Econ 488

4/8/21

Abe Burton, Jacob Lawrence, Mena Villanueva

Charlotte Housing Shifts from Bankers' Bonuses

I. INTRODUCTION

What is the effect of bankers' bonuses on Charlotte housing market activity? Charlotte, North Carolina is in many ways an unremarkable city, infrequently appearing in discussions of the major metropolises of the United States. Nevertheless, the city's financial district, while not on the scale of such centers as New York or Chicago, is home to Bank of America, Wells Fargo, and other large banks. BofA and WF alone employ some 40,000 people in the area and have a reputation for giving out exorbitant bonuses. This reputation has led many Charlotteans to the conclusion that it is these bonuses that cause a massive uptick in local housing market activity in the month of February. The theory goes that the bonuses are received in mid to late January and thus are subsequently turned into down payments on homes the following month. Academic literature has been completely silent on the subject up until now and so this paper seeks to shed some light on the matter. If true, this knowledge could benefit those in the housing market, perhaps inducing those looking to sell property to do so in February or those looking to buy to do so in a different month to avoid the competition.

II. DATA

The data collected falls into three groups: housing, unemployment, and population. First, the housing data comes from Movoto.com, an online real estate marketplace. The website contains publicly accessible, month by month and city by city data for the housing market, from 2016 through 2021. The data is collected automatically as buyers and sellers make purchases through the site and is then made available to inform potential buyers and sellers of market trends. The particular variable of interest taken from this data set is "Median Days on Movoto," which simply gives the average measure of how long listings remain on the site before being sold. One weakness of this data is that since we are relying on only one data source, it could be that the trends reflected in homes sold on Movoto are not true of general real estate trends. However, this is not much of a concern because the site caters to the general market and not to say only the higher or lower end and thus the homes on the site are representative of general availability and price dispersion.

Second, the unemployment data comes from the website of the Bureau of Labor Statistics (BLS), a division of the Department of Labor. This is a national database with a range of available data through 2019. The federal government hires field economists to gather data through interviews and then uses analysts to clean and sort the acquired data. This is a highly reputable dataset and is used for a variety of purposes by labor economists and policy makers. The key variable used for this research is the unemployment rate, which is given by city and month. Additional controls for the net change in unemployment and percentage change of unemployment allow for further control of unemployment trends. This data provides a useful control to the Movoto housing trend information by providing a measure of economic trends which may have a confounding effect on the results.

Third, the population data is from the national census. This data is collected by the federal government for a variety of purposes, including apportioning funding and for dividing the states into districts. Population levels are an important control for the housing market as they, along with employment levels, are an important indicator of demand. Unfortunately, while this data is available at the city level, it is only available on a yearly, and not monthly, basis. Additionally, this data is only accessible for the years 2017 and 2018 and not through 2019 as the other data is.

In order to overcome these weaknesses, two primary regressions are performed: a regression of 2017-2018 data with a population control added (to get monthly measures, a linear growth rate between years was assumed and then divided by twelve to estimate a monthly population) and a regression with data from 2016-2019 but without the population control. These two separate regressions (along with other minor iterations of them) allow for a check of the robustness of the findings. Collectively, the three data sets combined are very well suited to answer the question, as they allow for isolation of housing market effects in a specific city (Charlotte) in a specific month (February).

Table 1: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Median days on Movoto	1080	90.12	47.744	0	884
Unemployment Rate	1080	4.695	1.299	2.5	13.7
Monthly net change of unemployment	1080	-.042	.493	-6.7	7.3
Monthly percent change of unemployment	1080	-.569	8.803	-48.9	114.1
Population	552	131382.1	178381.4	20041	885708

Table 2: Summary Statistics in February

Variable	Charlotte			Not Charlotte		
	Obs	Mean	Std. Dev	Obs	Mean	Std. Dev
Median days on Movoto	3	42.333	2.082	69	97.116	35.051
Unemployment rate	3	4.4	.458	69	4.894	1.143
Monthly net change of unemployment	3	-.1	.1	69	-.165	.215
Monthly percent change of unemployment	3	-2.267	2.411	69	-3.058	4.07

III. IDENTIFICATION STRATEGY

A fixed effect regression is well suited to identify the causal effect of interest. The primary estimating equation is as follows:

$$Y_{imt} = \beta_0 + \gamma_i + \tau_m + \eta_t + \theta(Charlotte * \tau_m) + \beta_1 Unemployment_{im} + \beta_2 Population_{im} + \epsilon_{imt}$$

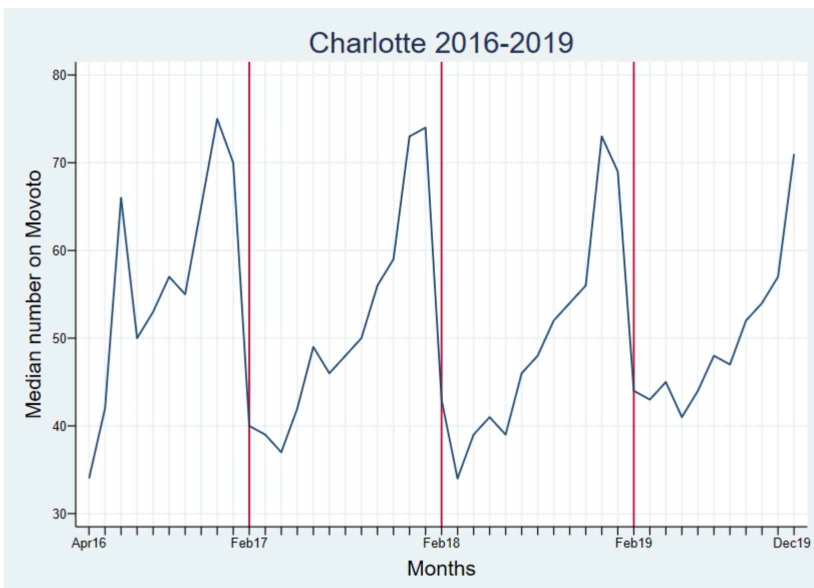
Where Y is equal to median days on Movoto, gamma, tau, and eta represent the fixed effects of city, month, and year respectively, theta is an indicator for the effect of Charlotte in each month, and betas one and two are controls for unemployment and population respectively. These variables are included and omitted to check for robustness.

The main assumption that has to be made is that the error terms are not correlated with the fixed effects and that there are no contemporaneous shocks, the risk of both of which is mitigated by the various controls put into place. The last assumption is that there is no endogenous treatment, which is fairly certain as bankers cannot change when they receive their bonuses or when the month of February happens. Once past those assumptions, the strategy for inference is a difference in means hypothesis test between Charlotte in January and Charlotte in

February against a zero null hypothesis. This method is the standard test for significance in a fixed effects regression and as such is appropriate for this use. Clustered standard errors at the city level allow further confidence in the accuracy of the causal inference.

IV. RESULTS

Plotting the median days a house is listed on Movoto in Charlotte shows a steep decrease from January to February. Houses don't sell as fast in the period between July to January, with a general increase in median listing time during this period.



Graph 1

1. Comparing Means

First, to demonstrate that there is in fact a significant change from January to February in Charlotte, without considering city or time fixed effects, a simple difference in means t test is performed. As seen in Table 3, there is a significantly lower mean for February than January. Further, the difference between February and March is not statistically significant.

Table 3: Difference of Means Between Jan-Feb-March

Variable	February	January		March	
	Mean	Mean	Diff	Mean	Diff
Median days on Movoto	42.33	71	28.67 (1.94)***	38.67	-3.67 (2.87)
Unemployment rate	4.4	4.5	0.1 (.37)	4.07	-0.33 (0.32)
U. rate monthly net change	-.1	.43	0.53 (.067)***	-0.33	-0.23 (0.11)
U. rate percent change	-2.27	10.83	13.1 (1.98)***	-7.4	-5.13 (2.06)*

*** $p<.01$, ** $p<.05$, * $p<.1$

The statistically significant differences between January and February for unemployment changes suggest that the two months differ in more aspects than just the variable of interest.

Table 4 shows that there are significant differences between Charlotte and other cities, although the fact that the unemployment change variables do not show significant differences suggest that Charlotte follows the same employment trends as do other cities.

Table 4: Difference of Means Between Charlotte and other cities (2016-2017)			
Variable	Charlotte	Others	Diff
Median days on Movoto	51.56 (1.49)	91.8 (1.49)	40.24 (2.28)***
Unemployment rate	4.14 (0.08)	4.72 (0.04)	0.57 (.09)***
U. rate monthly net change	-.038 (0.04)	-0.42 (0.02)	-0.004 (.04)
U. rate percent change	-.74 (1.00)	-.56 (.28)	13.1 (1.98)

*** *p<.01,* ** *p<.05,* * *p<.1*

These clear differences between cities and months are what motivate the inclusion of fixed effects which control for them.

2. Fixed Effects Regressions

The data for 2016-2019 reveal a significant month by city effect for Charlotte in February. Appendix Table 1 shows the estimates of the fixed effects for individual cities, months, and years, and for covariates of unemployment rate, monthly net change, and monthly percent change. Panel D of that table gives the estimate for how the location of Charlotte in a given month affects how quickly a house sells. The treatment variable is the second in this panel. Not every month yielded a significant estimate but February did and was the largest registered (fixed effect value of -20.88 as compared to the reference group of December, with the next largest effect being only about 14). Despite not being as significant as the February value, the fact that there are several other months with a level of significance weakens the claim by suggesting there may be more going on in Charlotte than just the bonuses.

The following tables and discussion show a variety of regressions with various inclusion and omission of certain data and controls.

- Fixed Effects Without Population Control

Table 5: Fixed Effects 2016-2019

Variable	(1)	(2)	(3)
Being in Charlotte in February	-21.37 (3.44)***	-21.43 (3.5)***	-20.87 (3.48)***
Unemployment rate	-	1.03 (4.27)	3.4 (5.9)
U. rate monthly net change	-	-	-3.24 (3.91)
U. rate percent change	-	-	-.07 (.31)

**** p<.01, ** p<.05, * p<.1*

Table 5 gives the estimates for fixed effects regressions as we add covariates. All columns control for fixed effects at a city, year, and month level. Column (1) does not include covariates, column (2) incorporates a control for monthly unemployment, and column (3) adds monthly net change and percent change of unemployment rate. The estimates for the variable of interest oscillated around -21 and were consistent at every regression. Adding regressors did not change the estimate significantly. The effect found in column 3 is slightly smaller than in the others, which suggests that excluding these covariates may produce an overestimated effect. Overall, table five shows that being in Charlotte in February decreases the median days a house is listed by an average of 20.87 as compared to controls (while unemployment variables did not show significant effects).

These regressions do not include population controls and thus have a potential source of omitted variable bias. However, they are robust to all unobserved city, month, and year fixed effects, plus the city by month fixed effects for Charlotte.

- Variation in Included Years

Table 6: Fixed Effects Regression, varying the sample size

Variable	2016-2019 (1)	2017-2018 (2)	2018-2019 (4)
Being in Charlotte in February	-20.87 (3.48)***	-24.98 (4.88)***	-23.63 (4.61)***
Unemployment rate	3.4 (5.9)	13.27 (3.14)***	7.27 (4.05)
U. rate monthly net change	-3.24 (3.91)	-19.13 (6.99)**	-7.76 (5.99)
U. rate percent change	-.07 (.31)	.23 (.39)	.07 (.31)

**** p<.01, ** p<.05, * p<.1*

2016-2017 was excluded because there are no observations before April 2016.

Table 6 shows the estimates from limiting the sample to smaller time frames. Column (1) corresponds to the full sample. Columns (2) and (3) constrain the sample to 2017-2018 and 2018-2019 respectively and have a fairly similar estimate and significance to the full sample. This allows more confidence when next the population control is added (by necessity) to just the 2017-2018 data, by suggesting that there are similar trends in that small sample as in the full sample (despite the different estimates of unemployment controls) and thus that, although the smaller sample has less power, it at least represents the same city by month fixed effect, which is the variable of interest.

- 2017-2018 with Varying Controls

Table 7 below shows the different estimates found as we incorporate controls. The estimate for being in Charlotte in February increases as more variables are included (note column 3 expresses the same result as column 2 of Table 6).

Table 7: Fixed Effects Regression (2017-2018), varying the controls

Variable	(1)	(2)	(3)	(4)
Being in Charlotte in February	-25.54 (5.29)***	-25.94 (5.12)***	-24.98 (4.88)***	-13.491 (6.75)
Unemployment rate	-	7.08 (3.55)*	13.27 (3.14)***	13.12 (3.19)***
U. rate monthly net change	-	-	-19.13 (6.99)**	-19.53 (6.89)**
U. rate percent change	-	-	.23 (.39)	.29 (.39)
Population	-	-	-	.001 (.0005)*

**** p<.01, ** p<.05, * p<.1*

Although the significant negative result is robust to unemployment controls, it is not to the population control. This suggests that the lack of population control generates an omitted variable issue, biasing those results. Indeed, population growth in Charlotte (which has seen a high amount of migration into the city) may be part of the cause for the housing market activity.

Table 2 in the Appendix shows the detailed results from controlling for fixed effects at city, month, and year level and clustering for autocorrelation within cities. The base units for the estimation were January, 2017, and Charlotte in December.

V. CONCLUSION

The results lead to an unsatisfying conclusion. There is in fact a significant result found in Charlotte in February when the population control is not included, which supports the argument that this could be due to bankers and their bonus timing. Sadly, once the population control is added, this effect is no longer significant, which suggests that this claim is a fallacy.

However, is this loss of significance due solely to the introduction of the control or to the limited window caused by lack of data? Once the population control is added, the Charlotte by February coefficient does get significantly smaller, but the standard error is also significantly larger than for the full sample regression. It is possible that the inclusion of more population data would lower the standard error enough to conclude a significant result and thus this question may be worth future study once more data is available. For now, this paper can at least show that there is indeed a significant uptick in housing market activity in February, but whether this effect is more pronounced in Charlotte due to the banking industry is still an open question.

VI. APPENDIX

APPENDIX TABLE 1: Charlotte in February (2016-2019)

Median Days	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Panel A: Month Fixed Effects (Base January)							
2	-14.624	3.585	-4.08	0	-22.041	-7.207	***
3	-33.349	3.937	-8.47	0	-41.494	-25.205	***
4	-33.048	4.389	-7.53	0	-42.128	-23.968	***
5	-29.642	3.905	-7.59	0	-37.72	-21.564	***
6	-21.242	6.672	-3.18	.004	-35.044	-7.44	***
7	-33.45	3.909	-8.56	0	-41.535	-25.364	***
8	-34.716	3.308	-10.50	0	-41.559	-27.873	***
9	-37.836	4.848	-7.80	0	-47.866	-27.807	***
10	-29.325	5.503	-5.33	0	-40.708	-17.942	***

11	-23.893	5.115	-4.67	0	-34.474	-13.312	***
12	-12.171	5.472	-2.22	.036	-23.491	-.85	**
Panel B: Year Fixed Effects (Base 2016)							
2017	-27.174	12.103	-2.25	.035	-52.212	-2.137	**
2018	-33.442	15.545	-2.15	.042	-65.599	-1.284	**
2019	-46.202	15.849	-2.92	.008	-78.988	-13.416	***
Panel C: Covariates							
unemployment	3.402	5.914	0.58	.571	-8.832	15.635	
m_net_change	-3.236	3.916	-0.83	.417	-11.337	4.865	
mo_perc_change	-.071	.312	-0.23	.823	-.717	.576	
Panel D: Charlotte in Month X (Base Charlotte in December)							
month#1 : base 1	-4.525	2.223	-2.04	.053	-9.123	.073	*
2	-20.88	3.479	-6.00	0	-28.076	-13.683	***
3	-5.805	3.951	-1.47	.155	-13.978	2.368	
4	-14.413	5.634	-2.56	.018	-26.068	-2.757	**
5	-13.355	5.811	-2.30	.031	-25.375	-1.335	**
6	-13.972	9.778	-1.43	.166	-34.2	6.256	
7	-4.951	2.857	-1.73	.097	-10.861	.96	*
8	-2.911	2.05	-1.42	.169	-7.151	1.33	
9	3.619	1.538	2.35	.028	.437	6.801	**
10	-.581	1.441	-0.40	.691	-3.563	2.401	
11	-1.249	1.138	-1.10	.284	-3.603	1.106	
Constant	128.616	39.949	3.22	.004	45.974	211.257	***

Mean dependent var	90.120	SD dependent var	47.744
R-squared	0.206	Number of obs	1080
F-test	.	Prob > F	.
Akaike crit. (AIC)	10955.085	Bayesian crit. (BIC)	11039.825

**** p<.01, ** p<.05, * p<.1*

APPENDIX TABLE 2: Charlotte in February (2017-2018)

Median Days	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Panel A: Month Fixed Effects (Base January)							
2	-19.356	3.821	-5.07	0	-27.279	-11.432	***
3	-35.48	4.837	-7.33	0	-45.512	-25.448	***
4	-34.167	4.978	-6.86	0	-44.491	-23.843	***
5	-29.384	4.491	-6.54	0	-38.698	-20.07	***
6	-31.675	3.562	-8.89	0	-39.063	-24.288	***
7	-34.349	4.53	-7.58	0	-43.744	-24.955	***
8	-35.625	4.919	-7.24	0	-45.826	-25.425	***
9	-32.716	8.176	-4.00	.001	-49.672	-15.759	***
10	-16.167	4.941	-3.27	.003	-26.414	-5.92	***
11	-11.237	4.844	-2.32	.03	-21.282	-1.192	**
12	-3.105	6.091	-0.51	.615	-15.738	9.528	
Panel B: Year Fixed Effects (Base 2017)							
2018	-2.137	3.067	-0.70	.493	-8.498	4.224	
Panel B: Covariates							
value	13.125	3.192	4.11	0	6.505	19.745	***
m_net_change	-19.534	6.899	-2.83	.01	-33.842	-5.227	***
mo_perc_change	.285	.388	0.74	.47	-.519	1.09	
pop_month_est	.001	.001	2.09	.049	0	.002	**
Panel C: Charlotte in Month X (Base Charlotte in December)							
month#1 : base 1	3.952	6.089	0.65	.523	-8.676	16.579	
2	-13.492	6.748	-2.00	.058	-27.487	.503	*
3	-2.947	6.309	-0.47	.645	-16.032	10.138	
4	1.585	5.48	0.29	.775	-9.781	12.951	

5	2.832	4.538	0.62	.539	-6.58	12.244	
6	4.984	3.932	1.27	.218	-3.171	13.139	
7	5.481	3.42	1.60	.123	-1.612	12.575	
8	6.307	2.995	2.11	.047	.095	12.519	**
9	5.983	2.874	2.08	.049	.023	11.944	**
10	-.713	2.115	-0.34	.739	-5.099	3.673	
11	-4.94	1.471	-3.36	.003	-7.991	-1.889	***
Constant	-91.522	67.425	-1.36	.188	-231.354	48.31	

Mean dependent var	87.529	SD dependent var	30.866
R-squared	0.480	Number of obs	552
F-test	.	Prob > F	.
Akaike crit. (AIC)	4475.930	Bayesian crit. (BIC)	4544.946

*** *p<.01,* ** *p<.05,* * *p<.1*