

Retirement Protection Gaps under Longevity Risk: Designing Effective Matching Incentives

Evidence from a Multi-Stakeholder Survey in Asia

Yu Fu

Nanyang Business School

8th September, 2026

THE MOTIVATION

Asia is ageing faster, and further, than any region in history

An unprecedented transition

Asians aged 60+ double from 568M (2022) to 1.2 billion by 2050 — 25% of the population.

The family buffer (annuity) is fading

Low fertility and labour mobility erode the informal, intergenerational care that once absorbed old-age risk

Insurance becomes a social pillar

As public fiscal space narrows and household risk rises, private cover shifts from a financial product to structural infrastructure.

COMPRESSED AGEING

Longevity stretches a 10–15 year retirement into 20–30 years — with far less time to prepare than the West had.

Years to move from 10% → 20% aged (65+)



The West took generations — Asia has decades.

THE PRESSING ISSUE

A protection gap opens when needs outgrow the resources that fund them



Retirement income

Public pensions + private annuities & saving

Risk: outliving fixed, un-indexed income.

Healthcare

Public schemes + commercial health cover

Risk: co-pays & chronic costs

Long-term care

Public LTC + private LTC / riders.

Risk: a wide 'missing middle'.

MACRO LENS

Six dimensions of each public scheme — accessibility, universality, affordability, adequacy, equity, sustainability.

MICRO LENS

Post-retirement income vs. living, health and premium spend — by population segment (formal, informal, rural, women).

A four-country, multi-stakeholder perspective

Countries

China Japan Singapore Indonesia

- Different stages of population ageing
- Different public pension architectures
- Different roles for employers, insurers, and household saving

Stakeholders

- Working-age individuals
- Retirees and households
- Policymakers and public institutions
- Employers and financial-services providers

The project links system-level protection gaps to individual responses to retirement-saving policy design.

Research questions

1. Where do public systems and commercial markets leave retirement-protection gaps across Asian economies?
2. Can government matching incentives raise intended retirement saving?
3. Does a higher match rate generate a larger response?
4. How do an annual cap, a minimum contribution, and a lock-in condition change the response?
5. Who responds, and are responses associated with income, baseline saving, or policy recognition?

THE MULTI-PILLAR MODEL

Five pillars, five risks — every system is a different mix

PILLAR	FINANCING MECHANISM	RISK IT CARRIES	EXAMPLES · DIFFERENT MIXES
0 Basic floor non-contributory, tax-financed	State redistribution	Fiscal / political	SG Silver Support · China resident floor
1 Mandatory public contributory, earnings-related	Pay-as-you-go (workers → retirees)	Demographic & political	Japan National + Employees' Pension · Indonesia JP
2 Mandatory funded individual accounts (DC)	Self-financing (you → future self)	Market / investment	Singapore CPF · China occupational · Indonesia JHT
3 Voluntary funded personal / occupational top-up	Self-financing + pooling (annuities)	Market; low uptake	Japan iDeCo/NISA · China personal pension · SG SRS
4 Informal family, home, health	Family transfers — the informal 'annuity'	Fades as families shrink	Indonesia family support · rural China (thinning)

As ageing strains pillars 1 & 4, the weight must shift to the funded pillars 2–3 — **the insurer's opening**.

Retirement Protection in Asia · Who Pays Whom?

10

Public pensions: Broad coverage — but adequacy and equity vary sharply

	China	Japan	Singapore	Indonesia
Scheme / model	UEBP + Resident (dual-track)	National + Employees' (PAYG)	CPF (funded accounts)	JHT / JP / THT (fragmented)
Coverage	Universal (UEBP 38% / Res. 57%)	Universal (Employees' 33%)	Near-universal, mandatory	Limited (JHT/JP 13%)
Contribution	8% employees / 20% flexible	9.15%	Mandatory CPF	1–3.25%
Avg. monthly benefit	US\$438 / US\$28	US\$917 / US\$356	CPF-balance dependent	US\$152–170 (+ lump sums)
Replacement rate	91% / 6%	38% / 15%	balance-dependent	13–15%

The divide: Japan and Singapore are near-universal; China is dual-track (formal 91% vs resident 6%); Indonesia's monthly pension is thin, cushioned by lump sums.

Public health insurance: the four systems at a glance

	China	Japan	Singapore	Indonesia
Model	Dual-pillar social insurance	Multi-tier social insurance	3M medical-savings model	Single national pool (JKN)
Coverage	94.6%	>99% (near-universal)	100% citizens / PRs	98.5%
Contribution	2% (informal 8%)	9.2%	Age-based premiums	1% or fixed
Out-of-pocket ratio	15% / 31% (resident)	19% (8% late-stage)	28.4%	61.8%
OOP / income	1.4–1.9%	1.7%	n.a.	9.8%

Near-universal — but adequacy diverges: Indonesia's out-of-pocket is 62% of health spend (9.8% of income); Singapore's premiums climb steeply with age.

Public long-term care: the four systems at a glance

	China	Japan	Singapore	Indonesia
Public scheme	LTCI pilots	Universal LTCI (since 2000)	CareShield Life	None — family-based
Coverage (% of population)	13% (pilot cities)	62.9% (all 40+ insured)	33% (covered cohorts)	0
Benefit trigger	Severe disability (Gr. 3–5)	Support / nursing need	≥3 of 6 ADLs	—
Monthly benefit	US\$185 (Beijing max 417)	US\$154	US\$628	—
OOP / income	16.5%	23.6%	n.a.	100% (informal care)

The least-developed pillar: Japan alone runs mature universal LTCI; China's pilot covers 13%; Singapore pays the highest cash benefit but for severe disability only; Indonesia has no public LTC.

Four archetypes — who bears the residual risk

China	Japan	Singapore	Indonesia
Broad basic net, thin depth Near-universal but stratified; dual-track urban / rural; LTC still in pilots.	Mature, integrated, ageing Comprehensive public architecture incl. universal LTC — but fiscally strained.	Pre-funded self-reliance Mandatory CPF + CareShield Life; strong backstop, deliberate residual risk.	Early-stage, informal JKN gives a health baseline; little annuitisation; no national LTC scheme.
KEY TENSION Adequacy & structural inequality	KEY TENSION Fiscal & intergenerational burden	KEY TENSION Adequacy tied to lifetime balances	KEY TENSION Informality & missing LTC pooling

Across all four, gaps arise not from the absence of systems but from how responsibility is split between state, market and household.

Matching experiment: common decision environment

Decision

Respondents choose what percentage of annual income to contribute to a government retirement account.

- A 5% default saving rate is pre-set
- Respondents may keep, reduce, or increase it
- The same response categories are used across scenarios

Experimental presentation

- Seven policy scenarios
- Presented in randomized order
- Each scenario framed as a separate decision
- A choice in one scenario does not affect another

Source: Survey instrument. Randomized scenario order was used in the Singapore and Japan questionnaires.

Seven independently presented policy scenarios

Scenario	Policy feature
1 No government match	No minimum, cap, or lock-in
2 25% flat match	SGD 25 per SGD 100 saved
3 50% flat match	SGD 50 per SGD 100 saved
4 Income-based match	100%, 50%, or 25% according to income
5 Income-based match + cap	Government match capped at SGD 3,600 per year
6 Income-based match + cap + minimum	Save at least 3% of income to qualify
7 Full design + time condition	Withdrawal within five years forfeits the match

Sample profile

Analytical sample

2800

respondents

- Complete relative-income classification
- Six groups defined relative to average income

Income relative to average	Share
Below 60%	23.1%
60–80%	17.5%
80–120%	25.1%
120–200%	19.7%
200–300%	9.0%
300%+	5.6%

Outcome construction

Saving-rate categories

0, 1–4, 5, 6–9, 10–14, 15–19, 20% or more

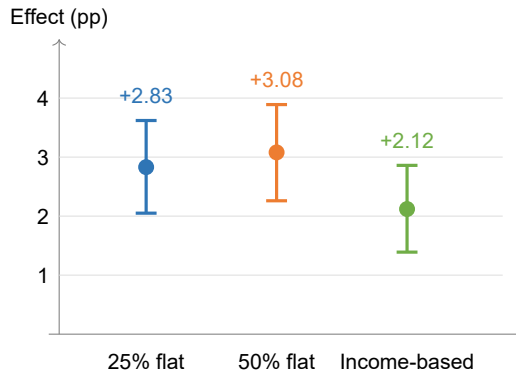
- Analysis uses category midpoints
- Numeric scale: 0, 2.5, 5, 7.5, 12, 17, 20
- “Not sure” is excluded from numeric outcomes

Within-respondent contrasts

$$\Delta_{ip} = S_{ip} - S_{i0}$$

- Main matching effects compare each design with no match
- Feature effects compare scenarios differing by one policy condition
- Effects are measured in percentage points

Matching incentives raise intended saving



- 25% flat match: **+2.83 pp**
- 50% flat match: **+3.08 pp**
- Income-based match: **+2.12 pp**

Inference

All three effects are significant under t-tests, Wilcoxon tests, and Holm adjustment: $p < 0.001$.

Error bars show 95% confidence intervals.

Doubling the flat match rate adds little

50% versus 25% flat match

+0.21 pp

95% CI: -0.34 to 0.76

t -test $p = 0.448$; Wilcoxon $p = 0.323$

- Both flat match designs raise intended saving
- The incremental gain from doubling the match rate is small
- Flat matching outperforms income-based matching by 0.83 – 0.98 pp

A simpler 25% flat match achieves nearly the same average response as a 50% flat match.

Additional policy conditions attenuate the response

Design relative to no match	Effect	95% CI	<i>p</i>	Holm <i>p</i>
Income-based match	+2.12 pp	1.39 to 2.86	< .001	< .001
+ annual cap	+0.95 pp	0.28 to 1.61	.005	.011
+ 3% minimum contribution	+0.28 pp	−0.36 to 0.92	.397	.397
+ five-year lock-in	+1.05 pp	0.39 to 1.71	.002	.005

Incremental feature effects

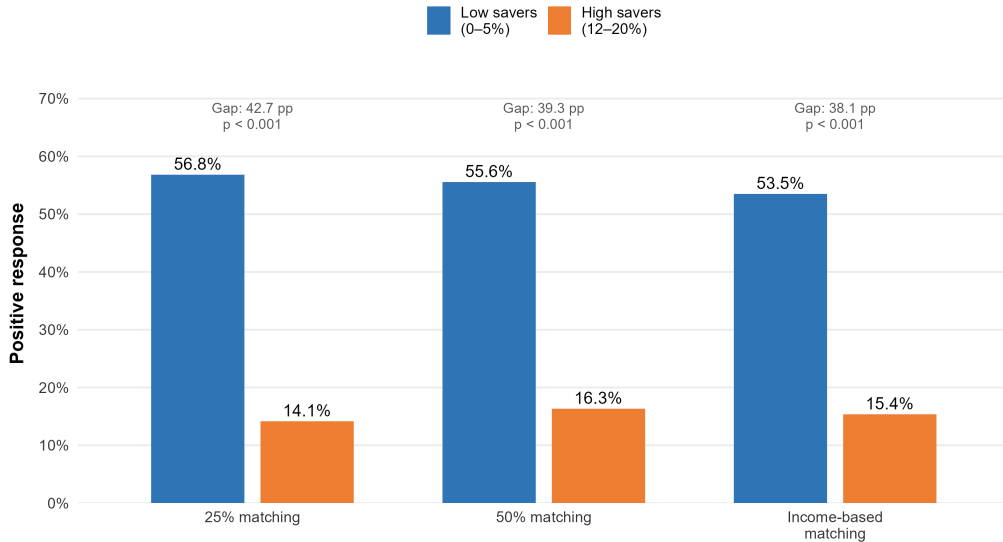
Cap: −1.17 pp Minimum: −0.72 pp
Time condition: +0.70 pp

Interpretation boundary

These are within-respondent contrasts between independently presented scenarios, not a temporal sequence of policy exposure.

Low baseline savers are more likely to respond

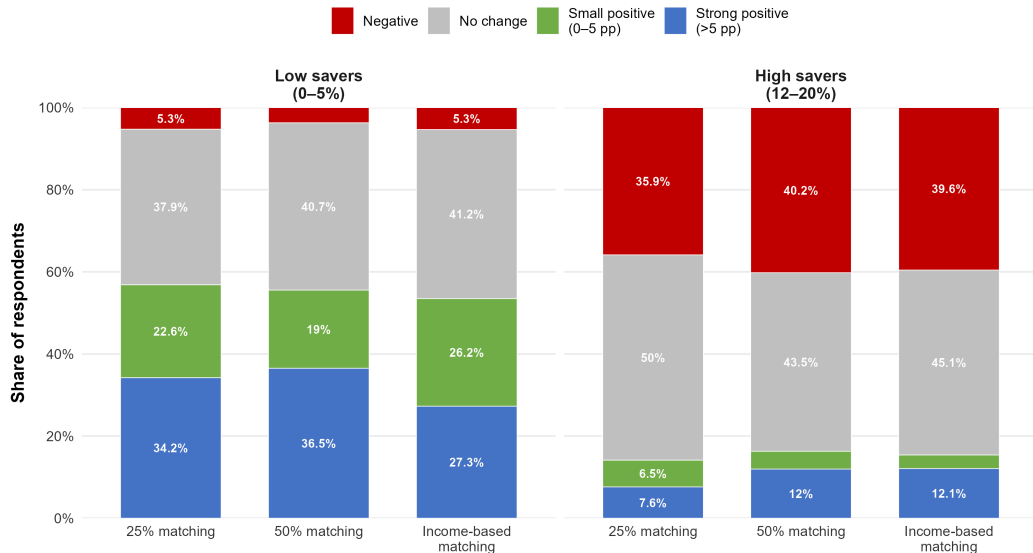
Share increasing saving relative to the no-policy baseline



Notes: A positive response is a saving-rate increase above zero. Reported gaps are Low minus High. P-values are from two-sample proportion tests.

Policy responses differ sharply in composition

Low baseline savers show more positive responses, while High savers show more negative responses



Notes: Small positive responses are above 0 and no more than 5 percentage points. Strong positive responses exceed 5 percentage points.

Conditional response magnitude and mechanical headroom

	Mean effect among responders		
	25%	50%	Income
Low baseline	9.66	10.40	8.49
High baseline	6.15	7.07	7.07
Mean gap	3.51	3.34	1.42
<i>t</i> -test <i>p</i>	< .001	< .001	.069
Wilcoxon <i>p</i>	.295	.373	.441

Upward headroom

Low baseline: **16.36 pp**

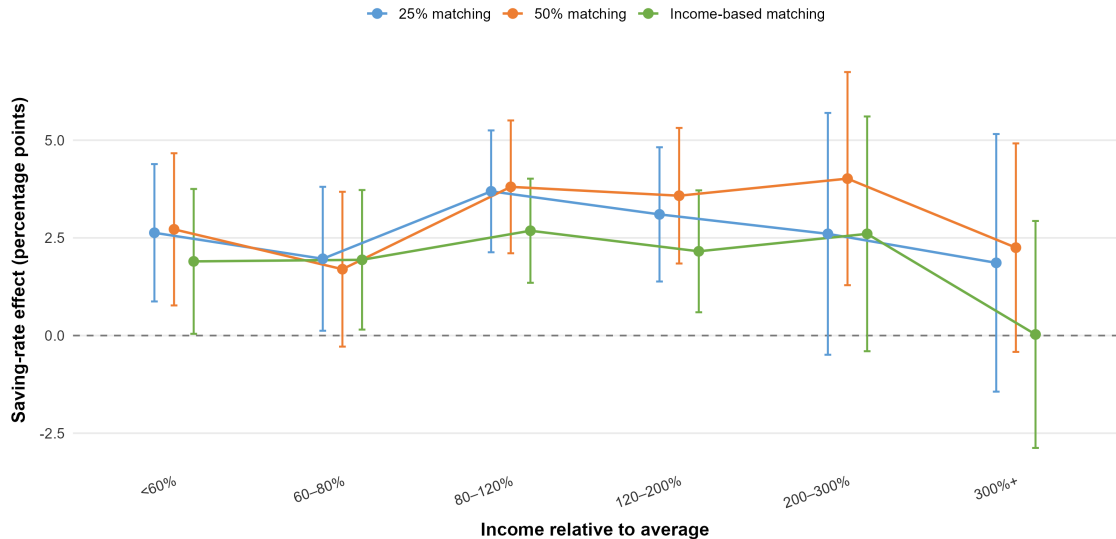
High baseline: **4.38 pp**

- Mean differences under flat matching are not matched by rank-distribution differences
- Low-group P90 reaches 15–20 pp; High-group P90 is 8 pp

The upper-tail pattern combines behavioral heterogeneity with mechanically different adjustment ranges.

Average matching effects across income groups

Points show group means; error bars show 95% confidence intervals



Notes: The figure is descriptive. Formal income-group heterogeneity is assessed using classical and Welch ANOVA in Module 06.

No detectable heterogeneity across income groups

Classical ANOVA

25% matching	$p = .768$
50% matching	$p = .554$
Income-based matching	$p = .751$

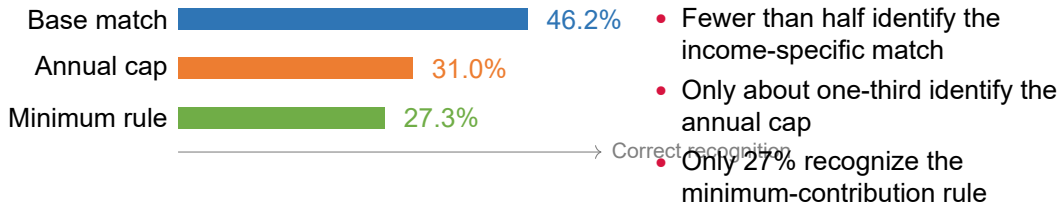
Welch ANOVA

25% matching	$p = .764$
50% matching	$p = .558$
Income-based matching	$p = .679$

- Point estimates vary, especially in smaller high-income groups
- No omnibus test detects systematic differences across the six groups
- All Holm-adjusted p -values equal 1.000

Response heterogeneity is much more pronounced by baseline saving than by relative income.

Conditional policy rules are not well recognized



Source: The final multi-select item is not shown because selecting a factually correct statement does not uniquely identify complete understanding of the time condition.

Recognition does not explain the weaker responses

Recognition measure	Correct–incorrect effect	95% CI	<i>p</i>
Base matching amount	+0.94 pp	−0.54 to 2.42	.211
Annual cap	−1.06 pp	−2.44 to 0.32	.131
Minimum contribution	+0.02 pp	−1.06 to 1.09	.974
Final factual-statement indicator	−1.01 pp	−2.29 to 0.27	.120

- No significant differences in mean policy effects
- No significant differences in positive-response probabilities
- No significant differences in strong-response probabilities after adjustment

Lower recognition and weaker responses coexist, but the data do not establish misunderstanding as the mechanism.

Robustness: positive direction is preserved

Specification	25% flat	50% flat	Income-based
Original midpoint coding	2.83***	3.08***	2.12***
Winsorized at 5th/95th percentiles	2.83***	3.15***	2.19***
Complete seven-scenario sample	2.81***	2.86***	2.02***
Exclude $ \Delta > 10$ pp	0.73***	0.76***	0.60*
Exclude $ \Delta > 5$ pp	0.29†	0.37*	0.47**

Threshold robustness

Low–High gaps for responses > 5 pp:

- 25% match: 26.6 pp, $p < .001$
- 50% match: 24.6 pp, $p < .001$
- Income-based: 15.2 pp, $p = .007$

Notes: *** $p < .001$; ** $p < .01$; * $p < .05$; † $p < .10$. Significance shown using raw two-sided t-tests.

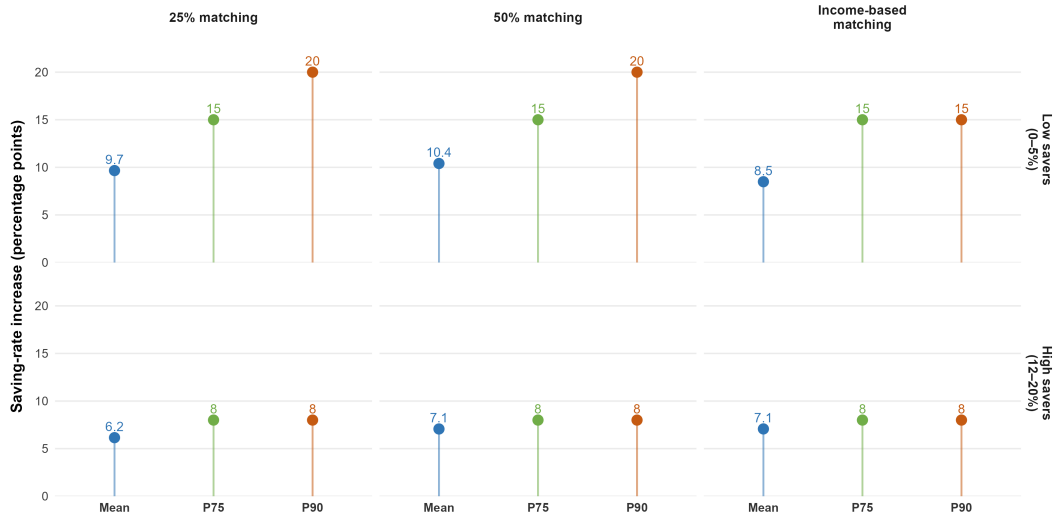
What changes?

Removing all changes above 5 pp removes a substantively meaningful responder group, not merely statistical outliers. The 25% estimate then falls to $p = .072$.

Response magnitude differs mainly in the upper tail

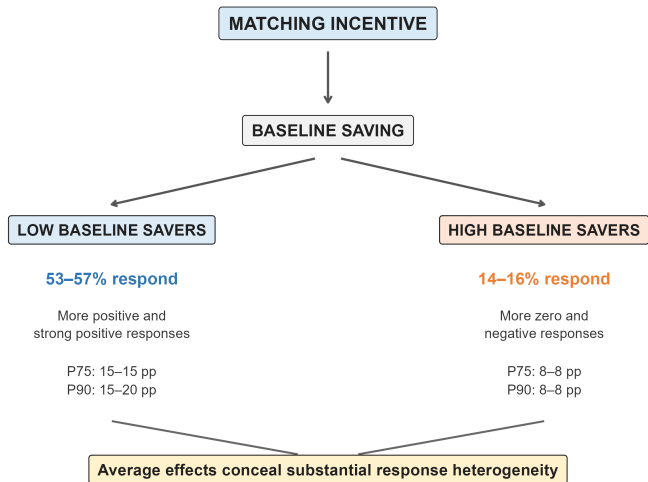
Saving-rate increases among positive responders only

Mean P75 P90



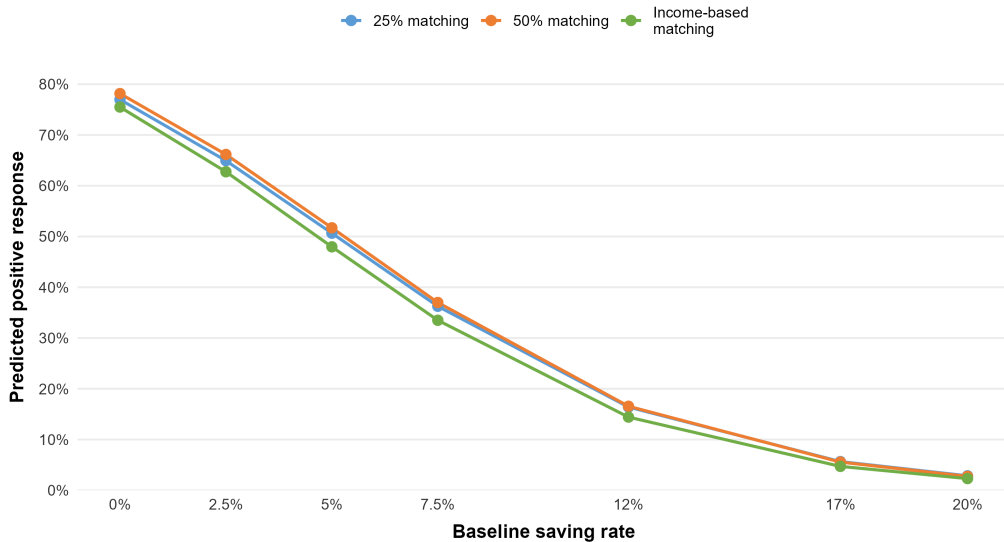
Notes: Mean, P75, and P90 are calculated only among respondents whose policy effect is above zero. The figure describes conditional response magnitude.

Who responds to matching incentives?



Response probability declines with baseline saving

Income-adjusted predictions from continuous-baseline logit models



Notes: Predictions are averaged over the observed income-group distribution. Baseline saving is entered as a continuous predictor in 5-percentage-point

Baseline-saving regressions

Policy	Low vs High OR	95% CI	OR per 5-pp baseline
25% matching	8.02	4.29 to 16.06	0.305
50% matching	6.67	3.64 to 12.93	0.293
Income-based matching	6.49	3.49 to 12.81	0.291

- Models adjust for relative-income group
- All displayed coefficients have $p < .001$
- Fully expanded models show signs of quasi-separation and are not emphasized
- Associations partly reflect the bounded outcome and baseline-defined response threshold

Direct matching-design comparisons

Comparison	Effect	95% CI	<i>p</i>	Holm <i>p</i>
50% vs 25% flat	+0.21 pp	−0.34 to 0.76	.448	.448
Income-based vs 25% flat	−0.83 pp	−1.47 to −0.18	.013	.025
Income-based vs 50% flat	−0.98 pp	−1.61 to −0.35	.002	.007

Recognition models

Recognition measure	Adjusted mean effect	Positive-response OR	Strong-response OR
Base matching amount	0.60 ($p = .370$)	1.28 ($p = .364$)	1.25 ($p = .445$)
Annual cap	-0.77 ($p = .214$)	1.17 ($p = .617$)	1.77 ($p = .313$)
Minimum contribution	-0.00 ($p = .998$)	0.68 ($p = .318$)	0.60 ($p = .440$)
Final factual statement	-0.75 ($p = .335$)	0.94 ($p = .862$)	0.38 ($p = .054$)

Linear models and logits adjust for baseline saving and relative-income group. Recognition was not randomized. The final multi-select indicator does not uniquely measure complete understanding of the time condition.

What the matching experiment shows

1. **Matching works**: intended saving rises by 2.1–3.1 percentage points.
2. **More is not necessarily better**: 50% and 25% flat matches generate statistically similar responses.
3. **Conditions can dilute incentives**: the cap and minimum requirement attenuate the response.
4. **Responses are heterogeneous**: increases concentrate among low baseline savers, although headroom matters.
5. **Income is not the main dividing line**: no systematic heterogeneity appears across relative-income groups.
6. **Complexity is visible but not explanatory**: recognition is limited, yet recognition does not predict the corresponding response.

Implications for retirement protection in Asia

Policy design

- Favor salience and simplicity
- Compare marginal benefit with administrative complexity
- Monitor whether eligibility rules deter participation

Market design

- Target gaps that public systems do not fully cover
- Use defaults and matching alongside communication
- Preserve flexibility for heterogeneous households

Effective retirement protection requires coordination between public incentives, household financial education, and commercially viable risk-pooling solutions.

Thank you!