Ethnic Investing and the Value of Firms*

Jonas Hjort
UCL
& BREAD & CEPR & U. Oslo

Changcheng Song
U. Florida

Christopher Yenkey
U. South Carolina

February 3, 2025

Abstract

We study ethnic investing, using transaction data from Kenya's stock exchange and CEO/board turnover. We show that a given investor invests more in a given firm when the firm is run by coethnics and earns lower risk-adjusted returns on such investments. We then model and empirically test for the aggregate impact of (i) the implied taste- or psychology-driven investor discrimination and (ii) counteracting demand- and supply-side forces. Our estimates imply that listed Kenyan firms could collectively be worth 38% more—with minority-run firms benefitting the most—if the neutral proportion of active investors increased from 4.6 to 50%.

*j.hjort@ucl.ac.uk, ccsong@smu.edu.sg, cyenkey@moore.sc.edu We are grateful to John Y. Campbell, Francesco D'Acunto, Kent Daniel, Stefano DellaVigna, James Fenske, Raymond Fisman, Andrew Hertzberg, Harrison Hong, Jianfeng Hu, Amit Khandelwal, Weikai Li, Roger Loh, Suresh Naidu, Michaela Pagel, Tommaso Porzio, Jonah Rockoff, Paul Tetlock, and seminar participants at AEA/ASSA 2022, CERGE-EI, Columbia, Harvard/MIT, HKUST, Marseille School of Economics, University of Namur, NYU, Ottawa, Paris School of Economics, UCLA, UCSD, USC, Warwick, and Wharton for helpful suggestions; to Kenya's Central Depository and Settlement Corporation, Ltd. (CDSC) for data access; and to Sawal Acharya, Josephine Gantois, Naman Garg, Sakshi Gupta, Alokik Mishra, and especially Junyi Que and Xinyi Shen for superb research assistance.

1 Introduction

People tend to invest more in others to which they are linked through ethnic ties.[1] This may be due to information asymmetries arising from easier communication or screening among coethnics, in which case investors will tend to earn higher returns on coethnic investments (Lang, 1986; Greif, 1993; Cornell & Welch, 1996; Fisman *et al.*, 2017). Alternatively, investors may have a taste for—or a psychological or social bias towards—investing in coethnics, in which case they will tend to earn lower returns on coethnic investments (Becker, 1957; Hjort, 2014; Fisman *et al.*, 2020).

The *aggregate* economic consequences of coethnic investing—an elusive target for economists since Banerjee & Munshi (2004) showed evidence that ethnic-majority firms benefit from easier access to capital—depend on the nature and magnitude of these investor biases. They also depend on market responses to favoritism (Arrow, 1973; Shleifer & Vishny, 1997; Van Nieuwerburgh & Veldkamp, 2009).

In this paper we study the extent, nature, and market-wide impact of coethnic investing in Kenya. We use complete 2006-2010 transaction level data from the Nairobi Securities Exchange. We first show that a *given Kenyan investor* invests considerably more *in a given firm* when its CEO and/or board is of the same ethnicity as the investor, and earns lower risk-adjusted returns on such "coethnic investments". We use a simple model to illustrate how supply- and demand-side responses counteract such investor taste for or bias towards coethnic firms, and show evidence consistent with these predictions—but also that market responses appear to far from offset the overall impact on Kenyan firms' value.

The Kenyan stock market is an ideal setting to study cothnic investing. First, ethnic divisions are salient in Kenyan society (Ndegwa, 1997; Barkan, 2004; Berge *et al.*, 2020). Second, some investment objects—here, large firms—in effect change ethnicity across time in Kenya, and we observe the investment behavior of tens of thousands of ethnically identifiable investors. We can thus estimate how coethnicity affects investment *within investor-investment object pairs*. Third, since we study "atomistic" investors whose returns are observed—pecuniary, unobserved dimen-

[1] See e.g. Fafchamps (2000); Rauch (2001); Banerjee & Munshi (2004); Guiso *et al.* (2009); Hjort (2014); Burgess *et al.* (2015); Beach & Jones (2017); Fisman *et al.* (2017); Burchardi *et al.* (2019); Fisman *et al.* (2020).

sions of returns are unlikely in retail stock market investing[2]—we can distinguish returns-increasing and -decreasing sources of discrimination. Finally, observing a delineated, complete market and the full extent of its supply and demand side allows us to study counteracting forces and the ultimate impact market-wide.

We start by documenting a positive and large coethnicity effect in investment decisions. To do so, we regress investments in a given firm on measures of the firm's CEO and/or board belonging to the same ethnicity as the investor in the month in question. We show that the particular parallel trends assumption required to interpret the estimate causally appears to hold.

We next show that the risk-adjusted return on coethnic investments is on average lower. This underpins a simple model characterizing how coethnic investing arising from investor preferences or biases "misallocates demand". The model also illustrates how supply- and demand-side market participants—firms, through their choice of (CEO) ethnicity, and neutral investors—can benefit from partially or fully equating demand for and supply of each type of firm.

Finally, we show evidence suggesting that coethnic investing markedly lowers firms' average value in Kenya. In the first of three ways we probe the model's predictions, we show that a given firm's price-to-book value increases when its "coethnic investor base"—the *proportion* of all portfolio wealth held by active investors of the same ethnicity as the firm's CEO—rises. Next, firm values are higher when the proportion of overall equity held by neutral—foreign and institutional—investors is higher.[3] Individual firms benefit less from neutral investors than from coethnic investors, but minority-ethnicity firms benefit more from neutral investors than majority-ethnicity firms do.

A sharper form of variation arises on the market's supply-side when a change in "firm ethnicity" resulting from CEO turnover changes and abruptly increases a firm's coethnic investor base. We find that the firm's price-to-book value then also increases. In contrast, when CEO turnover abruptly decreases a firm's coethnic investor base, firm value also decreases.

[2]The mean and median of the share of firms owned by each of their 10 largest individual Kenyan shareholders is 0.13% and 0.05% in our data.

[3]For evidence on institutional investors' role in rich countries' stock markets, see e.g. Gabaix *et al.* (2006); Boehmer & Kelley (2009); Campbell *et al.* (2009); Basak & Pavlova (2013); Edelen *et al.* (2016). Do *et al.* (2024) show compelling evidence that investments in firms with Jewish connections earned higher returns during a period of increased antisemitism in 19th century France.

These results imply that demand- and supply-side forces counteract but do not offset the impact of coethnic investing on firm value. One of our back-of-the-envelope calculations suggests that the total value of the firms on Kenya's stock exchange could be 38% (or USD 5.35 billion in 2010) higher if the proportion of neutral investors in the market was one-half rather than the monthly average of 4.6%.

Economists have long been interested in market-wide economic costs of discrimination (Becker, 1957; Arrow, 1973; Phelps, 1972; Charles & Guryan, 2008; Hsieh *et al.*, 2019). We make progress by analyzing a complete market associated with comparatively efficient capital allocation; wherein a specific form of "micro" (investor×firm) level discrimination can be identified, yielding theoretical predictions for how value creation should be affected in the absence of fully compensating supply- or demand-side responses; and in which such responses are also observed.

The body of work on ethnic ties and investment is related but distinct (see e.g. Rauch & Trindade, 2002; Banerjee & Munshi, 2004). Focusing on a stock market and large firms—on both of which there is almost no research from developing countries[4]—we estimate the causal effect of coethnicity *within a given investor-investment object pair*. Existing studies instead capture shared identity combined with correlated, unobserved match characteristics, by comparing a given investor when "assigned" to a coethnic versus a non-coethnic investment opportunity (Hjort, 2014; Fisman *et al.*, 2017, 2020) or vice versa (Burgess *et al.*, 2015; Burchardi *et al.*, 2019).

Finally, this paper relates to the literature that studies the *nature* of discrimination (see List & Rasul (2011); Charles & Guryan (2013); Bertrand & Duflo (2017) for overviews, and Bohren *et al.* (2019); Rose (2023); Bohren *et al.* (2024) for recent related work), and the parallel finance literature on discrimination and "home bias" in investing (see, for overviews, Lewis, 1999; Coeurdacier & Rey, 2013; Cooper *et al.*, 2013; Ardalan, 2019). We show that in a context where investors are "atomistic" and risk-adjusted financial returns are plausibly fully observed, the primary explanation underlying coethnic investing appears to be preferences or psychology. This may constrain regions and firms with small or poor investor bases (Teoh

[4] Anagol & Kim (2012); Anagol *et al.* (2018, 2021); Yenkey (2015, 2018a,b) are important exceptions.

et al., 1999; Banerjee & Munshi, 2004; Banerjee & Duflo, 2005).[5]

2 Data

Detailed information on the data we use is in Appendix A1; we now provide an overview. The version of the NSE's Transactions Registry we have access to reports the firm's ticker id, the number of shares traded, the price, the seller's (masked) id, the buyer's (masked) id, and the date for all trades that occurred on the NSE from January 1, 2006 through December 31, 2010. Short-selling was not allowed during this period.

We do not observe shares that an investor had bought before the NSE "went digital" in 2006 and did not trade thereafter.[6] To construct a measure of an investor's portfolio, we thus assume that all investors have zero holdings as of 2006. We thereafter simply add any observed purchases to investor i's inferred holdings, and subtract any observed sales (see Appendix A1). Our results are very similar if we instead focus only on investors who opened their NSE account in 2006 or later, in which case we observe investors' full portfolio at every point in time. The fact that we do not observe pre-2005 holdings is also not relevant for the "flow" measure of coethnic investing that, as we describe in Section 3, is our preferred measure.

The version of the NSE's Investor Registry we have access to reports the investor's (masked) id, account creation year, and—crucially—last name. In addition, the names of listed firms' CEO and board-members are publicly available. Information on firms' book value, outstanding shares, etc, come from their financial statements.

[5] Returns are usually unobserved. We follow Cohen *et al.* (2008); Bandiera *et al.* (2009); Hjort (2014); Fisman *et al.* (2017, 2020) in using returns to tell apart discrimination motivations. However, "non-atomistic" investors who favor coethnics may benefit outside of the observed market or workplace. In the finance literature, this paper is most closely related to Teoh *et al.* (1999). They find that boycotts of South Africa in the 1980s had little discernible effect on the valuation of firms operating in the country and financial markets because corporate involvement with South Africa was small. Our analysis of individual investors' decisions is also closely related to Hong & Kacperczyk (2009), Kumar *et al.* (2015), and Barber *et al.* (2021). They show that firms that do not promote vice or funds that promise "impact" or are run by managers with American-sounding names command a higher willingness-to-pay from particular investors.

[6] About 36% of stock held at the beginning of 2006 was sold during the 2006-2010 period, or conversely; 64% of *starting-point* holdings were not traded during our data period. However, total value traded from 2000 to 2005 was about 20% of that from 2006 through 2010, a period when the number of investors on the NSE grew rapidly (Yenkey, 2015). The Privatization Act of 2005 lowered entry barriers to retail investing by digitizing the trading system and by requiring firms to make a higher proportion of newly issued shares accessible to domestic, small-scale investors via smaller lots.

Appendix Table A1 provides summary statistics on our analysis sample. We restrict attention to investors who trade five or more times at least one year. There are about 55,000 such investors for which we can also infer ethnicity, with average portfolio values of around USD 6,000 in 2006.[7] The 47 firms that appear on the NSE cover a range of sectors, with 53 percent in "Services", 38 percent in "Industrial", and 9 percent in "Agriculture". They are large, with an average total market capitalization of around USD 261 million in 2006.

We probabilistically assign ethnicities to investors, CEOs, and board-members using their last names. The starting point is name×ethnicity match probability information recorded by Yenkey (2015, 2018a,b). We then construct four measures of an investor's ethnic proximity to a firm's CEO and board; Appendix A2 has detailed information on their construction. The first, CoethnicCEO$_{ijt}$, is an indicator equal to one if investor i and the CEO running firm j in month t are relatively likely to belong to the same ethnicity—they share a *Likely Ethnicity*. The second, CEOCoethnicityIndex$_{ijt}$, is a 0 (minimum proximity) to 1 measure of the expected ethnic proximity between the investor's and the CEO's name, given each person's expected probability of belonging to each ethnicity. Specifically, the index is equal to the inner product of the investor and the CEO's name×ethnicity match probabilities. In this case we can make use of the full sample, and we avoid restricting attention to the investor's and CEO's most likely ethnicity and the judgment required to define a *Likely Ethnicity*.[8]

One board measure, BoardCoethnicityIndex$_{ijt}$, is the proportion of board-members that are coethnic (measured as for CoethnicCEO$_{ijt}$) with the investor. The other, CoethnicBoard$_{ijt}$, is an indicator, and essentially repeats the construction of CoethnicCEO$_{ijt}$ twice, first between individual board-members and the investor, then for the board as a whole vis-a-vis the investor. This measure of investor-board coethnicity is strict: for CoethnicBoard$_{ijt} = 1$ in month t, each board member must be classified as either a likely coethnic or non-coethnic, and the board as a whole must be more likely to share the investor's ethnicity than any other. We will see that all four

[7] This includes a small number of brokers and institutional investors for which we can infer the ethnicity of the individual listed as account owner.

[8] CEOCoethnicityIndex$_{ijt}$ is "assumptions-free" in that it follows directly from the raw data from the RAs. The reason why this measure also allows us to make use of a larger part of our sample is that it does not require leaving out observations for which we cannot assign a name to a given ethnicity with confidence. CoethnicCEO$_{ijt}$ is e.g. missing if either the investor or the CEO does not have a *Likely Ethnicity*.

definitions of CoethnicFirm$_{ijt}$ give similar results.[9]

In Panel C of Table A1 we show statistics at investor×firm×month level. Of particular interest is coethnicity between investors and firms. With ethnicity defined as our preferred measure of *Likely Ethnicity* (see Appendix A2 for details), 48.8% of investors in our sample are classified as Kikuyu. Kikuyus are the biggest ethnic group in Kenya—roughly 17% of the population—and their disproportionate representation among investors on the stock market may in part reflect their comparatively strong economic position. 5.8% of investors are Luo, 5.4% Kamba, 4.2% South Asian, 3.6% Luhya, and 2.9% Kalenjin. On the CEO (or "firm type") side, 26.9% of CEOs are classified as Kikuyu, 25.4% as Anglo, 7.5% as South Asian, 4.5% as Luo, 3.0% as Kamba, 3.0% as Meru, and 1.5% as Luyha. With coethnicity measured as CoethnicCEO$_{ijt}$=1, the investor belongs to the same ethnicity as the CEO in 27% of investor×firm×month observations in our analysis dataset.

3 Ethnic Investing in Kenya

We observe which particular investors belong to the same ethnicity as each firm's management at each point in time, and coethnicity status changes when CEOs and board-members are replaced by others of another ethnicity. We first run:

$$\text{Investment}_{ijt} = \alpha + \beta \text{CoethnicFirm}_{ijt} + \gamma_i + \delta_j + \psi_{c(jt)} + \theta_t + X_{jt} + \varepsilon_{ijt} \qquad (1)$$

where Investment$_{ijt}$ is the value of the investment investor i holds in firm j in month t, normalized by the total value of all her investments—a "portfolio weight" (Cohen *et al.*, 2008; Hvide & Døskeland, 2011). We also exploit the granularity of transactions data to construct OrderImbalance$_{ijt}$: the value of shares in firm j purchased by investor i in month t minus the value of shares sold, divided by the sum of purchases and sales by i in j at t (see e.g. Chordia *et al.*, 2002).[10]

[9] A sensible further alternative to the eight coethnicity codings we show results from (the four preferred ones used in our main tables and figures and four alternative ones using different probability cutoffs to define ethnicity in Appendix Table B1) is to define an investor-CEO (or board) pair as "coethnic" if all the Kenyan research assistants who coded the names in our data unanimously agreed that the two names represent the same ethnic group. This yields estimates of the effect on Investment, OI, and risk-adjusted returns that are consistent with—and if anything bigger than—those resulting from our preferred, broader coethnicity definitions, but this approach relies on a small number of coethnic "matches" (only about 4% of investor-CEO pairs in the sample are defined as "perfectly coethnic" in this sense).

[10] OrderImbalance$_{ijt}$ is a *flow* measure of investment. Unlike Investment$_{ijt}$, it also is not influenced by the evolution of (the values of) an investor's holdings. Normalizing by volume traded is standard: doing so controls for potential liquidity

In addition to month fixed effects θ_t, we include investor, firm, and CEO ethnicity fixed effects γ_i, δ_j, and $\psi_{c(jt)}$ so that our results are not driven by differences across investors, firms, or the various ethnic groups in our data. We also include a measurable "value control" (X_{jt}) that varies at the firm-month level, the return-on-equity (ROE) over the past 12 months. We cluster the error term ε_{ijt} at the investor ethnicity×CEO ethnicity level.

The share of investments held in a given firm is 1.8% higher if the firm is managed by a coethnic CEO at the relevant point in time (CoethnicCEO$_{ijt} = 1$). This is shown in the top panel of Table 1. Similarly, the fraction of her investments an investor holds in a firm is 2% greater when she has maximum ethnic proximity to the firm's CEO (CEOCoethnicityIndex$_{ijt} = 1$ vs. $= 0$).[11] Columns 3 and 4 show that Investment$_{ijt}$ is 3.5% higher if firm j is managed by a coethnic board at time t (CoethnicBoard$_{ijt} = 1$), and 8.5% greater when investor i has maximum ethnic proximity to the firm's board (BoardCoethnicityIndex$_{ijt} = 1$ vs. $= 0$).

We also find that investor i's normalized net investment in a given firm—OrderImbalance$_{ijt}$—is 11% greater if the firm is managed by a coethnic CEO in the month in question; 18% greater with maximum ethnic proximity to the CEO; 70% greater if the firm is managed by a coethnic board; and 167% greater with maximum ethnic proximity to the board. This is shown in the bottom panel of Table 1. Appendix Table A2 shows that the results are very similar if we restrict attention to investors' *buys*, ignoring their sells, or if we exclude Anglo and South Asian investors (see Appendix Table A3). The same is true if we exclude the largest investors in the sample.[12] We also find some evidence that coethnic bias in investment decisions is somewhat greater for investments in Services than in "Industrial". We leave a deeper investigation of heterogeneity in ethnic investing across individual investors to future research.

The estimates in Table 1 capture how "coethnicity itself" and any correlated match characteristics of investor-firm pairs affect investment. They rely both on cross-sectional variation of the form used in existing studies of ethnic discrimination in markets and workplaces—loosely, comparing the investment of investors

differences across observations (see e.g. Chordia *et al.*, 2002).

[11] Ethnic concentration increases during our data period: "the other side of the market" in transactions that increase the portfolio weight of coethnic firms is often simply non-coethnics of the CEO or board. The period we study is also one in which many small investors joined the stock market.

[12] For example leaving out the 10% of investors with highest portfolio value, or the 10% biggest investors in each firm.

A and B in firm 1 relative to firm 2, when one investor shares an ethnicity with one of the two firms and the other with neither—and on "changes in coethnicity" *within investor-investment object pairs*.[13] We next replace γ_i and δ_j with an investor-firm fixed effect. By only using CEO/board turnover to estimate β, we isolate a more precisely defined coethnicity effect—how shared identity affects investment—under a particular identifying assumption.[14] That assumption is that investment in particular firms—those that switch from being managed by a CEO/board of ethnicity A to one(s) of ethnicity B—would have evolved similarly, relative to investment in other firms, for investors of ethnicity A and B compared to other investors, absent such switches.

Before regression results, we examine this assumption. Thirteen out of the 47 firms in our sample "change ethnicity" (and three do so multiple times). In Figure 1 we see that investment from "post-coethnics"—coethnics of the incoming CEO—rises markedly in the month the new CEO takes over relative to investment from investors who are coethnic with neither the outgoing nor the incoming CEO.[15] In the subsequent month, the *flow* of investment from post-coethnics is again similar to that of "others". From month 2 onwards we see indications of post-coethnics investing more than others again. Most importantly, we see no indication of concerning non-parallel pre-trends in Figure 1.

Table 2 shows that overall investment increases significantly within a given investor-firm pair when a CEO or board-member of a different ethnicity than the investor is replaced by a coethnic. We lose some power when restricting to investor-firm pairs that change CEO coethnicity status and including investor-firm fixed effects. However, the estimates in Table 2 suggest that investors if anything adjust the share of their investments that is held in a given firm somewhat more when coethnicity "turns on" within a given investor-firm pair. The patterns in Figure 1 provide direct, visual support for a causal, shared identity-based interpretation of the results in Table 2. This is important for interpreting the estimates

[13] Table 1 combines both forms of variation because a comprehensive notion of coethnicity is most relevant for aggregate economic consequences.

[14] Unobserved match effects between managers and investors that are correlated with coethnicity are—even among retail investors—a possibility, but less plausible. They would be hard to reconcile with Section 4's results on returns, and managers tend to have much less influence on firm activities than features of the firm itself do (Bertrand & Schoar, 2003).

[15] Figure 1 depicts results from a dynamic version of the (OrderImbalance$_{ijt}$ version of) specification (1) estimated on the investor×"switcher-firms" sample. It therefore maps to the Table 1 market-wide results, which motivate the model in Section 5.

in Table 1. Together, the three indicate that the coethnicity effect we uncover is driven by investor-management *match effects*. Various underlying phenomena may contribute to or explain these match effects, and they may also be heterogeneous across ethnic groups. An example is the possibility that investors from larger ethnic groups expect firms run by CEOs from their group to benefit from political or consumer favoritism and that their investment behavior also contributes substantially to the estimates in tables 1 and 2.

In this section we showed that Kenyan investors invest considerably more in a given firm when the firm is run by coethnics, and that such coethnic investing appears to be driven to a large extent by shared identity itself.

4 Understanding Ethnic Investing

Ethnic investing may broadly be due to *information asymmetries* or investor *preferences or biases*. Both are consistent with the results in Section 3, but the former imply higher returns, and the latter lower or equal returns. The documented role of shared identity may point towards preferences or biases. To investigate, we run:

$$\text{RiskAdjReturns}_{bijt} = \alpha + \beta \text{CoethnicFirm}_{ijt} + \gamma_i + \delta_j + \psi_{c(jt)} + \theta_t + X_{jt} + \varepsilon_{bijt} \quad (2)$$

For stocks bought and sold during our sample period, we define (risk-unadjusted) returns as the realized return based on the buy and sell price. For stocks that were bought but not sold during our data period, we compute returns as unrealized paper returns on the 31st of December 2010.[16] We then measure $\text{RiskAdjReturns}_{bijt}$, the risk-adjusted return on investment b made by investor i in firm j in month t, in several ways; our preferred measure is simply the Sharpe Ratio—the difference between the returns on the investment and the risk-free return, divided by the standard deviation of the difference. Within a given firm-month pair—that is, for "buys" of the stock of a given firm made in a given month—$\text{RiskAdjReturns}_{bijt}$ varies across investors. This is because different investors invest in the firm on different days within the month and sell their stocks at different times.

A given investor's risk-adjusted return on investments in a given firm in a given

[16] We later show results restricted to realized returns.

month are respectively 9.6 and 21.8% lower if the investment is made when the firm is run by a coethnic CEO or when the investor has maximum ethnic proximity to the CEO, relative to when the CEO is a non-coethnic.[17] This is shown in the top panel of Table 3. A different way to illustrate the magnitude of the estimates is as follows. Column 1 shows that an investor's return on investments is 1.12% lower for one unit of risk taken if the investment is made when the firm is run by a coethnic CEO, relative to when the CEO is a non-coethnic. The median of risk (standard deviation of excess returns) in our sample is 1.55. Therefore, for investors taking median risk, return on investments is 1.74% lower if the investment is made when the firm is run by a coethnic CEO, relative to when the CEO is a non-coethnic. In many columns of the bottom panel the negative estimates are even larger in magnitude; there we estimate how coethnic investing affects returns by exploiting changes in coethnicity within investor-firm pairs.

The measures of returns and risk we use are common in the finance literature. Alternative measures generally give similar findings. The conclusions are for example very similar to those from Table 3 if we restrict to investments both bought and sold or to firms whose CEO ethnicity remains constant during our data period[18], and if we focus on end-of-first-year returns, as shown in Appendix tables A4, A5, and A6. For the latter exercise, the end-of-first-year return is calculated based on transaction price and the price of the last day in the first calendar year (see Appendix A3 for more details). Appendix figures A1 and A2 show that monthly returns are on average lower soon after a firm "changes ethnicity" and that "post-coethnics" additionally earn lower returns compared to others in the period after such switches. In Figure 1 we saw that post-coethnics are especially likely to invest at such times. In Appendix A3 we provide more details and additional results.

Our results so far indicate that Kenyan investors engage in coethnic investing, and that a taste for or psychological bias towards coethnic firms appears to be the most common motivation. In a setting where individual investors are generally small, these average behaviors and motivations are the natural starting point

[17] Similarly, an investor's risk-adjusted return on investments made when the firm's board is generally of the same ethnicity as the investor are 44.8% lower. For the $\text{BoardCoethnicityIndex}_{ijt}$ measure of $\text{CoethnicFirm}_{ijt}$, the point estimate is small and noisily estimated, but positive.

[18] For these firms the estimated differential return on coethnic investments cannot be due to any stock price dynamics associated with CEO (ethnicity) turnover.

for a theoretical framework focusing on the aggregate economic consequences of coethnic investing. In the next section we present such a model; corresponding empirical tests are in Section 6. A primary focus is how neutral investors affect the aggregate impact of the costly form of favoritism that individual investors in our sample display on average, accounting for variation in investor group size. We theoretically conceptualize neutral investors as a *different category* than biased-on-average individual investors, and empirically proxy for them simply with foreign and institutional investors.[19]

5 Theoretical Framework

If each group of investors primarily invests in firms of a specific type, this will—relative to a scenario in which investors are neutral—tend to lower the average value of firms. Investors as a whole could create more value by investing in firms with a smaller investor base. This "clientele" prediction, first emphasized by Merton (1987), only holds if responses to coethnic investing on the demand- and supply-side of the market are limited in scope, however.

We consider a model of financial markets where firms differ in ethnicity and some investors favor coethnic firms. We focus on the two-ethnicities case. We describe the set-up, results, and intuition here: details and proofs are in Appendix A4.

5.1 Firms

We study a one-period world with two types of firms, which differ in ethnicity. Firms of a given type have the same production technology, characterized by a normally distributed cash flow with mean μ_i and variance σ_i^2, where $i = 1, 2$. Cash flows within each firm type are perfectly correlated, while the covariance between cash flows of different types is σ_{12}. Besides firm stocks, there is a riskless asset with rate of return normalized to zero and perfectly elastic supply. Borrowing is allowed, but short-selling of risky assets is not.

[19] We show in the appendix that the model's key results hold also in the case where only a subset of local investors are ethnically biased.

5.2 Investors

There are three investor types. Biased investors only invest in firms run by co-ethnics, while neutral investors invest in both firm types.[20] Let I denote the total number of investors, α the share of neutral investors, and β the share of biased investors in ethnic group 1. All investors have constant absolute risk aversion preferences with risk tolerance τ.

5.3 Equilibrium and Results

We solve for investors' optimal portfolio choices and equilibrium prices in Appendix A4, assuming for simplicity that the two types of firms differ only in their ethnicity. In the same appendix, we also analyze the differences between the mixed scenario (biased and neutral investors) and the scenario where all investors are neutral.

Prices are determined by four primary effects. Consider stock prices for firm type 1 (p_1):

- **Biased investors demand effect:** A higher share of biased investors in group 1 (β) increases demand for these firms' stocks, pushing p_1 upward.
- **Crowd-out demand effect:** Neutral investors avoid overpriced stocks from group 1 when β is large, creating downward pressure on p_1.
- **Diversification demand effect:** Lower cash flow correlations enable neutral investors to diversify more effectively, increasing demand and prices for both share types. Note that this demand effect applies only to neutral investors, as biased investors invest exclusively in firms from their ethnic group.
- **Supply effect:** A large presence of group 1 shares (N_1) implies high supply, lowering p_1. If cash flows are positively correlated, a large number of shares from the other ethnic group (N_2) also reduces p_1, while a negative correlation causes N_2 to increase p_1.

[20] When we test the model's predictions empirically, this implies treating all individual, Kenyan investors as biased, since we observe a particular group of investors—institutional investors and foreigners—that are plausibly less biased (no firms are coethnic with such "neutral" investors). Appendix A5 shows that key results hold even when only a subset of investors in each ethnicity are biased.

Proposition 1. *A firm's stock price is increasing in the share of biased investors of the firm's ethnicity under reasonable conditions.*

Proof. See Appendix A4. □

To see the intuition and to focus on demand-side effects, assume $N_1 = N_2$.[21] Proposition 1 always holds when the correlation between cash flows (ρ) is negative or when it is positive and $\beta \leq \frac{1}{2}$. However, it may fail when α is very low, β is large, and $\rho > 0$. Intuitively, when most investors are biased, increases in β (which in that case is already high), slightly raise the price in group 1 but significantly lower the price in group 2. Neutral investors adjust to avoid overpriced group 1 shares, causing the crowd-out effect to dominate and preventing the price in group 1 from increasing with β. Conversely, with larger α fewer investors are biased, limiting demand responses to changes in β. As a result, share prices remain relatively stable, and neutral investors adjust less. The crowd-out effect weakens, causing Proposition 1 to hold.

Proposition 2. *A firm can benefit from changing its ethnicity from that of the smaller (investor) group to that of the larger (investor) group under reasonable circumstances.*

Proof. See Appendix A4. □

When a firm changes its ethnicity, it alters its investor base and the stock supply. For small firms with minimal impact on total supply, the benefits of switching from minority- to majority-ethnicity are unambiguous. However, large firms can exert significant downward pressure on stock prices through increased supply, making the switch unprofitable.[22]

Proposition 3. *Total market value is increasing in the share of neutral investors.*

Proof. See Appendix A4. □

[21] Inequality (11) from Appendix A4 then simplifies to

$$1 > \frac{(2\beta - 1)\rho\alpha(1-\alpha)}{\alpha + \alpha(1-\alpha)(1-2\beta) + (1-\beta)^2(1-\alpha)^2(1-\rho^2)}. \tag{3}$$

The following discussion of and take-aways from Proposition 1 generally hold also for the case where total shares in the two firm types are not equal.

[22] With only neutral investors, the benefit of switching depends solely on the relative share supply, unlike with biased investors, where correlation also matters. Without bias, there is no need to counteract distortions by leveraging cash flow correlation.

Except in extreme cash flow correlation cases, total market value is consistently higher under complete investor neutrality compared to a mixed scenario with biased and neutral investors (see Appendix A4). Coethnic bias reduces risk-sharing, distorts share prices, and increases firms' average cost of capital. Positive price distortions in one firm group are offset by negative distortions in the other, lowering total market value. As neutral investors increase, these distortions diminish, and aggregate market value rises.

Proposition 4. *A marginal increase in the share of neutral investors has a larger effect on the stock price of firms of the minority ethnicity.*

Proof. See Appendix A4. □

When outstanding shares are the same, the price of firms of the minority-ethnicity is lower and they are therefore more attractive to neutral investors. An increase in the share of neutral investors consequently affects the market value of these firms more.

5.4 Supply- and demand-side responses to coethnic investing

We have described a partial equilibrium with ethnically-biased investing of the form we documented in sections 3 and 4. Proposition 1 then implies that the price of majority-ethnicity firm shares will be higher than that of otherwise similar minority-ethnicity firm shares. We might then expect both demand- and supply-side responses.

First, unbiased investors may enter the market. Proposition 3 and 4 then predict an increase in total market value and especially in the value of minority-run firms. Second, undervalued minority-ethnicity firms may strategically respond to coethnic investing. Proposition 2 states that they can do so by "becoming" a majority-ethnicity firm, for example by appointing a CEO from the larger ethnic (investor) group.

For demand- and supply-side responses to *eliminate* the difference in investment objects' value and the impact on aggregate value creation due to ethnic investing, such responses would need to be of comparable magnitude to investor biases themselves. They may not be because markets—even text-book ones like

stock exchanges—often display costs associated with market responses. Limiting costs of arbitrage are well-established (Gromb & Vayanos, 2010), and neutral investors may have easy access to other appealing markets. Similarly, firms tend to experience significant CEO transition costs, and the labor market for CEOs is thin in Kenya (Hjort *et al.*, 2024).

6 The Consequences of Ethnic Investing

6.1 Ethnic investing and the value of a firm: empirics

To test how coethnic investing affects the price-to-book value of firms, we first run:

$$\text{PriceToBook}_{jt} = \alpha + \beta \text{CoethnicInvestorBase}_{jt} + \delta_j + \theta_t + X_{jt} + \varepsilon_{jt} \qquad (4)$$

We include firm fixed effects δ_j, month fixed effects θ_t, a value control that varies at firm$\times$month level (ROE), and cluster the error term ε_{jt} at firm level. We measure $\text{CoethnicInvestorBase}_{jt}$ simply as the portfolio value investors that are active—that is, that trade—at time t and who belong to the same ethnicity as firm j's CEO hold, *relative* to that of all potentially active coethnic investors.[23]

We restrict the sample to firms whose ethnicity remains constant during our data period; those which do not change their CEO to someone belonging to a different ethnicity. Variation in $\text{CoethnicInvestorBase}_{jt}$ thus arises on the demand-side, from investors joining or leaving the market and changes in their activity. The inclusion of firm and month fixed effects, and the focus on coethnic *potential* investors, leave room only for very particular non-causal interpretations of the results from (4).[24] Since we are not able to exploit relevant, *market-wide* exogenous variation in $\text{CoethnicInvestorBase}_{jt}$, interpreting the results through the lens of the model in Section 5 and the investor behavior documented in sections 3 and 4 that motivated it will be helpful.

[23] We define potentially active investors as all individual investors who have invested on the NSE up to and including the month in question.

[24] Two patterns are arguably necessary for the estimated coefficient on $\text{CoethnicInvestorBase}_{jt}$ in (4) to not reflect a causal effect of the relative size of firms' coethnic investor bases. First, that large numbers of retail investors of a given ethnicity become active or inactive on the stock market at times when particular firms of the same ethnicity (but not the market as a whole) whose market value is in fact not responsive to coethnic demand, would in any case have seen a notable increase or decrease in their stock price for other reasons. And second, that such dynamics are either highly correlated across firms of a given coethnicity, or driven by firms' whose "attraction" is big enough to spill over onto other firms of the same ethnicity (as we "assign" active investors of a given ethnicity to all firms of the same ethnicity).

We find that a firm's price-to-book value rises significantly relative to other firms when its coethnic investor base increases in size, consistent with Proposition 1. This result is in the first column of Panel A of Table 4. The estimate implies, for example, that we would expect the price-to-book value of a firm that is led by a CEO from an ethnic group that has the same proportional number of investors as the group with the biggest investor base observed in our data to be 67% greater than an otherwise identical firm led by a CEO from an ethnic group with a base of the same size as the smallest one we observe.[25]

To test Proposition 3, we next add the additional regressor $\text{NeutralInvestorBase}_t$ to (4). We proxy for neutral investors with foreign and institutional investors. Since they are not coethnic with any firms, $\text{NeutralInvestorBase}_t$ varies only across months.[26] The average proportion of active neutral investor is 4.6%.

Firms' average price-to-book value increases significantly when the proportion of active investors that are neutral rises. We show this in the second column of Panel A of Table 4. A doubling of the share of neutral investors is associated with 3% higher average price-to-book firm value. However, we also find—again consistent with the framework in Section 5—that neutral investors influence individual firms' value notably less than coethnic investors do. Investor favoritism is a different phenomenon than demand itself.

We next show that minority-ethnicity firms especially benefit from neutral investors. We simply add the interaction between $\text{CoethnicInvestorBase}_{jt}$ and $\text{NeutralInvestorBase}_t$ to the regression. The results in Column 3 of Panel A of Table 4 imply that we would expect the price-to-book value of a firm that is led by a CEO from an ethnic group that has the same proportional number of investors as the group with the smallest coethnic investor base observed in our data to increase 36% more in response to a doubling of the share of neutral investors than that of an otherwise identical firm with a coethnic base as large as the biggest one we ob-

[25] This calculation uses the average investor base size of firms in our data. The biggest investor base size is thus equal to the investor base size of the firm which has the maximum average size, which is 0.38. The smallest investor base size is defined analogously and is 0.00001 in our data. Since the referenced estimate in Table 4 is 1.77, the percent change corresponds to $1.77 \times (0.38 - 0.00001) \times 100\% \approx 67\%$.

[26] To test the impact of an explanatory variable defined at the market×month level, we naturally rely on variation at the same level. Since θ_t is collinear with $\text{NeutralInvestorBase}_t$, it is left out of this regression. Similar to $\text{CoethnicInvestorBase}_{jt}$, we measure $\text{NeutralInvestorBase}_t$ as the portfolio value of neutral investors that are active—that is, that trade—at time t, relative to that of all potentially active investors. We now define potentially active investors as all individual, Kenyan investors and neutral investors who have invested on the NSE up to and including the month in question.

serve.[27] This finding is consistent with Proposition 4 of the model in Section 5 and illustrates how coethnic investing "misallocates demand" across firms.

In Panel B of Table 4 we estimate the coethnic investor base effect in an alternative way. Thirteen of the 47 firms in our sample "change ethnicity" during our data period. We code changes in a firm's investor base as 0/1 up-or-down events resulting from CEO (ethnicity) turnover, and restrict attention to 12 month windows around such events in the spirit of an event study analysis. In this way we test Proposition 2, which considers a particular supply-side response—that a firm can benefit from changing its ethnicity to that of a larger investor group. We run:

$$\text{PriceToBook}_{jt} = \alpha + \beta \text{I(CEO switched} \rightarrow \Delta\text{CoethnicInvestorBase)}_{jt} + \delta_j + \theta_t + X_{jt} + \varepsilon_{jt} \tag{5}$$

Here, $\text{I(CEO switched} \rightarrow \Delta\text{CoethnicInvestorBase)}_{jt}$ is an indicator for firm j increasing or a decreasing its coethnic investor base by changing its CEO from an individual belonging to one ethnicity to someone else belonging to another ethnicity. The indicator equals one in any month t after the switch.

A firm that changes its ethnicity to one with a larger investor base sees a significant and large—33.2%—increase, while a firm changing its ethnicity to one with a smaller investor base sees a corresponding—albeit proportionally smaller, at around 20%—decrease, in price-to-book value. These results are in Panel B of Table 4. They exploit abrupt and large changes in firms' coethnic investor base coming from the supply (firm) side of the market, and thus support the demand (investor) side evidence in Panel A.

The evidence in this section points towards three conclusions. The first is that the *relative* funds of potential investors of the same ethnic group as a given firm influence the value of large firms in Kenya, as the model in Section 5 predicts when investors have a taste for or psychological bias towards coethnic firms. The second is that counteracting demand- and supply-side responses affects firms' value as theory predicts. Finally, the first take-away holds despite of the second one: the magnitude of counteracting market responses isn't large enough to offset the

[27] From Column 3 of Panel A in Table 4, we find that the expected percent change in price-to-book value for the groups with the smallest and largest coethnic investor bases are, respectively, $0.91 + 0.00001 \times (-20.66)$ and $0.91 + 0.38 \times (-20.66)$. Taking the difference and multiplying it by 0.046 (since we are doubling the neutral investor share) yields approximately 36%.

impact of investor favoritism on market-wide value creation.

6.2 The cost of ethnic investing

We can now estimate the overall cost of coethnic investing. We focus on counterfactual demand-side scenarios. The proofs of propositions 3 and 4 characterize how an increase in the share of neutral investors and corresponding decrease in ethnically biased investors—as might occur over time—will affect market-wide value creation. We consider two counterfactual scenarios.

We first assign all firms a neutral investor base as large as the largest one we observe, corresponding to the the month when foreign and institutional investors made up the largest share of all potentially active investors. In the second scenario, we increase the share of neutral investors in the market to half, or in other words, set $\text{NeutralInvestorBase}_t = 0.5$ and $\text{CoethnicInvestorBase}_{jt} = 0.5$ for all firms. We calculate expected changes in firms' valuation using the estimated coefficients on $\text{NeutralInvestorBase}_t$ and $\text{NeutralInvestorBase}_t \times \text{CoethnicInvestorBase}_{jt}$ in Panel A of Table 4, and firms' information (book value and outstanding shares) at the end of our data period.

Suppose that the estimated coefficient on firms' neutral investor base in Column 2 of Panel A of Table 4 is $\hat{\gamma}$. $(\text{NeutralInvestorBase}^C_{jt} - \text{NeutralInvestorBase}_{jt}) \times \hat{\gamma}$ is thus the difference between the firm's price-to-book value under the counterfactual less-coethnic-investing scenario and the observed state of the world. The decrease in the expected value of a firm due to coethnic investing is then $(\text{NeutralInvestorBase}^C_{jt} - \text{NeutralInvestorBase}_{jt}) \times \hat{\gamma} \times \text{BookValue}_{jt} \times \text{TotalShares}_{jt}$. Computing this quantity for the last month observed in our data suggests that listed Kenyan firm could collectively be worth USD 6.02 billion or 42% more if the proportion of neutral investors in the market was as high as the maximum in our data period. If instead we use the estimates from Column 3 and thus account for differential effects of neutral investors on firms with larger and smaller coethnic investor bases, this counterfactual scenario is predicted to increase the value of listed Kenyan firms by USD 8.96 billion or 63%.

In the other counterfactual scenario, in which half of investors are neutral, listed Kenyan firms are predicted to jointly be worth USD 5.35–7.90 billion or 38–55%

more.

Both scenarios are far out-of-sample compared to the share of neutral investors in almost all months during our data period. The counterfactual calculations we present therefore rely on substantial extrapolation of the linearly estimated effect of neutral investors in Panel A of Table 4. They nevertheless underscore the massive market-wide value loss that likely results from coethnic investing in Kenya.

7 Conclusion

We use transaction data from Kenya's stock exchange and CEO/board turnover to document the surprising extent of "excess" investing in coethnic investment objects in a large, anonymous type of market associated with efficient capital allocation. Coethnic investments earn lower returns, pointing towards a taste-based or psychological explanation. Taking advantage of the complete market nature of a stock exchange and variation over time in firms' coethnic investor bases and neutral investor activity, we show that while both demand-side and supply-side market responses counteract ethnic investing, they far from offset the impact on total stock market value creation. In our model, firm fundamentals (e.g., share numbers and cash flow correlations) do not explicitly depend on the parameters that capture ethnic bias. Thus, the model and corresponding empirical results illustrate how bias distorts prices away from the fundamentals-based benchmark (complete investor neutrality). Such distortions plausibly in turn influence fundamentals—e.g. because ethnic-majority firms gain better capital access and can more easily invest and recruit good workers (Banerjee & Munshi, 2004; Hales *et al.*, 2015), or because stock market wealth affects consumer spending which in turn affects local economic activity and employment (Chodorow-Reich *et al.*, 2021)—but modeling and testing for such "downstream" impacts on fundamental value is beyond the scope of this paper.

References

ANAGOL, S, & KIM, H. 2012. The Impact of Shrouded Fees: Evidence from a Natural Experiment in the Indian Mutual Fund Market. *American Economic Review*.

ANAGOL, SANTOSH, BALASUBRAMANIAM, VIMAL, & RAMADORAI, TARUN. 2018. Endowment Effects in the Field: Evidence from India's IPO Lotteries. *Review of Economic Studies*.

ANAGOL, SANTOSH, BALASUBRAMANIAM, VIMAL, & RAMADORAI, TARUN. 2021. Learning from noise: Evidence from Indiaâs IPO lotteries. *Journal of Financial Economics*, **140**(3), 965–986.

ARDALAN, K. 2019. Equity Home Bias: A Review Essay. *Journal of Economic Surveys*, **33**, 949–967.

ARROW, KENNETH. 1973. The Theory of Discrimination. *In:* ASHENFELTER, ORLEY, & REES, ALBERT (eds), *Discrimination in Labor Markets*. Princeton University Press.

BANDIERA, O., BARANKAY, IWAN, & RASUL, IMRAN. 2009. Social Connections and Incentives in te Workplace: Evidence from Personnel Data. *Econometrica*.

BANERJEE, ABHIJIT, & DUFLO, ESTHER. 2005. Growth Theory Through the Lens of Development Economics. *In:* DURLAUF, STEVE, & AGHION, PHILIPPE (eds), *Handbook of Economic Growth: 2005, Vol. 1A*. Elsevier Science Ltd, North Holland.

BANERJEE, ABHIJIT, & MUNSHI, KAIVAN. 2004. How efficiently is capital allocated? Evidence from the knitted garment industry in Tirupur. *Review of Economic Studies*.

BARBER, BRAD M.B, MORSE, ADAIR, & YASUDA, AYAKO. 2021. Impact investing. *Journal of Financial Economics*, **139**, 162–185.

BARKAN, JOEL D. 2004. Kenya after Moi. *Foreign Affairs*.

BASAK, SULEYMAN, & PAVLOVA, ANNA. 2013. Asset Prices and Institutional Investors. *American Economic Review*.

BEACH, BRIAN, & JONES, DANIEL B. 2017. Gridlock: Ethnic Diversity in Government and the Provision of Public Goods. *American Economic Journal: Economic Policy*.

BECKER, GARY S. 1957. *The Economics of Discrimination*. Chicago: University of Chicago.

BERGE, LARS IVAR OPPEDAL, BJORVATN, KJETIL, GALLE, SIMON, POSNER, EDWARD MIGUEL DANIEL N., TUNGODDEN, BERTIL, & ZHANG, KELLY. 2020. Ethnically Biased? Experimental Evidence from Kenya. *Journal of the European Economic Association*.

BERTRAND, M., & DUFLO, E. 2017. Field Experiments on Discrimination. *In:* BANERJEE, A., & DUFLO, E. (eds), *Handbook of Economic Field Experiments*. Elsevier.

BERTRAND, MARIANNE, & SCHOAR, ANTOINETTE. 2003. Managing with style: The effect of managers on firm policies. *Quarterly Journal of Economics*.

BOEHMER, EKKEHART, & KELLEY, ERIC K. 2009. Institutional Investors and the Informational Efficiency of Prices. *Review of Financial Studies*.

BOHREN, J AISLINN, IMAS, ALEX, & ROSENBERG, MICHAEL. 2019. The dynamics of discrimination: Theory and evidence. *American economic review*, **109**(10), 3395–3436.

BOHREN, J AISLINN, HULL, PETER, & IMAS, ALEX. 2024. *Systemic discrimination: Theory and measurement*. Tech. rept. National Bureau of Economic Research.

BURCHARDI, KONRAD, CHANEY, THOMAS, & HASSAN, TAREK A. 2019. Migrants, Ancestors, and Foreign Investments. *Review of Economic Studies*.

BURGESS, ROBIN, JEDWAB, REMI, MIGUEL, EDWARD, MORJARIA, AMEET, & I MIQUEL, GERARD PADRO. 2015. The Value of Democracy: Evidence from Road Building in Kenya. *American Economic Review*.

CAMPBELL, JOHN Y., RAMADORAI, TARUN, & SCHWARTZ, ALLIE. 2009. Caught on Tape: Institutional Trading, Stock Returns, and Earnings Announcements. *Journal of Financial Economics*.

CHARLES, KERWIN, & GURYAN, JONATHAN. 2008. Prejudice and Wages: An Empirical Assessment of Becker's The Economics of Discrimination. *Journal of Political Economy*.

CHARLES, KERWIN, & GURYAN, JONATHAN. 2013. Taste-Based or Statistical Discrimination: The Economics of Discrimination Returns to its Roots. *Economic Journal*.

CHODOROW-REICH, GABRIEL, NENOV, PLAMEN T, & SIMSEK, ALP. 2021. Stock market wealth and the real economy: A local labor market approach. *American Economic Review*, **111**(5), 1613–1657.

CHORDIA, TARUN, ROLL, RICHARD, & SUBRAHMANYAM, AVANIDHAR. 2002. Order imbalance, liquidity, and market returns. *Journal of Financial Economics*.

COEURDACIER, NICOLAS, & REY, H'EL'ENE. 2013. Home Bias in Open Economy Financial Macroeconomics. *Journal of Economic Literature*, **51**, 63–115.

COHEN, LAUREN, FRAZZINI, ANDREA, & MALLOY, CHRISTOPHER. 2008. The Small World of Investing: Board Connections and Mutual Fund Returns. *Journal of Political Economy*.

COOPER, IAN, SERCU, PIET, & VANP'EE, ROSANNE. 2013. The Equity Home Bias Puzzle: A Survey. *Foundations and Trends in Finance*, **7**, 289–416.

CORNELL, BRADFORD, & WELCH, IVO. 1996. Culture, Information, and Screening Discrimination. *Journal of Political Economy*.

DO, QUOC-ANH, GALBIATI, ROBERTO, MARX, BENJAMIN, & SERRANO, MIGUEL A ORTIZ. 2024. J'Accuse! Antisemitism and financial markets in the time of the Dreyfus Affair. *Journal of Financial Economics*, **154**, 103809.

EDELEN, ROGER M., INCEB, OZGUR S., & KADLECC, GREGORY B. 2016. Institutional Investors and Stock Return Anomalies. *Journal of Financial Economics*.

FAFCHAMPS, MARCEL. 2000. Ethnicity and credit in African manufacturing. *Journal of Development Economics*.

FISMAN, RAYMOND, PARAVISINI, DANIEL, & VIG, VIKRANT. 2017. Social Proximity and Loan Outcomes: Evidence from an Indian Bank. *American Economic Review*.

FISMAN, RAYMOND, SARKAR, ARKODIPTA, SKRASTINS, JANIS, & VIG, VIKRANT. 2020. Experience of Communal Conflicts and Inter-group Lending. *Journal of Political Economy*.

GABAIX, XAVIER, GOPIKRISHNAN, PARAMESWARAN, PLEROU, VASILIKI, & STANLEY, H. EUGENE. 2006. Institutional Investors and Stock Market Volatility. *Quarterly Journal of Economics*.

GREIF, AVNER. 1993. Contract enforceability and economic institutions in early trade: The Maghribi traders' coalition. *American Economic Review*.

GROMB, DENIS, & VAYANOS, DIMITRI. 2010. Limits of Arbitrage. *Annual Review of Financial Economics*.

GUISO, LUIGI, SAPIENZA, PAOLA, & ZINGALES, LUIGI. 2009. Cultural Biases in Economic Exchange? *Quarterly Journal of Economics*.

HALES, JEFFREY, WANG, LAURA W, & WILLIAMSON, MICHAEL G. 2015. Selection benefits of stock-based compensation for the rank-and-file. *The Accounting Review*, **90**(4), 1497–1516.

HJORT, JONAS. 2014. Ethnic Divisions and Production in Firms. *Quarterly Journal of Economics*.

HJORT, JONAS, MALMBERG, HANNES, & SCHOELLMAN, TODD. 2024. *The Missing Middle Managers: Labor Costs, Firm Structure, and Development*. NBER Working Papers 30592, National Bureau of Economic Research, Inc.

HONG, HARRISON, & KACPERCZYK, MARCIN. 2009. The Price of Sin: The Effects of Social Norms on Markets. *Journal of Financial Economics*.

HSIEH, CHANG-TAI, HURST, ERIK, JONES, CHAD, & KLENOW, PETER. 2019. The Allocation of Talent and U.S Economic Growth. *Econometrica*.

HVIDE, HANS, & DØSKELAND, TROND. 2011. Do individual investors have asymmetric information based on work experience? *Journal of Finance*.

KUMAR, ALOK, NIESSEN-RUENZI, ALEXANDRA, & SPALT, OLIVER G. 2015. What's in a Name? Mutual Fund Flows When Managers Have Foreign-Sounding Names. *The Review of Financial Studies*, **28**, 2281–2321.

LANG, KEVIN. 1986. A Language Theory of Discrimination. *Quarterly Journal of Economics*.

LEWIS, KAREN K. 1999. Trying to Explain Home Bias in Equities and Consumption. *Journal of Economic Literature*, **37**, 571–608.

LIEBERSON, STANLEY. 1969. Measuring population diversity. *American Sociological Review*, 850–862.

LIST, JOHN A., & RASUL, IMRAN. 2011. Field Experiments in Labor Economics. *In:* ASHENFELTER, ORLEY, & CARD, DAVID (eds), *Handbook of Labor Economics Vol. 4a*. Elsevier.

MERTON, ROBERT C. 1987. Presidential address: A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*, **42**(3), 483–510.

NDEGWA, S. 1997. Citizenship and Ethnicity: An Examination of Two Transition Moments in Kenyan Politics. *American Political Science Review*.

PHELPS, EDMUND S. 1972. The statistical theory of racism and sexism. *American Economic Review*.

RAUCH, JAMES E. 2001. Business and Social Networks in International Trade. *Journal of Economic Literature*.

RAUCH, JAMES E., & TRINDADE, VITOR. 2002. Ethnic Chinese Networks in International Trade. *Review of Economics and Statistics*.

ROSE, EVAN K. 2023. A constructivist perspective on empirical discrimination research. *Journal of Economic Literature*, **61**(3), 906–923.

SHLEIFER, A., & VISHNY, R.W. 1997. The limits of arbitrage. *Journal of Finance*.

TEOH, SIEW HONG, WELCH, IVO, & WAZZAN, C. PAUL. 1999. The Effect of Socially Activist Investment Policies on the Financial Markets: Evidence from the South African Boycott. *The Journal of Business*, **72**, 35–89.

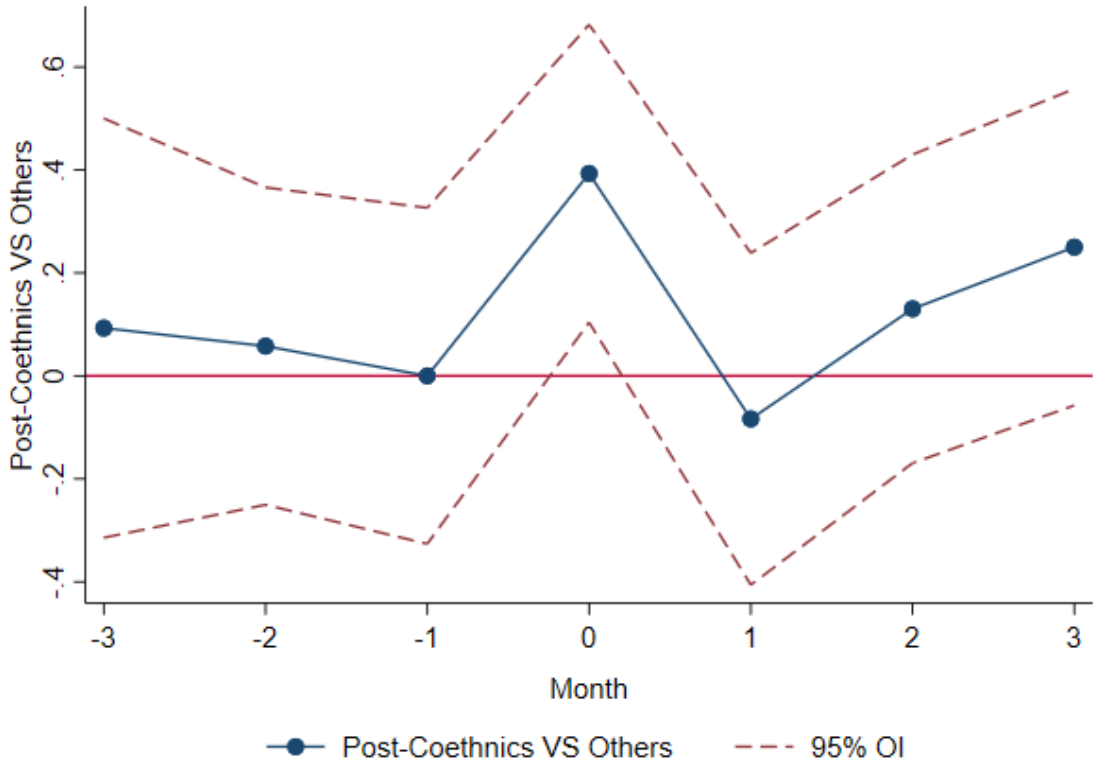
VAN NIEUWERBURGH, STIJN, & VELDKAMP, LAURA. 2009. Information Immobility and the Home Bias Puzzle. *Journal of Finance*.

YENKEY, CHRISTOPHER. 2018a. Fraud and Market Participation: Social Relations as a Moderator of Organizational Misconduct. *Administrative Science Quarterly*.

YENKEY, CHRISTOPHER. 2018b. The Outsider's Advantage: Distrust as a Deterrent to Exploitation. *American Journal of Sociology*.

YENKEY, CHRISTOPHER B. 2015. Mobilizing a Market: Ethnic Segmentation and Investor Recruitment into Kenya's Nascent Stock Market. *Administrative Science Quarterly*.

FIGURE 1: INVESTMENT FLOWS FROM "POST-COETHNICS" VS OTHERS WHEN A FIRM "CHANGES ETHNICITY" DUE TO CEO TURNOVER



We regress the monthly OI between post-coethnics and others. Post-coethnics mean the investor and the firm are coethnic after the firm switches CEO. Others mean the investor and the firm aren't coethnic both before and after the firm switches CEO. The sample uses only those firms where the ethnicity of the CEO changes at least once, and we delete the pre-coethnics sample. The change occurs at month 0. This figure is consistent with Figure A1.

TABLE 1: INVESTOR-FIRM COETHNICITY AND INVESTMENT

	(1) Investment	(2) Investment	(3) Investment	(4) Investment
CoethnicCEO	0.00979***			
	(0.00349)			
CEOCoethnicityIndex		0.0110***		
		(0.00392)		
CoethnicBoard			0.0197***	
			(0.00438)	
BoardCoethnicityIndex				0.0461***
				(0.00991)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.547	0.546	0.560	0.543
R2	0.399	0.393	0.431	0.390
N	273466	399457	187355	429519
	(1) OI	(2) OI	(3) OI	(4) OI
CoethnicCEO	0.00881*			
	(0.00467)			
CEOCoethnicityIndex		0.0128**		
		(0.00531)		
CoethnicBoard			0.0708***	
			(0.00996)	
BoardCoethnicityIndex				0.117***
				(0.0185)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.0805	0.0731	0.101	0.0700
R2	0.331	0.325	0.344	0.317
N	409290	602420	280488	648131

The specification is estimated on investor-firm-month-level data. The sample consists of all months in which a trade is made by any investor in any firms stock. Panel A shows the outcome investment, which is the proportion of the investor's portfolio that is held in the share. Panel B shows order imbalance, which measures how much the investor net buys or sells a particular firm's stock, as a proportion of the investor's total traded stock of the same stock during the same month. All specifications in both panels include investor, firm, month, and CEO ethnicity fixed effects and we control for the value control return on equity (ROE) in the prior 12 month period. Standard errors are clustered at the investor ethnicity $\times$ CEO ethnicity level. The dataset spans January 2006-December 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE 2: INVESTOR-FIRM COETHNICITY AND INVESTMENT WITHIN INVESTOR-FIRM PAIRS

	(1) Investment	(2) Investment	(3) Investment	(4) Investment
CoethnicCEO	0.0123			
	(0.0200)			
CEOCoethnicityIndex		0.0265		
		(0.0228)		
CoethnicBoard			0.0622***	
			(0.0125)	
BoardCoethnicityIndex				0.231***
				(0.0467)
Value Controls	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.533	0.529	0.546	0.525
R2	0.606	0.606	0.629	0.606
N	204928	295741	134914	316152
	(1) OI	(2) OI	(3) OI	(4) OI
CoethnicCEO	-0.0353			
	(0.0419)			
CEOCoethnicityIndex		0.0384		
		(0.0457)		
CoethnicBoard			0.157***	
			(0.0280)	
BoardCoethnicityIndex				0.647***
				(0.0931)
Value Controls	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.125	0.115	0.140	0.112
R2	0.444	0.445	0.466	0.441
N	306914	449148	201232	481154

The specification is estimated on pair-month-level data. Pair is defined as a unique investor-firm grouping. The sample consists of all months in which a trade is made by any investor in any firms stock. Panel A shows the outcome investment, which is the proportion of the investors' portfolio that is held in the share. Panel B shows order imbalance, which measures how much the investor net buys or sells a particular firm's stock, as a proportion of the investor's total traded stock of the same stock during the same month. All specifications in both panels include pair, month, and CEO ethnicity fixed effects and we control for the value control return on equity (ROE) in the prior 12 month period. Standard errors are clustered at the investor ethnicity $\times$ CEO ethnicity level. The dataset spans January 2006-December 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE 3: INVESTOR-FIRM COETHNICITY AND RETURNS

	(1) Risk-adjusted Returns	(2) Risk-adjusted Returns	(3) Risk-adjusted Returns	(4) Risk-adjusted Returns
CoethnicCEO	-0.0112**			
	(0.00447)			
CEOCoethnicityIndex		-0.0195***		
		(0.00561)		
CoethnicBoard			-0.0592***	
			(0.00775)	
BoardCoethnicityIndex				0.0110
				(0.0184)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.117	0.0893	0.132	0.0726
R2	0.583	0.568	0.638	0.550
N	216531	318345	150091	342730
	(1) Risk-adjusted Returns	(2) Risk-adjusted Returns	(3) Risk-adjusted Returns	(4) Risk-adjusted Returns
CoethnicCEO	0.00930			
	(0.0180)			
CEOCoethnicityIndex		-0.0875*		
		(0.0491)		
CoethnicBoard			-0.129***	
			(0.0291)	
BoardCoethnicityIndex				-0.104
				(0.104)
Value Controls	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.162	0.137	0.189	0.120
R2	0.755	0.751	0.787	0.745
N	137215	196784	92344	209102

The specifications are estimated on investor-firm-month-transaction level data. Risk-adjusted returns is the Sharpe Ratio, which is defined as the difference between the risk unadjusted returns and the treasury bill rates in Kenya, divided by the standard deviation of the difference. The sample consists of all transactions initiated during the period. The month indicates origination of the transaction. Specifications in Panel A include investor, firm, month, and CEO ethnicity fixed effects while specifications in Panel B include pair, month, and CEO ethnicity fixed effects. We control for the value control return on equity (ROE) in the prior 12 month period in both panels. Standard errors are clustered at the investor ethnicity $\times$ CEO ethnicity level. The dataset spans January 2006-December 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE 4: AGGREGATE CONSEQUENCES OF COETHNIC INVESTING

	(1)	(2)	(3)
	Log Price-to-book	Log Price-to-book	Log Price-to-book
Coethnic Investor Base	1.770**	2.303**	11.17***
	(0.732)	(1.101)	(2.187)
Neutral Investor Base		0.657**	0.910***
		(0.306)	(0.255)
Coethnic Investor Base × Neutral Investor Base			-20.66***
			(3.634)
Value Controls	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Month FE	Yes	No	No
CEO ethnicity switch	No	No	No
Mean of Dep. Var.	0.853	0.853	0.853
R2	0.883	0.729	0.747
N	1828	1828	1828

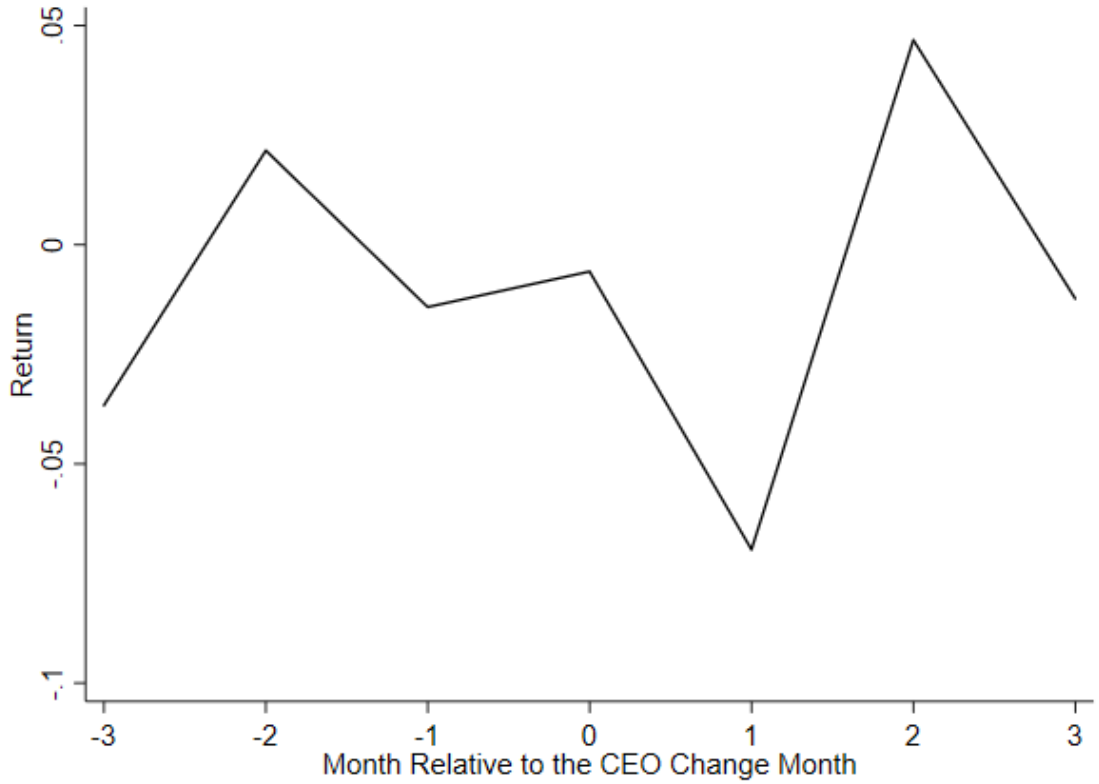
	CEO switch → Investor base ↑	CEO switch → Investor base ↓
	(1)	(2)
	Log Price-to-book	Log Price-to-book
I(CEO switched → Δ CoethnicInvestorBase)	0.332**	-0.206***
	(0.125)	(0.0751)
Value Controls	Yes	Yes
Month FE	Yes	Yes
Firm FE	Yes	Yes
Mean of Dep. Var.	0.966	0.802
R2	0.817	0.848
N	1655	2319

(1) Top panel: Column 1 includes only Biased Investor Base Value, which refers to the aggregate value traded by those coethnic investors in the month as a proportion of total value traded in the same month. Column 2 includes both Biased Investor Base Value and Neutral Investor Base Value, and the latter refers to the aggregate value traded by those neutral investors in the month as a proportion of total value traded in the same month. Column 3 adds the interaction of the two variables. The specifications are estimated on firm-month level data. The dataset spans January 2006-December 2010 and covers only those firms listed on the NSE where the ethnicity of the CEOs remained constant throughout the period. All specifications include firm and month fixed effects and we control for the value control return on equity (ROE) in the prior 12 month period. Standard errors are clustered at the firm level.
(2) Bottom panel: The specifications are estimated on firm-month level data. All specifications include firm and month fixed effects. Switched CEO is an indicator equal to 1 if the ethnicity of the firm CEOs change during the period. Investor base size has the same definition as in the top panel. Post switch is an indicator equal to one after the change in CEOs. The sample looks at a 12 month window around the switch, 6 months prior and 6 months following. Col (1) limits the sample to those firms in which the new CEO has a higher investor base size than the old CEO, and col (2) limits the sample to those firms in which the new CEO has a lower investor base size than the old CEO. All specifications include firm and month fixed effects and we control for the value control return on equity (ROE) in the prior 12 month period. Standard errors are clustered at the firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Online Appendix

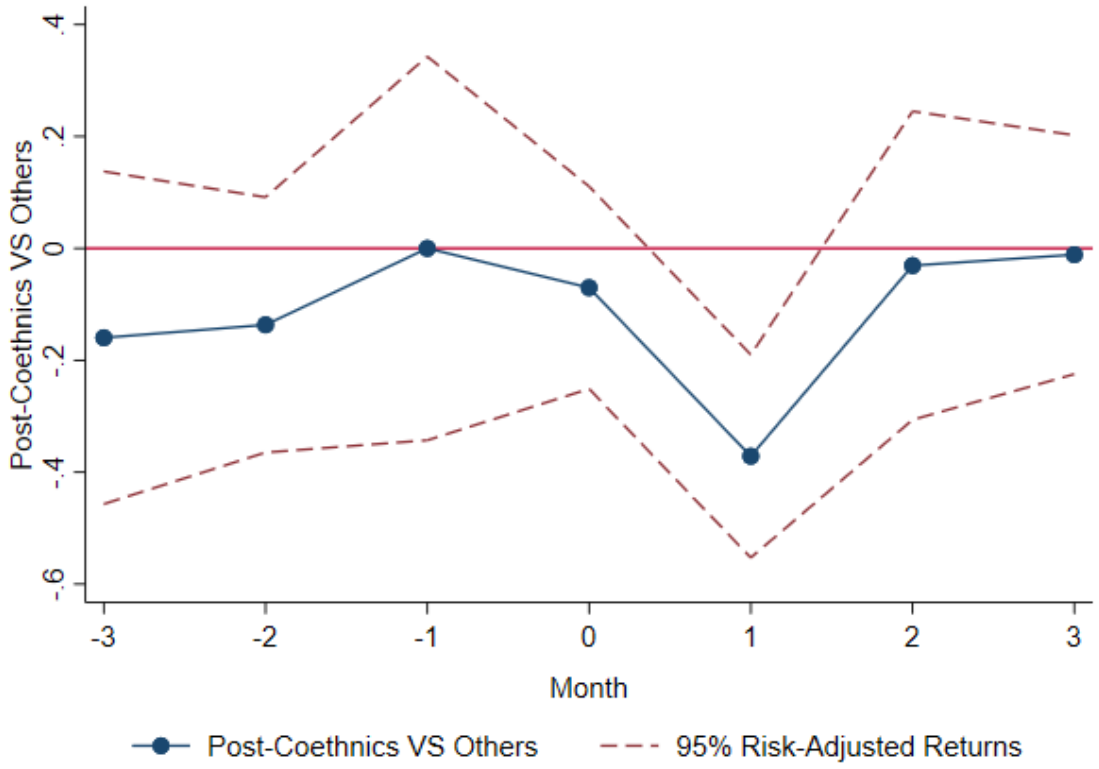
Appendix figures

FIGURE A1: ONE MONTH RETURNS WHEN A FIRM "CHANGES ETHNICITY" DUE TO CEO TURNOVER



The average monthly return over the change month of CEO. The sample uses only those firm where the ethnicity of the CEO changes at least once. The change occurs at month 0.

FIGURE A2: RISK-ADJUSTED RETURNS FROM "POST-COETHNICS" VS OTHERS WHEN A FIRM "CHANGES ETHNICITY" DUE TO CEO TURNOVER



We regress the monthly Risk-adjusted returns between post-coethnics and others. Post-coethnics mean the investor and the firm are coethnic after the firm switches CEO. Others means the investor and the firm aren't coethnic both before and after the firm switches CEO. Risk-adjusted returns correspond to the Sharpe Ratio, which is defined as the difference between the risk unadjusted returns and the treasury bill rates in Kenya, divided by the standard deviation of the difference. The sample uses only those firms where the ethnicity of the CEO changes at least once, and we delete the pre-coethnics sample. The change occurs at month 1.

Appendix Tables

TABLE A1: SUMMARY STATISTICS

Variable	Mean	Std. Dev.
Panel A: Investor level		
N = 54915		
Average portofolio value 2006 (USD)	5999	66832
Average portofolio value 2010 (USD)	4570	47340
Panel B: Firm level		
N = 47		
Listed by 2006	.894	.312
Agricultural	.085	.282
Service	.532	.504
Industrial	.383	.491
Market cap. 2006 (USD 000's)	260599	466847
Market cap. 2010 (USD 000's)	285579	488948
Panel C: Investor $\times$ firm $\times$ month level		
N = 658188		
Investment	.547	.405
Order Imbalance	.069	.985
CoethnicCEO	.271	.445
CoethnicBoard	.406	.491
CEOCoethnicityIndex	.184	.294
BoardCoethnicityIndex	.152	.168
Risk-adjusted Returns	.094	4.706

The dataset spans January 2006-December 2010. The data consists of all investors observed over the period that have made at least five trades (buying or selling) in a given year, as well as 47 firms that were listed on the NSE during some part of the period. These firms include ACCS, BAMB, BAT, BBK, CABL, CMC, DTK, EABL, EQTY, EVRD, HFCK, ICDC, JUB, KCB, KEGN, KENO, KNRE, KPLC, KQ, MSC, NBK, NIC, NMG, OCH, PORT, REA, SCAN, SCBK, SCOM, SGL, TOTL, TPSE, ARM, SASN, FIRE, PAFR, UNGA, BERG, CFC, UCHM, COOP, CandG, MASH, KUKZ, BOC, UTK, CARB. The trades have been aggregated to the investor-firm-month level. For any given investor and firm, only those months where a trade has been made are included. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE A2: INVESTOR-FIRM COETHNICITY AND BUYING STOCKS

	(1) Buy	(2) Buy	(3) Buy	(4) Buy
CoethnicCEO	0.00422*			
	(0.00238)			
CEOCoethnicityIndex		0.00628**		
		(0.00276)		
CoethnicBoard			0.0350***	
			(0.00517)	
BoardCoethnicityIndex				0.0590***
				(0.00945)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.542	0.538	0.553	0.537
R2	0.337	0.331	0.351	0.323
N	395691	583348	271310	627549

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The specification is estimated on investor-firm-month-level data. The sample consists of all months in which a trade is made by any investor in any firms stock. This table shows the outcome buy, which is a dummy variable measuring whether the investor purchases the stock during that month. All specifications include investor, firm, month, and CEO ethnicity fixed effects and we control for the value control return on equity (ROE) in the prior 12 month period in both panels. Standard errors are clustered at the investor ethnicity $\times$ CEO ethnicity level. The dataset spans January 2006-December 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE A3: INVESTOR-FIRM COETHNICITY AND INVESTMENT: ETHNICALLY KENYAN INVESTORS

	(1) Investment	(2) Investment	(3) Investment	(4) Investment
CoethnicCEO	0.0117***			
	(0.00418)			
CEOCoethnicityIndex		0.0123***		
		(0.00473)		
CoethnicBoard			0.0134*	
			(0.00706)	
BoardCoethnicityIndex				0.0393***
				(0.0111)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.548	0.544	0.567	0.542
R2	0.413	0.406	0.450	0.399
N	166462	263388	111470	290821
	(1) OI	(2) OI	(3) OI	(4) OI
CoethnicCEO	0.0100*			
	(0.00535)			
CEOCoethnicityIndex		0.0152**		
		(0.00628)		
CoethnicBoard			0.0290**	
			(0.0133)	
BoardCoethnicityIndex				0.0553***
				(0.0194)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.0734	0.0628	0.102	0.0584
R2	0.349	0.338	0.343	0.328
N	251677	401730	168550	444124

The specification is estimated on investor-firm-month-level data. The sample consists of all months in which a trade is made by any investor in any firms stock. The sample is restricted to ethnically Kenyan investors. Panel A shows the outcome investment, which is the proportion of the investors' portfolio that is held in the share. Panel B shows order imbalance, which measures how much the investor net buys or sells a particular firm's stock, as a proportion of the investor's total traded stock of the same stock during the same month. All specifications in both panels include investor, firm, month, and CEO ethnicity fixed effects and we control for the value control return on equity (ROE) in the prior 12 month period. Standard errors are clustered at the investor ethnicity $\times$ CEO ethnicity level. The dataset spans January 2006-December 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE A4: INVESTOR-FIRM COETHNICITY AND RETURNS: REALIZED RETURN

	(1) Risk-adjusted Returns	(2) Risk-adjusted Returns	(3) Risk-adjusted Returns	(4) Risk-adjusted Returns
CoethnicCEO	-0.0116**			
	(0.00482)			
CEOCoethnicityIndex		-0.0132**		
		(0.00539)		
CoethnicBoard			-0.101***	
			(0.0153)	
BoardCoethnicityIndex				-0.0355
				(0.0224)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.110	0.0906	0.120	0.0860
R2	0.562	0.544	0.605	0.527
N	86720	128777	61070	139721

The specifications are estimated on investor-firm-month-transaction level data. The sample is restricted to those accounts with a realized return who have both buy and sell. Risk-adjusted returns correspond to the Sharpe Ratio, which is defined as the difference between the risk unadjusted returns and the treasury bill rates in Kenya, divided by the standard deviation of the difference. The month indicates origination of the transaction. All specifications include investor, firm, month, and CEO ethnicity fixed effects. We control for the value control return on equity (ROE) in the prior 12 month period in both panels. Standard errors are clustered at the investor ethnicity $\times$ CEO ethnicity level. The dataset spans January 2006-December 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE A5: INVESTOR-FIRM COETHNICITY AND RETURNS: SAMPLE WITH NO CEO ETHNICITY CHANGE

	(1) Risk-adjusted Returns	(2) Risk-adjusted Returns	(3) Risk-adjusted Returns	(4) Risk-adjusted Returns
CoethnicCEO	-0.0158**			
	(0.00626)			
CEOCoethnicityIndex		-0.0176**		
		(0.00759)		
CoethnicBoard			-0.0496***	
			(0.00813)	
BoardCoethnicityIndex				-0.00823
				(0.0230)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.147	0.131	0.160	0.106
R2	0.637	0.618	0.658	0.593
N	171683	243379	133967	267735

The table shows results from the regression, which is estimated on investor-firm-month-transaction-level data. The sample is restricted to those firms for which the (ethnicity of the) CEO did not change during our data period. Risk-adjusted returns are defined as the difference between the return on investment of the transaction and the risk-free return, divided by the risk or standard deviation of the monthly returns over the holding period. The sample consists of all transactions initiated during the period. The month indicates origination of the transaction. All specifications include investor, firm, month of origination, and CEO ethnicity fixed effects and we control for the value control return on equity (ROE) in the prior 12 month period. Standard errors are clustered at the investor ethnicity $\times$ CEO ethnicity level. The dataset spans January 2006-December 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE A6: INVESTOR-FIRM COETHNICITY AND RETURNS: ONE YEAR RETURNS

	(1)	(2)	(3)	(4)
	One Year Return	One Year Return	One Year Return	One Year Return
CoethnicCEO	-0.00396**			
	(0.00159)			
CEOCoethnicityIndex		-0.00409**		
		(0.00183)		
CoethnicBoard			-0.0454***	
			(0.00527)	
BoardCoethnicityIndex				-0.0193**
				(0.00810)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.0315	0.0294	0.0340	0.0274
R2	0.526	0.518	0.553	0.505
N	220803	323944	152852	348844

The specifications are estimated on investor-firm-month-transaction level data. One Year Return is calculated based on transaction price and the price of last day in the first calendar year. The sample consists of all transactions initiated during the period. The month indicates origination of the transaction. Specifications in both Panel A and Panel B include investor, firm, month, and CEO ethnicity fixed effects. We control for the value control return on equity (ROE) in the prior 12 month period in both panels. Standard errors are clustered at the investor ethnicity $\times$ CEO ethnicity level. The dataset spans January 2006-December 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A1 Data and variables

A1.1 Data

We use the following data sources. The NSE's Transactions Registry is recorded by the Central Depository and Settlement Corporation, Ltd. (CDSC), the "back office" that manages the clearing and settlement of NSE transactions. The CDSC also maintains a Registry of NSE Investor Accounts. They gave us access to a de-identified version that contains, in addition to a scrambled id, the investor's gender, residential location (typically a town or city), account creation year, account type (individual/institutional investor/broker), nationality (Kenya/East African Community (Burundi, Rwanda, South Sudan, Tanzania, and Uganda)/"foreign"), and last name. Information on firm characteristics (book value, outstanding shares, etc) comes from the firms' financial reports.

A1.2 Variables definition

To construct a measure of an investor's portfolio, we begin by assuming that all investors have zero holdings as of 2006. We thereafter simply add any observed purchases to investor i's inferred holdings, and subtract any observed sales. What we term **Investment**, or holdings imbalance, ranges from 0 to 1. It measures, at the investor-firm-month level, the value of a particular investor's holdings of a particular stock, as a proportion of the value of the investor's total portfolio.[28]

Order Imbalance ranges from -1 to 1. It measures, at the investor-firm-month level, how much the investor net buys or sells a particular firm's stock, as a proportion of the investor's total traded stock of the same stock during the same month (see e.g. Chordia *et al.*, 2002). Specifically,

$$\text{Order Imbalance} = \frac{(\text{Total value of stocks bought}) - (\text{Total value of stock sold})}{\text{Total volume traded within the month}}$$

In the sample of investors who bought and sold the same stocks during our sample period, we define **Risk Unadjusted Returns** as the realized return based on the buy and sell price during the holding period. In the sample of investors

[28] Recall that the NSE was much less active before 2006: our results are very similar if we instead focus only on investors who opened their NSE account in 2006 or later, in which case we observe investors' full portfolio at every point in time.

who bought but not subsequently sold before the end of our data period, the 31st of December 2010, we compute the **Risk Unadjusted Returns** as unrealized paper returns at the 31st of December 2010.

Sharpe Ratio is defined as the difference between the returns of the investment and the risk-free return, divided by the standard deviation of the difference, which represents the additional amount of return that an investor receives per unit of increase in risk. Specifically,

$$\text{Sharpe Ratio} = \frac{E[R - R_b]}{\sqrt{var[R - R_b]}}$$

where R is the risk unadjusted returns, and R_b is the risk-free return. We use the treasury bill rates in Kenya as the risk-free return here.

CoethnicInvestorBase$_{jt}$ is the portfolio value investors that are active—that is, that trade—at time t and who belong to the same ethnicity as firm j's CEO hold, relative to that of all potentially active coethnic investors. We define potentially active coethnic investors as all investors who are Kenya individual investors and have invested on the NSE up to and including the month in question.

NeutralInvestorBase$_t$ is the portfolio value of neutral investors that are active—that is, that trade—at time t, relative to that of all potentially active investors. We define potentially active investors as all investors who are Kenya individual investors and neutral investors, and have invested on the NSE up to and including the month in question. We proxy for neutral investors with foreign and institutional investors.

Alpha is another risk-adjusted returns we define as abnormal return (alpha) based on standard CAPM. In this specification, the risk-free return is defined as the treasury bill rates in Kenya and the market return is calculated based on the Nairobi Securities Exchange 20 Share Index(NSE20). NSE20 is a major stock market index which tracks the performance of 20 best performing companies listed on the Nairobi Securities Exchange. Then, we estimate β and alpha using the return of each stock, the risk-free return in Kenya, and the market return in Kenya.

A2 Coding ethnicity and coethnicity

We probabilistically assign ethnicities to investors, CEOs, and board-members using their last names. As described in Section 2, the starting point is name×ethnicity match probability information recorded by Yenkey (2015, 2018a,b). The author hired eight Kenyan research assistants (RAs), each of whom reported if they were highly confident that a given name could belong to a given ethnicity or not.[29] For each last name, each RA was asked to assign a 1 to any ethnicity that the RA felt 75 percent confident that the name was likely to belong to, and a 0 otherwise. There is overlap in the names used by some ethnicities so that the RAs could assign a given name to multiple ethnicities. We start by taking the average of the 1's and 0's across all RAs for each name to arrive at a single number for each name n and ethnicity e, p_{en}.

From this information we need to construct measures of whether an individual investor is likely to be of the same ethnic group as a given CEO and board. We say that ethnicity e is name n's *Likely Ethnicity* if $p_{en} \geq 0.4$ and p_{en} is ≤ 0.3 for all other ethnicities.[30] If this it not true for any ethnicity, n does not have a *Likely Ethnicity*.

We construct four measures of an investor's ethnic proximity to a firm's CEO and board respectively.[31] As described in Section 2, the first CEO measure, CoethnicCEO_{ijt}, is an indicator variable equal to 1 if investor i and the CEO running firm j in month t share a Likely Ethnicity, and 0 if not.

The second CEO measure, $\text{CEOCoethnicityIndex}_{ijt}$, is a 0 (minimum proximity) to 1 measure of the expected ethnic proximity between the investor's and the CEO's name, given each person's expected probability of belonging to each ethnicity. More precisely, the index is equal to the inner product of the investor and the CEO's name×ethnicity match probabilities, or 1 minus Lieberson (1969)'s index of population diversity.

The first board measure, $\text{BoardCoethnicityIndex}_{ijt}$, is equal to the proportion of

[29] The ethnicities the RAs were asked about, and that we observe, are Anglo, Embu, Kalenjin, Kamba, Kikuyu, Kisii, Luhya, Luo, Maasai, Meru, Somali, South Asian, and Swahili.

[30] These cut-offs were chosen with the goal of minimizing both type 1 and type 2 errors. We also wish to make use of a high proportion of the sample of investors; for this reason the 0.4 threshold is relatively low and the 0.3 threshold relatively high, given considerable overlap in the names used by some Kenyan ethnic groups. In sub-section B1 of this appendix we show that our results are qualitatively very similar if we vary the thresholds.

[31] There are several potential reasons why board coethnicity may affect investment somewhat more (or less) than CEO coethnicity. It could for example be that changes in which ethnic group dominates a board are less frequent than changes in the identity of the CEO and hence provide a more deeply rooted measure of a firm's perceived identity.

board-members that are coethnic with the investor, where coethnicity is measured as for the CoethnicCEO$_{ijt}$.

The other board measure, CoethnicBoard$_{ijt}$, is a 0/1 variable, and essentially repeats the construction of CoethnicCEO$_{ijt}$ twice, first between individual board-members and the investor, then for the board as a whole vis-a-vis the investor. To set CoethnicBoard$_{ijt} = 1$ in month t, we require, first, each individual board-members to be relatively likely to belong to the same ethnicity and relatively unlikely to belong to a different ethnicity than the investor, or vice versa, and second, for the board as a whole—given the expected individual board-member/investor co-ethnicity/non-coethnicity statuses—to be relatively likely to belong to the same ethnicity as the investor and relatively unlikely to belong to another ethnicity.

A3 Robustness checks

In Appendix Table B1 we show that our results from Section 3 of the paper are qualitatively very similar if we vary the thresholds used to define investors' and managers' ethnicities. The coethnicity variables are defined differently than in Table 1: the cutoffs, both to define individual and board level ethnicity are a high of 0.3 and low of 0.2, compared to 0.4 and 0.3, respectively in the main analysis.

In Appendix Table B2 we restrict our sample to investors who open their stock market accounts during our data period so that we have their full transaction history after the account opening. We find that the results are similar to Table 1. The results imply that lack of transaction history for investors before 2006 will be unlikely to affect our results.

In Appendix Table A4 we show that the results are very similar to those in Table 3 if we restrict our sample to investors who bought and sold during our sample period and study the relationship between coethnicity and realized returns. In Appendix Table A5, we restrict the sample to firms whose CEO ethnicity remains constant during our data period.

To investigate returns over different horizons, in Appendix Table A6, we show the relationship between coethnicity and one year return. One year return is defined based on transaction price and the price of last day in the first calendar year. We show that the results are similar with our main Table 3. We also investigate

very short-run (1-day and 5-day) returns on coethnic investments in Appendix Table B3. We find that the one- and five-day return on coethnic investments is—in terms of point estimates—extremely close to that of non-coethnic investments. The only somewhat larger and statistically significant difference we find is for CoethnicBoard measure, which is lower for coethnic investments.

In Appendix Table B4, we define our risk-adjusted returns as abnormal return (alpha) based on standard CAPM. We estimate β and alpha using the return of each stock, the risk-free return in Kenya, and the market return in Kenya. The risk-free return is defined as the treasury bill rates in Kenya and the market return is calculated based on the Nairobi Securities Exchange 20 Share Index(NSE20). NSE20 is a major stock market index which tracks the performance of 20 best performing companies listed on the Nairobi Securities Exchange. We show that the results are similar to those in Table 3.

We focus on the differential returns individual investors make on coethnic investments on *average*. This is the appropriate basis for investigating the most common motivations underlying Kenyan stock market investors discriminating against non-coethnic firms on average, as we saw in Section 3 that they do. However, it would be surprising if there wasn't considerable heterogeneity in the extent to which investors favor coethnic firms, or their reasons for doing so.

A4 Model notation, details, and proofs of propositions

A4.1 Equilibrium and results details

Let I denote the total number of investors; x_i and x_{ni} denote the number of shares of type i owned by biased and neutral investors, respectively, and p_i the price per share of firm type i. The total outstanding shares of stocks in the market are given by N_i. Firms of a given type have the same production technology, characterized by a normally distributed cash flow with mean μ_i and variance σ_i^2, where $i = 1, 2$.

Given CARA preferences and normally-distributed cash flows, the optimal portfolio choices satisfy the following first order conditions:

$$x_i = \frac{\tau(\mu_i - p_i)}{\sigma_i^2} \tag{6}$$

$$x_{n1} = \frac{\tau[\sigma_2^2(\mu_1 - p_1) - \sigma_{12}(\mu_2 - p_2)]}{\Delta} \tag{7}$$

$$x_{n2} = \frac{\tau[\sigma_1^2(\mu_2 - p_2) - \sigma_{12}(\mu_1 - p_1)]}{\Delta} \tag{8}$$

where $\Delta = \sigma_1^2\sigma_2^2 - \sigma_{12}^2$.

Equilibrium prices are solved by imposing the constraints:

$$\alpha I x_{n1} + (1-\alpha)\beta I x_1 = N_1,$$
$$\alpha I x_{n2} + (1-\alpha)(1-\beta) I x_2 = N_2.$$

which give:

$$p_1 = \mu_1 - \frac{\sigma_1^2[(1-\alpha)(1-\beta)N_1\Delta + \alpha(N_1\sigma_1^2 + N_2\sigma_{12})\sigma_2^2]}{I\tau[\beta(1-\beta)(1-\alpha)^2\Delta + \alpha\sigma_1^2\sigma_2^2]}$$
$$p_2 = \mu_2 - \frac{\sigma_2^2[(1-\alpha)\beta N_2\Delta + \alpha(N_2\sigma_2^2 + N_1\sigma_{12})\sigma_1^2]}{I\tau[\beta(1-\beta)(1-\alpha)^2\Delta + \alpha\sigma_1^2\sigma_2^2]}$$

With only neutral investors ($\alpha = 1$), prices become:

$$p_1^N = \mu_1 - \frac{N_1\sigma_1^2 + \sigma_{12}N_2}{I\tau}, \quad p_2^N = \mu_2 - \frac{N_2\sigma_2^2 + \sigma_{12}N_1}{I\tau}.$$

We now derive the results, assuming the two firm types differ only in ethnicity—their return structures are the same ($\sigma_1 = \sigma_2 = \sigma$ and $\mu_1 = \mu_2 = \mu$). The equilibrium prices simplify to:

$$p_1 = \mu - \frac{\sigma^2[N_1(1-\rho^2)(1-\beta)(1-\alpha) + \alpha(N_1 + N_2\rho)]}{I\tau A}$$
$$p_2 = \mu - \frac{\sigma^2[N_2(1-\rho^2)\beta(1-\alpha) + \alpha(N_1\rho + N_2)]}{I\tau A}$$

where $A = (1-\rho^2)\beta(1-\beta)(1-\alpha)^2 + \alpha$ and ρ denotes the correlation coefficient.

The respective prices with only neutral investors are:

$$p_1^N = \mu - \frac{\sigma^2(N_1 + \rho N_2)}{I\tau}, \quad p_2^N = \mu - \frac{\sigma^2(N_2 + \rho N_1)}{I\tau}.$$

Barring extreme correlation cases, group 1 share prices are generally higher under complete investor neutrality than with both biased and neutral investors when β is small and α is large, as the crowd-out effect dominates. A low cash flow correlation also contributes to this outcome by strengthening the diversification demand effect. A high relative supply of group 1 shares amplifies the supply effect, further favoring the complete neutrality scenario. We formally substantiate these claims in the next subsection.

A4.2 Prices in the only-neutral and mixed scenarios

We now examine conditions under which prices are higher under complete investor neutrality compared to the mixed scenario. First, note that when cash flows are perfectly negatively correlated ($\rho = -1$), prices are equal under both scenarios:

$$p_1^N = p_1 = \mu - \frac{\sigma^2(N_1 - N_2)}{I\tau}, \quad p_2^N = p_2 = \mu - \frac{\sigma^2(N_2 - N_1)}{I\tau}$$

Opposing risks offset naturally and firms are valued symmetrically based on share quantities. Diversification eliminates risk asymmetry, making investor composition irrelevant for equilibrium prices.

Similarly, when cash flows are perfectly positively correlated ($\rho = 1$), both firm types have identical prices:

$$p|_{\rho=1} = \mu - \frac{\sigma^2(N_1 + N_2)}{I\tau}$$

Identical return structures and perfect cash flow correlation lead firms to be perceived as interchangeable, so investor composition has no impact on equilibrium prices.

When $-1 < \rho < 1$, firm group 1's price is higher under complete investor neutrality if:

$$N_1 > \frac{\beta(1-\alpha)\rho}{1-\beta(1-\alpha)} N_2. \tag{9}$$

Firm group 2's price is higher under complete neutrality if:

$$N_2 > \frac{(1-\beta)(1-\alpha)\rho}{1-(1-\beta)(1-\alpha)} N_1. \tag{10}$$

From expressions (9) and (10), the price of shares from firm group $i = 1, 2$ is more likely to be higher under complete neutrality when:

- **The share of biased investors in group i is smaller:** In group 1, the share of biased investors, $\beta(1-\alpha)$, decreases as $\beta \to 0$, making (9) more likely to hold. Similarly, in group 2, $(1-\beta)(1-\alpha)$ decreases as $\beta \to 1$, increasing the likelihood of (10)

  The biased investors demand effect is weak with a low share of biased investors from the corresponding group. Consequently, the crowd-out effect dominates in the scenario with both biased and neutral investors, leading to higher prices under complete investor neutrality.

- **The share of neutral investors is larger:** As $\alpha \to 1$, the share of neutral investors increases, reducing the overall share of biased investors. This makes the right-hand side of inequalities (9) and (10) approach zero, increasing the likelihood that the conditions hold.

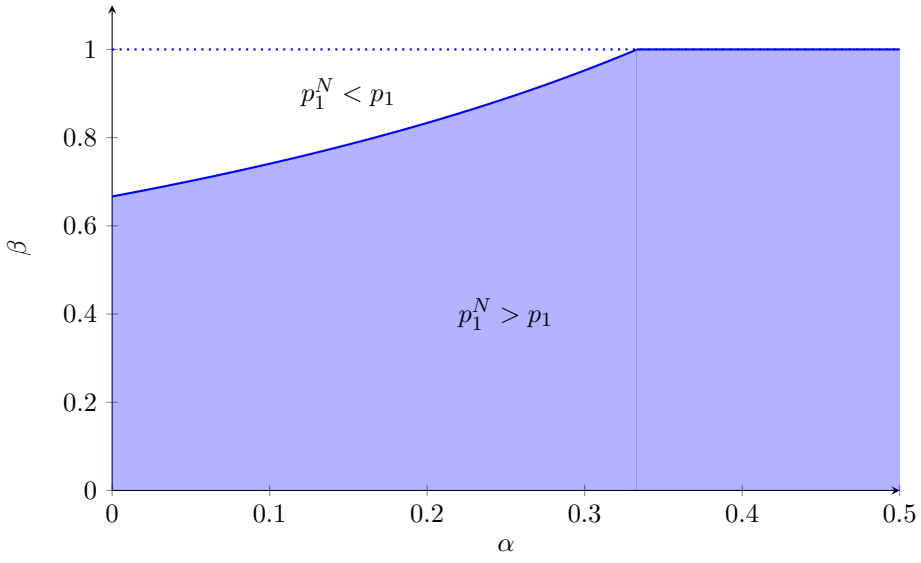
  As the share of neutral investors grows, more investors avoid overpriced shares due to bias (crowd-out effect), potentially increasing demand for the risk-free asset. The reduced influence of biased investors prevents prices in the mixed scenario from exceeding those under complete neutrality.

- **The correlation between cash flows is lower:** when $\rho \leq 0$, (9) holds if $N_1 > 0$, and (10) holds if $N_2 > 0$. For $0 < \rho < 1$, smaller ρ increases the likelihood of both inequalities. Low correlation enables neutral investors to diversify more effectively, increasing the demand and prices of both share types. Biased investors do not exploit these benefits, as they only invest in coethnic firms. Hence, for small values of ρ, prices are more likely to be higher under complete neutrality.

- **The number of shares from group i is large relative to the other group:** When group 1's share supply is large, the supply effect is strong. If this supply exceeds the ratio of biased investors in group 1 to all other investors in the mixed scenario, downward pressure on p_1 is intensified by biased investors' preference for Firm 2 shares or the risk-free asset. As a result, the price of group 1 is more likely to be higher under complete investor neutrality.

Figure A3 compares p_1 and p_1^N across α and β values for $N_1 = N_2$ and $\rho = 0.5$ (substituting into (9)). The blue region indicates $p_1^N > p_1$, while the area above the curve shows $p_1^N < p_1$. For low α, most investors are biased, so a sufficient share in group 1 strengthens the biased investors' demand effect, making prices higher in the mixed scenario when β is large. For $\alpha > 1/3$, neutral investors dominate, and large β cannot make the biased demand effect outweigh the crowd-out effect, resulting in higher prices under complete neutrality.

FIGURE A3: COMPARISON OF p_1 AND p_1^N WHEN $\rho = 0.5$ AND $N_1 = N_2$



A4.3 Proof of Proposition 1

Proof. Taking derivatives of p_1 with respect to β gives:

$$\frac{\sigma^2(1-\rho^2)(1-\alpha)}{I\tau A^2}\{N_1[\alpha+\alpha(1-\alpha)(1-2\beta)+(1-\beta)^2(1-\alpha)^2(1-\rho^2)] \\ +(1-\alpha)(1-2\beta)\rho\alpha N_2\}$$

which is positive if and only if

$$N_1 > \frac{(2\beta-1)\rho\alpha(1-\alpha)N_2}{\alpha+\alpha(1-\alpha)(1-2\beta)+(1-\beta)^2(1-\alpha)^2(1-\rho^2)} \tag{11}$$

□

Inequality (11) holds for a wide range of parameters, including for example when the prices of the two types of stocks are uncorrelated.

The value of ρ that maximizes the right-hand side of (11) (when it is positive) is $\rho = 1$. Assuming $N_1 = N_2 > 0$, $0 < \alpha < 1$, and $\rho = 1$, the inequality holds if and only if

$$\beta < \frac{1}{2} + \frac{1}{4(1-\alpha)}.$$

It follows that as $\alpha \to 0$, β must remain below $3/4$ for p_1 to increase with β.

A4.4 Proof of Proposition 2

Proof. Let ΔN denote the number of shares issued by the firm and suppose $\beta > \frac{1}{2}$. The stock price for the firm before the CEO switch is simply p_2. The stock price after the switch is

$$\tilde{p_1} = \mu - \frac{\sigma^2[(N_1+\Delta N)(1-\rho^2)(1-\beta)(1-\alpha)+\alpha(N_1+\rho N_2+\Delta N(1-\rho))]}{I\tau A}$$

The firm benefits from the switch if and only if $\tilde{p_1} > p_2$, that is

$$N_2 > \frac{(1+\rho)(1-\alpha)(1-\beta)+\alpha}{(1+\rho)(1-\alpha)\beta+\alpha}(N_1+\Delta N) \tag{12}$$

□

Inequality (12) is more likely to hold when N_2, the total outstanding shares of

minority-ethnicity firms, is large compared to $N_1 + \Delta N$, the sum of outstanding shares of majority firms and the switching firm, and when β is large. When N_2 is large relative to N_1 and when β is large, the stock price for type 1 firms tends to be higher than that for type 2 firms before the ethnicity switch. In this case there is greater demand for the stocks of type 1 firms and relatively smaller supply. Moreover, when ΔN is small, the additional supply of stocks of type 1 firms is marginal, so the switch won't reduce the stock price for type 1 firms by much.

With only neutral investors, (12) simplifies to $N_2 > N_1 + \Delta N$. Hence, the condition for a firm to benefit from switching from firm 2 to firm 1 only depends on the relative supply of shares, leaving the correlation term out.

A4.5 Proof of Proposition 3

Proof.

$$TMV = N_1 p_1 + N_2 p_2$$

$$= \mu(N_1 + N_2) - \frac{\sigma^2}{I\tau A}\left[(1-\rho^2)(1-\alpha)[(1-\beta)N_1^2 + \beta N_2^2] + \alpha(N_1^2 + 2\rho N_1 N_2 + N_2^2)\right]$$

Case when $\rho \in \{-1, 1\}$. In this case, total market value is the same under complete investor neutrality (TMV^N) and under neutral and biased investors (TMV).

$$TMV^N = TMV = \mu(N_1 + N_2) - \frac{\sigma^2(N_1 + N_2)^2}{I\tau}, \quad \text{when } \rho = 1.$$

$$TMV^N = TMV = \mu(N_1 + N_2) - \frac{\sigma^2(N_1 - N_2)^2}{I\tau}, \quad \text{when } \rho = -1.$$

When cash flows are perfectly negatively correlated, bias has no effect, as the market offsets opposing risks, valuing both firms based solely on share supply. Conversely, with perfectly positively correlated cash flows, shares from both firms are interchangeable, making market value unaffected by bias.

Case when $-1 < \rho < 1$.

$$\begin{aligned}
\frac{\partial TMV}{\partial \alpha} =& \frac{\sigma^2(1-\rho^2)}{I\tau A^2}\{[(1-\beta)N_1^2 + \beta N_2^2][1-(1-\rho^2)\beta(1-\beta)(1-\alpha)^2] \\
& - (N_1^2 + 2N_1N_2\rho + N_2^2)\beta(1-\beta)(1-\alpha^2)\} \\
=& \frac{\sigma^2(1-\rho^2)}{I\tau A^2} M \\
\frac{\partial M}{\partial \alpha} =& 2[(1-\beta)N_1^2 + \beta N_2^2](1-\rho^2)\beta(1-\beta)(1-\alpha) \\
& + 2\alpha(N_1^2 + 2N_1N_2\rho + N_2^2)\beta(1-\beta) \\
\geq& 0
\end{aligned}$$

To prove $M \geq 0$, it suffices to show $M \geq 0$ when $\alpha = 0$.

$$\begin{aligned}
M|_{\alpha=0} =& (1-\beta)^2[1-(1-\rho^2)\beta](N_1 - \frac{N_2\rho\beta}{(1-\beta)[1-(1-\rho^2)\beta]})^2 \\
& + \frac{N_2^2\beta^3(1-\beta)(1-\rho^2)^2}{1-(1-\rho^2)\beta} \\
\geq& 0
\end{aligned}$$

□

We now provide a formal proof for $TMV^N > TMV$ when $-1 < \rho < 1$. First, note that in this case:

$$TMV^N = \mu(N_1 + N_2) - \frac{\sigma^2}{I\tau}\left(N_1^2 + 2\rho N_1 N_2 + N_2^2\right).$$

For $TMV^N > TMV$, we require:

$$\frac{\sigma^2}{I\tau[(1-\rho^2)\beta(1-\beta)(1-\alpha)^2 + \alpha]}\left[(1-\rho^2)(1-\alpha)[(1-\beta)N_1^2 + \beta N_2^2] + \alpha(N_1^2 + 2\rho N_1N_2 + N_2^2)\right] >$$

$$\frac{\sigma^2}{I\tau}\left(N_1^2 + 2\rho N_1 N_2 + N_2^2\right)$$

$$(1-\beta)N_1^2 + \beta N_2^2 > \beta(1-\beta)(1-\alpha)\left(N_1^2 + 2\rho N_1 N_2 + N_2^2\right)$$

To demonstrate that this last inequality always holds for $-1 < \rho < 1$ and $0 \leq \alpha < 1$, note that the value of α that maximizes the right-hand side is $\alpha = 0$. For

any other value $0 < \alpha < 1$, the right-hand side decreases. Therefore, we will first show that the left-hand side exceeds the right-hand side when $\alpha = 0$; it will then follow that the inequality remains true for all other values of α.

If $\alpha = 0$, the inequality is:

$$(1-\beta)N_1^2 + \beta N_2^2 > \beta(1-\beta)\left(N_1^2 + 2\rho N_1 N_2 + N_2^2\right).$$

$$(1-\beta)^2 + \beta^2\left(\frac{N_2}{N_1}\right)^2 - 2\rho\beta(1-\beta)\left(\frac{N_2}{N_1}\right) > 0.$$

Let $x \equiv \frac{N_2}{N_1}$, then:

$$f(x) \equiv (1-\beta)^2 + \beta^2 x^2 - 2\rho\beta(1-\beta)x > 0.$$

The expression above is a quadratic function in x of the form $ax^2 + bx + c$, where $a = \beta^2 \geq 0$. If $\beta = 0$, then $f(x) = 1$ and the inequality holds trivially. If $\beta = 1$, $f(x) = (N_2/N_1)^2 > 0$ and the inequality also holds trivially.

Now suppose $0 < \beta < 1$ and let $D_f(x)$ be the discriminant of $f(x)$. Then $D_f(x) = (-2\rho\beta(1-\beta))^2 - 4\beta^2(1-\beta)^2$. Since $\beta > 0$, $a > 0$. Hence, if we can show that the discriminant is negative, it follows that $f(x) > 0$. Expanding $D_f(x)$:

$$D_f(x) = 4\rho^2\beta^2(1-\beta)^2 - 4\beta^2(1-\beta)^2$$

$$D_f(x) = 4\beta^2(1-\beta)^2(\rho^2 - 1)$$

Since $\rho \in (-1, 1)$, $\rho^2 < 1$, and hence $D_f(x) < 0$. Thus, we conclude that $TMV^N > TMV$ when $\alpha = 0$. For $0 < \alpha < 1$, the right hand side is smaller than when $\alpha = 0$, and hence the inequality also holds. Thus, $TMV^N > TMV$ holds for $-1 < \rho < 1$.

A4.6 Proof of Proposition 4

Proof.

$$\begin{aligned}\frac{\partial P_1}{\partial \alpha} =& \frac{\sigma^2}{I\tau A^2}\{N_1(1-\rho^2)(1-\beta)[1-(1-\rho^2)\beta(1-\beta)(1-\alpha)^2]- \\ & (N_1+N_2\rho)(1-\rho^2)\beta(1-\beta)(1-\alpha^2)\} \\ \frac{\partial P_2}{\partial \alpha} =& \frac{\sigma^2}{I\tau A^2}\{N_2(1-\rho^2)\beta[1-(1-\rho^2)\beta(1-\beta)(1-\alpha)^2]- \\ & (N_1\rho+N_2)(1-\rho^2)\beta(1-\beta)(1-\alpha^2)\}\end{aligned}$$

$\frac{\partial P_1}{\partial \alpha} > \frac{\partial P_2}{\partial \alpha}$ if and only if the following inequality holds:

$$\begin{aligned}& N_1(1-\beta)[1-(1-\rho^2)\beta(1-\beta)(1-\alpha)^2-\beta(1-\rho)(1-\alpha^2)] \\ >& N_2\beta[1-(1-\rho^2)\beta(1-\beta)(1-\alpha)^2-(1-\beta)(1-\rho)(1-\alpha^2)]\end{aligned}$$

If $N_1 = N_2$ the condition can be simplified to $\beta < \frac{1}{2}$ □

A5 Alternative Parametrization

The model and propositions above correspond most directly to a situation in which investors' bias is is observable. But even in the case where only ethnicity is observable, the main results of our model still hold. To see this, it's more convenient to reparametrize the model in the following way.

As before, let I denote the total number of investors. But we group investors by their ethnicity first this time. Let α' denote the share of all investors that belong to ethnic group 1 and β_i the share of type i investors that are neutral. The reparametrization can thus be summarized by

$$\begin{aligned}\alpha &= \alpha'\beta_1 + (1-\alpha')\beta_2 \\ \beta &= \frac{\alpha'(1-\beta_1)}{1-\alpha'\beta_1-(1-\alpha')\beta_2}\end{aligned}$$

With no other information, we assume the proportion of biased investors is the same across different ethnic groups, i.e., $\beta_1 = \beta_2 = \beta'$. Thus the reparametrization

can be simply given by

$$\alpha = \beta'$$
$$\beta = \alpha'$$

Given the additional assumption, Proposition 1 above can be interpreted in an alternative manner.

Proposition 5 (Proposition 1′)**.** *The stock price of firms is increasing in the share of total investors who have the same ethnicity as their CEOs under reasonable conditions.*

Proof. In this case, inequality (11) is replaced by

$$N_1 > \frac{(2\alpha' - 1)\rho\beta'(1-\beta')N_2}{\beta' + \beta'(1-\beta')(1-2\alpha') + (1-\alpha')(1-\beta')^2(1-\rho^2)^2}$$

□

TABLE B1: INVESTOR-FIRM COETHNICITY AND INVESTMENT: ALTERNATIVE ETHNICITY CODING

	(1) Investment	(2) Investment	(3) Investment	(4) Investment
CoethnicCEO	0.00983**			
	(0.00401)			
CEOCoethnicityIndex		0.0121***		
		(0.00389)		
CoethnicBoard			0.00252	
			(0.00593)	
BoardCoethnicityIndex				0.0166
				(0.0103)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.554	0.546	0.522	0.543
R2	0.395	0.393	0.446	0.389
N	183754	399457	68062	429519
	(1) OI	(2) OI	(3) OI	(4) OI
CoethnicCEO	0.00309			
	(0.00563)			
CEOCoethnicityIndex		0.0182***		
		(0.00521)		
CoethnicBoard			-0.00563	
			(0.0156)	
BoardCoethnicityIndex				0.0158
				(0.0182)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.0838	0.0731	0.0404	0.0700
R2	0.333	0.325	0.396	0.317
N	274656	602420	109361	648131

The specification is estimated on investor-firm-month-level data. The sample consists of all months in which a trade is made by any investor in any firms stock. The coethnicity variables are defined differently than in Table 1 from the main tables. The cutoffs, both to define individual and board level ethnicity are a high of 0.3 and low of 0.2, compared to 0.4 and 0.3, respectively in the main analysis. Panel A shows the outcome investment, which is the proportion of the investors' portfolio that is held in the share. Panel B shows order imbalance, which measures how much the investor net buys or sells a particular firm's stock, as a proportion of the investor's total traded stock of the same stock during the same month. All specifications in both panels include investor, firm, month, and CEO ethnicity fixed effects and we control for the value control return on equity (ROE) in the prior 12 month period. Standard errors are calculated at the investor ethnicity $\times$ CEO ethnicity level. The dataset spans January 2006-December 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE B2: INVESTOR-FIRM COETHNICITY AND INVESTMENT: NEW INVESTORS

	(1) Investment	(2) Investment	(3) Investment	(4) Investment
CoethnicCEO	0.0146***			
	(0.00376)			
CEOCoethnicityIndex		0.0185***		
		(0.00478)		
CoethnicBoard			0.0202***	
			(0.00482)	
BoardCoethnicityIndex				0.0525***
				(0.0110)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.574	0.570	0.588	0.567
R2	0.400	0.395	0.431	0.391
N	169029	245884	115888	264094
	(1) OI	(2) OI	(3) OI	(4) OI
CoethnicCEO	0.0110**			
	(0.00499)			
CEOCoethnicityIndex		0.0180***		
		(0.00582)		
CoethnicBoard			0.0644***	
			(0.0113)	
BoardCoethnicityIndex				0.112***
				(0.0194)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.0764	0.0667	0.0929	0.0642
R2	0.385	0.379	0.393	0.371
N	259665	380841	178311	408938

The specification is estimated on investor-firm-month-level data. The sample consists of all months in which a trade is made by any investor in any firms stock. The sample is restricted to the investors opening accounts during our sample period so we have the full transaction information of them. Panel A shows the outcome investment, which is the proportion of the investors' portfolio that is held in the share. Panel B shows order imbalance, which measures how much the investor net buys or sells a particular firm's stock, as a proportion of the investor's total traded stock of the same stock during the same month. All specifications in both panels include investor, firm, month, and CEO ethnicity fixed effects and we control for the value control return on equity (ROE) in the prior 12 month period. Standard errors are clustered at the investor ethnicity $\times$ CEO ethnicity level. The dataset spans January 2006-December 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE B3: INVESTOR-FIRM COETHNICITY AND SHORT-RUN RETURNS: ONE DAY AND FIVE DAY

	(1) Return_1day	(2) Return_1day	(3) Return_1day	(4) Return_1day
CoethnicCEO	0.000176			
	(0.000259)			
CEOCoethnicityIndex		-0.0000622		
		(0.000342)		
CoethnicBoard			-0.00164***	
			(0.000332)	
BoardCoethnicityIndex				-0.000840
				(0.00102)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.00137	0.00100	0.00165	0.000955
R2	0.374	0.368	0.394	0.357
N	366041	526864	247205	561911
	(1) Return_5day	(2) Return_5day	(3) Return_5day	(4) Return_5day
CoethnicCEO	0.00000619			
	(0.000413)			
CEOCoethnicityIndex		0.000273		
		(0.000534)		
CoethnicBoard			0.000368	
			(0.000633)	
BoardCoethnicityIndex				0.000657
				(0.00156)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.00240	0.00165	0.00203	0.00158
R2	0.368	0.365	0.400	0.354
N	261485	378507	175056	403516

The specifications are estimated on investor-firm-month-transaction level data. We calculate Returns_1day using the price of the ticker 1 day from the transaction date divided by the price of the buying transaction, and Returns_5day using the price of the ticker 5 days from the transaction date divided by the price of the buying transaction. The sample consists of all transactions initiated during the period. The month indicates origination of the transaction. Any investor may have multiple transactions for a given firms stock in a given month, if there are different shares bought are sold in multiple different future months and thus may result in varying returns. The sample includes both transactions that were closed (sold in full) during the period, as well as those open at the end of the period. For those open at the end of the period, we assume the transactions were closed in the last month. Specifications in both Panel A and Panel B include investor, firm, month, and CEO ethnicity fixed effect. We control for the value control return on equity (ROE) in the prior 12 month period in both panels. Standard errors are clustered at the investor ethnicity $\times$ CEO ethnicity level. The dataset spans January 2006-December 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE B4: INVESTOR-FIRM COETHNICITY AND RETURNS: ALPHA

	(1) Risk-adjusted Returns	(2) Risk-adjusted Returns	(3) Risk-adjusted Returns	(4) Risk-adjusted Returns
CoethnicCEO	-0.00109			
	(0.000870)			
CEOCoethnicityIndex		-0.00145		
		(0.00103)		
CoethnicBoard			-0.0147***	
			(0.00211)	
BoardCoethnicityIndex				-0.0190***
				(0.00344)
Value Controls	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
CEO Ethnicity FE	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.00256	0.00312	0.00998	0.00310
R2	0.607	0.618	0.673	0.619
N	216214	318295	150788	342721

The specifications are estimated on investor-firm-month-transaction level data. Risk-adjusted returns is abnormal return (alpha) based on CAPM, where the risk-free return is defined as the treasury bill rates in Kenya and the market return is defined as NSE20 (the Nairobi Securities Exchange 20 Share Index). The sample consists of all transactions initiated during the period. The month indicates origination of the transaction. All specifications include investor, firm, month, and CEO ethnicity fixed effects. We control for the value control return on equity (ROE) in the prior 12 month period in both panels. Standard errors are clustered at the investor ethnicity $\times$ CEO ethnicity level. The dataset spans January 2006-December 2010. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$