there is no evidence that the Miami index recovered after the 1994-1995 supply shock. The continued decline in Miami's index after the supply shock can perhaps be attributed to the steady increase in Cuban immigration after 1996.

---

[45] The cities in the synthetic control are Los Angeles (0.448), New Orleans (0.116), New York (0.178), Providence (0.022), and San Bernardino (0.237).

[46] The relative decline of the post-1995 Miami HWI could have resulted from a faster rate of adoption in online ads in Miami, with a corresponding decline in printed ads. Anastasopoulos et al (2018) estimate a regression model that uses data on a metropolitan area's intensity of computer use (available in CPS supplemental surveys) to examine this hypothesis. The adjustment barely affects the coefficients reported in Table B2.

Figure B1. The Help-Wanted Index in Miami, 1954-1974

A. Relative to regional-national aggregates
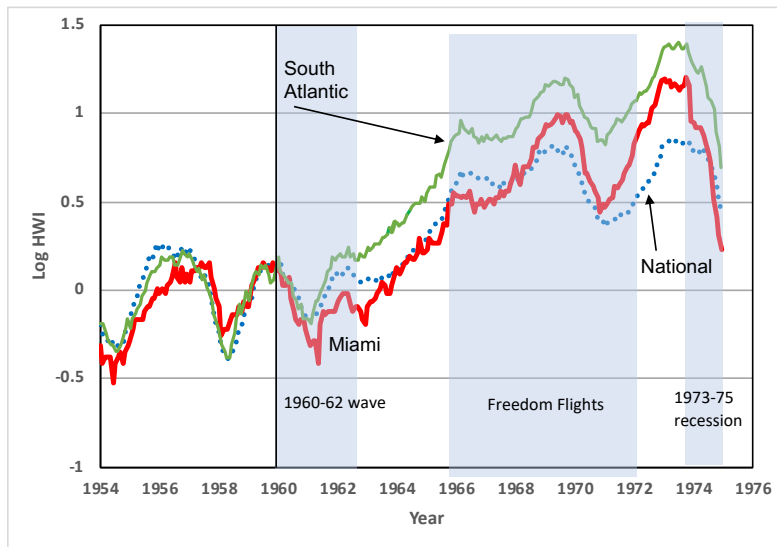
B. Relative to the synthetic control
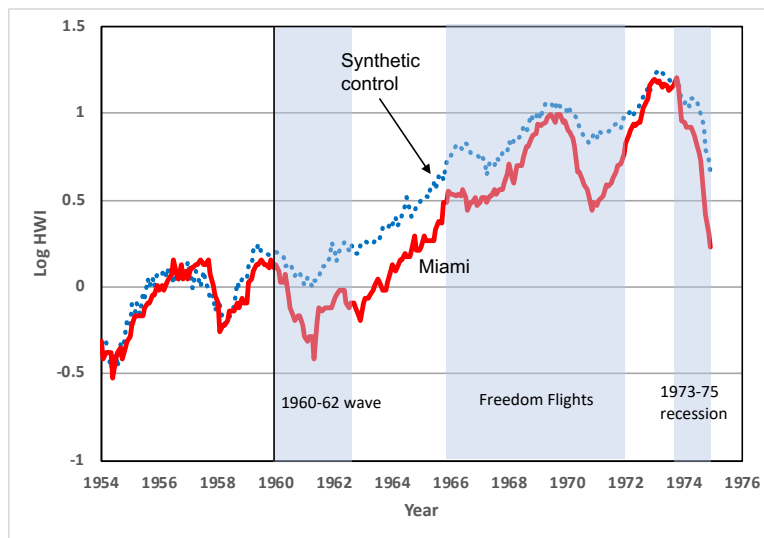
Notes: The HWI for each city/region is rescaled to equal 1 in 1956-1958. The treatment line is drawn as of January 1960.

Figure B2. The Help-Wanted Index in Miami, 1990-1999

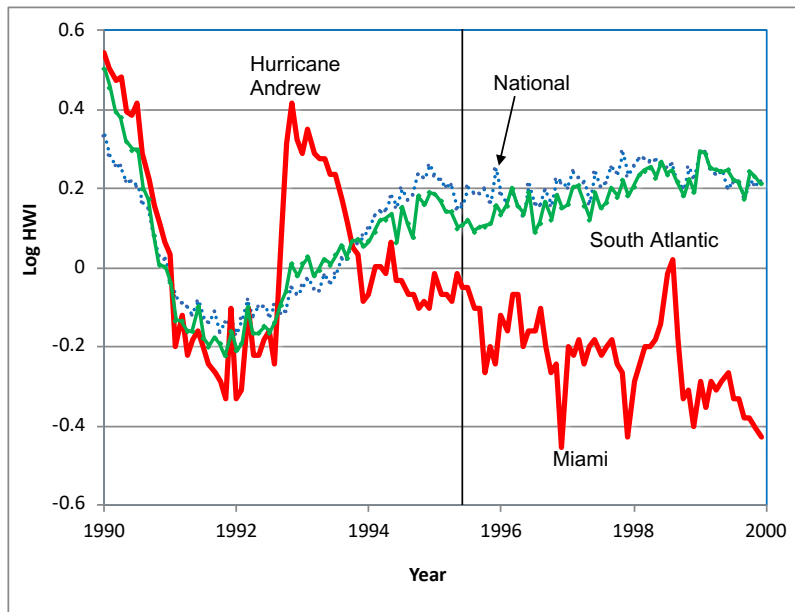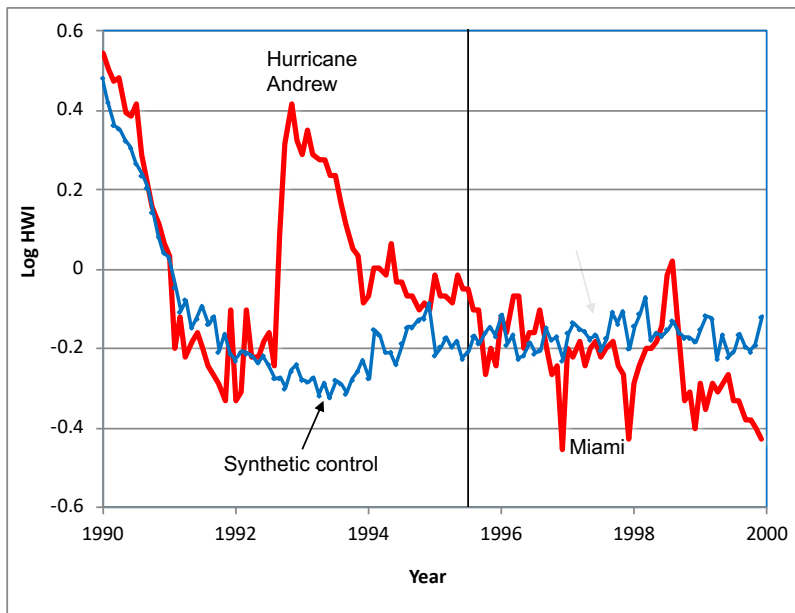### A. Relative to regional-national aggregates

B. Relative to the synthetic control

Notes: The HWI for each city/region is rescaled to equal 1 in January 1991-August 1992. The treatment line is drawn as of June 1995.

Table B1. Characteristics of other supply shocks of Cuban immigrants

Episode:	High school dropouts	High school graduates	Some college	College graduates	All workers
A. First wave, 1960-62					
No. of workers in Miami, 1960	172.1	111.0	43.1	35.5	361.8
No. of Cuban immigrants	21.7	15.5	7.7	15.0	59.9
Percent increase in supply	12.6	14.0	17.9	42.2	16.6
Education in occupation employing:					
Average native	10.2	11.6	12.2	14.6	11.7
Average Cuban immigrant	10.0	10.7	11.8	13.2	11.2
B. Guantanamo, 1994-1995					
No. of workers in Miami, 1990	246.9	222.9	260.8	193.7	932.3
No. of Cuban immigrant	13.5	11.5	4.5	6.8	36.2
Percent increase in supply	5.5	5.2	1.7	3.5	3.9
Education in occupation employing:					
Average native	12.1	12.8	13.4	15.2	13.6
Average Cuban immigrant	11.8	12.2	12.5	13.5	12.3

Notes: The pre-existing number of workers in Miami reported in Panel A is calculated from the 1960 census; the number of Cuban immigrants (at least 18 years old as of 1962) comes from the 1970 census; and a small adjustment is made because the 1970 census only identifies immigrants who arrived between 1960 and 1964. The pre-existing number of workers in Miami reported in Panel B is obtained from the 1990 census; the number of Cuban immigrants (at least 18 years old as of 1995) is obtained from the 2000 census.
.

Table B2. Difference-in-differences impact of the 1960-62 and 1994-1995 supply shocks (relative to synthetic control)

A. 1960-1962 supply shock		B. 1994-1995 supply shock	
Relative to 1954-1959	(1)	Relative to Jan. 1994-May 1995	(2)
1960-1961	-0.210	January 1990-August 1992	-0.044
	(0.021)		(0.039)
1962-1963	-0.270	September 1992-December 1993	0.366
	(0.014)		(0.046)
1964-1965	-0.218	June 1995-December 1997	-0.141
	(0.015)		(0.027)
1966-1969	-0.141	January 1998-December 1999	-0.215
	(0.015)		(0.030)
1970-1971	-0.250	---	---
	(0.020)		
1972-1974	-0.098	---	---
	(0.023)		

Notes for Panel A: Robust standard errors are reported in parentheses. The data consist of monthly observations for Miami and the synthetic control between 1954 and 1974. 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-shock period. The regression has 502 observations.

Notes for Panel B: Robust standard errors are reported in parentheses. The data consist of monthly observations for each city between 1990 and 1999. 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-shock period. The regression has 240 observations.

Appendix C: Supply Shocks and the Beveridge Curve in the Shimer Mismatch Model

Shimer (2007) proposes an alternative micro foundation for a Beveridge curve: Vacancies arise because of a "mismatch" between the cities where the unemployed are looking for jobs and the cities where the jobs are located (rather than the search-theoretic framework underlying the matching function). The mismatch hypothesis also predicts that a supply shock produces an inward shift in the Beveridge curve.

We reinterpret the mismatch hypothesis in terms of a single local labor market that has multiple skill-based sectors. The mismatch then arises because the talents that unemployed workers are offering are not suitable for the vacancies that the city's employers are advertising.[47] Following Shimer, let L denote the number of workers per sector and N denote the number of jobs per sector in the local labor market. Suppose further that the number of sectors is very large and that the actual number of workers and jobs across sectors varies according to a Poisson distribution. The fraction of sectors with exactly i workers and j jobs can be written as:

$$\pi(i; L) = \frac{e^{-L} L^i}{i!}, \tag{C1}$$

$$\pi(j; N) = \frac{e^{-N} N^j}{j!}. \tag{C2}$$

A sector where $i > j$ has $(i - j)$ unemployed workers; a sector where $j > i$ has $(j - i)$ vacancies. Summing these differences across all sectors (and assuming independence between the sectors where firms have vacancies and the sectors where workers are looking), the number of unemployed workers per sector and the number of job vacancies per sector are:

$$U(L, N) = \sum_{i=1}^{\infty} \sum_{j=0}^{i} (i - j)\, \pi(i; L)\, \pi(j; N), \tag{C3}$$

$$V(L, N) = \sum_{j=1}^{\infty} \sum_{i=0}^{j} (j - i)\, \pi(i; L)\, \pi(j; N). \tag{C4}$$

The aggregate unemployment rate u equals the ratio U/L, and the aggregate vacancy rate v is V/N. To simplify, suppose that the number of workers per sector, L, is exogenous. In this simplified version of Shimer's mismatch model, productivity shocks manifest themselves through variations in N, with a fixed level of L. The downward-sloping Beveridge curve is derived by choosing a fixed value of L, varying $N \in (0, \infty)$, and calculating the corresponding values of u and v that trace out the curve. For example, a productivity shock that increases N, holding L constant, raises the vacancy rate and lowers the unemployment rate (Shimer, 2007, Proposition 3). We wish to examine how the resulting Beveridge curve shifts when L increases exogenously due to a supply shock.

---

[47] Our reinterpretation is influenced by research on the sectoral shifts hypothesis; see Lilien (1982) and Abraham and Katz (1986). It is also related to the work of Lazear, Shaw, and Stanton (2018), who provide a conceptual framework for the derivation of a cross-sectional Beveridge curve, where vacancies are concentrated in occupations requiring high levels of skills while unemployment mainly targets low-skill workers.

Consider the initial Beveridge curve B in Figure C1, where $L = \ell > 0$, and an arbitrary point $p = (u^*, v^*)$ on that curve. A supply shock where $L = \ell + k$ (with $k > 0$), moves the labor market to a point on a different Beveridge curve, say B'.

We can derive the new Beveridge curve by varying N with L fixed at $\ell + k$. Let n be the value of N corresponding to the initial point (u^*, v^*) on curve B, and increase n by the same proportion as the increase in m, so $N = \frac{\ell+k}{\ell} n$. The new pair of values for L and N generates a point on the new Beveridge curve B'. Shimer (2007, footnote 7) shows that a proportional increase in L and N decreases both u and v, so that B and B' cannot be the same downward-sloping curve.

Let $p' = (u', v')$ be the point on B' resulting from the proportional increase in L and N, with $u' < u^*$ and $v' < v^*$. We can derive another point $p'' = (u^*, v'')$ on curve B' by decreasing N in order to raise u until it equals the initial u^* (with L fixed at $\ell + k$). This decrease in N must reduce v. Because $v'' < v' < v^*$, B' represents an inward shift of the Beveridge curve.[48]

Although the mismatch model predicts an inward shift in the Beveridge curve after a supply shock, the model does not predict the curve's eventual return to its original position. In the mismatch model, a larger labor market simply provides more opportunities for successful matches. Barring other changes that would affect the fundamentals (i.e., the number of workers per market or the number of jobs per market), there would be no reason for Miami's Beveridge curve to return to its original (relative) position.

---

[48] Put differently, an increase in M, holding the unemployment rate constant, reduces the vacancy rate. We are grateful to Jonathan Arnold for his assistance in outlining the proof.

Figure C1. A supply shock and the Beveridge curve in the mismatch model

Vacancy rate

$v*$

v'

v''

p

p'

p''

$B(\ell)$

$B(\ell + k)$

u' $u*$

Unemployment rate
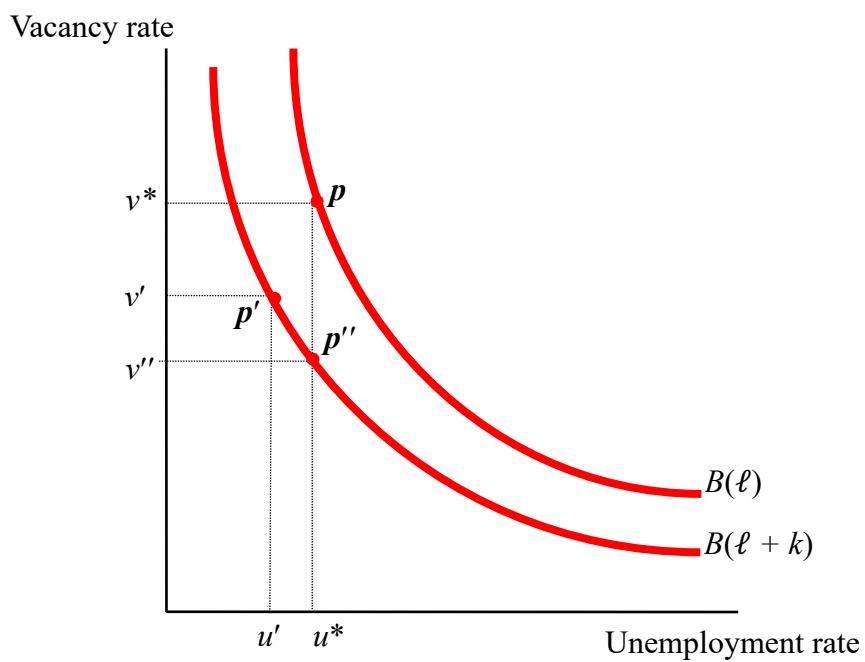
Appendix D. Construction of Samples of Help-Wanted Classifieds

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.[49] 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 D1 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 D2 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-

---

[49] 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.

We compared the annual trend in the number of ads in our random sample from the *Miami Herald* to the "population counts" reported in the HWI. To make this comparison, however, we need to make two minor adjustments. First, three months from the 1978-1984 period do not appear in our data as the Mechanical Turk workers were unable to transcribe text from the digital images of the *Herald*. We adjusted for the missing months by running a regression of the number of ads in our sample on the HWI with vectors of year and month fixed effects, and then using the prediction to "fill in" the three missing data points. Second, our sampling scheme implies that a small number of months appear twice in the sample. We adjusted for the "double counting" by using the average number of ads in the affected month to represent the monthly count. As the top panel of Figure D3 shows, the (adjusted) annual number of ads in our sample closely tracks the trend in the annual HWI. The correlation coefficient for the seven data points is 0.88.

B. Control Newspapers

As noted in the text, we obtained a file from 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
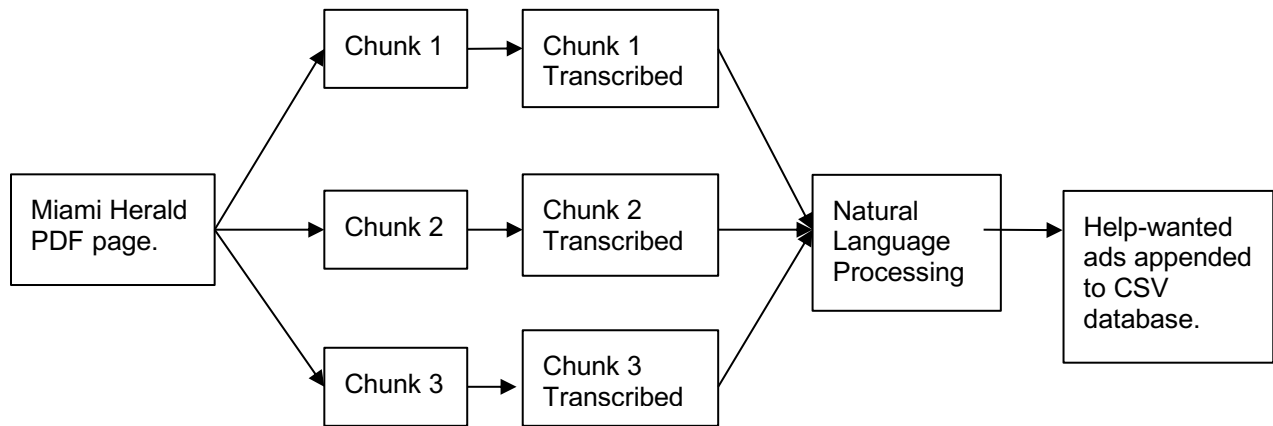
and bigrams, and the removal of stop words using the *quanteda* package in R (Benoit et al, 2018). The resulting text was then converted into a high-dimensional document-feature matrix that was used to estimate a series of topic models on the pages of each newspaper. Topic models are an unsupervised 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 D1 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 D1. Example outline of the process of transcribing
a single *Miami Herald* page of help-wanted ads**

```
                  ┌──────────┐     ┌──────────────┐
              ┌──▶│ Chunk 1  │───▶│ Chunk 1      │──┐
              │   └──────────┘     │ Transcribed  │  │
              │                    └──────────────┘  │
┌──────────────┐  ┌──────────┐     ┌──────────────┐  │  ┌──────────────┐     ┌────────────────┐
│ Miami Herald │─▶│ Chunk 2  │───▶│ Chunk 2      │──┼─▶│ Natural      │───▶│ Help-wanted    │
│ PDF page.    │  └──────────┘     │ Transcribed  │  │  │ Language     │     │ ads appended   │
└──────────────┘                   └──────────────┘  │  │ Processing   │     │ to CSV         │
              │   ┌──────────┐     ┌──────────────┐  │  └──────────────┘     │ database.      │
              └──▶│ Chunk 3  │───▶│ Chunk 3      │──┘                        └────────────────┘
                  └──────────┘     │ Transcribed  │
                                   └──────────────┘
```

Figure D2. 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

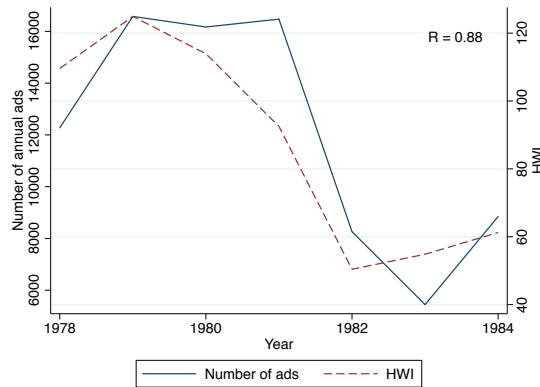salespersons exp better dresses top salary import style ctr

--**Step 4**---

UPLOAD your .txt file with the transcriptions to a Dropbox folder using this link*: https://www.dropbox.com/request/GnTR9D2c9OJ5HgLyXRJJ
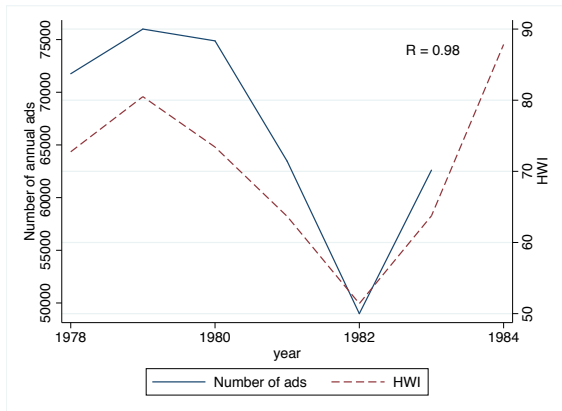
Figure D3. Trends in the local HWI and in the number of ads in sample
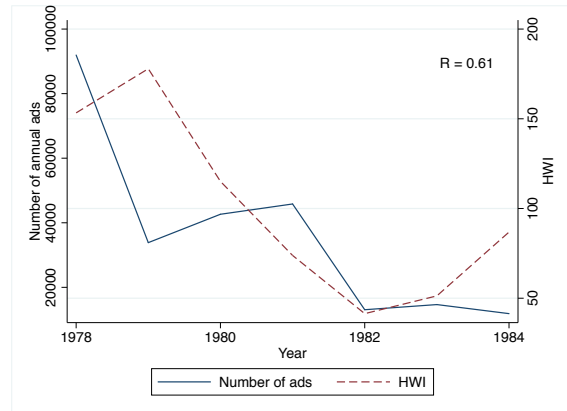
A. *Miami Herald*
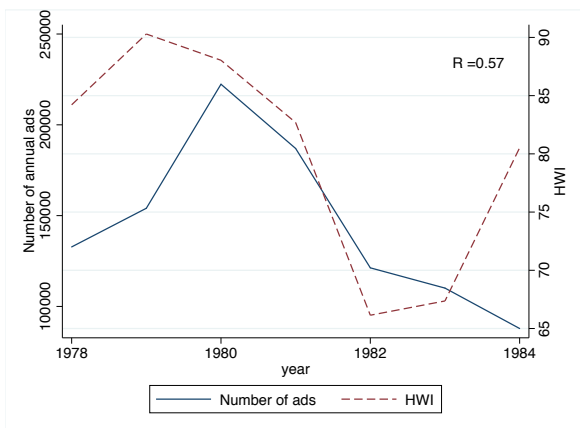

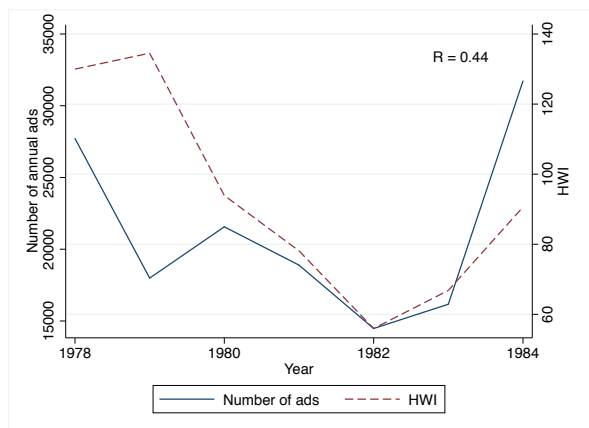## B. Control group

1. *Boston Globe*

2. *Minneapolis Star-Tribune*

3. *New York Times*

4. *St. Louis Post-Dispatch*

Notes: See Appendix C for details.

**Table D1. 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.