



HARVARD Kennedy School
JOHN F. KENNEDY SCHOOL OF GOVERNMENT

Still More on Mariel: The Role of Race

Faculty Research Working Paper Series

George J. Borjas
Harvard Kennedy School

August 2017

RWP17-029

Visit the **HKS Faculty Research Working Paper Series** at:
https://www.hks.harvard.edu/research-insights/publications?f%5B0%5D=publication_types%3A121

The views expressed in the **HKS Faculty Research Working Paper Series** are those of the author(s) and do not necessarily reflect those of the John F. Kennedy School of Government or of Harvard University. Faculty Research Working Papers have not undergone formal review and approval. Such papers are included in this series to elicit feedback and to encourage debate on important public policy challenges. Copyright belongs to the author(s). Papers may be downloaded for personal use only.

Still More On Mariel: The Role of Race
George J. Borjas
NBER Working Paper No. 23504
June 2017
JEL No. J0,J61

ABSTRACT

Card's (1990) study of the Mariel supply shock remains an important cornerstone of both the literature that measures the labor market impact of immigration, and of the "stylized fact" that immigration might not have much impact on the wage of workers in a receiving country. My recent reappraisal of the Mariel evidence (Borjas, 2017) revealed that the wage of low-skill workers in Miami declined substantially in the years after Mariel, and has already encouraged a number of re-examinations. Most recently, Clemens and Hunt (2017) argue that a data quirk in the CPS implies that wage trends in the sample of non-Hispanic prime-age men examined in my paper does not correctly represent what happened to wages in post-Mariel Miami. Specifically, there was a substantial increase in the black share of Miami's low-skill workforce in the relevant period (particularly between the 1979 and 1980 survey years of the March CPS). Because African-American men earn less than white men, this increase in the black share would spuriously produce a drop in the average low-skill wage in Miami. This paper examines the robustness of the evidence presented in my original paper to statistical adjustments that control for the increasing number of black men in Miami's low-skill workforce. The evidence consistently indicates that the race-adjusted low-skill wage in Miami fell significantly relative to the wage in other labor markets shortly after 1980 before fully recovering by 1990.

George J. Borjas
Harvard Kennedy School
79 JFK Street
Cambridge, MA 02138
and NBER
gborjas@harvard.edu

For acknowledgments and financial disclosure information, see:
<http://www.nber.org/papers/w23504>

Still More on Mariel: The Role of Race

George J. Borjas*

Since its public release in September 2015, my reappraisal of the wage impact of the Mariel supply shock (Borjas, 2017) has attracted considerable attention. My study showed that the long-believed “stylized fact” that the *Marielitos* did not have a wage impact on Miami’s workers, first reported in Card (1990), did not truly represent what the data actually said. Instead, my reappraisal indicated that the wage of Miami’s high school dropouts, the group most affected by the low-skill *Marielitos*, had declined substantially after 1980 but had fully recovered by 1990.

Not surprisingly, my reappraisal has itself inspired a number of re-reexaminations, including Peri and Yasenov (2015) and Clemens and Hunt (2017). Peri and Yasenov argued that the problem with my original analysis resulted from the rules I used for sample selection. I examined wage trends using a sample of non-Hispanic men aged 25-59. Peri and Yasenov proposed expanding the sample to include women, non-Cuban Hispanics, and teenagers. As I showed in Borjas (2016), each of those sampling decisions introduced highly questionable changes in sample composition that contaminated observed wage trends.

Clemens and Hunt criticize my reappraisal from a different angle. Instead of claiming that the sample of non-Hispanic prime-age men does not correctly represent what the ideal sample should be, they instead argue that, from an operational standpoint, this sample is unrepresentative. There seems to be a data issue, particularly in the March CPS, that biases wage trends in this sample. In particular, there was a substantial increase in the black share of the low-skill workforce in Miami in the relevant period. Because African-American men earn less than white men, the increase in the black share of the low-skill workforce would spuriously produce a drop in the average low-skill wage in Miami—a wage drop that I incorrectly attributed to Mariel.

* Robert W. Scrivner Professor of Economics and Social Policy, Harvard Kennedy School; Research Associate, National Bureau of Economic Research; and Program Coordinator, IZA Program on Labor Mobility.

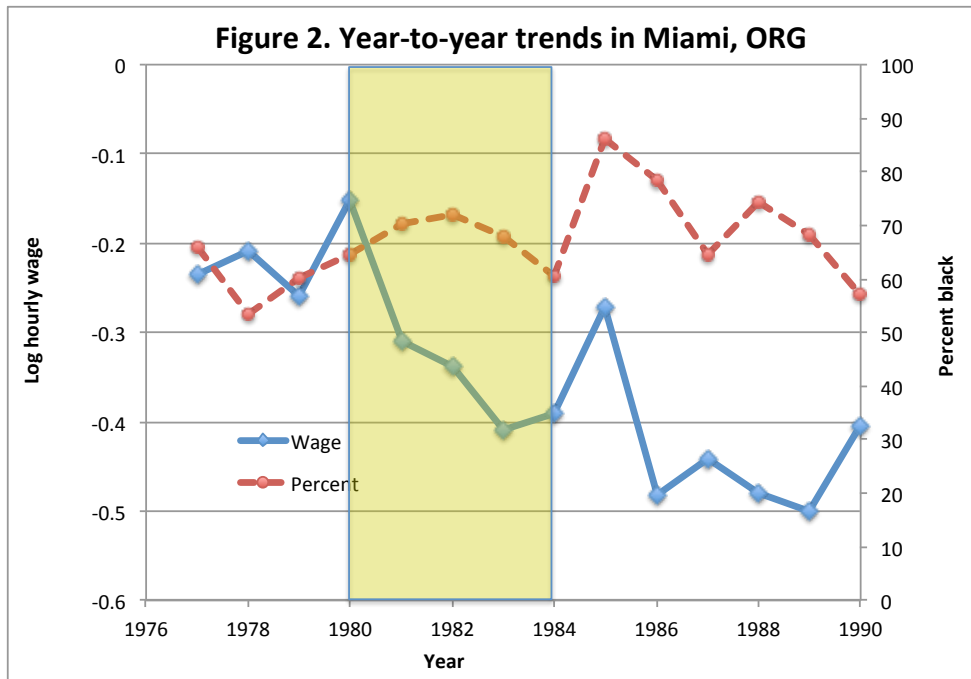
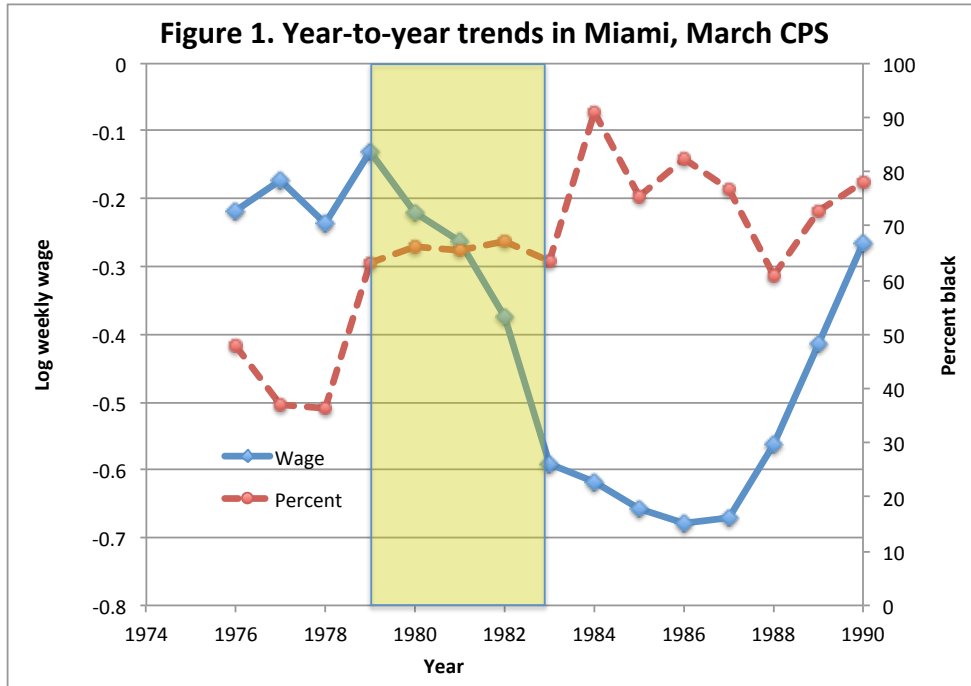
This short paper examines the robustness of the Mariel evidence presented in my original paper to statistical adjustments that control for the increasing number of black men in Miami's low-skill workforce. The evidence reported in this paper uses *exactly* the same sample examined throughout much of Borjas (2017): non-Hispanic men aged 25-59 who live in one of the 44 metropolitan areas identifiable in the CPS during the relevant period. As a result, any changes in the statistical evidence reported below and the comparable findings in my original paper can be attributed solely to the biases created by the changing racial composition of the low-skill workforce.

The statistical analysis clearly shows that although the black share of the low-skill workforce did rise in Miami, particularly in the March CPS and strikingly so in one single year (survey year 1980), the evidence consistently indicates that the race-adjusted low-skill wage in Miami fell significantly relative to those in other labor markets shortly after 1980 before fully recovering by 1990.

1. The timing of the increase in the black share

There is a very simple and intuitive way of showing that the rising black share in Miami's low skill workforce cannot be *the* reason for the average wage drop observed among low-skill workers in Miami after Mariel. The timing of the two events simply does not coincide.

Figure 1 shows the year-by-year trend in the age-adjusted log weekly wage in the March CPS between 1976 and 1990, while Figure 2 shows the respective trend in the age-adjusted log hourly wage in the ORG. All calculations reported throughout the paper use sampling weights. It is worth emphasizing that the wage trends illustrated in Figures 1 and 2 are *exactly* the ones reported and examined in my Mariel paper.



The figures also plot the fraction of the low-skill workforce that is black. It is evident that something strange happened to this fraction in calendar year 1979 in the March CPS

(which corresponds to survey year 1980), when it rose from 36.3 percent in 1978 to 63.0 percent in 1979. It is extremely unlikely that this shift reflects what was actually happening in Miami's low-skill labor market. Notably, the ORG does not show such a large increase in the black share, as the ORG black share was already high prior to 1979.¹ As a result, it seems likely that the jump in the March statistic reflects a sampling issue, an imputation problem, or an error in the construction/manipulation of the data by the BLS or IPUMS.

Note, however, that nothing whatsoever happened to the black share of the workforce in the March CPS between the years 1979 and 1983. Notably these are the years that witnessed the largest drop in the average wage of Miami's low-skill workforce. To emphasize again, the wage trend in Figure 1 is the wage trend used in the Borjas (2017) paper—it *includes blacks*. The figure convincingly shows that there is no relationship whatsoever between the fraction of blacks in the low-skill workforce and the average low-skill wage in those years that the average low-skill wage fell most. Therefore, it is *impossible* to explain the steep wage drop in terms of a statistical spuriousness created by a rising black share in the low-skill workforce. And this leads to an obvious inference: the Clemens-Hunt argument is not consistent with the *timing* of the increase in the black share and the drop in the average low-skill wage.

The ORG data in Figure 2 is equally striking. Again, the average low-skill wage in Miami (including blacks) fell dramatically between 1980 and 1984 while the black share rose slightly and then declined slightly over the period—ending up pretty much at the same place it started. There is simply no way that the large wage drop observed in the first few years after Mariel could somehow be related to a change in the racial composition of the low-skill workforce in the ORG.

¹ Interestingly, the 1980 census, conducted on April 1, 1980, just weeks before the Mariel supply shock, reports that the black share in the sample of low-skill, non-Hispanic working men aged 25-54 was 54.5 percent. This evidence is consistent with the data from the 1980 March CPS and the corresponding data from the ORG. The consistency suggests that the problem may lie with a systematic undercounting or misallocation of low-skill black workers in the March CPS prior to the 1980 survey year.

2. Trends in the age- and race-adjusted wage

The most obvious way of isolating the importance of the change in the racial composition of the sample on the average wage is to examine the trend in the *race-adjusted* wage.

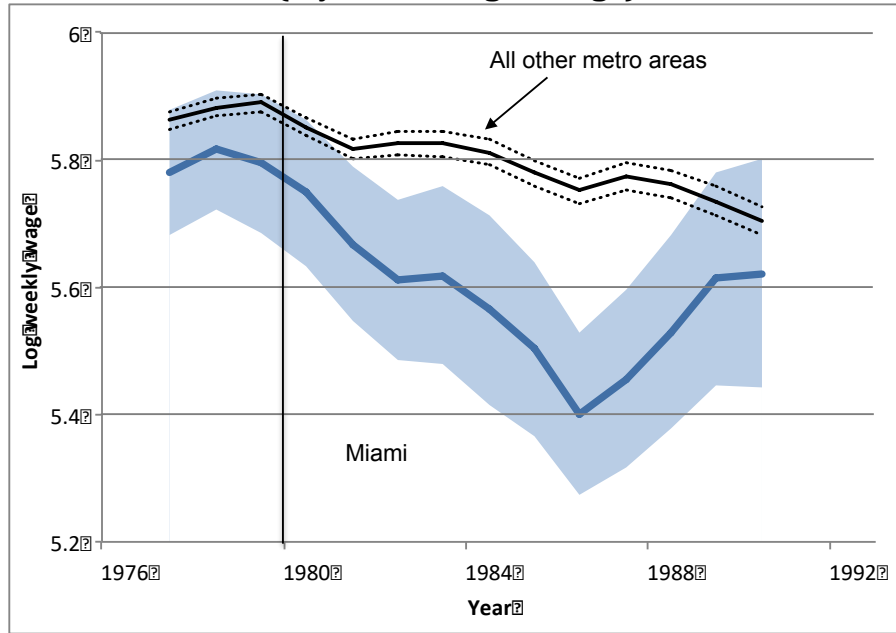
My original paper examines age-adjusted wage trends throughout much of the analysis. I now conduct a parallel exercise that examines the trend in the age-adjusted *and* race-adjusted wage. In particular, I used the sample of high school dropouts to estimate the following regression model separately in each cross-section of the March CPS or ORG files:

$$(1) \quad \log w_{it} = \alpha_t + \beta_t A_{it} + \gamma_t R_{it} + \lambda_{ct} C_{it} + \varepsilon_{it}$$

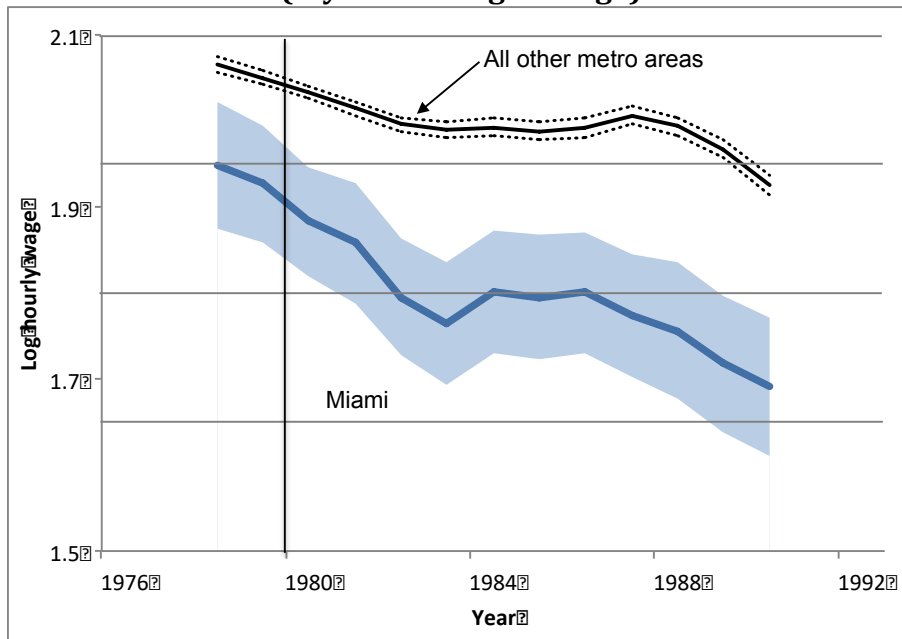
where w_{it} is the weekly wage of worker i in cross-section t in the March CPS (or the hourly wage of the worker in the ORG); A_{it} is a vector of fixed effects indicating the worker's age; C_{it} is a vector of fixed effects indicating the city of residence; and R_{it} is an indicator variable telling us if the worker is African-American.

The residual of this regression, or e_{it} , measures the age- and race-adjusted wage of worker i at time t . These residuals obviously have zero mean. It is trivial to redefine the residual so that it better reflects the actual wage level in a particular city at a particular time. In particular, let λ_{ct} be the estimate of the fixed effect for city c in year t . The measure of the age- and race-adjusted wage is then given by the sum $(e_{it} + \lambda_{ct} + \alpha_t)$. The variables added to the residual effectively allow for a person's age- and race-adjusted wage to differ systematically across cities (due to city fixed effects) and for that wage level to differ systematically over time (due to year fixed effects).

**Figure 3. Age- and race-adjusted 95% confidence interval, March CPS
(3-year moving average)**



**Figure 4. Age- and race-adjusted 95% confidence interval, ORG
(3-year moving average)**



Figures 3 and 4 illustrate the trend in this measure of the age- and race-adjusted wage measure, using a 3-year moving average (as comparable graphs did in my original paper). The graph also shows the 95 percent confidence interval. It is visually obvious that

something did indeed happen in Miami in the early 1980s in both the March CPS and the ORG. Is it also obvious that whatever it is that happened in Miami led to a substantial and significant decline in the age- and race-adjusted low-skill wage.

3. Regression analysis

What would happen to the regression results in the key Table 5 of my original paper if one were to redo the entire analysis using the age- and race-adjusted wage of workers? This remarkably easy and intuitive exercise, which would obviously be the first step in *any* empirical analysis that attempts to determine the role played by racial composition, is curiously not reported in Clemens and Hunt (2017).

Table 1. Difference-in-differences impact of the *Marielitos* on the age- and race-adjusted wage of high school dropouts

	Card placebo	Employment placebo	Low-skill placebo	Synthetic control	All cities
A. March CPS, men					
1981-1983	-0.121 (0.078)	-0.194 (0.070)	-0.101 (0.083)	-0.183 (0.053)	-0.117 (0.063)
1984-1986	-0.202 (0.072)	-0.301 (0.053)	-0.233 (0.068)	-0.290 (0.059)	-0.251 (0.041)
1987-1989	-0.202 (0.093)	-0.237 (0.071)	-0.099 (0.068)	-0.190 (0.039)	-0.193 (0.065)
1990-1992	0.094 (0.086)	0.025 (0.117)	0.129 (0.076)	-0.112 (0.133)	0.062 (0.044)
B. CPS-ORG, men					
1981-1983	-0.047 (0.029)	-0.104 (0.046)	-0.054 (0.024)	-0.123 (0.015)	-0.052 (0.010)
1984-1986	-0.024 (0.067)	-0.079 (0.060)	-0.033 (0.062)	-0.115 (0.055)	-0.045 (0.055)
1987-1989	-0.074 (0.039)	-0.137 (0.064)	-0.072 (0.035)	-0.114 (0.023)	-0.114 (0.021)
1990-1992	0.069 (0.044)	-0.041 (0.070)	0.077 (0.042)	-0.056 (0.024)	0.021 (0.048)

Notes: Robust standard errors are reported in parentheses. The data consist of annual observations for each city between 1977 and 1992 (1980 excluded). All regressions include vectors of city and year fixed effects. The table reports the interaction coefficients between a dummy variable indicating if the metropolitan area is Miami and the timing of the post-Mariel period. All regressions are weighted by the number of observations used to calculate the dependent variable. The number of observations of the synthetic control is a weighted average of the sample size in the actual cities that make up the synthetic city.

Table 1 reports the regression coefficients resulting from this analysis. Let me emphasize yet again: This exercise only adds a race indicator to the micro-level regression model estimated in my original paper—as summarized by equation (1) in that paper. It does not change anything else about the sample or regression specification.

It is evident that the regression coefficients—which measure the wage impact after accounting for the changing black share of the workforce in Miami and elsewhere—still show a significant wage drop in Miami relative to any placebo one cares to pick. And this is true both in the March CPS data as well as in the ORG. So the key inference from the key regression table in my paper is unchanged. Something happened to Miami’s low-skill wage after 1980.

4. Changing the baseline

As Figure 1 makes clear, it is the change in the black share in the March CPS between calendar years 1978 and 1979 (corresponding to survey years 1979 and 1980) that can potentially represent a problem when comparing averaged pre-1980 data with averaged post-1980 data. There is an alternative and very simple way of determining whether this data quirk creates a problem for estimating the wage impact of Mariel: simply discard all the March CPS data prior to calendar year 1979. By starting the series in calendar year 1979, the huge shift in the black share would no longer be able to contaminate the regression model.

This approach, of course, requires that the baseline—the pre-Mariel period—be defined differently. The simplest “trick” is to use the CPS data for calendar years 1979 and 1980 as the pre-Mariel years. The 1980 data is obviously contaminated by Mariel, as half that year’s earnings were observed after the *Marielitos* arrived. But the use of 1980 data to define the pre-treatment period should bias the results against finding any impact of Mariel.

I re-estimated the regression models in the March CPS and the ORG using the calendar years 1979-1980 years as the pre-treatment period. Table 2 summarizes the regression coefficients.

Table 2. Difference-in-differences impact of the *Marielitos* on the age- and race-adjusted wage of high school dropouts, using the 1979-1980 calendar years as baseline

	Card placebo	Employment placebo	Low-skill placebo	Synthetic control	All cities
A. March CPS, using age-adjusted wage					
1981-1983	-0.186 (0.096)	-0.220 (0.069)	-0.178 (0.089)	-0.144 (0.045)	-0.183 (0.078)
1984-1986	-0.344 (0.086)	-0.384 (0.049)	-0.388 (0.064)	-0.346 (0.041)	-0.385 (0.029)
1987-1989	-0.304 (0.099)	-0.233 (0.066)	-0.224 (0.070)	-0.181 (0.031)	-0.278 (0.061)
1990-1992	-0.002 (0.094)	0.019 (0.116)	0.041 (0.074)	-0.076 (0.144)	0.005 (0.036)
B. March CPS, using age- and race-adjusted wage					
1981-1983	-0.162 (0.091)	-0.192 (0.070)	-0.149 (0.084)	-0.143 (0.049)	-0.169 (0.061)
1984-1986	-0.238 (0.087)	-0.300 (0.050)	-0.279 (0.069)	-0.250 (0.055)	-0.302 (0.037)
1987-1989	-0.237 (0.105)	-0.237 (0.070)	-0.145 (0.072)	-0.151 (0.032)	-0.245 (0.063)
1990-1992	0.058 (0.099)	-0.031 (0.109)	0.082 (0.078)	-0.072 (0.133)	0.011 (0.040)
C. ORG, using age- and race-adjusted wage					
1981-1983	-0.055 (0.029)	-0.089 (0.041)	-0.061 (0.022)	-0.094 (0.028)	-0.069 (0.020)
1984-1986	-0.034 (0.066)	-0.064 (0.056)	-0.040 (0.062)	-0.086 (0.061)	-0.062 (0.058)
1987-1989	-0.085 (0.038)	-0.122 (0.061)	-0.079 (0.035)	-0.084 (0.033)	-0.131 (0.027)
1990-1992	0.060 (0.043)	-0.026 (0.067)	0.071 (0.041)	-0.026 (0.033)	0.004 (0.051)

Notes: Robust standard errors are reported in parentheses. The data consist of annual observations for each city between 1979 and 1992. All regressions include vectors of city and year fixed effects. The table reports the interaction coefficients between a dummy variable indicating if the metropolitan area is Miami and the timing of the post-Mariel period. All regressions are weighted by the number of observations used to calculate the dependent variable. The number of observations of the synthetic control is a weighted average of the sample size in the actual cities that make up the synthetic city.

It is obvious that excluding from the analysis all the CPS March surveys that reported a low share of the black workforce barely changes the results. In fact, an “eyeball”

comparison of the regression coefficients reported in the top panel in Table 2 with the corresponding coefficients in my original paper (Borjas, 2017, Table 5) indicates that *the results barely change!* As an example, consider the estimated wage impact on Miami relative to all other cities in the 1984-1986 period. In my original paper, the estimated effect was -0.393 (with a standard error of 0.022). Excluding all pre-1979 March CPS observations leads to an estimated impact of -0.385 (0.029). There is no doubt, therefore, that the post-1980 relative wage drop experienced by Miami's low-skill workforce was significantly steeper than that observed in all other comparison groups. This finding is not surprising given the visual evidence reported earlier, as Miami's low-skill wage did not change much in the years leading up to 1980.

5. Combinatorial analysis

As I explicitly noted in my original paper, one of the troubling things about the statistical methodology involved in Mariel-type research is that the researcher can influence the answer by picking an appropriate set of placebo cities for comparison. The increasingly popular synthetic control method is a very clever attempt to get rid of that influence, but it is still problematic. The researcher is still free to select whichever control variables he or she wants to use to create the synthetic city. And, as I show in Borjas (2017), different sets of control variables lead to very different measured impacts.

So my favorite part of my Mariel paper was towards the end, where I chose to eliminate researcher influence entirely by doing a combinatorial-type analysis. Why not simply compare what happened in Miami to *every single conceivable placebo* and see what the distribution of impacts across potential placebos looks like?

I operationalized this idea by constructing every single four-city placebo. Because there are 43 cities besides Miami that can be identified in the CPS during that period, that means that there are 123,410 conceivable four-city placebos. If the resulting distribution across the 123,410 measured impacts had a lot of positive impacts, a lot of zero impacts, and a lot of negative impacts, that information would be very informative about what the "real" impact was. If, on the other hand, the distribution is composed mostly of negative impacts, that too would be informative.

Figure 5. Frequency distribution of short-run age- and race-adjusted wage effects, March CPS

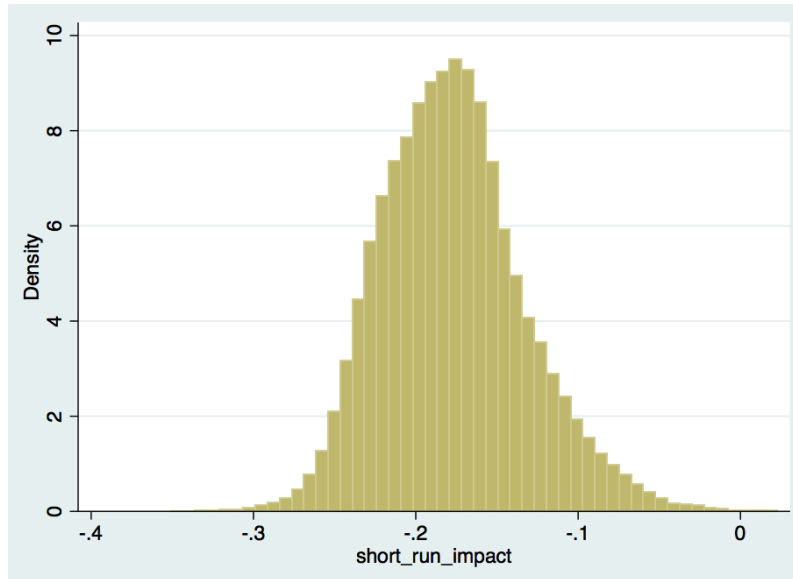
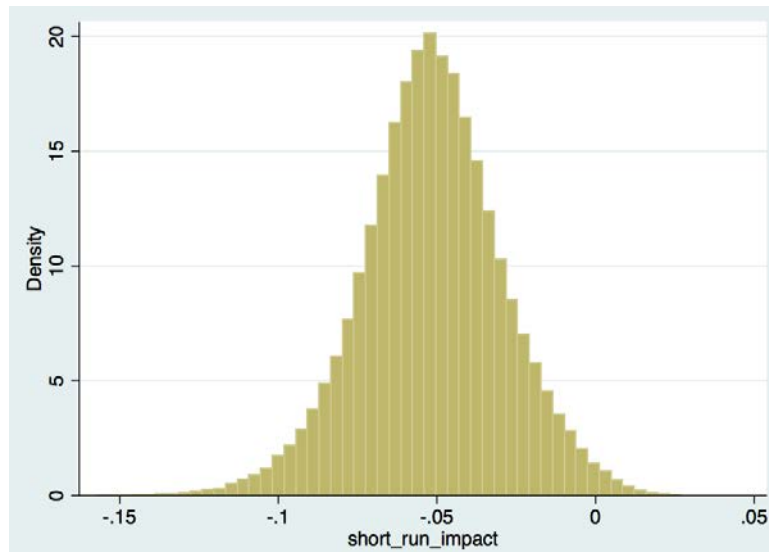


Figure 6. Frequency distribution of short-run age- and race-adjusted wage effects, ORG



I re-estimated the frequency distributions reported in Figure 8 of Borjas (2017) using the age- and race-adjusted wage data, and using exactly the same regression specification for estimating the “short-run wage impact.” The short-run effect gives the relative wage change in Miami between the pre-Mariel period (1977-1979) and the post-Mariel period (1981-1986). In short, Figures 5 and 6 are identical to what was presented in my original paper, except that they use age- *and* race-adjusted wage data.

It is again obvious that adjusting the data for race does not change the basic insight. In the March CPS, the median impact is -0.18, which means the wage drop in Miami before and after Mariel was around 18 percent larger than it was in the median placebo, and the 95 percent confidence interval covers the range between -0.25 and -0.08. Similarly, in the ORG, the median impact is -0.05, and 95 percent of the effects lie between -0.10 and -0.01. In sum, the short-run wage effect is negative, although the numerical magnitude has a wide range.

I also redid the combinatorial analysis by using the 1979-1980 calendar years as the baseline (as in the previous section), and then estimated the short-run wage impact for each of the 123,410 potential placebos. This approach is completely unaffected by the quirk in the reported black share of the low-skill workforce in the March CPS between calendar years 1978 and 1979. The median wage effect in the March CPS was -0.23, and the range of the 95 percent confidence interval was from -0.33 to -0.12. The median effect in the ORG was -0.07, and the 95 percent confidence interval went from -0.11 to -0.02. The Appendix presents the figures for the two frequency distributions that use the 1979-1980 calendar years as the baseline.

6. Calculating the race-adjusted wage

In a recent blog post, Clemens (2017) noted that the regression evidence summarized in Table 1 is indeed correct. The race adjustment performed in equation (1) by simply adding a race fixed effect to the micro-level regression model in each cross-section leads to wage trends that still indicate a decline in Miami's low-skill wage after Mariel. This very simple type of race adjustment is, by far, the most straightforward (and parsimonious) way to fix concerns about sampling.

The blog post, however, also notes that an alternative approach to adjusting for race—where the race variable *in each cross-section* is introduced not only as a fixed effect, but as a fixed effect interacted with metropolitan area and education (so that the racial wage gap for an education group in a particular city is allowed to vary over time)—takes away much of the significant wage drop. In an important sense, this response moves the “goalpost” of the Clemens-Hunt criticism of my Mariel paper. It is no longer that the change in the black share of the workforce induced a spurious correlation that led to lower wages

in post-Mariel Miami; the evidence reported in the previous sections shows that this particular argument is just plain wrong. It is now instead that the measured wage impact of Mariel could be zero if we calculate the race-adjusted wage by allowing for the racial wage gap for low-skill workers to vary over time in a city like Miami.

Let's put aside momentarily the conceptual issue of whether it is proper to calculate the race-adjusted wage by netting out the time-series variation in the racial wage gap within a city (a variation that may have arisen *because* of immigration). If we had a lot of data available, the conventional approach to estimating the race-adjusted wage would be to include a fully interacted set of fixed effects in the regression model. The fully interactive regression specification used to calculate the age- and race-adjusted wage for worker i would be given by:

$$(2) \quad \log w_{it} = \alpha_t + \lambda_t C_{it} + \varphi S_{it} + \beta_t A_{it} + \gamma_t R_{it} \\ + \kappa_{1t} (C_{it} \times S_{it} \times A_{it} \times R_{it}) + \varepsilon_{it},$$

where the vector $(C_{it} \times S_{it} \times A_{it} \times R_{it})$ denotes all possible two-, three-, and four-way interactions among all variables—the city fixed effects (C_{it}), the education fixed effects (S_{it}), the age fixed effects (A_{it}), and the variable indicating the worker's race (R_{it}). Because the micro-level regression is estimated separately in each CPS cross-section, the model is also fully interactive with respect to time. Therefore, this model allows variation in the racial wage gap across age-education-city cells at a point in time and within each cell over time. The age- and race-adjusted wage is given by the residual of this regression, and I use this residual to calculate the mean age- and race-adjusted wage in each city in each year for high school dropouts.

Table 3. Difference-in-differences impact of the *Marielitos* on the age- and race-adjusted wage of high school dropouts (using alternative interaction models)

	March CPS			ORG		
	Card placebo	Synthetic control	All cities	Card placebo	Synthetic control	All cities
A. Full interaction model						
1981-1983	-0.011 (0.016)	0.017 (0.028)	-0.013 (0.010)	-0.031 (0.011)	-0.023 (0.021)	-0.022 (0.010)
1984-1986	0.000 (0.015)	0.011 (0.008)	-0.003 (0.007)	-0.049 (0.015)	-0.036 (0.021)	-0.033 (0.010)
1987-1989	-0.004 (0.023)	-0.017 (0.013)	-0.017 (0.013)	-0.043 (0.013)	-0.033 (0.021)	-0.033 (0.007)
1990-1992	0.003 (0.013)	0.008 (0.012)	0.005 (0.010)	-0.037 (0.014)	-0.028 (0.023)	-0.023 (0.011)
B. (Age × race × year), (education × race × year), (city × race × year)						
1981-1983	-0.184 (0.070)	-0.256 (0.057)	-0.187 (0.055)	-0.050 (0.026)	-0.127 (0.023)	-0.052 (0.008)
1984-1986	-0.231 (0.065)	-0.339 (0.055)	-0.273 (0.036)	-0.074 (0.053)	-0.152 (0.033)	-0.081 (0.042)
1987-1989	-0.274 (0.073)	-0.281 (0.058)	-0.243 (0.044)	-0.102 (0.039)	-0.133 (0.030)	-0.124 (0.025)
1990-1992	-0.014 (0.071)	-0.201 (0.158)	-0.025 (0.045)	0.011 (0.043)	-0.083 (0.028)	0.020 (0.043)

Notes: Robust standard errors are reported in parentheses. The data consist of annual observations for each city between 1977 and 1992 (1980 excluded). All regressions include vectors of city and year fixed effects. The table reports the interaction coefficients between a dummy variable indicating if the metropolitan area is Miami and the timing of the post-Mariel period. All regressions are weighted by the number of observations used to calculate the dependent variable. The number of observations of the synthetic control is a weighted average of the sample size in the actual cities that make up the synthetic city.

The top panel of Table 3 summarizes the regression results using this alternative measure of the age- and race-adjusted wage. The magnitude of the estimated wage impacts is far smaller (particularly in the March CPS) than in the comparable regressions that excluded all the interactions in Table 1. This is not surprising because the use of a fully interactive model nets out much of the wage variation that exists across the various “markets.” The typical March CPS cross-section in the 1980s, for instance, has 8,865 observations. The fully interactive model adds 2,816 regressors to the regression (or 44

metro areas \times 4 education groups \times 8 age groups \times 2 race groups). Given the relatively small sample sizes available in the March cross-sections and the large number of regressors, it is also not surprising that this kitchen-sink approach reduces the precision of the estimates in Table 3 sufficiently so that the null hypothesis of “no effect” can no longer be rejected. Remarkably, many of the wage impacts remain negative and significant in the ORG, particularly when there is more variation in the data (as in the regression that compares Miami to all other cities). So there is still a suggestion that low-skill workers in post-Mariel Miami experienced a steeper wage drop than comparable workers in other cities.

However, it is far from clear that one should interpret the regression coefficients resulting from this fully interactive model as measuring the wage impact of immigration. As I noted above, the variation in the racial wage gap for a particular education group across cities and over time might have arisen *because* of immigration. It seems plausible, for instance, that a supply shock might affect the wage and employment of low-skill black and white workers differently.² There are substantial differences in the jobs the two groups hold, in the occupations they enter, and in the industries that employ them. A particular supply shock—such as Mariel—inevitably penetrated some sectors more than others, affecting the magnitude of the racial wage gap for a particular education group in Miami relative to other cities. The residual from a regression that nets out this differential impact removes much of the effect that immigration might have had on the local labor market. As a result, it would not be surprising if models that calculate the race-adjusted wage *after* netting out the impact that immigration might have had on the racial wage gap generate smaller, and perhaps near zero, measured impacts of immigration.

To illustrate the quantitative importance of this bias, the second panel of Table 3 uses an alternative model that relies on binary interactions.

² Very few studies estimate the potentially disparate impact of immigration on black and white workers. Borjas, Grogger, and Hanson (2010) model the source of the different effects, but do not separately estimate the impact of immigration on low-skill blacks and whites. The wage elasticity—estimated using data from all education groups—is essentially the same for the two race groups, but immigration has a much greater adverse effect on black employment. Smith (2012, Table 8) also finds that low-skill immigration leads to a much greater reduction in annual hours worked by black young men than by white young men. The early work of Altonji and Card (1991, Tables 7.8 and 7.9, column 4) does estimate the elasticity separately for the two low-skill groups. The negative wage effect is almost twice as large for blacks as it is for whites.

$$(3) \quad \log w_{it} = \alpha_t + \lambda_t C_{it} + \varphi S_{it} + \beta_t A_{it} + \gamma_t R_{it} \\ + \kappa_{1t} (C_{it} \times R_{it}) + \kappa_{2t} (S_{it} \times R_{it}) + \kappa_{3t} (A_{it} \times R_{it}) + \varepsilon_{it},$$

This model includes all two-way interactions—(city \times race), (education \times race), and (age \times race)—when calculating the race-adjusted wage in each cross-section. Let me emphasize that this regression model is estimated separately in each cross-section, allowing for the average racial wage gap in each city to vary over time, and for the average racial wage gap for each education (or age) group to also vary over time. *The crucial property of the binary model is that its residual (i.e., the measure of the race-adjusted wage) does not net out that the racial wage gap for high school dropouts in a particular city (like Miami) might have changed over time.* As Table 3 shows, this alternative approach to calculating the race-adjusted wage leads to a much larger wage drop in post-Mariel Miami than in comparison cities, and the estimated wage effects, as in my original paper, are significant.

Table 4. Numerical example of bias from alternative calculations of the race-adjusted wage

	Before		After		% drop in average wage
	Whites	Blacks	Whites	Blacks	
Wage of low-skill workers:					
A. Actual data:					
Miami	\$10	\$7	\$9	\$4	23.5
New York	\$11	\$8	\$11	\$8	0.0
B. Race-adjusted wage, using city-race-year interaction:					
Miami	\$10	\$10	\$9	\$9	10.0
New York	\$11	\$11	\$11	\$11	0.0
C. Race-adjusted wage, using race-year interaction					
Miami	\$10	\$10	\$9	\$6	25.0
New York	\$11	\$11	\$11	\$11	0.0

Note: The calculation of the average wage assumes that 50 percent of Miami's low-skill workforce is black.

It is easy to illustrate the source of the bias induced by a fully interactive model with a trivial numerical example, summarized in Table 4. Suppose there are two cities, Miami and New York. New York did not receive any immigrants, but Miami did. The table shows the average wage of black and white low-skill workers in the two cities before and after the supply shock. The top panel gives the unadjusted wage data (i.e., the data that would be available in the CPS). By construction, immigration had a much larger impact on black workers in Miami, reducing their wage from \$7 to \$4, while the wage of white workers fell by only \$1, from \$10 to \$9.

The middle panel of the table shows the race-adjusted wage in each city. The regression that includes a complete set of interactions between city, education, race, and year would “see” that there is a \$3 racial wage gap among low-skill workers in Miami prior to the supply shock, and use that information to conclude that the race-adjusted wage of a black worker should be \$10. After the supply shock, the regression model would “see” a \$5 racial wage gap, and use that information to conclude that the race-adjusted wage of a black worker should be \$9.

Suppose that half of Miami’s workforce is black. The average race-adjusted wage in Miami fell from \$10 to \$9, or 10 percent. In fact, however, the average wage in Miami fell from \$8.50 to \$6.50, or nearly a 25 percent drop. Ignoring the fact that the supply shock might have increased the racial wage gap greatly underestimates the wage impact of immigration.

The bottom panel of Table 4 shows what would happen if we used an alternative regression specification that did not “throw the baby out with the bathwater.” Suppose, in particular, that Miami is a small city relative to New York, so that the national racial wage gap in the post-migration period is \$3. The regression model that includes only a race-year interaction would then yield a race-adjusted post-migration wage for black workers in Miami of \$6 (or \$3 less than what whites get), indicating that the average race-adjusted wage in Miami fell from \$10 to \$7.50, or 25 percent.

I can even extend this numerical example in a more subtle direction. Perhaps the \$3 racial wage gap in Miami in the pre-migration period represents the “true” (and permanent) discriminatory penalty suffered by low-skill black workers in the Miami labor market. We could then use that information to calculate the race-adjusted wage of a black

worker in Miami in the post-migration period, and that race-adjusted wage would again be \$6. The average race-adjusted wage in Miami would again fall by 25 percent.

In short, the empirical exercise used to compute the race-adjusted wage in a city at a point in time should not follow blindly from a kitchen-sink approach to regressions. Careful thought must be given to *why* racial wage differences might arise within skill groups and across cities, how the time trend of those racial differences might be affected by immigration, and exactly which value of the racial “wage penalty” should be used in the calculation of a race-adjusted wage. To the extent that immigration may have contributed to racial wage differences, particularly in post-Mariel Miami, it is incorrect to calculate a race-adjusted wage that nets out that variation. In the end, the use of a simple linear race fixed effect that varies over time, and perhaps even across education groups, seems to be an important virtue of a more sensible research strategy, both because of the small sample sizes involved and because it does not introduce any risk that it inadvertently wipes away the impact we are trying to estimate.

7. Summary

It is evident that adjusting for the racial composition of the workforce in either the March CPS or the ORG files does not change the key insight of the Mariel reappraisal in Borjas (2017). The relative wage of low-skill workers in Miami fell significantly in the period after Mariel, reaching a nadir somewhere around the mid-1980s and recovering fully by 1990. The claim that changes in the racial composition of Miami’s workforce created a spurious correlation that explains this wage trend is demonstrably false.

References

Altonji, Joseph G., and David Card. "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives," in John M. Abowd and Richard B. Freeman, editors, *Immigration, Trade, and the Labor Market*. Chicago: University of Chicago Press, 1991, pp. 201-234.

Borjas, George J. "The wage impact of the Marielitos: Additional evidence", Working Paper No. w21850, NBER Working Paper No. 21850, January 2016.

Borjas, George J. "The wage impact of the Marielitos: A reappraisal", *Industrial and Labor Relations Review*, forthcoming 2017.

Borjas, George J., Jeffrey Grogger, and Gordon H. Hanson. "Immigration and the Economic Status of African-American Men," *Economica* 77 (April 2010): 255-282.

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

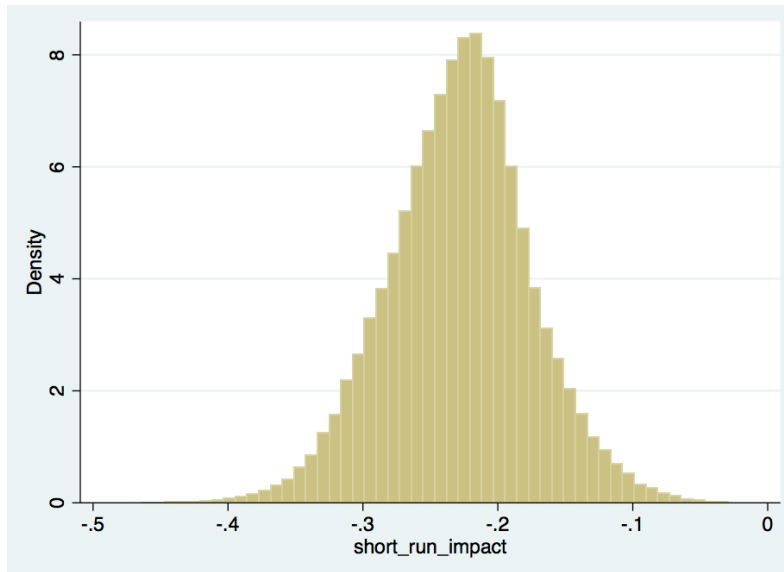
Clemens, Michael A. "What Economists Can Learn from the Mariel Boatlift, Part Two: Answering Questions about Our Research," <https://www.cgdev.org/blog/what-economists-can-learn-from-mariel-boatlift-part-two?callout=1-1>, May 30, 2017.

Clemens, Michael A., and Jennifer Hunt. "The Labor Market Effects of Refugee Waves: Reconciling Conflicting Results," NBER Working Paper No. 23433, May 2017.

Peri, Giovanni and Vasil Yassenov. "The labor market effects of a refugee wave: Applying the synthetic control method to the Mariel boatlift," NBER Working Paper No. 21801, December 2015.

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

Appendix Figure A1. Frequency distribution of short-run age- and race-adjusted wage effects, using the 1979-1980 calendar years as baseline, March CPS



Appendix Figure A2. Frequency distribution of short-run age- and race-adjusted wage effects, using the 1979-1980 calendar years as baseline, ORG

