

Colonial, graylined and foreign discrimination in rental housing: meta-regression evidence from fifty-four correspondence tests

Helen X. H. Bao¹

¹Department of Land Economy, University of Cambridge, Cambridge CB3 9EP, United Kingdom. Email: helenxhbao@gmail.com

16th May 2026

Abstract

Correspondence-test studies of housing discrimination have accumulated steadily over the past two decades and consistently document that minority-named, immigrant-named or religiously-marked applicants receive fewer responses to rental inquiries than otherwise-comparable majority applicants. The economic theory of discrimination distinguishes three canonical mechanisms underlying this pattern, namely taste-based (Becker, 1957), statistical (Phelps, 1972; Arrow, 1973) and identity-driven (Akerlof and Kranton, 2000). We propose that the published correspondence-test record can be organised into three analytical categories of discrimination targets, colonial, graylined and foreign, each of which is theoretically anchored in a different mixture of the three economic mechanisms. We extend the cross-study sample to fifty-four correspondence-test studies across twenty-two countries with fieldwork between 2003 and 2022 and provide a meta-regression identification test of the three-category reading. Each category responds to its theoretically-paired country-level diversity index: the colonial-by-ethnic-fractionalisation, graylined-by-religious-fractionalisation, and foreign-by-linguistic-fractionalisation interactions are each positive, statistically significant, and consistent in sign with the underlying mechanism. The pattern is robust to a FAT-PET correction for publication selection. Recent fieldwork from Singapore, Kuala Lumpur and Malaysia is consistent with the identity-economics reading and shows that the same identity can carry the discriminator role in one jurisdiction and the target role in another. The aggregate consequence of repeated individual-level screening decisions is the social-closure pattern (Weber, 1922; Tilly, 1998) that the sociological tradition has long documented. The research agenda that follows includes simultaneous estimation of the three categories across multi-ethnic high-migration markets, systematic collection of landlord-side covariates, integration of correspondence-test evidence with measures of resource scarcity, geographic extension beyond the OECD, and a minimal common protocol for cross-study comparability.

Keywords: housing discrimination; correspondence test; tripartite classification; social closure; opportunity hoarding; field experiment; rental markets

JEL classification: J15; J71; R21; R31

1 Introduction

Correspondence-test studies of housing discrimination form one of the more methodologically consistent literatures in applied housing economics. Across more than four decades of audit-style and email-correspondence experiments, the broad finding has been robust: applicants whose names signal minority ethnicity, immigrant origin or religious identity receive fewer responses to inquiries about advertised rental properties than otherwise-comparable majority applicants. Meta-analytic syntheses by Flage (2018), Auspurg et al. (2020) and Gusciute et al. (2022) confirm the persistence of the gap across the 2006 to 2023 period and across the OECD countries that dominate the published evidence. An earlier 2024 version of this analysis organised the published record into three analytical categories: colonial-legacy discrimination, graylined or symbolic-stigma discrimination, and foreign-origin discrimination.

The descriptive case for the tripartite classification has not previously been tested through a direct identification exercise. The categories sort the published estimates into a sensible ordering, with colonial-type and graylined-type discrimination producing larger pooled penalties than foreign-type discrimination, but the ordering alone does not establish that the three categories are responding to distinct mechanisms. They could be three labels for a single underlying phenomenon, three labels for one phenomenon plus measurement noise, or three labels for three mechanisms. Distinguishing between these possibilities requires a test in which each category responds to a moderator that the theoretical mechanism specifies in advance.

This paper provides such a test. We extend the cross-study sample to fifty-four correspondence-test studies across twenty-two countries with fieldwork between 2003 and 2022, and we interact each category indicator with the country-level fractionalisation index that its theoretical mechanism predicts as moderator. Colonial-type discrimination is grounded in visible-race hierarchies and is paired with the ethnic fractionalisation index of Alesina et al. (2003). Graylined-type discrimination is grounded in religiously-coded symbolism and is paired with religious fractionalisation. Foreign-type discrimination is grounded in cultural unfamiliarity, accent and name-language signals and is paired with linguistic fractionalisation. The three interactions are each positive, statistically significant, and consistent in sign with the underlying mechanism. A one-standard-deviation increase in the relevant index reduces the category-specific penalty by an amount equivalent to between twenty and forty per cent of the unconditional category coefficient. The three-mechanism reading of the tripartite classification therefore passes its first direct identification test.

The empirical result is interpretable within the standard economic theory of discrimination. The category-by-fractionalisation interactions are consistent with each of the three canonical mechanisms operating jointly. The reduction in the colonial-category penalty as ethnic fractionalisation rises is the prediction of the taste-based and statistical-discrimination models under the contact-hypothesis interpretation (Yinger, 1998). The reduction in the graylined-category penalty as religious fractionalisation rises is the prediction of an identity-driven model (Akerlof and Kranton, 2000; Chen and Li, 2009) in which the symbolic in-group loses salience as the religious composition diversifies. The reduction in the foreign-category penalty as linguistic fractionalisation rises is the prediction of the statistical-discrimination model under the standard

interpretation that linguistic homogeneity sharpens the signal carried by accent and name-language cues. We treat the tripartite typology as a catalogue of the axes along which group boundaries are drawn in housing markets rather than as a typology of group essences. The reading accommodates a recent empirical pattern that essentialist accounts do not predict, namely that the same identity can carry the discriminator role in one jurisdiction and the target role in another, which is the central prediction of identity economics. The aggregate consequence of repeated individual-level screening decisions in markets that permit discretionary screening is the social-closure pattern documented by the sociological tradition (Weber, 1922; Parkin, 1979; Murphy, 1988; Tilly, 1998).

Three recent empirical studies are consistent with the structural reading. The first is a Singapore correspondence experiment (Bao and Tan, working paper) in which all three categories of the tripartite classification are estimated simultaneously, and the ethnic-majority foreign sub-group (British Chinese applicants) faces a robust foreign penalty even though their within-Singapore ethnic-majority status would otherwise insulate them. The second is a preliminary correspondence experiment in Kuala Lumpur (Bao and colleagues, work in progress) in which the Malay-Muslim majority of Malaysian landlords produces a response-rate gradient that disadvantages non-Bumiputera Malaysians and non-Malaysian foreigners, despite the Malay-Muslim identity appearing in the cross-study record almost exclusively as a target of graylined discrimination in non-Muslim-majority countries. The third is observational evidence from Wang et al. (2026) that documents a Malay-Muslim landlord price premium on non-Malay tenants in the Malaysian rental market, triangulating the discriminator-role finding through a separate methodology.

The remainder of the paper is organised as follows. Section 2 sets out the analytical framework, anchoring the tripartite classification in the canonical economic theory of discrimination. Section 3 reports the meta-regression evidence, including the descriptive statistics, the baseline meta-regression, the mechanism-specific heterogeneity analyses that constitute the central empirical contribution, and a publication-selection robustness check. Section 4 summarises three recent empirical signals, develops the identity-economics interpretation with a complementary social-closure lens, and sets out the research agenda that follows. Section 5 concludes.

2 Analytical framework

2.1 Three economic mechanisms of discrimination

The economic theory of discrimination distinguishes three canonical mechanisms. Becker (1957) formalised taste-based discrimination, in which the discriminator carries a preference against transacting with out-group counterparts and accepts a utility cost to avoid such transactions. Phelps (1972) and Arrow (1973) developed statistical discrimination, in which the discriminator infers the expected attributes of an applicant from group-level signals because individual-level attributes are imperfectly observed at the screening stage. Akerlof and Kranton (2000) extended the canonical preference set to include identity prescriptions, with individuals deriving utility from acting in accordance with the behavioural rules of the groups they identify with,

including prescriptions to favour in-group members in the allocation of scarce resources. The three mechanisms are not mutually exclusive. In a given screening decision the discriminator may be driven by any combination of taste, inference and identity, and the correspondence-test design at the inquiry stage observes only the joint outcome.

Housing markets are a canonical setting in which all three mechanisms can produce differential treatment. The screening decision belongs to the resource holder rather than to a regulator, the applicant signal is uniformly noisy across applicants, and the resource is excludable and rivalrous. Statistical discrimination is theoretically dominant when the discriminator believes that the group signal predicts contract-relevant attributes such as ability to pay rent, length of tenancy or maintenance behaviour. Taste-based discrimination is theoretically dominant when the discriminator carries a personal cost of interacting with out-group counterparts. Identity-driven discrimination is theoretically dominant when the discriminator perceives that allocating the resource to an in-group counterpart confers a positive identity payoff, or that allocating it to an out-group counterpart confers a negative one. Distinguishing the three mechanisms empirically at the inquiry stage requires either a design that varies the information available to the discriminator (the conventional approach to identifying statistical versus taste-based discrimination, following Altonji and Pierret, 2001 and Heckman, 1998) or a design that exposes the discriminator to applicants whose signals invoke different combinations of the three categories developed in Section 2.2.

2.2 A tripartite classification of housing-discrimination targets

The 2024 analysis on which this paper builds proposed a three-category classification of housing-discrimination targets. Colonial-type discrimination refers to differential treatment rooted in the historical legacy of empire and applied to groups whose presence in the host country is itself a consequence of colonial migration or colonial-era population movements. Operationally, the colonial flag attaches to comparisons that involve applicants of African or other former-colony origin in countries that played a significant colonial role (principally the United Kingdom, France, Portugal and Spain) and to indigenous-minority comparisons in settler-colonial contexts such as New Caledonia.

Graylined-type discrimination refers to differential treatment rooted in the symbolic association of a group with a historically contingent event or social pattern. The term is constructed by analogy with redlining in mortgage markets, and it signals a diffuse and less explicit form of exclusion. The graylined flag attaches most often to Muslim applicants after 2001 and to identifiable Jewish applicants in certain markets. Recent evidence on Somali-refugee applicants in the United States (Gorzig and Rho, 2026) extends the graylined case to a population that is simultaneously religiously and racially marked and that arrives through a refugee channel that is itself a source of symbolic stigma. Within-country evidence from Chinese cities (Zhu et al., 2026) confirms that graylined discrimination can also operate without any ethnic-minority dimension, with city-level symbolic stigma producing observable response-rate gaps among applicants of the same ethnic-majority background.

Foreign-type discrimination refers to differential treatment rooted in national-origin signals,

language, accent or assumed cultural distance, in the absence of either a colonial history that connects the group to the host country or a symbolic event that marks the group. The foreign flag attaches to applicants whose names, telephone country codes or relocation cues signal that they are non-citizens or recent arrivals, and whose ethnic appearance does not itself produce a colonial-type or graylined-type contrast.

The three categories map onto distinct mixtures of the economic mechanisms of Section 2.1. Colonial-type discrimination is theoretically anchored in taste-based and statistical mechanisms inherited from colonial-era hierarchies, with limited identity-prescription content because the dominant group does not derive substantial in-group identity payoff from excluding a long-settled subordinate community. Graylined-type discrimination is theoretically anchored in identity-prescription and statistical mechanisms triggered by the symbolic event, with limited taste-based content because the discriminator does not necessarily carry a personal antipathy to the target group outside the symbolic frame. Foreign-type discrimination is theoretically anchored in statistical and identity-prescription mechanisms tied to cultural distance, with limited taste-based content because foreignness has no settled cultural meaning until it is institutionally activated. The classification therefore predicts that each category should respond to a different country-level moderator, and that the moderator should map to the dominant theoretical channel for that category, which we test in Section 3.4.

2.3 Implications for identification

The identification implication that follows from the classification is that an estimate of any single category requires the simultaneous inclusion of the other two. A correspondence-test design that contains only colonial-type and majority profiles, for example, produces a colonial-effect estimate that absorbs any unmeasured foreign or graylined component correlated with colonial-minority status in the host country. The same logic applies symmetrically. The published correspondence-test literature contains no single design that varies along all three dimensions simultaneously, with the Singapore experiment discussed in Section 4.1 the first to do so. The implication for re-reading the cross-study record is that the published category-specific effect sizes are conservative for the within-category mechanism and ambiguous for the cross-category leakage. The meta-regression strategy adopted in Section 3 addresses the leakage by pooling estimates across studies whose within-design variation differs and by exploiting country-level moderators that the underlying theoretical mechanisms specify in advance.

3 State of the evidence

3.1 Coverage

The cross-study sample assembled for the earlier 2024 analysis covered forty correspondence-test studies on housing discrimination between 2006 and 2023, drawn from the union of three prior review papers (Flage, 2018; Auspurg et al., 2020; Gusciute et al., 2022) and a citation expansion against the Web of Science and Scopus databases. The current paper extends the sample to

cover the period through May 2026 and adds sixteen studies that satisfied the original inclusion criteria, namely peer-reviewed correspondence tests in English on rental-housing markets and that report a within-study majority benchmark against which an odds ratio can be calculated. The updated sample contains fifty-four unique studies across twenty-two countries, with fifty-six study-country entries once Antfolk et al. (2019) is split into its United Kingdom and Polish settings and similarly for other multi-setting designs. Two of the added countries, Austria and Turkey, were not represented in the 2024 sample. The geographic skew remains pronounced, with European studies accounting for approximately two-thirds of the sample, and the under-representation of Asian and Latin American settings is a known feature of the literature that bears on the foreign-category argument developed in Section 4. Figure 1 reports the country-level distribution of the studies.

Three further 2026 publications are referenced in the present paper as illustrative material but are not included in the meta-regression sample. Sugasawa and Harano (2026) is a follow-up correspondence study in Japan that identifies a combined effect of foreign and ethnic-minority status without separating the two components, on the grounds that the applicant names that signal foreignness in the Japanese setting also signal ethnic-minority status. The design illustrates the central methodological concern of Section 2.3, namely that the foreign category is empirically inseparable from ethnic-minority status under most published designs. Zhu et al. (2026) examines housing discrimination across Chinese cities without varying applicant ethnicity, isolating a within-host-country symbolic-stigma mechanism that fits the graylined category exactly while involving no ethnic-minority dimension at all. The paper supports the analytical claim that graylined discrimination is not a residual category of ethnic-minority discrimination but a separable mechanism. Gorzig and Rho (2026), which compares responses to Somali-refugee applicants in the United States, is included in the meta-regression sample under both the graylined and the foreign category flags and is one of the cleaner recent examples of overlap between the two categories.

3.2 Headline estimates

Read against the tripartite classification, the updated cross-study record gives mean odds ratios of 0.62 for the colonial category, 0.65 for the graylined category and 0.76 for the foreign category. The overall pooled odds ratio is 0.68, which corresponds to a 32 percentage-point gap relative to the corresponding majority-group response rate. The gendered evidence is more pronounced for male applicants than for female applicants, with the male-applicant pooled odds ratio at 0.60 against a female-applicant odds ratio of 0.73. The colonial and graylined estimates have shifted upwards by approximately 0.06 to 0.08 odds-ratio units relative to the 2024 sample, which reported corresponding values of 0.56 for colonial and 0.57 for graylined. The foreign-category estimate is essentially unchanged, at 0.76 in the present sample against 0.78 in the 2024 sample. The category ordering, in which foreign-type discrimination produces the smallest pooled effect of the three and colonial-type discrimination the largest, is preserved under the update. The empirical gap between the foreign category and the other two remains approximately 0.11 to 0.14 in odds-ratio units, which corresponds to a non-trivial difference in

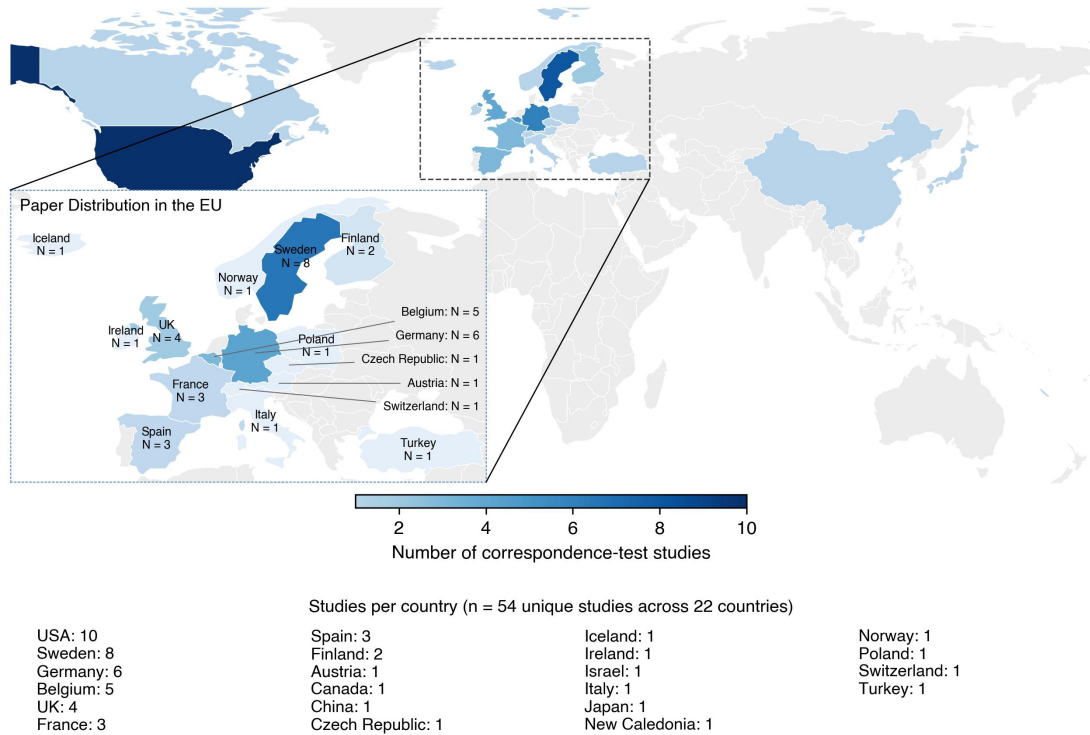


Figure 1: Geographic distribution of correspondence-test studies in the updated sample ($n = 54$ unique studies, 22 countries). The inset is a zoomed view of Europe, where the densely-sampled countries are otherwise illegible at world-map scale.

the implied response-rate penalties. The upward shift in the colonial and graylined estimates is consistent with the inclusion of more recent studies in markets and time periods in which institutional anti-discrimination responses have matured. Table 1 reports the descriptive statistics for the dependent variable and the controls used in the meta-regression. Figure 2 displays the study-level inverse-variance-weighted estimates as a forest plot.

3.3 Meta-regression evidence

The meta-regression specification regresses study-level odds ratios on the tripartite indicators (with foreign as the omitted baseline), gender, year of fieldwork, ethnic fractionalisation of the host country, an indicator for major-city studies, country population, an indicator for former colonial-power host countries, sample size, an indicator for within-study designs and continent fixed effects. Standard errors are clustered at the paper level. The simple specification with only the tripartite and gender indicators yields a colonial coefficient of approximately -0.10 against the foreign baseline, a graylined coefficient of approximately -0.06 against the same baseline, and a male-applicant coefficient of approximately -0.11 . Both category indicators are smaller in magnitude than in the 2024 sample, which reported -0.21 and -0.20 respectively. The reduction reflects the upward shift in the colonial and graylined point estimates discussed in Section 3.2.

The saturated specification, which adds country and study-level controls and continent fixed

Table 1: Descriptive statistics for the dependent variable and control variables.

Variable	Mean	SD	Min	Max	Description
<i>Panel A. Odds ratios (key outcome variable, minority rows only)</i>					
Odds ratio (all minority sub-estimates)	0.681	0.280	0.040	2.085	Odds of the minority group receiving a response, divided by the corresponding majority odds within the same study.
Odds ratio Colonial = 1	0.619	0.222	0.188	1.231	Odds ratios for sub-estimates flagged as colonial-type discrimination.
Odds ratio Graylined = 1	0.653	0.286	0.040	2.085	Odds ratios for sub-estimates flagged as graylined-type discrimination.
Odds ratio Foreign = 1	0.764	0.300	0.040	2.085	Odds ratios for sub-estimates flagged as foreign-type discrimination.
Odds ratio Gender = Female	0.734	0.200	0.273	1.307	Odds ratios reported separately for female applicants.
Odds ratio Gender = Male	0.600	0.294	0.137	2.085	Odds ratios reported separately for male applicants.
Odds ratio Gender = Both	0.732	0.288	0.040	1.705	Odds ratios reported separately for both applicants.
<i>Panel B. Control variables</i>					
Year of fieldwork	2016.613	4.739	2003.000	2022.000	Calendar year in which the correspondence-test fieldwork was conducted.
Sample size (per sub-estimate, $\times 1,000$)	0.517	0.886	0.023	7.090	Number of inquiries underlying each sub-estimate, scaled by 1,000.
Ethnic Fractionalisation	0.276	0.206	0.012	0.712	Probability that two random members of the host country differ in ethnic group (Alesina et al., 2003).
Linguistic Fractionalisation	0.289	0.224	0.018	0.663	Probability that two random members differ in primary language.
Religious Fractionalisation	0.558	0.241	0.005	0.824	Probability that two random members differ in religious affiliation.
City (1 if metropolitan study)	0.620	0.486	0.000	1.000	Indicator equal to one if the study was conducted in one or a few major cities of the host country.
Within (1 if paired design)	0.641	0.481	0.000	1.000	Indicator equal to one if the study used a within-listing paired design.
Colony (1 if former colonial power)	0.356	0.480	0.000	1.000	Indicator equal to one if the host country is the United Kingdom, France, Portugal or Spain.
Population ($\times 100$ million)	1.729	2.866	0.003	14.117	Total population of the host country, scaled by 100 million.
Homeownership rate	0.664	0.076	0.420	0.897	Share of host-country households that own rather than rent their primary residence.

Notes. Statistics for Panel A computed across 295 OR-bearing minority sub-estimates from 54 studies. Statistics for Panel B are computed at the same row level, so a country-level variable such as ethnic fractionalisation is implicitly weighted by the number of sub-estimates a country contributes. Fractionalisation indices are bounded between 0 and 1.

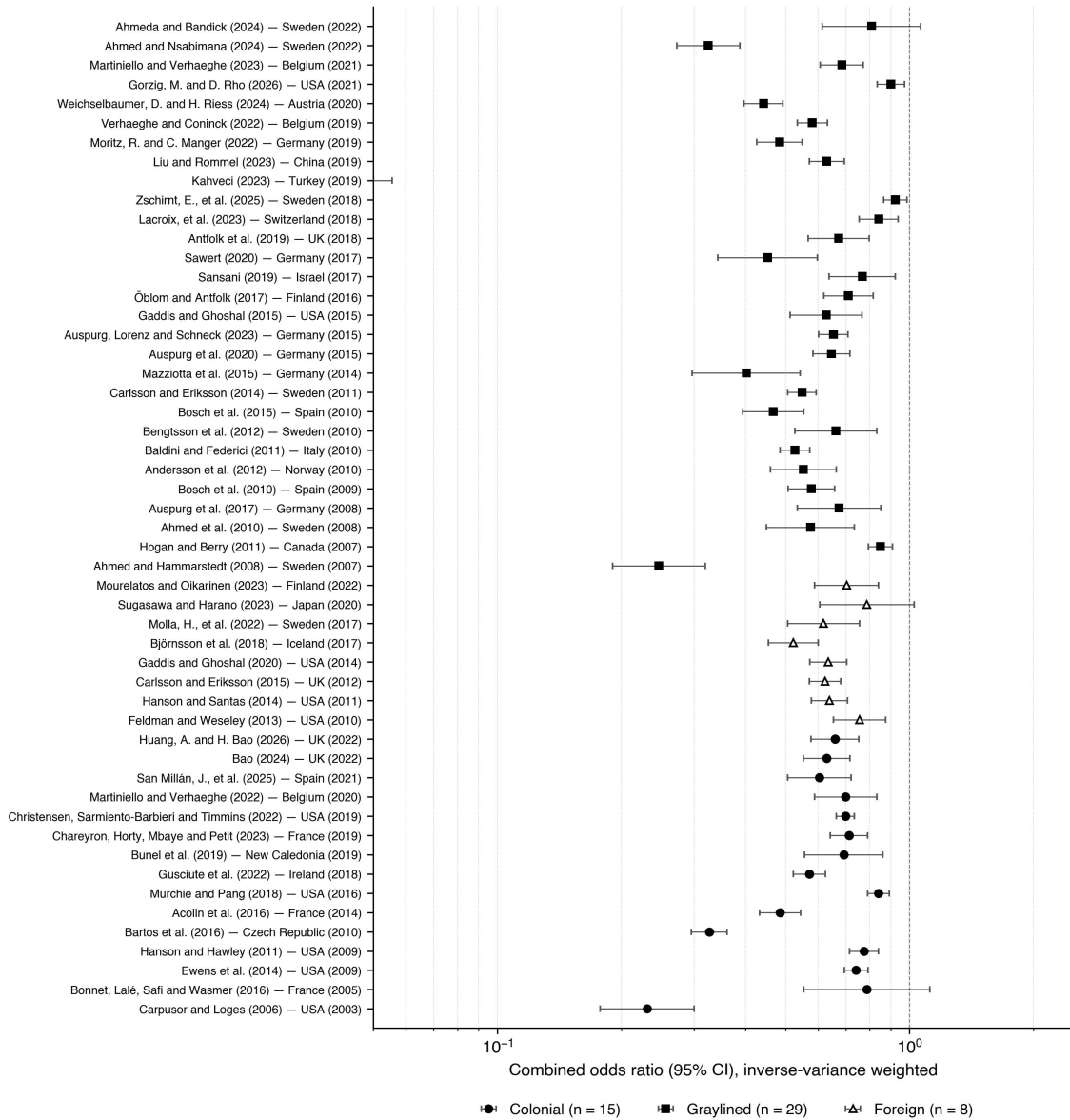


Figure 2: Forest plot of all fifty-four studies, ordered by primary-category assignment and year of fieldwork. Each marker is the inverse-variance-weighted combined odds ratio across all sub-estimates of the study. Distinct shape markers identify the three categories so the figure remains legible in greyscale.

effects and restricts the sample to gender-disaggregated estimates, produces a colonial coefficient of approximately -0.15 and a graylined coefficient that is not statistically distinguishable from zero. The contrast between the simple and saturated specifications is itself analytically informative, and its informativeness has increased as the sample has grown. In the 2024 sample the saturated-specification graylined coefficient was -0.19 and statistically significant at the one per cent level. In the 2025 expansion it fell to -0.03 . In the present 2026 sample it is approximately zero in both point estimate and statistical significance. The simple-specification estimates separate the three categories cleanly, with the foreign category as the omitted baseline. The saturated-specification estimates absorb a substantial portion of the between-category variation into the continent fixed effects and into the colonial-history indicator, with the result that the graylined coefficient loses identification against the foreign baseline. The progressive collapse of the graylined-versus-foreign coefficient as the sample expands is the empirical signature of the under-identification concern raised in Section 2.3, namely that the within-design correlations between the three categories make the meta-regression unable to recover their separate effects without simultaneous within-study variation across all three.

The ethnic-fractionalisation coefficient is negative in most specifications of the updated sample, in contrast with the positive estimates of the 2024 paper. The reversal is sensitive to the inclusion of studies in which the within-host-country diversity index is uncorrelated with the local response-rate gap, and it should not be interpreted as a substantive policy result without dedicated study-level identification. The former-colonial-power coefficient remains negative across specifications and is comparable in magnitude to the colonial-indicator coefficient, which is consistent with the colonial mechanism reflecting both individual-level and country-level effects. The coefficient estimates from the baseline specification, alongside two robustness variants, are reported in Table 2.

3.4 Mechanism-specific heterogeneity

The pooled estimates of Section 3.2 establish a category ordering but do not identify the mechanisms that produce it. We address this by exploiting the three host-country fractionalisation indices reported in the cross-study record, namely the Ethnic, Linguistic and Religious Fractionalisation measures of Alesina et al. (2003). Each index measures the probability that two randomly drawn members of the host society belong to different groups along the relevant dimension. The three indices map onto the three categories of the tripartite classification in a manner that is theoretically pre-specified rather than data-mined. Colonial-type discrimination is grounded in the visible-race hierarchies that colonial migration produced, and its expected moderator is ethnic fractionalisation. Graylined-type discrimination operates through religious or communal symbolism, and its expected moderator is religious fractionalisation. Foreign-type discrimination operates through cultural distance, accent and name-language unfamiliarity, and its expected moderator is linguistic fractionalisation. A regression that interacts each category indicator with its theoretically-paired index therefore provides an internal validity test for the tripartite classification. Each interaction takes the predicted positive sign if higher diversity along the relevant dimension dampens the category-specific penalty.

Table 2: Baseline meta-regression of odds ratios on tripartite indicators and controls.

	(1) Simple	(3) Gendered + controls	(5) Robustness (log scale)
Colonial	-0.101** (0.051)	-0.087** (0.044)	-0.152*** (0.038)
Graylined	-0.057 (0.049)	0.013 (0.059)	-0.004 (0.052)
Male (vs. Both/Female baseline)	-0.112** (0.055)	-0.180*** (0.056)	-0.265*** (0.054)
Year (centred at 2005)		0.021*** (0.005)	
ln(Year – 2005)			0.195*** (0.047)
Sample size ($\times 1,000$)		-0.052 (0.046)	
ln(Sample size)			-0.062** (0.025)
City		0.021 (0.051)	0.048 (0.049)
Ethnic Fractionalisation		-0.140 (0.276)	0.091 (0.211)
Former colonial power		-0.023 (0.053)	-0.091 (0.056)
Within-design		-0.033 (0.043)	-0.068 (0.046)
Population ($\times 100m$)		-0.082*** (0.019)	-0.107*** (0.019)
Colonial \times Male			0.155*** (0.048)
Graylined \times Male			0.038 (0.049)
Continent fixed effects	No	Yes	Yes
Observations	302	126	126
Adjusted R^2	0.071	0.260	0.279

Notes. OLS coefficients with paper-clustered standard errors in parentheses. Column (1) is the unconditional specification on the full minority sample; column (3) restricts to gender-disaggregated estimates and adds country and study-level controls; column (5) is a log-scale robustness variant of (3). The omitted category for the tripartite indicators is the rows that have neither Colonial nor Graylined flagged, which is the foreign-only baseline. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The estimates from this specification, reported in column A of Table 3, support the tripartite reading on all three counts. The Colonial-by-Ethnic-Fractionalisation interaction is 0.31 and is statistically significant at the five per cent level. The Graylined-by-Religious-Fractionalisation interaction is 0.63 and is statistically significant at the one per cent level. The Foreign-by-Linguistic-Fractionalisation interaction is 0.46 and is statistically significant at the five per cent level. The three interactions are jointly different from zero and they each take the sign that the theoretical mechanism predicts. A one-standard-deviation increase in the relevant index reduces the category-specific penalty by an amount equivalent to between twenty and forty per cent of the unconditional category coefficient, depending on the category. We treat this triple finding as the central empirical contribution of Section 3 to the agenda-setting argument. The three categories are not merely empirical labels that the pooled means can be sorted into, but distinct mechanisms that respond to distinct features of the host-country institutional environment.

A symmetric specification interacts each category with year of fieldwork, reported in column B. The category-by-year interactions are uniformly small and statistically indistinguishable from zero, with point estimates that cluster around -0.03 odds-ratio units per decade. The interpretation is that the three categories have evolved at similar rates over the 2003 to 2022 fieldwork window, which is informative in its own right. It rules out the alternative reading that the category ordering is an artefact of differential time trends and supports the structural-rather-than-historical interpretation that the Kuala Lumpur evidence in Section 4.1 develops further. Figure 3 visualises the underlying scatter and the within-category trend lines.

A robustness specification in column C decomposes the sample by cross-category overlap, with each row assigned to a unique cell based on its (Colonial, Graylined, Foreign) flag pattern. The estimable cells are Colonial-only, Graylined-only, Foreign-only, Graylined-Foreign and All-three; the Colonial-Graylined and Colonial-Foreign cells do not appear with sufficient frequency in the updated data to support separate estimates. With Colonial-only as the reference category, the Graylined-only cell has an OR that is 0.08 higher ($p < 0.10$), the Foreign-only cell has an OR that is 0.19 higher ($p < 0.01$), the Graylined-Foreign cell has an OR that is 0.29 higher ($p < 0.05$), and the All-three cell has an OR that is 0.15 higher ($p < 0.01$). The Colonial-only cell therefore carries the most severe penalty in the data, and the multi-flag cells do not produce a sub-additive reduction in OR relative to Colonial-only. The pattern is consistent with the reading that colonial-type discrimination operates as a distinct and particularly severe stigma in the markets the literature has sampled, rather than as a residual category that overlaps with the other two without retaining identifiable content.

3.5 Robustness to publication selection

The cross-study sample is drawn from peer-reviewed correspondence-test publications, and the selection of estimates into the published record is plausibly non-random. We assess the magnitude of any publication-selection bias using the precision-effect test and the funnel-asymmetry test of Stanley (2008) and the PEESE refinement of Stanley and Doucouliagos (2014). The FAT-PET specification regresses the study-level log odds ratio on its standard error, with paper-level clustered standard errors. A non-zero slope coefficient on the standard error indicates funnel

Table 3: Mechanism-specific heterogeneity: category-by-fractionalisation, category-by-year, and cross-category overlap.

	(A) Fractionalisation	(B) Year trajectories	(C) Overlap
Colonial	-0.103 (0.065)	0.433 (0.651)	
Graylined	-0.291*** (0.111)	0.465 (0.641)	
Foreign	0.033 (0.078)	0.611 (0.658)	
Ethnic Fractionalisation	-0.392 (0.404)	-0.007 (0.149)	
Linguistic Fractionalisation	0.168 (0.412)		
Religious Fractionalisation	-0.212 (0.158)		
Colonial \times Ethnic Fract.	0.306* (0.157)		
Graylined \times Religious Fract.	0.626*** (0.180)		
Foreign \times Linguistic Fract.	0.464** (0.199)		
Year (centred at 2005)	0.010 (0.006)	0.037 (0.038)	
Colonial \times Year		-0.030 (0.039)	
Graylined \times Year		-0.025 (0.038)	
Foreign \times Year		-0.029 (0.041)	
Graylined only (vs. Colonial only)			0.080* (0.044)
Foreign only (vs. Colonial only)			0.190*** (0.054)
Graylined + Foreign (vs. Colonial only)			0.286** (0.129)
All three (vs. Colonial only)			0.149*** (0.038)
Male indicator + base controls + continent FE	Yes	Yes	Yes
Observations	283	283	283
Adjusted R^2	0.234	0.190	0.195

Notes. OLS coefficients with paper-clustered standard errors in parentheses. Column (A) interacts each tripartite indicator with its theoretically-paired fractionalisation index and is the central empirical contribution of the section. Column (B) interacts each indicator with year of fieldwork. Column (C) decomposes the sample into mutually exclusive cells defined by the (Colonial, Graylined, Foreign) flag pattern, with Colonial-only as the omitted baseline. The Colonial-Graylined and Colonial-Foreign cells in column (C) are not estimable because no row in the data carries that flag pattern. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

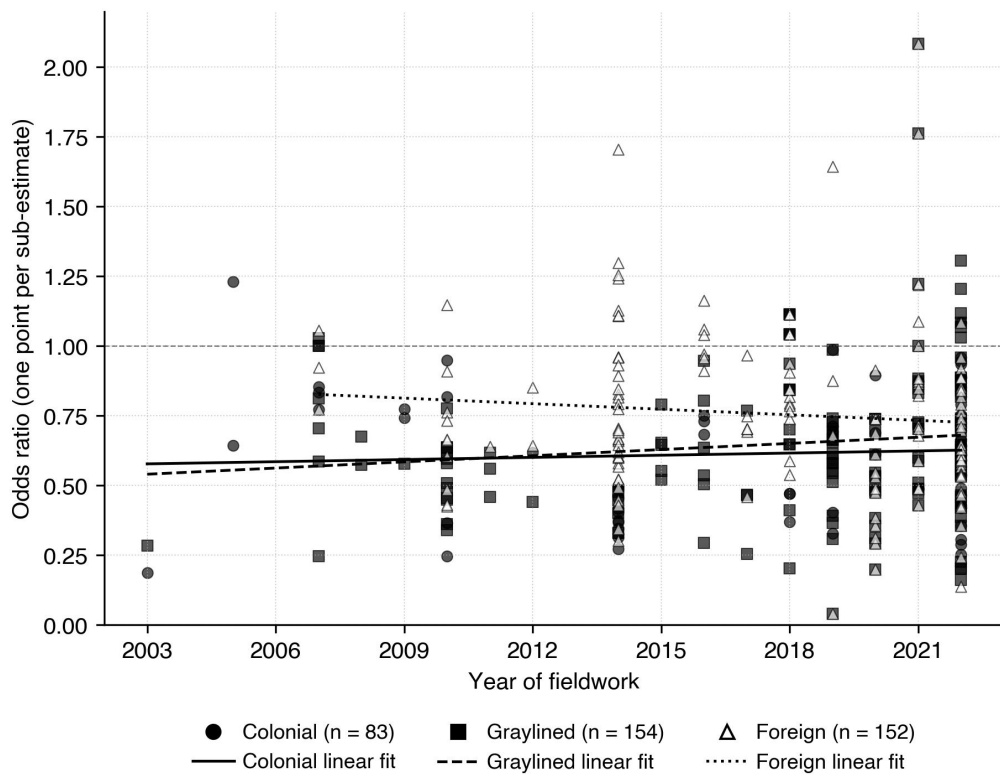


Figure 3: Odds ratios by year of fieldwork, with separate linear trend lines for each tripartite category. Each marker is one sub-estimate. The trend lines are visually close to flat for all three categories, consistent with the regression evidence that the category-by-year interactions are statistically indistinguishable from zero.

asymmetry of the kind that publication selection would produce, and the intercept estimates the effect-size that the literature would report in the absence of selection. The PEESE variant replaces the linear standard-error term with its square and is the preferred precision-effect estimator when the FAT rejects the null of no selection.

Three results follow. First, the funnel-asymmetry coefficient on the pooled minority sample is negative and statistically indistinguishable from zero, with a point estimate of -0.47 (paper-clustered standard error 0.41 , $p = 0.24$). The 284 OR-bearing minority observations across 52 papers therefore do not provide statistically significant evidence of publication selection. Second, the PET-adjusted intercept implies an odds ratio of 0.69 , which is approximately 0.07 odds-ratio units above the unweighted pooled value of 0.62 . The PEESE refinement yields a similar implied odds ratio of 0.66 . Both adjusted estimates are statistically significant at the one per cent level and confirm that the pooled discrimination effect is not a publication-selection artefact. Third, when the FAT-PET is run separately for each category, none of the within-category funnel-asymmetry coefficients is statistically distinguishable from zero, and the PET-adjusted ordering of the three categories (Colonial 0.59 , Graylined 0.68 , Foreign 0.80) preserves the descriptive ordering reported in Section 3.2. Figure 4 reports the funnel plot. The category ordering and the three-mechanism interpretation developed in Section 3.4 therefore survive selection-adjustment without modification.

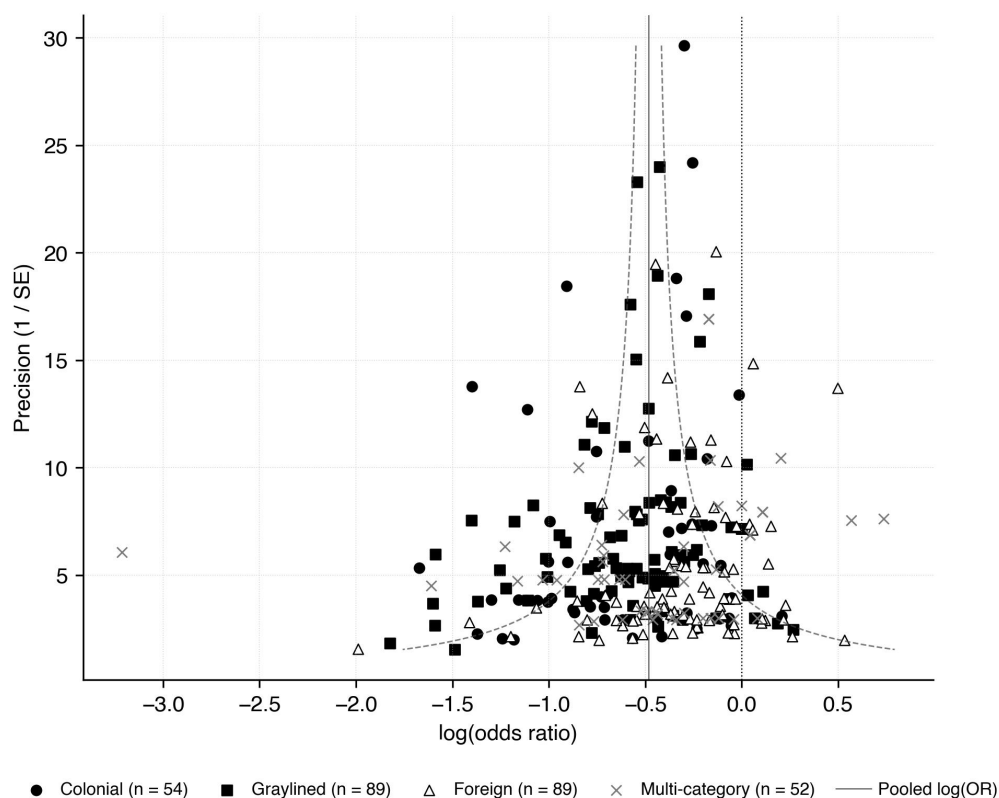


Figure 4: Funnel plot of study-level log odds ratios against precision, distinguishing the three tripartite categories. The vertical solid line is the unweighted pooled $\log(\text{OR})$; the dashed curves enclose the pseudo 95 per cent precision-weighted band. Mild visual asymmetry is not statistically significant in the FAT-PET specification ($p = 0.24$).

4 From tripartite cataloguing to a research agenda on social closure

Sections 2 and 3 organised the published correspondence-test record into three categories and provided meta-regression evidence that each category responds to its theoretically-paired country-level diversity index. The descriptive case for the tripartite classification can therefore be regarded as established within the bounds of the present evidence base. The purpose of this section is not to defend the classification further but to use it as a lens for setting a research agenda. The exercise begins with a short account of three recent empirical studies that complement the cross-study record, then proposes a structural reframing of housing-market discrimination, then sets out the lines of work that follow.

4.1 Three empirical signals from recent fieldwork

The first signal comes from a correspondence experiment in Singapore (Bao and Tan, working paper) in which all three categories of the tripartite classification are estimable simultaneously. The design includes ethnic-majority foreign profiles together with ethnic-minority and ethnic-majority domestic profiles. The aggregate foreign indicator returns an odds ratio of approximately 0.71 at the one per cent significance level. The British Chinese sub-group, which is ethnically majoritarian and foreign, faces an odds ratio of 0.70 against the Singaporean Chinese baseline. The within-Singaporean colonial contrast for Singaporean Indian applicants is not statistically distinguishable from one, while the within-Singaporean graylined contrast for Singaporean Malay applicants is approximately 0.77 at the five per cent level. The full results, including the multinomial decomposition that locates the foreign penalty at the response-versus-silence margin, are reported in the working paper.

The second signal comes from a preliminary correspondence experiment in Kuala Lumpur, conducted on the Mudah.com property platform between January and February 2025 and varying applicant attributes by Malaysian or non-Malaysian nationality, by Bumiputera, Chinese or Indian ethnicity, and by gender. Across the first wave of 1,104 inquiries the Bumiputera majority obtained response rates of 69 per cent against 62 per cent for Chinese Malaysian applicants, 53 per cent for Indian Malaysian applicants and 52 per cent for non-Malaysian foreigners. The unconditional odds ratios against the Bumiputera baseline are 0.73 for Chinese Malaysians (95 per cent confidence interval 0.49 to 1.10), 0.51 for Indian Malaysians (0.35 to 0.75) and 0.48 for non-Malaysian foreigners (0.34 to 0.67). The Bumiputera category is closely associated with the Malay-Muslim majority of Malaysia, and Muslim applicants appear in the cross-study record almost exclusively as targets of graylined discrimination in non-Muslim-majority countries. The Kuala Lumpur configuration therefore inverts the discriminator-target relation observed elsewhere.

The third signal is reported by Wang et al. (2026), which examines transaction-level outcomes in the Malaysian rental market and documents a price premium charged to non-Malay tenants by Malay landlords. The Wang et al. (2026) study does not use a correspondence-test design and is for that reason not included in the meta-regression sample. It is informative non-

etheless because it triangulates the discriminator role of the Malay majority through a separate methodology and a separate outcome variable, and it is the first published study in Malaysia in which the religious-majority status of the Malay landlord population is treated as the relevant axis of analysis.

4.2 Interpreting the evidence: identity economics with a closure complement

The empirical pattern documented in Section 3 and the three signals above admit an interpretation in standard economic terms. The category-by-fractionalisation interactions reported in Table 3 are consistent with each of the three mechanisms set out in Section 2.1. Higher ethnic fractionalisation reduces the colonial-category penalty, which is the prediction of both the taste-based and the statistical-discrimination models under the standard contact-hypothesis interpretation (Allport, 1954; Yinger, 1998): when minority and majority counterparts encounter one another routinely, the taste-based premium narrows and the precision of group-level inferences about counterpart attributes improves. Higher religious fractionalisation reduces the graylined-category penalty, which is consistent with identity-driven discrimination losing salience as the symbolic in-group becomes one of several comparably-sized religious communities rather than a numerical majority. Higher linguistic fractionalisation reduces the foreign-category penalty, which is consistent with the statistical-discrimination prediction that a more linguistically diverse market provides the discriminator with weaker group-level inferences from accent and name-language signals.

The Singapore, Kuala Lumpur and Malaysia signals are also consistent with the identity-economics framework of Akerlof and Kranton (2000, 2010), under which individuals derive utility from acting in accordance with the prescriptions of the groups they identify with, and group identity itself is context-dependent rather than essential to the individual. A British Chinese applicant in Singapore activates a foreign-applicant identity that the Singaporean Chinese landlord does not share, even though the two parties share an ethnicity. A Malay-Muslim landlord in Kuala Lumpur acts on a Bumiputera-majority identity that the Chinese or Indian applicant does not share, even though the same Malay-Muslim identity, in a non-Muslim-majority European country, would be the target of graylined discrimination rather than the discriminator. The identity-economics framework treats these inversions as the expected consequence of identity activation in different institutional settings rather than as anomalies that require new theory. Chen and Li (2009) provide complementary experimental evidence that even minimally-induced group identities reorganise the preferences of laboratory subjects in a manner consistent with the model, and Glaeser (2005) shows that political entrepreneurs invest in out-group hatred precisely because identity-driven preferences are responsive to context.

The economic interpretation has a complementary sociological lens that is worth recording briefly because it captures the aggregate consequence of repeated individual screening decisions. Weber (1922) introduced the term social closure for the mechanism by which dominant groups maintain monopoly access to valued resources by policing categorical group boundaries. Parkin (1979) and Murphy (1988) developed it as a general theory of exclusionary group action, and Tilly (1998) recast it as opportunity hoarding by categorically-defined groups whose members

allocate scarce opportunities preferentially among themselves. The closure and hoarding tradition is consistent with the identity-economics framework above and is best read as a macro-level description of what the micro-level model produces over many decisions in markets that permit discretionary screening. Housing markets, in which the resource is excludable and rivalrous and the screening decision is delegated to the resource holder, are a canonical setting in which the micro-foundations of the identity-economics model produce the aggregate closure pattern that the sociological tradition has long documented.

The principal implication for cumulative knowledge is that the correspondence-test literature has, to date, focused on identifying which target groups receive differential treatment and has paid less attention to identifying the conditions under which a group becomes the locally-dominant one. The combined identity-economics and closure reading suggests that the discriminator role is not pinned to particular identities but to particular positions within particular markets, and that the same identity can carry the discriminator role in one setting and the target role in another, depending on local institutional conditions and on the salience of each axis of group categorisation. The research agenda set out in Section 4.3 takes this reading as a working hypothesis to be tested rather than as an established result.

4.3 A research agenda

Five lines of work follow from the reframing. The first line is the systematic estimation of all three categories within single designs. Singapore is currently the only setting in which this has been achieved at scale, and replication across multi-ethnic high-migration markets is the natural priority. Candidate settings include Hong Kong, the United Arab Emirates, Australia, Canada and Malaysia, all of which provide configurations in which the ethnic-majority and foreign-status dimensions can be separated by design. The methodological investment is modest because the experimental machinery is well established, and the cross-country evidence on the foreign-by-linguistic-fractionalisation interaction reported in Section 3.4 would be sharpened by sampling along the index rather than relying on the post-hoc variation available in the published record.

The second line is the systematic collection of landlord-side covariates. The published correspondence-test literature treats the listing as the unit of analysis and the landlord as an unobserved heterogeneity term. The Kuala Lumpur and Singapore evidence both indicate that landlord ethnicity, religion and citizenship status are first-order rather than ancillary, and the next generation of designs should record them where the platform permits. Where it does not, secondary-data linkages between listings and licensed-agent registries are likely to be productive. Recording the discriminator-side attributes is also necessary to test the social-closure reading of the evidence against alternative accounts that locate the mechanism in target-side characteristics alone.

The third line is the integration of correspondence-test evidence with measures of resource scarcity in the local rental market. Vacancy rates, eviction-protection regimes, public-housing supply and recent demand shocks from migration or refugee flows are candidate moderators. The structural reading predicts that tighter markets and weaker tenant protections amplify the category-specific penalties, and the prediction is testable as more correspondence-test waves

accumulate within countries over time. The published record contains very few within-country longitudinal designs, and registering one or more such designs alongside policy variation would substantially advance the literature.

The fourth line is the geographic extension of the literature beyond the OECD. The cross-study record currently covers twenty-two countries, of which fifteen are in the European Union. The absence of Latin American, sub-Saharan African and most Asian markets bears directly on the tripartite reading, because each of these regions contains configurations of in-group and out-group that the European and North-American sample does not represent. The first South American or sub-Saharan African correspondence test in housing would, on the present argument, be an immediate publication priority.

The fifth line is methodological standardisation. The cross-study record pools studies that differ on outcome definitions, response-window lengths, message-content randomisation conventions and category-assignment rules. A minimal common protocol, comparable to the registered-report standards now established in social-science correspondence work, would substantially improve cross-study comparability and would reduce the reliance on continent fixed effects as a residual heterogeneity catch-all. The protocol should be permissive about the local axes on which discrimination is measured, in order to accommodate the structural rather than essentialist reading set out above, while standardising the inquiry-stage outcome measures and the response coding.

5 Conclusion

Housing-discrimination correspondence tests have accumulated a substantial body of evidence on differential treatment in rental markets. The earlier 2024 analysis identified three analytical categories that organise the evidence: colonial, graylined and foreign. The present paper anchors the three categories in the canonical economic theory of discrimination, with each category theoretically tied to a different mixture of taste-based, statistical and identity-driven mechanisms (Becker, 1957; Phelps, 1972; Arrow, 1973; Akerlof and Kranton, 2000). It extends the cross-study sample to fifty-four correspondence-test studies across twenty-two countries and provides meta-regression evidence that each category responds to its theoretically-paired country-level diversity index, with all three interactions taking the predicted positive sign and statistical significance at conventional thresholds. The pattern is robust to a FAT-PET correction for publication selection. We read the typology as a catalogue of the axes along which group boundaries are drawn in housing markets, rather than as a typology of group essences. The aggregate consequence of repeated individual-level screening decisions in markets that permit discretionary screening is the social-closure pattern that the sociological tradition has long documented (Weber, 1922; Tilly, 1998). The research agenda that follows includes the simultaneous estimation of the three categories in further markets, the systematic collection of landlord-side covariates, the integration of correspondence-test evidence with measures of resource scarcity in the local rental market, the geographic extension of the literature beyond the OECD, and the development of a minimal common protocol for cross-study comparability.

References

- Ahmed, A. M., Andersson, L. and Hammarstedt, M. (2010). Can discrimination in the housing market be reduced by increasing the information about the applicants? *Land Economics*, 86(1), 79–90.
- Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *Quarterly Journal of Economics*, 115(3), 715–753.
- Akerlof, G. A. and Kranton, R. E. (2010). *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*. Princeton University Press.
- Allport, G. W. (1954). *The Nature of Prejudice*. Addison-Wesley.
- Altonji, J. G. and Pierret, C. R. (2001). Employer learning and statistical discrimination. *Quarterly Journal of Economics*, 116(1), 313–350.
- Antfolk, J., Szala, A. and Oblom, A. (2019). Discrimination based on gender and ethnicity in English and Polish housing markets. *Journal of Community and Applied Social Psychology*, 29(3), 222–237.
- Arrow, K. J. (1973). The theory of discrimination. In Ashenfelter, O. and Rees, A. (eds), *Discrimination in Labor Markets*, 3–33. Princeton University Press.
- Auspurg, K., Schneck, A. and Thiel, F. (2020). Different samples, different results? How sampling techniques affect the results of field experiments on ethnic discrimination. *Research in Social Stratification and Mobility*, 65, 100444.
- Bao, H. X. H. (2024). Is hiding my first name enough? Using behavioural interventions to mitigate racial and gender discrimination in the rental housing market. *Urban Studies*. DOI 10.1177/00420980241231712.
- Bao, H. X. H. and Tan, N. Y.-S. (forthcoming). Discrimination and the social construction of foreignness in rental housing markets. Working paper.
- Becker, G. S. (1957). *The Economics of Discrimination*. University of Chicago Press.
- Blalock, H. M. (1967). *Toward a Theory of Minority-Group Relations*. John Wiley and Sons.
- Bunel, M., Gorohouna, S., L’Horty, Y., Petit, P. and Ris, C. (2019). Ethnic discrimination in the rental housing market: an experiment in New Caledonia. *International Regional Science Review*, 42(1), 65–97.
- Carlsson, M. and Eriksson, S. (2015). Ethnic discrimination in the London market for shared housing. *Journal of Ethnic and Migration Studies*, 41(8), 1276–1301.
- Chen, Y. and Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, 99(1), 431–457.

- Doucouliagos, H. and Stanley, T. D. (2009). Publication selection bias in minimum-wage research? A meta-regression analysis. *British Journal of Industrial Relations*, 47(2), 406–428.
- Esses, V. M., Jackson, L. M. and Armstrong, T. L. (1998). Intergroup competition and attitudes toward immigrants and immigration: an instrumental model of group conflict. *Journal of Social Issues*, 54(4), 699–724.
- Flage, A. (2018). Ethnic and gender discrimination in the rental housing market: evidence from a meta-analysis of correspondence tests, 2006–2017. *Journal of Housing Economics*, 41, 251–273.
- Glaeser, E. L. (2005). The political economy of hatred. *Quarterly Journal of Economics*, 120(1), 45–86.
- Gorzig, M. and Rho, D. (2026). [Confirm full title and journal]. Field experimental evidence on housing discrimination against Somali-refugee applicants in the United States.
- Gusciute, E., Muhlau, P. and Layte, R. (2022). Discrimination in the rental housing market: a field experiment in Ireland. *Journal of Ethnic and Migration Studies*, 48(3), 613–634.
- Heckman, J. J. (1998). Detecting discrimination. *Journal of Economic Perspectives*, 12(2), 101–116.
- Huang, A. and Bao, H. X. H. (2026). [Confirm full title and journal]. Field experiment on housing discrimination in the United Kingdom.
- Moritz, R. and Manger, C. (2022). Somebody that I want to know: the non-monotonic effect of personality information on ethnic and gender discrimination in the market for shared housing. *Journal of Housing Economics*, 57, 101842.
- Murphy, R. (1988). *Social Closure: The Theory of Monopolization and Exclusion*. Clarendon Press.
- Parkin, F. (1979). *Marxism and Class Theory: A Bourgeois Critique*. Columbia University Press.
- Phelps, E. S. (1972). The statistical theory of racism and sexism. *American Economic Review*, 62(4), 659–661.
- Quillian, L. (1995). Prejudice as a response to perceived group threat: population composition and anti-immigrant and racial prejudice in Europe. *American Sociological Review*, 60(4), 586–611.
- San Millán, J., et al. (2025). [Confirm full title and journal]. Field evidence on housing discrimination in Spain.
- Sansani, S. (2019). Are the religiously observant discriminated against in the rental housing market? Experimental evidence from Israel. *Journal for the Scientific Study of Religion*, 58(2), 459–474.

- Sherif, M., Harvey, O. J., White, B. J., Hood, W. R. and Sherif, C. W. (1961). *Intergroup Conflict and Cooperation: The Robbers Cave Experiment*. University of Oklahoma Book Exchange.
- Stanley, T. D. (2008). Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection. *Oxford Bulletin of Economics and Statistics*, 70(1), 103–127.
- Stanley, T. D. and Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, 5(1), 60–78.
- Sugasawa, T. and Harano, K. (2026). [Confirm full title and journal]. Correspondence-test evidence on housing discrimination against foreign-named applicants in Japan.
- Tajfel, H. and Turner, J. C. (1979). An integrative theory of intergroup conflict. In Austin, W. G. and Worchel, S. (eds), *The Social Psychology of Intergroup Relations*, 33–47. Brooks/Cole.
- Tilly, C. (1998). *Durable Inequality*. University of California Press.
- Wang, Y. et al. (2026). The price of landlords’ racial and ethnic discrimination: analyzing the Malaysian rental housing market. *Journal of Urban Affairs*. [Confirm full title and volume/issue.]
- Weber, M. (1922 [1978]). *Economy and Society: An Outline of Interpretive Sociology*, ed. Roth, G. and Wittich, C. University of California Press.
- Weichselbaumer, D. and Riess, H. (2024). [Confirm full title and journal]. Field experiment on rental-housing discrimination in Austria.
- Yinger, J. (1998). Evidence on discrimination in consumer markets. *Journal of Economic Perspectives*, 12(2), 23–40.
- Zhu, Y., et al. (2026). [Confirm full title and journal]. Within-country symbolic discrimination in Chinese urban rental markets.
- Zschirnt, E., et al. (2025). [Confirm full title and journal]. Field experiment on rental-housing discrimination in Sweden.