NBER WORKING PAPER SERIES

THE LABOR MARKET EFFECTS OF A REFUGEE WAVE: APPLYING THE SYNTHETIC CONTROL METHOD TO THE MARIEL BOATLIFT

Giovanni Peri
Vasil Yasenov

Working Paper 21801
http://www.nber.org/papers/w21801

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2015

We thank David Card, Colin Cameron, Thomas Lemieux, David Green, Michael Clemens, Pat Kline, Doug Miller, Joan Monras, Enrico Moretti, David Roodman, Shu Shen and participants in seminars at NBER winter meetings, University of British Columbia and Georgetown University for useful comments. We are grateful for the criticism contained in Borjas (2015b) to the first draft of this paper, as it has helped our current revision. We have not received any financial support for this paper. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Labor Market Effects of a Refugee Wave: Applying the Synthetic Control Method to the Mariel Boatlift

Giovanni Peri and Vasil Yasenov

NBER Working Paper No. 21801
December 2015, Revised May 2016
JEL No. J3,J61

ABSTRACT

We apply the synthetic control method to re-examine the wage and employment effect of the Mariel Boatlift, a large inflow of Cuban refugees to Miami in 1980. This method improves on previous studies by choosing a control group for Miami so as to best match its labor market features in the eight years before the Boatlift. Given the presence of significant measurement error for average city wages we emphasize the importance of using the May-ORG CPS sample rather than the March-CPS. The first includes a more reliable measure of weekly wages, has larger sample size and smaller measurement error. Analyzing wages and unemployment rates we find no significant departure between Miami and its control between 1980 and 1983. Using the March-CPS data, however, one could find negative wage effects in small subsamples after 1979 as pointed out in George Borjas (2015a). However those estimates are imprecise and very sensitive to the choice of sample and of the outcome variable.

Giovanni Peri
Department of Economics
University of California, Davis
One Shields Avenue
Davis, CA 95616
and NBER
gperi@ucdavis.edu

Vasil Yasenov
Department of Economics
University of California, Davis
One Shields Avenue
Davis CA, 95616
viyasenov@ucdavis.edu

1 Introduction

The Syrian refugee crisis that reached Europe during the Summer of 2015 has once again ignited the discussion about immigration and its economic effects. It has also rekindled the interest in looking at the past history of refugee waves to learn from these. How did receiving countries absorb sudden waves of immigrants? What were their immediate effects on wages and employment? How long did these effects last? Over the past three decades the United States has experienced large and slow inflows of millions of less educated (many of them undocumented) as well as highly educated immigrant workers. The skills and labor they supply has been absorbed by the United States economy that has adjusted in terms of specialization, capital and technology (see Peri and Sparber 2009 and Lewis, 2011). Immigrants have contributed to GDP growth, to the variety of skills available and their long run economic impact has been studied in several papers (e.g. Ottaviano and Peri, 2012). In terms of impact and economic effects, these large and unevenly distributed inflows of immigrants since the 1980s are a very important phenomenon to analyze.

A single episode, however, has held a special place in the minds of the American economists as it has been an unique example of sudden, unexpected and large refugee inflow on American soil. On April 20, 1980, Fidel Castro publicly announced he would open the ports of Mariel, 25 miles away from Havana, enabling anyone who wanted to leave the country to freely do so. Consequently, between April and September of the same year, almost 125,000 Cubans fled to the United States shores in what is known as "the Mariel Boatlift". The majority of them settled in Miami, increasing its labor force by about 8%. Most of these immigrants were low-skilled. This event provides an unique quasi-experimental environment where labor economics theories can be tested. More specifically, the nature of this sudden inflow guarantees high potential for short-run consequences on wages in the Miami local labor markets if other factors (technology, efficiency, physical capital) did not respond immediately. The Mariel Boatlift was a rather unique "natural experiment" for the United States. Several studies have examined similar settings in Europe (see Hunt (1992), Carrington and De Lima (1994) and Freidberg (2001) among others).

An early study by David Card (1990) analyzed the Mariel Boatlift and his results showed that the impact on employment and wages of low skilled non-Cubans in Miami was insignificant. This made the Boatlift a prominent example of how the predictions

of the simplistic canonical model of labor demand and labor supply do not work well in analyzing the consequences of immigration even in the short run[1]. Moreover, from a methodological point of view, the experimental design of Card's paper has profoundly influenced the direction of research in the field (see Angrist and Pischke 2010). Prominent textbooks in Labor Economics, both at the undergraduate (Borjas 2012, Laing 2011) and at the graduate levels (Cahuc et al. 2014), still use the Mariel Boatlift study by Card (1990) to illustrate the "difference-in-differences" empirical method.

Given the importance of this historical episode and of the study by David Card, one reason to revisit it is that since 1990 we have improved our methodological toolbox. The "Synthetic Control Method" (SCM), an econometric technique developed and used in a series of papers by Alberto Abadie and coauthors (Abadie and Gardeazabal 20003; Abadie, Diamond and Heinmueller 2010, 2015) is better suited than the original approach for addressing this type of case-studies. The SCM is based on the idea that a linear combination of labor markets is a better control group than any single one. Hence one should weight all available United States cities by minimizing the pre–1979 difference with Miami for a set of relevant labor market characteristics and create a single "synthetic" city which will serve as the control group. Consequently, this synthetic city will constitute a labor market, as similar to Miami as possible, that did not experience a large immigrant arrival. Relative to the analysis of Card (1990) this method has several advantages. First, this formalized procedure reduces the "ad-hoc" nature of choosing the control group. Second, it allows to validate the quality of the control group by checking the pre-treatment differences between the outcome variable in the treated and in the synthetic units. Finally, by applying this method to each city and simulating a distribution of effects, we can construct a p-value for how significant is the post-treatment difference for Miami relative to the whole distribution. This produces a confidence level for our inference that accounts for idiosyncratic variation in the data. All these are important improvements on Card (1990) whose standard errors were incorrect as they only accounted for sampling error in wage measurement. Accounting more formally for the uncertainty introduced by idiosyncratic

[1]Several subsequent papers suggested how different channels for absorbing the Mariel Cubans might have worked, rationalizing the results within richer models. Lewis (2004) showed that less skilled Cubans were absorbed by industries that chose more "unskilled-intensive" technology and less automation. In addition, Bodvardson et al. (2008) argued that the immigrants increased significantly local demand for services, and hence also labor demand and not only labor supply. Monras (2015) revisits Card (1990) analysis of this immigration shock and confirms the results of no significant wage effects on less educated natives when using the Merged Outgoing Rotation Group (ORG-CPS).

factors is also crucial in addressing some of the criticism moved to Card (1990) by later studies such as Angrist and Krueger (1999).

In revisiting this episode we also shine light on how important the choice of the dataset and sample is in this case due to the small size and significant measurement error for variables calculated at the metropolitan area level. With the goal of comparing both data sources, we develop a way to quantify the measurement error in average wages in the ORG-CPS and March-CPS datasets. We find that the measurement error in March-CPS is so large that differences of 15-30% in average city wages can arise purely from it. We also review the literature (in particular Lemieux 2009) that emphasize the imprecise nature of March-CPS data when measuring wages, especially for workers who are paid by the hour. Overall, we find ORG-CPS to be superior in this sense. Furthermore, to help alleviate the measurement error problem we consider the largest possible group likely to be negatively affected in its labor market outcomes by Mariel immigrants. Namely, these are Non-Cuban workers, with *no high school degree* between 19 and 65 years of age, not self employed and in the labor force. As labor market outcomes we consider log wages (annual, weekly and hourly) and unemployment rates of this group, relative to the control.

Our results show no significant difference in the post-1979 labor market outcomes of high school dropouts between Miami and Synthetic Miami. Neither wages (annual, weekly or hourly) nor unemployment of high school dropouts differ significantly between Miami and the control between 1980 and 1983. Similarly, considering wages in the bottom 15th or 20th percentile of the distribution for non-Cuban workers shows no significant departure. We run "difference-in-differences" type of regressions for Miami and the Synthetic control to show the lack of statistically significant differences. Again, we do not find any systematic deviation post 1979 and the point estimates of the Miami-Synthetic control differences in wages are usually *positive*. The statistical inference using all the possible permutations of other 43 cities also show that the change in wages and unemployment of Miami high school dropouts relative to the control group in 1979-1982 was within the distribution of other cities' idiosyncratic variation. In some specifications the wage deviation was close to the limit of the existing range, but in the *positive direction*. Our method, therefore, confirms the early results of Card (1990).

Then we focus on understanding the different results presented in Borjas (2015a). Replicating his results we find that a large negative deviation of wages of high school dropouts in Miami arise *only when using the March CPS data*, and such deviation is

significant *only in the sub-sample obtained by eliminating women, non-Cuban Hispanics and selecting a short age range (25-59 years old)* among high school dropouts. This very drastic choices leaves the sample in Miami as small as 15-20 observations per year. Measurement error in average Miami wages can be in the order of 20-30% for such sample. Moreover, the restrictions are, in our view, very problematic. They eliminate groups of workers on which the effect of Mariel should have been particularly strong. Non-Cuban Hispanics, many of whom are United States born should be included as potentially affected group. Similarly young workers, age 19 to 24 can be inexperienced and vulnerable to competition. If one is worried of the different national trends for these groups there are ways to address those (as we show in section 5) without dropping so many observations. We show that by adjusting Borjas' sample in minor ways, to include these sub-groups of high school dropouts or looking at alternative outcomes (yearly wages, 15th percentile wage, employment), the post 1979 Miami-control differences vary widely (from negative to zero to positive) when using March-CPS data. To the contrary these changes do not affect much the finding in the May-ORG data confirming their larger precision and reliability. We finally revisit the "1994 Mariel Boatlift That Did Not Happen" analyzed by Angrist and Krueger (1999) also using the synthetic control method.

The rest of the paper is organized as follows. In Section 2 we begin by describing the data and the relevant group of workers that should be analyzed. We also present the timing of the events and the characteristics of the Mariel Cubans and we make a case for strongly preferring the ORG-CPS data rather than the March-CPS because of its larger size, smaller measurement error and more precise measurement of wages. We continue in section 3 where we calculate, as a benchmark, the largest short-run wage effect of the Mariel Boatlift predicted by a very simple two-skill model without adjustment of capital or technology. Section 4 briefly describes the Synthetic Control Method and it discusses the period and variables matched. Section 5 contains our main results. Then, in section 6 we consider the small and noisy March-CPS sample and we account for what determines the large negative wage effects estimated in Borjas (2015a, 2015b). Next, in section 7 we reconcile the odd result of some apparent labor market effects in 1994 with no Boatlift happening, pointed out by Angrist and Krueger (1999). Finally, section 8 concludes the paper.

2 Measuring the Boatlift and Miami Labor Markets

2.1 Number and Demographics of the Mariel Cubans

In order to identify the workers who were most likely to be affected by the Mariel immigrants we show in Table 1 the aggregate numbers and summary statistics of demographic characteristics of immigrants and of Miami workers. In the first column we show data relative to the labor force in Miami as of 1980. In the second column we present the data on all Mariel Cubans, identified in the Census 1990 as Cubans who arrived in the United States in 1980 and were 19 or older at arrival. The last column shows the number and demographic characteristics of Mariel Cubans from Census 1990 who were still living in Miami as of 1990, and hence likely settled there.

Limiting our attention to individuals between 19 and 65 years old upon arrival, and assuming that the percentage arriving in Miami was equal to the percentage of those Cubans still in Miami in 1990, we obtain a total of 54,196 working-age Mariel Cubans. Our calculations shown in Table 1 say that 56% of them lacked a high school degree. Furthermore, 62% of them were in still in Miami as of 1990. Hence either at arrival or in the successive years 40% of them located in other places. As we show below there is some evidence in the (noisy) CPS data for Miami that some of the Cubans who arrived in 1980 might have left the city in the following 2-3 years. The share of Cubans in Miami, in fact, peaks in 1981 and then declines between 1981 and 1985[2].

Overall, the Mariel Boatlift produced an 18% increase in the number of high school dropouts in the Miami labor market, while for the other education groups the increase was only 5% and for the total labor force it was 8.4%. As this change took place in few months it was certainly exceptional. The most significant change analyzed in other "quasi experiments" literature is the inflow of Russians to Israel (Friedberg 2001) which was equal to 12% of initial population but took place over 5 years (between 1989 and 1994). Looking at other demographic characteristics, namely gender and share of young individuals, we see that those are similar in the Mariel population and in the 1980 Miami labor force. This is true even conditional on high school dropouts. The change in supply for high school dropouts due to Mariel Cubans seems well balanced between genders and age groups. Hence, in order to minimize the measurement error (maximize the sample

[2]The share of Cubans in the Miami Labor force follows a very similar pattern.

size) it seems reasonable to pool all high school dropouts, male and female in the age range 19 to 65 together. As we want to identify the impact on existing Miami workers we exclude those who self-identified as Cubans (from the question on Hispanic origin in the CPS)[3].

In summary, our preferred sample of workers consists of non-Cuban high-school dropouts, age 19-65, not self-employed individuals in the labor force, with non-missing earnings and with positive sample weight. This choice ensures that the workers are not attending school. Consequently, unless otherwise noted, all the graphs below will use this sample. The next step in analyzing the impacts of the Mariel Boatlift is the choice of a dataset.

2.2 Sample Size and Measurement Error: March-CPS and May-ORG CPS

Table 2 shows the number of observations for Miami workers with these characteristics in the main data sources. Using the March-CPS (first column) our sample includes only 60-80 observations for per year. This could certainly raise concerns of significant measurement error. The May-ORG CPS, instead, is relatively small between 1973 and 1978 (comparable to the March CPS), but beginning in 1979 it consists of usually around 150 observations. While still not too large, these numbers of observations are closer to comfort. In the last two columns of Table 2 we show the number of observations included in Borjas (2015a) by using the sample restricted to male, non-Hispanic between 25 and 59 years old. One can notice a stark difference in the numbers, that decrease for the March-CPS data to 15-24 observations in the relevant period, a point to which we will come back to in Section 6. Next, let us characterize more precisely the measurement error and discuss potential problems of using the March-CPS.

Besides the larger sample size of ORG-CPS, there are several additional reasons we believe it to be superior to the March-CPS for our purpose. First, the variation of measurement error in average wage across cities in the March-CPS sample is much larger than in the May-ORG CPS sample. One way to show this is to assume that, for year 1979, the Census 1980 allows to calculate the "correct" average city wages[4]. We then compare these

[3]The numbers presented in this section are in accordance with previous studies including Card (1990) and Borjas (2015a).

[4]This is a reasonable assumption as the Census uses a 5% sample of the population and it includes

with the ones estimated from March-CPS data (from 1980) and the ORG-CPS data (from 1979). We calculate average log weekly wages for our preferred sample of non-Cuban high school dropouts in the age range 19-65, in each of the 41 metropolitan areas available in all three datasets. Then, we calculate the difference of these average log wage in each metropolitan area, between the March-CPS and the Census and we do the same between the ORG-CPS and Census. We call these deviations "measurement error".

Figure 1 shows the distribution (kernel density) of this measurement error (deviation from Census) using the March-CPS (grey) or the May-ORG CPS (black). It is evident from the picture that the variance of the measurement error is much larger in the March-CPS which implies that two cities' average log wages may differ from each other by a very significant amount in this sample, simply because of it. The measurement error in the March-CPS has a standard deviation of 0.12 logarithmic points (about 12 %) while the one in ORG-CPS, has a standard deviation of only 0.07 log points (7%). Consequently, in the March-CPS, a difference in average wages in the order of 15% between two cities could easily arise by pure error[5]. As long as the reasonable wage differences to be identified are smaller than 15% the noise of the March-CPS sample is simply too large to have any power to identify such an effect[6]. While the ORG-CPS sample also displays a non-negligible variation of measurement errors its standard deviation is about half of that in the March-CPS.

A second reason to be skeptical about the use of the March-CPS in measuring weekly wages of low skilled individuals was illustrated by Lemieux (2006). He showed that March-CPS wage data, based on the recollection of previous year annual salary compound a recall and a division error that are particularly severe for people who are paid by the hour (which includes a large fraction of the high school dropouts). People may have a hard time recalling the actual number of hours worked last year and then in calculating their annual wage income from those. To the contrary, the May-ORG CPS sample based on weekly wage recall from the last week of work produces a less noisy and more reliable estimate of earnings for individuals who are paid by the hour. Both Lemieux (2006) and

thousands of observations for each metro area.

[5]On average, the difference between measurement errors for two randomly chosen cities will be 1.12*standard deviation measured in the distribution.

[6]In Figure 1 one also notices a negative bias of the ORG-CPS and a positive bias of CPS (average of errors is not 0) data. If one chooses the ORG-CPS data from 1980 (rather than 1979) assuming that people may report in Census and CPS a yearly wage more similar to what received in 1980, the average of the error in the ORG-CPS becomes almost 0 but the standard deviation (that matters to us) remains the same.

Bollinger (1998) illustrate the severity of this measurement error and how it could affect the estimate of wage dispersion at the national level. Its impact can only be magnified for the very small sample at the metro area level. A third argument to prefer May-ORG CPS is the fact that March-CPS makes available fewer cities to be included in the control group (31) relative to the May-ORG CPS (44). All else equal, this reduces the probability of finding a good control group for Miami pre-1979 even when using the Synthetic control method. This poorer quality of pre-1979 match when using the March-CPS data causes several specifications to fail some falsification tests (presented in Figure 9C). The same variables calculated on the May-ORG CPS sample pass this test, (see Figure 5C).

In summary, we are convinced that May-ORG CPS is superior to March-CPS and that, when analyzing labor market outcomes in metro areas during the 1980's, the March-CPS data should be avoided. However, we will use the latter sample in some robustness checks because only using this sample we can replicate and explain the results of Borjas (2015a) and (2015b)[7].

2.3 Measured Labor Supply Shift in Miami

Identifying the exact conditions right before 1980, and right after that, helps to maximize the chance of identifying the largest possible short-run effect. We are considering a one-time, unexpected shift in supply that took place between March and September 1980, and not a persistent policy change. The adjustment dynamics, then, would determine its persistence after 1981, however the bulk of the effect should be detected in 1980 and 1981. We adopt the convention in all figures, to call "1979-Pre" the data relative to the last observation before, and as close as possible to, the Mariel Boatlift. This is usually year 1979[8]. We call "1980" the data relative to the period during the Boatlift and we call "1981-Post" the first data point after all Mariel Cubans had arrived. We also take the convention, in each figure, of showing a vertical bar exactly at the *last "pre-shock period"* (hence on "1979-Pre"). This notation helps to visually identify the last period of the status quo, right before the shock. To the immediate right of the bar we can see the impact of the sudden shock. To its left we can see the trend and variation during the pre-treatment period.

[7] Both Card (1990) and Angrist and Krueger (1999) use the ORG-CPS sample in their analysis.

[8] For wage and unemployment in the March-CPS we use data collected in 1980, which is relative to the previous year. For the ORG-CPS we use data collected in year 1979.

Figure 2 shows the share of Cubans in Miami's population, age 19-65 (dark lines) and in the population with no high school degree age 19-65 (lighter line), between 1973 and 1985. Panel A shows the March-CPS data, while Panel B uses the May-ORG CPS. Notice that for the March-CPS, as the data on demographics are relative to the month of March-CPS, the last pre-treatment observation is the one collected in March-CPS 1980, and it is called ("1979- Pre"), and it is differentiated from the 1979 (March-CPS 1979). The 1981-Post is the observation for March-CPS 1981, while the "1980-shock" is simply the linear interpolation of 1979 and 1981[9]. For Panel B, the year corresponds to when the ORG-CPS survey (and before 1979 the May-CPS survey) was done. Paying attention to the pre-post details around 1980 allows us to align our data precisely around the Miami Boatlift shock. Identifying a clear jump upwards of the Cuban share from " Pre" to "Post" would be the "mark" of the Boatlift on the CPS data.

Three facts emerge from Figure 2. First, the March-CPS and May-ORG CPS data show Cuban shares close to each other for the total population but they are more noisy and less consistent with each other for high school dropouts. We see significant noise, especially before 1980, when the samples were rather small[10]. Second, both lines and samples show a significant increase between "1979-Pre" and "1981-Post". Compared to trends and year to year movements before and after, the 1979-1981 increase does not seem particularly large, however. Considering May-ORG CPS figures, the 1979-1981 increase as percentage of the population equals about 6 points and as percentage of high school dropouts the increase was around 12 points. These percentages are consistent with the ones obtained from the Census and described in section 2.1. Third, after the increase in the share of Cubans between 1979 and 1981, in the following 4 years that share decreased according to all the samples and especially as share of high school dropouts[11]. This could be because some of them left the city or because more non-Cuban immigrants arrived. In 1985 Cubans as share of high school dropouts seem to be back at percentages comparable to those of 1979. This emphasizes strongly the temporary nature of the shock that happened suddenly in 1980 and was absorbed so as to be completely indiscernible in

[9] This is done only for this graph, due to the timing of the March-CPS enumeration, that in 1980 was just before the Boatlift. The wage and employment data, however, are relative to the whole year and will be attributed to the relevant year in the other graphs.

[10] Notice an isolated spike in the share of Cubans among dropouts in 1978, only shown in the May CPS.

[11] The sudden 1979-81 increase, the peak reached in 1981 and the following decline are also features of the Cuban share of the *labor force* in Miami (rather than population, as represented here). The full inflow of Mariel people into the Miami labor market happened in the 1979-81 period.

the CPS data by 1985.

3 Predicted Short-run Effects with no Adjustment

The time profile of the shock in Figure 2 suggests that the strongest impact on the Miami labor market should be between 1980 and 1981, with not much residual impact as of 1985. Knowing the magnitude of the labor supply shift due to the Mariel Boatlift allows us to calculate the short-run prediction of a simple model with everything else fixed (capital, efficiency and productivity). Recent papers surveyed in Peri and Lewis 2015 have emphasized the importance of several margins of adjustment affecting capital, technology, efficiency, specialization in response to immigration and explaining the small effects on wages. Other studies (e.g. Ottaviano and Peri 2012) have indicated that imperfect substitution and heterogeneity between natives and immigrants can further attenuate wage competition effects. Here, it is useful to calculate what would be the magnitude of the negative wage effect predicted by a simple skilled-unskilled labor model with fixed capital and technology and with perfect substitutability between immigrants and natives. This effect is *the most negative* that a reasonable partial equilibrium model can predict and it will serve as a benchmark for our estimates. It will also help us determine if our empirical analysis will be likely to detect such an effect, given the size of the measurement errors described in section 2.2. We use a simplified version of the model in Borjas (2003), assuming a production function for the city of Miami where output Y is a function of total factor productivity, A, physical capital K, and a labor aggregate, N made of workers with no diploma, L and workers with high school diploma H as follows:

$$Y = AK^{\alpha}N^{1-\alpha} \text{ where } N = \left(\theta_H H^{\frac{\sigma-1}{\sigma}} + (1-\theta_H)L^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \quad (1)$$

The parameter α is the output elasticity to physical capital, σ denotes the elasticity of substitution between skilled and unskilled workers and θ_H captures the skill-bias of technology. In this exercise we consider physical capital K, total factor productivity A and skill-biased technology θ_H as fixed in the short run. We then calculate the effect of an increase in aggregate labor, N, by 8% and in the supply of L (high school dropouts) by 18% which correspond to the changes in Miami labor supply due to the Mariel Boatlift

as measured in section 2.1[12]. One can easily obtain the short-run (partial) wage effect from the model above, considering that the wage of L workers equals their marginal productivity and differentiating it with respect to percentage (logarithmic) change in N and L. The expression is as follows:

$$\frac{\Delta w_L}{w_L} = -\alpha \frac{\Delta N}{N} + \frac{1}{\sigma}\left(\frac{\Delta N}{N} - \frac{\Delta L}{L}\right)$$

Considering the commonly used value for the parameter α (the share of income to physical capital) of 0.33 and choosing for σ, the elasticity of substitution between H and L, a value equal to 2, (in the range estimated by Katz and Murphy (1992), Angrist (1995), Johnson (1997), and Krusell et al. (2000)) it is easy to calculate that the predicted "short-run" effect is equal to -7.6% This value is certainly as large (in absolute value) an effect as one can expect as it implies no adjustment, perfect native-immigrant substitution and uses the largest estimate of supply change due to Mariel immigrants. The model highlighted above says that a negative wage effect on high school dropouts in the order of 7-8% in year 1981 is the largest negative impact that a model can predict. Such an effect should also be strongest in 1981-1982, right after the inflow, and dissipate by 1985 when the share of Cubans was back to pre-1989 levels. This is an upper bound against which we can evaluate the estimated effects. This calculation makes also clear that using the March-CPS data, in which 10-15% wage differences across metropolitan areas are due to the *average* measurement error, would deliver no power at all to detect any reasonably sized effect. So in the light of the largest magnitude of effects that one can plausibly expect it would seem indispensable to avoid using the March CPS.

4 The Synthetic Control Method

The Synthetic control method, first introduced by Abadie and Gardeazabal (2003) and then further developed in Abadie et al. (2010) provides a systematic way to analyze the impact of an event in case-studies such as the Mariel Boatlift. Typically, in these settings a single unit (metropolitan area, state or a country), experiences the event ("treatment") while the others do not. In order to evaluate whether this had an impact on some outcomes

[12] Using the data in Figure 2 the supply changes in 1979-81 would rather be 6% for N and 12% for L so the values chosen in our exercise are at the upper end of the likely range.

in the treated unit, relative to what that outcome would have been in absence of the treatment, the method identifies a control group (called the Synthetic control). To fix ideas, we consider $J+1$ metropolitan areas indexed by $j = 0, 1, 2...J$ and denote Miami as 0 while we call the group of all 43 other cities "the donor pool". Then define a vector G_0 of dimension $k \times 1$ whose elements are equal to the values of variables that help predict the wages of high school dropouts including the values of the outcome variable itself for the city of Miami, in the years from 1972 to 1979. Then we define, similarly, a $k \times J$ Matrix, G_J in which row j is the sequence of values for the same variables and years relative to city j in the "donor pool". The Synthetic control method identifies the vector of weights W^* that produce a convex combination of variables in cities in the donor pool, G_J, so as to approximate as close as possible, in terms of quadratic error, the pre-treatment vector of variables chosen for metro area 0, G_0[13]. Once we have identified $W*$ we can use those weights to calculate the post-treatment outcome variables for the "Synthetic control". Comparing the pre-post 1979 change in the outcome variable for Miami, relative to the pre-post change for the Synthetic control is the basis to evaluate if the treatment has had any effect on Miami. As there is discretion in choosing what variables to match in the pre-treatment period, it is important to validate the choice of the control group. To do so we can check the pre-intervention (1972-1979) levels and trends of the outcome variable to see how closely the treated unit and the Synthetic control group track each other before the event. Large differences in the pre-treatment path between treated and Synthetic control would cast doubts on the validity of the chosen group as control. As a more formal test we will also check in a regression environment whether the pre-1979 and post-1979 difference between Miami and the Synthetic control are statistically significant. We tried several different combinations of variables to be included in the pre-treatment distance minimization. We finally selected variables that capture important features of the low skilled labor market, namely the share of dropouts, the share of Hispanics and the share of manufacturing workers in the labor force, besides also including the outcome variable for some pre-1979 years. When using the March-CPS we included the 1972-1979 period and when using the May-ORG - the 1973-79 period, to allow for a reasonably long pre-

[13]In our estimation the quadratic form to be minimized includes V is a $k \times k$ diagonal, positive definite matrix that determines the weight for the contribution of each element of the vector in the objective function. We use STATA's default option for the matrix V which is chosen among all diagonal and positive definite matrices to minimize the average squared prediction error of the outcome variable during the pre-shock period.

treatment[14]. The footnotes in Figures 3 to 12 list the cities that constitute the Synthetic control for each figure and the weights that each of them is given by the Synthetic control method. Typically, between two and four cities are chosen to have positive weight. Most results we show are robust to (small) variations in the selection of the Synthetic control. In Appendix Figure A1 we also show the analysis keeping the metro areas and weights in the Synthetic control fixed to the group that is identified to best match the log weekly wages of high school dropouts (1972-1979) between Miami and Synthetic control.

5 Empirical Estimates

5.1 Main Results

We begin with showing, in Figure 3, the main results from the Synthetic control method using May-ORG CPS data. The four panels of this figure tell the basic story that will be then confirmed, time and again, in all the robustness checks and in the subgroups. Panel A shows log weekly wages (gross wage and salary) for Miami (solid line) and for the "Synthetic Miami" (dashed line). The footnote indicates which cities enter in the Synthetic control and their associated weights. Panel C, shows the same figure for log hourly earning as outcome. These are calculated dividing the weekly earning by reported hours worked per week. One may argue that this measure is closer to capturing the marginal productivity (hence price) of labor. Panels B and D, show the same variables and the same sample of workers except that they include the group of 16-18 years old (who are still potentially in school) and exclude the older group, focusing on the 16-61 age interval as done in Card (1990). While the sample used in Panels B and D is larger, some of the younger workers may still be in school which may pollute the estimated wage effects if we think this group is very different from the rest (Borjas 2015b). The only discernible effect of including this group is decreasing the year to year noise of the data a bit, but the main features of the figures are unchanged. Let us first comment on the pre-1980 time path for the considered wage outcomes in Miami and in the Synthetic control. A

[14]Most studies (including, for instance, Abadie et al. 2010, Bohn, Lofstrom and Raphael 2014) use at least ten pre-treatment years and have less noisy data. We do the best we can, given our data constraints, by extending the analysis back to 1972 (or 1973). Before that year the number of Metropolitan areas sampled in CPS was simply too small.

reasonably good match of the pre-1980 trend is needed to consider the Synthetic control as a good placebo group. Overall the Synthetic controls in Panel A-D do a good job in tracking the dropout wages and matching the 1972-1979 Miami trend in each of the Panels. Still, deviations between Miami and Synthetic control in the order of 0.01 to 0.05 logarithmic points are common. This implies that such level of noise could make it hard to identify deviations of average wages between Miami and the Control in the order of 1 to 4%. Nevertheless, in spite of the noise, we should be able to discern if differences (in the order of 7-8%) suddenly arise between Miami and its control in the aftermath of the Cuban inflow between 1979 and 1981. Importantly all graphs show a clear long-term downward trend for wages of high school dropouts in Miami and in the Synthetic control since 1972 (and perhaps earlier, but we do not go further back) all the way to 1991. Matching this preexisting time-trend between Miami and Synthetic control is crucial to claim that we have identified a good placebo and it is an improvement on Card (1990). The 1980s were a period of large increase in wage inequality and poor performance of the wages of unskilled workers. Hence identifying these features as common to Miami and the control is important.

The key takeaway from the four panels is that average wage of high school dropouts does not show any negative break or jump in 1980-1981 for Miami relative to the Synthetic control. If anything, especially in the Panel A and C (from the sample without individuals still potentially in school), the Miami wage rises slightly above that in the "Synthetic control" between 1979 and 1981, and the post-1979 wage trend in Miami seems less negatively sloped than the post 1979 trend for the Synthetic control. In fact, inspecting all panels and the whole post-1979 period there seem to be two instances in which the difference between Miami and Synthetic control is not negligible. They take place in the periods 1985-1989 and show deviations of Miami *above* its Synthetic control for weekly and hourly wages. However, the fact that such deviations occurs five or more years after the Boatlift, that the year-to-year variation is large, that after 1985 the Cuban share had returned to pre-1980 levels implies, in our view, that they have absolutely nothing to do with Mariel inflow.

5.2 Controlling for National Trends

Average wages of different demographic groups such as women and men, young and old, Hispanic and non-Hispanic workers had different national trends in the 1970s and 1980s. Depending on the demographic composition of cities, these trends can affect them differentially, introducing some confounding factors in the analysis. We adjust our wage measures to account for this potential bias. A method commonly used in the literature to control for labor force demographic characteristics of different labor markets, and to reduce the potential confounding effect of age, gender and ethnic heterogeneity across them is to first "adjust" individual log wages by running the following regression:

$$\begin{aligned} \ln w_{ijt} &= \beta + \beta_{AGE} * (Age\_Group)_{it} * (Year)_t + \\ &\quad + \beta_{female} * (Female)_{it} * (Year)_t + \beta_{Hisp} * (Hisp)_{it} * (year)_t + \varepsilon_{ijt} \end{aligned} \tag{2}$$

The wage of (high school dropout) individual i in metro area j in year t is regressed on a set of five-year age dummies $(Age\_Group)_{it}$ interacted with year dummies, $(Year)_t$, on a female dummy, $(Female)_{it}$ interacted with year dummies and on a Hispanic dummy, $(Hisp)_{it}$ interacted with year dummies. This produces the residual ε_{ijt} that captures individual (and city) log wage variation once those aggregate trends are accounted for. We then implement the Synthetic control method on those residuals, averaged by city. The resulting graphs, for choices of sample and of wage definitions analogous to those of Panel A-D in Figure 3, are reported in Figure 4. Each panel shows very significant noise, in the form of year to year fluctuations both before and after 1979 and the match of Miami and Synthetic control before 1979 is not very accurate. This is reasonable as we are now trying to match a residualized wage that is a noisier measure than the average wage. However, if we align the 1979 observations for Miami and its Synthetic control, and we look for the "difference in difference" effect of the Boatlift, we always observe a positive deviation of Miami for the year 1980 and 1981 and small remaining deviation in 1982 in each panel. Once again, the post 1979 period does not show any negative deviation between Miami and control and, other than increasing variability, there is no other clear effect of using the residualized wage.

5.3 Additional Labor Market Outcomes

The effect of unskilled immigrants on native labor markets can be particularly strong at the low end of the wage distribution. While certainly workers with no high school degree belong to this segment of the labor market, some economists (e.g. Dustmann, Frattini and Preston (2013)) have argued that the best way to capture this competition effect is to look directly at natives in the bottom percentiles of the wage distribution. Panels A and B of Figure 5 show the Synthetic control results considering as outcome the wage at the 15th or at the 20th percentile of the wage distribution of natives. One advantage in choosing the wage percentile, rather than the average wage of a small group (such as the high school dropouts), is that the sample used is larger (the whole non-Cuban labor force) in Miami and this statistic should be less sensitive to extreme values of wages in the city and hence somewhat less volatile. This is reflected in a somewhat smaller year-to-year volatility, observed in Panels A and B (especially before 1980) relative to the Panels of Figure 3 and in an slightly improved match of the pre-1980 trend between Miami and the Synthetic control. In this case the during the 1979-1981 period both Panel A and B show no deviation of the Miami wage relative to the Synthetic Miami, with small positive differences between Miami and control in the 1981-84 period. This is true both for the wage at the 15th percentile (shown in Panel A) as well as for that at the 20th percentile (shown in Panel B). As a falsification check we also show, in Panel C, the time evolution of the 90th native wage percentile in Miami and Synthetic control. The theory would argue that this group of workers does not experience any effect or possibly a somewhat positive effect from complementarity. Aligning the observations in 1979 for Miami and Synthetic control we see that the 90th native wage percentile in Miami had a small positive deviation from control after 1979 and not much deviation after 1981. This is similar to the behavior of the bottom percentiles, suggesting that some of the adjustment mechanisms must have been in place as not even the wage at the top relative to the bottom percentiles increased in Miami. One possible explanation for the small wage effects observed is that, in Miami, wages were rigid downward in the years 1979-1981 and hence a negative demand shock for native workers did not translate into lower wages. An alternative is that displacement of natives, rather than wage adjustment, took place. If either of these explanation is correct then the inflow of Mariel Cubans must be associated with an increase in the unemployment rate of non-Cuban high school dropouts in Miami. Panel D of Figure 5 shows the unemployment rate of the non-Cuban high

school dropouts 19-65 for Miami and for the Synthetic control. First, let us notice that the year-to-year volatility of unemployment in Miami before 1979 was quite large. In particular, we see a spike in unemployment in Miami in 1975 that is not matched by the Synthetic control. This should make us cautious, as other factors differentiating Miami from its control, existed in the pre-1979 period. With this caveat in mind we observe that the 1980-85 behavior of Miami unemployment rate relative to the Synthetic control does not show any significant departure. Even immediately after the shock in 1980 and 1981 no significant difference between the unemployment rate in Miami and Synthetic control arises.

5.4 Subsamples

By restricting the focus to sub-samples of high school dropouts one faces the serious risk of introducing very substantial measurement error as the sample size include only few dozens individuals in each metropolitan area. However, where possible, it can be beneficial to separate the impact by group. First, while Mariel Cubans were divided between men and women in similar proportions as in the preexisting labor force, one may think that their impact was differential. This may be the case if, for instance, they took jobs and specialized in occupations that were in competition with the male labor force. Panels A and B of Figure 6 show the Synthetic control analysis, when separating men and women. Second, one may think that the non-Cuban Hispanics in Miami were also likely to be prior immigrants and they should be separated when evaluating the impact of Mariel Cubans. Moreover, they or the African-American workers could be particularly exposed to the new immigrant competition. Panels C and D show impact for those two groups, separately. In line with the results obtained using aggregate data, each group exhibits negligible deviation of average log wage between Miami and control post 1979, once we align the 1979 observations. In fact, one may be tempted to argue that a *positive temporary* effect of Mariel Cubans is discernible on Hispanics. In this sample a small divergence appears in 1980 and 1981, it persists for several years until 1986 and then the noise of the control group becomes quite large. If one is *really* determined to find an effect that matches the temporary characteristics of a shock and the timing of the Mariel Boatlift, this could be it. Except that it is positive, and hence one could argue about the important complementaries between the new Cuban group and the existing group of

Hispanics in Cuba. Or possibly that the Mariel immigrants attracted business and created new opportunities for them. However, as this case is not confirmed in all subsamples, and as the point estimates of the deviations are small, it is more likely that such a deviation could be a by-product of noise that in these small samples is significant[15].

5.5 Regression Analysis

There are two types of uncertainty that affect inference with the Synthetic control method and it is not immediately clear how to produce standard errors and confidence intervals that account for both. The first type is simply due to "sampling variation", stemming from the fact that we are measuring the average outcome (wage, unemployment) in a metropolitan area, with error. This type of uncertainty would be eliminated if we had data on the whole population of workers. As shown in Section 2, given the small sample size in our setting, this error can be large. However, we can use time series observations in a regression framework over several years to estimate the difference between Miami and Synthetic control, and its standard error. The second type of uncertainty is due to the fact that even if we could measure the average outcome exactly, the pre-post 1980 differences in wages between Miami and its Synthetic Control can be affected by unobserved factors generating significant variation. In order to deal with this second type of uncertainty we produce a simulated test of the significance of the difference between the post-1980 outcome in Miami relative to its Synthetic control vis-a-vis the distribution of that statistics for all other cities in the sample. We will address this issue in section 5.6 below. The regression analysis conducted here allows us to construct standard errors in a familiar environment. We consider the time series observations for Miami and its Synthetic control between 1972 and 1991 and estimate coefficients as it is done in a regression context for a classic difference-in-differences analysis. The pre-event differences Miami-Control in this regression will also provide a confidence level in the choice of the comparison group done by the Synthetic control method. We estimate the following regression:

[15]Let us emphasize that if one is looking for a sample and specification that delivers a *positive* post-1979 deviation and a positive post-1979 trend changes in Miami, arguing for strong complementarity between the Mariel Cubans and local low skilled one can find it in Figure A2. This shows the Synthetic control analysis when the sample is the same as in Figure 3, and the Synthetic control is selected by also matching pre-1979 employment growth and high school dropouts employment growth. The figure is not very different from Panel A of Figure 3 but the positive deviation post 1979 is even clearer.

$$y_{it} = Miami_i + \sum_{P \in PRE-79} \alpha_P D_P + \sum_{P \in POST-79} \alpha_P D_P + \\ + \sum_{P \in PRE-79} \beta_P (D_P * Miami_i) + \sum_{P \in POST-79} \beta_P (D_P * Miami_i) + \varepsilon_{it} \quad (3)$$

The variable y_{it} is the outcome of interest (e.g. average log of weekly wages of high school dropouts) in unit i which can take only two values, either "Miami" or its "Synthetic control" and in year t, between 1972 and 1991. The variable $Miami_i$ is a dummy equal to one for Miami and 0 for the Synthetic control. D_P is a set of 3-year dummies that span the whole period but omit 1979, which is absorbed in the constant and hence serves as reference year, right before the shock took place. In the pre-1979 period the dummies are D_{73-75} and D_{76-78} in the post-1979 period they are $D_{80-82}, D_{83-85}, D_{86-88}$ and D_{89-91} and they equal one in the years indicated in the subscript and 0 otherwise. α_P is the set of coefficients corresponding to the period dummies and β_P is the set of coefficients associated to the interaction between the dummy $Miami_i$ and the period dummies. The term ε_{it} captures the classical error term, uncorrelated with the observables and with 0 average that we interpret as residual measurement error for each metropolitan area. The method of estimation used is Feasible Generalized Least Squares allowing the measurement errors to be autocorrelated in an AR(1) process[16]. The coefficients of interest are β_P's. In particular, if the Mariel shock had any effect, this should be captured by the coefficient β_{80-82}. It captures the average difference between Miami and Synthetic control arising in 1980, 1981 and 1982 once the 1979 difference is standardized to 0. The subsequent coefficients $\beta_{83-85}, \beta_{86-88}$ and β_{89-91} complete the picture. Theory predicts that if their value is affected by the Boatlift, it should be smaller than β_{80-82}. Several economic shocks (including an important recession and a worsening of the war on drugs in 1982 which involved Miami in a very intensive way) took place during the decade and Miami could have responded differently from the control group. Hence, the farther away from 1979 we look the more likely it is that other factors could have affected the difference between Miami and its Synthetic control. Just as importantly, our framework allows us to estimate the pre-1979 differences between the two cities. The estimates of β_{73-75} and

[16] Assuming that errors are independent over time does not change much the estimates of the coefficients and standard errors.

β_{76-78}, provide validation for how well the two cities track each other before the shock. Statistically significant pre-1979 differences would cast doubts on our control group as they will imply systematic deviations between the Miami and control, even before the treatment (Boatlift).

Table 3 shows all estimated coefficients. The header of each column indicates the Panel and Figure corresponding to the estimated regression. The analyzed dependent variables are the log of weekly wages for high school dropouts in column (1), log of hourly wages for high school dropouts in column (2) log of weekly wages at the 15th, 20th, and 90th percentiles of the native distribution in columns (3), (4) and (5), and unemployment in the last column. Each regression is estimated on 38 or 40 observations[17], and hence any result should be taken with a grain of salt. The estimates are simply a quantification of the deviations between the time series represented in figures 3 and 5 with the provision that the regression standardizes this to 0 in 1979, while the graphs minimize the whole pre-1980 distance (along with other labor market characteristics) without setting it to 0 in 1979. Some features of the estimates are consistent and robust. First, Miami and Synthetic control move together, to a reasonable extent, in the pre-1979 period: the coefficients of the pre-period interactions ($73-75$ and $76-78$) are not significant except in one case (20th wage percentile in 72-75). This is a validation of the choice of the control group and an an indicator of goodness of fit. The standard errors for the wage regressions, however, (between 0.04 and 0.06 log points) are large enough and deviations in the order of few percentage points would be hard to find. Second, none of the deviation coefficients for the period right after the Boatlift, β_{80-82} is significantly different from 0. Moreover the point estimates of all β_{80-82} relative to a wage outcome (Column 1-5) are positive and they reveal that Miami had a small departure upwards relative to its Synthetic control after the Boatlift. Given the estimated coefficients and standard errors we can rule out a negative effect larger in absolute value than 2 or 3%. Hence the prediction of the simple model without adjustment seems strongly rejected[18].

[17] The May-ORG CPS does not include year 1972 and has, therefore 38 observations.

[18] Extending the consideration to the 1983-85 period, most of these coefficients (deviations) for wages are also positive and some of them significant, and for unemployment the coefficient is very close to 0. In some cases these deviations are larger than in 1980-82 and this is a sign that we should not consider them in any way as consequences of the Mariel Boatlift, unless we have some theory of why that shock should have affected the labor markets with 2-3 years delay. Finally, the only negative departure for wages (in point estimates) are estimated for the 86-88 and 87-91 intervals, but again the large imprecision of the data and the long time since the Boatlift suggests that these deviations are completely uninformative of the effects of the Boatlift.

Overall, both the inspection of figures 3, 4 and 5 and the regression analysis in Table 2 suggest that: (i) there is no significant deviation of wages and unemployment level for high school dropouts in Miami relative to the control after the Mariel Boatlift (ii) most of the point estimates for the deviation of wages are actually *positive* (iii) a negative effect as large as the one predicted by the short-term model without adjustment can be ruled out at standard confidence level, in most cases.

5.6 Inference Using Permutations

While the regression approach has its appeal and simplicity, the small number of time series observations and the imprecision of the estimates limit credibility and potential. A more accurate way of doing inference with the Synthetic control method proposed by Abadie et al. (2010) is based on permutations. The core idea is to simulate a distribution of deviations between each city in the donor pool and its Synthetic control and examine whether Miami shows a post-1979 deviation from its Synthetic control that is large relative to the whole distribution. Panels A-D of Figure 7 do exactly this, analyzing log weekly wages, log hourly wages, log wages at the 15th percentile and unemployment rate for non-Cuban high school dropouts, respectively. The dark line corresponds to Miami, while each of the lighter lines correspond to one of the 43 control cities. All panels reveal that Miami is a rather average city in the pre-1979 deviations from its Synthetic control in any outcome. Then Panel A and B show that Miami average wage had a positive deviation from its control which is in the high end of the simulated range, in 1980-82 while its deviations look within the range of idiosyncratic variation after that year. Panel C and D show instead that for the 15th wage percentile and for the unemployment rate Miami is well into the range of simulated deviations any year post 1979. Notice that the range of simulated idiosyncratic noise can be quite large in the sample. For instance log weekly and hourly wages show a range of noise spanning the interval between -20% and +20%. Let us emphasize once again that with this degree of noise it may be hard to identify effects in the order of few percentage points. However, let us also notice that of the 44 simulated cases Miami is one showing a *relatively high and positive* wage deviation from Synthetic control in the 1980-82 period. This is not at all consistent with the prediction of the simple model of section 3. A reasonable explanation for this is that the complementarity between new immigrants and natives was significant and the changes of techniques and specialization

that allow an economy to absorb immigrants with no wage changes for natives were at work already in the short run.

In Table 4 we show test statistics based on the simulations reported in Figure 7. First we calculate the Pre-Post ratio in the average absolute deviation of Miami from its control, considering 1980-82 as the post-period and, alternatively, either the 1972-79 (upper panel in Table 4) or the more recent 1977-79 interval (lower panel in Table 4) as pre-period. Notice that by taking the absolute deviation of Miami from the Synthetic control after the Boatlift (1980-82) we are considering if any significant deviation arises, and standardizing for the pre-1979 average absolute deviation we adjust this value for the idiosyncratic deviations contained in pre-1979 existing factors. We then do the same for all other cities in the sample of 44 and we produce the same statistics. In Table 4 we show the value of this ratio statistics. We also show the rank of Miami statistics in the distribution of 44 cities (1 being the lowest value and 44 the highest) and the p-value of a one-sided test that uses the distribution of these statistics for 44 cities, for the probability of a city in the distribution having a statistics larger than Miami[19]. A low value of the rank and a value of the p-statistics higher than 0.10 indicate that Miami deviations are not unusual relative to the other cities. Column (1) and (2) show quite high values of the statistics (and of the ranking of Miami) and low p-value, which indicate that Miami average wage of dropouts was somewhat unusual in its post 1979 wage deviation ... *but in the positive direction*! In column (3) and (4) instead, Miami tends to be at low or intermediate levels of the distribution of the statistics revealing that it is well within the range simulated for the other 43 cities that did not receive the Boatlift. This means that idiosyncratic variability produced by many other factors and by measurement error, and likely to exist in any unit-control pairing, fully explains the post-1979 behavior for Miami as that city does not appear to be an outlier.

6 March-CPS Sample and the Results in Borjas (2015a)

An important question to address is how to reconcile our estimates of small and non-significant effects from the Mariel Boatlift, which confirm the original findings of Card

[19] We implement the correction technique of Ferman and Pinto (2015) who derive conditions under which inference in Synthetic Control corrects for heteroskedasticity. Namely, it requires normalizing the post-RMSE by dividing it by the pre-RMSE. See the paper for an in depth discussion.

(1990), with the results of Borjas (2015a). The latter study presents estimates of the wage effects of the Mariel Boatlift that are negative and much larger than the prediction of the most conservative model of short-run effects described in 3. Borjas (2015a) argues that a decline of 30-40% in the wages of high school dropouts in Miami relative to a control group, that reached a negative peak in 1985, five years after the shock (when the share of Cubans in Miami was back to pre-1979 levels), was a consequence of the Boatlift. The model of section 3 with no adjustment, perfect substitution immigrant-natives and the largest estimates of the Mariel inflow can at most predict a wage decline of 7-8% (about 5 times smaller) in 1981.

In this section we show how the negative estimates of Borjas (2015a) can only be found when using the March-CPS sample that we argued to be too imprecise, too noisy and too small to address any reasonable wage effect. Moreover the negative wage deviation of Miami from Controls is significant only in a very specific and carefully selected sub-sample of high school dropouts: male, non-Hispanic, in the 25-59 age range. This sample excludes about two thirds of low skilled Miami workers. One also needs to focus only on one labor market outcome (weekly wages of high school dropouts) to find such effect. The fragility of Borjas' (2015a) results to changes in sample and in variable definition confirms how noisy and unreliable the March-CPS data are especially when analyzing a sub-group in metropolitan labor markets. Perhaps this is not too surprising as Column 3 of Table 2, showed that the sample selection and the choice of March-CPS leaves Borjas (2015a) with only 17 to 24 observations per year in Miami during the period 1980-89 after the Boatlift. The restrictions also produce very small samples in the cities used as controls. Cincinnati has a number of observations between 36 and 108 and New Orleans, between 20 and 67 in the period . What Borjas (2015a) calls the average wage of native dropouts in Miami, a group that included about 120,000 individuals as of 1980 (see Table 1) is estimated in his study by 17 to 24 individuals per year! In the remaining of this section we present the results obtained from the largest possible sample and then from sub-samples of the March-CPS and then we replicate and show the sensitivity and limitations of Borjas' (2015a) results.

6.1 Estimates using March-CPS and sub-samples

Figure 8 shows the Synthetic control analysis on the sample of non-Cuban high school dropouts using March-CPS data. Panel A and C consider (log) weekly wages as outcome and differ between each other as Panel A does not include the group of individuals still potentially in school, focusing on workers between 19 and 65 years old. Panel B includes the younger workers and looks at the age range 16-61. As the weekly wage is only obtained indirectly in the March-CPS (dividing the yearly wages by the number of weeks worked) we also consider the direct measure of (log) yearly wages in Panel B (age 19-65) and D (age 16-61). One observes significant year to year variability of average wages both for Miami and for the control group, especially after 1985. Moreover the graphs appear somewhat sensitive to small changes in sample and variable. For instance, Panel A shows a *small negative deviation* of Miami from control in 1981 (but not in 1980 or in 1983) with larger negative deviations after 1985. To the contrary Panel D shows a *small positive deviation* after 1979, especially in 1980, and *larger and more significant positive deviations after 1983*. In general, however, for weekly wages one does not see significant departures of Miami wages from the placebo until 1984, with the largest deviation in 1986. The yearly wage analysis, moreover, does not seem to show any consistent deviation at all but instead very substantial year to year variation and noise. Aware of the very large noise and imprecision in the March-CPS we are very reluctant in considering any deviation shown in Figure 8 as effects of the Mariel Boatlift. Another set of good reasons not to take the small deviations observed in Figure 8 after 1979 as effects of the Boatlift are illustrated in Figure 9. In this figure we show two alternative measures of low skill non-Cuban wages in Miami, namely the wage at the 15th (Panel A) and at the 20th (Panel B) percentile of the non-Cuban wage distribution. In both cases one observes a small positive relative deviation in 1980 and a small negative one in 1981 and no relative deviation in 1982. This are noisy patterns and likely due to idiosyncratic error. This is confirmed by Panel C in which we show, as a falsification test, the 90th percentile of the non-Cuban wage in Miami and Synthetic control. The high-skilled workers' wages should show no negative effect from the Mariel Boatlift, and possibly a small positive effect due to complementarity. We see, instead, similar deviation patterns (positive in 1980, negative in 1981 and no deviation in 1982) in this statistics as observed for the 15th and 20th percentile. So the March-CPS wage data seem to show a degree of imprecision and idiosyncratic noise for Miami in 1980-1982, but no deviation consistent with the negative prediction of the short-run

effects of immigrants on wages of less educated workers. To complete the picture Panel D shows the unemployment rate of non-Cuban high school dropouts and no significant deviation appears in the post-1979 period, relative to the idiosyncratic deviation present throughout the sample.

More systematic measures of the (lack of) statistical significance of post-1979 deviations in these Synthetic control samples are shown in Table 5. The table reports the regression coefficients relative to the high school dropout wages represented in Panels A and C of Figure 8 (first two columns) and for all the Panels of Figure 9 in the remaining columns. Three results emerge. First, no coefficient relative to the 1980-82 deviation of Miami from control (row 3) is statistically significant. Second the standard errors when considering dropout wages as outcome (columns 1 and 2) are in the order of 0.07 to 0.17 log points, which makes them two to three times larger than those obtained for the ORG-CPS sample in Table 3. Third, when we estimate a large value for the 80-82 deviation in the order of 10%-20% such as in columns 1 and 2, we also estimate large deviations before 1979 and well in later years (e.g. 1986-88). This also suggest that it would not be reasonable to attribute the non-significant large deviation after 1979 to the one-time Mariel Boatlift but rather that the idiosyncratic noise probably dominates the exercise.

Finally, Figure 10 shows the Synthetic control analysis when considering sub-groups of the non-Cuban high school dropouts. The very big caveat is that in all cases these groups include less than 55 observations per year in Miami and in most control cities in the 1980s and, in several years, the number of observations drops well below 20. The measurement error here can be very large. We separate the groups of men (Panel A), women (Panel B), Hispanics (Panel C) and African-American (Panel D) workers. What stands out is the very large and irregular year to year variation of average wages in several samples (Women and Hispanics) and the fact that the post-1979 pattern is different across sub-groups. One can spot a negative deviation in 1981 and 1982 for men, but not much for women, a positive deviation for Hispanics and a negative one for African-American. The match of the pre-1979 trend and fluctuations between Miami and Synthetic control is not precise in any of these sub-groups and in all of them one observes more significant deviation of Miami from control after 1984 that would hardly be reasonable to explain as a consequence of a 4-year old supply shock, by then fully absorbed.

6.2 Sensitivity of Borjas (2015a) results

The results of Borjas (2015a) are obtained using March-CPS data for non-Hispanic males 25-59 years of age, a sub-sample that is smaller than any of those used in Figure 10. Figure 11 shows our replication of Figure 3A in that paper, extended back to 1972. That figure and the corresponding regression analysis in Column 1 of Table 6 capture the essence of the findings in Borjas (2015a). The figure shows the high school dropout log wages in Miami and in the "Employment Control" (chosen by Borjas to be constituted by an average of Anaheim, Nassau-Suffolk, Rochester and San Jose). We have "aligned" Miami and the employment control in 1979, to give a cleaner visual impression: the picture conveys a strong idea of a large and protracted wage drop in Miami relative to the control starting in 1980. A few features are surprising. First, the picture, although it uses a very small sample made of 17-24 observations per year, has much less year-to-year variation than all the previous we produced. This is due to a smoothing procedure that we will discuss later. Second, Miami and Control diverge from 1980 in a progression that peaks in 1985 and continues up to 1987, and this is a feature that no other previous figure showed. Third, one can see that the "employment control" chosen by Borjas does not match very well Miami in the pre-1979 period, being flatter in the 1977-79 period and steeper before. The impression from this graph is that something very major started in 1980 and continued for 7 years affecting negatively Miami wages.

Panels A-D of Figure 12 illustrate the sensitivity of the result shown in Figure 11 to the definition of sample and variable. We also want to attract attention to the smoothing and the timing of the Miami-Control departure. First in Panel A of Figure 12 we introduce two straightforward modifications to Figure 11. First, Borjas (2015a) smooths the time series using a 3-year moving average. This procedure seems to run contrary to the identification idea, based on exploiting the suddenness and exact timing of the temporary shock (April-June 1980) and of its consequences. By using a moving average we confound data of the pre-shock observation with 1980 data. We also include later dynamics of 1982, in the "post" observation of 1981, which should capture just the response to the shock. Borjas (2015a) argues that this procedure is adopted to increases the yearly sample size. Certainly, the 15-25 observations used in his analysis need some improvement, but using a moving average builds autocorrelation in the time series and one cannot consider the post-1979 observations as independent. Hence, we will not use the moving average. Second, Borjas (2015a) only includes non-Hispanic prime-age (25-59) males. We add to these

workers the Hispanic non-Cuban individuals and we extend the age range to include the more often used working-age period for high school dropouts (age 19-65). Hispanic workers were mainly United States born and were more similar in their jobs and occupations to the newly arrived Cubans. Hence this broader choice should improve the precision and detect a stronger effect. Individuals with weaker job protection and shorter labor market attachment (young) could also be more vulnerable to new immigrant competition. In Panel A of Figure 12 we include the two changes described above and we report the Synthetic control constructed by matching values of 1972-1979 variables as in our previous analysis (rather than matching only employment growth in 77-80 as done in Borjas). The figure has already changed importantly. First, the downward trend of dropout wages 1972-1986 common to both Miami and Synthetic control, is now clear, albeit with noise. Second, the dropout wages in 1979-81, when the shock took place and should have produced the largest wage effects in Miami relative to control, show no deviation of Miami. Third, we notice how the Synthetic control group matches much better the pre-1979 behavior of Miami than the "employment control" did. In this frame, the fact that we observe no departure between Miami and control up to 1981 is the strongest hint that the Boatlift did not have any significant effect on non-Hispanic male high school dropouts. We notice in Panel B that a departure of Miami from the Synthetic control arises in 1982 and lasts until 1986. This departure was not present in any of the Synthetic control analyses using the May-ORG CPS and hence can be due to pure noise. Nevertheless, to explore whether this feature may be due to something happening in Miami in 1982 (not in 1980!) to the male, non-Cuban group of high school dropouts we take the exact same specification, sample and outcome variable (log weekly earnings) of Panel A and we simply use the larger May-ORG CPS sample. Doing this we obtain Figure 11 Panel D, showed in the right-bottom corner of Figure 12. The departure of Miami log wages from the control between 1979 and 1983 becomes totally negligible, and in 1984 and 1985 a *positive* departure appears. This is a sign that the residual departure observed in 1982 for Panel A was simply measurement error. The larger sample does not show it. In fact, even more directly, we keep the exact sample definition of male non-Hispanic dropouts age 25-59 as in Borjas (2015a), and we use the ORG-CPS sample and the Synthetic control matching 1972-79 variables, and show this analysis in Panel C, at the left bottom corner of Figure 8. The divergence between Miami and Synthetic control between 1979 and 1983 disappears completely again. Using ORG-CPS data on samples very similar to Borjas we do not find any deviation at all after

1979. Finally, to confirm that the original data in the March-CPS sample are extremely noisy in Panel B, top-right of Figure 12, we consider the same sample as Panel A, and we use the yearly (rather than weekly) wage definition. Interestingly the small size of the sample (with no smoothing) shows very dramatic, sharp and erratic year to year variation in log wages in Miami. It is not uncommon to have 0.20 log points variations in opposite directions in two successive years. Clearly this variance is too large to say anything precise on wage impacts in the order of 7-8%.

The regressions of Table 6 show the regression coefficients for the Miami-control specifications as presented in Figure 11 (column 1) and in Figure 12, Panel A-D. These estimates illustrate in a formal way how the findings in Borjas (2015a) sample quickly disappear and give raise to extreme noise as soon as we introduce the small modifications discussed above. Even in column (1) that uses exactly Borjas specifications of Figure 11, we see that the deviations in 1980-82 are large (-0.18 log points) but extremely noisy (standard deviation 0.09) and not very different from pre-1979 deviation (-0.12 in the period 72-75)[20]. Already in column (2) where we do not smooth the data and include Hispanics and 19-65 years old, reduces the 1980-82 estimated deviation within one standard error (-0.13) of zero and improves somewhat the pre-1979 match. Column (3) shows the very large and extremely imprecise deviations that one obtains pre and post-1979 between Miami and controls when using yearly wage as a outcome (and identical sample as in Borjas 2015). Finally, columns (4) and (5) show that less noise and no deviation whatsoever is detected post 1979 when we use the ORG-CPS data either on Borjas exact sample (column 4) or on that sample plus Hispanics and extended age group (column 5). Almost any departure from the very strict definition of sample, and of the dependent variable used in Borjas (2015a) either eliminates the negative deviation of Miami or results in incredibly large noise before and after 1979. We do not think it is reasonable to interpret a result based on a small and very noisy sample, only on one outcome, likely arising after 1982 and not confirmed in most other samples, as an impact of the Boatlift.

[20] Borjas (2015a) does not present the estimates of pre-1979 deviations between Miami and control and does not test, nor discusses, whether they are different from 0.

7 The Boatlift That Did Not Happen

In 1994 Fidel Castro once again announced that Cubans who wanted to flee the country would not be stopped. However, this time the United States Coast Guard diverted the majority of the flow to the naval base in Guantanamo Bay. A big inflow of refugees to Miami was about to take place but did not[21]. Nevertheless, Angrist and Krueger (1999) show that between 1993 (pre non-shock) and 1995 (post non-shock) the unemployment rate for Black workers in Miami increased by 3.6 percentage points, while in the control group (the one of Card 1990) it decreased by 2.7 percentage point. If a researcher were to analyze the impacts of this "non-event" she will estimate a fake "treatment effect" of +6.3 percentage points. While illustrating the classic difference-in-difference methodology, they argue that this "false positive" is a cautionary tale when utilizing a small number of units. Just as the coincidence of the non-event and the change in unemployment was a "false positive", the findings of Card (1990) could also be "false negative".

However, while idiosyncratic deviations between treatment and control groups, especially in the short run, can be pervasive and need to be kept in mind, this is only a partial tale. First, in order to have a complete story one should look at several groups and at wider set of outcomes. In so doing we already see that even using the data reported in Angrist and Krueger (1999) the unemployment rate for white workers shows a Difference-in-Difference change for Miami relative to control equal to only 0.3 percentage points from 1993 to 1995. The unemployment of Hispanics shows a difference-in-difference change of +1.4 percentage points. Both these differences are within two standard errors for the average Miami unemployment rate. Hence neither of them confirms the presence of an effect. Second, one should carefully verify the pre-event match between the treated unit and the control group. Looking at the deviations of Miami relative to the control group, for unemployment rates of black workers, one realizes that there are many instances of quite large deviations before 1994 (e.g. +4.7 percentage points in 88-89, or -2.7 percentage points in 1991-1992). This speaks strongly against the validity of the control group chosen. So, we mostly learn about the need to stay cautious and possibly to require an array of results to converge in one direction, before claiming an effect [22].

[21] By looking at Miami CPS-ORG data, the share of Cubans in Miami either declined or remained unchanged between 1993 and 1995. We show this in Figure A2 in the Appendix. Hence we maintain the assumption of Angrist and Krueger (1999) that this Boatlift did not change the supply of Cubans in Miami significantly.

[22] Importantly, let us also point out that in 1994 the CPS underwent a major redesign and several

The Synthetic control method makes significant progress on the problems raised by Angrist and Krueger (1999). It eliminates the arbitrary choice of control and it allows the control cities for Miami in 1994 to be different from those chosen by Card (1990). It also allows to produce a validation, checking how good the pre-1994 fit of Miami and control was. We apply the Synthetic method to high school dropouts weekly wages and hourly wages and to wages at the 20th percentile, for the preferred sample using the ORG-CPS, which are among the variables we considered in our previous analysis[23]. Figure 13 Panels A-B show the behavior of hourly and weekly wages of high school dropouts in Miami and Synthetic control between 1989 and 2001. In order to have a balanced panel of control cities (consistently defined over the whole period) we keep the pre-1994 period to 6 years only. The vertical line is on year 1993, preceding the Boatlift that ended up in Guantanamo. In Panel A, using hourly wages, we do not observe any departure in 1994 and a small positive one in 1995-96, however the size of the departure is within the typical variation between Miami and control along the pre-period. In Panel B, using log weekly wages, we similarly do not observe significant deviations. Using the 15th wage percentile as alternative measure of low skill (non-Cuban) wages, in Panel C and D, we do not observe any significant difference between Miami and control arising in 1994 or 1995 neither for hourly nor for weekly wages. Overall, these pictures do not produce any evidence of a downward wage movement for low skilled in Miami post 1994. As for the unemployment rate of minorities (Black and Hispanics) that is the only variable considered in Angrist and Krueger (1999), we show their behavior vis-a-vis the Synthetic control in Figure A3 of the appendix. While unemployment of Black individuals still shows an increase relative to the Synthetic control in 1994 and 1995, that difference is less dramatic and it is reversed by 1996. The unemployment of Hispanic individuals experienced actually a decline relative to the Synthetic control in 1994-1995. The unemployment of Hispanics in the Synthetic control grew significantly more than that of Miami in that period. Overall the Synthetic control analysis done on wages and unemployment for Miami 1994-95 would bring the researcher to recognize significant noise of the data and not to identify consistent signs of effects on wage and employment of the non-existent 1994 Boatlift.

measures of employment, especially for males and subgroups were significantly affected (see Polivka and Miller 1995). Hence focusing on changes exactly around 1994 can be very risky.

[23] In this case, to keep computational time within a reasonable amount, we limit the "donor pool" for the control group to cities with at least 20 observations in the relevant group of high school dropouts per year. This produces a pool of about 40 cities.

8 Conclusions

This paper applies the Synthetic control method to the Miami Boatlift episode. That inflow of refugees increased labor supply of Cubans in Miami possibly up to 18% of the high school dropout group and by 8% of the total labor force, between April and September 1980. We use a wide variety of labor market outcomes for high school dropouts non-Cubans in Miami and for several sub-groups and we look for a significant and sudden impact of this shock on Miami labor markets in the period 1980-1982. We do not find any consistent evidence of a short-run depressing effect on low-skilled labor demand nor any dynamics after that. In applying this method we learned that noise and measurement error due to small samples from the March-CPS data are too large for comfort. Hence the choice of the more accurate and less noisy ORG-CPS survey seems necessary. Noise remains large enough that log wage deviations in the order of few percentage points between Miami and control can be hard to identify. If one also adds the fact that several other idiosyncratic shocks hit the United States economy in the period between 1980 and 1982 (a strong recession in 1982, an increase in minimum wage in 1981) the need to rely on a good control is particularly important. The presence of other Miami-specific shocks in this period (i.e. the intensifying of the drug war in 1982, the 1980 riots for the death of Arthur McDuffie, a black salesman, at the hand of a white police officer) should also all be kept in mind as potential caveats in the attempt of isolating the effects of the Boatlift.

Overall, we learned that by applying the Synthetic control method, and including extensive checks (for different samples, variables and groups), we can improve on Card (1990). The matching of the pre-event trend by the control group is improved, we are given a less ad-hoc method to choose it and we produce correct standard errors and confidence level for our inference. We also understood that by choosing a small enough sample in the March-CPS data one can obtain erratic, noisy and large deviation of Miami unskilled wages from control that would be unwise to interpret as effects of the Boatlift. The ORG-CPS sample shows small, not significant and positive deviation of low-skilled wages in Miami post 1979, and rejects the possibility of negative deviation in the order of those predicted by a simple model. Hence, even using the more current econometric methods, we do not find any significant evidence of a negative wage and employment effect of the Miami Boatlift. This finding is consistent with the recent literature emphasizing mechanisms that allow absorption of immigrants through complementarity, technology adjustment and efficiency and go beyond the naive canonical model that is too partial in

the understanding of the labor market effects of immigrants.

References

[1] 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 (2010), 493–505.

[2] Abadie, Alberto, and Javier Gardeazabal, (2003) "The Economic Costs of Conflict: A Case Study of the Basque Country," *American Economic Review* 93, 113–132.

[3] Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. (2015) "Comparative politics and the Synthetic control method." *American Journal of Political Science* 59.2 : 495-510.

[4] Angrist, Joshua (1995). "The Economic Returns to Schooling in the West Bank and Gaza Strip." *American Economic Review*, 85, 1065–1087.

[5] Angrist Joshua and Alan Krueger (1999) "Empirical Strategies in Labor Economics" in O. Ashenfelter & D. Card (ed.), "Handbook of Labor Economics," Handbook of Labor Economics, Elsevier, edition 1, volume 3, number 3.

[6] Angrist, J. D., and Pischke, J. S. (2010). "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics." *Journal of Economic Perspectives*, 24(2), 3-30.

[7] Bodvarson Orn, H.F. Van der Berg and J.J. Lewer (2008) "Measuring Immigration's effect on Labor Demand: A Reexamination of the Mariel Boatlift" *Labor Economics*, 15:4.

[8] Borjas, George (2012) "Labor Economics" 6th Edition, McGraw-Hill Companies.

[9] Borjas, George, (2015a) "The Wage Impact of the Marielitos: a Reappraisal" National Bureau of Economic Research, Working Paper # 21588, Cambridge Mass.

[10] Borjas, George (201b) "The Wage Impact of the Marielitos: Additional Evidence". National Bureau of Economic Research, Working Paper # w21850, Cambridge Mass.

[11] Bohn Sara, Magnus Loefstrom and Steven Rapahel (2014) "Did the 2007 Legal Arizona Workers Act Reduce the State's Unauthorized Immigrant Population?," *The Review of Economics and Statistics*, MIT Press, vol. 96(2), pages 258-269, May.

[12] Bollinger, Christopher R. (1998) "Measurement error in the Current Population Survey: A nonparametric look." *Journal of Labor Economics* vol. 16 issue3 : 576-594.

[13] Card, David (1990) "The Impact of the Mariel Boatlift on The Miami Labor Market" *Industrial and Labor Relation*, Vol 43 n. 2.

[14] Card, David (2001) "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, University of Chicago Press, vol. 19(1), pages 22-64, January.

[15] Card, David (2009) "Immigration and Inequality," *American Economic Review*, American Economic Association, vol. 99(2), pages 1-21, May.

[16] Cahuc, Pierre, Stephan Carcillo, A. Zylberberg and Andre' Zylberberg (2014) "Labor Economics" Second Edition, MIT Press, Cambridge Mass.

[17] Carrington, W. J., and De Lima, P. J. (1996). "The impact of 1970s repatriates from Africa on the Portuguese labor market." *Industrial and Labor Relations Review*, 49(2), 330-347.

[18] Dustmann, Christian, Tommaso Frattini and Ian Preston (2013) "The Effect of Immigration along the Distribution of Wages," *Review of Economic Studies*, Oxford University Press, vol. 80(1), pages 145-173.

[19] Ferman, Bruno, and Christine Pinto (2015) "Inference in differences-in-differences with few treated groups and Heteroskedasticity."

[20] Foged, Mette and Giovanni Peri (2016) "Immigrants' Effect on Native Workers: New Analysis on Longitudinal Data," *American Economic Journal, Applied Economics*, Vol 8, Issue 2, pages 1-35

[21] Flood Sarah, Miriam King, Steven Ruggles, and J. Robert Warren. Integrated Public Use Microdata Series, Current Population Survey: Version 4.0. [Machine-readable database]. Minneapolis: University of Minnesota, 2015.

[22] Friedberg, R. M. (2001). "The impact of mass migration on the Israeli labor market." *Quarterly Journal of Economics*, 1373-1408.

[23] Hunt, J. (1992). "The impact of the 1962 repatriates from Algeria on the French labor market." *Industrial and Labor Relations Review*, 45(3), 556-572.

[24] Johnson, George E., (1997). "Changes in Earnings Inequality: The Role of Demand Shifts." *The Journal of Economic Perspectives*, 11, 41–54.

[25] Laing Kevin (2011) "Labor Economics" W.W. Norton and Company, New York-London.

[26] Katz, Larry and Kevin Murphy (1992). "Changes in Relative Wages 1963–1987: Supply and Demand Factors." *Quarterly Journal of Economics*, 107, 35–78.

[27] Krusell, Per, Lee Ohanian, Victor Rios-Rull, and Giovanni Violante (2000). "Capital–Skill Complementarity and Inequality: A Macroeconomic Analysis." *Econometrica*, 68, 1029–1053.

[28] Lemieux, Thomas. (2006) "Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill?." *The American Economic Review* : 461-498.

[29] Lewis, Ethan (2004) "How did the Miami labor market absorb the Mariel immigrants?," Working Papers 04-3, Federal Reserve Bank of Philadelphia.

[30] Lewis Ethan and Giovanni Peri (2015) "Immigration and the Economy of Cities and Regions," in Gilles Duranton, J. Vernon Henderson and William C. Strange Eds., Handbook of Regional and Urban Economics, Volume 5,

[31] Monras, Joan (2015) "Immigration and Wage Dynamics: Evidence from the Mexican Peso Crisis." Manuscript, Paris Science Po.

[32] Ottaviano Gianmarco I.P. and Giovanni Peri (2012) "Rethinking The Effect Of Immigration On Wages," *Journal of the European Economic Association, European Economic Association*, vol. 10(1), pages 152-197, 02.

[33] Peri Giovanni (2014) "Do Immigrant Workers Depress the Wage of Native workers?", IZA World of Labor, Evidence-Based Policy Making, May 2014, available at http://wol.iza.org/articles/do-immigrant-workers-depress-the-wages-of-native-workers/long (accessed November 20th 2015).

[34] Polivka, Anne E. and Stephen M. Miller (1995) "The CPS After the Redesign: Refocusing the Economic Lens." United States Bureau of Census, 1995; available at "http://www.bls.gov/ore/abstract/ec/ec950090.htm" (accessed on November 20th, 2015)

Tables and Figures

Table 1:
Demographics of Mariel Immigrants and of Miami Labor Force in 1980

	Miami Labor Force in 1980	Mariel immigrants, measured from the 1990 Census	Mariel Immigrants still in Miami as of 1990
Total in Labor Force (16 to 65 years of age)	644, 860	87, 347	54,196
Share with no HS degree	26.28	55.77	56.34
Share with HS degree	32.11	25.18	24.22
Share with some college	22.37	12.53	12.47
Share with college	19.24	6.52	6.97
Share of female	45.79	37.80	41.97
Share of young (<25 years old)	16.35	16.34	14.01
Only individuals with no High School degree			
Total in labor force	169,440	48,714	30,532
Percentage female	43.33	39.79	44.41
Percentage young (<25 years old)	11.36	12.90	10.24

Note: The values for the Miami Labor force are obtained from the 1980 census. Those on the Mariel Immigrants are obtained from the 1990 census as people born in Cuba who arrived in the US in 1980 and 1981 and were at least 19 years of age at the time of arrival. Labor force is defined as individual 19-65, not in school, and working or looking for a job.

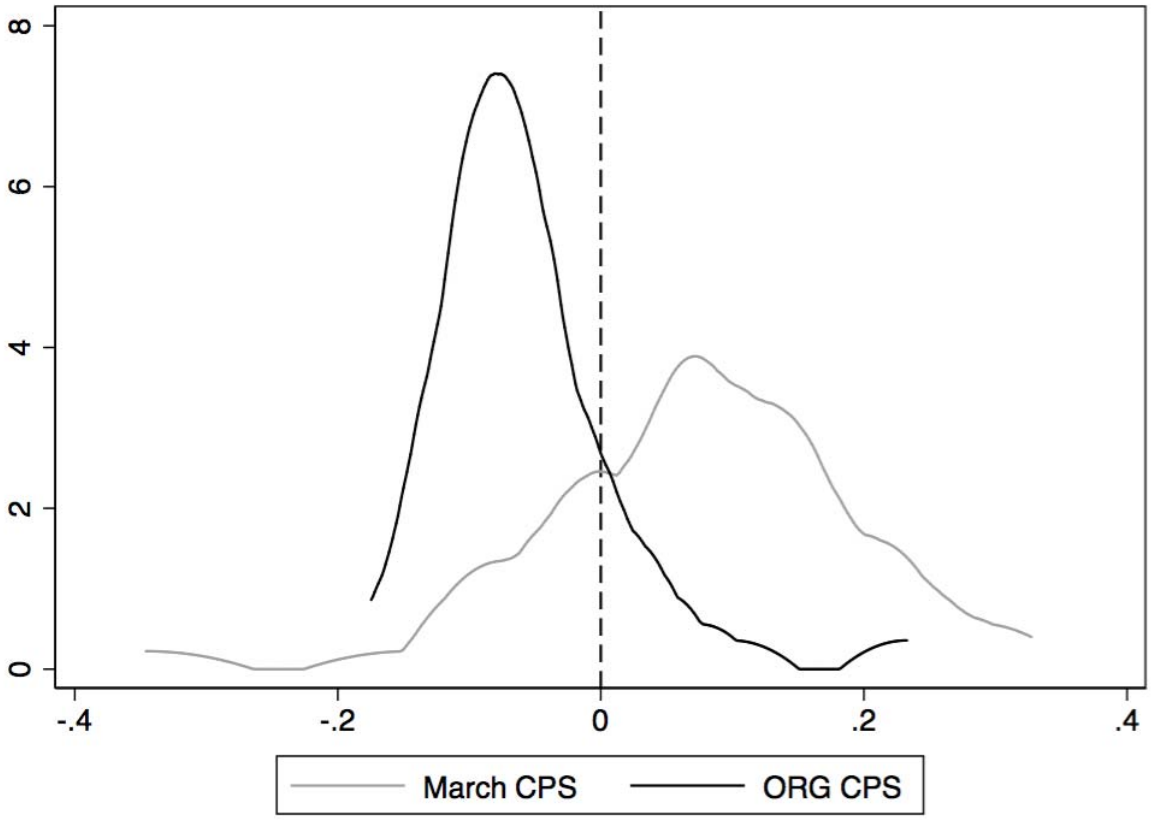
Table 2:
Number of observations for High School Dropouts in Miami, different samples

Year	Our sample selection, March CPS	**Preferred:** Our sample selection, May-ORG	Borjas' Preferred Sample, March CPS	Borjas' Preferred Sample, May-ORG
1973	70	42	30	17
1974	69	32	25	11
1975	66	41	16	13
1976	62	43	22	18
1977	64	39	25	19
1978	61	37	21	11
1979-Pre	62	**145**	17	56
1980-Shock	68	**161**	16	55
1981-Post	72	**145**	18	51
1982	62	**135**	24	39
1983	59	**149**	17	50
1984	55	**145**	15	48
1985	55	**72**	17	26
1986	61	**183**	17	61
1987	66	**221**	17	78
1988	86	**222**	17	72
1989	96	**222**	17	64
1990	74	**250**	4	63
1991	72	**175**	9	41

Note: Our sample includes individuals with no high school degree, non-Cuban, with positive earnings, not self-employed, in the labor force in the age range 19-65. Borjas' sample includes individuals with no high school degree, with positive earnings, not self-employed, in the labor force, only male, only non-Hispanic and in the age range 25-59. A one year adjustment is made to the March CPS numbers as previous year earnings are reported.

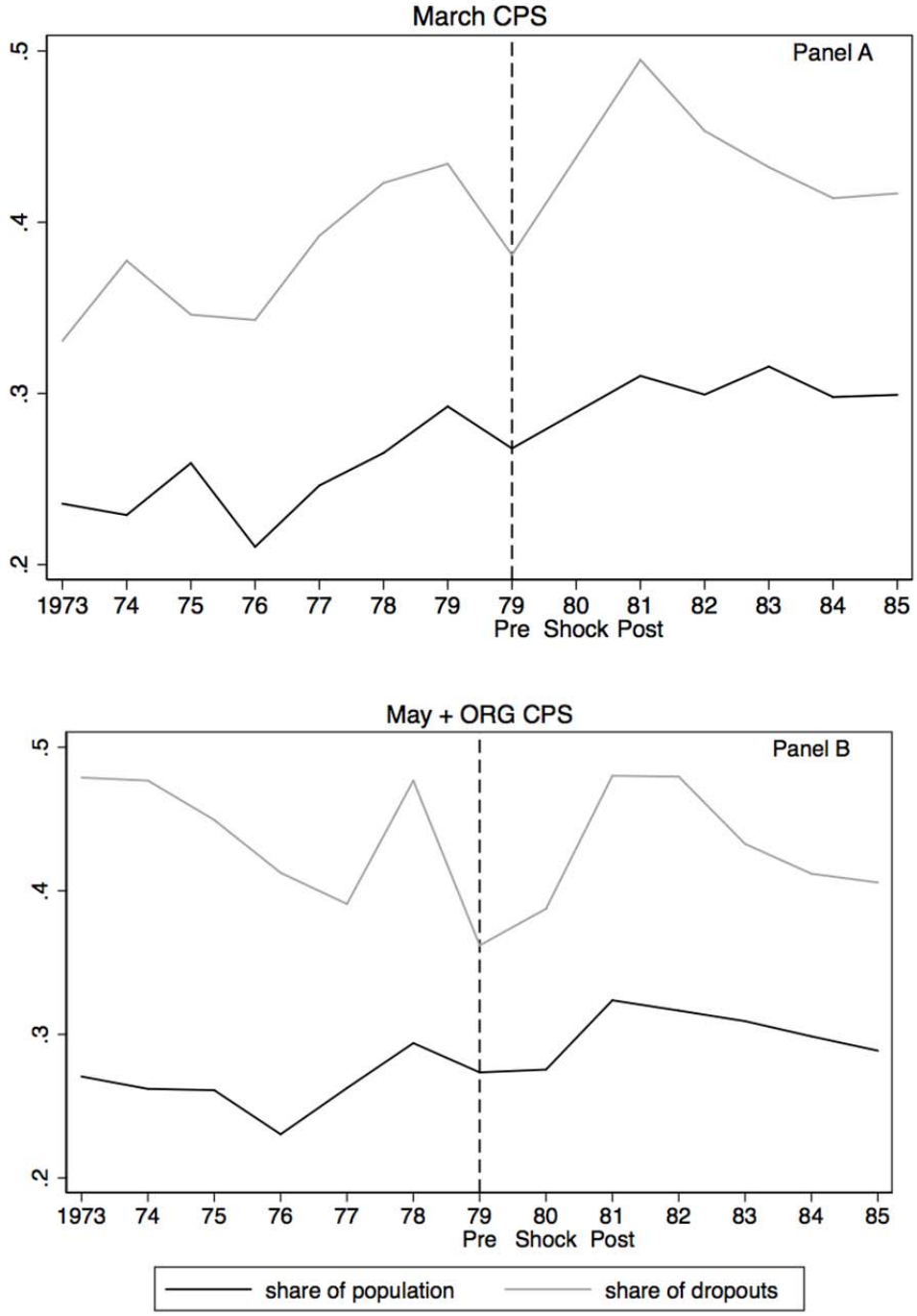
Figure 1:

Distribution of measurement error for the average log wage of high school dropouts across 41 cities



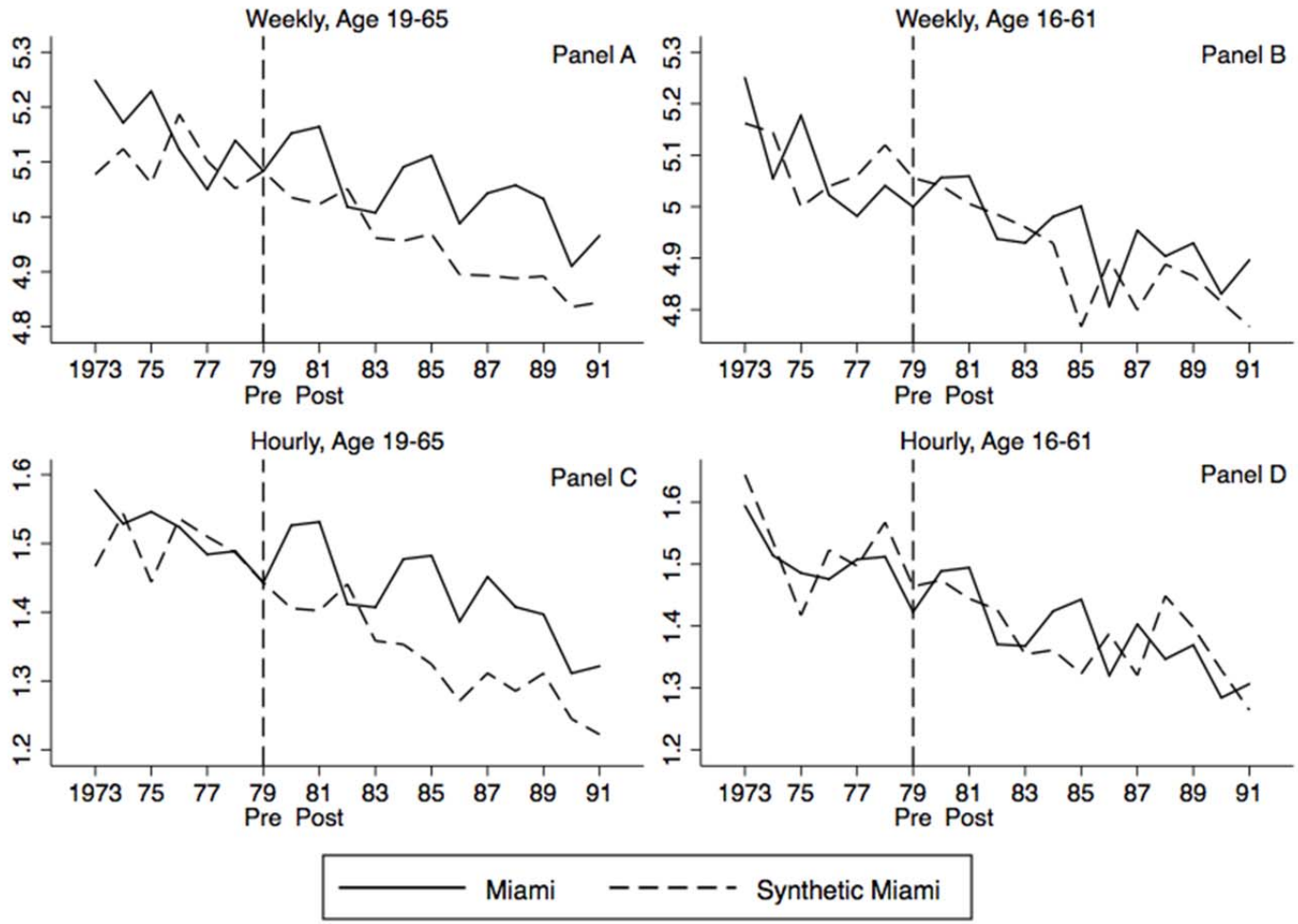
Note: Distribution of the measurement error for the average logarithm of weekly wage across 41 cities in 1979 for the March-CPS (grey line) and the ORG-CPS (black line). The sample is high school dropouts, non-Cuban, age 16-65. The measurement error is defined as the deviation from the same statistics (average log wage) for the same groups and city calculated using the Census 1980.

Figure 2:
Cubans in Miami as share of total and of the high-school-dropout population



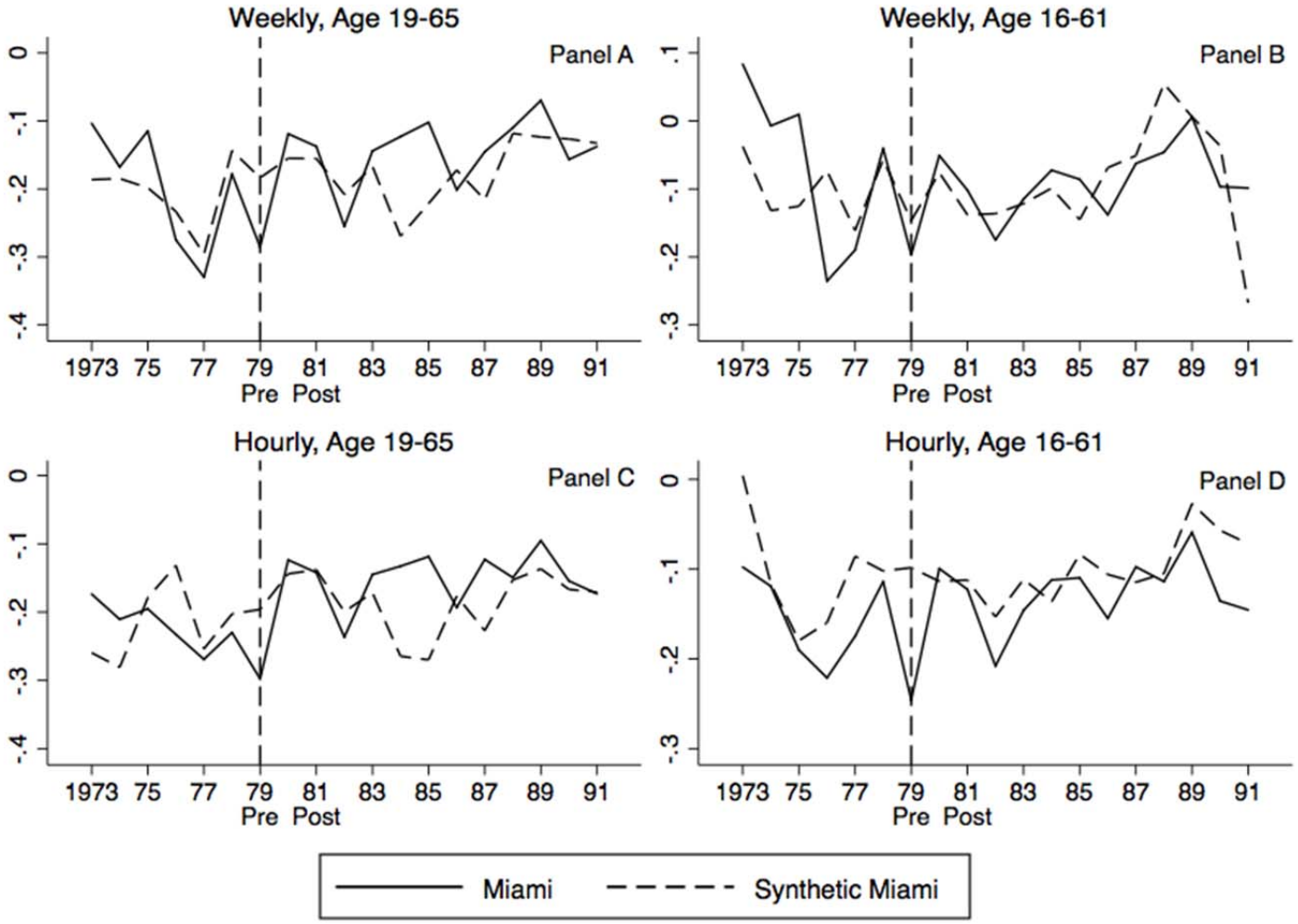
Note: We calculate the share of all those who define themselves as "Cuban" in the ethnicity question of the CPS. The population considered is the total number of individuals between 19 and 65. The high school dropout population is constituted by those who do not have a high school degree in the age range 19 to 65. For the March CPS, we include the figure for March 1980 as "1979-Pre" and we interpolate the figure for "1980-Shock", between 1979-Pre and 1981-Post. The vertical dashed bar corresponds to the last observation before the Mariel Boatlift happened.

Figure 3:
Log Wage Measures, High School dropouts Miami and Synthetic Control, 1973-1991
May + ORG CPS, Preferred Sample



Note: Each Panel shows the outcome variable for Miami (solid line) and Synthetic control (dashed line) in the period 1972-1991. The outcome shown and the sample used are noted in the title of each panel (A through D). Preferred sample means: non-Cubans, high school dropouts, not self-employed, in the labor force. The age group varies upon panel. The vertical line is drawn between the data points of 1979 and 1980 and it identifies the interval in which the Mariel Boatlift took place. The cities with positive weight in the synthetic control are as follows. Panel A: New Orleans, LA 43.2%, New York City, NY, 30.1%, Baltimore, MD 24.9%; Panel B: Cincinnati, OH, 35.5%, New York City, NY, 26.5%, Tampa Bay-St. Petersburg, FL 23.2%, Sacramento, CA 14.7%; Panel C: New Orleans, LA 43.9%, New York City, NY, 29.9%, Baltimore, MD 24.8%; Panel D: Tampa Bay-St. Petersburg, FL 45.5%, Sacramento, CA 44%, Los Angeles, CA 7.5%, Greensboro, SC 3.1%.

Figure 4:
Regression Adjusted Log Wage Measures, High school dropouts, Miami and Synthetic Control, 1972-1991
Regression Adjusted, May + ORG CPS, Preferred Sample

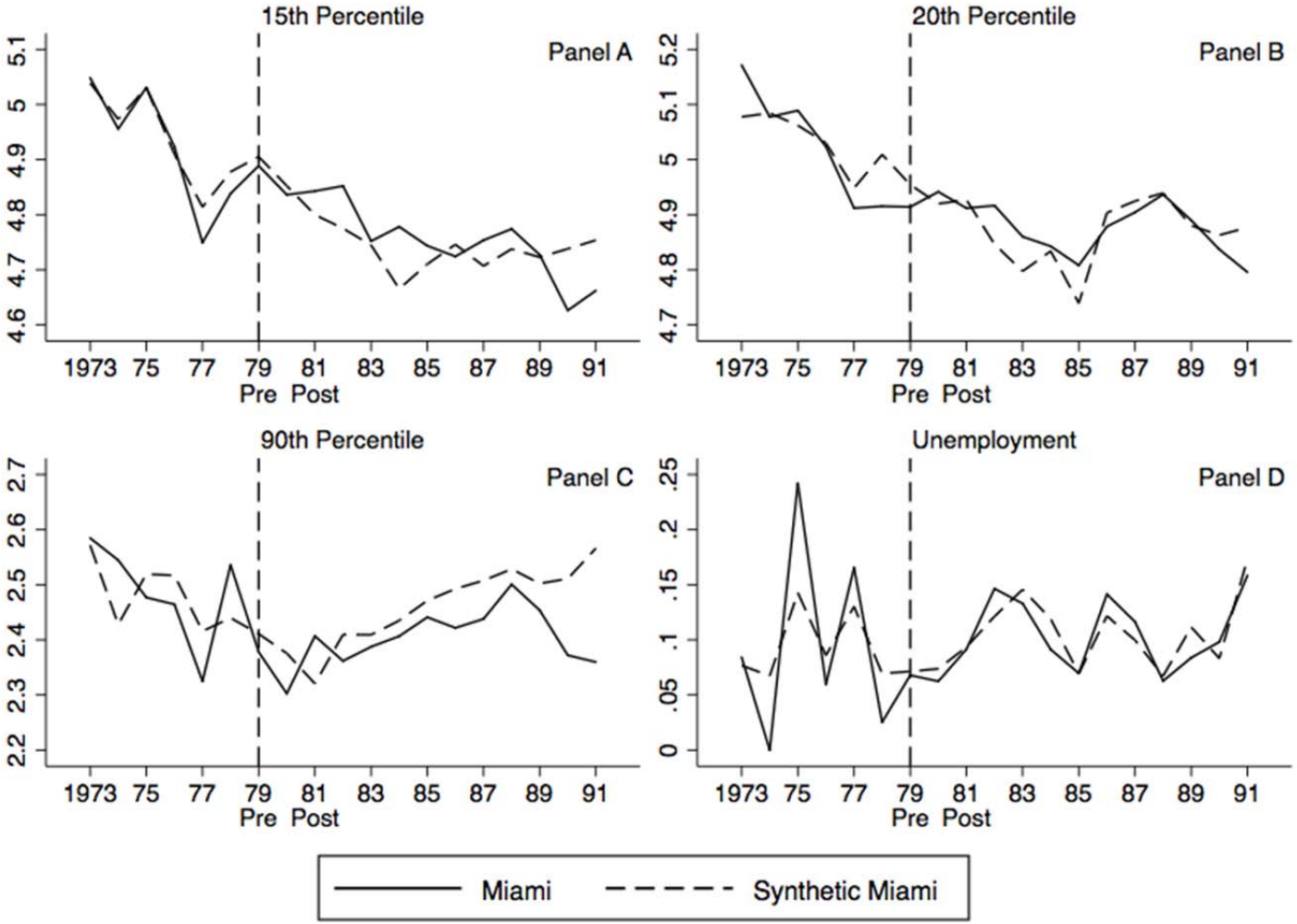


Note: Each Panel shows the outcome variable for Miami (solid line) and Synthetic control (dashed line) in the period 1972-1991. The outcome shown and the sample used are noted in the title of each panel (A through D). Preferred sample means: non-Cubans, high school dropouts, not self-employed in the labor force. The cities with positive weight in the synthetic control are as follows. Panel A: Tampa-St. Petersburg, FL 52%, Birmingham, AL 40.5%, New York City, NY 7.5%; Panel B: Tampa-St. Petersburg, FL 45.4%, New York City, NY 24.6%, Sacramento, CA 20.8%, Albany-Schenectady-Troy, NY 9.2%; Panel C: Tampa-St. Petersburg, FL 54%, Birmingham, AL 46%; Panel D: Greensboro, NC 48.3%, Norfold-Portsmouth, VA 30.5%, Cincinnati, OH 21.2%.

Figure 5:

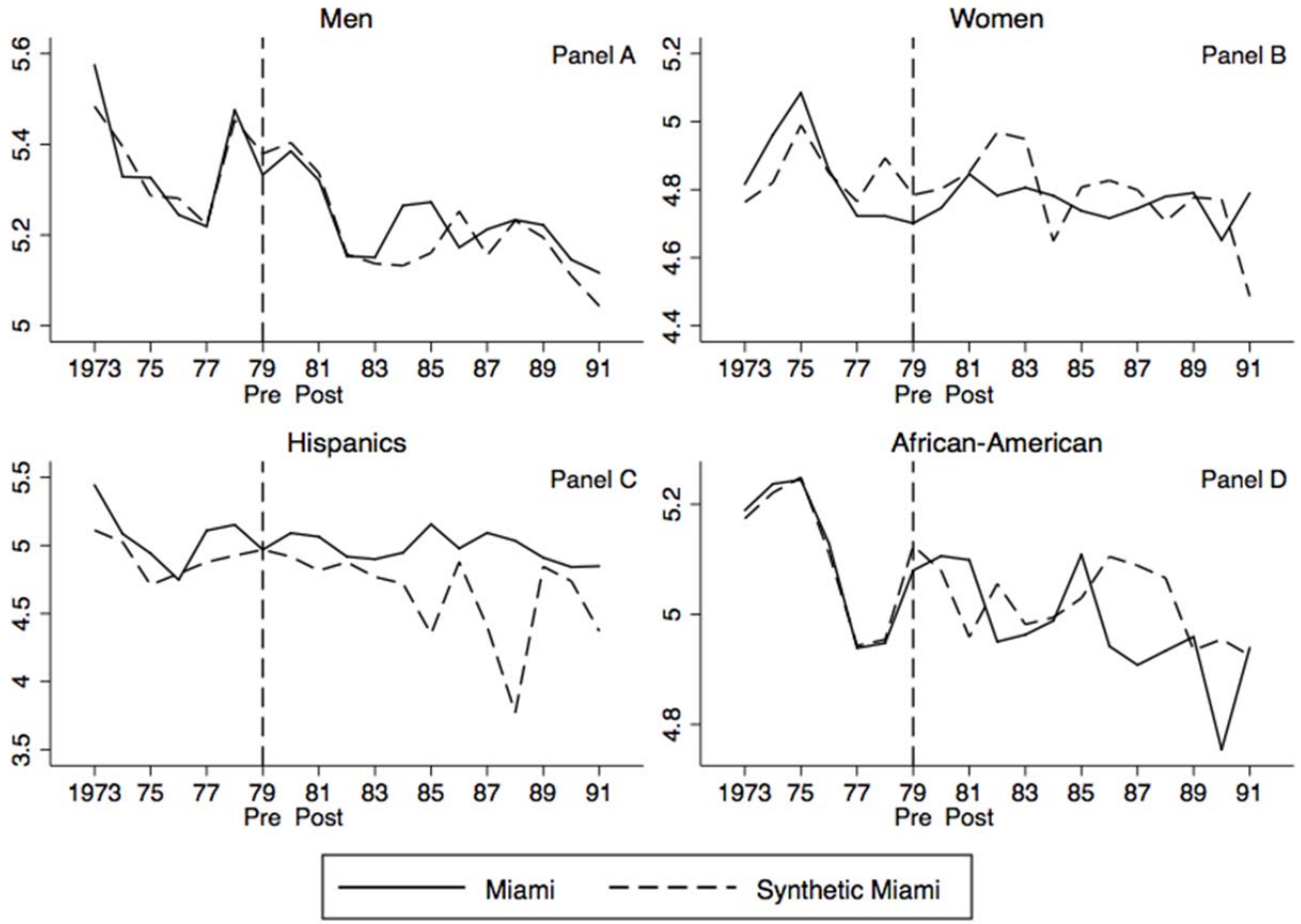
Additional Labor Market Outcomes: Log Weekly Wages at different percentiles and Dropouts' Unemployment

Miami and Synthetic Control, 1973-1991, May + ORG CPS, Preferred Sample



Note: Each Panel shows the outcome variable for Miami (solid line) and Synthetic control (dashed line) in the period 1972-1991. The outcome showed for each Panel and the sample used are noted in the title of each panel. Preferred sample means: non-Cubans, not self-employed individuals, in the labor force, age 19-65 for Panels A, B and C and non-Cubans, high school dropouts not self-employed, in the labor force, age 19-65 for Panel D.-The cities with positive weight in the synthetic control are as follows. Panel A: Birmingham, AL 60.6%, Rochester, NY 28.6%, Nassau-Suffolk, NY 10.4%; Panel B: San Diego, CA 57.7%, Birmingham, AL 28.4%, Nassau-Suffolk, NY 13.8%; Panel C: Tampa-St Petersburg, FL 67.3%, Nassau-Suffolk, NY 32.7%; Panel D: New Orleans, LA 48.4%, New York City, NY 30.9%, Albany-Schenectady-Troy, NY 19.5% Cincinnati, OH 1.1%

Figure 6:
Subsamples of High School Dropouts, Log Weekly Wages, Miami and Synthetic Control,
1973-1991, May + ORG CPS: By Gender and Ethnicity



Note: Each Panel shows the outcome variable for Miami (solid line) and Synthetic control (dashed line) in the period 1972-1991. The outcome variable and the sample are noted in the title of each panel. Preferred sample means: non-Cubans high school dropouts in the labor force, not self-employed, age 19-65. Panel A and B restrict the preferred sample to males and females only, respectively. Panel C and Panel D further restricts the preferred sample to Hispanic and Blacks only. The corresponding minimum and maximum values of the sample sizes are: Panel A (May CPS 23-33; ORG 54-153), Panel B (May CPS 16-20; ORG 26-141), Panel C (May CPS 5-14; ORG 16-97), Panel D (May CPS 14-25; ORG 47-99). The cities with positive weight in the synthetic control are as follows. Panel A: Tampa-St Petersburg, FL 92.6%, Greensboro, SC 6.3%, New York City, NY 1.1%; Panel B: Cincinnati, OH 66.8 %, Pittsburgh, PA 29.4%, Indianapolis, IN 3.8%; Panel C: Sacramento, CA 49.8%, Houston, TX 30.1%, Philadelphia, PA 20.1%; Panel D: Greensboro, NC 39.6%, Cincinnati, OH 19.8%, New York City, NY 15.8%, Seattle, WA, 9.7%, Birmingham, AL, 8.5%, New Orleans, LA 6.7%

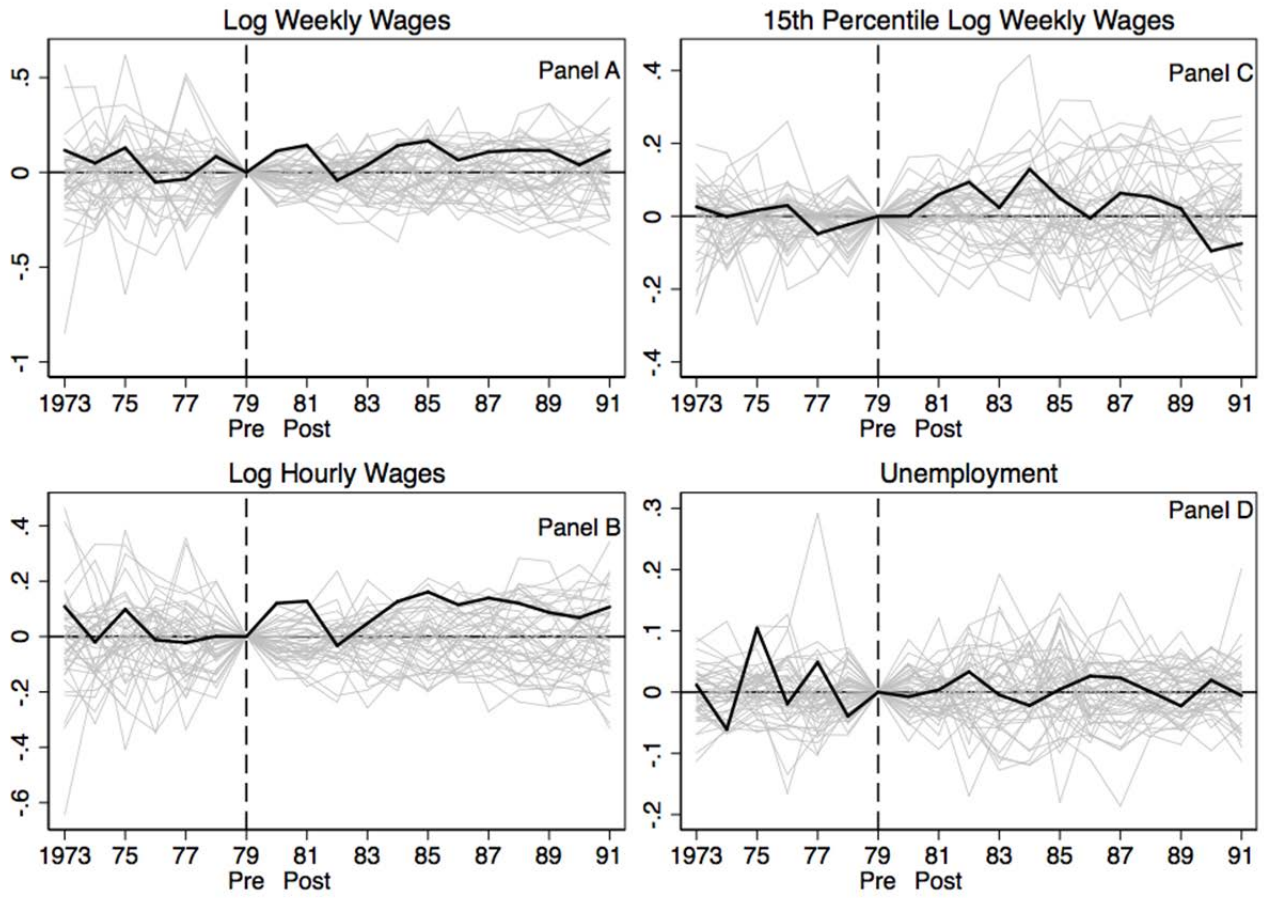
Table 3: Difference in Difference Regressions using Miami and Synthetic Miami observations 1972-1991

Preferred Sample, non-Cuban, high school dropouts 19-65.

	(1) Figure 3 Panel A	(2) Figure 3 Panel C	(3) Figure 5 Panel A	(4) Figure 5 Panel B	(5) Figure 5 Panel C	(6) Figure 5 Panel D
Dependent Variable:	**Ln(Weekly Wages)**	**Ln(Hourly Wages)**	**Ln Wages, 15th percentile**	**Ln Wages, 20th percentile**	**Ln Wages, 90th percentile**	**Unemployment Rate**
Miami X('73-'75)	0.089	0.044	0.037	0.089	0.047	0.028
	(0.055)	(0.045)	(0.043)	(0.055)*	(0.055)	(0.057)
Miami X('76-'78)	-0.042	-0.027	0.006	0.008	-0.018	0.023
	(0.058)	(0.047)	(0.047)	(0.059)	(0.059)	(0.061)
Miami X ('80-'82)	0.043	0.057	0.071	0.077	0.007	0.027
	(0.058)	(0.047)	(0.047)	(0.059)	(0.059)	(0.061)
Miami X ('83-'85)	0.069	0.088	0.093	0.099	-0.020	0.010
	(0.055)	(0.045) *	(0.043)*	(0.055)*	(0.055)	(0.057)
Miami X ('86-'88)	0.111	0.114	0.064	0.040	-0.039	0.033
	(0.055) *	(0.045)*	(0.044)	(0.055)	(0.055)	(0.057)
Miami X ('89-'91)	0.080	0.066	-0.031	0.022	-0.115*	0.014
	(0.055)	(0.045)	(0.044)	(0.056)	(0.056)	(0.057)
Observations	38	38	38	38	38	38

Note: Each column represents a regression of annual observations for Miami and the corresponding synthetic counterfactual between 1973 and 1991. Each specification includes vectors of city and year bins dummies. Each period dummies extends for three years. The bin for 1979 is excluded so as to standardize the value of that interaction to 0. The interaction coefficients between a dummy variable for Miami and a corresponding year bin are reported. The method of estimation is FGLS with AR1 process for the error term assumed * p<0.05; ** p<0.01; *** p<0.001.

Figure 7:
Inference, Simulated permutations, 44 metropolitan areas
Actual – Synthetic difference, All Metro Areas, May + ORG, Age 19-65



Note: Each graph reports deviations between synthetic control and treated group, assuming a treatment in 1980, for 44 metropolitan areas. The bold line represents Miami. Panel A shows the graph for the logarithm of weekly wages, Panel B shows it for the logarithm of hourly wages. Panel C for the 15th percentile of log weekly wages and Panel D the unemployment rate. The sample in Panel A, B and D includes non-Cuban, high school dropouts, 19-65 years old from the May and ORG CPS. In panel C the 15th percentile is calculated on all non-Cuban workers between 19 and 65 years old from the May and ORG CPS.

Table 4:

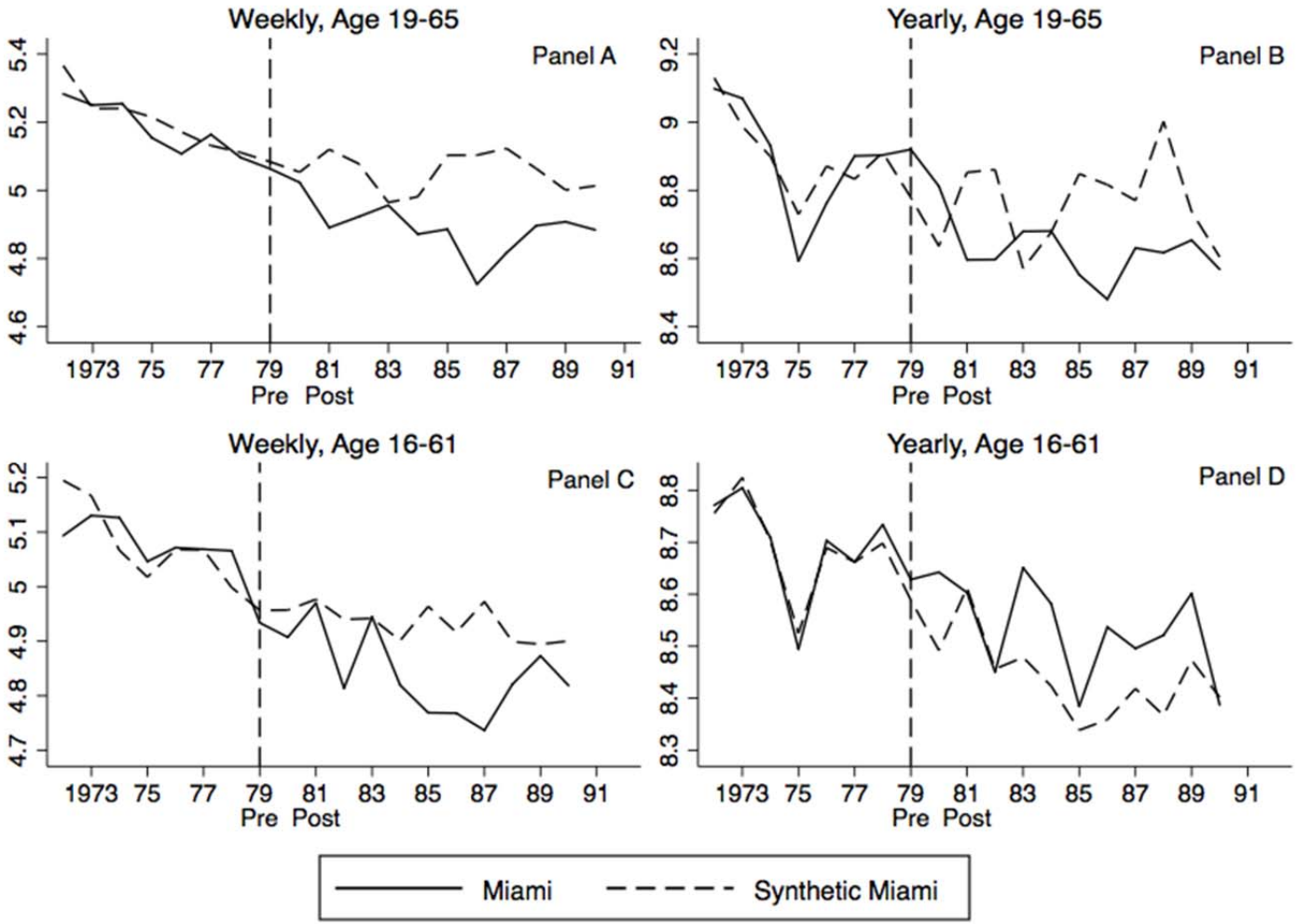
Distribution of 80-82 deviations of city outcomes from their synthetic control, relative to pre-1980 deviations

Permutations of 43 metropolitan areas

	Outcome variable			
	(1) Figure 7 Panel A	**(2) Figure 7 Panel B**	**(3) Figure 7 Panel C**	**(4) Figure 7 Panel D**
Dependent Variable	**Ln Weekly Wages**	**Ln Hourly Wages**	**Ln Weekly 15th Percentile**	**Unemployment Rate**
Analysis relative to Pre-period 72-79				
Ratio of Post-Pre MSPE	1.88	3.34	6.08	0.15
Rank, lowest to highest	39/42	41/43	41/43	2/44
P-value, one tailed test $P(\Delta>\Delta_{MIAMI})$	0.07	0.05	0.05	0.96
Analysis relative to Pre-period 77-79				
Ratio of Post-Pre MSPE	4.28	65.79	4.26	0.30
Rank, lowest to highest	38/42	43/43	23/43	4/44
P-value, one tailed test $P(\Delta>\Delta_{MIAMI})$	0.10	0.00	0.47	0.91

Note: The "Ratio of Post-Pre" equals the absolute value of the ratio of the average Miami-Synthetic control square deviation in 80-82 divided the average Miami-Synthetic control square deviation in the pre-period. In the upper panel the pre-period is the whole period 72-79, in the lower panel it is the last two years 77-79. We also calculate the same ratio for each city in the donor pool and construct a distribution of the 32 ratio statistics. The "rank" entry shows were Miami ranks in the distribution of 44 values (bottom to top) the p-value is a test of the probability that a random draw from the donor pool takes a higher than Miami value. We follow Ferman and Pinto (2015) who show that dividing the post-RMSE by the pre-RMSE corrects for heteroskedasticity in this setting. See their paper for more details.

Figure 8
Log Wage Measures, Miami and Synthetic Control,
1972-1991, March CPS, Preferred Sample

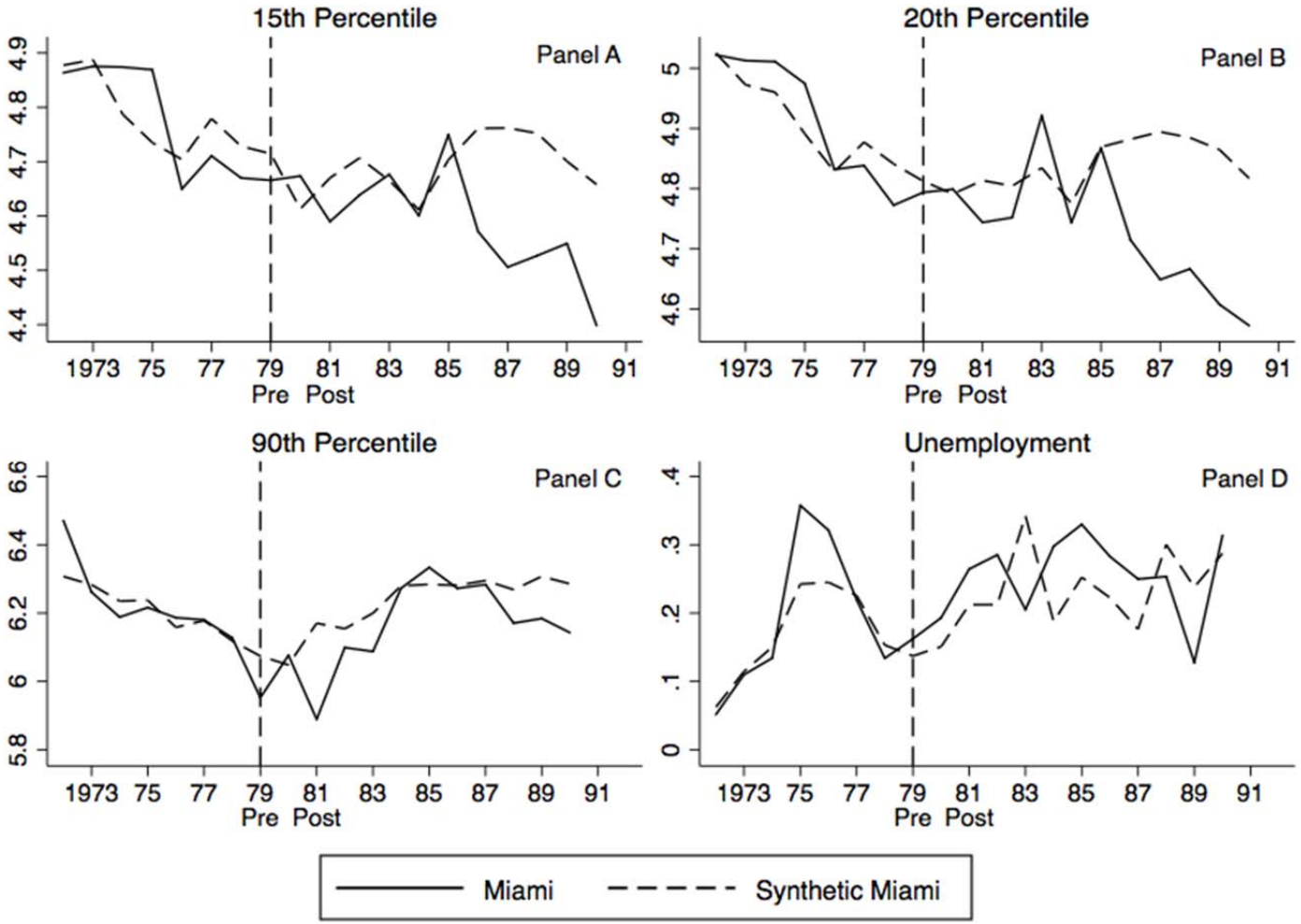


Note: Each Panel shows the outcome variable for Miami (solid line) and Synthetic control (dashed line) in the period 1972-1991. The outcome shown and the sample used are noted in the title of each panel (A through D). Preferred sample means: non-Cubans, high school dropouts, not self-employed, in the labor force, age 19-65. The cities with positive weight in the synthetic control are as follows. Panel A: Cincinnati, OH 37.1%, Atlanta, GA 32.2%, New Orleans, LA 18.5%, Tampa-St Petersburg, FL 12%; Panel B: Los Angeles, CA 80.1%, Dallas, TX 14.4%, New York City, NY 5.5%; Panel C: Atlanta, GA 61.2%, New Orleans, LA 30.6%, Cincinnati, OH 8.1%; Panel D: Houston, TX 50.8%, New York City, NY 24.9%, San Diego, CA 24.2%

Figure 9

Other Labor Market Outcomes: Log Weekly Wages and Unemployment,

Miami and Synthetic Control, 1972-1991, March CPS, Preferred Sample



Note: Each Panel shows the outcome variable for Miami (solid line) and Synthetic control (dashed line) in the period 1972-1991. The outcome showed for each Panel and the sample used are noted in the title of each panel. Preferred sample means: non-Cubans, high school dropouts, not self-employed, in the labor force, age 19-65. The cities with positive weight in the synthetic control are as follows. Panel A: New York City, NY 51.8%, San Diego, CA 38.2%, Riverside-San Bernadino, CA 10%; Panel B: San Diego, CA 49.9%, New York City, NY 39.7%, Tampa-St. Petersburg, FL 5.9%, Los Angeles-Long Beach, CA, 4.4%; Panel C: New York City, NY 59.5%, Tampa-St. Petersburg, FL 40.5%; Panel D: New Orleans, LA 39.3%, Denver-Boulder-Longmont, CO 30.4%, Tampa-St. Petersburg, FL 30.3%.

Table 5
Regression analysis for the March Sample and Control Group: Miami and Control 1972-1989

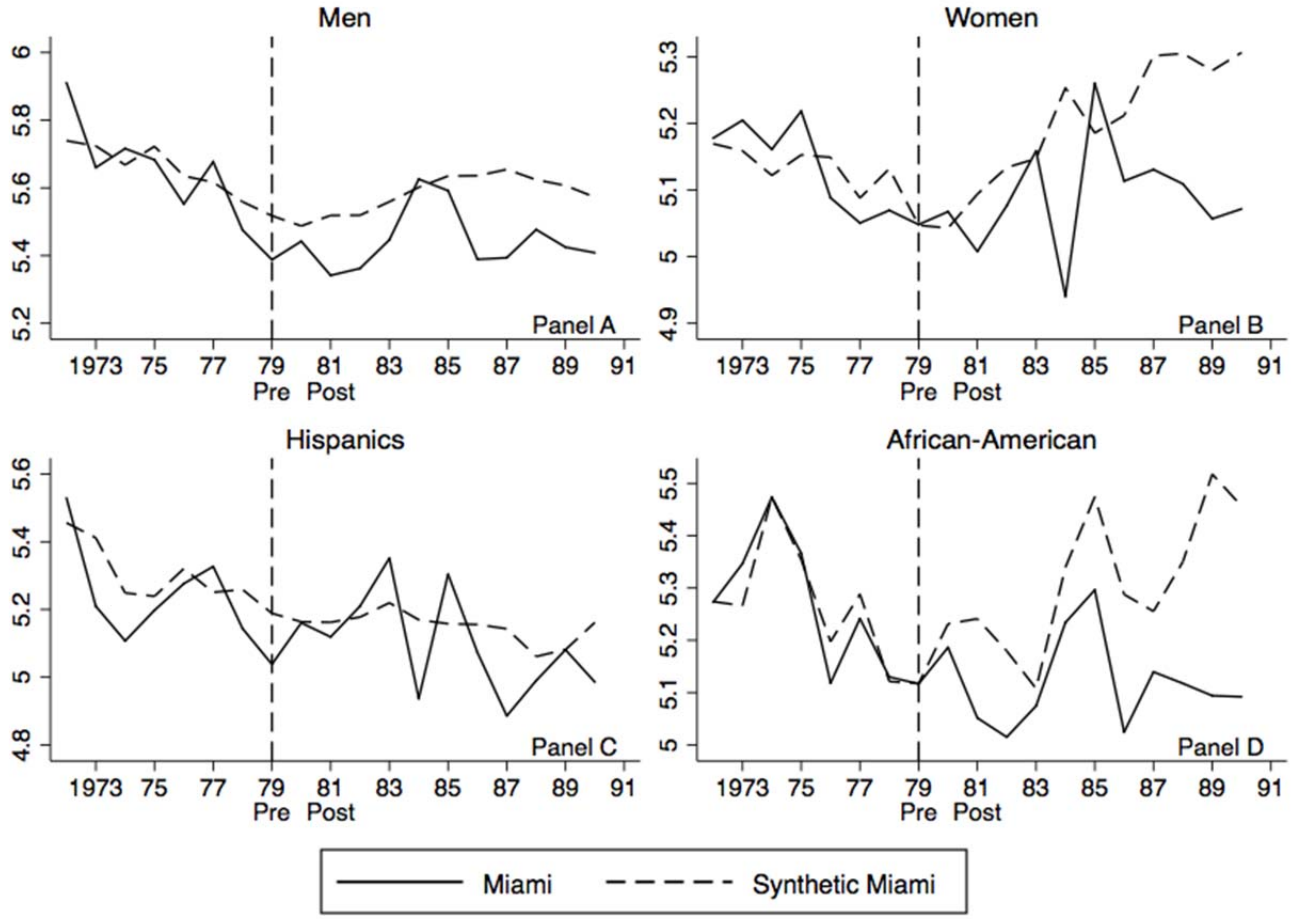
	Dependent variable, Sample, Control					
	(1) Figure 8 Panel A	(2) Figure 8 Panel C	(3) Figure 9 Panel A	(4) Figure 9 Panel B	(5) Figure 9 Panel C	(6) Figure 9 Panel D
Dependent Variable	**Ln Weekly Wages**	**Ln Weekly Wages**	**Ln Weekly, 15th perc.**	**Ln Weekly, 20th perc.**	**Ln Weekly, 90th perc.**	**Unemployment Rate**
Miami X (72-75)	-0.021	-0.134	0.078	0.062	0.101	-0.005
	(0.067)	(0.170)	(0.059)	(0.045)	(0.087)	(0.099)
Miami X (76-78)	-0.008	-0.135	-0.028	-0.016	0.096	-0.009
	(0.071)	(0.172)	(0.063)	(0.045)	(0.094)	(0.102)
Miami X (80-82)	-0.131	-0.235	-0.004	-0.019	-0.033	0.031
	(0.071)	(0.172)	(0.063)	(0.049)	(0.094)	(0.102)
Miami X (83-85)	-0.108	-0.177	0.044	0.033	0.071	-0.010
	(0.069)	(0.175)	(0.060)	(0.046)	(0.090)	(0.102)
Miami X (86-88)	-0.281	-0.403*	-0.191	-0.198	0.055	0.003
	(0.069)***	(0.175)	(0.060)**	(0.046)***	(0.090)	(0.102)
Miami X (89-91)	-0.097	-0.175	-0.175	-0.235	-0.051	-0.068
	(0.073)	(0.186)	(0.064)**	(0.049)***	(0.95)	(0.108)
Observations	**40**	**40**	**40**	**40**	**40**	**40**

Note: Each column represents a regression of annual observations for Miami and the corresponding Borjas (2015) control between 1972 (1973 for May+ORG sample) and 1991. Each specification includes a Miami fixed effect, period dummies and the interaction between Miami dummy and period effects. Each period dummies extends for three years, except for the beginning of the period in which 4 years are included in each dummy. The bin for 1979 is excluded so as to standardize the value of that interaction to 0. The interaction coefficients between a dummy variable for Miami and a corresponding year bin are reported. The method of estimation is FGLS with AR1 process for the error term assumed. * p<0.05; ** p<0.01; *** p<0.001.

Figure 10

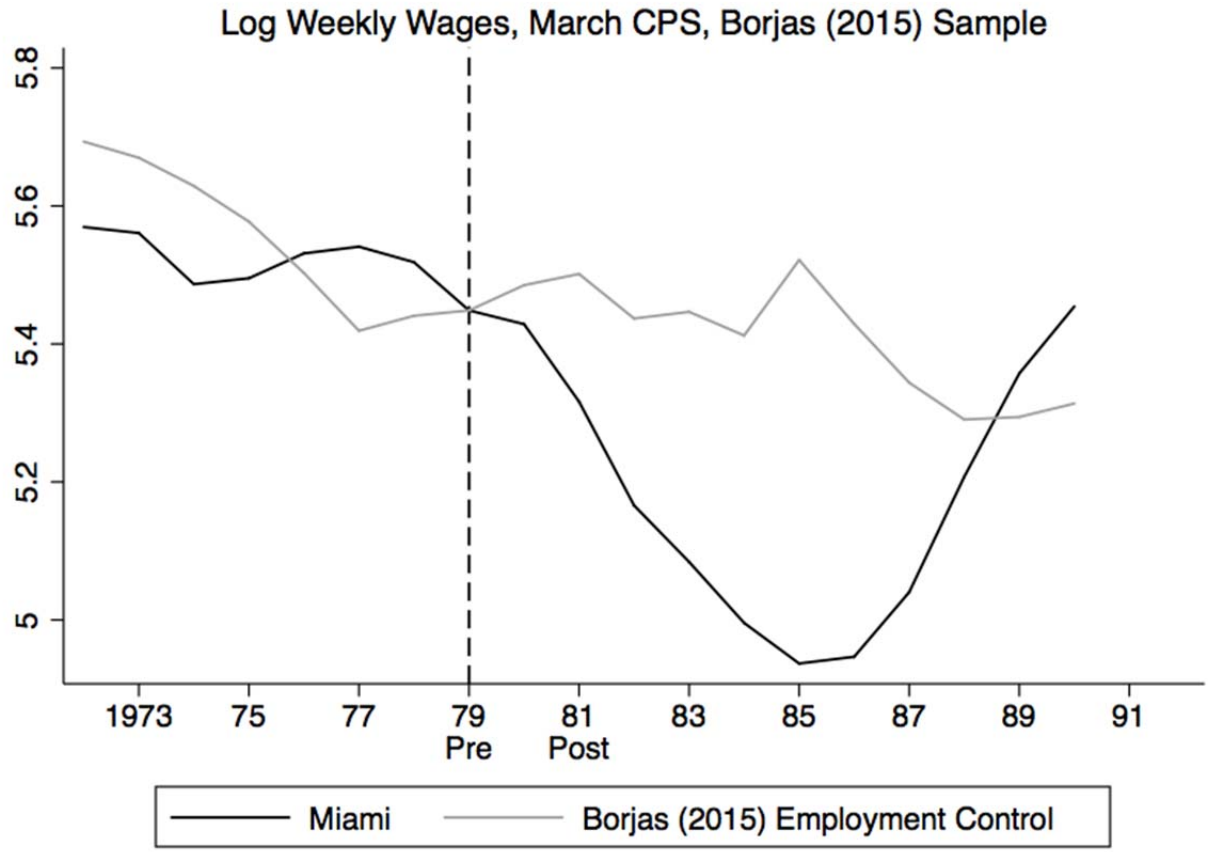
Subsamples of High School Dropouts: By Gender and Ethnicity, Miami and Synthetic Control,

1972-1991, Log Weekly Wages, March CPS, Preferred Sample



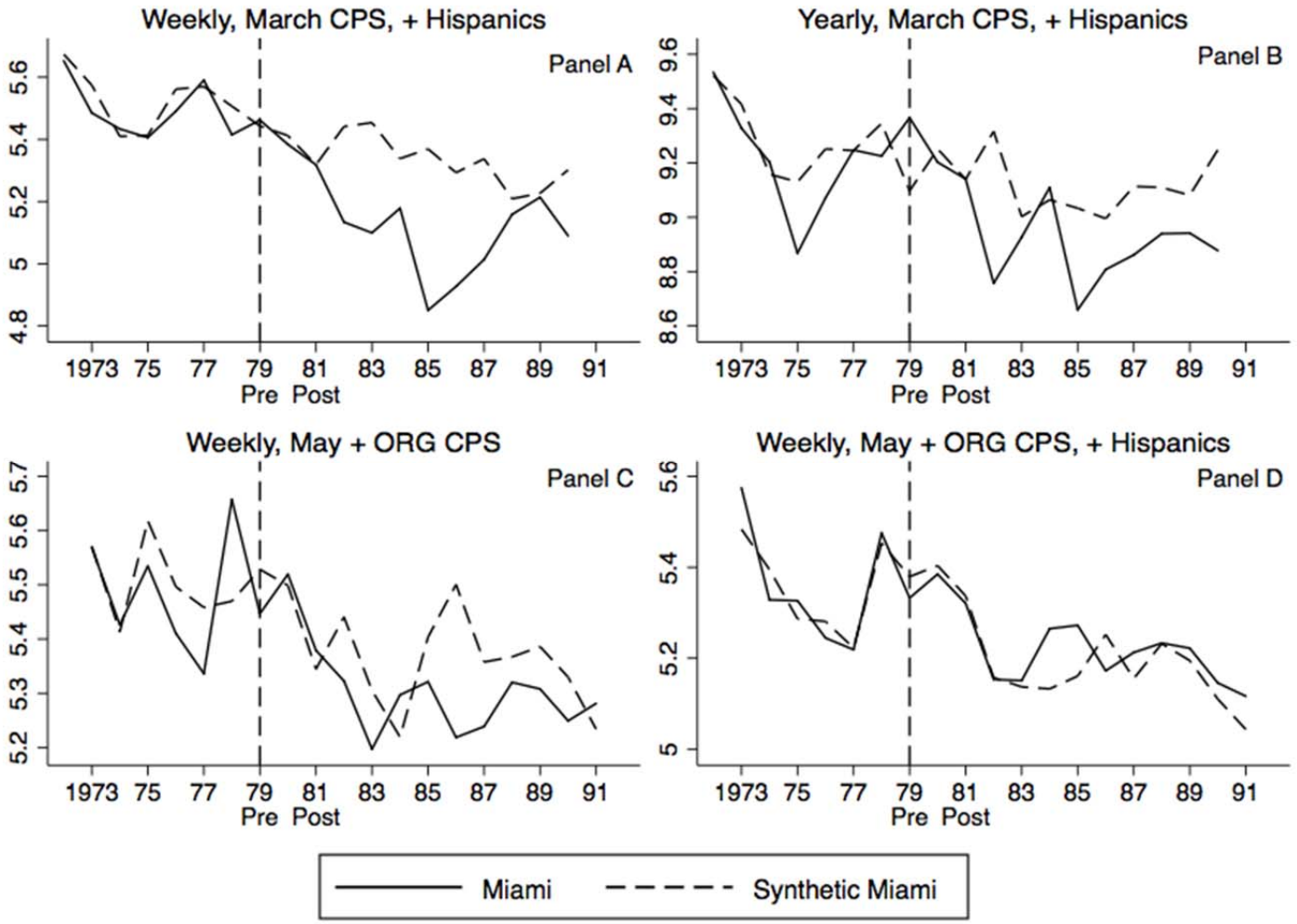
Note: Each Panel shows the outcome variable for Miami (solid line) and Synthetic control (dashed line) in the period 1972-1991. The outcome showed for each Panel and the sample used are noted in the title of each panel. The corresponding minimum and maximum values of the sample sizes are: Panel A (31-56), Panel B (24-49), Panel C (15-55), Panel D (19-37). Preferred sample means: high school dropouts, not self-employed, in the labor force, age 19-65. The cities with positive weight in the synthetic control are as follows. Panel A: New York City, NY 71.3%, Tampa-St. Petersburg, FL 28.7%; Panel B: Tampa-St. Petersburg, FL 47.8%, San Jose, CA 28.5%, New York City, 13.8%, Washington, DC/MD/VA 9.9%; Panel C: Los Angeles-Long Beach, CA 74.9%, Dallas, TX 25.1%; Panel D: Cincinnati, OH 58.9%, Riverside-San Bernadino, CA 29.3%, Los Angeles-Long Beach, CA 10.4%, Tampa-St. Petersburg, FL 1.4%.

Figure 11
Borjas (2015) Results
Our replication extended back to 1972.



Note: This reproduces Figure 3A in Borjas (2015). The dark line represents the log weekly wages trajectory of Miami and the grey one does the same for the employment placebo of consisting of Anaheim-Santa Ana, CA, Nassau-Suffolk, NY, Rochester, NY, and San Jose, CA.

Figure 12
Sensitivity of the Borjas (2015a) Results
Sample, Smoothing and Variable Definition



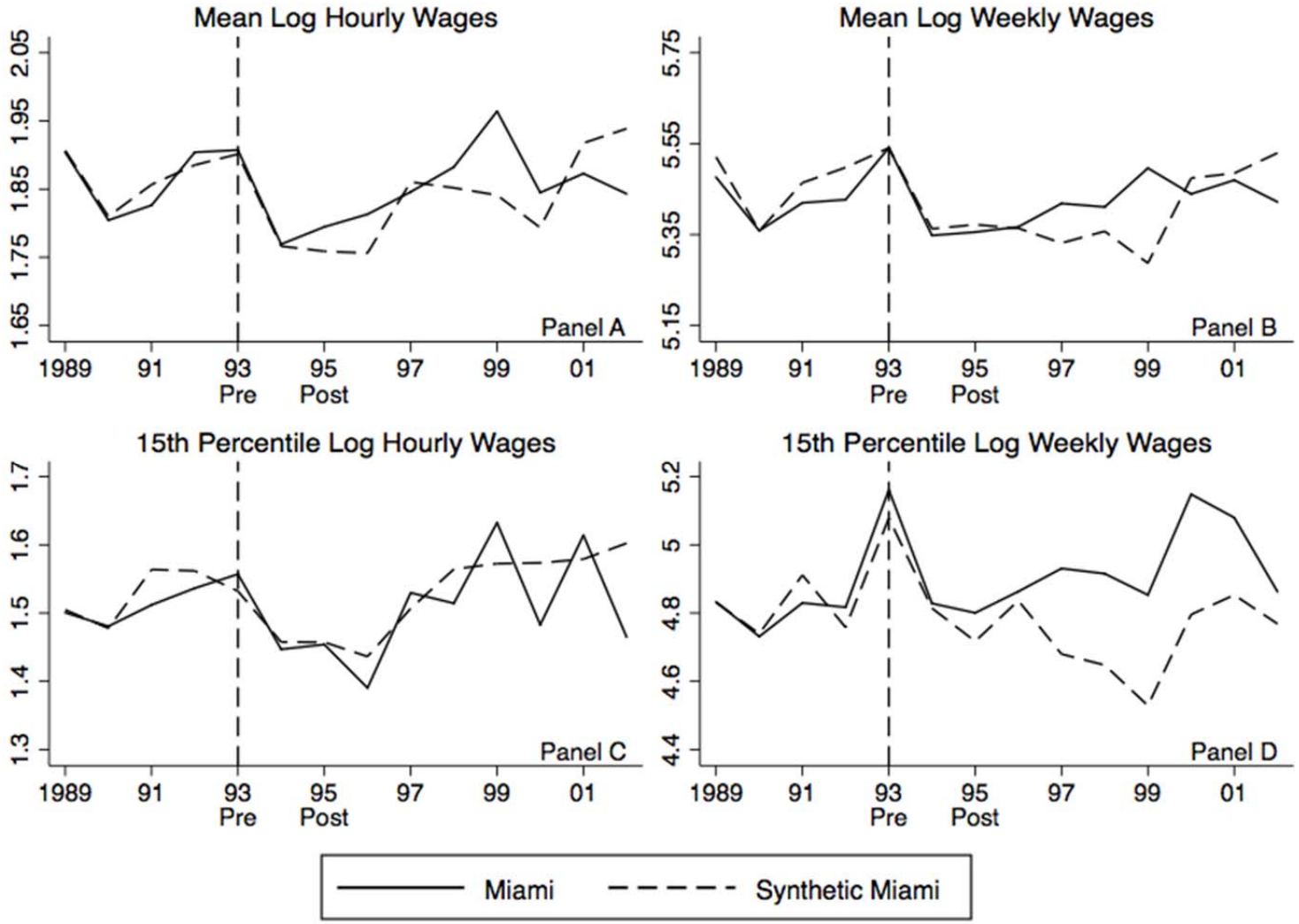
Note: Panel A reproduces Figure 3A in Borjas (2015) with yearly wages. The logarithm of weekly wages for males, high school dropouts, non-Hispanic age 25-59, smoothed with a 3-years moving average in the March CPS sample for Miami and the "Employment control" are shown. Panel B extends the sample to non-Cuban, males 19-65, drops the smoothing and introduces the synthetic control. Panel C goes back to Borjas original sample, removes the smoothing and uses ORG-CPS data. Panel D uses the same group definition of Panel B, on ORG data. The cities with positive weight in the synthetic control are as follows. Panel A: Dallas, TX 67.8%, Baltimore, MD 18.8%, New York City, NY 13.5%; Panel B: Dallas, TX 60.4%, Atlanta, GA 39.6%; Panel C: Greensboro, NC 64.8%, Tampa-St. Petersburg, FL 35.2%; Panel D: Tampa-St. Petersburg, FL 92.6%, Greensboro, NC 7.4%

Table 6: Regression analysis for the Borjas Sample and Variations: Miami and Control 1972-1989

	Dependent variable, Sample, Control				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
	Figure 11	**Figure 12 Panel A**	**Figure 12 Panel B**	**Figure 12 Panel C**	**Figure 12 Panel D**
Dependent Variable	**Ln Weekly Wages**	**Ln Weekly Wages**	**Ln Yearly Wages**	**Ln Weekly Wages**	**Ln Weekly Wages**
Miami X (72-75)	-0.123	-0.042	-0.334	-0.013	0.065
	(0.088)	(0.131)	(0.207)	(0.099)	(0.125)
Miami X (76-78)	0.070	-0.066	-0.358	-0.008	0.040
	(0.092)	(0.135)	(0.210)	(0.105)	(0.126)
Miami X (80-82)	-0.177	-0.131	-0.469*	-0.012	0.032
	(0.092)	(0.135)	(0.210)	(0.105)	(0.126)
Miami X (83-85)	-0.467***	-0.362**	-0.391	-0.031	0.131
	(0.091)	(0.135)	(0.214)	(0.099)	(0.124)
Miami X (86-88)	-0.301***	-0.265	-0.459*	-0.139	0.040
	(0.091)	(0.135)	(0.214)	(0.100)	(0.124)
Miami X (89-91)	0.101	-0.132	-0.516*	-0.021	0.091
	(0.096)	(0.143)	(0.227)	(0.100)	(0.124)
R^2	**0.914**	**0.787**	**0.658**	**0.746**	**0.658**
Observations	**40**	**40**	**40**	**38**	**38**

Note: Each column represents a regression of annual observations for Miami and the corresponding Borjas (2015) control between 1972 (1973 for May+ORG sample) and 1991. Each specification includes a Miami fixed effect, period dummies and the interaction between Miami dummy and period effects. Each period dummies extend for three years, except for the beginning of the period in which 4 years are included in each dummy. The bin for 1979 is excluded so as to standardize the value of that interaction to 0. The interaction coefficients between a dummy variable for Miami and a corresponding year bin are reported. The method of estimation is FGLS with AR1 process for the error term assumed. * p<0.05; ** p<0.01; *** p<0.001.

Figure 13: The Non-Event of 1994, Log Wage Measures, Miami and Synthetic Control, 1989-2001



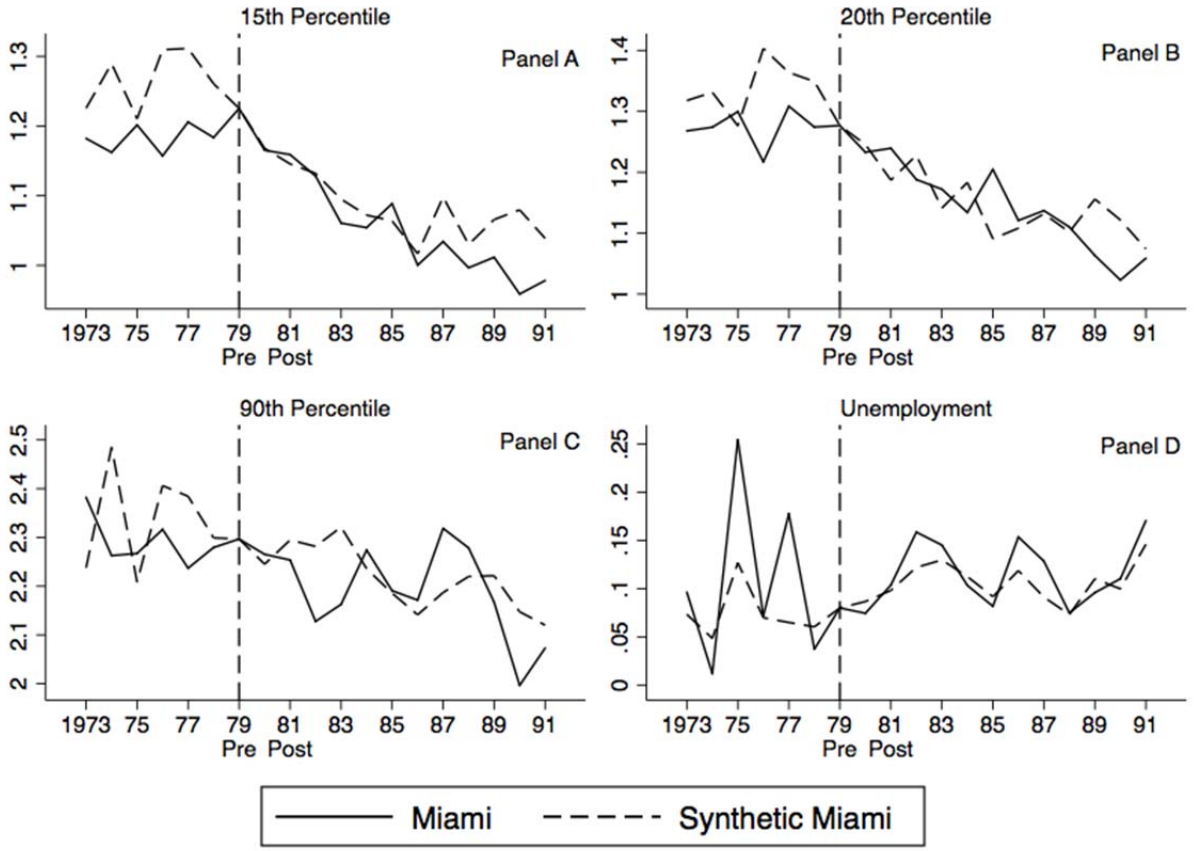
Note: Each Panel shows the outcome variable for Miami (solid line) and for the synthetic control (dashed line) in the period 1989-2001. The variable and sample are noted in the title of each panel. Preferred sample means: non-Cubans, high school dropouts in the labor force, age 19-61. The cities with positive weight in the synthetic control are as follows. Panel A: Bakersfield, CA 35.4%, New Orleans, LA 21.8%, El Paso, TX 16.8%, Jackson, MS 13.1%, Visalia-Tulare-Porterville, CA 12.9%; Panel B: Jersey City, NJ 31.8%, Sioux Falls, SD 30.4%, Bakersfield 27.6%, San Antonio, TX 10.2%; Panel C: Bakersfield, CA 38.4%, Lakeland-Winter Haven, FL 28.6%; New Orleans, LA 17.5%, El Paso, TX 10.9%, Jackson, MS 4.4%; Panel D: Jersey City, NJ 50.5%, Sioux Falls, SD 32.1%, San Antonio, TX 17.4%.

Tables and Figures Appendix

Figure A1: Fixed Synthetic Control
Log Hourly Wages at different percentiles and Unemployment

Miami and Synthetic Control, 1973-1991, May + ORG CPS, Preferred Sample

The control pool is restricted to: New Orleans, Baltimore, New York City. Same as in Figure3 A, C

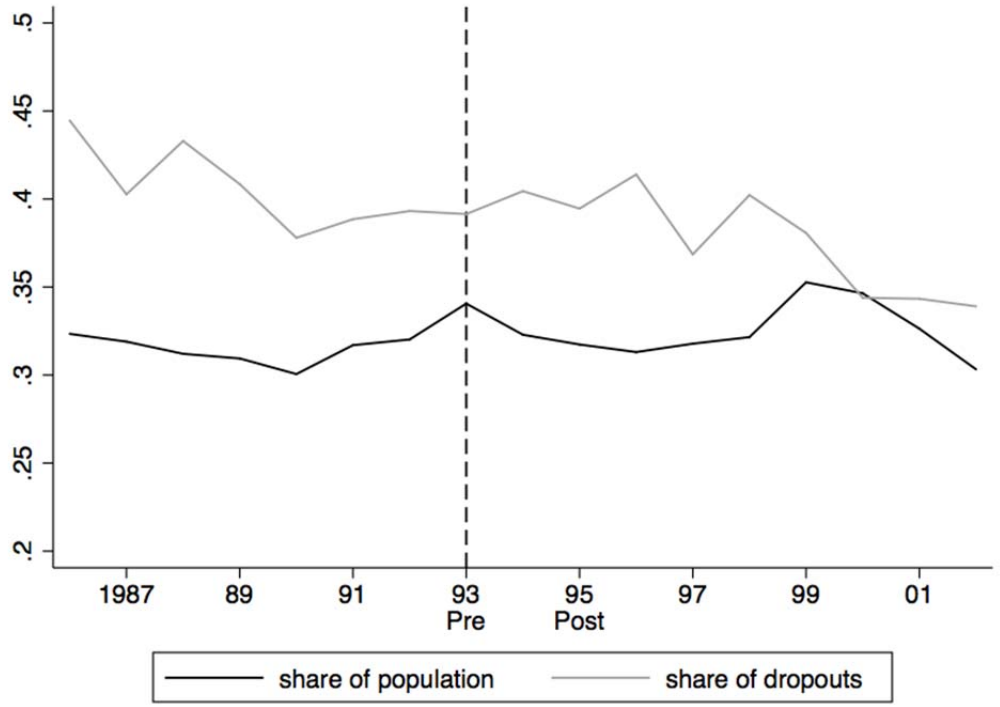


Note: Each Panel shows the outcome variable for Miami (solid line) and Synthetic control (dashed line) in the period 1972-1991. The outcome shown and the sample used are noted in the title of each panel (A through D). Preferred sample means: non-Cubans, high school dropouts, not self-employed, in the labor force. The age group varies upon panel. The vertical line is drawn between the data points of 1979 and 1980 and it identifies the interval in which the Mariel Boatlift took place. The synthetic control is always kept fixed and the cities with positive weights are: New Orleans, LA 44%, New York City, NY 31%, Baltimore, MD 25%.

Figure A2:

The Non-Event of 1994

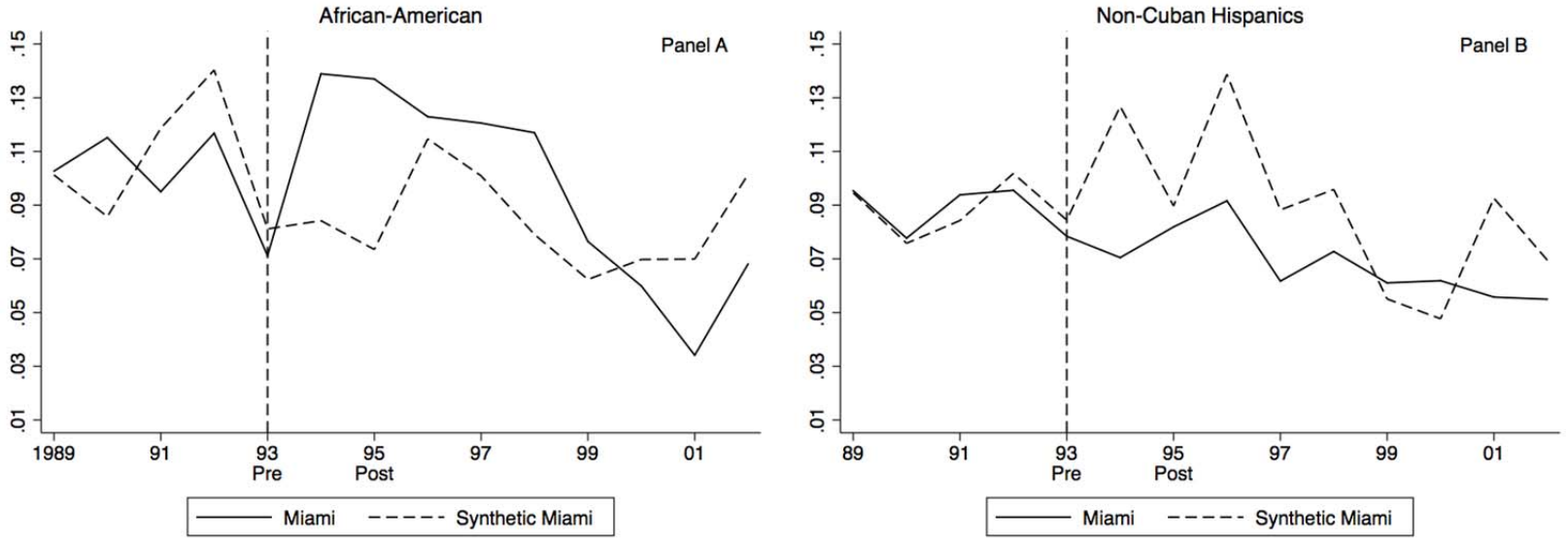
Cubans in Miami as a share of population and of the high school dropout population, ORG CPS



Note: We calculate the share of Cuban people as all those who define themselves as "Cuban" in the ethnicity question of the CPS. The population considered is the total number of individuals between 19 and 65. The high school dropout population is constituted by those who do not have a high school degree between 16 and 61. The vertical dashed bar is drawn at 1993.

Figure A3:

The Non-Event of 1994
Unemployment in Miami and in the Synthetic control



Note: Each Panel shows the behavior of the outcome variable for Miami (solid line) and for the synthetic control (dashed line) in the period 1989-2001. The vertical line corresponds to year 1993, immediately before the non-event of 1994. The variables are noted in the title of each panel. The sample here includes all non-Cubans, in the labor force, age 19-61 either of Black ethnicity (Panel A) or of Hispanic ethnicity (Panel B).