Routine Empirical Choices Affect Replicability and Inference: Evidence from Mutual Fund Flows Research

John Adams

April 6, 2025

Abstract

Kumar, Niessen-Ruenzi, and Spalt (2015) report that mutual fund managers with foreign-sounding names attract less investor flow, a pattern which is consistent with taste-based discrimination. While I can reproduce their main finding using their sample, the result does not hold under independent sample construction, alternative name classifications, and outlier-robust methods. My analysis finds that the original result is sensitive to a small number of extreme observations and classification decisions. This highlights how routine empirical choices can affect replicability and inference, and underscores the importance of economically motivated design strategies, particularly in filtering and classification.

Keywords: Replication; research methods; discrimination; mutual fund performance; flows
JEL Classifications: C18, G11, G20, J71

Adams is an associate professor of finance at the University of Texas at Arlington (email: jcadams@uta.edu). I thank Yi-Ju Chien for invaluable research assistance. I also thank Ivo Welch, an anonymous referee, Shiu-Yik Au, Spencer Barnes, Susan Crane, Arati Kale, Juhani Linnainmaa, Yu Liu, Sattar Mansi, Takeshi Nishikawa, David Rakowski, and seminar participants at the 2020 Southern Finance Conference, the 2024 Southwestern Finance Conference, and at University of Texas at Arlington for insightful comments and suggestions.

Seemingly minor empirical choices can meaningfully affect conclusions in financial research. I revisit Kumar, Niessen-Ruenzi, and Spalt (2015), who observe that mutual fund managers with 'Foreign'-sounding names attract less investor flow, which is consistent with taste-based discrimination. This paper does not focus on evaluating their discrimination hypothesis. Rather, I use their study to examine how methodological decisions influence replicability and inference. Substantial variation in reported fund flows across studies covering similar periods (Table 1) suggests that empirical outcomes often reflect researcher decisions as much as the underlying data.

[Table 1 here]

I can reproduce the main result of Kumar, Niessen-Ruenzi, and Spalt (henceforth KNS) using their dataset. However, their result does not hold when I do any of the following: (1) trim outliers differently, (2) construct independent samples, (3) apply outlier-robust regression techniques, (4) reclassify manager names with a new survey (i.e., foreign-sounding or not), or (5) use an AI-based name classification tool. In sum, the KNS result hinges on a small number of high-flow observations remaining in the sample and being classified differently than they are by the AI tool or another survey.

I follow Welch (2019) to construct a reproduction sample following their data descriptions and a replication sample that mitigates potential data and sampling errors identified in the initial reanalysis. Like KNS, I conduct an online survey to classify fund manager names as being familiar (i.e., American) or foreign sounding. I closely match, within 0.1 percentage points (4.6% versus KNS 4.5%) for the replication sample and within 1 percentage point (5.5%) for the reproduction sample, their average incidence of funds managed by individuals with foreign sounding names.

I find that survey-based classifications are inconsistent, with foreign-sounding designations differing across studies (KNS, my reanalysis) and from an artificial intelligence taxonomy that infers cultural or geographic name origins. These inconsistencies highlight challenges in using survey-based classifications in empirical finance, particularly for measuring subjective perceptions. Consistent with prior research, I use artificial intelligence taxonomy classifications as a proxy for U.S. perceptions of name foreignness. Baseline regressions show that the KNS finding disappears when using artificial intelligence classifications. Reclassifying a single high-flow manager's name eliminates statistical significance, suggesting that the discrimination effect is fragile and depends on subjective classifications and outlier influence.

I am unable to corroborate the KNS finding with the reproduction and replication samples. The main cause for the difference in KNS and my results appears to be how observations with extremely large fund flows are handled. KNS appear to drop the smallest and largest 1% of flow observations from their full dataset (or similar level) that includes observations that are not included in their analysis due to missing fund information. This choice results in the retention of a few extreme observations, all from American-sounding name funds. Their findings do not hold when the smallest and largest 1% of flow observations are instead excluded from the testable sample that contains all necessary variables for their main regressions.

This paper contributes to the empirical finance and replication literature by using a prominent behavioral finance paper as a case study in methodological sensitivity. Rather than focusing solely on the original claim that investors discriminate against fund managers with foreign-sounding names, I examine how standard empirical decisions—such as outlier treatment, sample construction, and classification strategies—shape the robustness and replicability of findings. The findings underscore how routine empirical decisions can materially affect research outcomes and offer guidance on implementing economically motivated filtering rules rather than relying on arbitrary statistical cutoffs. My work complements recent large-scale research by Gelman et al. (2023), which demonstrates that even when researchers test the same hypothesis using a shared dataset, their conclusions often diverge due to differences in empirical design. This variation underscores the need for theoretically grounded research strategies.

The remainder of the study is organized as follows. Section 2 describes the sample construction processes, variable definitions, and descriptive statistics. Section 3 provides my reanalysis using the KNS, reproduction, and replications samples on the relation between manager names and fund flows. Section 4 reconciles KNS and my findings. Section 5 concludes with recommendations for future studies.

2 Data

2.1 Sample Constructions

KNS provide their full sample of 4,804 funds and 29,001 fund-year observations.[1] The full sample includes trimmed fund flows (at the 1% level), managers foreign name classifications, and other variables to reproduce their baseline regression. The full sample reduces to 2,553 funds and 13,091 useable observations after dropping observations with missing data. For ease of exposition, the reduced file is hereafter labeled their final sample. KNS also provide a mapping file linking their manually created portfolio codes to CRSP share class identifiers. The KNS sample does not include flows estimates for trimmed or dropped observations, but KNS do provide their code to estimate fund flows. However, at the time of this writing they were unable to provide their source data and the complete code used to construct their full and final samples.

KNS's Table 2, Column 3 (on page 2290 of their published paper) presents baseline coefficient estimates–their main result–obtained from regressing fund flow on a foreign sounding name dummy and fund attributes using their final sample. My reanalysis focuses on reproducing and replicating this baseline result using data provided by KNS and sample reconstructions following their data descriptions. Specifically, I construct three samples. The first is the KNS reproduction sample that uses the KNS manually created portfolio codes and KNS manager foreign name classifications. The second reproduction sample uses my own manually created portfolio codes and foreign manager name classifications. The two reproduction samples, hereafter KNS reproduction and reproduction, use the CRSP mutual fund database for fund information. The third sample, hereafter the replication sample, replicates KNS by intersecting the CRSP and Morningstar databases to improve data accuracy (Berk and van Binsbergen 2015; Pástor, Stambaugh, and Taylor 2015) and using my own foreign name classifications. Unless otherwise noted, the flows are trimmed at the 1% level in the reproduction and replication samples to match KNS's final sample. Detailed descriptions of the reproduction and replication construction processes are in this paper's appendix. The KNS final sample, KNS reproduction, reproduction and replication sample sizes vary considerably.

The differences in sample sizes appear driven by how KNS and I define domestic equity funds and how funds are classified as team or solo-managed. The KNS reproduction and my reproduction

[1] KNS report 4,805 funds in their full sample (their page 2287) but the file they provide includes two observations with missing fund portfolio codes. There are 4,804 unique fund portfolio codes in their full sample.

sample processes are outlined in Appendix Tables 1 and 2. KNS reports 4,804 domestic equity funds and provides a portfolio code-to-share-class mapping file, which I use in Step 2 of Appendix Table 1. In contrast, I identify 9,120 domestic equity funds with my portfolio codes (Step 6, Appendix Table 2). KNS's criteria for defining domestic equity funds are not explicitly stated, but their approach appears more restrictive than mine. My methodology identifies a broader set of domestic equity funds, likely leading to the larger sample size. After annualizing, my sample remains about twice as large (8,609 vs. 4,804, Step 7, Appendix Table 2).

In Step 6 of the KNS reproduction (Appendix Table 1), most funds are retained. In contrast, my reproduction process requires team-managed funds to have a clearly identified lead manager, resulting in a sample reduction of about 50%. At this stage, my sample is comparable to KNS (4,049 vs. 4,430 funds). However, prior to dropping observations with missing data my final reproduction sample contains 17,422 observations (Step 8, Appendix Table 2), significantly fewer than the 25,104 observations in KNS (Step 6, Appendix Table 1). The difference may stem from how fund management changes over time are recorded. I track solo or team-managed status annually, while KNS's approach is unclear. They focus on solo-managed funds, but I cannot approximate their numbers without including lead managers from team-managed funds.

Finally, I retain 1,447 funds (Step 9, Appendix Table 2) after dropping funds with missing data, keeping 36% of funds from Step 8, compared to the KNS reproduction's 52% retention rate.

In sum, only the KNS reproduction sample approximates the KNS provided sample's size. The differences appear mostly driven by how KNS and I define domestic equity funds and how funds are classified as team or solo-managed. There are fewer funds, managers, and fund-year observations in the reproduction and replication samples.

2.2 Detailed Flows Statistics

Table 2 reports detailed annual flow statistics for the KNS final, KNS reproduction, reproduction, and replication samples. The incidences of foreign manager name funds (reported in brackets) are similar across the samples at 4.5%, 4.6%, 5.5%, and 4.6%. The mean annual flows values for each sample range from .6% per year for the replication sample to 20% per year for the KNS final sample. The mean annual flow for the KNS sample is 20.5% for other ('American' sounding manager names) funds compared to approximately 11% for the KNS reproduction sample and 12.6% for the reproduction sample. However, the most notable difference in flows across the four samples are the maximum values and skewness. The KNS final sample exhibits much higher positive skewness in annual flows estimates than the other samples.[2] The high degree of skewness occurs in their other funds sub-sample (8.132) which has a maximum flow value of 17.457 (1,745.7% of prior year net assets). Also, the 99th percentile flows for the KNS final sample is 4.014, which is greater than or equal to the maximum values for the reproduction and replication samples.

[Table 2 here]

The median values for funds with foreign- and American-sounding manager names in the KNS sample are similar. Figure 1 plots the median and maximum values for the KNS name classifications over their sample period. The maximum annual flow values for funds with American-sounding manager names consistently exceed those of funds with foreign-sounding manager names. In contrast, median flows between the two groups show no noticeable difference. Thus, the large flow means and differences reported in KNS appear to be driven by extreme positive outliers in funds managed by individuals with American-sounding names.

[Figure 1 here]

3 Reproduction and Replication

KNS provide univariate results in their Table 1, which reports differences in fund flows and other fund attributes between managers with foreign- and non-foreign sounding names. Their main results are in their Table 2, which presents coefficient estimates obtained from regressing fund flow on a foreign sounding name dummy and fund attributes.

3.1 KNS Table 1: Univariate Differences

Table 3 presents selected results of the KNS Table 1 univariate difference tests using the four samples. Columns 1 and 2 report mean characteristics for foreign and other (non-foreign or 'American' sounding) manager name funds while Columns 3 and 4 provide differences and t-statistics. KNS reports that funds managed by individuals with foreign sounding names have 10.9 percentage points

[2] Flows are naturally positively skewed because the lower limit is around -100% and, theoretically (although not realistically) there is no upper limit.

lower annual fund flows which I confirm using their final sample (KNS). In contrast, the results from the other samples do not indicate foreign funds receive significantly less flow.

[Table 3 here]

The reproduction and replication results for the fund attributes also differ. For example, foreign funds in the reproduction sample attract more flow, exhibit better performance (PRank), are younger, and have higher Morningstar ratings. The mean Morningstar rating is about 3.4, which is consistent with Morningstar's quantitative performance classification scheme that assigns five stars to the top 10% of funds according to risk-adjusted performance, the next 22.5% are assigned four stars, the middle 33% get three stars, 22.5% receive 2 stars, and the lowest performing 10% of funds get a one-star rating. KNS reports mean Morningstar ratings of about 1.8, which suggests the funds in their sample are, on average, poor performers. However, they report PRank (performance rank by style segment ranging from 0 to 1) mean values of 0.47 and 0.48 indicating their average fund nearly matches the performance of the average domestic equity fund. I examine whether KNS replaced missing ratings with zeros and compute an average rating of about 2.3. KNS does not address this anomaly, but they also do not include ratings as a control for their baseline regression. In sum, the published univariate flows results can only be reproduced using the KNS sample.

3.2 KNS Table 2: Regressions

Table 4 reproduces and replicates the KNS Table 2 baseline specification (main result) by regressing fund flow on Foreign75 and fund attributes using the four samples. Column 1 reports results using the KNS final sample. Columns 2, 3, and 4 provide results for the KNS reproduction, reproduction, and replication samples. The Foreign75 and control variables estimated coefficients in Column 1 exactly match those reported in KNS (their Table 2, Column 3 result on page 2290).

However, the estimated coefficient for Foreign75 is insignificant using the other samples. The KNS and the other samples' results are mostly similar for prior performance (PRank or PRank^2), fund size, fund age, and lagged fund flow. The results suggest that investors direct flows to better performing funds, smaller funds, younger funds, and to growing funds.

[Table 4 here]

In sum, the KNS main finding (Foreign75) can be exactly reproduced using their sample data. However, their results do not hold for the KNS reproduction, reproduction, and replication samples.

4 Reconciliation

This section reconciles my results with those of KNS. However, due to the unavailability of the complete code and source data, it is not possible to definitively identify all sources of discrepancy. Instead, my reconciliation analysis focuses on determining the most likely first-order explanations.

4.1 Outlier Robustness

Table 1 indicates that the KNS mean equity fund flow estimate is notably higher than in other studies, while Table 2 shows that fund flows in the KNS sample exhibit much greater positive skewness than those in the reproduction and replication samples. This discrepancy suggests that KNS flow estimates may be influenced by extreme observations. Building on this, Table 4 further reveals that the estimated coefficient for Foreign75 is negative and significant only in the KNS final sample regression. This finding suggests that the observed negative relationship between foreign-sounding manager funds and investor flows may be driven by a subset of influential outliers—observations whose removal renders estimates insignificant—that persist despite KNS's exclusion of the most extreme 1% of flows. Alternatively, there is no theoretical reason why influential observations are limited to 1% of each side of the flows distribution so trimming at the 1% level may simply be an ineffective outlier mitigation strategy.

It is well known that OLS regression estimates are sensitive to skewness and extreme outliers. To test whether outliers drive the KNS results, I follow Lou (2012) by applying the natural logarithm of (1 + fund flow) to reduce skewness and mitigate the undue influence of extreme values. Additionally, I conduct outlier-robust regressions using fund flow as the dependent variable, as in Table 4, Column 1, to assess the robustness of the results.[3]

Table 5 presents the regression estimates using the KNS final sample. For ease of comparison, Column 1 reports the KNS published and confirmed main result (and the skewness of fund flow from

[3] See Adams et al (2019) for a discussion and comparison of techniques to identify and treat outliers. Quantile regressions provide unbiased results in the presence of large y values (here fund flows) and MM regressions offer protections against extreme y and x values (here flows and control variables).

Table 2), while Columns 2-5 present results from alternative specifications, including log-transformed flows (OLS), quantile (median) regression, and MM regressions at two efficiency levels (28.7% and 95% relative to OLS). Columns 3-5 use fund flow as the dependent variable. All specifications include the controls in Table 4. The estimated coefficients on Foreign75 are not negatively significant across the outlier-robust specifications, indicating that the original finding is not robust.

[Table 5 here]

Thus, the KNS conclusion that foreign-sounding manager funds receive lower investor flows appears to be driven by extreme flow estimates. The negative relationship between foreign-sounding manager funds and investor flows disappears when using outlier-resistant specifications, suggesting that the original published result was an artifact of outliers rather than evidence of a meaningful economic relationship.

4.2 Outlier Mitigation

Because flows are trimmed at the 1% level (i.e., dropping observations with the smallest and largest 1% of flows) in the four samples, it is possible that how the sample construction processes identify outliers drive the disparate results. Therefore, I next examine how trimming outliers at different stages of the sample construction process affects flows estimates. Table 6 presents detailed flows statistics and estimated regression coefficients for the KNS full sample prior to trimming, for the KNS final sample that produces their main result, with 1% outlier trimming at the KNS full sample level, and with 1% trimming of the KNS usable sample observations. I use the KNS provided code and fund returns and asset values from CRSP to estimate flows for their full sample and confirm their estimates are accurate.[4]

[Table 6 here]

[4] KNS annualizes CRSP monthly share class returns and then takes the prior year TNA weighted averages to aggregate returns to the fund (portfolio) level. KNS then estimates fund level flows. In contrast, my reproduction and replication samples aggregate share classes monthly and then annualizes prior to estimating flows. Appendix Table 5 reports almost identical (and statistically insignificant) Foreign75 coefficient estimates using both approaches for observations common to the KNS final and KNS reproduction samples.

Panel A reports detailed flows statistics for the KNS provided sample data. Their full sample of 4,804 funds reduces to 4,362 funds with estimable flows (KNS full sample). The KNS final sample produces the published main results. Panel A also reports flows statistics after 1% trimming at the full sample level and after 1% trimming at the useable sample level. For trimming at the full sample level, the smallest and largest 1% of flows (1% each side) are dropped after which observations with missing variables necessary to run the baseline regression are dropped. The KNS 1% trimming at the useable sample level observations are obtained by first dropping observations without the necessary variables to run the baseline regression (i.e., dropping unusable observations) and then identifying and dropping the smallest and largest 1% of the remaining flow observations. Panel A also provides flow estimates for the intersection of the KNS final and the KNS 1% trimming at the useable level samples.

The flow estimates of the KNS full sample are highly skewed–possibly due to confounding events and data errors–and justify an outlier mitigation strategy. The KNS final sample's maximum (minimum) flow is 17.457 (–0.982) which is comparable to the KNS full sample's 99th (1st) percentile. The KNS final sample and 1% trimming the full sample level flows have similar mean, standard deviation, skewness, and other distribution values. Thus, it appears likely that KNS drops the smallest and largest 1% flows observations from the full sample. The 1% trimming at the useable sample level treatment has lower mean flows that are less skewed. In terms of reconciliation, the 1% trimming at the useable sample level treatment has distributional values much like my KNS reproduction and reproduction samples reported in Table 2 (which also trim at the useable sample level).

In terms of observations, the 1% trimming at the full sample level treatment yields 2,764 funds and 14,046 observations compared to 2,553 funds and 13,091 observations for the KNS sample, while the 1% useable sample trimming treatment yields 13,752 observations. This suggests KNS drops funds from the full sample or dataset for reasons unrelated to missing data for the main baseline regression.

The maximum values for the different treatments are also informative with respect to their validity (i.e., whether they should be in the sample). The KNS final sample records a maximum flow value of 17.457, which appears unlikely to be solely driven by investor purchases. I manually examine this observation (Federated MDT Mid-Cap Growth Fund) and find the large flow estimate is not the result of investor flow but an unreported (in CRSP) fund merger that occurred in 2007.[5] In this case, the fund name points to a possible family level merger as the cause for the large flow. In a fund family merger, acquisition, or fund liquidation, assets may be reallocated internally, creating large inflows to certain funds without new investor capital (see, e.g., Lou 2012). While CRSP and Morningstar report official fund mergers, they do not report 'unofficial' ones like this example. This observation also reflected a potentially incubated fund in that it is less than three years old, net assets are less than $25 million, and it does not have a ticker symbol. Overall, the observed flow for this outlier observations does not appear reflective of typical investor behavior.

The maximum value for the 1% full sample trimming treatment of 13.561 occurred in 1997 (Dreyfus Midcap Value Fund). The fund's inception date was September 1995 and in 1997 its net asset value was less than $25 million, so it also represents a potential incubated fund. Incubated funds can be problematic because some or all the initial flows may come from the fund sponsor and not outside investors (Evans 2010). In sum, KNS appear to remove extreme flow observations at the beginning of the sample construction process instead of as the final step. This leads to excessively skewed flows estimates in the larger subsample of other funds (those whose managers have 'American sounding' names) and leaves some bad data in their final sample. Trimming at the full sample level would also explain how their main results are unaffected if they trim, winsorize, or use raw flows (page 2286 of their paper). This is because all three outlier mitigation strategies are ineffective if employed at the full sample level.

Panel B provides Foreign75 regression coefficients for the trimming treatments (all controls from KNS Table 2 included but not reported). As before, Column 1 reports the KNS final sample result (i.e., published and confirmed). Column 2 reports results for the 1% full sample trimming treatment and Column 3 uses the 1% useable sample trimming treatment. Column 4 provides results for

[5]Federated acquired MDT Advisors in 2006 and merged the fund in 2007. However, the merger is not reported in CRSP. See https://www.sec.gov/Archives/edgar/data/1363526/000131814807000036/form.txt.

observations that are common to the KNS final sample and the 1% trimming at the useable sample level.

The Foreign75 estimated coefficient for the 1% full sample trimming observations (Column 2) is close in economic and statistical significance to the published results (Column 1) despite the difference in sample sizes. However, the Foreign75 coefficients are insignificant with 1% useable sample trimming (Column 3) and the intersected KNS final sample and 1% useable sample trimming regression (Column 4).

The evidence is consistent with KNS dropping outlying observations at the full sample or higher stage of sample construction. This approach retains questionable observations with large positive flow estimates that unduly influence coefficient estimates. Trimming at the useable sample level (testable observations) yields fewer extreme flow estimates and minimal skewness. More importantly, the estimated coefficient on Foreign75 is insignificant when outliers are trimmed at the useable sample level. The reproduction and replication samples are likewise trimmed at the useable sample level and also produce insignificant Foreign75 coefficients. In conclusion, the different results in KNS and my reanalysis appears to be how outliers are dropped.

4.3 Influential Outliers

Panel C compares the 1% trimmed sample and the KNS final sample matched observations (n = 12,747) to unmatched observations that appear only in the KNS final sample (n = 344). Altogether, and consistent with KNS not applying filters to avoid incubation bias, observations removed by 1% trimming (Column 2) exhibit significantly higher flows and are younger and smaller. In unreported baseline regressions using the KNS final sample, I find that the estimated coefficient for Foreign75 remains negative and statistically significant after dropping funds younger than 3 years old and with TNA less than $25. Thus, incubation bias alone does not appear to drive the published Foreign75 result.

Another unreported regression finds that Foreign75 is insignificant after dropping funds without CRSP-assigned portfolio identifiers. In the KNS final sample, these identifiers are generally missing before 1999 but appear in 70% to 90% of funds thereafter. CRSP began assigning portfolio identifiers in July 2003, retroactively applying them to active multi-share-class funds. However, funds that were single-share-class before this date may not have received an identifier.

Funds more likely to be single-share-class and missing portfolio identifiers include institutional funds, separately managed accounts (SMAs), and collective investment trusts (CITs), which are not publicly traded or marketed to retail investors and often lack ticker symbols. In fact, over 20% of KNS observations from 1995 to 2000 lack tickers, suggesting they were likely unavailable to most retail investors. After dropping potentially incubated funds and restricting the sample to post-2000 observations, the estimated coefficient on Foreign75 becomes statistically insignificant.

In other unreported baseline regressions, I find 67 influential observations with fund flows exceeding 600%. That is, the published estimated coefficient on Foreign75 (–0.089 with a t-statistic of −4.06) reported in KNS Table 2 becomes insignificant after removing these 67 observations (out of a sample of 13,091). Of the 67, 25 were less than 3 years old, 43 had assets values less than $25 million, 8 had no tickers for one or more share classes, 8 are funds that are only in the sample once, 17 did not have a Morningstar rating for any share class, 33 did not have CRSP assigned portfolio identifiers, and 8 lacked 12 months of returns necessary to accurately estimate annual fund flows. This suggests many of the 67 funds were likely incubated or not widely available to retail investors. In addition, 8 fund observations appear to team-managed although KNS focuses on solo manager funds.[6] All the fund managers for the 67 influential observations have KNS assigned American sounding names.

4.4 Alternative Name Classifications Using Artificial Intelligence

The mean incidence of foreign sounding names in the replication and reproduction samples closely match, within 0.1% and 1%, the rate in KNS. This would suggest that overall survey respondents in my reanalysis classified names similarly as in KNS. However, KNS and this paper's survey respondents classified some names much differently. Classifications for about 92% of the names that appear in both surveys are the same. However, when classifying names as foreign sounding, the match rate drops to just 16%. This disparity is consistent with recent evidence that MTurk data quality is declining (Ahler, Roush, and Sood 2021; Kennedy et al 2020). If so, my survey's names classifications may be less accurate than those in KNS. It is also possible that MTurk data quality for mundane tasks such as name classifications is low in both surveys.

[6] I obtain manager names from CRSP for these observations because Morningstar.com does not provide them.

The next set of tests uses the Namsor artificial intelligence name classification taxonomy to identify foreign sounding names.[7] Namsor's accuracy has been validated by Science Metrix for Elsevier and the European Commission and several other studies.[8] Following LeRoy (2018), Miles (2004), and McLaren and Torres (1999), managers with names of Northern, Southern, and Western European origins are classified as familiar-sounding (i.e., American). For consistency with the surveys, Foreign 75 in this analysis takes on a value of one if the probability that a manager's names does not originate in Northern, Southern, or Western Europe exceeds 75%. About 4% of managers in the KNS final sample are classified as foreign using this approach.

Table 7 reports classification differences in Panel A which compares Foreign75 assignments of the KNS final sample (KNS MTurk Foreign75), my reproduction and replication samples (My MTurk Foreign75), and two Namsor assignment sessions for my samples and KNS, Namsor Foreign75 December 2022 and Namsor Foreign75 February 2025, respectively. There are 1,138 manager names common to the four classification schemes. Column 1 reports the mean and number of Foreign75 assignments and Columns 2-5 report the proportion of different Foreign75 classifications. Comparing the KNS MTurk Foreign75 classification to My MTurk Foreign75 classification, about 63% of the manager names identified as foreign-sounding by MTurk in KNS are deemed American-sounding by MTurk in my reanalysis. This compares to disagreement rates of 49% and 72% for Namsor in 2022 and 2025. For names classified as foreign-sounding in my MTurk survey, about 53% were identified as American-sounding by KNS. The Namsor 2022 and 2025 Foreign75 assignments are less different. For names classified as foreign-sounding by Namsor in 2025 (Namsor Foreign75 February 2025), only 19% were classified as American-sounding by Namsor in 2022. Overall, Panel A illustrates variability in identifying foreign-sounding names across classification methods and time.

---

[7] Following Mateos, Longley, and O'Sullivan, 2011, I also use the NamePrism API algorithm to find ethnicities for each manager name and find similar results. NamePrism (http://www.name-prism.com/) was developed by academics from Stony Brook University and researchers from Yahoo! Research, Amazon AI, and NEC Labs America. It has been used as an objective tool to determine ethnicities in over 300 research papers.

[8] Science-Metrix. (2018). Analytical Support for Bibliometrics Indicators. Open access availability of scientific publications (https://www.science-metrix.com/sites/default/files/science-metrix/publications/science-metrix_bibliometric_indicators_womens_contribution_to_science_report.pdf), Bursztyn et al. (2024), and Sebo (2022) validate the accuracy of NamSor.

[Table 7 here]

Panel B provides results for the baseline regression. Columns 1, 3, and 5 report selected results obtained from using MTurk to classify names (e.g., the results from this paper's Table 4), while Columns 2, 4, and 6 use Namsor to classify names. The KNS estimated coefficient for Foreign75 is insignificant using Namsor classification (Column 2) as is the replication coefficient (Column 6). However, the Foreign75 coefficient is negative and marginally significant at the 5% level for my reproduction sample. However, in unreported analysis the Foreign75 coefficient is not significant in robust specifications (i.e., the log transformation, quantile, and MM regressions in Table 5). Panel B shows that the significance of Foreign75 depends on the classification method, with the main result becoming insignificant when using Namsor and in more robust regression specifications.

I sort managers in descending order by fund flows for those with different name classifications between KNS MTurk and Namsor. Panel C reports the top five. Anthony G. Orphanos, manager of the Warburg Pincus Growth & Income Fund, appears three times, with a maximum flow of 17.147 (in 1994).[9] MTurk (KNS MTurk) classifies his name as American-sounding, while Namsor identifies it as likely originating in Cyprus or Greece (Western Asia) and foreign-sounding.

Panel D reports baseline regression results when substituting MTurk name classifications with Namsor name classifications for the five managers listed in Panel C. As before, Column 1 reports the baseline result from KNS using MTurk name classifications. Column 2 reports the results when substituting Anthony G. Orphanos' KNS MTurk classification with his Namsor classification for his maximum flow observation. Column 3 reports results for substitutions for the 3 observations of Mr. Orphanos. Column 4 and 5 repeat Columns 2 and 3 for the five managers listed in Panel C.

The estimated coefficient on Foreign75 is insignificant in Columns 2-5. Most notably, changing a single name classification for one observation reduces the t-statistic on Foreign75 from –4.06 (Column 1) to –1.66 (Column 2), indicating that the result is highly sensitive to this classification. OLS minimizes a squared-error loss function, meaning it disproportionately penalizes large residuals.

---

[9] Orphanos took over the fund in January 1992. In 1993, the fund beat the Standard & Poor's 500 by 27 percentage points and one of the few diversified domestic stock funds that managed to produce positive returns in 1994. https://www.deseret.com/1994/6/5/19112456/success-runs-in-the-fund-family-for-2-superstar-stock-pickers/

As a result, extreme values in the dependent variable (here, fund flows) can exert outsized influence on coefficient estimates. Because OLS has a breakdown point of 0%, even a single influential data point can significantly alter regression results.

Panel D illustrates how observations with unusually high dependent variable values can disproportionately influence regression results, particularly when the variable of interest is sparsely represented in the upper tail. Given this, the Foreign75 result in KNS likely suffered from sample imbalance in the upper tail, making its originally strong t-statistic highly fragile and dependent on a few extreme observations. This sensitivity is further compounded by the subjective nature of Foreign75's classification. While this does not imply that the original authors' classifications are necessarily incorrect, it underscores that the subjective nature of name classification introduces unavoidable ambiguity. This measurement variability amplifies the instability of the Foreign75 effect, further increasing its susceptibility to small data changes. In sum, the KNS study suffers from sensitivities to both name classification and extreme values, such that reclassifying a single observation's name from American-sounding to foreign-sounding is enough to eliminate the original result. This dependence on one observation in a large sample indicates that the KNS reported effect of Foreign75 is not robust.

5 Conclusion

The finding by Kumar, Niessen-Ruenzi, and Spalt (2015)—that fund managers with foreign-sounding names receive less investor flow—provides evidence of taste-based discrimination. While I can replicate this result using their sample, the finding fails to hold under alternative data constructions, classification methods, and outlier-robust estimators. Notably, the patterns I observe are more consistent with a rational expectations framework, where investors allocate capital based on observable performance and fund characteristics, rather than name-based bias. This shift in interpretation underscores a broader point: empirical conclusions—and the hypotheses they appear to support—are often highly sensitive to methodological choices. By demonstrating how small but routine decisions can reverse a key empirical finding, this paper aims to provide practical guidance on how to evaluate and implement research design choices that promote replicability and inference consistency.

A fundamental challenge in empirical research is determining how to handle extreme observations in a way that preserves meaningful economic information without introducing bias. This reanalysis demonstrates that researchers should adopt an economically informed selection framework when making data adjustments rather than arbitrary statistical thresholds (e.g., trimming or winsorizing at the 1% level).

In the context of mutual funds, it is unlikely that a fund would attract inflows several times larger than its total assets under management. Extreme fund flows often arise from confounding events—such as large one-time capital injections resulting from family-level mergers, fund liquidations, or reallocations—or from measurement inconsistencies, rather than genuine investor purchase decisions. Another factor is fund incubation, where asset managers privately operate funds before public launch and selectively report only the successful ones. This practice distorts fund flow measurements by overrepresenting high-performing funds while omitting less successful counterparts, artificially inflating apparent investor demand. Additionally, fund families often seed new funds with internal capital, but because outsiders cannot distinguish between internal and external flows, interpretation becomes more complex. A reasonable, economically informed selection criterion might include limiting flows to below 400% of prior-year assets under management, requiring at least three years of fund history, and setting a minimum of $25 million in total net assets. These criteria help focus the analysis on economically plausible observations.

Researchers should also report comprehensive descriptive statistics to facilitate reliability and validity. Any analysis plan should account for key data distribution characteristics, such as skewness, outliers, and group representation revealed in the descriptive statistics. Although OLS regression is widely used in finance, its sensitivity to outliers makes it problematic for analyzing mutual fund flows, which exhibit strong positive skewness due to infrequent but extreme inflows. To address this, researchers should consider robust statistical methods, such as log transformations to mitigate skewness, quantile regression to capture distributional effects, or robust estimators like MM-estimation to reduce outlier sensitivity.

Finally, classification results should be validated across multiple methodologies, including both survey- and AI-based approaches. Variation across methods may reflect measurement error or differing underlying constructs—surveys aim to capture subjective perception, while AI tools infer name origin. However, if investor perceptions broadly align with linguistic or cultural cues that AI captures, algorithmic classifications may offer a more consistent and scalable proxy for perceived identity than limited or inconsistent survey responses.

In closing, I thank Kumar, Niessen-Ruenzi, and Spalt for sharing their data and code. I acknowledge that, despite my efforts, this reanalysis may contain errors. To support transparency and facilitate verification, all replication code and data used in this paper are available on the Open Science Framework (OSF).

References

Adams, J., D. Hayunga, and S. Mansi. 2018. Diseconomies of scale in the actively-managed mutual fund industry: What do the outliers in the data tell us? *Critical Finance Review, 7*(2), 273-329.

Ahler, D. J., C.E. Roush, and G. Sood. 2021. The micro-task market for lemons: Data quality on Amazon's Mechanical Turk. *Political Science Research and Methods*, 1-20.

Barber, B. M., X. Huang, and T. Odean. 2016. Which factors matter to investors? Evidence from mutual fund flows. *The Review of Financial Studies*, 29(10), 2600-2642.

Barber, B. M., T. Odean, and L. Zheng. 2005. Out of sight, out of mind: The effects of expenses on mutual fund flows. *The Journal of Business*, 78(6), 2095-2120.

Ben-David, I., J. Li, A. Rossi, and Y. Song. 2022. What do mutual fund investors really care about?. *The Review of Financial Studies*, 35(4), 1723-1774.

Berk, J. B. and R. C. Green. 2004. Mutual fund flows and performance in rational markets. *Journal of Political Economy*, *112*(6), 1269-1295.

Berk, J. B. and J. H. Van Binsbergen. 2015. Measuring skill in the mutual fund industry. *Journal of Financial Economics, 118*(1), 1-20.

Brown, D. P. and Y. Wu. 2016. Mutual fund flows and cross-fund learning within families. *The Journal of Finance,* 71(1), 383-424.

Bursztyn, L., T. Chaney, T. A. Hassan, and A. Rao. 2024. The immigrant next door. *American Economic Review*, *114*(2), 348-384.

Chevalier, J. and G. Ellison. 1997. Risk taking by mutual funds as a response to incentives. *Journal of Political Economy*, *105*(6), 1167-1200.

Choi, D., B. Kahraman, and A. Mukherjee. 2016. Learning about mutual fund managers. *The Journal of Finance*, *71*(6), 2809-2860.

Cooper, M. J., H. Gulen, and P. R. Rau. 2005. Changing names with style: Mutual fund name changes and their effects on fund flows. *The Journal of Finance,* 60(6), 2825-2858.

El Ghoul, S. and A. Karoui 2017. Does corporate social responsibility affect mutual fund performance and flows?. *Journal of Banking and Finance*, 77, 53-63.

Evans, R. B. 2010. Mutual fund incubation. *The Journal of Finance, 65*(4), 1581-1611.

Franzoni, F. and M. C. Schmalz. 2017. Fund flows and market states. *The Review of Financial Studies*, 30(8), 2621-2673.

Gelman, A., S. DellaVigna, D. G. Pope, and E. Vivalt, et al. 2023. Many Analysts, One Dataset: Replicability and Robustness of Financial Research. *The Journal of Finance,* 78(2), 553–594.

Jain, P. C. and J. S. Wu. 2000. Truth in mutual fund advertising: Evidence on future performance and fund flows. *The Journal of Finance*, 55(2), 937-958.

Jank, S. 2012. Mutual fund flows, expected returns, and the real economy. *Journal of Banking and Finance*, 36(11), 3060-3070.

Kacperczyk, M., C. Sialm, and L. Zheng. 2008. Unobserved actions of mutual funds. *The Review of Financial Studies, 21*(6), 2379-2416.

Kennedy, R., S. Clifford, T. Burleigh, P. D. Waggoner, R. Jewell, and N. J. Winter. 2020. The shape of and solutions to the MTurk quality crisis. *Political Science Research and Methods*, *8*(4), 614-629.

Kostovetsky, L. 2016. Whom do you trust?: Investor-advisor relationships and mutual fund flows. *The Review of Financial Studies, 29*(4), 898-936.

Kumar, A., A. Niessen-Ruenzi, and O. G. Spalt. 2015. What's in a name? Mutual fund flows when managers have foreign sounding names. *The Review of Financial Studies*, *28*(8), 2281-2321.

LeRoy, M. H. 2018. Targeting white supremacy in the workplace. *Stanford Law and Policy Review*, *29*, 107.

Lou, D. 2012. A flow-based explanation for return predictability. The Review of Financial Studies, 25(12), 3457-3489.

Massa, M., J. Reuter, and E. Zitzewitz. 2010. When should firms share credit with employees? Evidence from anonymously managed mutual funds. *Journal of Financial Economics*, *95*(3), 400-424.

Mateos, P., P. A. Longley, and D. O'Sullivan. 2011. Ethnicity and population structure in personal naming networks. *PLoS ONE, 6*(9), e22943.

McLaren, P. and R. Torres. 2005. Racism and multicultural education: Rethinking 'race' and 'whiteness' in late capitalism. In S. May (Ed.), *Critical Multiculturalism* (pp. 54-91). Routledge.

Miles, R. (2004). *Racism*. Routledge.

Niessen-Ruenzi, A. and S. Ruenzi. 2019. Sex matters: Gender bias in the mutual fund industry. *Management Science*, *65*(7), 3001-3025.

Patel, S. and S. Sarkissian. 2017. To group or not to group? Evidence from mutual fund databases. *Journal of Financial and Quantitative Analysis*, 52(5), 1989-2021.

Pástor, L., R. F. Stambaugh, and L. A. Taylor. 2015 Scale and skill in active management. *Journal of Financial Economics, 116*(1), 23-45.

Sebo, P. 2022. NamSor's performance in predicting the country of origin and ethnicity of 90,000 researchers based on their first and last names. *Research Square*. https://doi.org/10.21203/rs.3.rs-1565759/v2

Sialm, C. and T. M. Tham. 2016. Spillover effects in mutual fund companies. *Management Science,* 62(5), 1472-1486.

Sirri, E. R. and P. Tufano. 1998. Costly search and mutual fund flows. *The Journal of Finance*, *53*(5), 1589-1622.

Spiegel, M. and H. Zhang. 2013. Mutual fund risk and market share-adjusted fund flows. *Journal of Financial Economics*, 108(2), 506-528.

Warther, V. A. 1995. Aggregate mutual fund flows and security returns. *Journal of Financial Economics*, 39(2-3), 209-235.

Welch, I. 2019. Reproducing, extending, updating, replicating, reexamining, and reconciling. *Critical Finance Review, 8*(1-2), 301-304.

Zhu, Q. 2020. The missing new funds. *Management Science, 66*(3), 1193-1204.

Table 1
Equity mutual fund flows literature

Research Paper (Listed by Year)	Pub. Status	Data Frequency	Annual Flow	Active/ Passive Mgmt.	Flows Outlier Treatment	Min. Size (MM)	Min. Age (Years)	Data Source	Years Covered	Number of Funds	Require Mgr Names
Warther (1995)	JFE	Monthly	0.6%	Active	NR	NR	NR	ICI	1984-92	107	N
Chevalier & Ellison (1997)	JPE	Annual	12.2%	Active	NR	$10	2	Morningstar/CRSP	1983-93	449	N
Sirri & Tufano (1998)	JF	Annual	NR	Active	NR	NR	NR	ICDI	1971-90	690	N
Jain & Wu (2000)	JF	Monthly	4.7%	A&P	NR	NR	NR	Morningstar	1994-97	2,324[a]	N
Barber, Odean, & Zheng (2005)	JB	Annual	6%[b]	A&P	NR	NR	NR	CRSP	1970-99	NR	N
Cooper, Gulen, & Rau (2005)	JF	Monthly	5.6%[c]	A&P	NR	NR	NR	Morningstar/CRSP	1994-01	296	N
Kacperczyk, Sialm, & Zheng (2008)	RFS	Quarterly & Monthly	30%	A&P	1% Winsor	$5	1	CRSP/CDA	1984-03	2,543	N
Jank (2012)	JBF	Quarterly	1.4%	A&P	NR	NR	NR	ICI	1984-09	NR	N
Lou (2012)	RFS	Quarterly	NR	Active	NR	$1	NR	CRSP/CDA	1980-96	2,989	N
Spiegel & Zhang (2013)	JFE	Monthly	NR[d]	A&P	2.5% Winsor	NR	NR	CRSP	1991-06	NR	N
Barber, Huang, & Odean (2016)	RFS	Monthly	–6%	Active	Truncate[e]	$10	5	CRSP	1996-11	3,900[e]	N
Brown & Wu (2016)	JF	Monthly	3.3%	Active	1% Trim	$5	3	CRSP	1999-11	2,053	N
Kostovetsky (2016)	RFS	Monthly	4.8%	Active	1% Winsor	$10	3[c]	Morningstar/CRSP	1995-12	1,600[f]	N
Sialm & Tham (2016)	MS	Monthly	1.3%	Active	1% Winsor	NR	2	Morningstar/CRSP	1992-09	1,203[g]	Y
El Ghoul & Karoui (2017)	JBF	Annual	14.7%	Active	0.5% Winsor	$5.9	1	CRSP	2003-11	2,168	N
Franzoni & Schmalz (2017)	RFS	Quarterly	–2.8%	Active	1% Trim	$5	3	CRSP	1980-13	3,390[h]	N
Patel & Sarkissian (2017)	JFQA	Annual	0.5%	Active	1% Winsor	NR	1	Morningstar	1992-10	3,935[i]	N
Ben-David, Rossi, & Song (2022)	RFS	Monthly	1.8%	Passive	5% Winsor	$10	6.5	CRSP	1997-11	328	N
Kumar, Niessen-Ruenzi, & Spalt (2015)	**RFS**	**Annual**	**20.5%**	**A&P**	**1% Trim**	**<$1**	**NR**	**Morningstar/CRSP**	**1993-11**	**2,553[j]**	**Y**

Description: This table summarizes literature on domestic equity mutual fund flows. Publications are Journal of Banking and Finance (JBF), Journal of Finance (JF), Journal of Financial Economics (JFE), Journal of Financial and Quantitative Analysis (JFQA). Journal of Political Economy (JPE), Management Science (MS), and Review of Financial Studies (RFS). Information obtained from each paper's data section or table of descriptive statistics. NR denotes not reported. A&P denotes sample includes actively and passively managed funds, the default assumption unless the authors explicitly include or exclude passively managed funds. Flows are annualized for ease of comparison (monthly flows multiplied by 12, quarterly flows multiplied by 4). Flows outlier treatment is the % of treated observations from each side of the distribution. Minimum size (in millions of dollars) and fund age (in years) are obtained from the data section or from minimum and maximum values reported in the table of descriptive statistics. Due to variability in sample construction and reporting in the literature, notes are provided to aid comparisons.

[a] Includes control group. Mean annual flow for 258 (2,066) treatment (control) funds is 5.3% (4.7%).
[b] Approximate. Flows reported in expense ratio ranges and by sales load. See Table 1 of the published paper.
[c] Flows for treatment group. Flows for control group are 13.8%. Number of control funds not reported.

[d] Overall sample mean flows not reported. Flows only reported for every other year.
[e] Flows truncated below -90% and above 1,000%. Number of funds not explicitly reported but somewhere between 3,900 and 4,000.
[f] Requires contemporaneous reporting in Morningstar and a non-blank fund name in CRSP at the beginning of each calendar year to mitigate incubation bias and drops funds younger than 3 years. Approximate as the number of unique funds is not reported, estimated by dividing the number of reported fund-month observations by the number of months in the study.
[g] Funds with manager names. From the published paper, *"As not all of our sample funds from CRSP have non-missing CUSIPs and some funds with available CUSIPs are not covered in the list of funds with manager names from Morningstar, we have a reduced sample of funds for the analyses pertaining to management turnover. Thus, we only have available fund manager data for 1,203 equity funds and for 633 bond funds."* The full sample without manager names includes 6,102 funds.
[h] Drops observations before ticker creation date to mitigate incubation bias and funds younger than 3 years old. The published paper unit of observation is the share class. For ease of comparison with the other papers in this table, the number of share classes is divided by 2.3 (the average number of share classes per my review of Morningstar data around their sample period).
[i] Includes team-managed funds even when managers are not identified (e.g., fund listed as "Team Managed").
[j] KNS reports 4,805 funds which reduces to 2,553 funds after dropping observations with missing information. Mean flows values are for 'Other' funds (American sounding names) which account for about 95% of the sample.

Interpretation: There is substantial variation in flows estimates across studies. KNS estimated flows are much larger than those of contemporary studies.

Table 2
Detailed flows statistics

Sample	# Funds	# Obs.	Mean	Std Dev.	Skewness	Minimum	1st Percentile	Median	99th Percentile	Maximum
KNS Final										
All Funds	2,553	13,091	0.2	0.96	8.221	–0.982	–0.598	–0.01	4.014	17.457
Foreign Funds [4.5%]	184	592	0.096	0.444	3.08	–0.964	–0.703	–0.006	1.977	3.785
Other Funds	2,487	12,499	0.205	0.977	8.132	–0.982	–0.595	–0.01	4.216	17.457
KNS Reproduction										
All Funds	2,258	11,438	0.11	0.488	3.202	–0.653	–0.501	–0.019	2.4	3.87
Foreign Funds [4.6%]	157	531	0.087	0.412	3.525	–0.608	–0.455	–0.008	1.828	3.785
Other Funds	2,209	10,907	0.111	0.491	3.186	–0.653	–0.502	–0.019	2.403	3.87
Reproduction										
All Funds	1,403	7,063	0.13	0.506	3.342	–0.617	–0.462	–0.006	2.484	4.015
Foreign Funds [5.5%]	104	391	0.192	0.56	2.635	–0.609	–0.51	0.015	2.759	3.158
Other Funds	1,346	6,672	0.126	0.503	3.395	–0.617	–0.461	–0.007	2.481	4.015
Replication										
All Funds	1,562	8,546	0.006	0.295	2.271	–0.594	–0.483	–0.055	1.237	1.909
Foreign Funds [4.6%]	101	394	0.02	0.314	2.147	–0.588	–0.509	–0.055	1.368	1.836
Other Funds	1,518	8,152	0.005	0.294	2.277	–0.594	–0.483	–0.055	1.237	1.909

Description: This table reports descriptive statistics for the reported annual fund flows estimates in the KNS, KNS reproduction, reproduction, and replication samples. The observations are at the fund-year level. The data cover the period 1991–2011. All flows estimates are trimmed at the 1% level. The percentage of observations with foreign sounding manager names are reported in brackets. Name classifications are obtained using the web-based Amazon Mturk survey platform.

Interpretation: Flows of the KNS sub-sample of funds whose managers have 'American' sounding names (Other Funds) are highly skewed.

Table 3
KNS Table 1 Reproduction: Fund manager names and select fund characteristics

Variable	Foreign funds (1)	Other funds (2)	Difference (3)	t-statistic (4)
Fund flow				
KNS	0.096	0.205	–0.109	–2.70
KNS Reproduction	0.087	0.111	–0.024	–1.10
Reproduction	0.192	0.126	0.066	2.51
Replication	0.020	0.005	0.015	0.97
PRank				
KNS	0.467	0.475	–0.008	–0.73
KNS Reproduction	0.509	0.515	–0.006	–0.47
Reproduction	0.561	0.531	0.029	1.97
Replication	0.534	0.501	0.033	2.16
Fund size				
KNS	5.204	5.279	–0.074	–0.89
KNS Reproduction	5.215	5.361	–0.146	–1.67
Reproduction	5.746	5.696	–0.050	0.51
Replication	6.393	6.441	–0.048	0.54
Fund age				
KNS	2.269	2.190	0.079	2.18
KNS Reproduction	2.317	2.247	0.070	1.87
Reproduction	2.141	2.299	–0.159	–3.78
Replication	2.631	2.660	–0.029	–0.78
Morningstar rating				
KNS	1.800	1.843	-0.043	–0.26
Reproduction	3.398	3.180	0.218	3.93
Replication	3.500	3.419	0.081	1.59

Description: This table reports mean values by foreign name classification using the KNS, KNS reproduction, reproduction, and replication samples. Foreign funds are U.S. domestic equity funds whose managers have foreign sounding names and Other funds have managers with American sounding names.

Interpretation: The published flow value can only be reproduced using the KNS final sample.

Table 4

KNS Table 2 baseline regressions

	KNS Sample (1)	KNS Reproduction Sample (2)	Reproduction Sample (3)	Replication Sample (4)
Foreign75	**–0.089*****	**–0.010**	**0.037**	**–0.005**
	(–4.06)	**(–0.49)**	**(1.20)**	**(–0.19)**
PRank	0.052	0.160***	–0.019	0.213***
	(0.41)	(2.64)	(–0.24)	(4.07)
$PRank^2$	0.583***	0.200***	0.407***	0.042
	(3.98)	(3.19)	(5.08)	(0.80)
Fund size	–0.060***	–0.023***	–0.023***	–0.000
	(–10.07)	(–8.58)	(–6.39)	(–0.14)
Turnover	0.037**	0.959**	0.637	–0.006
	(2.15)	(2.02)	(1.24)	(–0.81)
Fund risk	0.017	–0.536**	–0.383	–0.023
	(0.04)	(–2.20)	(–1.15)	(–0.07)
Expense ratio	–2.106	–0.059	0.079	0.976
	(–1.90)	(–0.16)	(0.06)	(0.73)
Fund age	–0.094***	–0.049***	–0.061***	–0.059***
	(–9.16)	(–8.66)	(–7.48)	(–7.63)
Segment flow	0.092	0.153***	0.001	0.002
	(1.70)	(6.07)	(0.72)	(0.85)
Family flow	0.014**	0.000***	–0.000	–0.000
	(2.22)	(4.29)	(–1.44)	(–0.07)
Mgr tenure	–0.001	0.190***	–0.003**	–0.000
	(–0.96)	(14.59)	(–2.57)	(–0.33)
Lag fund flow	0.098***	–0.182***	0.145***	0.000
	(9.05)	(–3.14)	(9.99)	(1.08)
Constant	0.405***	–0.010	–0.200***	0.031
	(6.970)	(–0.49)	(–3.75)	(0.89)
Obs.	13,091	11,438	7,063	5,302
Adj. R-squared	0.102	0.193	0.178	0.118

Description: This table provides coefficients from OLS regressions of domestic equity mutual fund flows on manager and fund attributes. Column 1 (KNS) reproduces the baseline (their Table 2, Model 3) results from Kumar, Niessen-Ruenzi, and Spalt (2015) using their data. Column 2 presents results using the KNS reproduction sample. Columns 3 and 4 provide results for the reproduction and replication samples. All columns trim the most extreme 1% of fund flow observations from both sides of the distribution. All specifications include year and style (segment) fixed effects and fund-level clustered standard errors. T-statistics are in parenthesis. *** $p<.01$, ** $p<.05$. Family and segment flows winsorized at the 1% level.

Interpretation: The KNS main finding (Foreign75) can be exactly reproduced using their sample. However, their results are not robust for the KNS reproduction, reproduction, and replication samples.

Table 5
Outlier robust regressions using KNS final sample data

	(1) Published Results (Skewness=8.221)	(2) Log Fund Flow (Skewness=0.978)	(3) Quantile Regression	(4) MM Regression (Efficiency=28.7%)	(5) MM Regression (Efficiency=95%)
Foreign75	–0.089***	–0.019	0.017	0.019	0.026
	(–4.06)	(–1.15)	(1.56)	(2.14)	(2.36)
Obs.	13,091	13,091	13,091	13,091	13,091
Adj. R-squared	0.102	0.155	0.090	0.100	0.108

Description: This table reports baseline regression results using the KNS provided sample data for different outlier mitigation methodologies (all controls from KNS Table 2 included but not reported). Column 1 shows the published OLS results, Column 2 reports OLS results regressing log(1+fund flow) on Foreign75 and controls, Column 3 reports the coefficient estimate using quantile regression. Columns 4 and 5 report results for outlier robust regressions with high resistance to outliers (efficiency of 28.7%) and efficient with some resistance to outliers (95% efficiency relative to OLS).

Interpretation: The KNS result that foreign sounding manager funds receive less flow is sensitive to outlier mitigation.

Table 6
Reconciliation analysis

Panel A: Detailed flows statistics of the KNS provided sample data

Sample	# Funds	# Obs.	Mean	Std Dev.	Skewness	Min.	1st Percentile	Median	99th Percentile	Max.
KNS full sample	4,362	24,053	11.868	686.926	82.254	–10.675	–0.681	0.008	17.147	74089.88
KNS final sample (KNS provided flow estimates)										
All funds	2,553	13,091	0.2	0.96	8.221	–0.982	–0.598	–0.01	4.014	17.457
Foreign funds	184	592	0.096	0.444	3.08	–0.964	–0.703	–0.006	1.977	3.785
Other funds	2,487	12,499	0.205	0.977	8.132	–0.982	–.595	–0.01	4.216	17.457
1% trimming at the full sample level										
All Funds	2,764	14,046	0.182	0.852	6.925	–1.016	–0.597	–0.013	3.799	13.561
Foreign Funds	202	638	0.096	0.444	3.043	–0.964	–0.608	–0.007	1.862	3.785
Other Funds	2,688	13,408	0.186	0.867	6.866	–1.016	–0.595	–0.013	3.859	13.561
1% trimming at the useable sample level										
All funds	2,730	13,752	0.138	0.554	3.498	–0.635	–0.508	–0.012	2.67	4.65
Foreign funds	200	629	0.104	0.43	3.312	–0.608	–0.444	–0.004	1.862	3.785
Other funds	2,654	13,123	0.14	0.56	3.488	–0.635	–0.51	–0.013	2.706	4.65
1% useable sample trimming matched to KNS final sample										
All funds	2,515	12,747	0.143	0.558	3.494	–0.632	–0.501	–0.01	2.706	4.65
Foreign funds	183	583	0.108	0.437	3.309	–0.608	–0.455	–0.003	1.977	3.785
Other funds	2,449	12,164	0.145	0.563	3.486	–0.632	–0.502	–0.01	2.729	4.65
1% useable sample trimming matched to KNS final sample (KNS provided flow estimates)										
All funds	2,515	12,747	0.144	0.559	3.490	–0.972	–0.502	–0.01	2.711	4.65
Foreign funds	183	583	0.108	0.437	3.311	–0.608	–0.454	–0.003	1.977	3.785
Other funds	2,449	12,164	0.146	0.564	3.481	–0.972	–0.503	–0.01	2.736	4.65

Panel B: KNS Table 2 baseline regressions and trimming

	(1) KNS final sample (published main result)	(2) 1% full sample trimming	(3) 1% useable sample trimming	(4) 1% useable sample trimming matched to KNS final sample
Foreign75	–0.089***	–0.073***	–0.015	–0.017
	(–4.06)	(–3.78)	(–0.85)	(–0.92)
Obs.	13,091	14,046	13,752	12,747
Adj. R-squared	0.102	0.115	0.179	0.179

Panel C: Comparing matched and unmatched observations

Variable	(1) Obs. in matched sample n=12,747	(2) Obs. in KNS final sample only n=344	(3) Difference	(4) t-statistic
Flow (KNS provided flow estimates)	0.144	2.286	–2.142	–43.733
Foreign75	0.046	0.026	0.020	1.724
PRank	0.474	0.493	–0.019	–1.307
Fund size	5.315	3.795	1.520	14.142
Fund age	2.205	1.761	0.444	9.472
Morningstar rating	3.150	3.191	–0.041	–0.579
Missing Morningstar rating	0.271	0.395	–0.124	–5.093
Missing CRSP portfolio identifiers	0.445	0.526	–0.081	–2.99

Description: This table features reconciliation analysis using sample data and flows estimation code provided by KNS. Panel A reports detailed flows statistics for different outlier trimming treatments. The KNS full sample includes useable and unusable observations and does not drop extreme flows observations. The KNS final sample drops the smallest and largest 1% of flows (2% total) and removes unusable observations. The 1% full trimming sample drops the smallest and largest 1% of flows from the KNS full sample and then drops unusable observations. The 1% useable trimming sample drops the 1% smallest and largest flow observations of the useable observations. Flows are provided by KNS for the KNS final sample. The flows for the KNS full sample, 1% useable sample trimming, and 1% useable sample matched to KNS final sample are independently computed using code provided by KNS and data from the CRSP mutual fund database. Panel A also provides detailed flows statistics for the 1% useable sample matched to KNS final sample using KNS provided flows. Panel B reports the reproduced baseline regression results using the KNS final sample in Column 1. Column 2 reports results for full sample trimming. Column 3 reports results for flows trimmed at the useable sample level. Column 4 reports results for flows trimmed at the useable sample level and matched to the KNS final sample. All controls from KNS Table 2 included but not reported. All specifications include year and style (segment) fixed effects and fund-level clustered standard errors. T-statistics are in parenthesis. *** $p<.01$, ** $p<.05$. Family and segment flows winsorized at the 1% level. Panel C reports mean values and differences for the 1% useable sample matched and unmatched to the KNS final sample using KNS provided flow estimates. Foreign funds are U.S. domestic equity funds whose managers have foreign sounding names and Other funds have managers with American sounding names.

Interpretation: KNS appear to have 1% trimmed outlier flows at the full sample level which includes unusable observations. Their published result is not robust to 1% trimming at the useable sample level. Observations that are only in the KNS final sample have much higher estimated flows and are smaller, younger, and more frequently do not have Morningstar fund ratings or CRSP assigned portfolio identifiers.

Table 7
Identifying foreign sounding names using Namsor Artificial Intelligence Taxonomy

Panel A: Classification differences

		Proportion of Different Foreign75 Classifications			
	Mean/Number	KNS MTurk Foreign75	My MTurk Foreign75	Namsor Foreign75 (December 2022)	Namsor Foreign75 (February 2025)
	(1)	(2)	(3)	(4)	(5)
KNS MTurk Foreign75	0.053/60	0	0.633	0.494	0.717
My MTurk Foreign75	0.041/47	0.532	0	0.787	0.851
Namsor Foreign75 (December 2022)	0.045/51	0.529	0.804	0	0.314
Namsor Foreign75 (February 2025)	0.038/43	0.605	0.837	0.186	0
# Managers	1,138				

Panel B: Baseline regressions using Namsor

	KNS MTurk	KNS Namsor	Reproduction MTurk	Reproduction Namsor	Replication MTurk	Replication Namsor
	(1)	(2)	(3)	(4)	(5)	(6)
Foreign75	–0.089***	–0.009	0.037	–0.052**	–0.005	–0.025
	(–4.06)	(–0.19)	(1.20)	(–2.03)	(–0.19)	(–1.28)
PRank	0.052	0.055	–0.019	–.016	0.213***	0.213***
	(0.41)	(0.43)	(–0.24)	(–0.21)	(4.07)	(4.07)
$PRank^2$	0.583***	0.580***	0.407***	0.406***	0.042	0.042
	(3.98)	(3.96)	(5.08)	(5.05)	(0.80)	(0.81)
Fund Controls	Y	Y	Y	Y	Y	Y
Obs.	13,091	13,091	7,063	7,063	5,302	5,302
Adj. R-squared	0.102	0.101	0.178	0.178	0.118	0.118

Panel C: Maximum KNS flows and foreign name classification differences

Maximum Fund Flow	Manager Name	Obs.	Foreign75 KNS MTurk	Foreign75 KNS Namsor	Namsor Name Origin Primary Country	Namsor Name Origin Secondary Country	Namsor Name Origin Region
17.147	Anthony G. Orphanos	3	0	1	Cyprus	Greece	Western Asia
6.033	Edwin Walczak	5	0	1	Poland	Germany	Eastern Europe
5.333	Paul W. Wojcik	7	0	1	Poland	Germany	Eastern Europe
3.785	Boniface Zaino	6	1	0	Kenya	Congo	Eastern Africa
3.079	Chad Fleischman	4	0	1	Israel	Lebanon	Western Asia

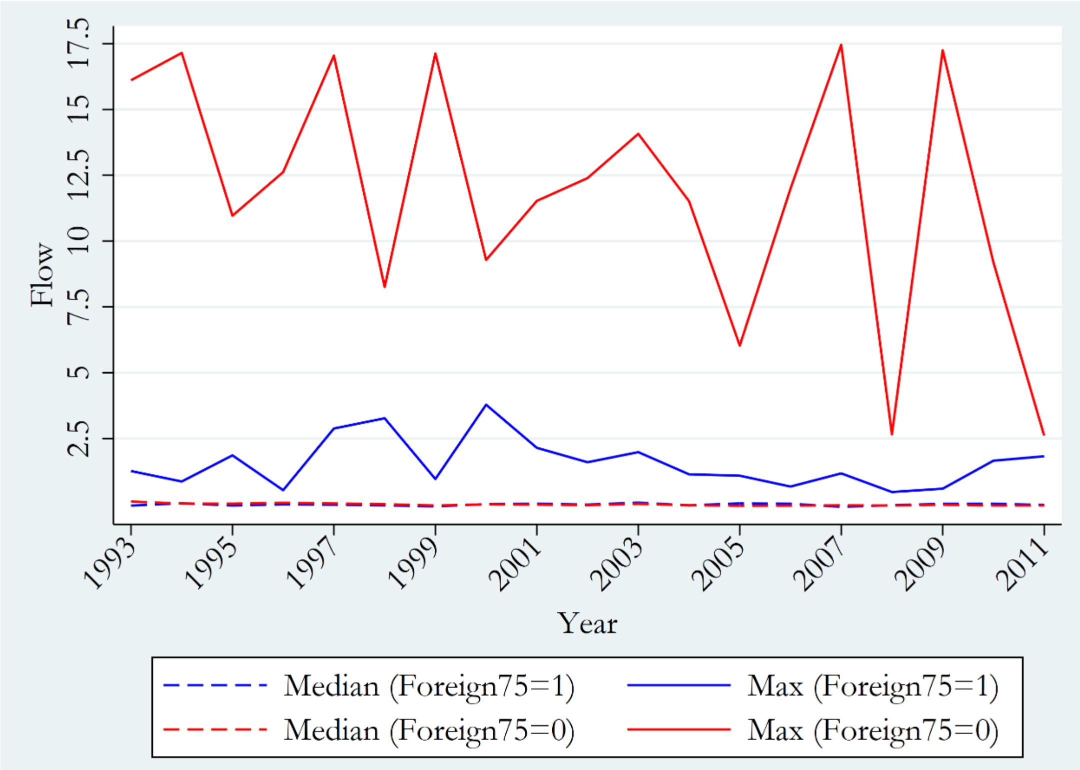
Panel D: KNS Baseline regressions substituting Namsor classifications for maximum flow managers

		Anthony G. Orphanos		All 5 managers in Panel C	
	KNS MTurk	Maximum Flow Observation # subs=1	All Observations # subs=3	Maximum Flow Observation # subs=5	All Observations # subs=25
	(1)	(2)	(3)	(4)	(5)
Foreign75	–0.089***	–0.06	–.063	–0.035	–0.04
	(–4.06)	(–1.66)	(–1.90)	(–0.91)	(–1.14)
PRank	0.052	0.053	0.053	0.054	0.053
	(0.41)	(0.42)	(0.41)	(0.42)	(0.42)
$PRank^2$	0.583***	0.582***	0.582***	0.581***	0.582***
	(3.98)	(3.98)	(3.98)	(3.97)	(3.98)
Fund Controls	Y	Y	Y	Y	Y
Obs.	13,091	13,091	13,091	13,091	13,091
Adj. R-squared	0.102	0.101	0.101	0.101	0.101

Description: Panel A presents classifications of foreign-sounding names—using both Mturk and Namsor AI—for managers appearing in the KNS, reproduction, and replication samples. Column 1 displays the mean and the number of managers classified as “Foreign75” (i.e., those with foreign-sounding names). Columns 2–5 show the differences in the proportion of “Foreign75” classifications across each methodology. Panel B regresses fund flows on manager names in the KNS, reproduction, and replication samples identified using the MTurk and Namsor Artificial Intelligence Taxonomy. Foreign75 in this table indicates a 75% probability using MTurk surveys where at least 75% of respondents indicated the manager’s name sounded ‘foreign’ in columns 1, 3, and 5. Columns 2, 4, and 6 report results using Namsor where Foreign75 indicates a 75% or higher probability that the manager’s name is not Northern, Western, or Southern European in origin. Panel C provides KNS MTurk and Namsor name classifications for the top 5 managers with the highest flows who also have different KNS MTurk and Namsor name classifications. Panel D reports baseline regression coefficients when changing the KNS MTurk Foreign75 to the Namsor Foreign75 value for the top 5 managers listed in Panel C. All regression specifications include the fund controls in Table 4 (baseline regressions) and year and style (segment) fixed effects and fund level clustered standard errors. T-statistics are in parenthesis. *** *p<.01, *** ** *p<.05*.

Interpretation: Foreign name classifications are inconsistent. The KNS main finding cannot be replicated using alternative methodology to identify foreign sounding names.

Figure 1
KNS Annual fund flows



Description: This figure depicts median (dashed lines) and maximum (solid lines) annual flows by foreign manager identification for the KNS final sample. Foreign funds (in blue) are funds managed by an individual with a foreign sounding name according to KNS's Foreign75 classification. Other funds (in red) are the funds with American sounding manager names.

Interpretation: The maximum flows values are consistently much higher for funds with American sounding manager names. There is not a noticeable difference in median flows between funds with foreign and American sounding names.

Appendix

1 Sample Constructions

This section details the reproduction and replication sample construction processes.

1.1. Identifying managers with foreign sounding names

Following KNS, I conduct an online survey through Amazon Mechanical Turk (MTurk) and ask participants to respond to the question, "Does this name sound foreign in the U.S.?" by choosing "Yes, No, or Unsure". Survey respondents are asked to record their perception of each name's geographical origin. For the replication sample, a total of 102 participants are recruited to each classify 765 of 1,485 randomly assigned names yielding about 50 perceptions of each manager's name. For the reproduction sample, I ask 55 participants to each classify 585 names that were not included in the replication sample, yielding about 50 perceptions per name. This compares to KNS who ask 150 participants to each classify 1,000 of 3,784 names for 30 perceptions per manager name. I also pay $15 as opposed to KNS's $7 payment to account for wage inflation, attract more conscientious respondents, respondents that live in the U.S., have at least a high school diploma, and are currently employed. Foreign75, my and KNS's main variable of interest, is a dummy variable that takes on a value of 1 if at least 75% of the survey participants classify a fund manager's name as foreign sounding. The mean Foreign75 value is 4.6% for the replication sample and 5.5% for the reproduction sample—close to the value of 4.5% reported in KNS (page 2287).

1.2. Variable measurements

Fund Attributes. The main fund variables of interest are the flows and performance estimates. Flow and all other measures are estimated following KNS. *Fund Flow* is the ratio of $TNA_{i,t}-TNA_{i,t-1}$ scaled by $TNA_{i,t-1}$ less $r_{i,t}$ where $TNA_{i,t}$ is total net assets and $r_{i,t}$ is fund i's annual return in year t. *Segment Flow* is the growth rate of fund i's market style due to flows in year t, excluding flows in fund i. *Family Flow* is the growth rate of fund i's fund family due to flows in year t, excluding flows in fund i. Fund performance is estimated using relative performance rankings. *Prank,* or performance rank, is the ranking of the fund in the previous year relative to all other funds in the same Morningstar investment category segment, scaled between 0 (lowest performance) and 1 (highest performance). *Fund size* as the natural logarithm of the fund's total net assets (in millions). *Turnover* is fund trading activity scaled by change in portfolio holdings, or the lesser of sales or purchases divided by average monthly total net assets in percentage. *Fund risk* is the standard deviation of fund monthly returns by calendar year and *Expense ratio* is the percentage of fund assets charged annually to pay for operating expenses including 12b-1 fees, management, administrative, distribution, and custodial services fees. *Fund age* is computed as the natural logarithm of the number of years since the fund's inception, and *Retail ratio* as the net assets of retail share classes scaled by total net assets of all fund share classes. *Morningstar rating* is the fund star rating designated by Morningstar and *No Load Fund* is a dummy variable that equals 1 if the fund does not charge sales fees.

Manager Attributes. Manager attributes include experience, gender, education, and certifications. *Tenure*, a proxy for fund management experience, is computed as the difference between the current year and a manager's starting year at a fund as reported in Morningstar Direct.[10] *Female*, a proxy for gender diversity, is a dummy variable that equals 1 if the manager is female. *MBA,* a proxy for enhanced career opportunities, greater skills, credibility, and respect, is a dummy variable that equals 1 if the manager has a master's degree. *Ph.D.*, a proxy for specialized education, is a dummy variable that equals 1 if the manager has a doctorate.

1.3 Constructing the KNS reproduction sample

KNS obtain fund data for individual-managed U.S. equity funds over the period 1993–2011 from the Center for Research on Security Prices (CRSP) mutual fund database, and manager names from Morningstar and from CRSP when Morningstar does not provide manager names. They drop the smallest and largest 1% of fund flow observations and note their results are robust to using raw flows, winsorizing, or dropping funds with TNA values below $1 million.

Appendix Table 1 details the KNS reproduction sample construction process. CRSP reports fund data monthly at the share class level and the KNS regressions include lagged fund flows so observations must have non-missing returns and TNA values for the prior two years. Therefore, step 1 reports monthly observations at the share class level for the 1991–2011 period. Although not reported in the published paper, the authors manually assign portfolio codes when CRSP identifiers are not available to avoid dropping funds in the earlier years of their sample.[11] So, Step 2 keeps only

---

[10] KNS use Morningstar for manager names but CRSP for manager tenure, likely resulting in some name-tenure mismatches.

[11] In response to an earlier version of this paper, KNS argue, "You can't simply aggregate share classes to the portfolio level without a manually constructed portfolio code for earlier sample years, for which CRSP portfolio identifiers are only

funds with KNS assigned portfolio codes. Steps 3-5 aggregate share classes to the fund level, drop duplicate observations, and annualize the monthly observations. Step 6 merges with the KNS provided manager name, classification, and tenure files. Step 7 drops observations missing information necessary to reproduce the KNS baseline regression result. Finally, step 8 drops observations with the smallest and largest 1% of flows. The KNS reproduction sample is approximately the same size as the KNS final sample of observations used for their baseline results.

1.4 Reproduction sample construction.

Appendix Table 2 details the reproduction sample construction process. This reconstruction uses the same CRSP data as step 1 of the previous table. Step 1 reports monthly observations at the share class level for the 1991–2011 period for funds with ticker symbols necessary to merge with the Morningstar database to obtain fund manager names. KNS does not specify how they identify domestic equity funds, only that their sample comprises funds that "invest predominately in U.S. domestic equities." I identify "predominantly equity" as having at least 50% of TNA invested in equities. However, some equity funds, particularly balanced funds, sometimes have less than 50% of total portfolio value invested in equities. To avoid dropping these funds, I compute a fund's equities holdings as a percentage of its TNA over the sample period. Step 2 keeps predominately equity funds with CRSP investment style codes ED (domestic equity) or M (balanced).

CRSP reports data at the share class level, so a portfolio-level identifier is necessary to aggregate to the fund level. Step 3 keeps observations with CRSP-assigned portfolio codes leaving 15,593 share classes and 7,893 funds. Step 4 uses Morningstar-assigned portfolio identifiers when available (matching with CRSP using tickers and CUSIP numbers) and otherwise manually assigns portfolio codes using fund names yielding 18,119 share classes for 9,120 funds.

Step 5 aggregates share classes to the fund level. KNS does not specify their method, so I compute the net asset value-weighted average of class-level returns and fees and sum the class-level net asset values. Step 6 drops duplicates arising from CRSP reporting errors. Step 7 annualizes the monthly

available for 65% of observations (see Table 7 in Zhu, 2020, Management Science: "The missing new funds.") These are smaller funds with higher growth rates, which also explains some of the differences between the authors' sample and ours. In this paper, we manually created a matching code ourselves, because too many funds would drop out of the sample otherwise, which is particularly relevant for minority research with only very few treated observations (in this case, only 4.5% of fund managers had a foreign sounding name)."

observations by dropping funds with fewer than 12 monthly observations each year (KNS compute fund risk as the annualized standard deviation based on 12 monthly returns) and then keeping only December. Step 8 drops team-managed funds and funds with missing manager names yielding 4,049 individual-managed funds, 3,811 fund managers, and 17,422 fund-year observations. Step 9 computes variables following the KNS appendix and drops observations with missing variables required to replicate the KNS Table 2 baseline regression. Almost all the approximate 10,000 observations dropped in step 9 (17,422 – 7,318) have missing flow information (current year fund flow, lagged year fund flow, or family flow).[12]

Finally, step 10 drops observations with the smallest and largest 1% of fund flows following KNS page 2286. This leaves 7,063 fund observations for 1,403 funds and 1,348 managers. This compares to the KNS sample of 13,091 fund-year observations, 2,553 funds, and 2,226 managers for their Table 2. Thus, I am unable to reproduce the KNS sample with my portfolio codes and following their data descriptions.

1.4 Replicating the KNS sample

Two fund data sources are used for the replication sample to minimize errors in flows estimates: The Center for Research in Security Prices (CRSP survivor-bias-free) database and the Morningstar Direct (Morningstar) database. The two databases are merged by matching funds on ticker symbol and CUSIP number and then dropping funds that are not in both databases. The intersected CRSP–Morningstar dataset facilitates the comparison of reported returns and asset values—which are used to estimate flows—across the two databases to improve data accuracy (Berk and van Binsbergen 2015; Pástor, Stambaugh, and Taylor 2015). Additionally, because the two databases are matched on ticker symbols and CUSIP numbers, the intersected dataset excludes funds without tickers and CUSIPs, namely, funds that are not generally available to outside investors (97% of fund observations have ticker symbols in either the CRSP or Morningstar databases).

Appendix Table 3 illustrates the importance of this approach. The ratio of monthly fund total net assets (TNA) reported in CRSP to the value reported in Morningstar Direct ranges from 0.110 to

[12] I drop single fund families as the fund manager may be the fund sponsor and KNS require family flows less fund flows. I also drop funds without family names in CRSP.

100,000 with a mean value of 1.919 times. Appendix Table 1 also shows considerable discrepancies in reported monthly returns where the CRSP database reports returns that range from 24% less to 21% more than what is reported in Morningstar Direct. These differences mean that one or both databases contain errors. Thus, calculating flows from either database alone—KNS uses CRSP—will result in numerous and large errors. These errors can be mitigated by matching on observations with similar reported TNA and return values in both databases. Appendix Table 3 provides comparisons (in brackets) after dropping funds with differences in fund TNA greater than 1% and differences in returns greater than 0.10% in the CRSP and Morningstar databases. This approach reduces the incidence of extreme flows estimate errors compared to using a single data source.

Appendix Table 4 details the sample construction process. Step 1 begins with the intersected CRSP–Morningstar monthly fund-level dataset that includes 11,278 funds compared to 11,374 in the CRSP only database (not reported). Step 2 drops observations without fund manager names in the Morningstar database following KNS and Massa, Reuter, and Zitzewitz (2010). For team-managed funds, only the manager with the longest tenure, who is classified as the lead manager, is included. This step reduces the number of funds in the sample from 11,278 to 6,834 because the lead manager is often not easily identifiable (e.g., multiple managers with the same start dates). In step 3, funds with missing segment or family names are dropped because subsequent analysis includes segment fixed effects and segment and family flows. This reduces the sample funds to 6,791 funds and 5,998 managers. Step 4 keeps only predominantly domestic equity funds (following KNS). Specifically, international funds, bond funds, real estate funds, target date funds, allocation funds with less than 50% equity holdings, and other non-equity funds are dropped using Morningstar's Investment Category assignments. Dropping non-domestic funds yields 3,042 funds and 3,183 managers.

Appendix Table 3 shows considerable differences in the CRSP and Morningstar TNA and returns data which, if uncorrected, can lead to erroneous flows estimates. Therefore, monthly fund observations with missing and non-matched TNA or return values are removed in steps 5 through 8. To avoid partial year returns, TNA, and flow measurement errors, funds with fewer than 12 monthly observations per year are removed in step 9. Because most mutual funds have multiple share classes that differ in expenses, loads, and clientele, the different classes are combined into a single fund. Specifically, I compute the net asset value-weighted average of class-level returns and sum the class-level net asset values. December values are retained to avoid errors arising from differences in fund fiscal year ends (step 10).

Step 11 drops fund-year observations with missing data necessary to replicate the baseline KNS flow regression while Step 12 trims fund flows at the 1% level (each side). This leaves a final sample of 5,302 fund-year observations, considerably less than the 13,091 fund-year observations in the published study.

Appendix Table 1
KNS reproduction sample construction

Step (1)	Procedure (2)	Frequency (3)	Classes (4)	Funds (5)	Observations (6)	Managers (7)
1	CRSP	Monthly 1991-2011	47,853		3,942,928	
2	Keep only funds with KNS assigned portfolio/fund identifiers	Monthly	13,511	4,804	1,293,417	
3	Aggregate to fund level	Monthly		4,804	565,252	
4	Less duplicates (CRSP errors)	Monthly		4,804	547,343	
5	Less funds missing 12 observations per year and keep only December	Annual		4,557	42,284	
6	Merge with KNS provided manager name, tenure, and classification files	Annual		4,430	25,104	3,607
7	Less dropping remaining non-domestic equity funds and funds with missing variables needed for Table 2 of KNS (2015)	Annual 1993–2011		2,299	11,884	2,117
8	Drop observations with smallest and largest 1% of flows	Annual 1993–2011		**2,258**	**11,438**	**2,085**
KNS final sample		Annual 1993–2011		**2,553**	**13,091**	**2,446**

Description: This table provides step-by-step details on how the KNS reproduction sample was constructed using KNS provided portfolio codes and manager names, tenures, and foreign sounding name classifications. Column 1 reports the steps, Column 2 reports the procedure, Columns 3 reports data frequency, Columns 4 to 7 report the number of unique share classes, unique funds, the number of observations, and the number of unique manager names.

Interpretation: The KNS sample size can be approximately reproduced using KNS provided portfolio codes.

Appendix Table 2
Reproduction sample construction

Step (1)	Procedure (2)	Frequency (3)	Classes (4)	Funds (5)	Observations (6)	Managers (7)
	CRSP	Monthly 1991-2011	47,853		3,942,928	
1	Keep only funds with tickers	Monthly	37,764		3,128,604	
2	Keep only domestic equity funds	Monthly	18,198		1,416,150	
3	Keep only funds with CRSP assigned portfolio/fund identifiers	Monthly	15,593	7,893	1,033,507	
4	Add Morningstar and manually assigned portfolio identifiers	Monthly	18,119	9,120	1,411,938	
5	Aggregate to fund level	Monthly		9,120	602,077	
6	Less duplicates (CRSP errors)	Monthly		9,120	598,300	
7	Less funds missing 12 observations per year and keep only December	Annual		8,609	47,865	
8	Less team managed and missing manager funds	Annual		4,049	17,422	3,811
9	Less dropping funds with missing variables needed for Table 2 of KNS (2015)	Annual 1993–2011		1,447	7,318	1,372
10	Drop observations with smallest and largest 1% of flows	Annual 1993–2011		**1,403**	**7,063**	**1,348**
KNS final sample		Annual 1993–2011		**2,553**	**13,091**	**2,446**

Description: This table provides step-by-step details on how my reproduction sample was constructed following KNS. Column 1 reports the steps, Column 2 reports the procedure, Columns 3 reports data frequency, Columns 4 to 7 report the number of unique share classes, unique funds, the number of observations, and the number of unique manager names.

Interpretation: The KNS sample size cannot be reproduced using my portfolio codes.

Appendix Table 3
CRSP and Morningstar TNA and return discrepancies

	Mean	Standard Deviation	Minimum	1st Percentile	Median	99th Percentile	Maximum
Ratio of CRSP to Morningstar monthly TNA	1.919	144.790	0.110	0.985	1.000	1.046	100,000.00
	(1.000)	(0.002)	(0.990)	(0.994)	(1.000)	(1.007)	(1.010)
CRSP less Morningstar monthly returns (%)	–0.000	0.052	–24.481	–0.044	0.000	0.045	21.282
	(0.000)	(0.011)	(–0.100)	(–0.038)	(0.000)	(0.040)	(0.100)

Description: This table reports descriptive statistics for the reported monthly TNA and return discrepancies between the CRSP and Morningstar mutual fund databases. The values in brackets are the descriptive statistics after dropping funds with differences in fund TNA greater than 1% and differences in returns greater than 0.10% in the CRSP and Morningstar databases. The data are at the fund level and cover the period 1991–2016.

Interpretation: Calculating flows only using CRSP fund size and return values can lead to large measurement errors.

Appendix Table 4
Replication sample construction

Step (1)	Procedure (2)	Frequency (3)	Funds (4)	Observations (5)	Managers (6)
1	CRSP/Morningstar intersection	Monthly 1991–2016	11,278	1,385,492	
2	Matching on Morningstar's single and lead manager files		6,834	789,826	6,033
3	Less missing segment and family names		6,791	787,446	5,998
4	Keep only domestic equity funds		3,042	348,555	3,183
5	Less funds missing CRSP or Morningstar TNA		3,040	343,236	3,179
6	Less CRSP/Morningstar TNA discrepancies >1%		3,005	330,370	3,158
7	Less observations with missing CRSP or Morningstar returns		3,000	323,546	3,152
8	Less CRSP/Morningstar return discrepancies >0.10%		3,000	322,641	3,152
9	Less funds missing 12 observations per year		2,541	235,800	2,644
10	Keep only December	Annual	2,541	19,650	2,644
11	Less dropping funds with missing variables needed for Table 2 of KNS (2015)	Annual 1995–2011	**1,216**	**5,408**	**1,175**
12	Drop observations with smallest and largest 1% of flows	Annual 1993–2011	**1,210**	**5,302**	**1,170**
KNS Test Sample		Annual 1993–2011	**2,553**	**13,091**	**2,446**

Description: This table provides step-by-step details on how the replication sample is constructed using the CRSP/Morningstar intersected database. Column 1 reports the steps, Column 2 reports the procedure, Columns 3, 4, and 5 report the number of unique funds, the number of observations, and the number of unique manager names.

Interpretation: The intersected CRSP/Morningstar database and correcting for TNA and return discrepancies yields a replication sample that is smaller than KNS despite.

Appendix Table 5
Estimating Flows

	(1) Annual Aggregation	(2) Monthly Aggregation
Foreign75	–0.010	–0.009
	(–0.51)	(–0.50)
Obs.	10,898	10,898
Adj. R-squared	0.182	0.186

Description: Appendix Table 5 reports baseline regression results (all controls from KNS Table 2 included but not reported) for the intersected KNS and KNS reproduction samples. Annual aggregation, where annual class level returns are computed and then aggregated to the fund/portfolio level prior to computing flows, is the method used by KNS. Monthly aggregation, is where class level monthly returns are aggregated to the fund level and then compounded to the annual level prior to computing flows, is the method used for the KNS reproduction, reproduction, and replication tests.

Interpretation: The estimated foreign sounding manager coefficients are insignificant for observations common to the KNS and KNS reproduction samples. Observations that are only in the KNS sample appear to drive the published finding that funds with managers having foreign sounding names garner less flow.