Online Appendix to

"An Analyst by Any Other Surname: Surname Favorability and Market Reaction to Analyst Forecasts"

Jay Heon Jung
Cass Business School

Alok Kumar
University of Miami

Sonya S. Lim
DePaul University

Choong-Yuel Yoo
KAIST College of Business

February 9, 2019

In this online appendix, we report results for several additional tests. Specifically, we identify underlying factors that form the favorability of a surname. We also examine the impact of surname favorability on post-revision price drifts. Lastly, we provide the robustness of our findings to using an alternative data source and measure of full name favorability.

1. Decomposing Surname Favorability

Several factors can influence surname favorability. For example, Americans' in-group bias against foreigners may be an underlying factor for surname favorability (e.g., Kumar et al., 2015). Also, cultural or ethnic similarities between the United States and the country associated with a surname or the perceived level of the country's corruption are likely to affect surname favorability.

Following prior studies (Hwang, 2011; Kumar et al., 2015), we construct five variables to capture the underlying factors of surname favorability: *Foreignness* captures whether the name of an analyst sounds foreign from American perspective. *Same ancestry* captures the similarity in ancestry between U.S. citizens and people in the countries associated with an analyst's surname. *Same language* captures the proportion of countries associated with an analyst's surname that speak English as the official language. *Cultural distance* captures the mean absolute difference in six dimensions of the Hofstede cultural index between the United States and the countries associated with an analyst's surname. Lastly, *Country corruption* captures the mean level of perceived corruption in countries associated with an analyst's surname.

We first examine the relation between surname favorability and the five aforementioned variables. Panel A of Table O.1 reports results from pooled OLS regressions in which the dependent variable is Americans' favorability of an analyst's surname (*FavSurname*), and the independent variables are each or all of the five underlying factors. Consistent with our conjecture, we find that surname favorability is positively associated with the ethnic and linguistic similarities

between the United States and an analyst's country of origin reflected in his or her surname, whereas it is negatively associated with the foreignness of an analyst's name, cultural distance, and the level of perceived corruption in the countries associated with the surname.

Next, we investigate which underlying factor of surname favorability plays the most important role in driving our findings. Using the coefficient estimates and the residual obtained from the pooled OLS regression of *FavSurname* on *Foreignness, Same ancestry, Same language, Cultural distance,* and *Country corruption*, we measure the six components of surname favorability explained by each of these five factors and the residual. For example, we measure the component of surname favorability explained by the foreignness of an analyst's name, *FavSurname* (*Foreignness*), as the coefficient estimate of *Foreignness* times the value of *Foreignness*. The residual component, *FavSurname* (*Residual*), is measured as the residual obtained from the regression.

Panel B of Table O.1 reports results for the baseline regressions using the six components of favorability. In column (1), we find that the component of surname favorability associated with the foreignness of an analyst's name has a significant and positive effect on market reactions to forecast revisions. This result suggests that Americans' in-group bias triggered by foreign-sounding names is an important driver for the estimated effects of surname favorability. We do not find significant effects of favorability components associated with other four factors.

Interestingly, we find significant results using the residual component of surname favorability, *FavSurname* (*Residual*), even after controlling for all other favorability components.[1] The evidence implies that the component of favorability not explained by foreignness, cultural proximity, and country corruption still has important explanatory power for investor reactions to

[1] In untabulated tests, we include these five factors as control variables in our baseline regression model. We still find significant results for the surname favorability effect.

forecast revisions. This finding clearly distinguishes our work from prior studies examining the effect of in-group bias and cultural proximity in the capital market.

2. Post-revision Price Drift

We now examine whether surname favorability mitigates investors' underreaction to forecast revisions. Gleason and Lee (2003) show that analysts' forecast revisions predict future abnormal returns over the next six months or one year, which is referred to as the post-revision price drift. If the market underreacts to forecast revisions, we should find that surname favorability is associated with weaker underreaction, because it has a positive effect on the immediate reaction to forecast revisions.

Following Gleason and Lee (2003), we measure the abnormal drift return to a forecast revision as the size-adjusted buy-and-hold return (*BHAR*) over the window from trading day +*m* (*m* = 2 or 11) to trading day +*n* (*n* = 127 or 253), where trading day 0 is the forecast revision date.[2] In Table O.2, we estimate the post-revision price drift regression in which the dependent variable is an abnormal drift return and the main independent variable of interest is the interaction term between *Revision* and *High FavSurname*. *High FavSurname* equals 1 if an analyst has a surname that ranks in the top tercile of surname favorability among all analysts in a calendar quarter and 0 otherwise.

In columns (1) to (4), we report results using our full sample. We do not find significant post-revision price drifts using the full sample, making it difficult for us to make a clear conclusion about the effect of surname favorability on investor underreaction.[3] Thus, as the next step, we

[2] We use *BHAR* as *CAR* implicitly assumes daily rebalancing and may lead to an upward bias in the return over long periods (Roll, 1983). In untabulated tests, we find that our results are robust to using *CAR* for the post-revision price drift tests.

[3] One potential explanation for not finding a significant drift using the full sample could be because our sample period spans more recent years from 1996 to 2014. Gleason and Lee (2003) use data in the 1990s. Another possible

confine our tests to a subsample of firms and analyst forecasts that are likely to have a significant underreaction to forecast revisions. According to Gleason and Lee (2003) and Zhang (2006), the post-revision price drift is likely to be greater as a firm's information environment is less transparent, when forecasts are issued by analysts who are less visible (i.e., non-all-star), and when forecasts provide more information to investors (i.e., non-herding forecasts). Thus, we construct a subsample by excluding firms in the top tercile of total analyst coverage or institutional ownership and dropping herding forecasts and forecast revisions issued by all-star analysts.[4] We re-estimate the drift regression using the subsample and report results in columns (5) to (8) of Table O.2.

Coefficients on *Revision* are positive and statistically significant across all four columns, suggesting significant drifts in this subsample. More importantly, we also find negative and significant coefficients on *Revision*×*High FavSurname*, suggesting that surname favorability is associated with a weaker underreaction to forecast revisions.

The results in Table O.2 show that surname favorability also affects price efficiency within a subsample with significant post-revision price drifts. However, the fact that we do not find significant post-revision price drifts nor the effect of surname favorability on drifts in the overall sample raises a question about the extent to which the surname favorability affects market mispricing.

3. Robustness to Alternative Measures

We further test the robustness of our findings to using an alternative measure and a different data source. First, we use a name-based ethnicity classification provided by OnoMap,

explanation is that we include more control variables and fixed effects.

[4] We find a significant surname favorability effect for the market's immediate reaction to forecast revisions in this subsample (regression models in Panel A of Table 4). In untabulated tests, we find that excluding firms in the top quartile or quintile of analyst coverage or institutional ownership does not change our results. In addition, we find that results are robust to removing the conditions on total analyst coverage or herding forecasts.

which derives data from telephone directories and electoral registers in 26 countries (Mateos et al., 2011). Using the ethnicity classifications in OnoMap, we construct an alternative measure of surname favorability, *FavSName* (*OnoMap*), and find that our main findings remain qualitatively the same. The results are provided in Panel A of Table O.3. Second, using a hand-collected dataset of analysts' full names and the ethnicity classifications in OnoMap, we construct a measure of full name favorability, *FavFName* (*OnoMap*). We find that our results are robust to measuring the favorability of an analyst's full name using a different data source. The results are provided in Panel B of Table O.3.

References

Gleason, C. A., Lee, C. M. C., 2003. Analyst forecast revisions and market price discovery. *Accounting Review* 78, 193-225.

Hwang, B., 2011. Country-specific sentiment and security prices. *Journal of Financial Economics* 100, 382-401.

Kumar, A., Niessen-Ruenzi, A., Spalt, O. G., 2015. What's in a name? Mutual fund flows when managers have foreign-sounding names. *Review of Financial Studies* 28, 2281-2321.

Mateos, P., Longley, P. A., O'Sullivan, D., 2011. Ethnicity and population structure in personal naming networks. *PLoS ONE* 6, e22943.

Roll, R., 1983. On computing mean returns and the small firm premium. *Journal of Financial Economics* 12, 371-386.

Zhang, X. F., 2006. Information uncertainty and stock returns. *Journal of Finance* 61, 105-137.

Table O.1
Decomposition of surname favorability

Panel A: Potential underlying factors of surname favorability

	Dependent variable: FavSurname					
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)
Foreignness	-0.174***					-0.021***
	(-26.026)					(-5.362)
Same ancestry		0.989***				-0.037
		(52.485)				(-1.252)
Same language			0.167***			-0.093***
			(57.277)			(-13.943)
Cultural distance				-0.010***		-0.014***
				(-86.795)		(-28.657)
Country corruption					-0.054***	-0.003*
					(-47.952)	(-1.677)
Intercept	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	31,443	31,443	31,443	31,443	31,443	31,443
Adjusted R^2 (%)	25.12	43.42	46.92	72.47	54.57	76.05

Panel B: Effects of individual components of surname favorability on market reactions to forecast revisions

	Dependent variable: Size-adjusted CAR [-1, +1]						
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Revision	1.144***	1.086***	1.109***	1.130***	0.941***	1.139***	1.488***
	(4.896)	(4.572)	(4.676)	(4.839)	(3.675)	(4.847)	(4.093)
Revision×FavSurname (*Foreignness*)	4.011*						4.525*
	(1.956)						(1.793)
Revision×FavSurname (*Same ancestry*)		-0.576					-1.668
		(-0.341)					(-0.619)
Revision×FavSurname (*Same language*)			0.102				1.383*
			(0.310)				(1.724)
Revision×FavSurname (*Cultural distance*)				0.124			0.622
				(0.913)			(1.628)
Revision×FavSurname (*Country corruption*)					-3.611		0.669
					(-1.640)		(0.170)
Revision×FavSurname (*Residual*)						1.027***	0.973***
						(4.007)	(3.900)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	753,714	759,544	762,780	762,244	762,568	750,044	750,044
Adjusted R^2 (%)	9.74	9.73	9.73	9.73	9.73	9.74	9.75

Table O.1 (*Continued*)
Decomposition of surname favorability

This table shows results from pooled OLS regressions. In Panel A, we estimate OLS regressions in which the dependent variable is Americans' favorability of an analyst's surname (*FavSurname*), and the independent variables are five potential underlying factors of surname favorability and year fixed effects. *Foreignness* is the percentage of the Amazon Mechanical Turk (AMT) workers who indicate that the name of the analyst is foreign sounding (Kumar et al., 2015). *Same ancestry* is the weighted average of the percentage of U.S. citizens whose ancestors came from countries associated with the analyst's surname. We obtain the ancestry of the U.S. citizens from the U.S. Census Bureau. *Same language* is the weighted average of *English* dummy for countries associated with the analyst's surname. The *English* dummy variable is equal to 1 if English is the official or the most popular language for a country according to the World Factbook by the Central Intelligence Agency (CIA) and 0 otherwise. *Cultural distance* is the weighted average of the culture difference for countries associated with the analyst's surname. The culture difference is measured as the mean value of the absolute differences in the Hofstede index between the United States and the country in question, across all six cultural dimensions. *Country corruption* is the weighted average of negative one times the Corruption Perception Index (CPI) published by Transparency International for countries associated with the analyst's surname. Weights are computed based on the frequency of the nationality of U.S. immigrants who have the same surname as an analyst. In Panel B, we estimate the pooled OLS market reaction regressions in which the dependent variable is the size-adjusted cumulative abnormal return (*CAR*) over the window from trading day -1 to trading day +1, where trading day 0 is an analyst's forecast revision date. *Revision* is the difference between an analyst's current and preceding earnings forecast for a firm, scaled by the stock price two trading days prior to the current forecast date. *FavSurname* (*Component*) is measured as the value of *Component* times the coefficient estimate of *Component*, obtained from the pooled OLS regression of *FavSurname* on *Foreignness*, *Same ancestry*, *Same language*, *Cultural distance* and *Country corruption*. *FavSurname* (*Residual*) is the residual value obtained from the OLS regression. In Panel B, for brevity we suppress coefficient estimates of the stand-alone variables, *FavSurname* (*Component*) and *FavSurname* (*Residual*). The set of control variables is identical to that in Panel A of Table 4. In parentheses below coefficient estimates are *t*-statistics based on standard errors clustered by analyst. All continuous variables are winsorized at the 1st and 99th percentiles. Variable definitions are provided in Appendix C. ***, **, * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

Table O.2
Post-revision price drift

	Dependent variable: Size-adjusted BHAR							
Return window:	[+2,+127]	[+11,+127]	[+2,+253]	[+11,+253]	[+2,+127]	[+11,+127]	[+2,+253]	[+11,+253]
Sample composition:	Full sample				Subsample with significant drifts			
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Revision	-0.570	-0.590	0.749	0.450	2.201**	2.661***	5.607***	6.121***
	(-0.952)	(-1.040)	(0.844)	(0.517)	(2.382)	(3.028)	(4.304)	(4.799)
Revision×High FavSurname	-0.044	-0.057	-0.087	-0.119	-0.209*	-0.201**	-0.324**	-0.330**
	(-0.549)	(-0.751)	(-0.713)	(-0.997)	(-1.951)	(-1.990)	(-2.029)	(-2.137)
High FavSurname	-0.000	0.000	-0.002	-0.002	0.000	0.000	-0.000	-0.000
	(-0.051)	(0.039)	(-0.847)	(-0.809)	(0.094)	(0.036)	(-0.062)	(-0.137)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	762,780	762,780	762,780	762,780	195,530	195,530	195,530	195,530
Adjusted R^2 (%)	15.12	14.16	23.86	23.16	21.88	20.52	32.61	31.67

This table shows results from pooled OLS regressions. The dependent variable is the size-adjusted buy-and-hold abnormal return (*BHAR*) over the next six-month or one-year window, starting from trading day *m* (*m* = 2 or 11) and ending on trading day *n* (*n* = 127 or 253), where trading day 0 is an analyst's forecast revision date. *Revision* is the difference between an analyst's current and preceding earnings forecast for a firm, scaled by the stock price two trading days prior to the current forecast date. *High FavSurname* is an indicator variable that equals 1 if an analyst ranks in the top tercile by surname favorability (*FavSurname*) among all analysts in a calendar quarter and 0 otherwise. In columns (5), (6), (7) and (8), we estimate the OLS regressions for post-revision price drifts after excluding firms in the top tercile of total analyst coverage or institutional ownership and dropping herding forecasts or forecasts issued by all-star analysts. We define a forecast as herding if it is between the analyst's own prior forecast and the consensus (Gleason and Lee, 2003). The set of control variables is identical to that in Panel A of Table 4. In parentheses below coefficient estimates are *t*-statistics based on standard errors clustered by analyst. All continuous variables are winsorized at the 1st and 99th percentiles. Variable definitions are provided in Appendix C. ***, **, * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

Table O.3
Alternative measures

Panel A: OnoMap-based measure of surname favorability

	Dependent Variable: Size-adjusted CAR							
	[-1,+1]		[-1,+3]		[-1,+5]		[-1,+10]	
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Revision	0.889***	0.892**	0.978***	0.984***	1.052***	1.039***	1.007***	0.879**
	(5.334)	(2.428)	(5.863)	(2.615)	(6.420)	(2.636)	(6.112)	(2.099)
Revision×FavSName (*OnoMap*)	0.416**	0.607***	0.390*	0.588***	0.325	0.542***	0.354*	0.587***
	(2.058)	(3.019)	(1.922)	(2.925)	(1.630)	(2.712)	(1.763)	(2.837)
FavSName (*OnoMap*)	0.001	0.001	0.001	0.001	0.000	0.000	-0.001	-0.001
	(0.902)	(0.914)	(0.622)	(0.597)	(0.168)	(0.144)	(-0.486)	(-0.363)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	436,821	436,821	436,821	436,821	436,821	436,821	436,821	436,821
Adjusted R^2 (%)	7.90	9.93	7.36	9.46	6.84	9.01	5.90	8.33

Panel B: OnoMap-based measure of full name favorability

	Dependent Variable: Size-adjusted CAR							
	[-1,+1]		[-1,+3]		[-1,+5]		[-1,+10]	
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Revision	0.456**	0.073	0.493***	0.110	0.598***	0.213	0.540***	-0.167
	(2.457)	(0.220)	(2.631)	(0.321)	(3.207)	(0.600)	(2.580)	(-0.427)
Revision×FavFName (*OnoMap*)	0.909***	1.143***	0.943***	1.199***	0.840***	1.127***	0.874***	1.208***
	(4.126)	(5.425)	(4.223)	(5.550)	(3.781)	(5.189)	(3.525)	(4.953)
FavFName (*OnoMap*)	0.001	0.001	0.001	0.001	0.000	0.000	-0.001	-0.001
	(0.849)	(0.740)	(0.550)	(0.438)	(0.327)	(0.149)	(-0.574)	(-0.592)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	604,536	604,536	604,536	604,536	604,536	604,536	604,536	604,536
Adjusted R^2 (%)	7.87	9.91	7.28	9.41	6.73	8.93	5.79	8.22

This table shows results for the market reaction regression estimates, using alternative measures of name favorability. We estimate the baseline OLS regressions in which the dependent variable is the size-adjusted cumulative abnormal return (*CAR*) over the window from trading day -1 to trading day *n* (*n* = 1, 3, 5, and 10), where trading day 0 is an analyst's forecast revision date. *Revision* is the difference between an analyst's current and preceding earnings forecast for a firm, scaled by the stock price two trading days prior to the current forecast date. In Panel A, we construct an alternative measure of surname favorability, *FavSName* (*OnoMap*), using the surname-ethnicity classifications in OnoMap. In Panel B, we construct an alternative measure of full name favorability, *FavFName* (*OnoMap*), using the full name-ethnicity classifications in OnoMap. Control variables and fixed effects are identical to those in Panel A of Table 4. In parentheses below coefficient estimates are *t*-statistics based on standard errors clustered by analyst. All continuous variables are winsorized at the 1st and 99th percentiles. Variable definitions are provided in Appendix C. ***, **, * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.