Mirrored Decisions: Facial Similarity in Venture Capital Matching

Emmanuel Yimfor

Columbia University
665 W 130th St, New York, New York 10027, U.S.A.

March 21, 2025

Abstract

This paper examines how facial similarity between venture capitalists and founders predicts investment matches, revealing a previously undocumented dimension of homophily in entrepreneurial finance. Using machine learning-based FaceNet embeddings from 86,409 potential lead investor-founder pairs between 2010-2020, I demonstrate that facial resemblance significantly increases the likelihood of a match occurring, even after controlling for demographic characteristics, network connections, and deal attributes. Through a matched sample design comparing actual matches to counterfactual pairs with similar characteristics, I establish that a one standard deviation increase in facial similarity raises match probability by 1.6 percentage points. This matching effect is substantially stronger when involving non-white investors (3.4× larger) and female investors (2.2× larger), suggesting facial similarity may be particularly influential in matches involving underrepresented groups. The findings contribute to our understanding of subtle factors in venture capital matching.

JEL: G24, J15, L26

Keywords: Venture Capital, Facial Similarity, Image Embeddings, Diversity, Bias, Investment Matching

⋆

Email address: emmanuel.yimfor@columbia.edu (Emmanuel Yimfor)

Abstract

This paper examines how facial similarity between venture capitalists and founders predicts investment matches, revealing a previously undocumented dimension of homophily in entrepreneurial finance. Using machine learning-based FaceNet embeddings from 86,409 potential lead investor-founder pairs between 2010-2020, I demonstrate that facial resemblance significantly increases the likelihood of a match occurring, even after controlling for demographic characteristics, network connections, and deal attributes. Through a matched sample design comparing actual matches to counterfactual pairs with similar characteristics, I establish that a one standard deviation increase in facial similarity raises match probability by 1.6 percentage points. This matching effect is substantially stronger when involving non-white investors (3.4× larger) and female investors (2.2× larger), suggesting facial similarity may be particularly influential in matches involving underrepresented groups. The findings contribute to our understanding of subtle factors in venture capital matching.

1. Introduction

The significant expansion of venture capital markets over the past two decades has transformed the innovation ecosystem, allocating billions to fund entrepreneurs and new ventures. Understanding the matching process between investors and entrepreneurs is crucial given its pivotal role in determining which innovations receive support.

Under conditions of extreme uncertainty that characterize early-stage investments, decision-makers often rely on heuristics and pattern recognition to navigate complex choices (Tversky and Kahneman, 1974; Gigerenzer and Gaissmaier, 2011). When evaluating unfamiliar ventures with limited track records, investors must form judgments based on incomplete information, creating fertile ground for various psychological biases to influence the matching process. In such environments, subtle cues like similarity and familiarity can unconsciously shape perceptions of trust, competence, and potential, affecting which investor-founder pairs ultimately match (Duarte et al., 2012; Huang et al., 2023).

This paper examines a previously unexplored dimension of homophily in entrepreneurial finance: facial similarity between venture capital investors and startup founders.

My empirical analysis uses data from PitchBook spanning 2010 to 2020, focusing on early-stage venture capital deals. For each deal, I identify the lead investment partner and the primary founder (typically the CEO) of the funded startup, collecting facial images for these individuals from public sources (Cook et al., 2024). To construct counterfactual matches, I implement a three-step procedure. First, I match on fundamental deal characteristics: deal type (exact match), industry (exact match), company founding year (exact match), and geographic location (exact match at state level). Second, I restrict potential matches to deals occurring within a 180-day window of the actual deal to ensure temporal comparability. Third, among these

potential matches, I select the deal with the closest deal size to ensure economic similarity. This process yields 43,204 actual deals and an equal number of counterfactual matches, allowing for a precise estimation of how facial similarity affects matching probability.

Using machine learning-based FaceNet embeddings from 86,409 potential investor-founder pairs between 2010-2020, I show that facial resemblance significantly increases the likelihood of investment matches occurring, even after controlling for demographic characteristics, network connections, and deal attributes. Through a matched sample design comparing actual investment matches to counterfactual pairs with similar characteristics, I find that a one standard deviation increase in facial similarity raises match probability by 1.6 percentage points.

The effect of facial similarity on investment matching varies significantly by investor demographics. The matching effect is much stronger for non-white investors (3.4× larger) and female investors (2.2× larger), suggesting facial similarity plays a bigger role in matches involving underrepresented groups. These findings challenge conventional wisdom about homophily in venture capital. If demographic attributes were the main driver of matching patterns, we would expect facial similarity to matter less for minority investors who already have fewer same-demographic founders to choose from. Instead, the stronger effect of facial similarity among underrepresented investors suggests a more complex matching process that goes beyond simple demographic categories to finer levels of perceived similarity.

The analysis also examines how facial similarity interacts with shared network connections between investors and founders. Shared network connections, which capture whether a founder and lead partner previously overlapped at the same company or university, significantly increase the likelihood of matching (Cook et al., 2024). Interestingly, the effect of facial similarity is even stronger when investors and founders share these network connections. This complementary relationship suggests that facial similarity may serve as an additional reinforcing factor when other familiarity

cues are present. Rather than substituting for network-based familiarity, facial similarity appears to amplify it, potentially by strengthening the sense of affinity and trust that comes from shared backgrounds or experiences. This finding further emphasizes how subtle psychological factors may compound more traditional network effects in venture capital matching.

My analysis makes three contributions to the literature. First, I introduce facial similarity as a novel dimension in the study of homophily and assortative matching in entrepreneurial finance. While prior research has identified various dimensions of homophily in venture capital—such as education, ethnicity, gender, and geography—the role of facial resemblance has remained unexplored. Second, using advances in computer vision and a carefully constructed counterfactual approach, I quantify the magnitude of this effect in a way that addresses selection concerns. Finally, by examining heterogeneity across investor demographics, I provide insights into how these effects may vary across different segments of investors.

The findings connect to a broader literature on appearance-based biases in economic decision-making. Duarte et al. Duarte et al. (2012) demonstrate that trustworthy-looking borrowers enjoy higher success rates in peer-to-peer lending, revealing implicit biases based on facial appearance. Similarly, Huang et al. (2023) find investors' decisions are significantly influenced by perceived personality traits inferred from entrepreneurs' photographs. Recent work by Guenzel et al. (2025) shows that personality traits extracted from facial images predict labor market outcomes, while Hu and Ma (2025) document how non-verbal cues in startup pitches affect investor decisions, often disadvantaging women entrepreneurs.

My empirical strategy relies on a carefully constructed counterfactual design. For each actual investor-founder pair that matched, I identify potential counterfactual matches—deals that could have happened but did not—by matching on deal characteristics including industry, founding year, location, and deal size within narrow time windows. This approach helps address concerns that facial similarity might be

correlated with other factors that drive the matching process. The results show that facial similarity predicts match probability even after controlling for demographic similarities (race, gender, age), network connections, and deal characteristics.

This research offers a unique perspective on the subtle psychological factors influencing the venture capital ecosystem. Rather than suggesting that either investors or founders are consciously choosing partners based on facial similarity, the findings point to an unconscious matching mechanism where pairs with greater facial resemblance are more likely to form investment relationships. The emphasis on matching rather than unilateral selection acknowledges the two-sided nature of the venture capital market, where both investors and founders have agency in the partnership formation process.

Understanding these subtle matching patterns provides important insights into how capital allocation decisions are made in entrepreneurial finance. As venture capital plays a crucial role in determining which innovations receive funding, identifying the factors that influence investor-founder matches beyond business fundamentals has important implications for addressing funding disparities and improving the efficiency of capital allocation.

2. Methodology

This paper employs facial embedding techniques to quantify similarity between venture capital investors and startup founders. To measure facial similarity, I use FaceNet, a state-of-the-art deep learning facial recognition system developed by Google Research. The methodology involves several steps: image preprocessing, facial embedding extraction, and similarity score calculation.

2.1. Image Preprocessing

Before extracting facial embeddings, images undergo standardized preprocessing to ensure consistent analysis. The process begins with face detection using RetinaFace,

a robust face detection algorithm that identifies the precise location of faces within images. Once detected, each face is aligned based on facial landmarks (eyes, nose, mouth) to normalize pose variations across images. The aligned faces are then cropped to focus solely on the facial region and resized to 160×160 pixels, which is the standard input size for the FaceNet model.

This preprocessing standardizes the facial images by removing background elements and focusing only on facial features, correcting for variations in head tilt and orientation, normalizing for distance from the camera, and creating uniform input dimensions for the neural network model.

The importance of this standardization is illustrated in Figure 1, which presents processed images of three female finance professors. As shown in the figure, despite differences in the original photos (background, lighting, pose), the standardized images isolate facial features in a consistent format that enables direct comparisons.

2.2. Facial Embedding Extraction

After preprocessing, the standardized facial images are passed through the FaceNet neural network model, which maps each face to a 128-dimensional embedding vector. FaceNet was trained on millions of faces using triplet loss, a method that optimizes the network to generate similar embeddings for images of the same person and dissimilar embeddings for images of different people.

The resulting 128-dimensional embedding represents a face's distinctive characteristics in a compact numerical format. The embedding captures both broad facial structures and fine-grained features, including aspects related to facial geometry and proportions, eye, nose, and mouth shapes, skin texture patterns, facial symmetry, and subtle contours and distinctive features.

These embeddings enable quantitative comparisons between faces, providing a mathematical representation of facial features that goes beyond what might be im-

mediately apparent to human observers.

2.3. Similarity Score Calculation

To quantify similarity between two faces, I calculate the cosine similarity between their respective 128-dimensional FaceNet embeddings. Cosine similarity measures the cosine of the angle between two vectors, yielding a value between -1 and 1, where 1 indicates perfect similarity (identical direction) and -1 indicates perfect dissimilarity (opposite direction). For facial embeddings, which are normalized to unit length, the cosine similarity ranges from 0 to 1, with higher values indicating greater facial resemblance.

For each potential investor-founder pair in the dataset, I calculate the facial similarity score as:

$$\text{similarity}(A, B) = \frac{A \cdot B}{||A|| \cdot ||B||} = \frac{\sum_{i=1}^{128} A_i B_i}{\sqrt{\sum_{i=1}^{128} A_i^2} \sqrt{\sum_{i=1}^{128} B_i^2}} \tag{1}$$

Where A and B represent the 128-dimensional embeddings of the investor and founder, respectively.

As I show in Figure 1, this method produces meaningful similarity scores even within demographic groups. For example, among the three female economics professors in Figure 1, the similarity score between Ulrike Malmendier and Laura Veldkamp (0.6264) is higher than between either of them and Kelly Shue (0.5173 and 0.4700, respectively). This demonstrates that the FaceNet similarity captures nuanced facial feature correspondence rather than just shared demographic characteristics.

After calculating similarity scores for all potential investor-founder pairs, I standardize these scores to facilitate interpretation and comparison. The standardized facial similarity measure serves as the key independent variable in the subsequent analyses of investment matching patterns.

3. Data and Estimation

3.1. Data

This study combines several datasets to analyze the relationship between facial similarity and venture capital investment matching. The primary data comes from PitchBook, a comprehensive database that tracks private capital markets including venture capital investments. I focus on early-stage deals (angel investments, seed rounds, and early-stage VC) occurring between 2010 and 2020, which represents a period of significant growth and evolution in the venture capital ecosystem.

For each deal, I identify the lead partner at the investing firm and the CEO or primary founder of the startup receiving funding. I then collect facial images for these individuals from public sources, primarily company websites and LinkedIn profiles. Following the approach in Cook et al. (2024), I achieve approximately 95 percent coverage of images for both investors and founders in the sample. These images are preprocessed and embedded using FaceNet as described in the methodology section.

To derive demographic information, I use the DeepFace package, which includes the VGG-Face classifier developed by Serengil and Ozpinar for race classification, gender detection, and age estimation (Cook et al., 2024). DeepFace is a hybrid face recognition framework that wraps state-of-the-art face recognition models including VGG-Face, Google FaceNet, OpenFace, Facebook DeepFace, DeepID, and ArcFace. The framework offers both verification (comparing faces) and recognition (determining identity) capabilities with reported accuracy rates exceeding 97% on standard benchmarks.

For each individual in the dataset, I use their image to classify them into demographic groups based on race (White, Black, Asian, Hispanic), gender (Male, Female), and age group (categorized as 25-35, 36-45, 46-55, 56-65, and 66-75 years). The deal characteristics data includes deal type (angel, seed, early-stage VC), deal size in millions of dollars, industry classification, company founding year, geographic location

(state), and deal date.

To construct the counterfactual sample, I implement a three-step procedure. First, I match on fundamental deal characteristics: deal type (exact match), industry (exact match), company founding year (exact match), and geographic location (exact match at state level). Second, I restrict potential matches to deals occurring within a 180-day window of the actual deal to ensure temporal comparability. Third, among these potential matches, I select the deal with the closest deal size to ensure economic similarity. This process yields 43,204 actual deals and an equal number of counterfactual matches, allowing for a precise estimation of how facial similarity affects matching probability.

3.2. Estimation

The empirical strategy relies on comparing actual investment matches to counterfactual matches that share similar observable characteristics. The main specification for analyzing the relationship between facial similarity and investment match probability is:

$$\text{Match}_{ij} = \alpha + \beta \cdot \text{FacialSimilarity}_{ij} + \gamma \cdot X_{ij} + \delta_t + \theta_p + \epsilon_{ij} \tag{2}$$

Where Match_{ij} is an indicator equal to 100 for actual matches and 0 for counterfactual matches between investor i and founder j, $\text{FacialSimilarity}_{ij}$ is the standardized facial similarity score, X_{ij} includes control variables such as same race, same gender, similar age, and shared network connections, δ_t represents year fixed effects, and θ_p represents partner fixed effects.

To examine heterogeneous effects across investor demographics, I estimate:

$$\text{Match}_{ij} = \alpha + \beta \cdot \text{FacialSimilarity}_{ij} + \lambda \cdot \text{FacialSimilarity}_{ij} \times \text{InvestorDemographic}_i + \gamma \cdot X_{ij} + \delta_t + \theta_p + \epsilon_{ij} \quad (3)$$

Where $\text{InvestorDemographic}_i$ represents binary indicators for investor characteristics such as being non-white or female.

Finally, to understand how much variation in facial similarity is explained by demographic factors, I estimate:

$$\text{FacialSimilarity}_{ij} = \alpha + \beta_1 \cdot \text{SameRace}_{ij} + \beta_2 \cdot \text{SameGender}_{ij} + \beta_3 \cdot \text{SimilarAge}_{ij} + \delta_t + \epsilon_{ij} \quad (4)$$

It is important to acknowledge that despite the careful matching procedure, this approach may not fully address all potential confounding factors. Unobserved variables may simultaneously affect facial similarity and investment decisions. For example, certain cultural backgrounds might influence both facial features and investment preferences. Additionally, the population of venture capitalists and founders may have selection biases that are not fully captured in the matching process. While the detailed matching procedure and extensive controls mitigate these concerns, the results should be interpreted as strong correlations rather than definitively causal relationships. At the very least, this paper tests the hypothesis that facial similarity predicts funding matches independent of other observable characteristics, including demographic similarities, network connections, and deal attributes.

4. Results

4.1. Summary Statistics

Table I presents summary statistics comparing actual investment matches with counterfactual matches. The data includes 43,204 actual investment pairs and an

equal number of counterfactual pairs constructed through the matching procedure described earlier. The table reveals several important patterns. First, facial similarity scores are significantly higher between founders and investors in actual matches compared to counterfactual matches, with a mean standardized facial similarity (standardized to have a mean of zero and standard deviation of one) of 0.03 in actual deals versus -0.04 in counterfactual deals (t-stat = 10.77). This difference suggests that facial similarity may play a role in investment decisions.

Second, demographic similarities are also more prevalent in actual matches. Same-race pairs appear in 65% of actual matches versus 60% of counterfactual matches, a statistically significant difference (t-stat = 15.73). Similarly, same-gender pairs are more common in actual matches (84% versus 82%, t-stat = 8.28), as are similar-age pairs (34% versus 31%, t-stat = 8.74). These patterns suggest homophily across multiple dimensions in venture capital matching.

Third, the distributions of investor demographics are balanced across actual and counterfactual matches, with non-white investors comprising 25% and female investors 9% of both samples. This balance confirms the effectiveness of the matching procedure in creating comparable groups that differ primarily in whether a match occurred.

Finally, the matching procedure successfully balanced deal characteristics across the two groups. Deal size averages $9.19 million in both groups, and the distributions of deal types (angel, seed, early-stage VC) and industry sectors are virtually identical. This balance strengthens the validity of subsequent analyses by ensuring that differences in matchinbg outcomes are not driven by underlying differences in deal attributes.

4.2. Facial Similarity Within and Across Groups

Figure 1 illustrates how FaceNet captures meaningful facial similarity even within demographic groups. The figure shows standardized facial images of three female

economics professors along with their pairwise similarity scores. Despite all being women, their facial similarity scores vary substantially, with the similarity between Ulrike Malmendier and Laura Veldkamp (0.6264) notably higher than between either of them and Kelly Shue (0.5173 and 0.4700, respectively). This demonstrates that FaceNet embeddings capture nuanced facial features beyond simple demographic categorization.

Figure 2 further demonstrates this point by showing the distribution of facial similarity scores separated by whether the pairs are of the same race or different races. While same-race pairs have higher average similarity scores, there is substantial overlap between the distributions. This overlap indicates significant within-race variation in facial similarity, confirming that FaceNet captures meaningful facial feature differences beyond broad racial characteristics.

4.3. Determinants of Facial Similarity

Table II examines the demographic factors associated with facial similarity. The results show that shared demographic characteristics significantly predict higher facial similarity. Specifically, same-race pairs have facial similarity scores that are 0.228 standard deviations higher than different-race pairs (Column 1). Same-gender pairs have similarity scores 0.385 standard deviations higher than different-gender pairs (Column 2), and similar-age pairs have scores 0.138 standard deviations higher (Column 3).

When including all three demographic factors together (Column 4), each remains significant with only slight reductions in coefficient magnitudes. This suggests that race, gender, and age contribute independently to facial similarity. The adjusted R-squared of 0.039 in the full model indicates that while demographic factors meaningfully predict facial similarity, most of the variation in facial similarity exists within demographic groups. This within-group variation is crucial for the subsequent analy-

ses, as it allows me to examine the relationship between facial similarity and investment decisions while controlling for demographic similarities.

4.4. Facial Similarity and Investment Match Probability

Table III presents the main results on how facial similarity affects investment match probability. Column 1 shows that a one standard deviation increase in facial similarity is associated with a 2.443 percentage point increase in the probability of an investment match (significant at the 1% level). This effect represents a substantial 4.9% increase relative to the baseline 50% probability in the matched sample.

Columns 2 through 5 progressively add demographic controls to assess whether facial similarity has an independent effect beyond demographic homophily. After controlling for same race, same gender, similar age, and shared network connections, the coefficient on facial similarity remains economically and statistically significant at 1.576 (Column 5). This indicates that facial similarity predicts investment matches even among pairs that share the same demographic characteristics and network connections.

The control variables themselves reveal substantial homophily effects. Having the same race increases match probability by 7.414 percentage points, same gender by 5.079 percentage points, and similar age by 3.826 percentage points. Shared network connections have the largest effect, increasing match probability by 26.997 percentage points. Despite these strong homophily effects, facial similarity remains a significant predictor, suggesting it captures an additional dimension of similarity that influences investment decisions.

4.5. Heterogeneous Effects Across Investor Demographics

Table IV explores how the effect of facial similarity varies across investor demographics. Column 1 shows a strong interaction between facial similarity and investor

race. While facial similarity has a positive effect for white investors (coefficient of 0.844), this effect is substantially larger for non-white investors (additional effect of 2.825, making the total effect 3.4 times larger). Column 2 shows a similar pattern for investor gender, with the effect for female investors (additional effect of 1.935) approximately 2.3 times larger than for male investors.

Column 3 examines the interaction with shared network connections. The coefficient on the interaction term (1.950) indicates that facial similarity has a stronger effect on match probability when the investor and founder share network connections. Column 4 includes all interaction terms simultaneously, and the results remain qualitatively similar, suggesting that each interaction represents an independent effect.

Figure 3 visualizes these heterogeneous effects using binned scatter plots. Panel A shows that the relationship between facial similarity and match probability is steeper for non-white investors compared to white investors, especially at higher levels of facial similarity. Similarly, Panel B shows a steeper relationship for female investors compared to male investors. These visualizations reinforce the finding that facial similarity has stronger effects on investment decisions for investors from underrepresented groups.

This finding is particularly surprising given the venture capital landscape. There are relatively few underrepresented investors, while many underrepresented founders remain unfunded by white male investors due to various factors. If this select group of underrepresented investors were primarily matching with previously overlooked underrepresented founders, we would expect facial similarity to be less important in their decisions, not more. The fact that facial similarity actually plays a stronger role for non-white and female investors contradicts this expectation, suggesting that subtle psychological factors influence matching patterns in ways that transcend straightforward demographic considerations.

5. Conclusion

This paper documents a novel dimension of homophily in entrepreneurial finance: facial similarity between venture capital investors and startup founders. Using machine learning-based facial embeddings from 86,409 potential investor-founder pairs, I find that greater facial resemblance significantly increases the probability of investment matches, even after controlling for demographic similarities, network connections, and deal characteristics. A one standard deviation increase in facial similarity raises match probability by 1.6 percentage points—an economically meaningful effect representing about one-fifth the magnitude of shared racial background. Surprisingly, facial similarity exerts a much stronger influence for underrepresented investors, with effects 3.4 times larger for non-white investors and 2.2 times larger for female investors compared to their white and male counterparts. This finding challenges conventional wisdom, as we would expect underrepresented investors, who face a larger pool of overlooked founders from similar demographic backgrounds, to rely less on subtle facial cues, not more. While demographic similarities strongly predict facial similarity, sufficient within-group variation enables identification of facial similarity effects independent of demographic homophily. These results highlight how subtle, often unconscious factors may influence economic matching processes, suggesting homophily mechanisms in venture funding are more complex and multifaceted than previously recognized. Although I emphasize caution in causal interpretation, these findings reveal the potential importance of subtle physical resemblance in shaping economic interactions and resource allocation decisions.

References

Cook, L.D., Marx, M., Yimfor, E., 2024. Funding Black High-Growth Startups. Working Paper 30682. Journal of Finance, Forthcoming.

Duarte, J., Siegel, S., Young, L., 2012. Trust and credit: The role of appearance in P2P lending. Review of Financial Studies 25, 2455–2484. doi:10.1093/rfs/hhs071.

Gigerenzer, G., Gaissmaier, W., 2011. Heuristic decision making. Annual review of psychology 62, 451–482.

Guenzel, M., Kogan, S., Niessner, M., Shue, K., 2025. Ai personality extraction from faces: Labor market implications .

Hu, A., Ma, S., 2025. Persuading investors: A video-based study. Journal of Finance .

Huang, X., Ivković, Z., Jiang, J., Wang, I., 2023. Angel investment and first impressions. Journal of Financial Economics .

Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. science 185, 1124–1131.

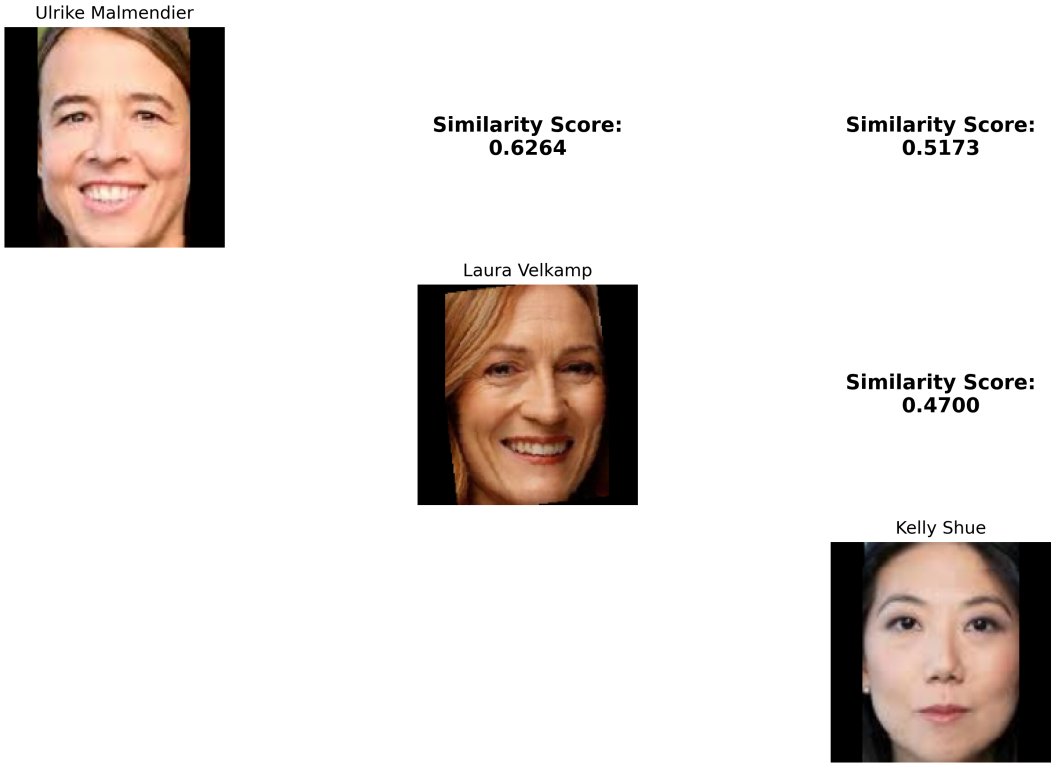


Figure 1: FaceNet Similarity Analysis of Female Economics Professors

This figure demonstrates the effectiveness of FaceNet embeddings in capturing meaningful facial similarity within gender groups. The diagonal shows the standardized face images that FaceNet processes before generating embeddings. The similarity scores between pairs (displayed in off-diagonal positions) are calculated using cosine similarity between the 128-dimensional FaceNet embeddings, scaled to a 0-1 range. Higher scores indicate greater facial similarity as determined by the neural network. Note that these similarity scores reflect genuine facial feature correspondence rather than just shared gender characteristics, showing that FaceNet can distinguish between individuals of the same gender based on their unique facial features.

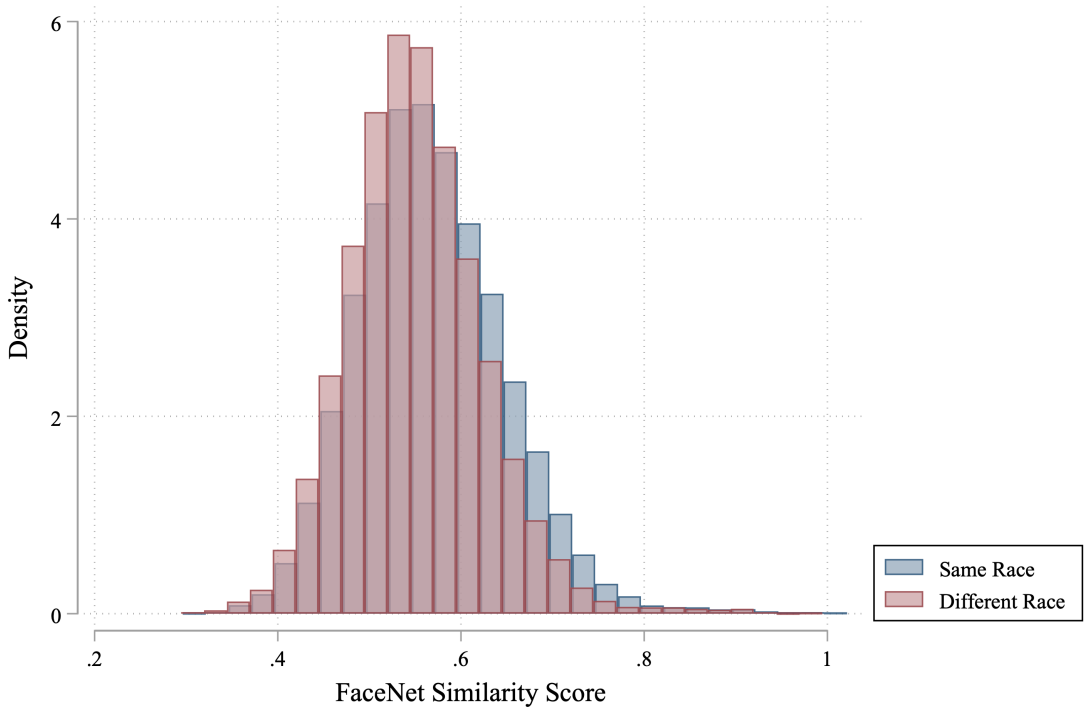
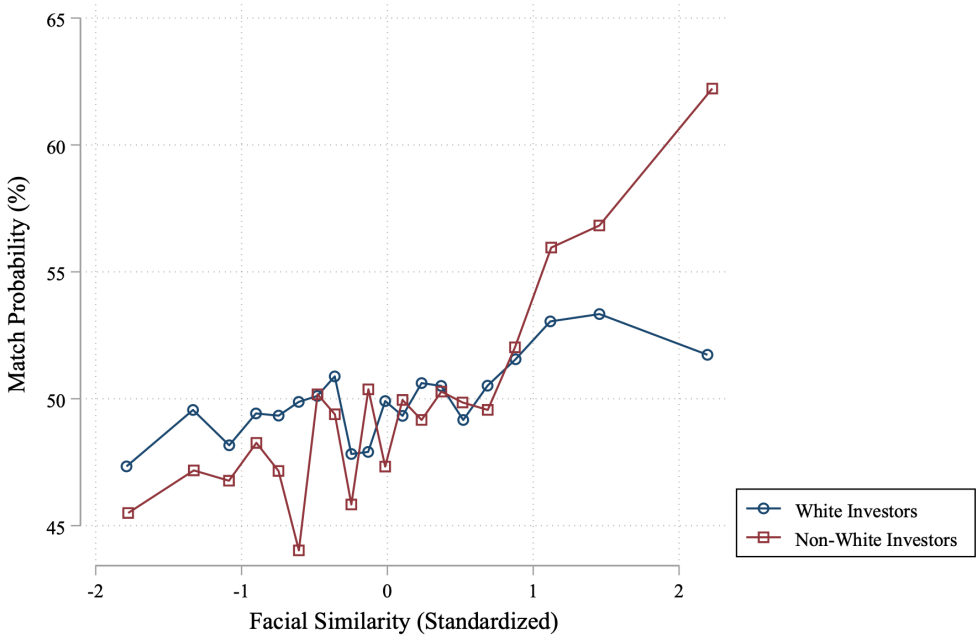


Figure 2: Distribution of Facial Similarity Scores Within and Across Racial Groups

This figure displays the distribution of FaceNet facial similarity scores between founder-investor pairs, separated by whether the individuals are of the same race (navy) or different races (maroon). The overlap between distributions demonstrates significant within-race variation in facial similarity, confirming that FaceNet captures meaningful facial feature differences beyond racial characteristics. Higher similarity scores indicate greater facial resemblance between pairs.

Panel A: Effect of Facial Similarity by Investor Race



Panel B: Effect of Facial Similarity by Investor Gender

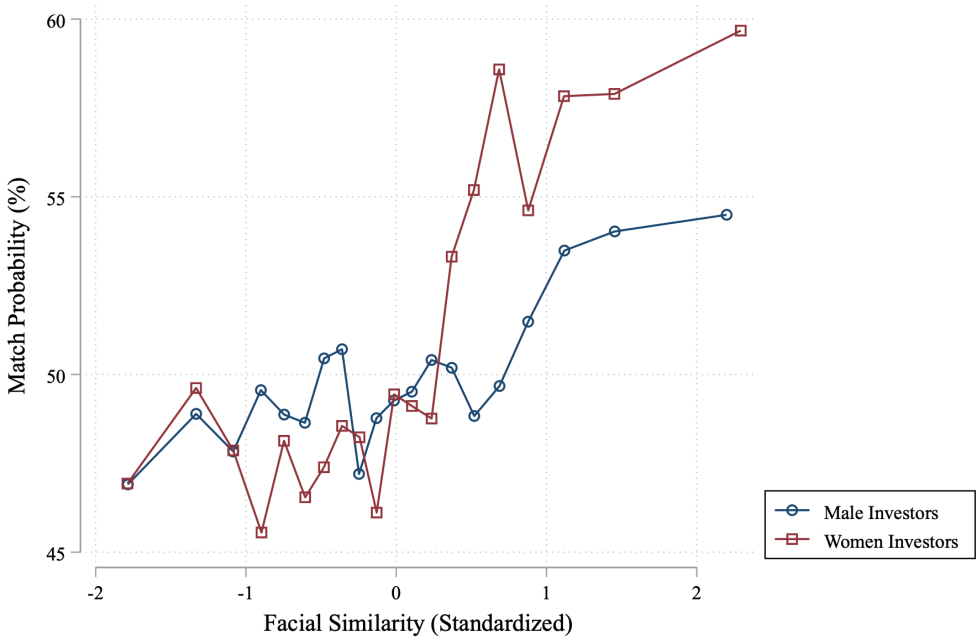


Figure 3: Facial Similarity and Investment Match Probability

This figure examines the relationship between facial similarity and venture capital investment decisions across investor demographics. Panel A presents a binned scatter plot showing how standardized facial similarity (FaceNet) between founders and investors affects the probability of an investment match, separately for white and non-white investors. Deals are grouped into bins along the horizontal axis based on facial similarity scores, with the vertical axis showing the percentage likelihood of a match. Panel B presents a similar analysis by investor gender, contrasting the effects for male versus female investors. Each point represents the average match probability for deals within each facial similarity bin.

Table I: Summary Statistics for Startups by Founder Race (2000-2020)

This table presents summary statistics comparing actual investment matches with counterfactual matches that could have happened but did not. Each observation represents a lead partner-founder pair, where the counterfactual pairs are constructed by matching on deal characteristics. Specifically, we match deals based on deal type, industry, founding year, and state, requiring deals to occur within a 180-day window, and selecting the closest match by deal size. This matching procedure creates a controlled comparison where the only difference between actual and counterfactual deals is the partner-founder pairing, allowing us to isolate homophily effects. The key variables are defined as follows: *Facial Similarity (Standardized)* is the standardized cosine similarity between partner and founder FaceNet embeddings (higher values indicate greater facial resemblance); *Same Race*, *Same Gender*, and *Similar Age* are indicators equal to one if the partner and founder share these demographic characteristics; *Non-White Investor* and *Female Investor* are indicators for partner demographics; *Deal Size* is the total capital raised in the financing round; *Angel Investment*, *Seed Round*, and *Early Stage VC* are indicators for deal type; industry indicators identify the startup's primary sector. The t-statistics test differences in means between actual and counterfactual deals. $^{***}p < 0.01$ denotes significance at the 1% level, $^{**}p < 0.05$ denotes significance at the 5% level, and $^{*}p < 0.10$ denotes significance at the 10% level. All continuous variables are winsorized at the 1% and 99% levels to minimize the influence of outliers.

	Actual Deal		**Counterfactual Deal**		
	N = 43,204		**N = 43,204**		
	Mean	Std. Dev.	Mean	Std. Dev.	t-stat
Facial Similarity (Standardized)	0.03	0.97	−0.04	0.95	10.77***
Same Race	0.65	0.48	0.60	0.49	15.73***
Same Gender	0.84	0.37	0.82	0.39	8.28***
Similar Age	0.34	0.47	0.31	0.46	8.74***
Non-White Investor	0.25	0.43	0.25	0.43	−0.01
Female Investor	0.09	0.29	0.09	0.29	0.00
Deal Size ($ Millions)	9.19	27.40	9.19	27.40	0.00
Deal Year	0.01	0.11	0.01	0.11	0.00
Angel Investment	0.46	0.50	0.46	0.50	0.00
Seed Round	0.53	0.50	0.53	0.50	0.00
Early Stage VC	2016.15	3.10	2016.15	3.10	0.00
Information Technology	0.13	0.33	0.13	0.33	0.00
Healthcare	0.19	0.39	0.19	0.39	0.00
Financial Services	0.00	0.06	0.00	0.06	0.00
Consumer Products & Services	0.02	0.14	0.02	0.14	0.00
Business Products & Services	0.09	0.29	0.09	0.29	0.00
Energy	0.56	0.50	0.56	0.50	0.00
Materials & Resources	0.00	0.04	0.00	0.04	0.00

Table II: Determinants of Facial Similarity Between Founders and Investors

This table presents results from our analysis of the demographic factors associated with facial similarity between founders and investors. We report coefficients from OLS regressions with standardized facial similarity (FaceNet) scores as the dependent variable. Standard errors are clustered at the partner level and shown in parentheses. Each column examines a different demographic factor that might relate to facial similarity: same race (Column 1), same gender (Column 2), similar age (Column 3), and all factors combined (Column 4). Year fixed effects are included in all specifications. Facial similarity is measured using FaceNet embeddings, which capture fine-grained facial features beyond demographic characteristics. $^{***}p < 0.01$ denotes significance at the 1% level, $^{**}p < 0.05$ denotes significance at the 5% level, and $^{*}p < 0.10$ denotes significance at the 10% level.

	Dependent Variable: Facial Similarity (Standardized)			
	(1)	(2)	(3)	(4)
Same Race	0.228***			0.217***
	(0.009)			(0.009)
Same Gender		0.385***		0.371***
		(0.011)		(0.011)
Similar Age			0.138***	0.136***
			(0.009)	(0.009)
Observations	86,409	86,409	86,409	86,409
Adjusted-R^2	0.013	0.023	0.005	0.039
Year Fixed Effects	Yes	Yes	Yes	Yes

Table III: Effect of Facial Similarity on Investment Match Probability

This table examines how facial similarity between founders and investors affects the probability of investment matching. We report coefficients from OLS regressions where the dependent variable is an indicator for whether an investment deal occurred (multiplied by 100). Standard errors are clustered at the partner level and shown in parentheses. Column (1) shows the baseline effect of facial similarity, while subsequent columns add controls for demographic similarities: same race (Column 2), same gender (Column 3), similar age (Column 4), and shared network connections (Column 5). All specifications include year and partner fixed effects to account for temporal trends and time-invariant partner characteristics. Facial similarity is measured using standardized FaceNet embeddings, which capture fine-grained facial features that may influence investment decisions through homophily or other mechanisms. $^{***}p < 0.01$ denotes significance at the 1% level, $^{**}p < 0.05$ denotes significance at the 5% level, and $^{*}p < 0.10$ denotes significance at the 10% level.

	Dependent Variable: Match Probability (%)				
	(1)	(2)	(3)	(4)	(5)
Facial Similarity (Standardized)	2.443***	1.967***	1.732***	1.639***	1.576***
	(0.232)	(0.230)	(0.232)	(0.231)	(0.229)
Same Race		7.611***	7.546***	7.528***	7.414***
		(0.551)	(0.552)	(0.551)	(0.547)
Same Gender			5.075***	5.060***	5.079***
			(0.754)	(0.754)	(0.748)
Similar Age				4.105***	3.826***
				(0.579)	(0.579)
Same Network					26.997***
					(0.998)
Observations	86,409	86,409	86,409	86,409	86,409
Adjusted-R^2	-0.164	-0.159	-0.158	-0.157	-0.144
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Partner Fixed Effects	Yes	Yes	Yes	Yes	Yes

Table IV: Heterogeneous Effects of Facial Similarity Across Investor Demographics

This table examines how the effect of facial similarity on investment matching varies across investor demographics and network connections. We report coefficients from OLS regressions with standard errors clustered at the partner level shown in parentheses. Column (1) shows the interaction between facial similarity and investor race, testing whether non-white investors respond differently to facial similarity than white investors. Column (2) shows the interaction with investor gender, testing for differential effects between male and female investors. Column (3) examines how shared network connections moderate the effect of facial similarity. Column (4) includes all interaction terms simultaneously to test their joint significance. All specifications include year and partner fixed effects. The dependent variable in all columns is Match Probability (percent), which indicates whether a real investment match occurred. $^{***}p < 0.01$ denotes significance at the 1% level, $^{**}p < 0.05$ denotes significance at the 5% level, and $^{*}p < 0.10$ denotes significance at the 10% level.

	Dependent Variable: Match Probability (%)			
	(1)	(2)	(3)	(4)
	Race	Gender	Network	All
Facial Similarity (Standardized)	0.844***	1.434***	1.469***	0.628**
	(0.264)	(0.236)	(0.235)	(0.271)
Facial Similarity × Non-White Investor	2.825***			2.766***
	(0.530)			(0.529)
Facial Similarity × Female Investor		1.935**		1.788*
		(0.932)		(0.929)
Facial Similarity × Shared Network			1.950**	1.820**
			(0.871)	(0.874)
Same Race	7.274***	7.419***	7.407***	7.276***
	(0.547)	(0.547)	(0.547)	(0.547)
Same Gender	5.145***	4.990***	5.078***	5.060***
	(0.748)	(0.744)	(0.748)	(0.745)
Similar Age	3.834***	3.829***	3.829***	3.838***
	(0.579)	(0.579)	(0.579)	(0.579)
Shared Network	26.979***	26.974***	26.904***	26.872***
	(0.995)	(0.998)	(1.001)	(0.998)
Observations	86,409	86,409	86,409	86,409
Adjusted-R^2	-0.14	-0.14	-0.14	-0.14
Year Fixed Effects	Yes	Yes	Yes	Yes
Partner Fixed Effects	Yes	Yes	Yes	Yes