

Diversification as a Service

Stanislav Sokolinski*

September 2026

Abstract

What would the asset-management industry sell if nobody could pick stocks? I study an industry in which investors agree about expected returns and risk, each investor can hold one asset for free and pays a fixed cost to trade any other, and funds sell standardized bundles of assets at proportional fees, each covering the same fixed operating cost, with free entry. When assets differ modestly in the common risk they carry and funds are cheap to run, competition produces a market fund that every investor holds. Every other fund must sell a correction to its risk mix, and one ratio, operating cost relative to the revenue that dispersion in risk loadings creates, determines how many such funds survive and how narrow they are. The paper explains why the broadest fund is the largest and the cheapest, why funds earn positive alpha after fees against it without any skill, why across fund types the smaller funds with higher active share earn more of it, and why a concentrated direct holding can sit inside a diversified portfolio.

Keywords: Asset management, diversification, mutual funds, performance evaluation.

JEL codes: G11, G12, G23.

*Broad College of Business, Michigan State University. Department of Finance, 315 Eppley Center, East Lansing, MI 48824. E-mail: sokolins@msu.edu. Web: <https://www.stanislavsokolinski.com>.

1 Introduction

The asset-management industry sells diversification in thousands of packages. At the end of 2025, U.S. investors could choose among 8,030 mutual funds and 4,813 exchange-traded funds ([Investment Company Institute, 2026](#)). The menu is layered. The first exchange-traded funds tracked broad indexes and offered diversification at low cost, while later products track niche portfolios and charge high fees ([Ben-David et al., 2023](#)). Across countries, larger funds charge lower fees, and index funds charge less than actively managed funds ([Khorana et al., 2009](#)). Many of these products hold rule-based portfolios of securities that investors could, in principle, buy themselves.

This paper asks what such an industry would sell if nobody could pick stocks. A natural conjecture is that competition among products that anyone can copy leaves one cheap index fund and little else. In the model I build, competition leaves a market fund that every investor holds and, beside it, up to two narrower funds. How many survive, and how narrow they are, depends on the covariance of returns and on a fixed cost per product. Which products survive? Why is the broadest product both the largest and the cheapest? What does a standard performance regression report about funds that have no skill? And which investors hold assets directly alongside their funds?

I answer these questions in a model of packaged diversification. Investors agree about expected returns and risk, and neither they nor any fund manager can pick stocks. The economy has four risky assets and four investor types. Each investor can hold one asset, her home asset, at no cost. To trade any other asset herself, she must pay a fixed access cost. Funds are standardized bundles of assets, sold at proportional fees. Every fund pays the same fixed operating cost, and entry is free. Because every fund holds assets that investors could trade themselves, funds add no new payoffs ([Allen and Gale, 1994](#)). Their service is to save investors the cost of assembling diversified positions, and a fund survives only if investors pay enough for that service to cover the cost of running it.

The risk structure of returns gives narrower funds something to sell. Returns have one common factor, and assets differ in their loadings on it. If all loadings were equal, an equal-weight portfolio of all assets would be the diversified portfolio every investor wants. When loadings differ, that portfolio holds too much of the high-loading asset, which carries the most common risk, and too little of the low-loading asset, the best diversifier. The two

middle assets have the average loading. Investors agree on the desired change in weights, which I call the correction. The correction grows with the dispersion of loadings, and so does what investors will pay for each unit of it, so the revenue a fund can earn by selling it grows with the square of dispersion. I measure the cost of running a fund against this revenue:

$$\phi = \frac{F}{\Theta(m) c^2}, \quad (1)$$

where F is the operating cost, c is the dispersion of loadings and $\Theta(m)$ depends only on the average loading m . For a given average loading, which narrower funds can survive depends on the primitives only through ϕ : doubling dispersion has the same effect as cutting operating cost to one quarter.

The first result is that competition produces a market fund. Every investor wants the same diversified portfolio, but each starts from a different home asset. A market fund, which holds every asset in equal weights, sells nearly that portfolio to all investors as one product. It spreads one operating cost over the entire common demand and can therefore charge a very low fee. Narrower funds can deliver the same exposure only in combination, and each fund in the combination must cover its own operating cost. When loading dispersion and operating cost are sufficiently low, a market fund can profitably enter whenever narrower funds alone supply the common demand, and it would profitably cut its fee to attract any investor who obtained her market exposure elsewhere. Every equilibrium therefore contains a market fund that every investor holds. No investor pays to trade other assets directly: the market fund already delivers nearly the portfolio she wants, and the remaining improvement is worth less than the fixed access cost.

Once the market fund is in place, other funds can sell only the correction. I call them specialists. A specialist lets an investor move exposure away from the high-loading asset and toward the diversifier, replacing part of her market-fund holding. I characterize the equilibrium organizations of the industry in a small-dispersion approximation, which keeps the leading terms of fees, holdings and revenues. All standardized products other than the market fund may enter, at any positive fee.

Three organizations survive, and ϕ selects among them through three cutoffs. Above the highest cutoff, the market fund operates alone. Between the middle and the highest, a pooled specialist operates beside it: one fund that holds every asset except the high-loading one

and serves investors with different home assets. Below the middle cutoff, a targeted pair of specialists operates, each fund combining the diversifier with one of the middle assets. The pair moves exposure exactly along the correction and lets investors tailor the middle weights, but it must finance two operating costs where the pooled fund finances one; between the lowest and the middle cutoff, either organization can operate. No other set of specialists survives, so every specialist in equilibrium leaves out the high-loading asset and holds the diversifier. Figure 1 draws the classification along the ϕ axis: the targeted pair at the low-cost end, the pooled specialist at the high-cost end, an overlap between them in which either can operate, and the market fund alone beyond the highest cutoff. Lower operating cost or greater dispersion moves an industry to the left, toward more specialists with narrower mandates, across economies or over time.

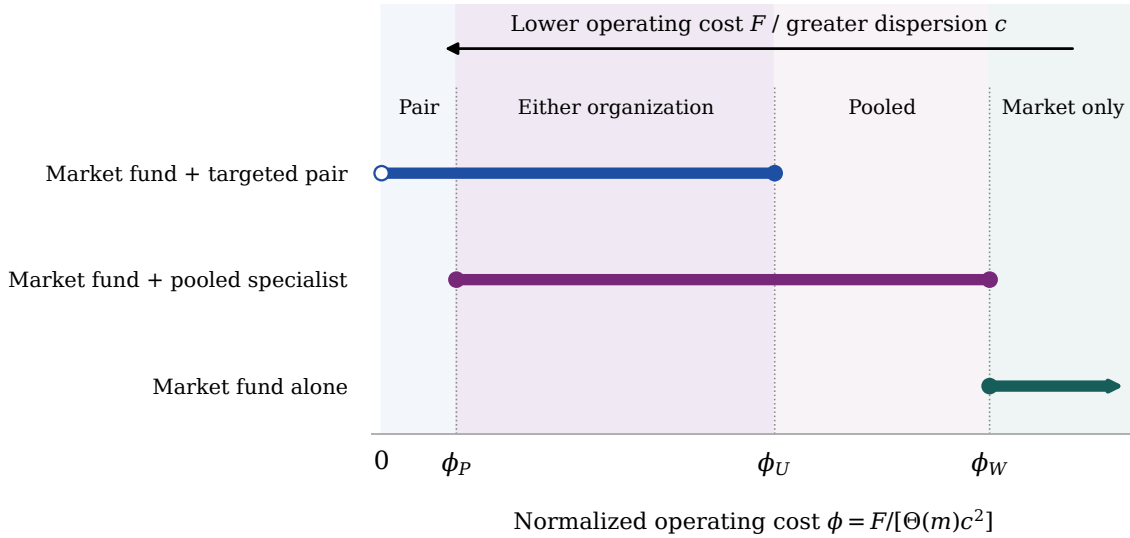


Figure 1: **Equilibrium fund variety.** Equilibrium organizations at each normalized operating cost ϕ in the small-dispersion approximation (Proposition 2). The market fund operates in every organization. Filled dots are included endpoints; the targeted pair requires positive cost. In the overlap $[\phi_P, \phi_U]$, the pooled and the targeted organizations are both equilibria; at ϕ_W , so are the market fund alone and the pooled organization. The arrow indicates that lower operating cost or greater loading dispersion lowers ϕ .

The equilibrium industry has implications for three things one can observe: fees and sizes, measured performance, and investors' portfolios. Start with fees and sizes. Free entry leaves each active fund's fee revenue exactly equal to its operating cost, so in every equilibrium fees

are inversely proportional to assets under management. The market fund sells the diversified position that every investor wants, while a specialist sells only an adjustment to it. The market fund is therefore much larger than any specialist and charges a much lower fee. In numerical solutions of the original economy at loading dispersions that U.S. stock returns suggest, it manages 4 to 25 times the assets of any specialist and charges a fee 4 to 25 times lower. Fixed operating costs alone favor large funds; demand explains why one fund is so much larger than the others. Observed fees fall far more slowly with size than this fixed-cost benchmark implies ([Khorana et al., 2009](#)); costs that grow with a fund's assets, which the model omits, would flatten the relation.

Next, performance. Every asset has the same expected return, so a fund's alpha against the market fund's return before fees comes from its beta alone. Each specialist carries less of the market fund's risk and earns positive alpha before fees, although nobody in the economy has skill. It keeps positive alpha after fees even though it only breaks even. The reason is on the demand side: an investor buys a specialist until its fee equals the value of one more unit to her, which is the fund's alpha less the extra risk the unit adds to her whole portfolio, so the fee stays below the alpha. In the numerical solutions, specialists keep 6 to 20 basis points of alpha a year after fees.

Against a costless portfolio with the fund's own weights, every fund instead has zero alpha before fees. Measured alpha therefore depends on the benchmark, in the model as in the data: a benchmark of index funds that investors can actually buy credits funds with the value of diversification services ([Berk and van Binsbergen, 2015](#)). The model gives an equilibrium account of that service: which specialists exist, what they charge, and how much alpha remains after fees.

Performance is also tied to how far a fund departs from the market portfolio. Measured against the market fund, the targeted funds depart furthest and the pooled fund less, and their alphas before fees rank in the same order. Where both organizations are equilibria at the same primitives, the pooled fund is larger than either targeted fund would be, and in every organization the market fund is the largest. Across fund types, smaller funds with higher active share therefore earn higher alpha. Studies of mutual funds find a related pattern and attribute it to stock-picking skill and to diseconomies of scale ([Cremers and Petajisto, 2009](#); [Chen et al., 2004](#)). In the model, the pattern reflects which correction a fund packages: a fund with fixed weights and return moments earns the same alpha before fees at any size.

Finally, portfolios. Households often combine concentrated direct stockholdings with mutual funds, which limits the losses from their concentrated stock accounts (Calvet et al., 2007). In the pooled organization, and in the targeted one when both pair funds charge the same fee, the investor whose home asset is the diversifier holds it alongside the market fund and buys less of each specialist than other investors do: her own asset already supplies part of what a specialist sells. Her direct holding is a single asset, yet her whole portfolio is diversified. Investors whose home asset does not diversify the market portfolio hold no direct position and buy the specialists; with the targeted pair they hold both pair funds, because either one alone would tilt their portfolio toward a middle asset.

Why does an industry with no skill to sell look so much like the actual one? Two forces do the work. The first is a fixed cost per product. It makes breadth cheap: a product that everyone wants spreads its cost over everyone, so the market fund exists, is the largest fund and charges the lowest fee, and every other product must find a revenue of its own. The second is the covariance of returns. Dispersion in loadings creates a correction that every investor wants, and the revenue it can earn grows with the square of dispersion; whether a product's fixed cost fits under that revenue is the question ϕ answers, and the answer is the number and breadth of the funds. Access frictions then decide who buys the correction and who supplies part of it from her own holdings. And because every asset earns the same expected return, whatever a fund adds to a portfolio's risk mix appears as alpha against the market fund. Alpha in this industry measures a covariance service, and the fee for that service stays below it.

The model is deliberately spare. Equal expected returns isolate covariance as the reason to package one asset rather than another. Funds are equal-weight bundles, standardized rules of the kind the industry sells, and the choice of the rules lies outside the model; every product has the same operating cost, so differences in fees and sizes come from demand alone. Section 6 adds two extensions. The first gives specialists a fixed cost premium over the market fund: the local structure of a market fund and specialists survives, with the fee ratio between the two tiers scaled by the premium. The second lets the broad fund tilt its weights within a mandate: it then delivers part or all of the correction itself, and the split between pooled and targeted specialists rests on standardized weights. The model is therefore a baseline: what covariance, access costs and free entry deliver on their own, before skill, information, trust or beliefs are added.

This paper contributes to the equilibrium theory of delegated asset management. [Berk and Green \(2004\)](#) explain fund flows and the absence of performance persistence with heterogeneous skill, decreasing returns to scale and learning from past returns; [Pástor and Stambaugh \(2012\)](#) explain the size of the active industry with industry-level decreasing returns, which [Pástor et al. \(2015\)](#) document; in [Gennaioli et al. \(2015\)](#), trust in managers lowers the risk investors perceive and supports fees above cost and, with biased beliefs, product proliferation, and [Gennaioli et al. \(2014\)](#) embed this view in a growth model. These papers explain what managers earn, how much capital they attract and why investors delegate. I hold skill, information and trust at zero and ask which standardized portfolios a competitive industry sells when investors differ only in the assets they can hold cheaply.

The closest models are [Gârleanu and Pedersen \(2018\)](#) and [Gârleanu and Pedersen \(2022\)](#). In the first, investors search for informed managers, asset-price efficiency is tied to the efficiency of the market for asset management, and an uninformed manager adds nothing over investing alone. The second adds many securities and costly self-directed investing: a competitive passive manager charges its cost, serves investors whose own cost of building the uninformed portfolio is higher, and tilts away from the securities whose supply is most uncertain, a step toward a theory of optimal indexes. My market fund earns its fee in the same way. In their model every uninformed investor holds the same portfolio; here investors start from different home assets and differ in how much of the correction they need to buy, and competition determines which corrections are sold as separate products.

A second literature asks why composite products exist at all: uninformed traders prefer baskets, where adverse selection costs are lower ([Subrahmanyam, 1991](#)); composites protect uninformed investors from insiders, and when those investors differ in nontradeable endowment risk, clientele size decides between clientele-specific composites and one market composite held with individual securities ([Gorton and Pennacchi, 1993](#)); fund trading styles follow from investors who must delegate trading ([Mamaysky and Spiegel, 2001](#)); and a manager who screens investors with different beliefs offers a menu of bundled funds ([Cvitanić and Hugonnier, 2022](#)). These papers explain why baskets are valuable and how investor heterogeneity shapes them. I ask which standardized baskets survive competition when access is costly and each fund pays a fixed operating cost.

Other work traces product lines to fund families and search: families exploit investor heterogeneity, and differentiation in non-performance characteristics raises the number of

funds (Massa, 2003); non-portfolio differentiation and search frictions explain the proliferation and fee dispersion of S&P 500 index funds (Hortaçsu and Syverson, 2004); specialized exchange-traded funds track niche portfolios, charge high fees and underperform after launch, consistent with catering to extrapolative beliefs (Ben-David et al., 2023). Here funds differ in the portfolios they hold and investors in the assets they can hold cheaply, so fee differences reflect the demand that covariance creates, with contestability disciplining prices (Baumol et al., 1982).

The access friction has a long history in asset pricing. In Merton (1987), investors hold only the securities they know about, and the shadow cost of incomplete information raises expected returns on less widely known securities; Gârleanu et al. (2020) show that such frictions are equivalent to investor- and asset-specific investment taxes, which generate underdiversification and justify fund-specific benchmarks; investors hold mostly domestic assets (French and Poterba, 1991; Coeurdacier and Rey, 2013), and learning can amplify a small home information advantage (Van Nieuwerburgh and Veldkamp, 2009). I take investor-specific access costs as given, hold returns fixed, and study the fund industry that forms around local direct portfolios. Investors do pay intermediaries for exposures they find costly to build: mutual fund fees rise with market beta above one, and more steeply when leverage constraints tighten (Hitzemann et al., 2022); and when a robo-advisor lowered its account minimum, middle-class investors entered, gaining, in the authors’ calibrated model, diversification they could not build themselves (Reher and Sokolinski, 2024).

The paper proceeds as follows. Section 2 sets out the assets, access costs, fund products and equilibrium concept. Section 3 derives the equilibrium industry: the market fund first, then specialist demand and fund variety. Section 4 develops the implications for fees, performance and investor portfolios. Section 5 reports the numerical evidence at finite dispersion, Section 6 relaxes the common operating cost and the fixed fund weights, and Section 7 concludes. The Online Appendix contains the proofs, exact certificates and supporting calculations.

2 The economy and equilibrium

This paper models the asset-management industry as a supplier of packaged diversification. Investors in the model agree about expected returns and risk. None of them can pick stocks, and neither can any fund manager. What differs across investors is only the asset each

of them can hold for free. Everything else has to be bought: an investor can pay a fixed cost to trade another asset directly, or she can buy a fund that bundles several assets at a proportional fee. Funds in this world sell convenience, not skill. They package exposures that investors could in principle assemble themselves, and they survive only if packaging is cheaper than self-assembly.

The industry I study is therefore deliberately stripped down. Its moving parts are covariance, access, and product cost. Covariance determines which corrections are worth packaging; access determines what investors produce themselves; and operating cost determines how finely a contestable industry can divide the remaining demand. The next three subsections define these primitives; Section 2.4 derives the fund demand they generate; Section 2.5 states the survival tests that free entry imposes and defines equilibrium.

2.1 Covariance and the residual correction

Four risky assets sit at equally spaced points on a circle, indexed $j \in \{0, 1, 2, 3\}$. Their excess returns follow a one-factor structure,

$$R_j = \mu + \sigma \left(a_j f + \sqrt{1 - a_j^2} \epsilon_j \right), \quad f, \epsilon_0, \dots, \epsilon_3 \stackrel{\text{iid}}{\sim} N(0, 1), \quad (2)$$

with $\mu > 0$, $\sigma > 0$, and $0 \leq a_j < 1$. Every asset has the same expected return and the same variance. I make this choice deliberately: with the return cross-section shut down, no asset is attractive because it pays more, and among payoff characteristics the only reason to hold or package any particular asset is how it covaries with the others. Covariance is the sole source of payoff differentiation; free home access separately differentiates investors.

The loadings are

$$a = (m + c, m, m - c, m), \quad 0 < c \leq m, \quad m + c < 1, \quad (3)$$

so the correlation matrix is

$$\Sigma = aa' + \text{diag}(1 - a_0^2, \dots, 1 - a_3^2), \quad (4)$$

with $\Sigma_{jk} = a_j a_k$ for $j \neq k$ and unit diagonal. Asset 0 is the *high-loading asset*: it carries and therefore repeats the most common risk in a portfolio. Asset 2 is the *low-loading diversifier*; assets 1 and 3 are the middle assets.

One property of this covariance structure matters for how the whole paper should be read. Because every asset keeps strictly positive idiosyncratic variance, Σ has full rank: no asset's payoff can be manufactured from the other three, and every asset adds an independent payoff direction. This has a sharp implication for what funds can and cannot do in this economy. A fund is a portfolio of the same four assets, so it cannot offer investors any payoff they could not, in principle, construct on their own — funds expand nobody's investment opportunity set. In the language of security design, there is no spanning value here to be created ([Allen and Gale, 1994](#)). Whatever value funds deliver must therefore come from a different source: they save investors the cost of assembling exposures, not the impossibility of it. This is a deliberate modeling discipline. If funds could span otherwise unavailable payoffs, they would trivially be valuable, and the theory would have assumed its conclusion. Here every dollar of fund value has to be earned by economizing on access costs, and the industry structure that emerges is the market's answer to a pure packaging problem.

The two covariance parameters play distinct roles, and much of the paper is about the second one. The average loading m sets how much common risk the market carries. The dispersion c makes some assets better diversifiers than others. At $c = 0$ the four assets are exchangeable, and an equal-weight bundle of all four has the mean–variance-efficient risky composition for everyone. At $c > 0$ that same bundle no longer matches the full-access mean–variance-efficient composition: it puts too much weight on the high-loading asset and too little on the low-loading diversifier, whose whole appeal is that it hedges the rest. This gap between the standardized market portfolio and the efficient target is what creates room for more than one fund.

Equal expected returns isolate this covariance channel. Under common frictionless access and unrestricted tangency weights, classical two-fund separation would make one risky fund sufficient. That comparison does not extend mechanically to the access economy: location-specific free assets make residual demands investor-specific. Access heterogeneity alone is still not enough here. In the low-cost region studied below, specialists need loading dispersion to survive: at $c = 0$, the market fund and each investor's home asset span the needed correction,

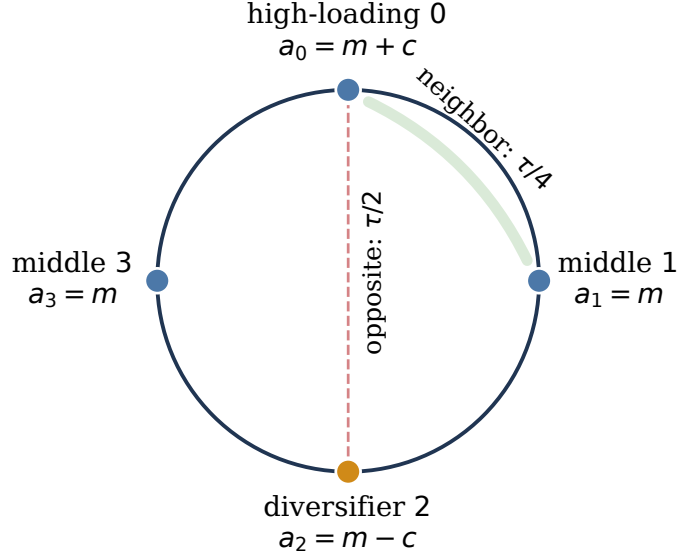


Figure 2: **Assets, investors, and direct access.** The four assets and investor types share one circle. An investor holds her home asset for free, pays $\tau/4$ to access a neighboring asset, and $\tau/2$ to access the opposite asset.

leaving no residual diversification service that could finance another product. Product variety requires both different starting points and loading dispersion.

2.2 Investors and access costs

There are four investor types, one located at each asset, each with mass $1/4$. Location is an investment technology, not a taste: type i can hold her *home asset* i directly at no cost, and she pays a fixed setup cost

$$\kappa(i, j) = \tau d(i, j) \quad (5)$$

to be able to trade any other asset j , where $d(i, j)$ is the distance around the circle. Reaching a neighboring asset costs $\tau/4$; reaching the opposite asset costs $\tau/2$. Once access is paid for, trading the asset is free at any scale. Figure 2 shows the setup: investors and assets share one circle, so every investor faces the same schedule of access costs but starts from a different free asset.

I think of τ as the cost of doing one's own investing in an unfamiliar market: opening the account, learning the instruments, monitoring the position. The home-bias literature

documents both the friction and its portfolio consequences (French and Poterba, 1991; Van Nieuwerburgh and Veldkamp, 2009; Coeurdacier and Rey, 2013); here I take the friction as primitive and derive the industry that forms around it. Two features of the specification matter. First, access is a fixed cost, so the decision to reach an asset is discrete, while the amount held afterward is continuous. Second, the home asset is always available even when the investor chooses not to hold it, so a home position that starts at zero and later turns positive is an ordinary portfolio adjustment, not a new access decision.

The graded form itself is a benchmark choice rather than a load-bearing assumption. More generally, an investor could face any pattern of asset-specific setup costs, with her home asset free and every non-home cost bounded away from zero. For the local equilibrium results below, those two properties — the zero at home and the positive floor elsewhere — are the essential ones.

Investors have CARA utility over terminal wealth with common risk aversion γ , and a risk-free asset is in perfectly elastic supply. Writing $\lambda = \gamma\sigma^2$, an investor with total risky exposure vector e enjoys the certainty equivalent $\mu\mathbf{1}'e - \frac{\lambda}{2}e'\Sigma e$, minus whatever fees and access costs she pays. Two remarks on constraints. Borrowing and lending at the risk-free rate are unrestricted, so leverage is allowed and wealth never constrains risky positions; “assets under management” below are therefore units of risky exposure, not shares of anyone’s wealth. Risky positions themselves are long-only. The long-only constraint is the natural one for a model of packaged diversification — retail investors do not short — and the next assumption ensures that it does not bind for an investor who can trade every asset for free.

I call the assets an investor can trade herself — her home asset plus any assets she has paid to reach — her *direct portfolio*.

Assumption 1 (Every asset is worth holding). *The loadings are such that an investor who could trade all four assets for free would hold all four: $a_0Q < 1 + R$, where $Q = \sum_j a_j/(1 - a_j^2)$ and $R = \sum_j a_j^2/(1 - a_j^2)$.*

The assumption selects the economies this paper is about. Its content is that diversification is genuinely four-dimensional: each of the four assets contributes enough independent diversification value that a portfolio is better with it than without it, even for the asset most exposed to common risk. Because asset 0 carries the most common risk, it is the asset whose place in a portfolio is most fragile—if any asset were dropped from a free portfolio,

it would be the high-loading asset—so one inequality on that asset is all that needs to be imposed. Without the condition, the frictionless portfolio could omit the high-loading asset, and product design would become a problem with effectively fewer assets. Within the economies the assumption selects, an investor who uses only direct assets holds every asset she can trade, so the value of each direct-access option takes a single formula (Appendix A verifies this for every asset subset). With funds available, her direct holdings can still be zero. The benchmark parameters used in all figures satisfy the condition comfortably.

Location is the only source of heterogeneity in the model. Every investor wants the same diversified exposure; they differ in which piece of it they start with. Type 2, the *diversifier investor*, is born with the best diversifier and has the least to gain from anyone’s help. Type 0, the *high-loading investor*, is born with the asset that repeats the most common risk and is farthest from the diversifier, so she depends most on packaged diversification. The middle investors lie between them.

2.3 Funds

A *fund* is a standardized product: an equal-weight portfolio of two or more adjacent assets on the circle, sold at a proportional fee. The product universe is the nine contiguous arcs,

$$\mathcal{M} = \{01, 12, 23, 30, 012, 123, 230, 301, 0123\}, \quad (6)$$

where the label lists the assets in the bundle; b_T denotes product T ’s unit-exposure portfolio vector. Any active fund T charges a proportional fee δ_T per unit of exposure and pays a fixed operating cost $F > 0$ per period, the same for every product. The common fixed cost is a deliberate neutrality choice: no product is cheaper to run by assumption, so if narrow and broad funds end up with very different fees and sizes, the difference is earned in equilibrium rather than built into the technology. Put differently, the supply side of this industry is deliberately featureless — identical costs, no skill, free entry — so that every result about which products exist, how large they are, and what they charge is driven by the demand side: by how the covariance structure and the access technology shape what investors want to buy.

I give four of the nine products names, for repeated reference; entry remains open to all nine. The *market fund* $B = 0123$ holds every asset. Under the model’s equal-supply normalization, its equal weights coincide with market weights. The market fund supplies the

equal weights over the assets in each arc					
$B = 0123$	0	1	2	3	the market fund
$C = 123$	0	1	2	3	the three-asset specialist
$P_L = 12$	0	1	2	3	two-asset specialist
$P_R = 23$	0	1	2	3	two-asset specialist
<i>the other five feasible arcs</i>					
012	0	1	2	3	
230	0	1	2	3	
301	0	1	2	3	
01	0	1	2	3	
30	0	1	2	3	

Figure 3: **The nine feasible products.** Each row is one contiguous arc, equal-weighted over its shaded assets. The four named products are highlighted: the market fund B , the three-asset specialist C , and the two-asset specialists P_L and P_R .

diversified base. The *three-asset specialist* $C = 123$ omits the high-loading asset and keeps everything else; it is the least concentrated way to correct the market portfolio's tilt. The two *two-asset specialists* $P_L = 12$ and $P_R = 23$ flank the diversifier and deliver the correction in concentrated form. Figure 3 displays all nine feasible products with these four highlighted. (The Online Appendix, the paper's formal layer, refers to B as the *broad fund*; both names denote the same product.)

Equal-weight contiguous arcs are a substantive standardization choice. Funds here are rules, not solutions: each mandate names its assets — the market, the market excluding its high-loading asset, a targeted diversifier pair — and weights them by a fixed convention rather than solving the covariance matrix for optimal weights. This mirrors the technology of the actual industry, whose products are publicly specified portfolio rules — capitalization weights, equal weights, sector sleeves — rather than solutions to each investor's portfolio problem. Investors choose how much of each available product to hold; intermediaries choose whether to offer a standardized product and at what fee. The paper therefore solves for

the active menu within a fixed product vocabulary. It does not solve unrestricted product design: because investors share beliefs, a single product that solved the covariance matrix outright would serve every investor at once, and the question of which imperfect standardized products coexist would not arise. Why intermediaries sell rules rather than solutions — estimation risk, agency, the scale economies of simple mandates — lies outside the model. Equal-weight diversification is also a recognizable benchmark in practice (DeMiguel et al., 2009). Section 6.2 relaxes the fixed weights as a robustness check: it lets a broad fund tilt its weights within a mandate and examines two specialist redesigns.

The timing is that of a contestable industry. Any product can enter or stay out; active funds post fees; investors then choose access and portfolios, taking the menu and fees as given. Section 2.5 turns free entry into two precise protection conditions — against cheaper clones of an active fund, and against entry by any omitted product — and defines equilibrium.

Units are normalized from here on. The equilibrium is homogeneous in μ and λ , so I set $\mu = \lambda = 1$: positions are measured in units of μ/λ , fees in units of μ , and the two fixed costs τ and F in units of μ^2/λ (Appendix A records the exact map). To restore economic units, multiply positions by μ/λ , fees by μ , and fixed costs by μ^2/λ .

2.4 The investor's problem

This subsection states the investor's complete problem — the joint choice of paid access, direct positions, and fund positions — and records the two facts about demand that the supply side uses: demand for a product falls in its own fee, and it is piecewise affine, with pieces given by portfolio regimes. Derivations and the formal demand lemma are in Appendix A.

Type i chooses a set D of assets to pay access for, nonnegative direct positions x_i supported on $D \cup \{i\}$, and nonnegative fund positions q_{iT} for the funds on offer. Her total risky exposure is $e_i = x_i + \sum_T q_{iT} b_T$, and she solves

$$\max_{\substack{D \subseteq \{0,1,2,3\} \setminus \{i\} \\ x_{ij} \geq 0, j \in D \cup \{i\} \\ q_{iT} \geq 0}} \left\{ \mathbf{1}' e_i - \sum_T \delta_T q_{iT} - \frac{1}{2} e_i' \Sigma e_i - \tau \sum_{j \in D} d(i, j) \right\}. \quad (7)$$

Holding other fund quotes fixed, revealed preference makes demand for a focal fund weakly decreasing in its own fee. Conditional on the access set and on the instruments used, the

investor faces a standard mean–variance problem. Its first-order conditions are linear, so positions are affine in fees within a portfolio regime. An instrument entering or leaving at zero creates a continuous kink. A change in the preferred paid-access set is discrete and can instead create a downward jump.

Two of her options require no industry at all: she can hold no risky exposure, or hold her free home asset alone, which with $\Sigma_{ii} = 1$ delivers certainty equivalent $1/2$. These outside options are available at every menu and every fee vector, and they anchor what any product must beat. What an investor can reach entirely on her own, with and without paid access, is characterized exactly in Appendix A.

A fund prices against the population average of these positions. For any fund T , assets under management are

$$\text{AUM}_T(\delta) = \frac{1}{4} \sum_{i=0}^3 q_{iT}(\delta), \quad (8)$$

measured in risky-exposure units. Averaging preserves weak monotonicity but does not smooth away discrete access choices. Aggregate demand for any product is therefore a chain of nonincreasing affine pieces joined by continuous kinks and possibly separated by downward jumps; the number, locations, sizes, and investor composition of the pieces are endogenous.

Several funds add one further complication: implementation ties. Conditional on a paid-access choice, strict concavity fixes final risky exposure and value, but redundant products can implement that exposure differently, leaving product-level AUM ambiguous. At a paid-access tie, even the exposure can differ. A selection convention, stated with the equilibrium definition in Section 2.5 and formally in Appendix A, resolves product-level demand at ties. Appendix A establishes this structure for any finite menu: away from regime boundaries, holdings are affine in fees; at an ordinary support boundary, demand is continuous though its slope may change; a change in paid access or in the selected implementation can create a discrete change in product-level demand; and, under the selection convention, selected aggregate demand for a product consists of finitely many affine pieces, so its revenue is quadratic in its own fee on each piece. The pricing consequence is immediate: a comparison between fees cannot be made locally. The same-product undercut test must check every lower quote, and Section 2.5 formalizes this global comparison.

2.5 Contestable product pricing

Funds face fixed costs and free entry, so pricing is a question of survival: which fees can persist when anyone can clone an active fund or launch an omitted one? This subsection defines the two survival tests and the resulting equilibrium notion. The first test protects an active product against a cheaper copy of itself. The second protects the menu against a different product that is currently absent.

The timing of Section 2.3 matters here: a deviating product posts its quote first, and investors choose access and portfolios after observing it, so a deviator competes against everything investors can rearrange in response — their access decisions included.

Both tests use demand from the complete investor problem in (7), including endogenous paid non-home access. Demand may be smooth, kink at a support change, or jump at an access switch. Figure 4 isolates the first test on a smooth one-regime schedule.

A fund's revenue is its fee times its assets under management $\text{AUM}_T(\delta)$, the aggregate demand defined in (8). A fund breaks even when fee revenue covers the operating cost,

$$\delta_T \text{AUM}_T(\delta) = F. \quad (9)$$

Break-even alone does not pin down a competitive fee. Revenue need not be monotone: the same fixed cost can be recovered at a low fee on a large customer base or at a high fee on a smaller base. Free entry distinguishes the two roots. A clone offers exactly the incumbent's portfolio and changes only its own fee. If some lower quote would give that clone revenue above F , it can enter, charge less, and displace the incumbent. An active fund is therefore protected only when

$$\eta \text{AUM}_T(\eta, \delta_{-T}) \leq F \quad \text{for every } 0 \leq \eta < \delta_T. \quad (10)$$

Here δ_{-T} holds every other product's quote fixed. The inequality compares the incumbent with every lower own quote, including quotes that move investors to another portfolio regime. A local first-order condition cannot make this comparison: a more distant quote may activate a home asset, change a replication selection, or cross an access boundary. The entire piecewise revenue curve matters: the definition is the global inequality in (10), not a first-crossing rule. Figure 4 shows the test on one smooth portfolio regime, where affine demand makes revenue

a single quadratic arc. The first break-even quote δ^- is protected: every cheaper quote earns less than F . The later crossing δ^+ is not: the intermediate lower quote η earns more than F , so a clone can undercut. The smallest protected break-even fee is the *lower root*.

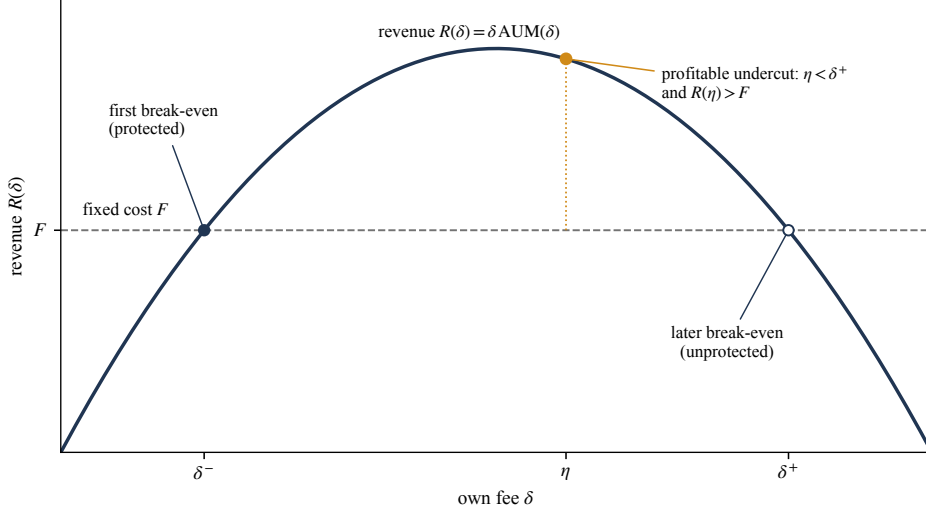


Figure 4: **Lower-root selection on a smooth revenue curve (schematic)**. The figure isolates one portfolio regime. At the first break-even quote δ^- , every cheaper quote earns less than the operating cost F . At the later break-even quote δ^+ , the strictly lower quote η earns more than F , so a clone can undercut. All horizontal locations are schematic.

Actual revenue can bend at a support change or jump at an access switch and skip the cost line. The global test in (10) still compares every lower quote.

The second survival test concerns products not yet on the shelf. An omitted product U can enter at any fee η , taking incumbent fees as given; its best attainable revenue is

$$\mathcal{R}_U = \sup_{\eta \geq 0} \eta \text{ AUM}_U(\eta, \delta), \quad (11)$$

and the menu is entry-proof when $\mathcal{R}_U \leq F$ for every omitted product. The supremum makes the test global in the entrant's fee. Entry is unilateral and product by product; coordinated launches of several products at once are outside the model.

Definition 1 (Free-entry product equilibrium). *An equilibrium is a menu $\mathcal{E} \subseteq \mathcal{M}$, fees $\delta^\mathcal{E}$, and investor portfolios such that:*

- (i) every investor solves the specified portfolio problem;
- (ii) every active fund has positive AUM, breaks even, and is protected against every lower quote (10); and
- (iii) no omitted product can earn revenue strictly greater than F at any fee.

At an implementation tie, incumbents use one common product-level allocation of the tied total exposure. A unilateral deviator changes only its own quote and is assigned the largest demand for that focal product among tied optimal implementations. Appendix A states the convention formally. Giving the deviator the largest tied demand makes incumbent protection harder, so the convention is conservative for the equilibrium menus.

This is contestable product competition (Baumol et al., 1982). Unlike models in which flows reward skill or search frictions split match surplus (Berk and Green, 2004; Gârleanu and Pedersen, 2018), any standardized bundle can be launched by anyone. Break-even and entry protection discipline prices, while asset returns remain exogenous.

Break-even ties each active product's fee to its size, and the link is exact at every c .

Lemma 1 (Equal revenue and inverse size–fee ratios). *In any equilibrium, every active product earns the same revenue: $\delta_T \text{AUM}_T = F$ for every active T . Hence, for any two active products T and U ,*

$$\frac{\delta_T}{\delta_U} = \frac{\text{AUM}_U}{\text{AUM}_T}.$$

The lemma restates Definition 1(ii) as accounting and needs no asymptotics: a common operating cost recovered from customer bases of different sizes must show up as inversely proportional fees. Relative fees measure relative sizes, one for one, in every equilibrium of the economy.

3 Equilibrium industry structure

The equilibrium argument has two steps. Competition first produces a market fund that serves the diversification demand shared by investors. The remaining demand then determines which specialists can operate alongside it.

3.1 Why competition produces a market fund

Investors share a demand for broad diversification. The question is how a contestable industry supplies it when each investor starts with a different free asset and any of nine standardized funds can enter. A market fund sells this diversification in one product. It spreads its operating cost over the entire common demand and can therefore charge a very low fee. Other funds can supply the same exposure to every investor only in combination, and each fund in the combination must cover its own operating cost. When c and operating cost are low enough, competition therefore produces a market fund that every investor holds.

3.1.1 A common demand for diversification

Investors want diversification for the same reason: combining assets reduces idiosyncratic risk without sacrificing expected return. Each investor can hold her home asset for free, but holding that asset alone leaves its idiosyncratic risk undiversified. Because investors have the same preferences and beliefs, they agree on the portfolio they would choose if all assets were freely accessible. Their home assets affect the cost of obtaining that portfolio. This shared demand gives a fund the opportunity to serve all investor types.

To identify this common demand, first set $c = 0$ and leave fees and access costs aside. All four assets then have the same factor loading, and every pair has correlation m^2 . Equal weights spread their independent risks evenly. Write $b_B = \frac{1}{4}\mathbf{1}$ for these weights, where $\mathbf{1}$ is the vector of four ones. The market portfolio's variance is then $m^2 + (1 - m^2)/4 = (1 + 3m^2)/4$. The common-risk component m^2 remains, while averaging reduces the idiosyncratic component to one quarter of its single-asset level. A unit of diversified exposure therefore offers the same expected return as the home asset with less risk.

The investor also chooses how much risky exposure to hold. Each additional unit brings the same expected return but adds progressively more risk. Equating marginal return and marginal risk cost gives total demand $Q_0(m) = 4/(1 + 3m^2)$ and the common frictionless portfolio

$$e_0^*(m) = \frac{1}{1 + 3m^2}\mathbf{1} = Q_0(m)b_B. \quad (12)$$

Equal weights describe the portfolio's composition; Q_0 describes its size. These are risky-exposure units, with the risk-free asset and borrowing providing financing as needed. Diversification lowers risk per unit invested, so the investor chooses a larger position: Q_0 units

of the market portfolio instead of one unit of her home asset. Even when factor loadings coincide, spreading idiosyncratic risk creates a service that every investor values.

At positive c , covariance-efficient weights shift toward the low-loading asset, which adds less common risk. Investors still share a preferred portfolio, but equal weights no longer deliver its exact composition. For sufficiently low c , they remain close to it, so the market fund sells nearly the portfolio that every investor wants.

3.1.2 Demand for the market fund and its fee

An investor can obtain the common portfolio in three ways: from a market fund, by paying for direct access to other assets, or by combining other funds, possibly with her free home asset. A combination works only if it reproduces equal weights. The complementary pair funds, 01 and 23 or 12 and 30, do so for every investor. A three-asset fund does so only for the investor whose home asset it omits. For example,

$$b_{01} + b_{23} = 2b_B, \quad 3b_{123} + \iota_0 = 4b_B, \quad (13)$$

where ι_0 holds asset 0 alone. I call such a combination a *replication* of the market portfolio. Its fee per unit of market exposure is $(\delta_{01} + \delta_{23})/2$ in the first example and $\frac{3}{4}\delta_{123}$ in the second, whose home-asset share is free. A home asset alone cannot replicate the market portfolio.

Because a replication delivers exactly the market fund's exposure, an investor compares the two by their fees.

Lemma 2 (Demand for the market fund). *Take all fees as given.*

- (a) *An investor holds none of the market fund if some replication available to her, built from offered funds and her own home asset, has a lower fee per unit than the market fund.*
- (b) *For sufficiently low c and a sufficiently low market-fund fee, whatever other funds are offered, no investor pays for direct access, and each investor holds a risky position close to the common portfolio $Q_0 b_B$. If the market fund's fee is below the fee per unit of every replication available to her, she holds about Q_0 units of the market fund, and her other holdings are small.*

Appendix B.2 proves the lemma and states its bounds.

Part (a) holds for every admissible c : an investor never pays more for an exposure she can buy more cheaply. Part (b) adds that her whole common position is at stake. Within this low-fee range, her demand for the market fund falls sharply, from about Q_0 to zero, when its fee rises past her cheapest replication cost. Below that cost her demand still responds to the fee, but only slightly.

The free home asset has two uses: saving fees and changing the risk mix. When the market fund is the only fund and $c = 0$, only fee saving matters: a low positive fee leads every investor to move part of her exposure into her home asset and to hold less in total. With the market fund still the only fund, a positive c makes the mix depend on the home asset. The low-loading home asset shares less risk with the market fund and can help even when the fund charges no fee. The high-loading home asset has a covariance disadvantage; fee savings must compensate for that disadvantage before the investor uses it.

A market fund can serve the common demand only if investors are willing to pay enough to cover its operating cost. I therefore restrict the cost of running a product.

Assumption 2 (Operating-cost bound). *Operating cost satisfies*

$$0 < F < \bar{F}(m, \tau),$$

where the positive ceiling $\bar{F}(m, \tau)$ is sufficiently low.

The role of this bound is to make cost recovery compatible with a fee investors are willing to pay. Their willingness to pay depends on two things: the benefit of diversification, governed by m , and the cost of obtaining it through direct access, governed by τ . If operating cost is too high, a fee that would cover it can drive investors toward their home asset or direct assembly, leaving too little demand to finance the fund. Direct access to even one other asset costs at least $\tau/4$. Part (b) of Lemma 2 uses this comparison: when c and the market fund's fee are low enough, even perfect self-assembly would improve on holding only the market fund by less than that charge. This is why τ enters the ceiling.

With the market fund as the only fund, no replication exists, so Lemma 2(b) applies to every investor: at a low fee, each holds about Q_0 units. At this offered menu, a break-even fee satisfies

$$\delta_B \text{ AUM}_B(c, \delta_B) = F.$$

Because every investor buys nearly the common quantity, this fee is about F/Q_0 : the cost spread over the common quantity. Revenue increases throughout the interval from zero to this break-even quote, so every lower quote earns less than F . This protects the quote against a cheaper identical fund. Appendix A.9 establishes these properties of the low break-even quote. Each investor then pays the market fund about F , and exactly F on average, because total investor mass is one.

A market fund that stands alone can therefore finance itself at a very low fee. The remaining question is whether other funds can keep it out.

3.1.3 Why one fund beats several

An omitted market fund can enter at one posted fee; an active market fund can lower its fee. A candidate equilibrium fails the corresponding test if the resulting revenue exceeds the fund's operating cost.

Without a market fund, an investor who holds nearly the common portfolio and pays for no direct access must hold it mostly through a replication. An investor who is far from that portfolio, or who pays for direct access, would gain from a low-fee market fund, which brings her close to the portfolio without paid access (Lemma 2(b)). A single narrower fund replicates the market portfolio for at most one investor, together with the home asset it omits. Supplying the common exposure to every investor therefore takes at least two funds, and a complementary pair is the only way to do it with two.

Suppose the complementary pair funds 01 and 23 supply the common exposure. Holding about $Q_0/2$ units of each delivers about Q_0 units of market exposure. With total investor mass one, each fund then has AUM of about $Q_0/2$ and needs a fee of about $2F/Q_0$ to break even. Investors therefore pay $2F$ in total for the common exposure: two products finance two operating costs. A market fund can instead quote $3F/(2Q_0)$, below the pair's fee per unit of market exposure. By Lemma 2(b), every investor then moves her common position to the market fund, which earns about $\frac{3}{2}F > F$. The offer leaves investors part of the cost saving and earns more than it costs to operate.

Appendix B carries out the entry and lower-quote tests across the full catalogue and establishes the following result.

Proposition 1 (Endogenous market fund). *Under Assumption 2, there is a positive bound $\bar{c}(m, \tau)$ such that, for every $0 < c < \bar{c}(m, \tau)$, every equilibrium contains B , held by every investor type. Its assets under management have a positive lower bound that does not depend on c or F . No investor pays for non-home access.*

The pair example is the simplest case. With more funds, investors bear more operating costs, but different investors can face different replication costs: an investor whose home asset completes a three-asset fund needs only three quarters of a unit of that fund per unit of market exposure. The competitive argument must therefore establish profitable entry at a single market-fund quote; adding up potential cost savings is not enough.

An active market fund faces the same discipline through its own fee reductions. If some investors obtained their market exposure elsewhere, some lower market-fund fee would attract them and earn more than F . In every equilibrium, therefore, every investor holds the market fund.

Competition selects the market fund because it economizes on operating costs: it supplies the demand that every investor shares with a single product. The proposition allows any fixed operating cost below its ceiling; F need not fall as c falls.¹

The market fund's equal weights need not deliver the covariance-efficient portfolio, so additional funds may help investors reallocate exposure. Their value depends on what investors can already obtain through the market fund and their home assets.

3.2 Specialist demand and fund variety

Section 3.1 established the market fund. Every investor uses it, and no investor pays for direct access to a non-home asset in the region studied there. The remaining question is what other funds can sell when this shared source of diversification is already available. Specialists sell an improvement to the market fund's risk mix. Every investor values it, but investors differ in how easily they can make it themselves. Loading dispersion raises both the amount of this improvement investors buy and what they are willing to pay for each unit. Specialist revenue therefore grows with the square of dispersion. Operating cost relative to that revenue determines *fund variety*: the number and composition of the funds offered alongside the

¹Proposition 1 characterizes every equilibrium in the stated region. Appendix B.8 separately establishes a market-fund-only equilibrium for each sufficiently low fixed positive operating cost and sufficiently low c .

market fund. Lower operating cost or greater dispersion makes room for more specialists, with narrower mandates.

3.2.1 What specialists add

The market fund puts one quarter of each dollar in every asset. Yet the assets contribute differently to common risk: asset 0 has the highest factor loading, asset 2 the lowest, and assets 1 and 3 lie between them. Starting from equal weights, investors can improve the risk mix by moving exposure from asset 0 toward asset 2. Expected returns are the same, so this change obtains a covariance benefit without giving up expected return. Investors agree on the desired change, just as they agree on the value of diversification in Section 3.1.1. I call this change the *correction*.

A specialist sells a way to make that change. Consider the three-asset fund $C = 123$, which excludes asset 0 and divides its holdings equally among the other three assets. An investor who replaces part of her market-fund holding with C reduces exposure to the high-loading asset. The replacement increases the weights of both middle assets as well as the diversifier. I call this the *pooled* specialist: one mandate can serve investors with different home assets.

The pair funds $P_L = 12$ and $P_R = 23$ offer a different implementation. Each puts half its investment in the diversifier and half in one middle asset. Equal holdings of the two funds, replacing market-fund investment, shift exposure from asset 0 to asset 2 while leaving the middle-asset weights unchanged. They move exposure in exactly the direction of the correction, and I call the two funds the *targeted* pair. Unequal holdings allow the investor to adjust those middle weights as well. The pair therefore provides more ways to tailor the portfolio, but it must finance the operation of two products. The pooled fund finances only one.

Measured against the market fund, each of these products earns positive alpha. A product's alpha is its expected excess return before fees, less its beta on the market fund times the market fund's expected excess return. Every asset has the same expected return, so alpha comes from beta alone. A product that holds more of the diversifier than of the high-loading asset shares less of the market fund's risk and has positive alpha. This alpha requires no private information and no stock-selection skill. It is a property of fixed portfolio weights and the commonly known covariance of the assets.

3.2.2 Demand for specialists

For tractability, I characterize demand and the organization of the industry for small loading dispersion c , holding m fixed. Fees, holdings and revenues are described by their leading terms in c ; I call this the *small-dispersion approximation*. Section 5 examines larger dispersion numerically.

A product's alpha is then determined by its weights on the diversifier and on the high-loading asset:

$$\alpha_T = 2(b_{T,2} - b_{T,0})\alpha, \quad \alpha = \frac{mc}{1 + 3m^2}, \quad (14)$$

where $b_{T,j}$ is the weight of asset j in product T 's portfolio b_T . Each pair fund has alpha α , and the pooled fund $2\alpha/3$. The formula applies to any portfolio: a single asset has alpha 2α if it is the diversifier, -2α if it is the high-loading asset, and zero otherwise.

The market fund's fee is too small to affect demand for specialists. The market fund covers its operating cost at a fee of about F/Q_0 , the cost spread over the common quantity (Section 3.1.2). A specialist can cover the same cost only if F is at most a constant times c^2 (Proposition B.9). Whenever a specialist operates, the market fee is therefore of order c^2 , whereas alphas are proportional to c . The home asset then no longer serves to save the market fund's fee; an investor holds it only to change her risk mix.

Specialists and home assets change the composition of an investor's portfolio. Write $d_T = b_T - b_B$ for product T 's active weights and $H_i = \iota_i - b_B$ for those of type i 's home asset, where ι_i has a one in position i and zeros elsewhere. If type i holds q_{iT} units of each specialist and x_{ii} units of her home asset, her *exposure adjustment* is $r_i = \sum_T q_{iT}d_T + x_{ii}H_i$. It is the difference between her risky exposure and the same total position held entirely in the market fund, so its entries sum to zero. At equal loadings such an adjustment carries no common risk. Its variance is $(1 - m^2)\|r_i\|^2$, the idiosyncratic variance of each asset times the sum of squared entries.

I call each specialist and the home asset a *position*. Position k has active weights d_k , alpha α_k and fee δ_k ; for the home asset, $d_k = H_i$ and $\delta_k = 0$. Positions k and l have covariance $(1 - m^2)d'_k d_l$. The investor collects the alpha of each holding, net of its fee, and bears a cost equal to one half the variance of her adjustment.

Lemma 3 (Demand for specialists). *In the small-dispersion approximation, fix the offered specialists and their positive fees. Type i 's optimal adjustment r_i is unique, and she can*

implement it with a set J of held positions whose active weights are linearly independent. Her holdings of these positions are

$$q_J = \Omega_J^{-1} (\alpha_J - \delta_J), \quad (15)$$

where q_J , α_J and δ_J list the holdings, alphas and fees of the positions in J , and Ω_J is the covariance matrix of their active weights. For every position k outside J , the value of an additional unit, $\alpha_k - (1 - m^2) d'_k r_i$, is at most its fee δ_k .

Appendix C.1 states the approximation precisely, proves the lemma, and explains how product demand is assigned when several sets of positions deliver the same adjustment.

The lemma gives each investor's holdings as a function of the fees. When she holds a single position, (15) divides its alpha net of fee by the variance of its active weights. This is the familiar mean–variance demand for an active position, with alpha in the role of expected excess return. The investor stops buying when the fee reaches the alpha. With several positions, an additional unit of a product is worth its alpha less the covariance of its active weights with her adjustment. A product is worth less to an investor whose adjustment already moves with it and more to one whose adjustment it offsets. Investors choose their portfolios as a whole, so demand for one product depends on the rest of the menu and on every fee.

The same covariances govern how demand responds to a fee. Within a given set of positions, holdings are linear in fees. When her other positions can replicate part of a product's active weights, her holdings of the product fall faster with its fee than if she held it alone. As the fee rises, she rearranges those positions to replace part of the product, instead of only cutting back. A free home asset can play this role. It can make an investor's demand for a specialist both smaller and more sensitive to the fee.

3.2.3 Revenue and operating cost

A specialist's fee revenue is its fee times its assets under management, $R_T = \delta_T \text{AUM}_T$, where AUM_T averages the four types' holdings of T . To operate, the specialist must cover its operating cost F .

Loading dispersion raises this revenue through two channels. The first is the price of the service. Alphas are proportional to α , and hence to c , and so is what investors will pay for a unit. The second channel is quantity. With free access to all assets, each investor would move $2\alpha/(1 - m^2)$ units of exposure from asset 0 to asset 2, relative to the same total position in

the market fund. Greater dispersion enlarges this correction, and at fees proportional to α the holdings in (15) are proportional to $\alpha/(1 - m^2)$. Revenue is therefore proportional to $\alpha^2/(1 - m^2)$, which I call the *revenue scale*. It grows with the square of loading dispersion.

I measure operating cost against this scale, as *normalized operating cost*:

$$\phi = \frac{(1 - m^2)F}{\alpha^2} = \frac{F}{\Theta(m)c^2}, \quad \Theta(m) = \frac{m^2}{(1 - m^2)(1 + 3m^2)^2}. \quad (16)$$

A larger ϕ means that operating a fund is more costly relative to the revenue that dispersion creates. For any positive c , this is a change of units. Operating cost remains fixed with respect to the fund's AUM, and the normalization does not assume that the product can cover it. Across economies with the same m , a higher operating cost F raises ϕ , while greater loading dispersion c lowers it.

3.2.4 The pooled fund: a worked example

Consider the pooled specialist $C = 123$ as the only fund offered alongside B . Take its fee as given, as in the market-fund demand calculation in Section 3.1.2. Investors choose how much of C to buy and how much of the desired exposure to obtain through their home assets.

Two customer groups. The fund has alpha $2\alpha/3$, and the variance of its active weights is $(1 - m^2)/12$. Investors differ in what they can do for themselves. The high-loading investor and the two middle investors leave their home assets unused at every positive fee. For the high-loading investor, the home asset's negative alpha outweighs its value as a hedge; for a middle investor, a home asset with zero alpha would only add risk. Each of these three types therefore holds C as her only position. The investor whose home asset is the diversifier has a useful substitute. Her free home asset has alpha 2α , three times that of C , and its active weights move with those of C . Whenever she buys C , she also holds her home asset. Applying Lemma 3 to these positions gives two demand schedules, plotted in Figure 5(a):

$$q_{0C} = q_{1C} = q_{3C} = \frac{[8\alpha - 12\delta_C]_+}{1 - m^2}, \quad (17)$$

$$q_{2C} = \frac{[6\alpha - \frac{27}{2}\delta_C]_+}{1 - m^2}, \quad (18)$$

where $[z]_+ = \max\{z, 0\}$.

The high-loading and middle investors buy less as the fee rises and stop buying when the fee reaches the fund's alpha. The diversifier investor buys less at every fee, and her demand falls faster. As C becomes more expensive, she shifts toward her home asset, which can replicate part of C 's active weights. She leaves the product at $\delta_C = 4\alpha/9$, while the others continue to buy. Above that fee she obtains the adjustment through her home asset alone. The demand difference arises despite identical preferences and beliefs: the investors have different free ways to implement the same desired portfolio. Thus a common desired portfolio need not produce common demand for the funds that help implement it.

From customers to revenue. Revenue is $R_C = \delta_C \text{AUM}_C$, with $\text{AUM}_C = (3q_{0C} + q_{2C})/4$. Raising the fee initially increases revenue, but eventually the loss of demand outweighs the higher fee per unit. The maximum occurs at $\delta_C = 10\alpha/33$, where revenue equals $(25/22)\alpha^2/(1 - m^2)$. In units of the revenue scale, this maximum is $\phi_W = 25/22$. Thus C cannot cover a normalized operating cost above ϕ_W , even when it is the only specialist available.

For a cost below the maximum, Figure 5 shows two fees that cover it. Only the first crossing can survive a cheaper offer of the same product. Every fee below that crossing earns less than cost. At the higher crossing, by contrast, lowering the fee attracts enough business to earn more than cost. Break-even alone therefore does not identify a competitively sustainable fee.

3.2.5 Equilibrium fund variety

A fee that protects C against a cheaper copy may still leave room for a differently composed fund to enter. I now let all eight non-market products compete² and look for equilibria as in Definition 1. Revenue at fees proportional to α is proportional to the revenue scale, so equilibrium fund variety depends on m , c and F only through ϕ .

²Besides C , P_L and P_R , the catalogue contains the pairs 01 and 30 and the triples 012, 230 and 301. The triples 012 and 230 have zero alpha, and the three products that hold asset 0 but not asset 2 have negative alpha. Alpha does not cap the fee an investor will pay, however: a product with zero alpha can be worth a positive fee when it offsets the risk of her other holdings. Appendix C therefore makes an exact, computer-assisted comparison across every menu, fee and tie. It covers the standardized catalogue, not every custom-designed portfolio.

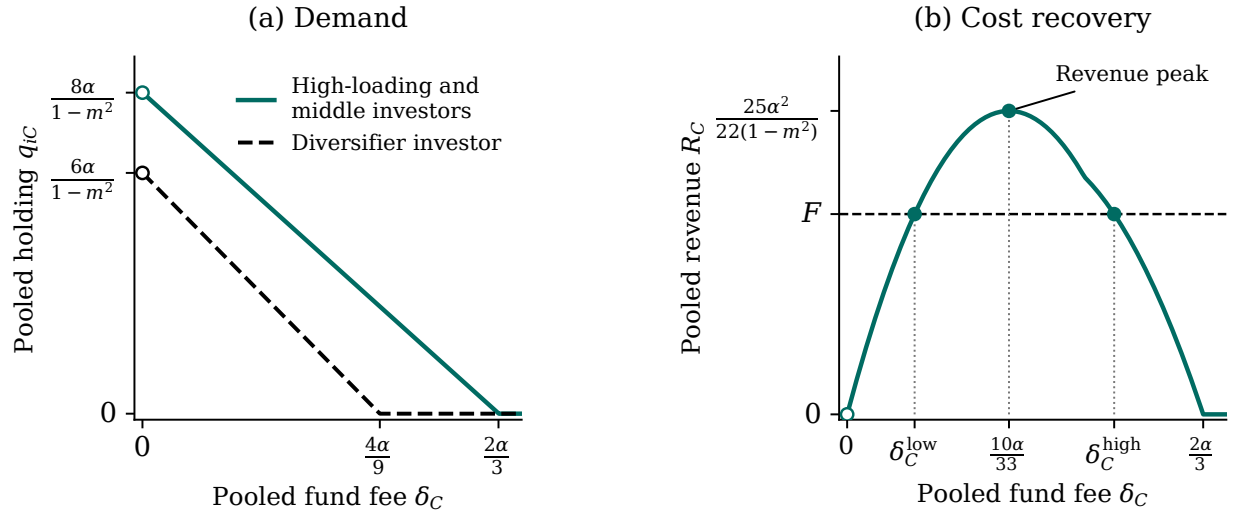


Figure 5: **Demand for the pooled fund and its revenue.** C is the only specialist. (a) Holdings of C by investor type. The high-loading and middle investors buy until the fee reaches the fund's alpha, $2\alpha/3$. The diversifier investor, who can substitute her free home asset, buys less and leaves at $4\alpha/9$. (b) Revenue, which peaks at $\delta_C = 10\alpha/33$. The horizontal line is an illustrative operating cost F . Of the two fees that cover it, only the lower one, δ_C^{low} , withstands every lower own fee. Open circles mark limits from positive fees.

Proposition 2 (Equilibrium fund variety). *In the small-dispersion approximation, let all eight non-market products compete at positive fees, and let $0 < \phi_P < \phi_U < \phi_W$ be the three cutoffs defined in Appendix C, equations (C.24), (C.25) and (C.26). The equilibrium organizations of the industry depend on normalized operating cost $\phi > 0$ as follows:*

- (i) *for $\phi > \phi_W$, the market fund operates alone;*
- (ii) *for $\phi_U < \phi < \phi_W$, the market fund operates with the pooled fund C;*
- (iii) *for $\phi_P \leq \phi \leq \phi_U$, the market fund operates either with C or with the targeted pair P_L and P_R , and both organizations are equilibria;*
- (iv) *for $0 < \phi < \phi_P$, the market fund operates with the targeted pair.*

*At $\phi = \phi_W$, the market fund alone and the market fund with C are both equilibria. No equilibrium has another set of specialists.*³

Lower cost and greater dispersion. Figure 1 displays the classification. Operating cost, measured against the revenue that dispersion creates, determines fund variety. At high relative cost the market fund stands alone. At intermediate cost one pooled fund serves several clienteles beside it. At low cost two targeted funds operate, each with a narrower mandate. As ϕ falls, the smallest and the largest number of specialists across equilibria weakly increase. Within the overlap both organizations remain possible, so a lower cost does not force a switch from pooling to targeting.

Lower operating cost and greater loading dispersion both lower $\phi = F/(\Theta(m)c^2)$. Only the ratio F/c^2 matters: doubling dispersion has the same effect as cutting operating cost to one quarter. Greater dispersion enlarges the correction investors want and raises what they will pay for each unit, so a given operating cost weighs less against the revenue a fund can earn. The model therefore links greater fund variety to cheaper fund operation and to more dispersed loadings. Conversely, with F fixed, less dispersion raises ϕ . Because a specialist can

³The proposition describes the small-dispersion approximation. In the original economy, fix m and positive ϕ in compact sets away from ϕ_P , ϕ_U and ϕ_W . For all sufficiently small positive c , no equilibrium has a set of specialists outside this classification. Each listed organization also occurs under the sufficient continuation conditions of Appendix B.9. For the pair, these include a condition that keeps the middle investors out of their home assets at the pair's equal fees. It holds at every pair cost when $m^2 \leq 1/5$; at larger m it covers only part of the pair region, or none of it. Proposition D.9 states the result. Fees, holdings and cutoffs at positive c can differ from their approximations.

cover its cost only if F is at most a constant times c^2 , specialists disappear once c is small enough. The market fund remains, selling the diversification that every investor shares.

The classification reflects three different constraints. A pooled fund needs enough revenue to finance operation; it must also leave too little business for a targeted entrant. A targeted pair must recover two operating costs at fees that neither member wants to undercut. The cost of operating one fund or two motivates this tradeoff. Demand substitution and pricing incentives determine its boundaries. The revenue boundary ϕ_W is the pooled fund's peak revenue from Section 3.2.4. No product earns more when it is the only specialist, so above ϕ_W no specialist can profitably enter against the market fund alone.

When can a targeted product enter against the pooled fund? A targeted fund puts more of each invested dollar in asset 2 than C does, and its alpha is α rather than $2\alpha/3$. Some investors can therefore improve their combined portfolio by using it even when C is already available. Entry is profitable only if those purchases generate enough fee revenue to cover the entrant's operating cost. A useful mandate can fail this test because the incumbent fund and home investment already supply much of the service. When normalized operating cost is sufficiently low, a targeted product can cover that cost and enter against the pooled incumbent. The threshold ϕ_P is the lowest cost at which the pooled fund blocks such entry. Below it, the pooled fund still covers its own cost, but a targeted entrant can cover its cost as well.

When can the targeted pair protect its fees? Each member must recover its own operating cost. But two break-even equations are not enough to establish equilibrium. If one fund lowers its fee, investors can shift business toward it while continuing to use its partner. At sufficiently high break-even fees, the additional business makes such a cut profitable.

The middle investors' home assets make this business elastic. At equal pair fees, each middle investor's home asset is worth exactly its zero fee at the margin. When P_L lowers its fee, the owner of asset 3 buys more of it and less of P_R , and she uses her free home asset to keep her exposure to asset 3 unchanged. A small cut therefore raises the cutting fund's revenue once the common fee is high enough, and the pair's break-even fee rises with operating cost. The threshold ϕ_U is the largest normalized cost at which the pair can withstand these

cuts. The upper boundary is therefore a pricing constraint, not simply a failure to collect enough revenue.

Why both organizations can survive at the same cost. The boundaries satisfy $\phi_P < \phi_U < \phi_W$. Hence there is a range in which either the pooled fund or the targeted pair can protect itself when already offered. With C as incumbent, it absorbs enough demand to prevent a targeted entrant from covering cost. With the pair as incumbent, its combined service prevents profitable pooled entry. These are alternative equilibria at the same primitives; no equilibrium includes all three specialists. The model does not choose between the two organizations in this overlap.

Fund variety in the data. The industry’s product line has this shape. Broad index funds are few, large and cheap; narrower funds are many, smaller and more expensive. In a sample of U.S. equity exchange-traded funds ending in 2017, ETFs classified as active in their design or use were 93 percent of the funds but 58 percent of the assets (Easley et al., 2021), and the broad products came first (Ben-David et al., 2023). In the model, the market fund can finance itself at costs at which no specialist can, and specialists appear only when operating cost is small relative to the revenue that dispersion creates. Search frictions (Hortaçsu and Syverson, 2004), fund-family strategies (Massa, 2003), pandering to investors’ beliefs (Gennaioli et al., 2015; Ben-David et al., 2023) and screening (Cvitanić and Hugonnier, 2022) also produce many funds, and in Gorton and Pennacchi (1993) clientele size determines whether uninformed investors hold clientele-specific composites or a market composite combined with individual securities. Some of these mechanisms sustain funds with identical portfolios, which contestable pricing rules out here. The model instead restricts which portfolios are packaged: every specialist omits the asset that carries the most common risk and holds the best diversifier.

These restrictions yield predictions for markets with the same average loading, access structure and product technology. Cheaper fund operation or more dispersed loadings lower ϕ and weakly raise both the smallest and the largest number of specialists across equilibria. The same comparison applies when an industry’s costs fall or its loadings disperse over time. At a high enough cost relative to dispersion only the market fund operates, which is consistent with the broad-first history without appeal to investors’ beliefs.

4 Fees, performance, and investor portfolios

Section 3 derived which funds survive competition: a market fund that every investor holds and, when operating cost is low relative to the revenue that dispersion creates, either one pooled specialist or two targeted ones. This section asks what that industry implies for three things one can observe: the fees and sizes of funds, their measured performance, and the portfolios investors hold. Evidence on each motivates the question; it does not test the model’s mechanism.

Throughout, I describe the equilibria of Proposition 2 in the small-dispersion approximation of Section 3.2.2, which gives fees, holdings and alphas by their leading terms in c . Revenue accounting holds exactly. Appendix E derives the results and records which of them hold exactly and which hold in every equilibrium for sufficiently low c .

4.1 Fund size and fees

Larger funds tend to charge lower fees. Across 46,580 fund share classes sold in 18 countries, Khorana et al. (2009) find that larger funds and larger fund families charge lower fees, and that index funds charge less than actively managed funds. Among exchange-traded funds, the first products tracked broad indexes and offered diversification at low cost, while later ones track niche portfolios and charge high fees (Ben-David et al., 2023).

Why would an industry look like this if every fund had the same operating technology and no manager had skill? Fixed operating costs already favor size, because a fund that sells more spreads its cost over more assets. What the cost argument leaves open is why one fund sells so much more than the others. In the model, the answer is demand.

The market fund sells the diversified position that every investor wants, so every investor holds a substantial amount of it. A specialist sells only a modest adjustment to that position’s risk mix (Section 3.2.1). The market fund therefore manages far more assets than any specialist.

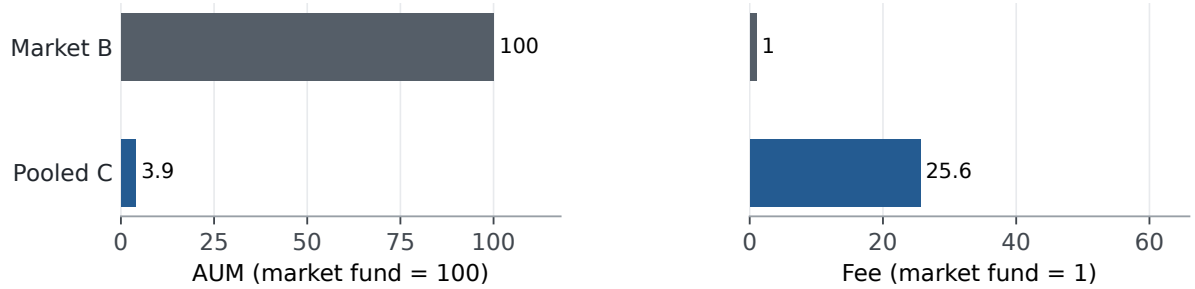
Every fund must nevertheless recover the same operating cost. Free entry leaves no fund with a profit, so each active fund’s fee revenue equals exactly F (Lemma 1), and fees are inversely proportional to size. The market fund spreads its cost over the common position and charges a tiny fee. A specialist recovers the same cost from a small adjustment and must charge far more.

Proposition 3 (Fund size and fees).

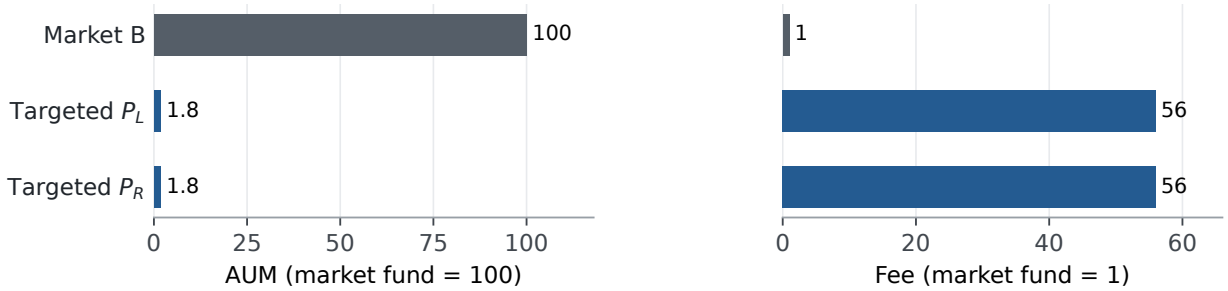
- (a) *Each active fund's fee times its assets under management equals its operating cost F . Fee ratios are therefore inverse size ratios.*
- (b) *The market fund is larger than every specialist and charges a lower fee.*
- (c) *When the two targeted funds charge different fees, the cheaper one is the larger.*

Figure 6 shows assets under management and fees in the pooled and the targeted organization, separately and at the same illustrative parameters.

Pooled organization



Targeted organization



Each fund covers the same operating bill: $\text{fee} \times \text{AUM} = F$.

Figure 6: **Fund size and fees.** Assets under management (market fund = 100) and fees (market fund = 1) in the pooled organization and, separately, in the targeted organization with equal fees. The two organizations are alternative equilibria at the same illustrative parameters (Appendix E.4). Each fund's fee times its assets under management equals the same operating cost. Small-dispersion approximation.

The inverse relation is a benchmark, not a fit. With only a fixed cost, a fund twice as large charges half the fee. Observed fees fall far more slowly with size. In the estimates

of [Khorana et al. \(2009\)](#), a fund at the 75th percentile of size, about sixteen times larger than one at the 25th percentile, has an expense ratio only 19 basis points lower. Costs that grow with a fund’s assets, which the model omits, would flatten the relation. What the model contributes is the difference in demand that makes one fund large and the others small. Because a specialist’s adjustment grows with dispersion, the difference narrows within an organization where loadings are more dispersed: at a given average loading and operating cost, more dispersed loadings make pooled and symmetric targeted specialists larger and cheaper, as a lower operating cost does at given dispersion (Appendix E).

Equal revenue does not require equal fees. The targeted organization also has equilibria in which the two pair funds charge different fees (Appendix E.1). The two funds hold different assets but have the same expected return, variance and beta. A middle investor whose home asset belongs to the more expensive fund can replace part of it with the cheaper fund and her home asset. The cheaper fund therefore serves more demand and is larger, and each fund still recovers F . These are different portfolios with equal return moments, not duplicate funds that track the same index. Duplicates cannot charge different fees here, since a cheaper copy would take all their investors; the fee dispersion among S&P 500 index funds that [Hortaçsu and Syverson \(2004\)](#) document requires non-portfolio differentiation or search frictions, both outside the model.

4.2 Performance without skill

Even passive indices receive nonzero measured alpha. [Cremers et al. \(2013\)](#) report that the Carhart four-factor model assigns the S&P 500 index an annual alpha of 0.82 percent and the Russell 2000 index an annual alpha of -2.41 percent over 1980–2005, so that index funds tracking them would appear to have significant skill, positive or negative. [Berk and van Binsbergen \(2015\)](#) measure a fund against the index funds its investors could actually buy. They note that this benchmark credits funds with the value of diversification services, and they separate that component from stock picking.

What would a standard evaluation report about an industry in which nobody has skill, and what would it be measuring?

In the model, every asset has the same expected return. This is an assumption, not a result, and it isolates what covariance does. A fund’s alpha against the market fund therefore

comes from its beta alone. Each specialist tilts toward the diversifier and away from the high-loading asset, so it carries less of the market fund's risk and earns positive alpha before fees (Section 3.2.1). A targeted fund tilts further than the pooled fund and has the higher alpha.

Measured performance also depends on the benchmark. Against a costless index with a fund's own weights, every fund has zero alpha before fees and alpha $-\delta_T = -F/\text{AUM}_T$ after fees: all funds underperform, and the smallest underperform most. That benchmark ignores what an investor would pay to build the index herself, which is exactly the service a fund sells. Against the market fund, the same specialists outperform.

Specialists keep positive alpha after fees, although funds merely break even. Each investor's demand is limited by the risk of her whole portfolio: she buys a specialist until its fee equals the value of one more unit, and that value is the fund's alpha less the extra risk the unit adds to her portfolio (Lemma 3). The fee therefore stays below the fund's alpha.

Why does competition stop there? Free entry disciplines revenue, and revenue is the fee times assets under management. A fee at which revenue exceeds operating cost invites a cheaper copy, and no fee below the equilibrium level earns more than operating cost, so no copy or entrant offers one (Section 2.5). The fee is therefore pinned at the lowest level that recovers the fixed cost. Alpha after fees is a different object: it is what the last unit of the fund is worth to the investor who holds it, and at the pinned fee it equals the extra risk that unit adds to her portfolio. Competition among funds drives profits to zero; investors' aversion to the risk of their last unit sets the alpha that remains.

Proposition 4 (Performance without skill).

- (a) *Against the market fund's return before fees, every specialist has positive alpha before and after fees, and a targeted fund's alpha before fees exceeds the pooled fund's. The market fund's own alpha is zero before fees and $-\delta_B$ after fees.*
- (b) *Against a costless portfolio with its own weights, every fund's alpha is zero before fees and minus its fee after fees.*
- (c) *When the two targeted funds charge different fees, their alphas before fees are equal, and the cheaper fund's alpha after fees is higher.*

Figure 7 compares the pooled fund and a representative targeted fund under the two benchmarks; the two funds belong to alternative organizations.

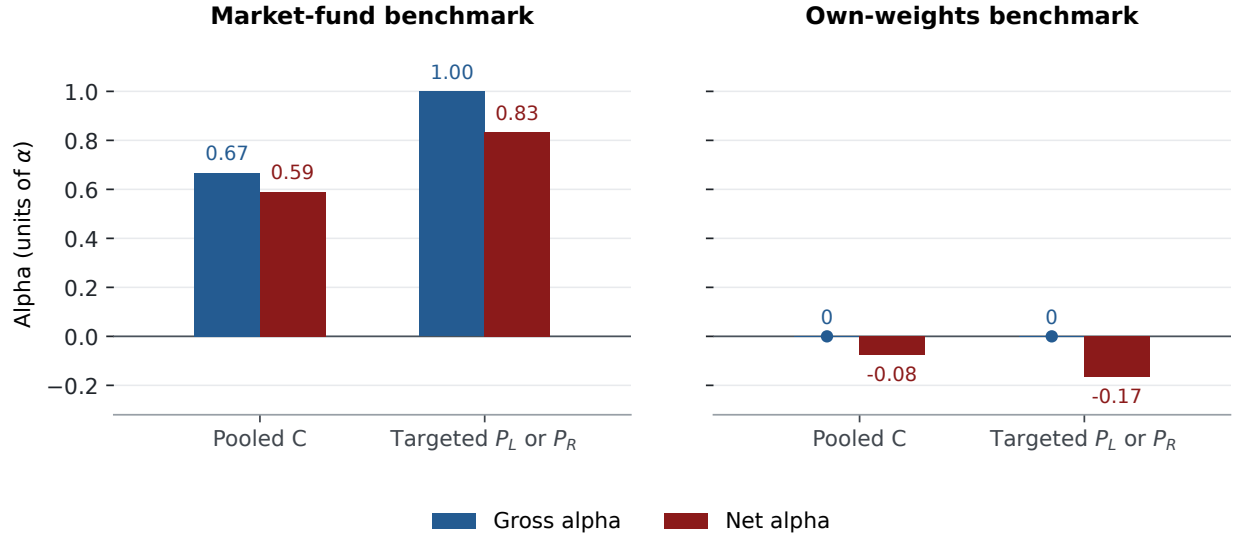


Figure 7: **Performance against two benchmarks.** Alpha before fees (gross) and after fees (net) of the pooled fund and a representative targeted fund, in units of α , measured against the market fund's return before fees (left) and against a costless portfolio with the fund's own weights (right). The pooled and targeted funds belong to alternative organizations. All assets have the same expected return. Small-dispersion approximation, at the parameters of Figure 6.

Positive alpha does not mean higher returns. Before fees every fund has the same expected return, so a specialist, with its higher fee, earns less than the market fund after fees. Its alpha measures a covariance comparison with the market fund's equal weights. A skill measure benchmarked only against a broad index would credit passive specialists with this diversification service, which is the component [Berk and van Binsbergen \(2015\)](#) separate. The model adds an equilibrium account of that service: which specialists exist, what they charge, and how much measured alpha remains after fees.

Active share, size and alpha. [Cremers and Petajisto \(2009\)](#) measure active share, the fraction of a fund's holdings that differs from its benchmark index. Funds with the highest active share tend to be small and outperform their benchmarks before and after expenses, and larger funds earn less after fees, as [Chen et al. \(2004\)](#) also find. These studies read the patterns as stock-picking skill and diseconomies of scale, and [Cremers and Petajisto \(2009\)](#) measure each fund against a fund-specific benchmark index. The model produces a related

pattern against a common benchmark with neither; it is an additional mechanism, not an account of their results.

Corollary 1 (Active share, size and alpha). *For positive loading dispersion, measure active share against the market fund's weights and alpha against its return before fees. The market fund, the pooled fund and each targeted fund have active shares 0, $1/4$ and $1/2$, and their alphas before fees rank in the same order. At the same primitives, the pooled fund is larger and cheaper than either targeted fund and has lower alpha after fees, and the market fund is larger and cheaper than both. Two targeted funds that charge different fees have the same alpha before fees, and the smaller one has the lower alpha after fees.*

Alpha measures how much of the correction one unit of a fund delivers, and a targeted fund delivers more per unit. The targeted organization, however, splits the correction between two funds, so each has a smaller customer base than the pooled fund. Each fund must still recover the same operating cost, so a targeted fund charges more and is smaller, yet it keeps more alpha after fees than the pooled fund (Appendix E.2).

The ranking is not a law that size erodes performance. Active share alone does not earn alpha: other catalogue products depart just as far from the market portfolio but tilt toward higher-loading assets and have negative alpha. Those products do not survive. Within one industry, the comparison is between the market fund and its specialists. A fund with fixed weights and return moments has the same alpha before fees at any size, so among funds with the same return moments the smaller one earns less after fees, as the corollary's last sentence shows. Within an organization, a lower operating cost makes pooled or symmetric targeted specialists larger with unchanged alpha before fees, and more dispersed loadings at a given average loading and operating cost make them larger with more alpha before fees.

A test could measure active share and alpha against one broad benchmark and ask whether alpha follows the covariance direction of each fund's departure rather than its size; in data this is a joint test with the model's premise that expected returns do not rise with loadings. A benchmark built from passive rules specified in advance, such as the indexes that specialists track, would absorb the alpha from those rules. Unlike a copy of the fund's realized holdings, it would not mechanically absorb returns from security choices within a rule. An empirical comparison could therefore examine how measured skill changes when specialist index products enter the benchmark set.

4.3 Combining direct investment and funds

Households combine concentrated direct holdings with diversified funds. Earlier studies found that households' portfolios of individual stocks are poorly diversified. Using complete household portfolios in Sweden, [Calvet et al. \(2007\)](#) show that mutual funds and cash dominate direct stockholdings in many portfolios, which limits the losses from concentrated stock accounts; 76 percent of the households that hold stocks directly also own mutual funds. Whether a household is diversified can only be judged from its whole portfolio.

Why would an investor hold a concentrated direct position alongside diversified funds, and why would she hold several funds rather than one?

Consider the pooled organization and the targeted organization with equal fees. The investor whose free home asset is the diversifier, the low-loading asset 2, can obtain part of the specialist's service directly: exposure that moves her portfolio away from common risk. She holds the home asset alongside the market fund and buys less of each specialist than other investors do. A specialist is worth less to her because her own alternative already contains part of its service. Her direct account holds a single asset, yet her total portfolio is diversified; judged by the direct account alone, it would look concentrated.

In these organizations, the other investors obtain the adjustment through funds. The high-loading asset adds common risk, and a middle asset does not by itself move a portfolio away from common risk, so these investors hold no direct position and buy the specialists. The same fund is therefore worth different amounts to different investors, depending on which assets they can hold directly at low cost.

The targeted organization shows why investors hold several funds. Each pair fund holds the diversifier together with one middle asset, so either fund alone would tilt the portfolio toward that middle asset. Equal holdings of both deliver the adjustment without the tilt. Funds sell standardized pieces, and investors combine them. When the pair's fee is high, the diversifier stops buying the pair funds and relies on her home asset instead; [Appendix E.3](#) gives the cutoff.

Proposition 5 (Investor portfolios). *In the pooled organization and in the targeted organization with equal fees:*

(a) *every investor holds the market fund;*

- (b) *the high-loading and middle investors hold every specialist on the menu and no direct position;*
- (c) *the diversifier holds her home asset and less of each specialist than the other investors. With the targeted pair, she holds neither pair fund when their fee exceeds a cutoff.*

Table 1: **Who holds what.** Positions each investor holds besides the market fund, in the pooled organization and in the targeted organization with equal fees. “Home” is the investor’s free home asset. Appendix E.3 gives the cutoff between low and high pair fees and the holdings at unequal fees.

Investor	Pooled	Targeted, low fee	Targeted, high fee
High-loading and middle	C	P_L, P_R	P_L, P_R
Diversifier	less C , home	less P_L, P_R , home	home

Within these organizations, the investor who can hold the diversifier for free buys less from the specialists. The prediction concerns covariance, not familiarity: a cheaply held asset replaces part of these specialists only when it supplies the diversification that they sell. Holding several funds has other explanations too; retirement-plan participants often split their contributions equally among the funds they choose (Benartzi and Thaler, 2007). Here, investors hold both targeted funds because each offsets the other’s tilt toward a middle asset, so at equal fees they hold equal amounts, and the diversifier investor holds less of both. Examining these predictions would require both the direct and the delegated holdings of each investor and information about which exposures she can hold cheaply. Wealth, information and preferences would remain alternative explanations for the same pattern.

5 Equilibrium at finite dispersion

The propositions describe the industry when loading dispersion is low. Stock loadings, however, differ materially: portfolios of stocks with similar volatility still have visibly different market betas. This section asks whether the organizations of Proposition 2, the size and fee ranking, the specialists’ positive alpha and the investors’ portfolios still describe the equilibria of the original economy at the dispersion that U.S. stock returns suggest.

To answer it, I solve the original economy numerically. For given fees, each investor chooses which assets to access directly and how much of each fund and asset to hold. Each

fund must break even, and no cheaper copy and no entrant may earn a profit (Definition 1). At every parameter point the search covers all 512 menus of the nine products. It uses many starting fees and dense grids of deviating quotes, but it is finite: I call an equilibrium it reports *computed*, and a menu at which it finds none is a nonfinding, not a proof of nonexistence. Appendix F documents the protocol and all results.

5.1 Parameters from stock returns

Two covariance parameters are tied to stock-return moments. In the model the average pairwise correlation of the four assets equals $m^2 - c^2/6$. Pollet and Wilson (2010) report an average correlation of .237 among the 500 largest U.S. stocks. Dispersion is measured relative to the average loading. Within rows of stocks sorted on volatility, the realized betas that Asness et al. (2020) report across correlation-sorted portfolios imply c/m between about one quarter and one third. Together the two moments give $m \approx .49$ and $c \approx .12$ to .16. I compute at $c = .12, .15$ and .16, rounded values in this range. At $c = .15$ the loadings are .64, .49, .34 and .49, and pairwise correlations range from .17 to .31.

The other inputs are not estimated from returns; Table 2 summarizes their roles. The access cost of .05 is an illustrative value: it translates the participation cost and risk aversion estimated by Gálvez and Paz-Pardo (2026) into the model’s units under stated conventions, and .02 and .10 serve as sensitivity checks. Operating cost is not estimated. I report results at twelve values of ϕ between .05 and 1.5, which span every organization in Proposition 2. To express fees and alphas in annual basis points, I multiply the model’s normalized fees by an equity premium of 4.5 percent, close to the long-horizon premium in the survey of asset managers by Dahlquist and Ibert (2024) and to the option-implied bound of Martin (2017). The premium converts units; it does not enter the equilibrium.

5.2 Results at the empirical anchor

At the empirical anchor the finite economy looks like the approximation. The anchor has 36 parameter points, three dispersions and twelve operating costs at $m = .49$ and access cost .05. At every point the search finds at least one equilibrium, and every equilibrium it finds is one of the three organizations of Proposition 2: the market fund alone, the market fund with the pooled fund, or the market fund with the targeted pair. Panel (a) of Figure 8

Table 2: **Parameters.** Inputs of the numerical computations at the empirical anchor. Only the average loading and the loading dispersion are tied to stock-return moments. Appendix F gives the mappings, the full source ledger and the alternative access-cost convention.

Model object	Values	Basis
Average loading m	.49	Average stock correlation (Pollet and Wilson, 2010), with the dispersion proxy
Loading dispersion c	.12, .15, .16	Beta dispersion within volatility-sorted portfolios (Asness et al., 2020)
Access cost	.05 (sensitivity .02, .10)	Illustrative translation of participation cost and risk aversion (Gálvez and Paz-Pardo, 2026)
Operating cost ϕ	Twelve values, .05–1.5	Varied; not estimated
Premium for annual units	4.5%	Survey and option-implied premia (Dahlquist and Ibert, 2024 ; Martin, 2017)
Assets, catalogue, investors	Four assets, nine equal-weight products	Maintained structure

shows where each occurs. The pattern is the same at all three dispersions. The targeted organization is computed at every tested cost up to .72, the pooled organization from .16 to 1.18 and the market fund alone at 1.5, so both specialist organizations are equilibria at the four tested costs from .16 to .72. The costs at which the organization changes lie slightly above the cutoffs of the approximation: at $c = .15$, near .157, .762 and 1.213, against .140, .735 and 1.136. At every tested cost but one, the computed organizations are those that the approximation predicts. The exception is $\phi = 1.18$, which lies between the approximate and the computed upper limit of the pooled organization.

Fees, sizes and alpha keep the pattern of Propositions 3 and 4. In every computed equilibrium the market fund is larger and cheaper than every specialist, managing between 4.1 and 25.3 times the assets of each. Because every fund’s fee revenue equals the same operating cost, these are also the ratios of the specialists’ fees to the market fund’s. Every specialist earns positive alpha after fees against the market fund’s return before fees, between 5.7 and 20.4 basis points a year. Table 3 compares the pooled and the targeted organization at the same primitives. The two columns are alternative equilibria, not funds observed together. The market fund manages 4.7 times the assets of the pooled fund and 10.8 times those of each targeted fund. Before fees, the pooled fund earns an alpha of 11.9 basis points a year and each targeted fund 19.4, and both keep most of it after fees. The fees themselves are far

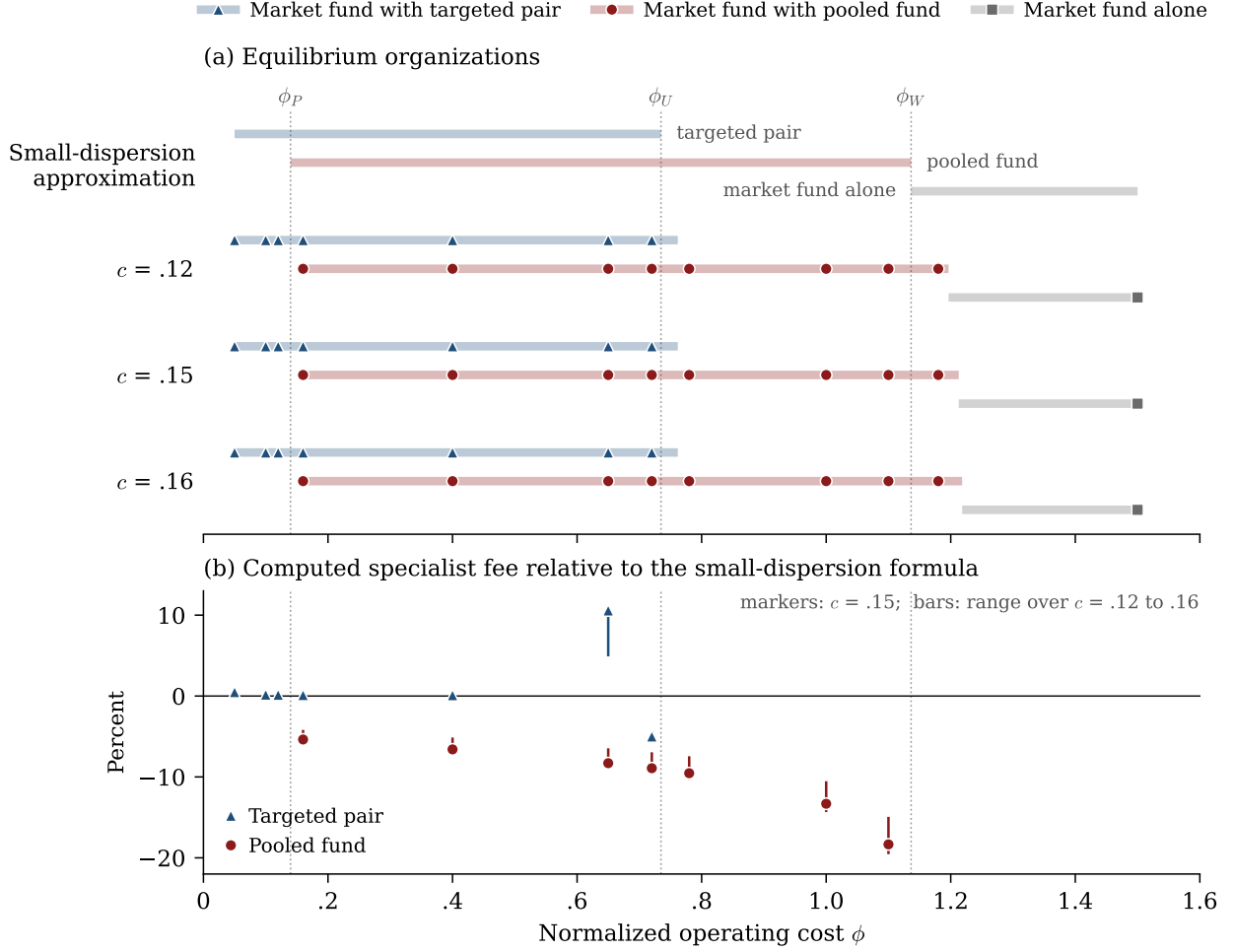


Figure 8: **Organizations and accuracy at the empirical anchor** ($m = .49$, access cost .05). Panel (a): equilibrium organizations by normalized operating cost ϕ . Markers show every organization computed at the twelve tested costs by the search over all menus. Bars show each organization's range between its transitions, within the tested range: for the small-dispersion approximation, the cutoffs ϕ_P , ϕ_U and ϕ_W of Proposition 2 (dotted lines); for $c = .12, .15$ and $.16$, computed transition brackets narrower than .01, which test only the menus of the three organizations. Panel (b): deviation of the computed specialist fee from its small-dispersion formula, in percent of the formula, in symmetric targeted and in pooled equilibria; markers at $c = .15$, bars spanning $c = .12$ to $.16$. At $\phi = 1.18$ the approximation has no pooled organization. Computed on finite grids, not certified.

below observed ones: the median exchange-traded fund charges 18 basis points (Khomyn et al., 2024). In the model a fund’s only cost is a fixed operating cost, so the computed fees are those of a pure diversification service.

Table 3: **Two organizations at the same primitives.** $m = .49$, $c = .15$, access cost .05, $\phi = .40$. Annual fees and alphas in basis points at a 4.5 percent premium; alphas against the market fund’s return before fees. The columns are alternative equilibria.

	Pooled organization	Targeted organization
Market-fund assets / specialist assets	4.71	10.76
Specialist fee / market-fund fee	4.71	10.76
Specialist fee	1.06	2.36
Market-fund fee	.226	.219
Specialist alpha before fees	11.91	19.35
Specialist alpha after fees	10.85	16.99

Investors’ portfolios follow Proposition 5 in its economic content, with small departures in detail. In every pooled equilibrium, every investor holds the market fund; the high-loading and middle investors hold the pooled fund and no direct position; and the diversifier investor holds her home asset and less of the pooled fund. In symmetric targeted equilibria the middle investors add small direct holdings of their home assets, at most 0.4 percent of their risky exposure, at tested costs up to .65 for $c = .12$ and up to .40 for $c = .15$ and .16. The condition of Appendix B.9 that keeps middle investors out of their home assets at equal pair fees holds at this loading only above about .66; its failure permits such holdings but does not require them. At the lowest tested cost, where the pair funds are cheapest, the diversifier investor holds none of her home asset and the same portfolio as the high-loading investor. No investor pays for direct access to another asset. The whole-portfolio account of Section 4.3 carries over; the exact holdings of each investor are finite-dispersion refinements.

Panel (b) of Figure 8 shows how accurate the small-dispersion formulas are, separately for the two organizations. For the targeted organization they are close: at tested costs up to .40, the computed specialist fee lies within 0.7 percent of its formula. Larger departures, up to 11 percent, occur at .65 and .72, near the upper limit of the targeted organization. For the pooled organization the computed specialist fee lies 4 to 19 percent below its formula, and the gap widens with operating cost. The market fund’s assets are within 3 percent of

the approximation for the computed organization, so its fee is close to operating cost divided by those assets. The limiting formula F/Q_0 , which ignores the assets that specialists draw from the market fund, is less accurate: the computed fee exceeds it by 8 to 24 percent in the specialist organizations. The small-dispersion specialist fees therefore describe the targeted organization well and the pooled organization roughly.

The access cost matters mainly at high operating cost. Sensitivity computations at access costs .02 and .10 solve the menus of the three organizations, not every menu, at $c = .15$ and .16 and five operating costs. At .10 they find the same organizations as at .05. At .02 the market fund alone becomes an equilibrium at high operating costs where only pooled equilibria are found at .05 (at 1.18, and at 1.00 for $c = .16$). In these equilibria the middle investors pay to hold the low-loading asset directly: cheaper access lets them buy part of the correction themselves.

The numerical evidence therefore supports the paper’s account at the loading and dispersion suggested by U.S. stock returns. Every computed equilibrium belongs to the three organizations, and their fees, alphas and portfolios behave as Sections 3 and 4 describe, with the pooled organization’s fees approximated less accurately than the targeted organization’s.

6 Robustness

Two assumptions shape the industry of Sections 3 and 4: every fund has the same operating cost, and every fund holds fixed, equal weights within its mandate. This section relaxes each assumption in turn. The first check asks whether specialists’ high fees could reflect higher operating costs rather than small demand. The second asks how much of the specialist business remains when funds may choose their weights within a mandate.

6.1 Higher specialist operating costs

A specialized mandate may require more fixed work than a broad index: its own index, its own data and its own portfolio operations. If so, a specialist’s high fee could reflect a higher operating cost rather than a small customer base, and the demand-based account of Section 4.1 would be incomplete. I ask which conclusions depended on every fund having the same operating cost.

Let the market fund cost F_B to operate, and let every non-market product cost

$$F_S = \kappa_F F_B, \quad \kappa_F \geq 1.$$

One premium κ_F applies to all eight products, whether active, a cheaper copy or an entrant. Costs remain fixed with respect to assets under management. Weights, access costs, the investor's problem and the pricing tests are those of Section 2.

Accounting comes first. Free entry still leaves each active fund with zero profit, so its fee revenue equals its own operating cost, $\delta_T \text{AUM}_T = F_T$. Between the two tiers, therefore,

$$\frac{\delta_S}{\delta_B} = \kappa_F \frac{\text{AUM}_B}{\text{AUM}_S}. \quad (19)$$

A specialist that costs twice as much to operate as the market fund and manages one tenth of its assets must charge twenty times its fee. With a common cost the fee ratio would equal the size ratio; the premium multiplies it by κ_F . Among specialists, which share one cost, revenue is still equal and fees are still inversely proportional to size.

The identity takes the difference in size as given. It does not say whether the specialist attracts those assets, or whether it can operate at all. Those are questions about demand, and investors never pay operating costs directly: they respond only to fees. The premium therefore enters the equilibrium only through the funds' break-even, cheaper-copy and entry tests.

In those tests the two tiers separate. The market fund's own tests involve only its own cost. The argument of Section 3.1.3 gains strength: investors who replace the market fund with several funds must finance several operating costs, and with a premium each of them costs $\kappa_F F_B$. The specialists' tests compare their revenue with F_S . Their normalized operating cost is the measure of Section 3.2.3 evaluated at their own cost,

$$\phi_S = \frac{F_S}{\Theta(m)c^2} = \frac{\kappa_F F_B}{\Theta(m)c^2}. \quad (20)$$

Proposition 6 (Specialist cost premium). *Let every non-market product cost $F_S = \kappa_F F_B$, with $\kappa_F \geq 1$ fixed.*

- (a) *In every equilibrium, each active fund's fee times its assets under management equals its own operating cost, and (19) holds.*
- (b) *Under Assumption 2 applied to F_B , with a ceiling that may depend on κ_F , there is a positive bound on c below which every equilibrium contains the market fund, held by every investor type. No investor pays for non-home access. The market fund charges about F_B/Q_0 , and every active specialist is smaller and charges a higher fee.*
- (c) *In the small-dispersion approximation, the equilibrium organizations are those of Proposition 2 with ϕ replaced by ϕ_S .⁴*

The proposition separates what the premium changes from what it leaves alone. Demand is unchanged, and so is fund variety at a given ϕ_S : the same cutoffs, the same organizations, and the same normalized specialist fees and holdings. What changes is the accounting between the tiers, and which organizations a given pair of costs can finance.

Two comparisons should be kept apart. In a given economy, with m , c and F_B fixed, a higher premium raises ϕ_S in proportion to κ_F . For the specialists' viability, and hence for the set of organizations, this acts like a fall in dispersion by the factor $\sqrt{\kappa_F}$; the analogy does not extend to fees and holdings in original units. The higher premium moves the organization toward fewer specialists, although the overlap in Proposition 2 admits more than one organization, so there is no unique path. The market fund's fee, about F_B/Q_0 , does not change to leading order. The local classification instead holds ϕ_S and κ_F fixed as c falls, so both costs fall with c^2 . With both costs fixed in levels below the ceilings of part (b), specialists disappear once dispersion is low enough, as they do with a common cost.

The premium steepens the relation between fee and size across the two tiers. With a common cost, a specialist with one tenth of the market fund's assets charges ten times its fee; with a premium it charges $10\kappa_F$ times the fee. A higher specialist cost therefore cannot

⁴Part (c) describes the small-dispersion approximation. In the original economy, fix m , κ_F and positive ϕ_S in compact sets away from ϕ_P , ϕ_U and ϕ_W . For all sufficiently small positive c , no equilibrium has a set of specialists outside the classification at ϕ_S . Each listed organization also occurs under the continuation conditions of Appendix B.9, with one change. In the condition that keeps the middle investors out of their home assets at the pair's equal fees, the market fund's fee enters through its own normalized cost, ϕ_S/κ_F . Appendix G gives the proofs.

explain why observed fees fall much more slowly with size than the fixed-cost benchmark of Section 4.1 predicts. Performance is unchanged where it depends on weights, and alphas before fees are the same. On the regions of the local classification, with ϕ_S in a compact positive set and c sufficiently small, specialists keep positive alpha after fees, and each specialist's alpha before fees times its assets exceeds twice its own operating cost.

6.2 Portfolio choice within mandates

The market fund leaves a correction for specialists because its weights are fixed. Equal weights overweight the high-loading asset (Section 3.2.1), and a sponsor free to improve them would sell part of the correction itself. How much of the specialist business would remain? The question also runs the other way: specialists that can redesign their own weights may displace the standardized organizations of Proposition 2. I take the broad fund first.

Suppose a broad fund's mandate lets it move a share s of its portfolio from the high-loading asset to the diversifier:

$$b_B(s) = b_B + s(\iota_2 - \iota_0), \quad 0 \leq s \leq \bar{s} \leq \frac{1}{4}. \quad (21)$$

The share s is a percentage-point transfer and equals the fund's active share against the equal-weight index b_B . The direction uses only the commonly known covariance, so choosing it requires no skill. Each choice of s is a distinct product, and every fund has the same operating cost F . A fund commits publicly to its weights and its fee before investors choose. Entry is unrestricted within the mandate. Any sponsor may offer any permitted weights at any fee or copy existing weights at a lower fee, and several broad funds with different weights may operate. The equal-weight index remains available but need not be chosen. Specialists keep their standardized weights.

The full correction requires the weight shift

$$s^* = \frac{mc}{2(1 - m^2)}. \quad (22)$$

Every investor holds about Q_0 units of a broad fund, so this shift moves the $2\alpha/(1 - m^2)$ units of exposure from asset 0 to asset 2 that she would choose with free access to all assets (Section 3.2.3). The shift is proportional to dispersion. I measure a broad fund's tilt as a share

of the full correction, $\theta = s/s^*$, and write $\bar{\theta} = \bar{s}/s^*$ for the cap. Throughout this subsection I use the small-dispersion approximation of Section 3.2.2, extended to broad funds with different weights. It maintains for the broad tier what Section 3.1 proves for the equal-weight market fund: broad funds carry the common position at fees of order c^2 , while specialists charge fees of order c .

Every investor wants the broad fund to tilt further, up to the full correction. A larger tilt delivers part of the correction inside the common position, without a specialist's fee. At given specialist fees, each investor's cost of completing her portfolio therefore falls as θ rises toward one, and rises beyond it. Competition among broad funds turns this shared preference into the design. Suppose that the only broad fund's tilt stops short of the cap and of the full correction, and that the fund breaks even. An entrant whose tilt is somewhat closer to the full correction, charging slightly more, is preferred by every investor. It takes the whole common position and earns more than its cost. A fund that overshoots the full correction loses in the same way to an entrant that tilts slightly less. Suppose instead that several broad funds operate. Each serves only part of the common position, so each must charge more than a single fund would need. An entrant with the best permitted weights and a fee between the two takes every investor. A fund with worse weights cannot undercut the selected one either. At the same or a higher fee no investor switches, and at a lower fee it cannot cover its cost, because the broad funds together serve a fixed common position.

Proposition 7 (Broad-fund weights). *In the small-dispersion approximation with broad funds of every permitted tilt, and with the cap $\bar{\theta}$ fixed relative to the full correction, every equilibrium has exactly one broad fund. Its tilt is the permitted tilt up to the full correction, $\theta^* = \min\{\bar{\theta}, 1\}$, and every investor holds it at a fee of about F/Q_0 . If $\theta^* < 1$, the specialists form an equilibrium of Proposition 2 at normalized operating cost*

$$\phi_{\text{eff}} = \frac{\phi}{(1 - \theta^*)^2}, \quad (23)$$

with that equilibrium's fees and holdings, evaluated at ϕ_{eff} , scaled down by the factor $1 - \theta^$. If $\theta^* = 1$, no specialist has positive demand. Conversely, every such configuration is an equilibrium.*

Appendix H gives the proof, together with the two specialist-redesign examples below.

The factor $(1 - \theta^*)^2$ has a simple source. With the broad fund supplying the share θ^* of the correction, specialists sell the remaining share $1 - \theta^*$. At fees scaled by $1 - \theta^*$, investors buy $1 - \theta^*$ times the quantities of Section 3.2.2, so a specialist's revenue potential falls with the square of the unmet correction. A fixed operating cost therefore weighs on specialists as if it were $\phi/(1 - \theta^*)^2$. A broad fund that supplies half of the correction leaves one quarter of the specialist revenue scale.

Every cutoff ϕ_X of Proposition 2 becomes $\phi_X(1 - \theta^*)^2$. In particular, for $0 < \phi < \phi_W$, specialists operate in every equilibrium when $\theta^* < 1 - \sqrt{\phi/\phi_W}$, and the broad fund stands alone in every equilibrium when

$$\theta^* > 1 - \sqrt{\phi/\phi_W}. \quad (24)$$

At equality, as at ϕ_W in Proposition 2, the broad fund alone and the broad fund with the pooled specialist are both equilibria. Specialists therefore exit before the broad fund removes the whole correction. Some demand for the correction remains, but it can no longer finance another operating cost. Only at $\theta^* = 1$ does the demand itself vanish.

Because the proposition fixes $\bar{\theta}$, the cap $\bar{s}$ in actual weights is proportional to c . A cap fixed in actual weights behaves differently. Because s^* is proportional to c , any fixed positive cap admits the full correction once dispersion is low enough, and the extended approximation then leaves no specialist demand. I do not establish this fixed-cap conclusion in the original economy, where the permitted designs become unbounded relative to the correction.

Standardized products can replicate the tilted broad fund: $b_B(s) = (\frac{1}{2} - 2s)b_{01} + 2sb_{12} + \frac{1}{2}b_{23}$. Its selection therefore does not rest on offering a new exposure. Its advantage is the one Section 3.1.3 identifies for the market fund: one fund supplies the common position at lower cost than several. The extended approximation maintains that comparison for the broad tier.

Two conclusions of Section 4 change. First, in the pooled organization and in the targeted organization with equal fees, specialists become smaller and more expensive as the broad fund's tilt grows at a given operating cost, until they exit. A larger tilt raises ϕ_{eff} , which raises the baseline fee and lowers baseline demand. Each specialist's assets fall, and its fee rises because its revenue must still cover F . Fees remain inversely proportional to size. Second, measured alpha now depends on which broad portfolio serves as the benchmark. Against the equal-weight index a specialist's alpha before fees is unchanged. Against the selected broad fund it is $1 - \theta^*$ times as large, and it remains positive after fees. The broad fund itself has alpha $4s\alpha$ against the index by (14), and this alpha exceeds its fee when $\phi < 8\theta^*$.

Now keep the broad fund at the equal-weight index and let every specialist mandate tilt in the same direction: toward the lower-loading assets it holds, in proportion to each asset's loading gap from the mandate's average. Each choice of weights is again a separate product. Two exact results in the small-dispersion approximation show which organizations of Proposition 2 depend on standardized weights.

First, arbitrarily small discretion displaces the equal-weight pooled fund. Take any pooled equilibrium fee in the interior of its range. An entrant holding the same three assets, with $2t/9$ more weight on the diversifier and $t/9$ less on each middle asset, can charge slightly less than C does. For small $t > 0$, investors drop C altogether and the entrant earns more than its cost. The example rules out the equal-weight pooled design, not pooling itself. I do not characterize the redesigned pooled fund that competition would select.

Second, a pooled mandate with enough discretion can bundle the targeted pair. The fund $C^* = (0, \frac{1}{4}, \frac{1}{2}, \frac{1}{4}) = \frac{1}{2}(b_{12} + b_{23})$ holds exactly what equal holdings of the two targeted funds hold. At equal pair fees every investor holds the two pair funds in equal amounts (Section 4.3), so one unit of C^* can replace half a unit of each while financing one operating cost instead of two. An entrant charging slightly less than the pair takes all of this business, even with the pair still on offer, and its revenue approaches twice its operating cost. The targeted organization with equal fees therefore does not survive when a pooled mandate can reach C^* . Reaching it moves the diversifier's weight from one third to one half and each middle weight from one third to one quarter, a large departure from equal weights. Targeted pairs with unequal fees, and the full equilibrium with flexible specialist mandates, remain open.

The two parts share one economic message. Specialists sell the part of the correction that broad funds are not allowed to deliver, and the permitted tilt, measured against dispersion, sets the size of that part. Standardized specialist weights, in turn, decide which specialist organization supplies it.

7 Conclusion

This paper studies an asset-management industry that sells diversification rather than skill. Investors agree about returns and risk but start from different home assets, and every fund holds assets that investors could trade themselves. When loading dispersion and operating

cost are sufficiently low, competition produces a market fund that every investor holds. One broad product spreads a single operating cost over the demand that all investors share, whereas supplying that demand through several narrower funds requires financing several operating costs. Once the market fund is in place, other funds can sell only a correction to its risk mix, away from the asset that carries the most common risk and toward the best diversifier. In the small-dispersion approximation, operating cost measured against the revenue that dispersion creates determines which corrections are sold: the market fund stands alone at high cost, a pooled specialist joins it at intermediate cost, and a targeted pair does so at low cost, with an overlap in which either specialist organization can operate. Cheaper fund operation or more dispersed loadings can therefore support more specialists, across markets or as an industry's costs fall or its loadings disperse.

The same forces shape what investors observe. Because every fund's fee revenue equals its operating cost, fees are inversely proportional to size, and the market fund, which sells the position every investor wants, is the largest and cheapest fund. Specialists earn positive alpha after fees against the market fund although nobody has skill: their alpha measures a covariance service, and each specialist's fee stays below it. Comparing fund types at the same primitives, with the pooled and targeted funds as alternative organizations, funds that depart further from the market portfolio toward the best diversifier are smaller and earn more alpha, without skill or diseconomies of scale. Investors whose home asset already supplies part of the correction hold it directly and buy less from specialists, so a concentrated direct holding can sit inside a diversified portfolio. At the loading and dispersion suggested by U.S. stock returns, every equilibrium computed in the original economy belongs to the same three organizations, with the market fund largest and cheapest and specialists keeping positive alpha after fees. The model thus offers a baseline for what covariance, access costs and free entry deliver on their own, before skill, information, trust or beliefs are added.

References

- Allen, F. and Gale, D. (1994). *Financial Innovation and Risk Sharing*. MIT Press, Cambridge, MA.
- Asness, C., Frazzini, A., Gormsen, N. J., and Pedersen, L. H. (2020). Betting against correlation: Testing theories of the low-risk effect. *Journal of Financial Economics*, 135(3):629–652.
- Bacher, A. (2025). The gender investment gap over the life cycle. *Review of Financial Studies*, 38(11):3205–3244.
- Baumol, W. J., Panzar, J. C., and Willig, R. D. (1982). *Contestable Markets and the Theory of Industry Structure*. Harcourt Brace Jovanovich, New York.
- Benartzi, S. and Thaler, R. H. (2007). Heuristics and biases in retirement savings behavior. *Journal of Economic Perspectives*, 21(3):81–104.
- Ben-David, I., Franzoni, F., Kim, B., and Moussawi, R. (2023). Competition for attention in the ETF space. *Review of Financial Studies*, 36(3):987–1042.
- Berk, J. B. and Green, R. C. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy*, 112(6):1269–1295.
- Berk, J. B. and van Binsbergen, J. H. (2015). Measuring skill in the mutual fund industry. *Journal of Financial Economics*, 118(1):1–20.
- Calvet, L. E., Campbell, J. Y., and Sodini, P. (2007). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy*, 115(5):707–747.
- Chen, J., Hong, H., Huang, M., and Kubik, J. D. (2004). Does fund size erode mutual fund performance? The role of liquidity and organization. *American Economic Review*, 94(5):1276–1302.
- Choukhmane, T. and de Silva, T. (2026). What drives investors’ portfolio choices? Separating risk preferences from frictions. *Journal of Finance*, 81(1):5–48.
- Coeurdacier, N. and Rey, H. (2013). Home bias in open economy financial macroeconomics. *Journal of Economic Literature*, 51(1):63–115.
- Cremers, K. J. M. and Petajisto, A. (2009). How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies*, 22(9):3329–3365.
- Cremers, M., Petajisto, A., and Zitzewitz, E. (2013). Should benchmark indices have alpha? Revisiting performance evaluation. *Critical Finance Review*, 2(1):1–48.
- Cvitanić, J. and Hugonnier, J. (2022). Optimal fund menus. *Mathematical Finance*, 32(2):455–516.

- Dahlquist, M. and Ibert, M. (2024). Equity return expectations and portfolios: Evidence from large asset managers. *Review of Financial Studies*, 37(6):1887–1928.
- DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, 22(5):1915–1953.
- Easley, D., Michayluk, D., O’Hara, M., and Putniņš, T. J. (2021). The active world of passive investing. *Review of Finance*, 25(5):1433–1471.
- French, K. R. and Poterba, J. M. (1991). Investor diversification and international equity markets. *American Economic Review*, 81(2):222–226.
- Gálvez, J. and Paz-Pardo, G. (2026). Richer earnings dynamics, consumption and portfolio choice over the life cycle. *Journal of Financial Economics*, 176:104206.
- Gârleanu, N. and Pedersen, L. H. (2018). Efficiently inefficient markets for assets and asset management. *Journal of Finance*, 73(4):1663–1712.
- Gârleanu, N. and Pedersen, L. H. (2022). Active and passive investing: Understanding Samuelson’s dictum. *Review of Asset Pricing Studies*, 12(2):389–446.
- Gârleanu, N., Panageas, S., and Yu, J. (2020). Impediments to financial trade: Theory and applications. *Review of Financial Studies*, 33(6):2697–2727.
- Gennaioli, N., Shleifer, A., and Vishny, R. (2014). Finance and the preservation of wealth. *Quarterly Journal of Economics*, 129(3):1221–1254.
- Gennaioli, N., Shleifer, A., and Vishny, R. (2015). Money doctors. *Journal of Finance*, 70(1):91–114.
- Gorton, G. B. and Pennacchi, G. G. (1993). Security baskets and index-linked securities. *Journal of Business*, 66(1):1–27.
- Hitzemann, S., Sokolinski, S., and Tai, M. (2022). Paying for beta: Leverage demand and asset management fees. *Journal of Financial Economics*, 145(1):105–128.
- Hortaçsu, A. and Syverson, C. (2004). Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 index funds. *Quarterly Journal of Economics*, 119(2):403–456.
- Investment Company Institute (2026). *2026 Investment Company Fact Book: A Review of Trends and Activities in the Investment Company Industry*. 66th edition. Investment Company Institute, Washington, DC. Figure 2.1, p. 13. <https://www.ici.org/system/files/2026-04/2026-factbook.pdf>.
- Khomyn, M., Putniņš, T. J., and Zoican, M. (2024). The value of ETF liquidity. *Review of Financial Studies*, 37(10):3092–3148.

- Khorana, A., Servaes, H., and Tufano, P. (2009). Mutual fund fees around the world. *Review of Financial Studies*, 22(3):1279–1310.
- Mamaysky, H. and Spiegel, M. (2001). A theory of mutual funds: Optimal fund objectives and industry organization. Working paper, Yale School of Management.
- Martin, I. (2017). What is the expected return on the market? *Quarterly Journal of Economics*, 132(1):367–433.
- Massa, M. (2003). How do family strategies affect fund performance? When performance-maximization is not the only game in town. *Journal of Financial Economics*, 67(2):249–304.
- Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance*, 42(3):483–510.
- Pástor, L. and Stambaugh, R. F. (2012). On the size of the active management industry. *Journal of Political Economy*, 120(4):740–781.
- Pástor, L., Stambaugh, R. F., and Taylor, L. A. (2015). Scale and skill in active management. *Journal of Financial Economics*, 116(1):23–45.
- Pollet, J. M. and Wilson, M. (2010). Average correlation and stock market returns. *Journal of Financial Economics*, 96(3):364–380.
- Reher, M. and Sokolinski, S. (2024). Robo advisors and access to wealth management. *Journal of Financial Economics*, 155:103829.
- Subrahmanyam, A. (1991). A theory of trading in stock index futures. *Review of Financial Studies*, 4(1):17–51.
- Van Nieuwerburgh, S. and Veldkamp, L. (2009). Information immobility and the home bias puzzle. *Journal of Finance*, 64(3):1187–1215.

Online Appendix

Diversification as a Service

Stanislav Sokolinski
September 2026

Appendix Contents

A	Primitives, demand, and pricing	A-3
A.1	Economic units and the normalized problem	A-3
A.2	Covariance, interiority, and direct outside options	A-3
A.3	Uniform local bounds and the common target	A-5
A.4	Investor demand at an arbitrary fund menu	A-6
A.5	Implementation ties and demand as a function of a fee	A-7
A.6	The competitive tests and exact revenue accounting	A-9
A.7	Exact demand at a given market-fund fee	A-10
A.8	Cheap market-fund access to the common target	A-12
A.9	Cost recovery and the symmetric fixed-cost benchmark	A-13
B	Market-fund selection and its consequences	A-16
B.1	Exact replication and the catalogue	A-16
B.2	Demand for the market fund	A-17
B.3	Profitable market-fund offers	A-20
B.4	Exact replacement and limiting market-fund demand	A-21
B.5	Entry and the market fund's own price reductions	A-24
B.6	Proof of Proposition 1 and its quantifiers	A-27
B.7	Quantitative consequences of selection	A-28
B.8	Existence when operating cost stays fixed	A-29
B.9	Uniformity and continuation domains	A-31
C	The normalized problem and its exact classification	A-33
C.1	The demand lemma and the appendix's units	A-33
C.2	The exact portfolio-adjustment problem	A-36
C.3	Exact demand, all quote intervals, and ties	A-40
C.4	Cutoffs and branch maps	A-42
C.5	The complete pair-price correspondence	A-43
C.6	The exact correspondence	A-45
C.7	Proof of Proposition C.7	A-45
C.8	Exact boundaries and exact nonclaims	A-57
D	Validity for positive c	A-58
D.1	What is transferred, and on which domain	A-58

D.2	Eliminating the market-fund coordinate	A-59
D.3	Demand limits, ties, and profitable deviation witnesses	A-60
D.4	The two quote tails and market-fund protection	A-61
D.5	Continuing break-even configurations and their supports	A-63
D.6	Protection against quotes close to the incumbent	A-66
D.7	Strict branch persistence under every price test	A-67
D.8	Closure of all primitive equilibria	A-68
D.9	Why the symmetric support condition cannot be omitted	A-70
D.10	Events and the limits of the strict continuation statements	A-71
E	Fees, performance, and investor portfolios	A-72
E.1	Fund sizes and fees	A-72
E.2	Performance and its benchmark	A-73
E.3	Portfolio implementation	A-76
E.4	Construction of the illustrations	A-79
F	Numerical protocol and results	A-80
F.1	What is computed	A-80
F.2	Parameters and sources	A-81
F.3	Results at the empirical anchor	A-82
F.4	Approach to the small-dispersion approximation	A-84
F.5	Replication files	A-85
G	A specialist operating-cost premium	A-85
G.1	Exact accounting	A-85
G.2	Market-fund selection, rates and scale	A-86
G.3	The reduced problem at the specialist cost	A-88
G.4	Positive dispersion: exclusion and continuation	A-88
G.5	Fees, performance and portfolios	A-90
H	Portfolio choice within mandates	A-91
H.1	Broad designs and the reduced design game	A-91
H.2	The value of tilt and conditional homogeneity	A-93
H.3	Selection of the broad design	A-94
H.4	Fees, sizes and alphas	A-96
H.5	Weight choice by specialists	A-97

A Primitives, demand, and pricing

A.1 Economic units and the normalized problem

This appendix supplies the derivations used in Sections 2 and 3. Appendix B proves Proposition 1 in the primitive nine-product economy, including endogenous access and the full deviation tests. The arguments concern the maintained model; claims about later specialist classification retain the additional assumptions stated with those results.

Write a tilde over a variable expressed in economic units. Gaussian returns and CARA preferences give, up to the constant certainty equivalent of initial wealth and risk-free investment,

$$\tilde{V}_i = \mu \mathbf{1}' \tilde{e}_i - \frac{\lambda}{2} \tilde{e}_i' \Sigma \tilde{e}_i - \sum_{T \in \mathcal{E}} \tilde{\delta}_T \tilde{q}_{iT} - \tilde{\tau} \sum_{j \in D} d(i, j), \quad \lambda = \gamma \sigma^2. \quad (\text{A.1})$$

Indeed, for Gaussian terminal wealth W , $-\gamma^{-1} \log \mathbb{E}[\exp(-\gamma W)] = \mathbb{E}[W] - \gamma \text{Var}(W)/2$. Fixed fees and access charges are deterministic expenditures. Unrestricted risk-free borrowing and lending make the risky holdings independent of initial wealth in this formulation.

Define

$$\begin{aligned} e_i &= \frac{\lambda}{\mu} \tilde{e}_i, & x_i &= \frac{\lambda}{\mu} \tilde{x}_i, & q_{iT} &= \frac{\lambda}{\mu} \tilde{q}_{iT}, & \delta_T &= \frac{\tilde{\delta}_T}{\mu}, \\ \tau &= \frac{\lambda}{\mu^2} \tilde{\tau}, & F &= \frac{\lambda}{\mu^2} \tilde{F}, & V_i &= \frac{\lambda}{\mu^2} \tilde{V}_i. \end{aligned} \quad (\text{A.2})$$

Substituting into (A.1) gives the objective in (7). The map is positive and invertible and leaves product weights, nonnegativity, access choices and optimality comparisons unchanged. With total investor mass one, $\widetilde{\text{AUM}}_T = (\mu/\lambda) \text{AUM}_T$, and hence

$$\tilde{\delta}_T \widetilde{\text{AUM}}_T - \tilde{F} = \frac{\mu^2}{\lambda} (\delta_T \text{AUM}_T - F). \quad (\text{A.3})$$

The signs of all entry and lower-quote profit comparisons are preserved. A proportional fee and an operating expenditure are connected by revenue, $\delta_T \text{AUM}_T = F$. They are distinct variables under either system of units.

A.2 Covariance, interiority, and direct outside options

Let $J = \{0, 1, 2, 3\}$, let ι_j be coordinate vector j , and set $d_j = 1 - a_j^2 > 0$. For a nonempty subset $S \subseteq J$, write

$$L_S = \sum_{j \in S} \frac{1}{d_j}, \quad Q_S = \sum_{j \in S} \frac{a_j}{d_j}, \quad R_S = \sum_{j \in S} \frac{a_j^2}{d_j}, \quad t_S = \frac{Q_S}{1 + R_S}. \quad (\text{A.4})$$

Here Q_S is a loading sum; it is distinct from the reference fund quantity $Q_0(m)$ in Section 3.

Lemma A.1 (Covariance and subset interiority). *The covariance matrix is positive definite. Under Assumption 1, every nonempty direct-asset subset has the strictly positive optimum*

$$e_j^S = \frac{1 - a_j t_S}{d_j} \quad (j \in S), \quad e_j^S = 0 \quad (j \notin S), \quad (\text{A.5})$$

and optimized gross value

$$W(S) = \frac{1}{2} \left(L_S - \frac{Q_S^2}{1 + R_S} \right). \quad (\text{A.6})$$

Proof. For every nonzero vector u , $u' \Sigma u = (a' u)^2 + \sum_j d_j u_j^2 > 0$. Each principal submatrix is therefore positive definite. Direct multiplication verifies the inverse

$$\Sigma_S^{-1} = D_S^{-1} - \frac{D_S^{-1} a_S a_S' D_S^{-1}}{1 + a_S' D_S^{-1} a_S}, \quad D_S = \text{diag}(d_j : j \in S). \quad (\text{A.7})$$

Consequently $\Sigma_S^{-1} \mathbf{1}$ has coordinates (A.5). Strict concavity makes this the unconstrained maximizer of $\mathbf{1}' e - \frac{1}{2} e' \Sigma_S e$.

It remains to verify its positivity for every subset, rather than only for J . Assumption 1 states $a_0 t_J < 1$, and $a_j \leq a_0$. Delete one index k from a set S for which $a_0 t_S < 1$. The identity

$$t_{S \setminus \{k\}} - t_S = - \frac{(a_k/d_k)(1 - a_k t_S)}{1 + R_{S \setminus \{k\}}} \leq 0 \quad (\text{A.8})$$

shows that the inequality is preserved under deletion. Iterating from J proves $a_j t_S < 1$ for every $j \in S$. Thus all coordinates in (A.5) are strictly positive, and the unconstrained solution also solves the long-only problem. At this solution $\Sigma_S e^S = \mathbf{1}$, so the gross value is $\frac{1}{2} \mathbf{1}' e^S$, yielding (A.6). $\square$

The exact value available to type i without any funds is therefore

$$V_i^{\text{direct}} = \max_{S \subseteq J: i \in S} \left\{ W(S) - \tau \sum_{j \in S \setminus \{i\}} d(i, j) \right\}. \quad (\text{A.9})$$

There are eight sets in this maximum. Each includes the free home asset; adding it to an accessible set entails no charge. The preceding lemma shows that every accessible asset is used when the investor holds only direct assets. A paid accessible asset with zero holdings could in any case be removed to save its charge. Home-only investing is the term $S = \{i\}$, with position one and value $1/2$. Holding no risky assets gives zero and is weakly inferior. When funds are present, the lemma does not assert that every accessible asset must still be held; fund fees and replication can make direct holdings bind at zero.

Positive definiteness also establishes the payoff-spanning statement in Section 2. If one underlying return were an exact linear combination of the others up to a risk-free payoff, a nonzero risky-asset combination would have zero variance, contradicting positive definiteness. Each fund return, on the other hand, is by construction a linear combination of these four returns. Funds change

implementation costs and accessible bundles, while leaving the span of the underlying risky payoffs unchanged.

A.3 Uniform local bounds and the common target

The selection argument fixes $m \in (0, 1)$ and $\tau > 0$, and varies c and F . The compact version fixes m in a compact subset $M \subset (0, 1)$. All constants below can then be chosen uniformly over M . The access comparisons require only a common positive lower bound on non-home charges.

Lemma A.2 (Uniform curvature and target expansion). *For all sufficiently small $c > 0$, the primitive loading constraints and Assumption 1 hold, and there are constants $\lambda_*, \lambda^*, C > 0$ such that $\lambda_* I \preceq \Sigma(m, c) \preceq \lambda^* I$. The gross optimum $e^*(c) = \Sigma(m, c)^{-1} \mathbf{1}$ is interior, and*

$$G^*(c) - G_c(e) = \frac{1}{2}(e - e^*(c))' \Sigma(m, c)(e - e^*(c)), \quad G_c(e) = \mathbf{1}'e - \frac{1}{2}e' \Sigma(m, c)e. \quad (\text{A.10})$$

Writing $w = 1 + 3m^2$, $b = \frac{1}{4}\mathbf{1}$, and $u = (1, 0, -1, 0)'$, the target satisfies

$$e^*(c) = \frac{1}{w}\mathbf{1} - c \frac{2m}{w(1 - m^2)}u + O(c^2). \quad (\text{A.11})$$

The optimized fee-free market-fund portfolio has a gross-value loss bounded by Cc^2 relative to $e^*(c)$.

Proof. At $c = 0$, $\Sigma_0 = (1 - m^2)I + m^2\mathbf{1}\mathbf{1}'$, whose eigenvalues are $1 - m^2$ on the zero-sum subspace and w on the span of $\mathbf{1}$. These are uniformly positive and finite on M . Also $e^*(0) = \mathbf{1}/w$ is strictly positive, uniformly on M . Continuity, compactness, and the strict loading inequalities at $c = 0$ give the asserted common neighborhood, eigenvalue bounds, and interiority. Completing the square around $\Sigma e^* = \mathbf{1}$ proves (A.10).

Differentiate $\Sigma(c)e^*(c) = \mathbf{1}$ at zero. Directly from the off-diagonal covariances, $\Sigma'(0)\mathbf{1} = 2mu$. Since $\mathbf{1}'u = 0$, $\Sigma_0^{-1}u = u/(1 - m^2)$. Thus $e^{*'}(0) = -2mu/[w(1 - m^2)]$, proving (A.11); smooth matrix inversion supplies the uniform Taylor remainder. The vector $e^*(0) = Q_0(m)b$, with $Q_0(m) = 4/w$, is feasible through the market fund and differs from $e^*(c)$ by $O(c)$. Equation (A.10) bounds its loss by Cc^2 , and optimizing fund quantity can only reduce that loss. $\square$

Proposition A.3 (The diversification calculation in Section 3.1.1). *At $c = 0$, the unique frictionless risky target is $e^*(0) = Q_0(m)b$. Its gross value is $2/w$, and its gain over home-only investing is $3(1 - m^2)/(2w) > 0$.*

Proof. The preceding lemma identifies the unique target. Alternatively, $b'\Sigma_0 b = w/4$, so a fund position q has value $q - wq^2/8$. The first-order condition gives $q = 4/w$ and value $2/w$. Home-only value is $1/2$; subtraction gives $(4 - w)/(2w) = 3(1 - m^2)/(2w)$. The risk-free asset finances the quantity choice, so neither Q_0 nor the coordinates of e^* are constrained to sum to one. $\square$

A.4 Investor demand at an arbitrary fund menu

Fix a finite offered menu $\mathcal{E}$, nonnegative fund fees δ , and an investor type i . For an access set $D \subseteq J \setminus \{i\}$, let A_{iD} have as its columns the direct-asset vectors ι_j , $j \in D \cup \{i\}$, and the offered fund vectors b_T , $T \in \mathcal{E}$. Write z for the corresponding nonnegative instrument holdings and f for the vector of per-unit charges: zero on direct assets and δ_T on fund T . Every column of A_{iD} is nonnegative and sums to one. Thus

$$e = A_{iD}z, \quad \mathbf{1}'e = \mathbf{1}'z.$$

Let $k_i(D) = \tau \sum_{j \in D} d(i, j)$. Conditional on D , the objective in (7) is

$$U_{iD}(z; \delta) = \mathbf{1}'z - f'z - \frac{1}{2}z'A'_{iD}\Sigma A_{iD}z - k_i(D).$$

There are eight possible access sets, including $D = \emptyset$. The complete investor problem maximizes these conditional values.

Lemma A.4 (Existence, bounds, and optimality conditions). *At every nonnegative fee vector, the conditional and complete investor problems have nonempty compact optimizer sets. If $\lambda_{\min}(\Sigma) \geq \lambda_* > 0$, every conditional optimizer satisfies*

$$\mathbf{1}'z = \mathbf{1}'e \leq \frac{8}{\lambda_*}.$$

For a fixed access set, z is optimal if and only if

$$\begin{aligned} z &\geq 0, & A'_{iD}\Sigma e &\geq \mathbf{1} - f, & e &= A_{iD}z, \\ z_j[(A'_{iD}\Sigma e)_j - (1 - f_j)] &= 0 & \text{for every instrument } j. \end{aligned} \tag{A.12}$$

Conditional on access, all optimizers have the same risky exposure and the same total fund-fee payment. Their individual instrument holdings need not be unique. Across tied optimal access sets, exposure itself can differ.

Proof. Put $s = \mathbf{1}'z = \mathbf{1}'e$. Nonnegative exposure and Cauchy–Schwarz give $\|e\|^2 \geq s^2/4$. Since fees are nonnegative, the objective before the fixed access payment is at most

$$s - \frac{\lambda_*}{8}s^2.$$

The feasible choice $z = 0$ has objective $-k_i(D)$. Consequently an optimal conditional portfolio cannot have $s > 8/\lambda_*$. Indeed, outside this bounded region its objective is strictly below that of zero holdings. Maximization can therefore be restricted to a compact set independent of fees; continuity gives an optimizer. The set of optimizers is closed and bounded. Maximizing over the finite access sets preserves existence and compactness. Lemma A.2 supplies a common λ_* in the local parameter region.

For fixed D , the objective is a differentiable concave quadratic on the nonnegative orthant. The coordinatewise marginal inequalities and complementarity conditions are precisely (A.12). They are necessary: a positive coordinate allows variations in either direction, and a zero coordinate allows increases. They are sufficient because concavity bounds every feasible improvement by the gradient inner product with the change in holdings, which is nonpositive under these conditions.

If two conditional optimizers had different exposures, their midpoint would be feasible for the same access set. Its linear return and fee terms would be the averages of theirs, whereas positive definiteness would give a strictly smaller quadratic risk cost than that average. This would strictly improve value, a contradiction. With exposure fixed, equality of optimized values also fixes the fee payment. Redundant instruments can nevertheless divide that payment and exposure in different ways. A midpoint argument does not apply across distinct access choices, which explains the last qualification. $\square$

Home-only investing is feasible at every menu and has value $1/2$. It follows that an optimal complete portfolio also attains at least $1/2$. The bound above did not require this stronger outside option; using zero holdings conditional on each access set makes the bound apply even to access sets that are never globally chosen.

A.5 Implementation ties and demand as a function of a fee

For a given menu and fee vector, let $\mathcal{Z}_i(\delta)$ denote the set of all complete optimal choices (D, z) of type i . A product that is absent from a particular implementation has holding zero in that implementation.

Definition A.5 (Common incumbent allocation and focal-upper demand). *At an incumbent menu, choose one array of complete investor optimizers, one for each type. All incumbent AUM and fee-revenue calculations use this same array.*

For a unilateral deviation by a focal product T , define its type- i demand by

$$q_{iT}^+(\delta) = \max\{q_{iT} : (D, z) \in \mathcal{Z}_i(\delta)\}, \quad \text{AUM}_T^+(\delta) = \frac{1}{4} \sum_{i=0}^3 q_{iT}^+(\delta).$$

The maximum includes all tied access choices as well as all tied implementations within an access choice.

The maxima exist by Lemma A.4. Investors choose independently, so each type's maximizing implementation can be selected simultaneously for one focal product. The corresponding productwise maxima for several different products need not be attainable by one common incumbent allocation. Definition A.5 does not require them to be.

Proposition A.6 (Monotonicity and the finite demand decomposition). *Fix every other product's fee and vary a focal fund fee $\eta \geq 0$. If $\eta_1 < \eta_2$, every optimal focal holding at η_1 is at least every optimal focal holding at η_2 . Focal-upper individual and aggregate demand consist of finitely many*

affine pieces on the fee axis, with their values at regime boundaries determined by Definition A.5. Each piece is nonincreasing. Revenue is quadratic on each open affine piece.

At a fee where the global optimizer set fixes a unique focal holding, selected demand is continuous. A support change with this uniqueness property therefore produces a continuous kink. A tie allowing distinct focal holdings, whether from access or implementation, can instead produce a downward jump as the fee rises.

Proof. For two complete choices, separate the focal fee from all other parts of utility and write the latter as H . Optimality at the two fees gives

$$H_1 - \eta_1 q_1 \geq H_2 - \eta_1 q_2, \quad H_2 - \eta_2 q_2 \geq H_1 - \eta_2 q_1.$$

Adding yields $(\eta_2 - \eta_1)(q_1 - q_2) \geq 0$. This argument permits different access sets and arbitrary optimal choices at either fee.

To prove the finite-piece statement without assuming unique implementations, first fix an access set. Its optimal exposure is unique. For that exposure e , the implementation set

$$\mathcal{P}(e) = \{z \geq 0 : A_{iD}z = e\}$$

is a compact polytope: $\mathbf{1}'z = \mathbf{1}'e$ bounds every coordinate. Its fee-minimizing implementations form a face, and those among them maximizing the focal holding form a further face. Select a vertex of this final face, which is also a vertex of $\mathcal{P}(e)$. The columns associated with its strictly positive coordinates must be linearly independent. Otherwise a nonzero dependence supported on those coordinates would permit small feasible perturbations in both directions, contradicting extremality.

There are finitely many such independent supports H . On one of them, the conditional KKT equalities give

$$z_H = (A'_H \Sigma A_H)^{-1}(\mathbf{1} - f_H), \quad z_{-H} = 0. \quad (\text{A.13})$$

Positive definiteness of Σ and independence of the columns make this Gram matrix invertible. Positivity of these coordinates and the unused-instrument KKT inequalities impose linear inequalities in fees. On the set where those conditions hold, the optimized value is

$$\frac{1}{2}(\mathbf{1} - f_H)'(A'_H \Sigma A_H)^{-1}(\mathbf{1} - f_H) - k_i(D). \quad (\text{A.14})$$

The empty support, when feasible by the marginal inequalities, has value $-k_i(D)$ and is included separately.

Thus each access set has a finite list of affine holding candidates with quadratic values. At least one of these candidates attains focal-upper demand for that access set. Taking all eight access sets covers the complete problem, including access ties. Along the one fee axis, each support's KKT region is an interval, possibly empty or a single point. Between their finitely many endpoints, comparisons of candidate values involve finitely many quadratic polynomials. A nonidentical pair has finitely many equality points. Identically equal values leave a tie, resolved by maximizing the finitely many affine focal holdings; those comparisons add only finitely many further breakpoints. This proves the claimed finite decomposition, including any exceptional endpoint values.

For completeness, focal demand is zero at every $\eta \geq 1$. All covariance entries are nonnegative in the primitive model. Removing a positive position q in a fund with fee at least one weakly increases its net-return contribution and strictly reduces risk: the variance term of that position alone falls by $q^2 b'_T \Sigma b_T > 0$, while the covariance with the remaining nonnegative exposure is nonnegative. Hence such a position cannot be optimal. It is therefore enough to apply the finite decomposition on the bounded interval $[0, 1]$, followed by the zero-demand ray.

The common compact bound in Lemma A.4 also proves the continuity assertion. Along any sequence of fees approaching a given fee, a subsequence of optimizers has a limit; after passing to a constant access set, continuity of the objective makes the limit a complete optimizer at the limiting fee. If all such optimizers agree on the focal holding, every subsequential holding limit is that same number. Finally, the revealed-preference inequality fixes the direction of every possible jump. Averaging over the four types preserves the finite affine decomposition and monotonicity. $\square$

This proposition supplies the demand curve used in unilateral entry and lower-quote tests. It does not impose a fee-dependent selection rule on all incumbent products simultaneously. Within a fixed access choice the exposure is continuous even when redundant instruments allow their individual holdings to change discontinuously.

A.6 The competitive tests and exact revenue accounting

For an active menu $\mathcal{E}$, let $\text{AUM}_T^{\text{inc}}$ be computed from the common allocation in Definition A.5. An active product's lower-quote test is

$$\eta \text{AUM}_T^+(\eta, \delta_{-T}; \mathcal{E}) \leq F \quad \text{for every } 0 \leq \eta < \delta_T. \quad (\text{A.15})$$

This is exactly the cheaper-clone comparison. If the incumbent were retained alongside a strictly cheaper identical bundle, an investor holding the incumbent could replace it one for one and save fees without changing exposure or access. Every optimal investor therefore assigns zero to the more expensive copy. Removing it and changing the product's quote produces the same focal-demand problem.

For an omitted product U , the entry test is

$$\mathcal{R}_U = \sup_{\eta \geq 0} \eta \text{AUM}_U^+(\eta, \delta; \mathcal{E} \cup \{U\}) \leq F. \quad (\text{A.16})$$

These are weak inequalities: an entrant earning exactly F is permitted by the protection condition. The equilibrium restriction excludes a strictly profitable entry quote. It does not require an omitted product's best revenue to be strictly below F .

Focal-upper demand is upper semicontinuous by the same compactness argument used in Proposition A.6: any limit of its maximizing optimizers is optimal at the limiting fee, whose focal maximum can only be larger. Consequently focal revenue is upper semicontinuous on $[0, 1]$. Revenue is zero beyond that interval, so the supremum in (A.16) is attained. Keeping the supremum notation emphasizes that the comparison checks all quotes, rather than only stationary points of individual

revenue pieces. Discrete changes in access or implementation prevent a single local first-order condition from replacing either test.

Lemma A.7 (Common cost and the inverse size–fee identity). *Suppose active products have positive AUM and satisfy break-even with the common cost $F > 0$, using one common incumbent allocation. Then every active fee is positive and*

$$\delta_T \text{AUM}_T^{\text{inc}} = F, \quad \frac{\delta_T}{\delta_U} = \frac{\text{AUM}_U^{\text{inc}}}{\text{AUM}_T^{\text{inc}}} \quad (T, U \in \mathcal{E}).$$

Proof. Positive AUM and positive F imply $\delta_T = F / \text{AUM}_T^{\text{inc}} > 0$. Dividing the two break-even equations gives the ratio identity. Neither a dispersion limit nor a demand regularity condition enters the calculation. The entry and lower-quote tests additionally determine whether these break-even prices are competitively protected; the accounting identity alone does not establish protection. $\square$

A.7 Exact demand at a given market-fund fee

First hold the offered menu fixed at $\{B\}$, impose no paid access, and write $\eta \geq 0$ for the market fund’s quoted proportional fee. This is a demand calculation at a given price; it imposes no equation between that price and operating cost. For type i , write

$$v = v_B(c) = b' \Sigma(c) b, \quad s = s_i(c) = (\Sigma(c) b)_i, \quad b = \frac{1}{4} \mathbf{1}.$$

For the loading vector $(m + c, m, m - c, m)$, summing covariance rows gives

$$\begin{aligned} v_B &= v_0 - \frac{c^2}{8}, & v_0 &= \frac{1 + 3m^2}{4}, \\ s_0 &= v_0 + \frac{mc}{2} - \frac{c^2}{4}, & s_2 &= v_0 - \frac{mc}{2} - \frac{c^2}{4}, \\ s_1 &= s_3 = v_0. \end{aligned} \tag{A.17}$$

For example, $4s_0 = 1 + (m + c)(3m - c) = 1 + 3m^2 + 2mc - c^2$. The variance is the average of the four row covariances. The primitive loading restrictions give $0 < s_i < 1$. The two instruments b and ι_i are linearly independent, so positive definiteness gives $v_B - s_i^2 > 0$.

Lemma A.8 (Fund and home demand at every nonnegative quote). *Define*

$$\eta_i^H = 1 - \frac{v_B}{s_i}, \quad \eta_i^{\text{off}} = 1 - s_i.$$

The conditional fund-and-home problem has a unique optimizer. Its fund holding is

$$q_{iB} = \begin{cases} (1 - \eta)/v_B, & 0 \leq \eta \leq \eta_i^H, \quad \eta_i^H \geq 0, \\ (1 - s_i - \eta)/(v_B - s_i^2), & \max\{0, \eta_i^H\} \leq \eta \leq \eta_i^{\text{off}}, \\ 0, & \eta \geq \eta_i^{\text{off}}, \end{cases} \quad x_i = [1 - s_i q_{iB}]_+. \tag{A.18}$$

The formulas agree wherever boundary cases overlap. In particular, when $q_{iB} > 0$, the equivalent expression is

$$x_i = \left[\frac{v_B - s_i + s_i \eta}{v_B - s_i^2} \right]_+, \quad q_{iB} = \frac{1 - \eta - s_i x_i}{v_B}. \quad (\text{A.19})$$

On a branch with both positions positive,

$$\frac{\partial x_i}{\partial \eta} = \frac{s_i}{v_B - s_i^2} > 0, \quad \frac{\partial q_{iB}}{\partial \eta} = -\frac{1}{v_B - s_i^2} < 0.$$

On the fund-only branch, $\partial q_{iB}/\partial \eta = -1/v_B$. On the home-only branch, $q_{iB} = 0$ and $x_i = 1$.

Proof. The covariance matrix of fund and home payoffs is positive definite. The objective

$$(1 - \eta)q + x - \frac{1}{2}(v_B q^2 + 2s_i q x + x^2)$$

is therefore strictly concave in (q, x) , and its optimizer is unique. Lemma A.4 gives its KKT conditions:

$$1 - \eta - v_B q - s_i x \leq 0, \quad 1 - s_i q - x \leq 0,$$

with equality for each strictly positive holding.

If $x = 0$ and $q > 0$, the first equality gives $q = (1 - \eta)/v_B$. The home marginal inequality is equivalent to $\eta \leq 1 - v_B/s_i = \eta_i^H$. If both positions are positive, solving the two equalities gives

$$x = \frac{v_B - s_i + s_i \eta}{v_B - s_i^2}, \quad q = \frac{1 - s_i - \eta}{v_B - s_i^2}.$$

These holdings are nonnegative exactly between the two stated thresholds. Their order is strict because

$$\eta_i^{\text{off}} - \eta_i^H = \frac{v_B - s_i^2}{s_i} > 0.$$

Finally, $q = 0$ forces $x = 1$, and its fund marginal inequality is $\eta \geq 1 - s_i$. These cases cover every nonnegative quote, and direct substitution shows equality at the boundaries. Differentiating within each branch gives the stated fee responses. $\square$

The covariance coefficients also verify the two home-investment motives in Section 3.1.2. At a zero fund fee, $v_B - s_2 = mc/2 + c^2/8 > 0$, so the low-loading home asset is used. For the high-loading home asset, $v_B - s_0 = -mc/2 + c^2/8 < 0$, since $c \leq m$; a positive fee saving must first offset this covariance disadvantage. The exact threshold formula, rather than an interior approximation, handles either participation decision.

Lemma A.9 (Symmetric demand and its uniform local extension). *At $c = 0$, the conditional demand for every type is*

$$x_0(\eta) = \frac{\eta}{1 - v_0}, \quad q_0(\eta) = \frac{1 - \eta/(1 - v_0)}{v_0}, \quad 0 \leq \eta < 1 - v_0. \quad (\text{A.20})$$

Fix $0 < \bar{\eta} < 1 - v_0$. For all sufficiently small $c \geq 0$, uniformly over $\eta \in [0, \bar{\eta}]$,

$$x_i(c, \eta) = x_0(\eta) + O(c), \quad q_{iB}(c, \eta) = q_0(\eta) + O(c). \quad (\text{A.21})$$

Every fund holding has a common positive lower bound on this region. Aggregate conditional demand

$$D_B(c, \eta) = \frac{1}{4} \sum_{i=0}^3 q_{iB}(c, \eta)$$

satisfies $D_B(c, \eta) = q_0(\eta) + O(c) = Q_0(m) + O(c + \eta)$ and is uniformly Lipschitz in η .

Proof. At $c = 0$, $v_B = s_i = v_0$, so the home threshold is zero and the fund exit threshold is $1 - v_0$. Substitution in Lemma A.8 gives (A.20). At zero fee this agrees with $Q_0 = 1/v_0$ and zero home investment. The total risky position is $q_0 + x_0 = (1 - \eta)/v_0$, so a higher fee reduces total exposure as well as reallocating part of it to the free home asset.

On the chosen compact fee interval, q_0 has a strictly positive minimum. Equation (A.17) shows that the covariance coefficients differ from their symmetric values by $O(c)$, and the denominators in (A.19) remain bounded away from zero. The positive-part map satisfies $|[a]_+ - [b]_+| \leq |a - b|$. Consequently the candidate holdings in that equation obey (A.21), uniformly including $\eta = 0$. For sufficiently small c , the candidate fund holding is positive; it satisfies the KKT conditions and hence is the unique conditional optimizer. This establishes the expansion without assuming a smooth home-participation pattern.

The derivatives in Lemma A.8 are uniformly bounded on this region. Demand is continuous at its support thresholds, so these piecewise derivative bounds give a common Lipschitz constant on the whole fee interval. Averaging preserves this bound and the positive lower bound. Finally, $q_0(\eta) = Q_0 - \eta/[v_0(1 - v_0)]$, which gives the joint $O(c + \eta)$ expression. $\square$

A.8 Cheap market-fund access to the common target

The next comparison restores every access choice. It also applies when funds besides B are available. Its conclusion concerns all optimal investors, including ties.

Lemma A.10 (A cheap market fund excludes paid access). *Fix $m \in (0, 1)$ and $\tau > 0$. There are positive constants $C, \bar{c}, \bar{\eta}$ such that any menu containing B at a fee $0 \leq \eta \leq \bar{\eta}$, with all other fees nonnegative, has the following property for $0 \leq c \leq \bar{c}$. Every complete optimizer (D, e, q) of every type satisfies*

$$\frac{1}{2}(e - e^*(c))' \Sigma(c)(e - e^*(c)) + \sum_T \delta_T q_T + k_i(D) \leq C(c^2 + \eta). \quad (\text{A.22})$$

The ceilings can be chosen so that $C(\bar{c}^2 + \bar{\eta}) < \tau/4$. Then every optimizer uses no paid access. Its exposure also satisfies

$$\|e - e^*(c)\| = O(c + \sqrt{\eta}).$$

When the offered menu is only $\{B\}$, the conditional demand in Lemmas A.8 and A.9 is therefore the complete optimal demand throughout the resulting low-fee region.

Proof. Write $G_c(e) = \mathbf{1}'e - \frac{1}{2}e'\Sigma(c)e$ and $G^*(c) = G_c(e^*(c))$. Lemma A.2 gives an interior target and

$$G^*(c) - \frac{1}{2v_B(c)} = O(c^2).$$

For $\eta < 1$, fund-only investment at the quoted fee has its optimum at $q = (1 - \eta)/v_B$, with net value $(1 - \eta)^2/(2v_B)$. Its gap from the frictionless value is

$$\begin{aligned} G^*(c) - \frac{(1 - \eta)^2}{2v_B(c)} &= \left[G^*(c) - \frac{1}{2v_B(c)} \right] + \frac{2\eta - \eta^2}{2v_B(c)} \\ &\leq C(c^2 + \eta) \end{aligned} \tag{A.23}$$

for a uniform constant on a sufficiently small parameter neighborhood. The fund-only choice is feasible without paid access at every menu under consideration. Therefore every complete optimizer satisfies

$$G_c(e) - \sum_T \delta_T q_T - k_i(D) \geq \frac{(1 - \eta)^2}{2v_B(c)}.$$

Subtract from $G^*(c)$ and apply the exact quadratic gap identity in Lemma A.2. This gives (A.22).

Every term on the left is nonnegative. A nonempty paid-access set costs at least $\tau/4$, so the strict ceiling condition rules it out for every optimizer. The eigenvalue floor in Lemma A.2 also implies $\|e - e^*(c)\|^2 \leq (2C/\lambda_*)(c^2 + \eta)$, proving the exposure bound. When B is the only fund, no paid access leaves exactly the strictly concave fund-and-home problem already solved. $\square$

All constants in the local demand and access lemmas can be chosen uniformly when m ranges over a compact subset of $(0, 1)$ and every non-home access charge has a common positive lower bound. Neither lemma imposes a relationship between the quoted fee and operating cost.

A.9 Cost recovery and the symmetric fixed-cost benchmark

A conditional demand schedule does not determine a fee until an operating cost is specified. The following result connects the given-price calculation to break-even and the lower-quote test. It holds the offered menu fixed at $\{B\}$; protection against omitted products is a separate condition.

Proposition A.11 (The locally protected break-even quote). *Fix $m \in (0, 1)$ and $\tau > 0$. There are positive constants $\bar{c}, \bar{F}, \bar{\eta}$ such that for $0 \leq c < \bar{c}$ and $0 < F < \bar{F}$, the market-fund-only menu has a unique break-even quote $\delta_B(c, F) \in (0, \bar{\eta})$. No investor pays for access at this quote or at any lower nonnegative quote, and every type holds the fund positively. Every strictly lower quote earns strictly less than F . Moreover,*

$$\delta_B(c, F) = \frac{F}{Q_0(m)} + O(F(c + F)), \quad D_B(c, \delta_B) = Q_0(m) + O(c + F). \tag{A.24}$$

The uniqueness assertion is within the displayed low-fee interval. It does not assert existence of a product equilibrium without the omitted-product entry checks.

Proof. Choose common fee and dispersion intervals on which Lemmas A.9 and A.10 apply. On this region, the complete demand $D_B(c, \eta)$ is continuous, bounded below by some $a > 0$, and Lipschitz in η with a common constant L . Reducing the fee ceiling if necessary, arrange $L\bar{\eta} < a/2$. For $0 \leq \eta_1 < \eta_2 \leq \bar{\eta}$, revenue satisfies

$$\begin{aligned}\eta_2 D_B(c, \eta_2) - \eta_1 D_B(c, \eta_1) &= (\eta_2 - \eta_1) D_B(c, \eta_2) + \eta_1 [D_B(c, \eta_2) - D_B(c, \eta_1)] \\ &\geq (a - L\bar{\eta})(\eta_2 - \eta_1) > 0.\end{aligned}$$

Thus revenue is strictly increasing throughout this interval, even when home participation changes. It starts at zero and is at least $a\bar{\eta}$ at the upper endpoint. Choosing $\bar{F} < a\bar{\eta}$, continuity gives exactly one break-even quote in the interval. The same strict increase proves protection against every lower quote. Absence of paid access and positivity of fund holdings hold on the entire interval by construction.

Break-even and the demand floor give $\delta_B \leq F/a = O(F)$. Lemma A.9 then yields $D_B(c, \delta_B) = Q_0 + O(c + F)$. Inverting this quantity in $F = \delta_B D_B$ gives (A.24). $\square$

This calculation explains why sufficiently low operating cost is compatible with a low proportional fee. In economic units the equation is still fee times AUM equals operating expenditure, as (A.3) shows. The reference quantity Q_0 gives the leading conversion from cost to fee; replacing actual AUM by Q_0 is an approximation, not an exact break-even identity.

At zero dispersion the conditional revenue equation can be solved without approximation. Write $v = (1 + 3m^2)/4$. From (A.20), revenue on $0 \leq \eta \leq 1 - v$ is

$$R_0(\eta) = \eta q_0(\eta) = \frac{\eta(1 - v - \eta)}{v(1 - v)}. \quad (\text{A.25})$$

This parabola is maximized at $\eta = (1 - v)/2$, with maximum $(1 - v)/(4v)$.

Proposition A.12 (Exact symmetric low-fee quantities). *For $0 < F < (1 - v)/(4v)$, the smaller root of the conditional symmetric break-even equation is*

$$\delta_0(F) = \frac{(1 - v) - \sqrt{(1 - v)^2 - 4Fv(1 - v)}}{2}. \quad (\text{A.26})$$

Its fund and home quantities are

$$x_0^{\text{eq}}(F) = \frac{\delta_0(F)}{1 - v}, \quad q_0^{\text{eq}}(F) = \frac{1 - \delta_0(F)/(1 - v)}{v}. \quad (\text{A.27})$$

At this root, optimized net value falls below the symmetric frictionless value by exactly

$$F + \frac{\delta_0(F)^2}{2v(1 - v)}. \quad (\text{A.28})$$

For F sufficiently small as in Proposition A.11, these are the no-paid-access quantities at the protected low-fee quote. The discriminant condition alone is only a conditional break-even feasibility bound.

Proof. Setting (A.25) equal to F gives $\eta^2 - (1 - v)\eta + Fv(1 - v) = 0$. The stated bound makes the discriminant positive. The smaller root is (A.26), lying in $(0, (1 - v)/2)$; the other root lies in $((1 - v)/2, 1 - v)$. Substitution into (A.20) gives the positions.

On the fund-and-home branch, the first-order equalities imply that optimized net value is half the net-return term:

$$V_0^{BH}(\eta) = \frac{1}{2}[(1 - \eta)q_0(\eta) + x_0(\eta)] = \frac{1}{2v} - \frac{\eta}{v} + \frac{\eta^2}{2v(1 - v)}.$$

Since the frictionless value is $1/(2v)$, the loss is $\eta/v - \eta^2/[2v(1 - v)]$. Break-even says $F = \eta/v - \eta^2/[v(1 - v)]$, giving (A.28).

For sufficiently small F , the smaller root lies within the no-access fee interval of Proposition A.11; it must be that proposition's unique break-even quote at $c = 0$. Alternatively, making the loss in (A.28) strictly less than $\tau/4$ directly rules out every paid-access choice at this benchmark. The local result verifies the entire lower-quote interval as well, which is what the protection test requires. $\square$

On the conditional no-access curve, the larger root has lower quotes with revenue greater than F . This curve alone is not a substitute for checking access at those quotes. The local protection result above uses an interval on which the complete demand problem has already been shown to choose no paid access.

For each fixed positive F sufficiently below the local cost ceiling, the conditional quote in Proposition A.11 converges to $\delta_0(F)$ as $c \rightarrow 0$. To see this, conditional demand and revenue are jointly continuous in (c, η) , including home support boundaries. Any subsequential limit of the low quotes solves the zero-dispersion break-even equation within their common strictly increasing interval. Its unique solution is $\delta_0(F)$. Continuity of the demand formulas then gives the limiting holdings in (A.27). Proposition B.10 supplies the additional omitted-product protection and the fixed-cost market-fund-only equilibrium conclusion. Conditional cost recovery alone does not supply that conclusion throughout the entire weak-cost rectangle.

Expanding in F , separately from taking a limit in dispersion, gives

$$\delta_0(F) = vF + O(F^2), \quad x_0^{\text{eq}}(F) = \frac{vF}{1 - v} + O(F^2), \quad q_0^{\text{eq}}(F) = \frac{1}{v} - \frac{F}{1 - v} + O(F^2). \quad (\text{A.29})$$

These follow by expanding the square root in (A.26) and substituting into the position formulas. In particular, the positive fee and the home position persist when c approaches zero at fixed $F > 0$. The terms vF and $vF/(1 - v)$ are first-order approximations in operating cost, not exact fixed-cost limits.

B Market-fund selection and its consequences

We prove Proposition 1 for the original finite catalogue, with investor access chosen after each quote. Fix $m \in (0, 1)$ and $\tau > 0$. No relation between F and c^2 is imposed. Write $A_{T,n} = \text{AUM}_{T,n}$ along a sequence of economies; a superscript zero denotes a limiting physical holding or AUM. When quantities are divided by $Q_0(m)$, we say so explicitly. Entry and lower-quote comparisons use the focal demand defined in Appendix A. Appendix B.2 first establishes Lemma 2 at given quotes; the equilibrium selection argument then follows.

B.1 Exact replication and the catalogue

Write $b = \frac{1}{4}\mathbf{1} = b_B$ and $Q_0 = 4/(1 + 3m^2)$, so the common frictionless exposure at $c = 0$ is $Q_0 b$. Let ι_i denote the unit vector for type i 's home asset, and define

$$d_T = b_T - b \quad (T \neq B), \quad H_i = \iota_i - b.$$

For a nonnegative vector $z = ((z_T)_{T \neq B}, z_H)$, set

$$D_i z = \sum_{T \neq B} z_T d_T + z_H H_i, \quad \ell(z) = \sum_{T \neq B} z_T + z_H, \quad K_i = \{z \geq 0 : D_i z = 0\}.$$

An unavailable product has coordinate zero. These definitions are homogeneous and apply either to physical holdings or to holdings divided by Q_0 .

Every instrument has weights summing to one. Consequently,

$$\sum_{T \neq B} z_T b_T + z_H \iota_i = \ell(z) b + D_i z. \tag{B.1}$$

In particular, a vector in K_i implements exactly $\ell(z)$ units of market exposure. If an investor's holdings contain such a vector coordinatewise, replacing it with $\ell(z)$ units of B preserves her entire risky exposure. This identity holds at every c ; it does not approximate the portfolio or discard any unreplaced holding.

Lemma B.1 (Catalogue replications).

1. *A single non-market product and a free home asset can reproduce positive market exposure only in the four cases*

$$(T, i) \in \{(123, 0), (230, 1), (301, 2), (012, 3)\}.$$

In each case $3b_T + \iota_i = 4b$, and the product and home quantities must be in the ratio 3 : 1.

2. *If two non-market products, together with each type's respective free home asset, can reproduce positive market exposure for all four types, their unordered pair is either $\{01, 23\}$ or $\{12, 30\}$. Every such replication uses zero home investment and equal quantities of the two products.*

Proof. For the first claim, suppose $qb_T + h\iota_i = tb$, where $q, h \geq 0$ and $t > 0$. Its right side has four positive coordinates. Since a non-market product contains either two or three assets, its support

and the home asset can cover all four coordinates only when T is the triplet excluding i . Comparing a coordinate in T with the home coordinate gives $q/3 = t/4 = h$. The four triplets in the catalogue give exactly the listed cases.

For the second claim, first suppose there are nonnegative u, v , not both zero, with $ud_T + vd_U = 0$. Both coefficients are positive because neither direction is zero, so the two directions are oppositely collinear. Each d_T is positive on T and negative outside T . Opposite collinearity therefore requires disjoint supports whose union contains all four assets. The catalogue's support sizes force both products to be pairs. Among the contiguous pairs, the only possibilities are $\{01, 23\}$ and $\{12, 30\}$; their directions are exact negatives.

If no such product-only relation exists, every type's market replication must use a positive home position. The relation $q_T d_T + q_U d_U + h H_i = 0$ then places $-H_i$ in the span of d_T, d_U . Universal replication would place all four home directions in a space of dimension at most two. This is impossible: the first three coordinates of H_0, H_1, H_2 form $I_3 - \frac{1}{4} \mathbf{1}_3 \mathbf{1}'_3$, whose determinant is $1/4$.

It remains to identify quantities for the two complementary pairs. Their common direction has two positive and two negative coordinates. Each H_i has one positive and three negative coordinates, so it is not collinear with that direction. A zero residual therefore requires zero home investment. The two opposite product directions then require equal product quantities. $\square$

Lemma B.2 (Removing an exact replication component). *Let $D_i z = \sum_{T \neq B} q_T (b_T - b) + h(\iota_i - b)$ and $\ell(z) = \sum_{T \neq B} q_T + h$, for $z = (q, h) \geq 0$. There is an $\epsilon_* > 0$, common across investor types, such that $\|D_i z\| \leq \epsilon_* \ell(z)$ with $\ell(z) > 0$ implies the existence of $k \geq 0$ satisfying*

$$D_i k = 0, \quad k \leq z, \quad \ell(k) \geq \frac{1}{2} \ell(z). \quad (\text{B.2})$$

Every nonzero such kernel vector satisfies $\sum_{T \neq B} k_T \geq 3\ell(k)/4$.

Proof. If the first assertion failed for some type, normalize a violating sequence to $\ell(z_n) = 1$, with $D_i z_n \rightarrow 0$. Compactness gives $z_n \rightarrow z^* \geq 0$, $D_i z^* = 0$, and $\ell(z^*) = 1$. On each positive coordinate of z^* , eventually $z_{n,j} \geq z_j^*/2$; on a zero coordinate the same inequality is automatic. Thus $k = z^*/2$ satisfies (B.2), a contradiction. Finiteness of the type set yields a common ϵ_* . For the paid-mass bound, the home coordinate of $\sum_T k_T b_T + k_h \iota_i = \ell(k)b$ gives $k_h \leq \ell(k)/4$, because all fund weights are nonnegative. The other coordinates of k therefore sum to at least $3/4$ of its total mass. $\square$

Lemma B.3 (One posted price for four types). *Let $\pi_1, \dots, \pi_4$ be finite nonnegative unit costs, sorted as $\pi_{(1)} \geq \dots \geq \pi_{(4)}$. Then*

$$\max_{1 \leq j \leq 4} \frac{j \pi_{(j)}}{4} \geq \frac{\frac{1}{4} \sum_{i=1}^4 \pi_i}{H_4}, \quad H_4 = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} = \frac{25}{12}.$$

Proof. Write $R = \max_j j \pi_{(j)}/4$. For each j , $\pi_{(j)} \leq 4R/j$. Summing these four inequalities and dividing by four gives $\frac{1}{4} \sum_i \pi_i \leq R H_4$. $\square$

B.2 Demand for the market fund

This subsection proves Lemma 2. Fix $m \in (0, 1)$, $\tau > 0$, a finite menu $\mathcal{E}$ containing B at the quote $\delta_B \geq 0$, and nonnegative fees δ_T for the other offered products. Use the notation of Appendix B.1: $b = b_B$, D_i , ℓ , and the exact replication map (B.1). For type i , let

$$\rho_i = \min \left\{ \sum_{T \in \mathcal{E} \setminus \{B\}} \delta_T y_T : y \geq 0, D_i y = 0, \ell(y) = 1, y_T = 0 \text{ for } T \notin \mathcal{E} \right\}, \quad (\text{B.3})$$

with $\rho_i = +\infty$ when no such y exists. The vector y has one coordinate for each non-market product and one for type i 's home asset, which carries no fee. A feasible y is a unit replication of the market portfolio available to type i , and ρ_i is the lowest fee per unit of market exposure among these replications, in the fee units of the body. The feasible set is a closed subset of the unit simplex, so the minimum is attained whenever the set is nonempty.

Lemma B.4 (Demand for the market fund).

- (i) In every primitive economy and at all nonnegative fees, if $\rho_i < \delta_B$, then no complete optimal choice of type i holds B .
- (ii) There are positive constants $\bar{c}$, $\bar{\eta}$ and K such that, whenever $0 \leq c \leq \bar{c}$ and $0 \leq \delta_B \leq \bar{\eta}$, every complete optimal choice of every type uses no paid access, and its exposure satisfies $\|e_i - e^*(c)\| \leq K(c + \sqrt{\delta_B})$.
- (iii) Under the conditions of part (ii), if $\rho_i > \delta_B$, then every complete optimal choice of type i satisfies

$$\ell(z_i) \leq K(c + \delta_B), \quad |q_{iB} - Q_0(m)| \leq K(c + \delta_B),$$

where z_i lists her holdings of the other offered products and of her home asset.

The constants depend only on m and τ . They do not depend on the menu, the other fees, or the gap $\rho_i - \delta_B$.

Proof. Part (i). Suppose an optimal choice holds $q_{iB} > 0$, and let y attain ρ_i . By (B.1) and $D_i y = 0$, $\sum_{T \neq B} y_T b_T + y_H \iota_i = b$. Replacing the q_{iB} units of B by $q_{iB} y$ therefore leaves her risky exposure and her access set unchanged and lowers her fee payment by $q_{iB}(\delta_B - \rho_i) > 0$, a contradiction. The argument uses only the exact identity and nonnegative holdings; it needs no additional restriction on c .

Part (ii). This is Lemma A.10. Its ceilings on c and on the market-fund quote apply to every menu containing B and to all nonnegative fees of the other products.

Part (iii). Fix a complete optimal choice of type i . By part (ii) it uses no paid access, so its exposure is $e_i = q_{iB} b + \sum_{T \neq B} q_{iT} b_T + x_{ii} \iota_i$. Write $M = \mathbf{1}' e_i = q_{iB} + \ell(z_i)$ and $r = D_i z_i$. By (B.1), $e_i = M b + r$ and $\mathbf{1}' r = 0$. Part (ii) and $\mathbf{1}' e^*(c) = Q_0 + O(c^2)$, from (A.11), give $M \geq Q_0 - 2K(c + \sqrt{\delta_B}) - O(c^2)$. After reducing the ceilings if necessary, $M \geq 3Q_0/4$.

First, z_i contains no nonzero exact replication. If some k with $0 \neq k \leq z_i$ had $D_i k = 0$, then $k/\ell(k)$ would be feasible in (B.3), so the fees on k would be at least $\rho_i \ell(k) > \delta_B \ell(k)$. Replacing k by $\ell(k)$ units of B would preserve her exposure and lower her fees. If $\rho_i = +\infty$, no such k exists at all. Lemma B.2 therefore gives

$$\ell(z_i) \leq \epsilon_*^{-1} \|r\|. \quad (\text{B.4})$$

Second, compare the choice with M units of B and nothing else. This alternative is feasible because $M > 0$, needs no paid access and has the same expected return M . Optimality of the choice gives

$$\frac{1}{2}r'\Sigma r \leq -Mb'\Sigma r + \delta_B \ell(z_i) - \sum_{T \neq B} \delta_T q_{iT} \leq M|b'\Sigma r| + \delta_B \ell(z_i).$$

Because $\Sigma(0)b = v_0 \mathbf{1}$ and $\mathbf{1}'r = 0$, we have $b'\Sigma(0)r = 0$, and smooth dependence on c gives $|b'\Sigma r| \leq C_1 c \|r\|$. Lemma A.4 bounds $M \leq 8/\lambda_*$, and Lemma A.2 gives $r'\Sigma r \geq \lambda_* \|r\|^2$. Together with (B.4),

$$\|r\| \leq \frac{2}{\lambda_*} \left(\frac{8C_1}{\lambda_*} c + \frac{\delta_B}{\epsilon_*} \right),$$

so $\|r\|$ and $\ell(z_i)$ are at most a constant times $c + \delta_B$.

Third, recover the market-fund holding. After reducing the ceilings further if necessary, the second step gives $\ell(z_i) \leq Q_0/4$, so $q_{iB} = M - \ell(z_i) \geq Q_0/2 > 0$. The market fund's first-order condition then holds with equality: $1 - \delta_B = b'\Sigma e_i = Mb'\Sigma b + b'\Sigma r$. Since $b'\Sigma b = v_0 - c^2/8$ by (A.17) and $Q_0 = 1/v_0$, it gives $|M - Q_0| \leq K_2(c + \delta_B)$. Hence $|q_{iB} - Q_0| \leq |M - Q_0| + \ell(z_i) \leq K(c + \delta_B)$.

Every constant comes from Lemmas A.2, A.4 and A.10, from the smooth dependence of Σ on c , or from the geometric constant ϵ_* , which is common to the finitely many types. None depends on the menu, the other fees or $\rho_i - \delta_B$. $\square$

Proof of Lemma 2. Part (a) is Lemma B.4(i). For part (b), Lemma B.4(ii) excludes paid access and gives $\|e_i - e^*(c)\| \leq K(c + \sqrt{\delta_B})$, while $e^*(c) = Q_0 b_B + O(c)$ by Lemma A.2; this is the sense in which each position is close to $Q_0 b_B$. A market-fund fee below the fee per unit of every available replication means $\rho_i > \delta_B$, and Lemma B.4(iii) gives the market-fund holding and the bound on her other holdings. $\square$

The lemma describes demand at given quotes. It imposes no relation between a quote and operating cost and is not itself a step in the proof of Proposition 1. That proof works along sequences of equilibria. It uses Lemma A.10, which is part (ii), and the limiting cost of Lemma B.6, together with break-even, the catalogue's replication geometry, a single posted quote and the lower-quote test. At a tie $\rho_i = \delta_B$ neither strict comparison applies, and product demand follows Definition A.5. Holdings need not vary continuously across such a tie. When type i 's cheapest replication uses her home asset, moving her common position between B and that replication changes both her fund and her home holdings discretely.

The cost ρ_i is measured in the body's fee units and ranges over every offered product. It differs from the normalized limiting cost π_i of Lemma B.6, which scales prices by F_n/Q_0 and uses only products with positive limiting AUM. When B is the only fund, no replication exists, so part (iii) applies to every type; Lemmas A.8 and A.9 then give that demand exactly and its local expansion.

B.3 Profitable market-fund offers

Fix $m \in (0, 1)$ and $\tau > 0$. The next two lemmas consider sequences with $c_n \rightarrow 0$ and $F_n \rightarrow 0$, where $F_n > 0$, without restricting their relative rates. Write

$$\Sigma_n = \Sigma(m, c_n), \quad G_n(e) = \mathbf{1}'e - \frac{1}{2}e'\Sigma_n e, \quad e_n^* = \Sigma_n^{-1}\mathbf{1}, \quad Q_0 = \frac{4}{1 + 3m^2}.$$

In particular, $e_n^* \rightarrow Q_0 b$. For an incumbent equilibrium, $e_{i,n}$ denotes type i 's exposure and $V_{i,n}$ her optimized net value. We set the AUM of an omitted product equal to zero. The finiteness of the catalogue and the boundedness in Lemma A.4 permit passage to a constant active menu and convergent holdings whenever needed.

Lemma B.5 (Broad-entry forcing). *Consider a sequence of equilibria of the primitive economy with $c_n, F_n \rightarrow 0$. Suppose that B is omitted throughout the sequence, or that it is active and $\text{AUM}_{B,n} \rightarrow 0$. Then*

$$e_{i,n} \rightarrow Q_0 b \quad \text{for every type } i,$$

and no type pays for non-home access for all sufficiently large n . Consequently, every subsequential limit z_i^0 of non- B and home holdings satisfies

$$z_i^0 \in K_i, \quad \ell(z_i^0) = Q_0.$$

Proof. By Lemma A.2, covariance has a common positive curvature bound $\lambda_* > 0$ for all sufficiently large n , and

$$G_n(e_n^*) - G_n(e) = \frac{1}{2}(e - e_n^*)'\Sigma_n(e - e_n^*) \geq \frac{\lambda_*}{2}\|e - e_n^*\|^2. \quad (\text{B.5})$$

Here and below, all fund fees and access payments are nonnegative. If some type's exposure does not converge to $Q_0 b$, pass to a subsequence and a fixed type i with $\|e_{i,n} - e_n^*\| \geq \epsilon > 0$. Her incumbent net value then satisfies

$$V_{i,n} \leq G_n(e_n^*) - L, \quad L = \lambda_* \epsilon^2 / 2 > 0.$$

If instead paid access occurs infinitely often, pass to a fixed type paying for at least one non-home asset. The same inequality holds with $L = \tau/4$, because even unrestricted gross value cannot exceed $G_n(e_n^*)$. These two arguments cover any persistent failure of either claimed conclusion.

Choose a fixed quote $\eta_0 > 0$ small enough that a portfolio using only B , without paid access, has net value at least $G_n(e_n^*) - L/2$ for all sufficiently large n . Such a quote exists by Lemma A.10: the loss of the feasible $Q_0 b$ portfolio is bounded by $Cc_n^2 + \eta_0 Q_0$. The same value bound holds at every quote $0 \leq \eta \leq \eta_0$. Thus adding B , or lowering its incumbent quote to such an η , raises this type's attainable net value by at least $g = L/2 > 0$.

This value gain forces a positive quantity of B in every deviating optimizer. To prove the assertion, suppose there were a sequence of deviating optimizers with $q_{iB,n}^{\text{dev}} \rightarrow 0$. Remove its entire B position while retaining its access set and all other holdings. The resulting portfolio is feasible at the incumbent menu and quotes, whether B was omitted or its quote was higher. The gross value change is bounded in absolute value by $Cq_{iB,n}^{\text{dev}}$, because exposures are uniformly bounded and G_n has a bounded gradient on that bounded set. Removing the position also eliminates its nonnegative

fee payment. The resulting incumbent-feasible value is therefore at least

$$V_{i,n} + g - Cq_{iB,n}^{\text{dev}} > V_{i,n}$$

for large n , a contradiction. There is accordingly a constant $a > 0$ such that every relevant deviating optimizer has $q_{iB,n}^{\text{dev}} \geq a$ for all sufficiently large n . The population demand for the deviating fund is at least $a/4$.

If B is omitted, entry at η_0 earns at least $\eta_0 a/4 > F_n$ eventually, contradicting entry protection. If B is active, use the strictly lower quote

$$\eta_n = \min\{\eta_0, \delta_{B,n}/2\}.$$

When $\eta_n = \eta_0$, the same fixed positive revenue bound applies. On the other branch, incumbent break-even gives

$$\frac{\eta_n \text{AUM}_B^{\text{dev}}}{F_n} \geq \frac{(\delta_{B,n}/2)(a/4)}{F_n} = \frac{a}{8 \text{AUM}_{B,n}} \rightarrow \infty.$$

Both branches violate the global lower-own-quote test. Hence neither an exposure failure nor paid access can persist.

With paid access absent, the exact catalogue decomposition is

$$e_{i,n} = (q_{iB,n} + \ell(z_{i,n}))b + D_i z_{i,n}.$$

Every individual B holding vanishes when aggregate B AUM does, because holdings are nonnegative and each type has positive population mass. Taking limits along any convergent holdings subsequence gives $D_i z_i^0 = 0$ and $\ell(z_i^0) = Q_0$, as claimed. $\square$

B.4 Exact replacement and limiting market-fund demand

The next lemma compares the cost of implementations that supply the same risky exposure. Its normalization divides quantities by Q_0 and fees by F_n/Q_0 . These are units for the comparison; no relation between F_n and c_n is imposed.

Lemma B.6 (Limiting cost and strict demand lifting). *Consider an equilibrium sequence with $c_n, F_n \rightarrow 0$, a constant active menu, no paid access, and $e_{i,n} \rightarrow Q_0 b$ for every type. Pass to a subsequence along which all holdings converge. Denote physical quantity and AUM limits by q_{iT}^0 and A_T^0 , and define*

$$y_{iT} = q_{iT}^0/Q_0, \quad a_T = A_T^0/Q_0 = \frac{1}{4} \sum_i y_{iT}, \quad \mathcal{O} = \{T \neq B : A_T^0 > 0\}.$$

For $T \in \mathcal{O}$, set

$$\beta_T = \frac{1}{a_T} = \frac{Q_0}{A_T^0}.$$

If $A_B^0 > 0$, use the same definition of a_B and β_B . Let the minimum cost of a unit-mass alternative implementation be

$$\pi_i = \min \left\{ \sum_{T \in \mathcal{O}} \beta_T y_T : y \in K_i, \ell(y) = 1, y_T = 0 \text{ for } T \notin \mathcal{O} \right\}, \quad (\text{B.6})$$

where the home coordinate carries zero fee and $\pi_i = +\infty$ if the feasible set is empty. Then:

- (i) Whenever finite, π_i is positive and attained. The selected limiting implementation minimizes fee cost among available unit-mass implementations through $\mathcal{O}$, home, and—when $A_B^0 > 0$ — B . In particular, if $A_B^0 = 0$, its normalized cost is π_i . If $A_B^0 > 0$,

$$\begin{aligned} y_{iB} > 0 &\implies \beta_B \leq \pi_i, \\ y_{iB} < 1 &\implies \pi_i \leq \beta_B, \\ 0 < y_{iB} < 1 &\implies \pi_i = \beta_B. \end{aligned}$$

- (ii) Fix a type i and a number $p > 0$ with $p < \pi_i$. Offer B at

$$\eta_n = p F_n / Q_0,$$

leaving every other product's quote fixed. If B is active with $A_B^0 > 0$, require additionally $p < \beta_B$; if B is active with $A_B^0 = 0$, no additional restriction is needed. The offer is then an admissible entry or a strictly lower own quote for all sufficiently large n . Every deviating optimizer eventually chooses no paid access, and the type's B holding converges to Q_0 , uniformly over its deviating optimizers. Consequently its focal-upper demand also converges to Q_0 .

Proof. All non- B and home quantity limits are nonnegative. The exact exposure decomposition and its unit-mass normalization give

$$\hat{z}_i = ((q_{iT}^0)_{T \neq B}, x_{ii}^0) / Q_0 \in K_i, \quad y_{iB} + \ell(\hat{z}_i) = 1.$$

A zero AUM limit implies a zero holding limit for every type, so $\hat{z}_i$ has product support contained in $\mathcal{O}$. Break-even gives the price limits

$$\frac{Q_0 \delta_{T,n}}{F_n} \longrightarrow \beta_T \quad (T \in \mathcal{O}), \quad \frac{Q_0 \delta_{T,n}}{F_n} \longrightarrow \infty \quad (T \text{ active, } A_T^0 = 0). \quad (\text{B.7})$$

The corresponding finite limit holds for B when $A_B^0 > 0$.

The feasible set in (B.6) is a closed subset of the unit simplex. Its minimum is therefore attained whenever it is nonempty. Every β_T in its objective is positive. A feasible vector of zero cost would have only a home coordinate, but $H_i = \iota_i - b \neq 0$, so a pure-home vector cannot belong to K_i with unit mass. The attained minimum is consequently positive.

To prove limiting cost optimality, suppose another feasible unit-mass implementation has a strictly smaller normalized fee cost than the selected limit $(y_{iB}, \hat{z}_i)$. When $A_B^0 = 0$, both implementations are restricted to a zero B coordinate. Multiply both vectors by Q_0 : each now implements exactly

Q_0b . Fix $\rho \in (0, 1)$. On every positive coordinate of the selected physical limit, finite incumbent holdings eventually exceed $(1 - \rho)$ times that limit. Zero coordinates require no subtraction. Remove this fraction of the selected limit from the finite portfolio and add the same fraction of the proposed alternative.

This replacement is feasible and preserves risky exposure exactly. It changes no access decision. Its fee change divided by F_n converges, by (B.7), to $(1 - \rho)$ times the strictly negative difference in normalized costs. Products with zero AUM limits have zero coordinates in both replacement vectors; their finite holdings and fee payments remain untouched. Thus the replacement strictly improves the finite investor objective for large n , a contradiction. This proves cost optimality. The stated comparisons between π_i and β_B follow by replacing any positive part of a more expensive implementation with the cheaper one.

We next prove (ii). When B has a positive AUM limit, (B.7) and $p < \beta_B$ make η_n strictly below its incumbent quote eventually. When its AUM limit is zero, the normalized incumbent quote diverges, so every fixed p eventually gives a strictly lower quote. Entry is admissible when B is omitted.

At the offered quote, holding Q_0 units of B is feasible for every type and loses at most $Cc_n^2 + pF_n$ relative to the unrestricted gross optimum. Every portfolio using paid access has net value at most $G_n(e_n^*) - \tau/4$. Hence all deviating optimizers choose no paid access for large n . Nonnegative fees and the curvature bound further imply, for every such optimizer,

$$\|e_{i,n}^{\text{dev}} - e_n^*\| = O(c_n + \sqrt{F_n}), \quad e_{i,n}^{\text{dev}} \longrightarrow Q_0b.$$

Only convergence is needed here; the bound does not impose a rate on F_n .

Choose any sequence of deviating optimizers and pass to a subsequence along which their holdings converge. Let $z_i^{\text{dev},0}$ be the physical non- B and home limit, and let $q_{iB}^{\text{dev},0}$ be the B limit. Then

$$D_i z_i^{\text{dev},0} = 0, \quad q_{iB}^{\text{dev},0} + \ell(z_i^{\text{dev},0}) = Q_0. \quad (\text{B.8})$$

The first identity is an exact catalogue identity. Because the product weights and home vectors are independent of c , it remains an exact replication in every finite economy along the sequence.

Suppose a coordinate T of $z_i^{\text{dev},0}$ is positive although its incumbent AUM limit is zero. Again fix $\rho \in (0, 1)$. The finite deviating holdings eventually contain $(1 - \rho)z_i^{\text{dev},0}$ coordinatewise. Remove that component and replace it by $(1 - \rho)\ell(z_i^{\text{dev},0})$ additional units of B . Equation (B.8) makes the replacement exposure-preserving. Its fee cost is

$$\frac{pF_n}{Q_0}(1 - \rho)\ell(z_i^{\text{dev},0}) = O(F_n).$$

The saving on coordinate T alone, divided by F_n , is

$$(1 - \rho)\frac{\delta_{T,n}}{F_n}z_{iT}^{\text{dev},0} \longrightarrow \infty.$$

Every other removed fee payment is nonnegative. This strictly cheaper implementation contradicts optimality. Therefore $z_i^{\text{dev},0}$ has product support contained in $\mathcal{O}$. This argument leaves all remaining finite holdings unchanged; in particular, it requires no $O(F_n)$ bound on the total deviating fee bill.

Set $L = \ell(z_i^{\text{dev},0})$. If $L > 0$, the vector $z_i^{\text{dev},0}/L$ is feasible in (B.6). Thus π_i is finite and

$$\sum_{T \in \mathcal{O}} \beta_T z_{iT}^{\text{dev},0} \geq \pi_i L.$$

Remove and replace $(1 - \rho)z_i^{\text{dev},0}$ as before. The removed fee cost minus the replacement cost, divided by F_n , converges to a quantity at least

$$\frac{(1 - \rho)L}{Q_0}(\pi_i - p) > 0.$$

This is again a strictly improving exact replacement. Hence $L = 0$ and (B.8) gives $q_{iB}^{\text{dev},0} = Q_0$. The case $\pi_i = +\infty$ has the same conclusion, since any $L > 0$ would itself establish a feasible alternative implementation and contradict $\pi_i = +\infty$.

Every sequence of deviating optimizers has bounded holdings, and every convergent subsequence has the same B -quantity limit Q_0 . If convergence were not uniform over optimizers, one could select a sequence staying a fixed distance from Q_0 and extract a contradictory subsequential limit. This proves the claimed uniform convergence and, in particular, the focal-upper demand conclusion. $\square$

B.5 Entry and the market fund's own price reductions

The following lemmas use an arbitrary sequence of equilibria with $c_n \rightarrow 0$, $F_n \rightarrow 0$, and $F_n > 0$, at the fixed m, τ of the appendix. Pass to subsequences with a constant active menu and convergent holdings whenever needed. Use the normalized limiting quantities and prices of Lemma B.6:

$$y_{iT} = \frac{q_{iT}^0}{Q_0}, \quad a_T = \frac{A_T^0}{Q_0}, \quad \beta_T = \frac{1}{a_T} = \frac{Q_0}{A_T^0}.$$

For a product with positive limiting AUM, break-even gives $\delta_{T,n}Q_0/F_n \rightarrow \beta_T$. Let $\mathcal{O}$ contain the non-market products with positive limiting AUM, and let π_i be type i 's minimum cost of one unit of exact market replication using $\mathcal{O}$ and her free home asset, at prices β_T . A quote with coefficient p means the physical fee $\eta_n = pF_n/Q_0$.

Lemma B.7 (The market fund cannot be absent or vanish). *Along any such equilibrium sequence, B is active eventually and $\liminf_n A_{B,n} > 0$.*

Proof. If either conclusion failed, a subsequence would have B absent throughout or $A_{B,n} \rightarrow 0$. Lemma B.5 then gives, for every type, exposure converging to $Q_0 b$ and no paid access eventually. All limiting market exposure must be implemented using $\mathcal{O}$ and home investment. In particular, every π_i is finite. If $N = |\mathcal{O}| = 0$, this would require a home asset alone to reproduce b , which is impossible.

Limiting cost optimality from Lemma B.6 and aggregate break-even give the identity

$$\frac{1}{4} \sum_{i=0}^3 \pi_i = \sum_{T \in \mathcal{O}} \beta_T \left(\frac{1}{4} \sum_{i=0}^3 y_{iT} \right) = \sum_{T \in \mathcal{O}} \beta_T a_T = N. \quad (\text{B.9})$$

This identity counts only products with positive limiting AUM. Products whose AUM vanishes can still collect their operating cost at finite n . They are not included in either side, and the exact replacements in Lemma B.6 leave their unreplaced holdings intact.

There are three remaining cases.

If $N = 1$, Lemma B.1 says the single product can reproduce market exposure with home investment for at most one type. It cannot serve all four types, a contradiction.

If $N = 2$, that lemma leaves only the complementary pairs $\{01, 23\}$ and $\{12, 30\}$. Every type uses half of each product and no home investment in normalized limiting units. Thus each product has $a_T = 1/2$, $\beta_T = 2$, and every type has $\pi_i = 2$. Quote $p = 3/2$. Strict lifting gives market-fund demand converging to Q_0 from every type, so

$$\frac{\eta_n \hat{A}_{B,n}(\eta_n)}{F_n} = p \frac{\hat{A}_{B,n}(\eta_n)}{Q_0} \longrightarrow \frac{3}{2} > 1.$$

Here $\hat{A}_{B,n}$ denotes aggregate focal demand following the quote. This profitable deviation contradicts protection.

If $N \geq 3$, sort the four costs and choose j attaining the maximum in Lemma B.3. By (B.9),

$$\frac{j\pi_{(j)}}{4} \geq \frac{N}{H_4}.$$

Choose the strictly lower coefficient $p = \frac{3}{4}\pi_{(j)} > 0$. The j types whose costs are at least $\pi_{(j)}$ satisfy $p < \pi_i$. Lemma B.6 therefore gives

$$\liminf_n \frac{\eta_n \hat{A}_{B,n}(\eta_n)}{F_n} \geq \frac{pj}{4} \geq \frac{3N}{4H_4} = \frac{9N}{25} \geq \frac{27}{25} > 1.$$

The strict margin does not rely on demand at a tied threshold.

When B is omitted, these are permitted entry quotes. When it is active with vanishing AUM, break-even gives $\delta_{B,n}Q_0/F_n = Q_0/A_{B,n} \rightarrow \infty$. Each fixed coefficient used above therefore also gives a strictly lower own quote for all sufficiently large n . Every possible case contradicts an equilibrium test. $\square$

Lemma B.8 (The market fund serves every type in the limit). *Along every equilibrium sequence with $c_n, F_n \rightarrow 0$ and $F_n > 0$,*

$$q_{iB,n} \longrightarrow Q_0 \quad \text{for all } i, \quad A_{T,n} \longrightarrow 0 \quad (T \neq B), \quad x_{ii,n} \longrightarrow 0.$$

No type pays for non-home access eventually.

Proof. Lemma B.7 permits passage to a subsequence with $A_B^0 > 0$. Break-even implies $\delta_{B,n} = O(F_n)$. The feasible fund-only portfolio at that quote has net value within $O(c_n^2 + F_n)$ of the frictionless optimum, by the quadratic comparison in Lemma B.5. Every investor can choose this portfolio. Any paid-access portfolio instead loses at least $\tau/4$ relative to the same gross optimum. Paid access is therefore absent eventually. Nonnegative fees and uniform positive curvature imply that every selected exposure converges to $Q_0 b$. This establishes the exposure and access hypotheses of Lemma B.6; they are not assumed from the market fund's positive AUM alone.

In normalized limiting units, each type's total risky exposure has mass one. Put $a_B = A_B^0/Q_0 > 0$, $\beta_B = 1/a_B$, and let r be the number of types with $y_{iB} > 0$. Necessarily $1 \leq r \leq 4$. Limiting cost optimality implies $\pi_i \geq \beta_B$ for each of these r types; $\pi_i = +\infty$ is allowed. Also

$$a_B = \frac{1}{4} \sum_i y_{iB} \leq \frac{r}{4}.$$

If this inequality were strict, then $\beta_B > 4/r$. Choose

$$p = \frac{1}{2} \left(\beta_B + \frac{4}{r} \right) < \beta_B.$$

It is also strictly below the r types' alternative costs. At the lower own quote pF_n/Q_0 , strict lifting makes each of these types hold Q_0 units of B in the limit. The resulting normalized revenue has lower bound

$$\frac{pr}{4} = \frac{1}{2} \left(\frac{\beta_B r}{4} + 1 \right) > 1,$$

contradicting lower-quote protection. Thus $a_B = r/4$, and every type with positive limiting market-fund demand holds its full unit of normalized exposure through B .

Suppose $r < 4$. Each remaining type holds no B in the limit, so her alternative replication is feasible and has finite $\pi_i \leq \beta_B$. Equality is impossible. If it held for one additional type, the quote

$$p = \frac{1}{2} \left(\beta_B + \frac{4}{r+1} \right)$$

would be strictly below $\beta_B = 4/r$, attract at least $r+1$ types by strict lifting, and earn more than F_n eventually. Consequently, all remaining costs are strictly below β_B .

Sort those $4-r$ costs in decreasing order. For every fixed $0 < p < \pi_{(j)}$, the quote pF_n/Q_0 is a lower own quote and captures at least $r+j$ types in the limit. Protection implies $p(r+j)/4 \leq 1$. Letting p increase to $\pi_{(j)}$ gives

$$\pi_{(j)} \leq \frac{4}{r+j}. \tag{B.10}$$

The r full-market-fund types have no limiting home or non-market holdings. Hence limiting cost optimality and break-even for the $N = |\mathcal{O}|$ non-market products now give

$$N = \sum_{T \in \mathcal{O}} \beta_T a_T = \frac{1}{4} \sum_{j=1}^{4-r} \pi_{(j)} \leq \sum_{j=1}^{4-r} \frac{1}{r+j} = H_4 - H_r.$$

If $r < 4$, at least one non-market product is needed: home investment alone cannot reproduce the remaining types' market exposure. Thus N is a positive integer. The possible cases are

r	$H_4 - H_r$	implication
1	13/12	$N = 1$
2	7/12	$N \geq 1$ is impossible
3	1/4	$N \geq 1$ is impossible.

The remaining case $r = 1, N = 1$ is also impossible: Lemma B.1 permits one non-market product to serve at most one type with her home asset, whereas three types here would require such a replication.

It follows that $r = 4$. Every type has $y_{iB} = 1$, so nonnegativity and total mass one force all her limiting non-market and home holdings to be zero. The argument applies to every convergent subsequence of the bounded holdings. Therefore the stated limits hold along the original sequence as well. $\square$

B.6 Proof of Proposition 1 and its quantifiers

Proof of Proposition 1. Fix $m \in (0, 1)$ and $\tau > 0$. Consider any sequence of primitive-economy equilibria with $c_n, F_n > 0$ and $(c_n, F_n) \rightarrow (0, 0)$. Investor holdings are uniformly bounded by Lemma A.4. From any subsequence, one may select a further subsequence with a constant active menu and convergent incumbent holdings, since the catalogue is finite. Lemma B.7 excludes an omitted market fund and a market fund with vanishing AUM. Lemma B.8 then implies, for every type,

$$q_{iB,n} \rightarrow Q_0(m), \quad x_{ii,n} \rightarrow 0, \quad q_{iT,n} \rightarrow 0 \quad (T \neq B), \quad (\text{B.11})$$

with paid access absent eventually. These conclusions hold for every selected equilibrium allocation and every convergent subsequence, hence for the original sequence.

Choose the desired demand floor $Q_0(m)/2 > 0$. If there were no positive pair of independent ceilings on c and F under which every equilibrium had B active, $q_{iB} \geq Q_0(m)/2$ for all types, and no paid access, one could select a violating equilibrium with $0 < c_n, F_n < 1/n$ for each sufficiently large n . This contradicts (B.11) and eventual access exclusion. Thus some positive $\bar{c}(m, \tau)$ and $\bar{F}(m, \tau)$ have the required properties. Their minimum with the primitive neighborhood from Lemma A.2 maintains all setup assumptions. Since each type has mass 1/4, $\text{AUM}_B \geq Q_0(m)/2$ throughout this rectangle, proving the stated uniform positive lower bound. $\square$

The joint sequences in this proof are a device for establishing two independent ceilings. They impose no functional relationship between F and c on an economy within the resulting rectangle. In particular, any fixed $F \in (0, \bar{F})$ is admissible for every $0 < c < \bar{c}$. The result concerns all equilibria that exist there; it does not, by this argument alone, establish existence at every point. The ceilings are sufficient existence bounds for the inequalities used in the proof, not sharp operating-cost or dispersion thresholds.

For uniformity in primitives, let m range over a compact $M \subset (0, 1)$, and let non-home access charges have a common positive floor. If common ceilings failed, take a counterexample sequence

and a subsequence $m_n \rightarrow m_* \in M$. Uniform curvature and continuity give the same proof with $Q_0(m_n) \rightarrow Q_0(m_*)$. The catalogue geometry is unchanged and the entry margins are strict. Hence common ceilings and the demand floor $\frac{1}{2} \min_{m \in M} Q_0(m) > 0$ can be chosen on these sets.

B.7 Quantitative consequences of selection

The following consequences make explicit what the weak operating-cost bound supplies when the later analysis asks whether additional products can operate. They are not assumptions in the proof of Proposition 1. All orders below are uniform on a sufficiently small independent (c, F) rectangle, at fixed m, τ , or on the compact primitive sets just described.

Proposition B.9 (Market-fund demand and necessary specialist viability). *After reducing the independent ceilings if necessary, every equilibrium satisfies*

$$q_{iB} = Q_0(m) + O(c + F), \quad \text{AUM}_B = Q_0(m) + O(c + F), \quad \delta_B = \frac{F}{Q_0(m)} + O(F(c + F)). \quad (\text{B.12})$$

Home holdings are $O(c + F)$. For every active non-market product, there is a common constant C_{viab} such that

$$F \leq C_{\text{viab}} c^2, \quad \delta_T = O(c), \quad \text{AUM}_T = O(c). \quad (\text{B.13})$$

These are necessary bounds, not sufficient conditions for product entry.

Proof. The sequential proof of Proposition 1 also implies that $\delta_T / \delta_B = \text{AUM}_B / \text{AUM}_T \rightarrow \infty$ for any active non-market product along a sequence tending to the origin. By another counterexample-sequence argument, reduce the rectangle so that all such fees exceed $2\delta_B$. The demand floor gives $\delta_B = F / \text{AUM}_B = O(F)$.

There is no paid access. Write total exposure as $e_i = M_i b + r_i$, where $M_i = \mathbf{1}' e_i$, $r_i = D_i z_i$, and $z_i = ((q_{iT})_{T \neq B}, x_{ii})$. If $\|r_i\| \leq \epsilon_* \ell(z_i)$ and $\ell(z_i) > 0$, Lemma B.2 supplies a removable kernel component. It costs more than $2\delta_B(3/4)\ell(k)$, whereas the same exposure costs $\delta_B \ell(k)$ through B . This violates optimality. Consequently

$$\ell(z_i) \leq \epsilon_*^{-1} \|r_i\|. \quad (\text{B.14})$$

The assertion also holds when $z_i = 0$.

Compare the chosen portfolio with the feasible portfolio holding its same total risky mass M_i entirely in B . Expected returns agree because $\mathbf{1}' r_i = 0$. Optimality gives

$$\frac{1}{2} r_i' \Sigma r_i \leq -M_i b' \Sigma r_i - \sum_{T \neq B} (\delta_T - \delta_B) q_{iT} + \delta_B x_{ii}. \quad (\text{B.15})$$

At $c = 0$, $b' \Sigma_0 r_i = 0$; smoothness implies $|b' \Sigma r_i| \leq C c \|r_i\|$. The M_i 's are uniformly bounded, the displayed fee differences are positive, and $x_{ii} \leq \ell(z_i)$. Combining uniform positive curvature, (B.14), and $\delta_B = O(F)$ in (B.15) yields $\|r_i\| = O(c + F)$ and $\ell(z_i) = O(c + F)$.

Every type holds B positively, so its first-order condition is $1 - \delta_B = b' \Sigma e_i = M_i b' \Sigma b + b' \Sigma r_i$. Since $b' \Sigma b = w/4 + O(c^2)$, this gives $M_i = Q_0 + O(c + F)$. Subtracting $\ell(z_i)$, averaging across types, and applying break-even proves (B.12).

For an active non-market product T , some type holds it positively. Subtracting its first-order condition from the market fund's gives $\delta_T = \delta_B - (b_T - b)' \Sigma e_i$. The covariance with $M_i b$ is $O(c)$, and the covariance with r_i is $O(c + F)$. Thus $\delta_T = O(c + F)$, while the quantity bound already gives $\text{AUM}_T = O(c + F)$. For a common $K > 0$, break-even therefore requires

$$F \leq K(c + F)^2 \leq 2Kc^2 + 2KF^2.$$

Reduce $\bar{F}$ to at most $1/(4K)$; absorption gives $F \leq 4Kc^2$. Substituting this necessary condition into the initial bounds gives both $O(c)$ bounds in (B.13), with $C_{\text{viab}} = 4K$. $\square$

Lemma 4 (Completion scale). *Fix a positive normalized operating cost ϕ , defined by (16). For sufficiently small positive c , every active non-market product in every primitive-economy equilibrium has both AUM and fee of order c , with positive lower and upper bounds after division by c , uniform on the compact parameter sets in Appendix B.9. Free home positions are $O(c)$, and final exposures are within $O(c)$ of the common frictionless target.*

Proof of Lemma 4. Use normalized cost ϕ from (16). Fix m and $\phi > 0$; then the recovered operating cost $F = \Theta(m)c^2\phi$ is below the independent cost ceiling for all sufficiently small c . The preceding bounds therefore apply. For an active specialist, break-even reads

$$\delta_T \text{AUM}_T = \Theta(m)c^2\phi.$$

The upper bound $\text{AUM}_T \leq Cc$ gives $\delta_T \geq \phi\Theta(m)c/C$; the upper fee bound gives the corresponding AUM lower bound. These bounds are uniform when m lies in a compact subset of $(0, 1)$, ϕ in a compact subset of $(0, \infty)$, and access costs have a common positive floor. In particular, the positive lower bound on normalized cost is needed for uniform lower fee and AUM bounds.

The general home bound is $O(c + F) = O(c)$ in these coordinates. Equation (A.11) and the preceding proof also give $e_i - e^*(c) = O(c)$. This establishes the scale lemma without changing Proposition 1's independent cost and dispersion ceilings. $\square$

B.8 Existence when operating cost stays fixed

The statement about every equilibrium in Proposition 1 is logically separate from constructing an equilibrium. The following result supplies existence in a region relevant for holding F fixed as dispersion decreases. It also gives the equilibrium interpretation of the conditional benchmark in Appendix A.9.

Proposition B.10 (A market-fund-only equilibrium at fixed cost). *After reducing the positive ceilings on c and F , there is an $\epsilon > 0$ such that a market-fund-only equilibrium exists whenever*

$$0 < c < \bar{c}, \quad 0 < F < \bar{F}, \quad c^2/F < \epsilon. \quad (\text{B.16})$$

For every fixed $F \in (0, \bar{F})$, a market-fund-only equilibrium therefore exists for all sufficiently small positive c . Moreover, every equilibrium is market-fund-only once $C_{\text{viab}}c^2 < F$, in addition to the maintained bounds.

Proof. Construction and lower quotes. Proposition A.11 supplies a low quote δ_B at the menu $\{B\}$, with no paid access, positive fund holdings for every type, and protection against every lower quote. Writing $D_B(\eta, c)$ for aggregate demand at this menu, it satisfies

$$\delta_B D_B(\delta_B, c) = F, \quad \delta_B = F/Q_0 + O(F(c + F)). \quad (\text{B.17})$$

The unique fund-and-home choices supply a common incumbent allocation. It remains to verify every omitted product's entry test.

Access after entry. The market-fund offer remains available following entry of any one omitted product. Its feasible net value is within $O(c^2 + F)$ of $G^*(c)$. Reducing the rectangle makes this gap less than $\tau/4$. Any portfolio paying for non-home access has net value at most $G^*(c) - \tau/4$. Hence paid access is absent for every optimizer, at every entrant quote. This comparison is uniform over entrant products and prices.

Entrant types without an exact replication. Consider an entrant T at any fee $\eta \geq 0$. For a type i outside the complementary-triplet case of Lemma B.1, the cone generated by $d_T = b_T - b$ and $H_i = \iota_i - b$ has no nonzero nonnegative kernel. Compactness of its unit-mass slice therefore gives a constant C such that $q_T + h \leq C\|q_T d_T + h H_i\|$ for all $q_T, h \geq 0$. There are finitely many entrant/type pairs, so one constant works for all these cases.

Write a deviating optimum as $e = Mb + r$, with $r = q_T d_T + h H_i$. Comparing it with the same total mass in B gives (B.15) with the single fee η in place of δ_T . Even if $\eta < \delta_B$, its fee-saving term is at most $\delta_B(q_T + h)$, because $\eta \geq 0$. The cone bound, $\delta_B = O(F)$, and positive curvature imply

$$\|r\| + q_T + h = O(c + F), \quad (\text{B.18})$$

uniformly over quotes. The feasible incumbent offer and the gross-value gap identity also give $e = Q_0 b + O(c + \sqrt{F})$. Thus $M \rightarrow Q_0$, and $q_B = M - q_T - h$ is bounded away from zero after reducing the rectangle. If $q_T > 0$, the two fund first-order conditions imply $\eta = \delta_B - d'_T \Sigma e = O(c + F)$. If $q_T = 0$, entrant revenue is zero. Each non-replicating type therefore contributes at most $C(c + F)^2$ to aggregate entrant revenue, including its population weight.

The complementary-home type of a triplet. Only one type can replicate market exposure with a given triplet. For that type, $3b_T + \iota_i = 4b$. If $\eta > 4\delta_B/3$, an optimum cannot have both $q_T > 0$ and $h > 0$: subtract $t(3, 1)$, where $t = \min\{q_T/3, h\} > 0$, and add $4t$ units of B . Exposure is unchanged and fees fall by $t(3\eta - 4\delta_B) > 0$. Thus at least one coordinate is zero. On either coordinate axis the corresponding direction is nonzero, so the same cone argument proves (B.18) and the $C(c + F)^2$ revenue bound. No quote gap bounded away from $4\delta_B/3$ is needed for this strict replacement.

If $0 \leq \eta \leq 4\delta_B/3$, the same type may hold a nonvanishing amount of the entrant. Nevertheless, $e_j = Q_0/4 + O(c + \sqrt{F})$ for each asset, by the feasible-offer value bound. For any $j \in T$, nonnegative market-fund holdings give $e_j = q_B/4 + q_T/3 \geq q_T/3$, and hence $q_T \leq 3Q_0/4 + O(c + \sqrt{F})$. The

type's mass is $1/4$, so its entrant-revenue contribution is at most

$$\frac{1}{4} \frac{4\delta_B}{3} \left(\frac{3Q_0}{4} + O(c + \sqrt{F}) \right) = \frac{F}{4} + O(F(c + \sqrt{F})), \quad (\text{B.19})$$

using (B.17). All other types are non-replicating and have the preceding quadratic revenue bound.

All-price entry protection. The preceding cases cover all eight omitted products, all prices, and all optimizers. Their aggregate revenue is bounded either by $C(c + F)^2$, or by

$$\frac{F}{4} + CF(c + \sqrt{F}) + C(c + F)^2.$$

Since $(c + F)^2/F \leq 2c^2/F + 2F$, first choose $\epsilon > 0$ sufficiently small, then reduce $\bar{c}, \bar{F}$, so that each bound is strictly less than F on (B.16). This proves entry protection even under focal-upper demand. Investor optimality, incumbent break-even and lower-quote protection were already verified, so the constructed menu is an equilibrium.

For fixed F , the wedge condition holds eventually as c decreases. Proposition B.9 excludes every active non-market product whenever $C_{\text{viab}}c^2 < F$. Together with Proposition 1, this proves the last assertion. It does not rule out multiple market-fund-only equilibria or assert existence outside the stated region. $\square$

B.9 Uniformity and continuation domains

Proposition 1 and the necessary viability bound use Assumption 2. Their constants can be chosen uniformly when m ranges over a compact set $M \subset (0, 1)$ and access costs have a common positive floor, with sufficiently small independent ceilings on c and F . In this region the general bounds are

$$\text{AUM}_B = \frac{4}{1 + 3m^2} + O(c + F), \quad \delta_B = \frac{1 + 3m^2}{4}F + O(F(c + F)).$$

Home positions are $O(c + F)$. Every active non-market product satisfies the upper order- c bounds after imposing its necessary viability condition. These facts do not give positive lower bounds after division by c .

For Lemma 4, use the normalized-cost coordinates of (16). Its bounds are uniform for $m \in M$, ϕ in a compact set $\Phi \subset (0, \infty)$, and a common positive access-cost floor. On these sets $\Theta(m)$ and ϕ are bounded above and away from zero. Recovering $F = \Theta(m)c^2\phi$ therefore gives two-sided cost bounds proportional to c^2 , and the independent selection ceiling holds for a common sufficiently small c . These uniform bounds do not extend to paths on which ϕ tends to zero or infinity. Selection, viability and these scale results do not require the numerical cap $m^2 < 44/69$ or the condition keeping middle investors' home holdings exactly zero.

The finite- c continuation claims use the same normalized-cost coordinates on the following sufficient domains. Take compact sets

$$M_{\text{cont}} \subset (0, \sqrt{44/69}), \quad \Phi_{\text{cont}} \subset (0, \infty),$$

with

$$\sup \Phi_{\text{cont}} < \inf_{m \in M_{\text{cont}}} \frac{2(1 - m^2)}{m^2},$$

and the branch and event separations listed in Appendix D. For scope equality on a domain where the pair is an exact equilibrium scope, impose in addition the strict symmetric-pair support condition

$$m^2[10 + \phi - 8p_s(\phi)] < 2, \quad (\text{B.20})$$

uniformly with a positive margin, where $p_s(\phi)$ is the symmetric pair quote in (C.34). This condition keeps both middle investors' home holdings zero on the constructed symmetric continuation. It is needed for the local lower-quote protection argument, not just for an exact statement about delegation. Appendix D derives the condition and gives a counterexample to symmetric persistence under the two numerical domain bounds alone. The empty and pooled continuations do not require (B.20); unequal continuations use their own strict support conditions.

These qualifications apply to the scope equality in Proposition D.9 and to the continued-branch clientele and direct-implementation statements. Exact classification in Proposition 2 holds for every $\phi > 0$. The all-equilibrium exclusion result in Appendix D does not require the additional symmetric support condition. Equality of scope sets uses that condition to supply a protected primitive pair equilibrium whenever the pair belongs to the exact correspondence. No assertion is made that pair scope is impossible outside this sufficient continuation domain.

How much of the pair region the support condition covers. Let $B(\phi) = 10 + \phi - 8p_s(\phi)$ on $0 < \phi \leq \phi_U$. On the lower symmetric branch, $\phi = 4p_s - 6p_s^2$, so $B = 10 - 4p_s - 6p_s^2$. On the upper branch, $\phi = 3p_s - 3p_s^2$, so $B = 10 - 5p_s - 3p_s^2$. Both expressions decrease with the positive quote. They agree at $\phi = 2/3$, where $B = 8$. Thus B decreases continuously from the limit 10 to $B(\phi_U) = 358/49$.

For $m^2 \leq 1/5$, (B.20) holds at every positive pair cost, with uniformity on the stated compact sets. For $1/5 < m^2 < 49/179$, it holds above the unique threshold $\phi_*(m) \in (0, \phi_U)$ satisfying $m^2 B(\phi_*(m)) = 2$. For $m^2 \geq 49/179$, it holds at no cost in the pair region. These statements concern this support test; all other continuation and event conditions remain in force.

For comparison with the loading $m = .49$ used in the numerical exercise, the support threshold is

$$\phi_*(.49) = -\frac{31238}{7203} + \frac{32\sqrt{526}}{147} \simeq 0.655781.$$

Only the upper part of the approximate pair-cost interval meets this sufficient support test at that loading. This does not rule out pair equilibria outside the test's domain. The finite-dispersion numerical evidence in Section 5 has its own stated range and does not follow from this continuation calculation.

C The normalized problem and its exact classification

C.1 The demand lemma and the appendix's units

Section 3.2 uses the fee and holding units of Sections 2 and 3.1. This appendix solves the same specialist problem after dividing fees and holdings by their common scales. The conversion below connects the body's demand formulas to the exact calculations that follow. Appendix D derives this problem from the primitive economy and states its domains.

Let $b_B = \mathbf{1}/4$, $d_T = b_T - b_B$, $H_i = \iota_i - b_B$, and $g = (-2, 0, 2, 0)'$. Write $\alpha = mc/(1 + 3m^2)$, as in the body. The scales in (D.1) satisfy

$$tc = \alpha, \quad Kc = \frac{\alpha}{1 - m^2}, \quad \Theta(m)c^2 = \frac{\alpha^2}{1 - m^2}. \quad (\text{C.1})$$

The following identities define the normalized variables; they are exact changes of units for the given positive m and c .

Object	Body units	Normalized variable used below
Specialist fee	$\delta_T = \alpha p_T$	p_T
Specialist holding	$q_{iT} = \frac{\alpha}{1 - m^2} u_{iT}$	u_{iT}
Free home holding	$x_{ii} = \frac{\alpha}{1 - m^2} h_i$	h_i
Exposure adjustment	$r_i = \frac{\alpha}{1 - m^2} y_i$	$y_i = \sum_T u_{iT} d_T + h_i H_i$
Fund AUM	$\text{AUM}_T = \frac{\alpha}{1 - m^2} Q_T$	$Q_T = \frac{1}{4} \sum_i u_{iT}$
Fund revenue	$R_T = \Theta(m)c^2 \mathcal{R}_T$	$\mathcal{R}_T = p_T Q_T$
Operating cost	$F = \Theta(m)c^2 \phi$	ϕ

The pooled quote denoted z below is $z = p_C = \delta_C/\alpha$. Dividing revenue and cost by the same positive quantity preserves break-even and deviation comparisons. The approximation concerns the investor problem derived from the primitive economy, not these unit identities.

The approximation and the market-fund fee. The body studies the limit with m fixed, the market-fund fee bounded by a constant times c^2 , and the specialist fee ratios $p_T = \delta_T/\alpha$ in a compact subset of positive values. Lemma D.2 gives the precise demand statement: every accumulation point of the primitive optimal holdings, divided by $\alpha/(1 - m^2)$, solves the normalized problem below. Its uniform bound excludes paid non-home access and applies to the incumbent and focal-deviation selections. This is an outer convergence statement; it does not assert convergence to every tied implementation.

The market-fee order has an equilibrium foundation. Proposition B.9 gives $\delta_B = F/Q_0 + O(F(c + F))$, and specialist viability requires $F \leq C_{\text{viab}} c^2$. Thus the fee is $O(c^2)$ in economies sustaining

specialists. Fixing positive ϕ also sets $F = \Theta(m)\phi c^2$; Lemma 4 then keeps active specialist fee ratios away from zero and infinity. For a conditional offered menu, the lemma uses this market-fee and quote domain. Small dispersion by itself does not impose that domain on arbitrary market-fund fees.

Proof of Lemma 3. Lemma D.1 optimizes the market-fund position and gives the leading specialist problem. Its desired exposure adjustment in the body's units is $r^* = \alpha g / (1 - m^2)$. Substituting the conversion above gives

$$\frac{1 - m^2}{2} \|r^* - r_i\|^2 + \sum_T \delta_T q_{iT} = \Theta(m)c^2 \left\{ \frac{1}{2} \|g - y_i\|^2 + \sum_T p_T u_{iT} \right\}. \quad (\text{C.2})$$

The expression in braces is (C.16). Its minimization is equivalent to (C.15), since $\|g\|^2/2 = 4$ is constant. Positive fees and nonnegative holdings are preserved by the conversion. For fixed positive fee ratios, Lemma D.2 excludes paid access, bounds the rescaled optimal holdings, and shows that every accumulation point solves (C.16).

Lemma C.3 establishes existence and uniqueness of the optimal exposure and total fee payment. Its necessary and sufficient first-order conditions, in body units, are

$$\delta_T \geq (1 - m^2)d'_T(r^* - r_i), \quad q_{iT}[\delta_T - (1 - m^2)d'_T(r^* - r_i)] = 0. \quad (\text{C.3})$$

For the home asset, replace d_T by H_i , set its fee to zero, and replace q_{iT} by x_{ii} . Since $(1 - m^2)d'_T r^* = \alpha g' d_T = \alpha_T$, (C.3) equates a held product's fee to the value of an additional unit. An unused product's fee is at least that value. Expanding the square in (C.2) gives the body's objective: alpha net of fees less half the adjustment variance, up to a constant.

To express demand as quantities, let J be an independent set of active directions and let D_J contain those columns. Include H_i when the home asset is used. Write q_J for their holdings, $\alpha_J = \alpha D'_J g$ for their alphas, and δ_J for their fees, with a zero home fee. Define $\Omega_J = (1 - m^2)D'_J D_J$. The active equalities give

$$\begin{aligned} q_J &= \Omega_J^{-1}(\alpha_J - \delta_J), & q_J &\geq 0, \\ \alpha_k - \delta_k - (1 - m^2)d'_k D_J q_J &\leq 0 & (k \notin J). \end{aligned} \quad (\text{C.4})$$

The last inequality includes an unused home asset, with direction H_i and fee zero. These are the necessary and sufficient conditions for the candidate to be optimal. They retain zero-quantity boundaries. Lemma C.3 shows that every extreme optimal implementation has such an independent support; the complete optimizer set follows from its extreme implementations. Redundant held positions need not have an invertible covariance matrix, so independence is a condition on this representation, not a restriction on the offered menu.

Within a region where the active set stays fixed, demand is affine in fees and the own-fee derivative is $-\Omega_J^{-1}]_{TT}$. Partition Ω_J into product T 's variance ω_T , its covariance row v' with the

other active directions, and their covariance matrix W . The Schur complement gives

$$[\Omega_J^{-1}]_{TT} = \frac{1}{\omega_T - v'W^{-1}v} \geq \frac{1}{\omega_T}. \quad (\text{C.5})$$

Positive definiteness makes the denominator positive. Equality holds when the covariance row is zero; for a single active position the coefficient is simply $1/\omega_T$. Thus joint reallocation can steepen the fee response. At a change of active set, use the adjacent demand regions and their boundary conditions. No claim of a common slope across all fees or all menus is needed. Strict convexity in exposure gives a unique adjustment; the objective then also fixes the fee bill. Different holdings can implement that adjustment at the same cost. The jointly compatible incumbent selection and focal deviation convention below implement the tie convention stated with Definition 1 (Appendix A). $\square$

The local alpha interpretation. Measured in Section 2.4's expected-return units, all products have the same expected excess return, one, before fees. Write $\rho_T = g'd_T = 2(b_{T,2} - b_{T,0})$, so that $\alpha_T = \alpha\rho_T$ in (14). The before-fee alpha of product T against B is therefore

$$1 - \frac{b'_T \Sigma b_B}{b'_B \Sigma b_B} = \frac{mc}{1 + 3m^2} g'd_T + O(c^2) = \alpha\rho_T + O(c^2). \quad (\text{C.6})$$

To obtain the expansion, put $\nu = (1, 0, -1, 0)'$. At zero dispersion, $b'_B \Sigma b_B = (1 + 3m^2)/4$; its first derivative is zero, while $\Sigma'(0)b_B = (m/2)\nu$. Since $g = -2\nu$, differentiating the covariance ratio gives (C.6). This concerns a product with fixed portfolio weights and requires no private information or stock-selection skill.

Both pair funds have $\rho_T = 1$. The pooled fund has $\rho_C = 2/3$ and $\|d_C\|^2 = 1/12$, which give the benefit and leading residual variance in the worked example. Its overlap with the diversifier's home direction is $d'_C H_2 = 1/12$, with $\|H_2\|^2 = 3/4$. Solving the joint first-order conditions subject to nonnegativity gives the two customer groups in Section 3.2.4. The general marginal value in (C.3) also depends on covariance with the investor's existing adjustment. A product's before-fee alpha alone is therefore not a general cap on its fee: a product can be useful because it offsets risk elsewhere in the combined position.

Pooled demand and revenue in the body's units. For the diversifier investor, the active columns (d_C, H_2) give

$$\Omega = (1 - m^2) \begin{pmatrix} 1/12 & 1/12 \\ 1/12 & 3/4 \end{pmatrix}, \quad \alpha_J = (2\alpha/3, 2\alpha)', \quad \delta_J = (\delta_C, 0)'.$$

Inverting this system gives $q_{2C} = (6\alpha - 27\delta_C/2)/(1 - m^2)$ and $x_{22} = (2\alpha + 3\delta_C/2)/(1 - m^2)$ while $0 < \delta_C < 4\alpha/9$. At and above that boundary, $q_{2C} = 0$ and $x_{22} = 8\alpha/[3(1 - m^2)]$. The other types leave home unused and hold $[8\alpha - 12\delta_C]_+/(1 - m^2)$, as verified in the pooled branch calculation

below. Aggregation gives

$$\text{AUM}_C = \frac{1}{1-m^2} \begin{cases} 15\alpha/2 - 99\delta_C/8, & 0 < \delta_C \leq 4\alpha/9, \\ 6\alpha - 9\delta_C, & 4\alpha/9 \leq \delta_C \leq 2\alpha/3, \\ 0, & \delta_C \geq 2\alpha/3. \end{cases} \quad (\text{C.7})$$

Multiplying by δ_C yields the revenue curve. Its global maximum is $25\alpha^2/[22(1-m^2)]$, attained at $\delta_C = 10\alpha/33$ in the first region. Revenue is decreasing throughout the second region. Equivalently, $F/\max R_C = \phi/\phi_W$, so the pooled fund can cover its cost exactly when $\phi \leq \phi_W$, as in Section 3.2.4. These demand and revenue statements concern that conditional menu; entry protection is established in the classification.

The home response behind the pair's pricing boundary. The local calculation is especially transparent in the normalized units above. Let the common pair fee be $p \in (1/3, 1)$, and let L make a sufficiently small cut to $\eta < p$. Types 0 and 1 hold only the two pair directions; type 2 uses her home asset alone. Type 3 uses (d_L, d_R, H_3) . Applying (C.4) on these supports gives

$$u_{0L} = u_{1L} = 4(1 - \eta), \quad u_{3L} = 4 + 4p - 8\eta, \quad u_{3R} = 4 - 8p + 4\eta, \quad h_3 = 4(p - \eta).$$

The home holding permits type 3 to move toward the cheaper fund without changing her exposure adjustment in asset 3. Aggregating gives $Q_L(\eta; p) = 3 + p - 4\eta$, so the derivative of unilateral revenue as η approaches p from below is $3 - 7p$. A small cut is therefore profitable when $p > 3/7$, or when the body's common fee exceeds $3\alpha/7$. Symmetric per-fund revenue at $p = 3/7$ is $3p(1-p) = 36/49$. This local calculation identifies the mechanism and the necessary boundary. The all-quote and omitted-entry comparisons in the pair branch proof below establish protection on the full claimed range.

Ties and zero fees. The independent-support calculation also covers tied positive-fee implementations through their extreme points. Incumbent allocations must be jointly compatible, while a focal deviation uses its own upper optimal quantity, as defined below. At zero product fees, the replication cycles in (C.19) can instead make the set of optimal holdings unbounded. This does not mean that no optimal exposure exists. Curves drawn at a zero specialist fee use limits from positive quotes.

C.2 The exact portfolio-adjustment problem

There are four equally weighted investor types and four underlying assets. Write

$$b_B = \frac{1}{4}(1, 1, 1, 1)' \quad \text{and} \quad \mathcal{S} = \{x \in \mathbb{R}^4 : \mathbf{1}'x = 0\}.$$

Section 3 establishes the market fund in the primitive economy. At fixed positive normalized cost ϕ , Appendix D optimizes its common component and derives the portfolio-adjustment problem

studied here. The market fund carries no fee at this order: its fee is order c^2 , whereas specialist fees are order c . Thus B is available as the reference implementation but has no additional pricing or break-even equation in this exact problem. An empty specialist menu means that investors adjust market-fund exposure using their free home assets alone. The classification below solves this derived economy for every $\phi > 0$; its transfer to positive dispersion is proved separately.

For any arc product T , let $d_T = b_T - b_B$. Use

$$L = P_{12} = P_L, \quad R = P_{23} = P_R, \quad C = P_{123},$$

and denote the two neutral triples by

$$N_L = P_{012}, \quad N_R = P_{230}.$$

The proposition imposes no product restriction. Its residual product universe is

$$\mathcal{P}_8 = \{P_{01}, L, R, P_{30}, N_L, C, N_R, P_{301}\}. \quad (\text{C.8})$$

Table C.1 records the directions and their completion-value coefficients in the independent basis (d_L, d_R, d_C) . The Gram matrix in these coordinates is

Table C.1: **The full residual product universe**

Product T	$[d_T]_{(L,R,C)}$	ρ_T
P_{01}	$(0, -1, 0)$	-1
L	$(1, 0, 0)$	1
R	$(0, 1, 0)$	1
P_{30}	$(-1, 0, 0)$	-1
N_L	$(2/3, 0, -1)$	0
C	$(0, 0, 1)$	$2/3$
N_R	$(0, 2/3, -1)$	0
P_{301}	$(-2/3, -2/3, 1)$	$-2/3$

$$G = \begin{pmatrix} 1/4 & 0 & 1/12 \\ 0 & 1/4 & 1/12 \\ 1/12 & 1/12 & 1/12 \end{pmatrix}. \quad (\text{C.9})$$

Thus a coordinate vector x has squared norm $x'Gx$, which is the ordinary Euclidean squared norm of its underlying four-asset residual vector. The useful replication identities are

$$d_{01} = -d_R, \quad d_{30} = -d_L, \quad (\text{C.10})$$

$$d_{NL} = \frac{2}{3}d_L - d_C, \quad d_{NR} = \frac{2}{3}d_R - d_C, \quad (\text{C.11})$$

$$d_{301} = d_C - \frac{2}{3}(d_L + d_R). \quad (\text{C.12})$$

Type i may also take a nonnegative amount of her free home direction

$$H_i = \iota_i - b_B.$$

In the same basis,

$$\begin{aligned} H_0 &= (0, 0, -3), & H_1 &= (0, -2, 3), \\ H_2 &= (2, 2, -3), & H_3 &= (-2, 0, 3), \end{aligned} \quad (\text{C.13})$$

and the associated completion-value coefficients are

$$(\rho_0^H, \rho_1^H, \rho_2^H, \rho_3^H) = (-2, 0, 2, 0). \quad (\text{C.14})$$

The home direction is investor-specific, free, and constrained to have a nonnegative quantity. Paid access to other assets is absent from the exact quotient.

Definition C.1 (Exact investor problem). *Let $E \subseteq \mathcal{P}_8$ be an active completion menu and let every $T \in E$ post a normalized fee $p_T > 0$. Type i solves*

$$\max_{\substack{u_{iT} \geq 0, \\ h_i \geq 0, \\ T \in E}} \left\{ \sum_{T \in E} (\rho_T - p_T) u_{iT} + \rho_i^H h_i - \frac{1}{2} \left\| \sum_{T \in E} u_{iT} d_T + h_i H_i \right\|^2 \right\}. \quad (\text{C.15})$$

The value coefficients have the common representation $\rho_T = g' d_T$ and $\rho_i^H = g' H_i$, where $g = (-2, 0, 2, 0)'$. Thus the objective is $g' y_i - \|y_i\|^2/2 - \sum_T p_T u_{iT}$, with $y_i = \sum_T u_{iT} d_T + h_i H_i$. This is the normalized adjustment in Appendix D; the body's exposure adjustment is $r_i = K c y_i$. Completing the square gives

$$\frac{1}{2} \|y_i - g\|^2 + \sum_T p_T u_{iT} = \frac{1}{2} \|g\|^2 - \left(g' y_i - \frac{1}{2} \|y_i\|^2 - \sum_T p_T u_{iT} \right).$$

Thus minimizing distance cost plus fees and maximizing the displayed covariance value net of fees select exactly the same portfolios. The equivalent problem is

$$\min_{u_{iT} \geq 0, h_i \geq 0} \left\{ \frac{1}{2} \|y_i - g\|^2 + \sum_{T \in E} p_T u_{iT} \right\}, \quad y_i = \sum_{T \in E} u_{iT} d_T + h_i H_i. \quad (\text{C.16})$$

This compares the risk cost of departing from the desired adjustment with the fee bill required to obtain it. Its restriction to $E = \{C\}$ is the worked example in Section 3.2.4. The optimized residual exposure and value in (C.15) are unique. Product quantities need not be unique when two implementations lie on a replication face. Let $\mathcal{U}_i(E, p)$ denote type i 's optimizer correspondence. An incumbent outcome selects one vector from every $\mathcal{U}_i(E, p)$, and the same four selected vectors must be used in all active products' break-even equations. For this jointly compatible selection, define

$$Q_T(p; E) = \frac{1}{4} \sum_{i=0}^3 u_{iT}, \quad \mathcal{R}_T(p; E) = p_T Q_T(p; E). \quad (\text{C.17})$$

Here $Q_T = \text{AUM}_T / (Kc)$ and $\mathcal{R}_T = R_T / [\Theta(m)c^2]$, where AUM_T and R_T are the body's AUM and fee revenue. The dictionary in Appendix C.1 gives the remaining conversions. In particular, tied demand cannot be maximized product by product using mutually incompatible incumbent implementations.

One-product deviations use a different, focal convention. For a deviation menu $G \ni T$, if T posts $\eta > 0$ while all nonfocal prices remain fixed, define

$$\bar{Q}_T(\eta; p_{-T}, G) = \frac{1}{4} \sum_{i=0}^3 \max_{u_i \in \mathcal{U}_i(G, (\eta, p_{-T}))} u_{iT}. \quad (\text{C.18})$$

This is the contestable upper tied demand for the focal product. It is the limit selected by lowering only the focal quote from η to $\eta - \gamma$, with $\gamma \downarrow 0$, while every nonfocal posted price stays fixed. A remaining nonfocal replication tie may leave several implementations, but it cannot change the greatest focal coordinate.

The restriction $\eta > 0$ is substantive. At zero price the normalized problem contains the nonnegative cycles

$$3C + H_0 = 0, \quad 3N_R + H_1 = 0, \quad 3P_{301} + H_2 = 0, \quad 3N_L + H_3 = 0. \quad (\text{C.19})$$

Quantities can therefore be unbounded along a zero-price implementation cycle. Every positive-quote revenue envelope is well defined. Whenever a zero-price value is useful as a diagnostic, it means the limit from positive quotes, not a separate zero-price demand convention.

Definition C.2 (Small-dispersion completion equilibrium). *At $\phi > 0$, an equilibrium consists of a menu $E \subseteq \mathcal{P}_8$, a fee vector $p \in (0, \infty)^E$, and one jointly compatible optimizer selection such that:*

- (i) *every selected portfolio solves (C.15);*
- (ii) *every $T \in E$ has positive selected demand and breaks even,*

$$Q_T(p; E) > 0, \quad p_T Q_T(p; E) = \phi; \quad (\text{C.20})$$

- (iii) *no active product earns more than fixed cost with a lower quote,*

$$\eta \bar{Q}_T(\eta; p_{-T}, E) \leq \phi \quad \text{for every } T \in E \text{ and } 0 < \eta < p_T; \quad (\text{C.21})$$

and

(iv) no omitted product earns more than fixed cost at any positive quote,

$$\sup_{\eta > 0} \eta \overline{Q}_U(\eta; p, E \cup \{U\}) \leq \phi \quad \text{for every } U \in \mathcal{P}_8 \setminus E. \quad (\text{C.22})$$

A positive-demand solution to (C.20) that violates (C.21) or (C.22) is only a break-even branch. A zero-price or zero-demand product is not active when $\phi > 0$.

Primitive reflection exchanges state and investor coordinates $1 \leftrightarrow 3$. It maps

$$L \leftrightarrow R, \quad N_L \leftrightarrow N_R, \quad P_{01} \leftrightarrow P_{30}, \quad H_1 \leftrightarrow H_3,$$

and fixes C, P_{301}, H_0, H_2 . This is a symmetry of the primitives and of Definitions C.1–C.2. It is not an assumption that the menu, prices, or allocations are symmetric.

C.3 Exact demand, all quote intervals, and ties

Lemma C.3 (Finite exact demand calculation). *At every positive fee vector, the optimizer set in (C.15) is nonempty and compact, and its residual exposure is unique. Every extreme optimal implementation uses at most three linearly independent directions, including the home direction if used. For a fixed such support J , let D_J contain its four-asset direction columns and let p_J include a zero home price. Its candidate quantities and weak KKT conditions are*

$$u_J = (D_J' D_J)^{-1} (D_J' g - p_J), \quad u_J \geq 0, \quad d_k'(g - D_J u_J) - p_k \leq 0 \quad (k \notin J). \quad (\text{C.23})$$

For an unused home direction, the last inequality uses $p_k = 0$. Enumerating independent supports and their weak boundaries gives a finite exact description of every optimizer. At a focal quote, the upper tied quantity is the largest focal coordinate among feasible extreme implementations. On each quote interval it is affine, and its revenue is quadratic. Its global supremum is obtained from the interval's endpoint limits and any interior stationary point.

Proof. The zero position is feasible. Completing the square in residual exposure gives $g'r - \|r\|^2/2 \leq \|g\|^2/2 = 4$. On the nonnegative upper-level set, $\|r\| \leq 2\|g\|$ and $\sum_T p_T u_T \leq 4$. Positive prices therefore bound all paid quantities. Since $H_i \neq 0$, the exposure identity then bounds the home quantity. These bounds are uniform on compact sets of positive prices. Continuity gives existence and compactness. If two optima had different residual exposures, a strict convex combination would improve their common value by strict concavity in exposure. Hence exposure is unique; its total fee payment is unique as well.

The KKT conditions are necessary and sufficient because the problem is a concave quadratic over a nonnegative orthant. On its positive coordinates, $p_k = d_k'(g - r)$, including the zero-price home coordinate. Thus the fee row restricted to the optimal face is a linear combination of the three residual-coordinate rows. An extreme implementation consequently has linearly independent direction columns, at most three in the three-dimensional zero-sum space. Solving its KKT equalities

yields (C.23); the displayed primal and inactive dual inequalities complete the test. Zero quantities are retained as weak equalities, so support boundaries are included.

Fix all nonfocal prices. Every candidate coordinate and KKT inequality is affine in the focal quote. Feasible quote sets are intervals with exact endpoints. The maximum focal coordinate over the optimal implementation polytope is attained at an extreme point. Subdivide also at intersections of candidate affine focal quantities. This produces all affine pieces, including changes of implementation on a replication face. On a piece $Q(\eta) = a - b\eta$, revenue is $a\eta - b\eta^2$. Checking the piece's endpoints, one-sided limits, and its stationary point if present therefore checks every positive quote. Moreover, a positive focal holding requires $\eta = d'_T(g - r) \leq 3\|d_T\|\|g\|$. All higher quotes have zero focal demand, so no unbounded high-quote region is left unchecked. The endpoint at zero means the limit from positive quotes; zero-price implementation cycles are never used to define demand. $\square$

Lemma C.4 (Focal demand and incumbent compatibility). *For a fixed positive focal quote η , demand as the quote approaches η from below converges to the upper tied demand in (C.18). In an equilibrium, no alternative optimal implementation can raise the aggregate holding of any active product above its selected holding. In particular, each investor's paid implementation is unique at an equilibrium, although it may lie on an ordinary zero-quantity support boundary.*

Proof. Let q^+ be the greatest focal coordinate among optimizers at η , and choose an optimizer at $\eta - \gamma$. Comparing it with an optimizer attaining q^+ at η shows that its focal holding is at least q^+ . Uniform compactness for $\eta - \gamma \geq \eta/2$ and continuity imply that every limiting optimizer belongs to the optimizer set at η . Its focal holding must therefore equal q^+ . This proves the one-sided limit for each type and hence for population demand.

If an active product has upper tied aggregate demand strictly larger than its selected demand, quotes approaching its current fee from below earn revenue converging to a number strictly above $p_T Q_T = \phi$. Such a quote violates lower-own-quote protection. Finally, two different optimal implementations have the same exposure and total fee payment. If every paid coordinate of an alternative were weakly below the selected coordinates, positive prices would force equality in all of them. The nonzero home direction would then force equality of the home holding too. Thus a genuinely different optimal implementation raises some paid coordinate. Changing only that investor's implementation gives the forbidden aggregate increase. The argument uses one feasible allocation at a time; it never combines incompatible product-specific maxima. $\square$

The menu exhaustion below uses an exact finite calculation rather than a price grid. For clarity, its elimination operations can be described directly. Write an affine condition as $a'x + b \geq 0$, marking whether it is strict. To eliminate a coordinate, retain conditions with zero coefficient and combine every positive- and negative-coefficient pair to cancel that coordinate, multiplying only by positive rational numbers. The combined condition is strict if either parent condition is strict. After the last coordinate, a negative constant, or a zero constant marked strict, proves infeasibility. These are Fourier–Motzkin operations over the rationals. Equal-revenue and quote-vertex substitutions leave univariate polynomial sign tests; factorization, exact root isolation, and signs on the intervening intervals decide them. Replication-price equalities are processed as equalities with feasible common

incumbent allocations, not inferred from neighboring open cells. The displayed matrices, KKT construction, quote-envelope rule, and menu ledger specify the finite proof; the replication archive supplies executable exact generators and sign checks.

C.4 Cutoffs and branch maps

The three scope cutoffs are

$$\phi_W = \frac{25}{22}, \quad (C.24)$$

$$\phi_U = \frac{36}{49}, \quad (C.25)$$

$$\phi_P = \frac{255 - 7\sqrt{621}}{576} = \frac{85 - 7\sqrt{69}}{192}. \quad (C.26)$$

$$(\phi_P, \phi_U, \phi_W) = \left(\frac{255 - 7\sqrt{621}}{576}, \frac{36}{49}, \frac{25}{22} \right). \quad (C.27)$$

They have different competitive meanings: ϕ_W is the wide-viability boundary against the empty completion menu, ϕ_U is the same-product no-undercut boundary of the symmetric pair, and ϕ_P is the pair-entry boundary against an incumbent wide product.

The pair-price correspondence contains four additional events:

$$\phi_D = \frac{2}{3}, \quad \phi_- = \frac{576}{841}, \quad \phi_0 = \frac{25}{36}, \quad \phi_+ = \frac{45}{64}. \quad (C.28)$$

Here ϕ_D is a symmetric-pair clientele boundary, ϕ_- a lower-branch pricing tangency, ϕ_0 a support join and pricing tangency, and ϕ_+ a merger of unequal and symmetric fee vectors. None is a product-scope cutoff. Their exact order is

$$0 < \phi_P < \frac{2}{3} < \phi_- < \phi_0 < \phi_+ < \phi_U < \phi_W. \quad (C.29)$$

The wide branch. For $\phi \in [\phi_P, \phi_W]$, the protected wide fee is

$$z(\phi) = \frac{10 - 2\sqrt{25 - 22\phi}}{33}. \quad (C.30)$$

It is the lower root of

$$\phi = z \left(\frac{15}{2} - \frac{99}{8}z \right), \quad 0 < z \leq \frac{10}{33}. \quad (C.31)$$

The upper root admits a profitable lower same-product quote. At the two scope boundaries,

$$z(\phi_P) = z_P = \frac{27 - \sqrt{621}}{108}, \quad z(\phi_W) = \frac{10}{33}. \quad (C.32)$$

For later use, the full wide-demand curve when it is the only specialist is

$$Q_C(z) = \begin{cases} \frac{15}{2} - \frac{99}{8}z, & 0 \leq z \leq 4/9, \\ 6 - 9z, & 4/9 \leq z \leq 2/3, \\ 0, & z \geq 2/3. \end{cases} \quad (\text{C.33})$$

Its revenue is maximized at $z = 10/33$, with value ϕ_W .

The symmetric pair branch. For every $0 < \phi \leq \phi_U$, the protected symmetric pair fee is

$$p_s(\phi) = \begin{cases} \frac{1 - \sqrt{1 - \frac{3}{2}\phi}}{3}, & 0 < \phi \leq \frac{2}{3}, \\ \frac{1 - \sqrt{1 - \frac{4}{3}\phi}}{2}, & \frac{2}{3} \leq \phi \leq \phi_U. \end{cases} \quad (\text{C.34})$$

The two formulas agree at $p_s(2/3) = 1/3$, and $p_s(\phi_U) = 3/7$. Each pair product has demand

$$Q_P(p) = \begin{cases} 4 - 6p, & 0 \leq p \leq 1/3, \\ 3(1 - p), & 1/3 \leq p \leq 1, \\ 0, & p \geq 1, \end{cases} \quad \mathcal{R}_P(p) = pQ_P(p). \quad (\text{C.35})$$

Thus the two lower-root equations are

$$\phi = p(4 - 6p), \quad p = \frac{1 - \sqrt{1 - \frac{3}{2}\phi}}{3}, \quad 0 < \phi < \frac{2}{3}, \quad (\text{C.36})$$

$$\phi = 3p(1 - p), \quad p = \frac{1 - \sqrt{1 - \frac{4}{3}\phi}}{2}, \quad \frac{2}{3} < \phi < \phi_U. \quad (\text{C.37})$$

C.5 The complete pair-price correspondence

Common fixed cost requires equal active-product revenue. It does not require equal active-product fees. The pair menu contains two reflected unequal-fee branches in addition to the symmetric branch.

Proposition C.5 (Complete unequal-fee pair correspondence). *Orient the pair so that $0 < \ell < r$. For every $\phi \in [\phi_-, \phi_0]$, there is a unique $\ell \in [9/29, 1/3]$ such that*

$$r = r_B(\ell) = \frac{3 + \sqrt{88\ell^2 - 56\ell + 9}}{8} \quad (\text{C.38})$$

and

$$\phi = \ell \left(\frac{7}{2} - \frac{11}{2}\ell + r \right) = r(3 + \ell - 4r). \quad (\text{C.39})$$

Equivalently, the branch has the one-to-one parameterization

$$\ell(s) = \frac{2(s+1)(5-4s)}{3(11-8s^2)}, \quad r(s) = \frac{(5-4s)(8s+11)}{12(11-8s^2)}, \quad (\text{C.40})$$

$$\phi(s) = \frac{(s+1)(5-4s)(9-8s)(8s+11)}{6(11-8s^2)^2}, \quad \frac{1}{8} \leq s \leq \frac{1}{2}. \quad (\text{C.41})$$

The function $\phi(s)$ is strictly increasing, with endpoints

$$(\ell, r, \phi) = \left(\frac{9}{29}, \frac{12}{29}, \frac{576}{841} \right) \quad \text{and} \quad \left(\frac{1}{3}, \frac{5}{12}, \frac{25}{36} \right). \quad (\text{C.42})$$

The aggregate demands on this lower-support segment are

$$Q_L = \frac{7}{2} - \frac{11}{2}\ell + r, \quad Q_R = 3 + \ell - 4r. \quad (\text{C.43})$$

On the upper-support segment

$$\phi_0 < \phi < \phi_+, \quad (\text{C.44})$$

the fees are

$$p_h = \frac{3}{8} + \frac{\sqrt{225-320\phi}}{40}, \quad p_\ell = \frac{3}{8} - \frac{\sqrt{225-320\phi}}{40}. \quad (\text{C.45})$$

Writing $(\ell, r) = (p_\ell, p_h)$, they satisfy

$$\ell + r = \frac{3}{4}, \quad \phi = 5\ell r, \quad \frac{1}{3} < \ell < \frac{3}{8}. \quad (\text{C.46})$$

Before imposing the off-symmetric break-even identity, the demands on this support cell are

$$Q_L = 3 - 4\ell + r, \quad Q_R = 3 + \ell - 4r. \quad (\text{C.47})$$

On the branch they reduce to

$$Q_L = 5r, \quad Q_R = 5\ell. \quad (\text{C.48})$$

The lower and upper formulas meet at $(\ell, r, \phi) = (1/3, 5/12, \phi_0)$. As $\phi \uparrow \phi_+$, both fees converge to $3/8$; the two reflected unequal branches then merge into the already existing symmetric fee vector.

Primitive reflection supplies the second orientation (r, ℓ) . Consequently the pair scope has two distinct reflected unequal-fee vectors exactly for

$$\phi_- \leq \phi < \phi_+. \quad (\text{C.49})$$

At ϕ_- the unequal equilibria are weak at the expensive product's own-price tangency. At ϕ_+ there is no additional unequal vector.

Corollary C.6 (Fee dispersion within a fixed scope). *On $[\phi_-, \phi_+)$, the pair menu supports one symmetric fee vector and two reflected unequal-fee vectors. Every active product nevertheless earns the same revenue ϕ . Thus common operating cost equalizes revenue, not price or quantity.*

C.6 The exact correspondence

Proposition C.7 (Exact equilibrium correspondence for the normalized problem). *This is the complete formal statement underlying Proposition 2. Under Definitions C.1 and C.2, the full set of equilibrium menu–fee pairs is exactly the following union:*

$$(\emptyset, \emptyset) \iff \phi \geq \phi_W, \quad (\text{C.50})$$

$$(\{C\}, z(\phi)) \iff \phi_P \leq \phi \leq \phi_W, \quad (\text{C.51})$$

$$(\{L, R\}, (p_s(\phi), p_s(\phi))) \iff 0 < \phi \leq \phi_U. \quad (\text{C.52})$$

In addition to the symmetric pair fee, the two reflected unequal vectors

$$(\{L, R\}, (\ell(\phi), r(\phi))) \quad \text{and} \quad (\{L, R\}, (r(\phi), \ell(\phi)))$$

exist exactly when $\phi_- \leq \phi < \phi_+$. Use (C.38)–(C.41) on $[\phi_-, \phi_0]$ and (C.45) on (ϕ_0, ϕ_+) . There are no other active menus and no other equilibrium fee vectors.

Equivalently, after suppressing price multiplicity, the exact scope correspondence is

ϕ region	equilibrium completion scopes	
$0 < \phi < \phi_P$	$\{L, R\}$	(C.53)
$\phi = \phi_P$	$\{C\}, \{L, R\}$	
$\phi_P < \phi < \phi_U$	$\{C\}, \{L, R\}$	
$\phi = \phi_U$	$\{C\}, \{L, R\}$	
$\phi_U < \phi < \phi_W$	$\{C\}$	
$\phi = \phi_W$	$\{C\}, \emptyset$	
$\phi > \phi_W$	$\emptyset$.	

C.7 Proof of Proposition C.7

The proof has two finite steps. First, investor optimality reduces the eight-product universe to five products. Second, exact break-even, lower-quote, and entry inequalities reduce the remaining menu lattice to the three scopes in the proposition. All comparisons use rational arithmetic and algebraic-number sign tests. Replication faces are checked separately with the compatible-incumbent and focal-upper conventions above.

Lemma C.8 (Support reduction). *No positive-fee investor optimum can assign positive quantity to P_{01} , P_{30} , or P_{301} . Hence every exact equilibrium menu is a subset of*

$$\mathcal{P}_5 = \{L, R, C, N_L, N_R\}. \quad (\text{C.54})$$

Proof. For each investor, fix a positive fee vector and a weak KKT support. The objective is strictly concave in residual exposure, so its optimizing exposure is unique. If a tied implementation assigned a positive quantity to one of the three products in the statement, maximizing that coordinate over the polytope of optimal implementations would produce an extreme optimal implementation with the same coordinate positive.

For clarity, the support-elimination system can be written without any unknown inactive-product fees. For each independent set of paid columns J , optionally including home, let $r = D_J u$. Require

$$u_j > 0 \ (j \in J), \quad d'_T(g - r) > 0 \ (T \in J \text{ paid}), \quad \begin{cases} H'_i(g - r) = 0, & \text{home included,} \\ H'_i(g - r) \leq 0, & \text{home excluded.} \end{cases}$$

The strict paid margin equals its positive fee. No assumption on an unused fee is imposed, so infeasibility of this smaller system rules out the support in every possible menu. These are rational affine inequalities in at most three quantities. Enumerate all independent paid subsets, with or without home, of total size at most three. Exact Fourier–Motzkin elimination of all 237 such systems that contain one of P_{01}, P_{30}, P_{301} makes its positive quantity incompatible with the remaining primal and dual inequalities. This closes both generic supports and dependent optimal faces. Removing those products leaves, for investor types 0, 1, 2, 3, respectively,

$$17, \quad 17, \quad 7, \quad 17$$

weak KKT cells over $\mathcal{P}_5$. These finite lists include all support equalities. Therefore an excluded product cannot be positive in any optimal implementation. $\square$

Lemma C.9 (Exhaustion of the five-product menu lattice). *If a menu $E \subseteq \mathcal{P}_5$ satisfies Definitions C.1–C.2, then*

$$E \in \{\emptyset, \{C\}, \{L, R\}\}. \quad (\text{C.55})$$

Proof. We first give the analytic exclusions and then the finite certificate for the remaining menu faces. Fees of L, R, C, N_L, N_R are written ℓ, r, z, a, b , respectively. A necessary local protection condition on a differentiable incumbent cell is $\partial(p_T Q_T)/\partial p_T \geq 0$: a negative derivative makes a sufficiently small price reduction profitable. This necessary condition is used for exclusion only. Sufficiency for surviving menus is proved below by the entire lower-quote envelope.

One pair cannot survive alone. Write its price as p . The home-inclusive KKT calculation gives

$$Q_L(p; \{L\}) = \begin{cases} 4 - 5p, & 0 < p \leq 1/3, \\ 7(1 - p)/2, & 1/3 \leq p \leq 1, \\ 0, & p \geq 1. \end{cases} \quad (\text{C.56})$$

A protected positive-demand fee must satisfy $p \leq 1/2$. For $0 < p \leq 1/3$, a pooled entrant quoting just below $2p/3$ has revenue approaching $4p - 9p^2/2$. Its gain over the incumbent's revenue is

$p^2/2 > 0$. For $1/3 < p \leq 1/2$, use the entrant cell

$$Q_C(z; \{L, C\}) = \frac{3}{8}(12p - 45z + 16), \quad z^* = \frac{2p}{15} + \frac{8}{45}.$$

Its KKT inequalities hold throughout this range. For instance, $p - 1/3 < z^* < 2p/3$, $z^* < (1 + p)/3$, $z^* < 4/9$, and $z^* < 4p/3$, which are its positivity and inactive-coordinate conditions. At the vertex,

$$z^* Q_C(z^*) - \frac{7}{2}p(1 - p) = \frac{114p^2 - 81p + 16}{30} > 0;$$

the numerator has positive leading coefficient and discriminant -735 . Thus $\{L\}$ is excluded, as is its reflection.

A pooled product and one pair cannot coexist. For $\{L, C\}$, set the pair price to p . Solving the independent investor supports yields exactly the five open positive-demand cells below. The support column lists the common support of types 0 and 1, followed by those of types 2 and 3; the home letter always denotes the respective investor's home direction.

Cell	Supports	Q_L	Q_C
I	(LC, H, LC)	$3(1 - 3p + 3z)/2$	$9(1 + p - 3z)/2$
II	(LC, H, LH)	$(5 - 9p + 6z)/2$	$3(1 + p - 3z)$
III	(LC, CH, LC)	$3(1 - 3p + 3z)/2$	$3(16 + 12p - 45z)/8$
IV	(LC, CH, LH)	$(5 - 9p + 6z)/2$	$3(12 + 8p - 33z)/8$
V	(LC, LCH, LH)	$(5 - 13p + 9z)/2$	$9(1 + p - 3z)/2$

These cells can be reproduced without guessing supports. Types 0 and 1 have possible paid supports L, C, LC , with the LC quantities $(2(1 - 3p + 3z), 6(1 + p - 3z))$. Type 2 has supports H, LH, CH, LCH , and type 3 has C, LC, LH . Imposing positivity and the inactive KKT inequalities in (C.23) on their Cartesian product, then requiring both aggregate demands positive, leaves the five rows above. The equal-revenue polynomials, up to a positive factor, are

$$\begin{array}{l|l} I & -3p^2 + p + 9z^2 - 3z \\ III & -12p^2 + 4p + 45z^2 - 16z \\ IV & -36p^2 + 20p + 99z^2 - 36z \\ V & -13p^2 + 5p + 27z^2 - 9z. \end{array}$$

In I, feasibility requires $z \geq 4/9$ and $p \geq 3z/2$. The first polynomial is decreasing in p there and is at most $3z(3z - 2)/4 < 0$. In II, simultaneous nonnegative own-revenue derivatives require $6z - 1 \leq p \leq (6z + 5)/18$, implying $z \leq 23/102 < 4/9$, contrary to its support.

In III, local protection requires $z \geq 2p - 1/3$, while the support requires $z \leq 2p/3$; hence $p \leq 1/4$. If $p \geq 1/5$, its equal-revenue polynomial is convex in z and negative at both endpoints $2p - 1/3$ and $2p/3$ for $1/5 \leq p \leq 1/4$. If $p < 1/5$, the missing pair R can quote $(1 - p + 3z)/4$. Direct KKT

substitution gives its revenue $3(1 - p + 3z)^2/8$, with strict gain

$$\frac{3}{8} \left\{ 13 \left(p - \frac{3(3z+1)}{13} \right)^2 + \frac{4}{13} (3z+1)^2 \right\} > 0.$$

In IV, convexity in z again reduces the comparison to support and local-protection endpoints. Substitution of $2p/3, 4p/3, (4p+6)/33$ gives, respectively,

$$4p(2p-1), \quad 28p(5p-1), \quad -\frac{4}{11}(95p^2 - 55p + 9),$$

which are negative whenever the associated endpoint is feasible. The last quadratic has discriminant -395 . In V, feasibility and local C protection give $4p/3 \leq z \leq (1+p)/6$, hence $p \leq 1/7$. The endpoint values of its convex equal-revenue polynomial are $7p(5p-1) < 0$ and $-(49p^2 - 20p + 3)/4 < 0$; the last discriminant is -188 . Ordinary weak support faces satisfy the same closed inequalities and the strict contradictions above. The sole positive-price replication circuit is $H_3 + 2L = 3C$, at $3z = 2p$. On its positive-quantity portion, a different optimal implementation raises one paid coordinate, violating Lemma C.4. Portions without a nontrivial implementation segment are the ordinary zero-quantity limits already covered. Reflection excludes $\{R, C\}$.

Elementary neutral exclusions. At the home-only optima, every neutral direction has nonpositive marginal value. For types 0, 1, and 3 the exposure is zero and the neutral value is zero. For type 2 the exposure is $8H_2/3$, giving neutral margin $-2/9$. Positive prices therefore make every neutral-only menu inactive. The support catalogue in Lemma C.8 further implies that type 2 never buys a neutral; a basis containing N_L contains C or R , and a basis containing N_R contains C or L . This immediately excludes $\{L, N_L\}$, $\{L, N_L, N_R\}$, and their reflections.

For the cross-neutral menu $\{L, N_R\}$, positive neutral demand requires $b < (1-p)/3$, and the aggregate demands are

$$Q_{N_R} = \frac{9}{2}(1-p-3b), \quad Q_L = \begin{cases} (11-13p-9b)/2, & p < 1/3, \\ 5-5p-9b/2, & p \geq 1/3. \end{cases}$$

No equal-revenue point has $p \geq 1/3$: writing $s = 1-p$ and $t = b/s \in (0, 1/3)$, equality would require $p(5-9t/2) = st(9/2-27t/2)$. The left side exceeds $7/6$, whereas the right side is at most $3/8$. In the lower-price cell, equal revenue requires

$$27b^2 - 9b + 11p - 13p^2 = 0.$$

Its discriminant restricts $p \leq (11 - \sqrt{82})/26 < 1/7$. At these prices, entry of C at $z = (1+p)/6$ satisfies all KKT inequalities and earns $3(1+p)^2/8 > 3/8$. Incumbent neutral revenue is strictly less than $3/8$, since

$$bQ_{N_R} < b(9/2 - 27b/2) = 3/8 - \frac{27}{2}(b - 1/6)^2.$$

This excludes the cross-neutral menu and its reflection.

For $\{C, N_L\}$ or $\{C, N_L, N_R\}$, the KKT solutions show that every N_L buyer holds at least as much C . For the two-product basis the difference is $u_C - u_{N_L} = 6 + 9a - 9z > 0$, using the positivity condition $3a + z < 2/3$. For the three-product basis the same difference and sign follow from $2a + b + z < 2/3$. Home-inclusive alternatives either have a negative neutral coordinate or violate home positivity. Thus the same-investor package $d_{N_L} + d_C = 2d_L/3$ is available for every unit of neutral demand. At the entry quote $\eta_L = 3(z + a)/2$, replacing this package by $2/3$ units of L preserves exposure and total fees. Upper tied L demand is at least $2Q_{N_L}/3$; its limiting revenue from a lower quote is at least

$$(z + a)Q_{N_L} = \phi + zQ_{N_L} > \phi.$$

Reflection gives the identical argument if only N_R is active. On $\{L, C, N_L\}$, the same identity gives a price trichotomy: if $p > 3(z + a)/2$, L is dominated by the package; if $p < 3(z + a)/2$, neutral holders can lower fees by replacing their available package with L ; at equality the nontrivial implementation tie contradicts Lemma C.4. This excludes the aligned dependent menu and its reflection.

The remaining finite faces. The remaining representatives require more support cells; their finite proof uses exactly the same matrix calculation, with no additional assumption on prices or investor choices. For each menu and each type, form all independent columns in (C.23), retain the affine KKT inequalities, and combine the four type supports. Require strictly positive aggregate demand for every present product. On an ordinary cell impose equal revenues and the necessary local protection inequalities. On a replication face first retain one common incumbent allocation, then compare its feasible focal-upper implementations using Lemma C.4. This order prevents an exclusion based on incompatible selected quantities.

The resulting exact certificate has the following terminal systems. Here “weak faces” includes the intersections of zero-quantity and KKT equalities, and a row with a reflected menu uses the primitive reflection already established above.

Representative	Exhaustion and terminal contradiction
$\{L, R, C\}$	4,536 support tuples reduce to seven locally protected affine systems. Equal-revenue substitution excludes each by a strict revenue difference or a one-product quote. The boundary calculation includes 228 oriented wall records (221 after identifying a wall with its negative), all wall intersections, and the three positive-price replication circuits.
$\{L, R, N_L\}$	Six aggregate cells with weak faces. Cells 1–3 have unequal L and N_L revenues even at upper tied neutral demand. In cells 4–5, equality of these revenues implies a profitable lower L quote. In cell 6, the two equal-revenue equations contradict a strictly positive polynomial identity on its feasible simplex.
$\{L, C, N_R\}$	Five aggregate cells with weak faces. The first two contradict equal incumbent revenues; the third and fourth admit omitted R entry or an own-product undercut; the fifth contradicts simultaneous local C and N_R protection. The positive-price incumbent replication face is tested with upper tied R demand.
$\{L, R, N_L, N_R\}$	Among 432 raw joint profiles, none has both neutrals active: their KKT conditions require simultaneously $\ell - r \geq 3a/2$ and $r - \ell \geq 3b/2$.
$\{L, C, N_L, N_R\}$	The weak support covers force $2\ell = 3(z + a)$. Converting a coheld (C, N_L) package to $2L/3$ preserves one investor’s exposure and payment and raises upper tied L demand, violating own-quote protection.
$\{L, R, C, N_L\}$	Seven aggregate cells, including weak faces, obey the trichotomy $2\ell \geq 3(z + a)$ or equality. The strict sides contradict positive demand, equal revenue, or a lower own quote. On equality the same-investor package conversion violates own-quote protection.
$\{L, R, C, N_L, N_R\}$	All 18 minimal support covers force both $2\ell = 3(z + a)$ and $2r = 3(z + b)$. Separate same-investor package conversions raise upper tied L or R demand. One such conversion suffices for a profitable lower quote.

The finite tests in this table are exact, not numerical searches. Every cell is constructed from the displayed rational directions. Affine feasibility uses the strict/weak Fourier–Motzkin procedure stated above; a revenue vertex is substituted only with its interval inequalities. The remaining polynomial inequalities are certified after equal-revenue substitution by factored signs, root isolation, or positive-coefficient representations in nonnegative cell slacks. At a genuine implementation tie the focal-upper conversion is checked before imposing local single-cell derivatives. The accompanying executable certificate generates the support lists and checks the substituted polynomials and sign identities. Thus the large rows of the table are a reproducible finite part of the proof, rather than an assumption that the generic cells extend to their boundaries.

For an explicit completeness check, the following ledger lists every subset of $\mathcal{P}_5$. “Pair” and “pooled” refer to the branches solved below; the named exclusions refer to the calculations just given. Mirrored entries use primitive reflection, not equal fees.

Menu	Disposition
$\emptyset$	empty branch
$\{C\}$	pooled branch
$\{L\}, \{R\}$	single-pair entry exclusion
$\{N_L\}, \{N_R\}, \{N_L, N_R\}$	neutral-only inactivity
$\{C, L\}, \{C, R\}$	five mixed cells and their replication face
$\{C, N_L\}, \{C, N_R\}$	package entry
$\{L, R\}$	complete pair branch
$\{L, N_L\}, \{R, N_R\}$	neutral support restriction
$\{L, N_R\}, \{R, N_L\}$	cross-neutral entry exclusion
$\{C, L, R\}$	three-product exact certificate
$\{C, L, N_L\}, \{C, R, N_R\}$	aligned package trichotomy
$\{C, L, N_R\}, \{C, R, N_L\}$	five-cell neutral certificate
$\{C, N_L, N_R\}$	package entry
$\{L, R, N_L\}, \{L, R, N_R\}$	six-cell neutral certificate
$\{L, N_L, N_R\}, \{R, N_L, N_R\}$	neutral support restriction
$\{C, L, R, N_L\}, \{C, L, R, N_R\}$	seven-cell package trichotomy
$\{C, L, N_L, N_R\}, \{C, R, N_L, N_R\}$	forced package tie
$\{L, R, N_L, N_R\}$	incompatible neutral fee inequalities
$\{C, L, R, N_L, N_R\}$	both package ties

The grouped rows contain, respectively, 1 empty menu, 5 singleton menus, 10 two-product menus, 10 three-product menus, 5 four-product menus, and the full menu. Hence all 32 menus have been considered, and the only possible scopes are $\emptyset, \{C\}, \{L, R\}$. $\square$

Completion of the proof of Proposition C.7. We solve and protect the three remaining menus. The calculations below also prove Proposition C.5.

Empty menu and the pooled product. With no specialist, the home optimum is zero for types 0, 1, and 3 and $h_2 = 8/3$ for type 2. Introducing C at fee z gives, for $0 < z \leq 4/9$,

$$u_{0C} = u_{1C} = u_{3C} = 8 - 12z, \quad u_{2C} = 6 - \frac{27}{2}z, \quad h_2 = 2 + \frac{3}{2}z.$$

The other home quantities are zero. Their unused-home KKT margins are $-2 + u_C/4 = -3z$ for type 0 and $-u_C/12 \leq 0$ for types 1 and 3. After these types leave C , the corresponding margins are $-2, 0, 0$. At $z = 4/9$, type 2 leaves the product; above this fee her home holding is $8/3$. The other types leave at $z = 2/3$. For type 2 on the home-only corner, the unused-product margin is $4/9 - z \leq 0$. Averaging yields (C.33). The first revenue piece reaches its global maximum at $10/33$, before the first clientele change, and the later piece is decreasing. Its maximum is therefore $25/22$. A pair entering alone has the demand curve (C.56) below and maximum revenue $7/8$. A neutral product alone has zero demand. Support reduction excludes the other three entrants. Thus the empty menu is protected if and only if $\phi \geq \phi_W$.

On a C -only menu the same demand calculation applies to lower own quotes. Protected break-even must lie on the increasing revenue branch, $0 < z \leq 10/33$; this gives (C.30). Any higher break-even fee admits a lower quote with greater revenue. Its own-revenue slope is $15/2 - 99z/4$, positive below the endpoint and zero at it.

To check all other products, apply (C.23) to a pair entering at fee η , with z fixed. The affine demand pieces and their quadratic revenue vertices give

$$\Pi_P(z) = \begin{cases} (1 + 3z)^2/8, & 0 < z \leq z_s, \\ 15z/4 - 45z^2/8, & z_s \leq z \leq 5/21, \\ (5 + 6z)^2/72, & 5/21 \leq z \leq 10/33, \end{cases} \quad z_s = \frac{2}{9} - \frac{\sqrt{10}}{18}. \quad (\text{C.57})$$

For example, in the first region the binding entrant cell has $Q_L(\eta; z) = \frac{3}{2}(1 - 3\eta + 3z)$: its revenue vertex is $\eta = (1 + 3z)/6$. In the middle region the binding quote is the replication boundary $\eta = 3z/2$, evaluated using its upper tied demand. In the last region the binding cell has $Q_L(\eta; z) = (5 - 9\eta + 6z)/2$, whose vertex is $(5 + 6z)/18$. Comparing all cell endpoints and vertices gives the two switches in (C.57). Reflection gives the same envelope for R .

Subtracting the first piece from incumbent revenue gives its first and only crossing in that region at $z_P = (27 - \sqrt{621})/108$. On the later pieces the gaps are, respectively,

$$\frac{3}{4}z(5 - 9z) > 0, \quad -\frac{927z^2 - 480z + 25}{72} > 0.$$

Each neutral entrant has peak revenue $(2 - 3z)^2/32$. Its gap below the three successive pair-envelope pieces is

$$\frac{9z(3z + 4)}{32}, \quad -\frac{(3z - 2)(63z - 2)}{32}, \quad \frac{(3z + 16)(21z + 4)}{288},$$

all positive on the relevant intervals. The other three products have zero demand by Lemma C.8. These are all-price comparisons, including the focal-upper value on the replication boundary. The pooled menu is consequently protected exactly on $[\phi_P, \phi_W]$.

The complete pair-price plane. Orient prices as $0 < \ell \leq r$. Positive R demand requires $r < 1$. In every cell below, types 0 and 1 hold $u_L = 4(1 - \ell)$, $u_R = 4(1 - r)$, and no home asset. The remaining types have the following possible portfolios, with coordinates ordered as (u_L, u_R, h_i) :

$$\begin{aligned} i = 2 : & \quad (4(1 - 2\ell - r), 4(1 - \ell - 2r), 4(\ell + r)), \\ & \quad (2(1 - 3\ell), 0, 2(1 + \ell)), \quad (0, 0, 8/3); \\ i = 3 : & \quad (4(1 - 2\ell + r), 4(1 + \ell - 2r), 4(r - \ell)), \\ & \quad (6(1 - \ell), 0, 2(1 - \ell)). \end{aligned}$$

The sign of the vanishing quantity and the inactive KKT inequality select the following exhaustive five-cell partition:

Cell	Interior conditions	Q_L	Q_R
A	$\ell < 1/3, r < (1 - \ell)/2$	$4 - 6\ell$	$4 - 6r$
B	$\ell < 1/3, (1 - \ell)/2 < r < (1 + \ell)/2$	$7/2 - 11\ell/2 + r$	$3 + \ell - 4r$
C	$\ell < 1/3, r > (1 + \ell)/2$	$4 - 5\ell$	$2(1 - r)$
D	$\ell > 1/3, r < (1 + \ell)/2$	$3 - 4\ell + r$	$3 + \ell - 4r$
E	$\ell > 1/3, r > (1 + \ell)/2$	$7(1 - \ell)/2$	$2(1 - r)$

The common quantities on ordinary boundaries are obtained from either adjacent formula. The columns d_L, d_R, H_i are independent for every type, so there is no positive-quantity implementation tie on this price plane. This fact does not remove the need to check one-sided revenue derivatives on the cell boundaries.

In cells C and E, $r > 1/2$, and the expensive product has own-revenue derivative $2 - 4r < 0$, ruling out protection. On the boundary $r = (1 + \ell)/2$, the derivative from lower r , computed from B or D, is $-1 - 3\ell < 0$; the boundary is excluded too. In A,

$$\ell Q_L - r Q_R = (\ell - r)\{4 - 6(\ell + r)\}.$$

The unequal root $\ell + r = 2/3$ violates the cell inequalities for $\ell < r$. In D the factorization is

$$\ell Q_L - r Q_R = (\ell - r)\{3 - 4(\ell + r)\}.$$

Thus unequal protected candidates can lie only on the B branch in (C.38) or the D branch with $\ell + r = 3/4$. The B equal-revenue equation is

$$-11\ell^2 + 7\ell + 8r^2 - 6r = 0.$$

Its root satisfying the B inequalities is precisely $r_B(\ell)$. The faces $\ell = 1/3$ and $r = (1 - \ell)/2$ give their adjoining limits; substitution produces no additional unequal branch. The diagonal is treated next.

Symmetric prices: every lower quote. On the diagonal, averaging the portfolios above gives (C.35). If $0 < p \leq 1/3$, holding the other price at p and lowering the focal quote gives simply $Q(\eta; p) = 4 - 6\eta$ for $0 < \eta \leq p$. Its revenue is increasing on this interval. For $1/3 \leq p < 1/2$, the complete lower-quote demand is

$$Q(\eta; p) = \begin{cases} 4 - 6\eta, & 0 < \eta \leq 1 - 2p, \\ 7/2 + p - 11\eta/2, & 1 - 2p \leq \eta \leq 1/3, \\ 3 + p - 4\eta, & 1/3 \leq \eta \leq p. \end{cases} \quad (\text{C.58})$$

The first two revenues are increasing throughout their displayed intervals for $p \leq 3/7$. On the third interval the final derivative is $3 - 7p$, so all lower quotes are protected precisely up to $p = 3/7$. If

$3/7 < p < 1/2$, the admissible quote $\eta^* = (3 + p)/8$ yields

$$\eta^* Q(\eta^*; p) - 3p(1 - p) = \frac{(7p - 3)^2}{16} > 0. \quad (\text{C.59})$$

For $1/2 \leq p < 1$, a sufficiently small own-price reduction already has a strict gain, since the local lower-quote derivative remains $3 - 7p < 0$. Higher prices have no demand. This proves both necessity and sufficiency of the protected range and selects exactly the roots in (C.34).

The omitted pooled product has the all-quote envelope

$$\Pi_C(p) = \begin{cases} 2p - 6p^2, & 0 < p \leq 1/6, \\ (12p + 1)^2/54, & 1/6 \leq p \leq 3/7. \end{cases}$$

On the first range its gap below incumbent revenue is $2p > 0$. On $[1/6, 1/3]$, the gap has numerator $-468p^2 + 192p - 1$ over 54; this concave quadratic is positive at both endpoints. On $[1/3, 3/7]$, the gap has numerator $-306p^2 + 138p - 1$ over 54; it is again positive at both endpoints. At the upper endpoint the gap is $95/2646$. The other five omitted products have no positive-price demand: their zero-quote KKT margins at the incumbent optimum are nonpositive, and adding a positive fee makes them strictly negative. Thus pair entry is not the constraint that sets ϕ_U ; its own lower-quote comparison is.

Unequal prices: every lower quote. For the B branch, the cheap product's lower-quote curve is

$$Q_L(\eta; r) = \begin{cases} 4 - 6\eta, & 0 < \eta \leq 1 - 2r, \\ 7/2 + r - 11\eta/2, & 1 - 2r \leq \eta \leq \ell, \end{cases}$$

whereas the expensive product's curve is

$$Q_R(\eta; \ell) = \begin{cases} 4 - 6\eta, & 0 < \eta \leq (1 - \ell)/2, \\ 3 + \ell - 4\eta, & (1 - \ell)/2 \leq \eta \leq r. \end{cases}$$

The cheap product's revenues increase throughout these intervals on the candidate protected branch. The expensive product's first piece has maximum at most $2/3$. Its final piece is increasing up to its posted quote if and only if $r \leq (3 + \ell)/8$. Substitution into the B equal-revenue equation gives the boundary factors $(3\ell - 1)(29\ell - 9)$. Together with the B inequalities this restricts ℓ to $[9/29, 1/3]$. Its common revenue is at least $\phi_- > 2/3$, so the earlier piece cannot defeat it. This proves global lower-quote protection, including the two endpoint tangencies. Away from these endpoints the expensive product's final slope $3 + \ell - 8r$ is strictly positive; the cheap product's slope $7/2 + r - 11\ell$ is strictly positive throughout.

The parameterization in (C.40) follows by rationally parameterizing the B equal-revenue conic. Direct substitution verifies all its equations and endpoints. Its revenue derivative is

$$\phi'(s) = \frac{448s^3 - 456s^2 + 264s - 209}{6(8s^2 - 11)^3} > 0 \quad (1/8 \leq s \leq 1/2).$$

For example, the numerator is bounded above on this interval by $56 + 132 - 209 < 0$, and the denominator is negative. Hence every cost between the stated endpoint costs has exactly one oriented B price vector.

For the D branch, let $r = 3/4 - \ell$, $1/3 < \ell < 3/8$. The cheap product's curve consists of $4 - 6\eta$, $7/2 + r - 11\eta/2$, and $3 + r - 4\eta$, on the intervals separated by $1 - 2r$ and $1/3$, up to ℓ . For the expensive product it consists of $4 - 6\eta$, $7/2 + \ell - 11\eta/2$, and $3 + \ell - 4\eta$, with breaks $1 - 2\ell$ and $1/3$, up to r . All these revenue pieces are increasing on their respective intervals. Their final slopes are $15/4 - 9\ell > 0$ and $9\ell - 3 > 0$. Thus every lower own quote is protected. Equal revenue is $\phi = 5\ell(3/4 - \ell)$, strictly increasing for $\ell < 3/8$. Solving gives (C.45); the limit at $\ell = 1/3$ joins B, while $\ell = 3/8$ is the symmetric merger. No separate unequal vector exists at that merger.

Unequal prices: every entrant quote. The following explicit envelopes verify entry protection, including its implementation boundaries. In the lower B range define

$$(a_1, a_2, a_3, a_4, a_5) = \left(\frac{\ell + 2r - 1}{3}, \frac{2\ell}{3}, \frac{\ell + r}{3}, \frac{4\ell}{3}, \frac{1 + \ell}{3} \right).$$

They are in increasing order, allowing equality at the joining endpoint. The pooled entrant's demand on the five successive open intervals $(0, a_1), (a_1, a_2), \dots, (a_4, a_5)$ is, respectively,

$$\frac{3}{8}(16 + 12\ell - 45z), \quad \frac{3}{8}\{4 + 24(\ell + r) - 81z\}, \quad \frac{3}{8}\{4 + 16(\ell + r) - 57z\}, \quad \frac{3}{8}(4 - 9z), \quad \frac{3}{2}(1 + \ell - 3z). \quad (\text{C.60})$$

It is zero above a_5 . At each endpoint its focal-upper value is the limit from the left, by Lemma C.4; this matters at a_2 , where the positive demand jump must not be replaced by the right-hand limit. Checking the five quadratic pieces and their endpoints gives

$$z^* = \frac{2 + 12(\ell + r)}{81}, \quad \Pi_C = \frac{\{1 + 6(\ell + r)\}^2}{54} \leq \frac{121}{216} < \phi_-.$$

Here z^* belongs to the second piece, and $\ell + r \leq 3/4$.

On the upper D range use

$$(b_1, b_2, b_3, b_4, b_5) = \left(\ell - \frac{1}{3}, \frac{1 - 2\ell}{6}, \frac{2\ell}{3}, \frac{1}{4}, \frac{4}{9} \right).$$

The pooled demand on the corresponding five intervals is

$$\frac{15}{2} - \frac{99}{8}z, \quad \frac{3}{8}(16 + 12\ell - 45z), \quad \frac{3}{8}(22 - 81z), \quad \frac{3}{8}(16 - 57z), \quad \frac{3}{8}(4 - 9z). \quad (\text{C.61})$$

Demand is zero above b_5 ; left-hand values apply on the boundaries, particularly the replication jump at b_3 . The global vertex is $z = 11/81$, and

$$\Pi_C = 121/216 < \phi_0. \quad (\text{C.62})$$

Only one neutral entrant, N_R , has positive revenue under this orientation. With $d = r - \ell > 0$, its demand is

$$\bar{Q}_{N_R}(z) = \begin{cases} \frac{3}{2}(1 - 5\ell + 4r - 15z), & 0 < z \leq d/3, \\ \frac{3}{2}(1 - \ell - 3z), & d/3 < z \leq 2d/3, \\ 0, & z > 2d/3. \end{cases}$$

The upper endpoint $2d/3$ retains its positive upper tied demand. The maximal revenue is consequently

$$\Pi_{N_R} = (r - \ell)(1 + \ell - 2r), \quad \Pi_{N_R} < 3/29 < \phi_- \text{ on B}, \quad \Pi_{N_R} \leq 1/24 < \phi_0 \text{ on D.} \quad (\text{C.63})$$

For completeness, the zero-quote marginal values of $(P_{01}, P_{30}, N_L, P_{301})$ at the pair incumbent are

$$\begin{array}{c|cccc} i = 0, 1 & -r & -\ell & (\ell - r)/3 & -(\ell + r)/3 \\ i = 2 \text{ on B} & (\ell - 1)/2 & -\ell & (\ell - 1)/3 & 0 \\ i = 2 \text{ on D} & -1/3 & -1/3 & -2/9 & 0 \\ i = 3 & -r & -\ell & 0 & -2r/3 \end{array}$$

They are all nonpositive; a positive entrant fee makes each corresponding KKT inequality strict. These products therefore have zero demand at every positive quote. Reflection covers the other orientation without requiring equal fees or compatible maxima across distinct products.

The all-quote calculations prove sufficiency of every displayed branch. The five-cell partition, its weak boundaries, equal revenue, and own-quote necessity prove that the pair menu has no other fee vectors. Together with Lemmas C.8 and C.9, this establishes the complete correspondence. $\square$

Lemma C.10 (Strict deviation gaps on interior branches). *Let K be a compact subset of one of the following cost intervals, following the indicated exact menu and fee branch:*

<i>branch</i>	<i>interval</i>
$\emptyset$	(ϕ_W, ∞)
C	(ϕ_P, ϕ_W)
<i>symmetric pair</i>	$(0, \phi_U)$
<i>lower unequal pair</i>	(ϕ_-, ϕ_0)
<i>upper unequal pair</i>	(ϕ_0, ϕ_+) .

Every omitted product's maximal positive-quote revenue lies uniformly strictly below ϕ on K . For each $\epsilon > 0$, all own quotes at least ϵ below the incumbent fee also earn uniformly strictly less than ϕ , whenever such quotes exist. On each fixed incumbent support cell, the derivative of own revenue with respect to its quote, evaluated from below at the incumbent quote, is strictly positive in these intervals. The derivative assertion excludes the symmetric support boundary $\phi = 2/3$.

Proof. For the empty menu, its greatest entry revenue is ϕ_W , strictly below every $\phi \in K$. For the pooled and pair menus, the explicit entry envelopes in the preceding proof have strict gaps on the stated intervals. Their maxima are obtained over finitely many quadratic pieces and one-sided endpoints; the inequalities include focal-upper values at every tied implementation. The displayed

own-quote curves increase to incumbent revenue, with the earlier expensive-product B piece bounded by $2/3 < \phi_-$. Hence their lower-quote inequalities are strict at every quote below the incumbent fee, not only locally. Their incumbent slopes, computed above, are respectively

C	$15/2 - 99z/4$
symmetric pair, $p \leq 1/3$	$4 - 12p$
symmetric pair, $p > 1/3$	$3 - 7p$
lower unequal pair	$7/2 + r - 11\ell, \quad 3 + \ell - 8r$
upper unequal pair	$15/4 - 9\ell, \quad 9\ell - 3.$

At the symmetric support boundary $p = 1/3$, the lower-quote slope is zero, rather than strictly positive; this single point is therefore excepted from the slope assertion. The revenue still increases strictly up to its endpoint there. All remaining slopes are positive on the stated strict branch interiors.

To make pointwise gaps uniform, the compact cost set gives bounded incumbent fees with positive minima for every present product. The finite affine-support formulas in Lemma C.3 give a common bound for focal quantities as the quote tends to zero from above; although the exactly zero-price problem may have cycles, none is used in these positive-quote formulas. Low-quote revenue consequently tends uniformly to zero. The high-quote bound in that lemma makes all sufficiently high quotes irrelevant. On the remaining compact positive quote set, upper focal demand is upper semicontinuous: any maximizing sequence has a convergent subsequence of portfolios, whose limit is an optimal portfolio at the limiting prices. Revenue therefore attains its upper bound. A bound equal to ϕ would contradict the pointwise strict inequalities just proved, after excluding the stipulated ϵ -neighborhood of an incumbent own quote. This proves the uniform claims. $\square$

Exact certification. The menu proof above is a finite, exact computer-assisted proof. The replication archive contains the weak-cell generators, executable rational and algebraic sign checks, and the complete menu ledger. The designated suite runs 15 proof-bearing files: support reduction, primitive reflection, the symmetric and unequal pair correspondence, all-active generic and boundary exclusion, mixed-menu exclusion, and all six neutral-menu closures. The proposition uses their composition; no individual computation is treated as an exhaustive proof by itself. The replication archive distributes that suite in `verification/exact_completion_archive`, together with a one-command runner, tested software versions, hashes, deterministic output, and a certificate-linked disposition ledger for all 32 residual menus.

C.8 Exact boundaries and exact nonclaims

At ϕ_P , the wide menu has fee $(27 - \sqrt{621})/108$ and its best pair entrant earns exactly ϕ_P ; the pair menu also exists. At ϕ_U , the symmetric pair fee is $3/7$ and its best lower same-product comparison is tangent; the wide menu also exists. At ϕ_W , the wide fee is $10/33$, its revenue curve is at its maximum, and it coexists with the empty completion menu.

Within the pair scope, $2/3$ changes the symmetric clientele, ϕ_- is a weak own-price boundary, ϕ_0 joins the two unequal support formulas, and ϕ_+ merges the unequal prices into the symmetric

vector. Because all four values lie strictly between ϕ_P and ϕ_U , the wide scope also exists at each one.

The exact problem has no remaining m, c, F, τ parameters and excludes $\phi = 0$. The market fund has already been established by Proposition 1; the theorem here classifies its possible specialists. Its thresholds are exact for the derived problem. They do not locate exact positive- c transition functions or protect against coordinated launches of several products.

D Validity for positive c

D.1 What is transferred, and on which domain

The exact problem in Appendix C describes the first nonconstant term in investors' portfolio choices and the first nonzero term in specialist revenue. This appendix establishes two distinct connections to the primitive economy. First, normalized equilibrium sequences are precompact on the stated domains, and every accumulation point is an equilibrium of the exact reduced problem. This rules out additional primitive scopes away from the exact scope boundaries. Second, suitably protected exact configurations can be continued to primitive equilibria. The second claim requires control of deviations arbitrarily close to an incumbent quote; it does not follow from convergence of investor objectives alone.

Use the normalized-cost coordinates from (16). Write

$$w = 1 + 3m^2, \quad Q^0 = \frac{4}{w}, \quad t = \frac{m}{w}, \quad K = \frac{m}{(1 - m^2)w}, \quad \Theta = tK, \quad \phi = \frac{F}{\Theta c^2}. \quad (\text{D.1})$$

At every positive c , the inverse identity $F = \Theta c^2 \phi$ recovers the same primitive operating cost. The local analysis holds m and positive ϕ fixed as c decreases. The general compactness and exclusion statements take m in a compact subset of $(0, 1)$, ϕ in a compact subset of $(0, \infty)$, and a common positive floor on paid non-home access costs. Constants and the sufficiently-small- c threshold are uniform on those sets; the primitive operating cost then satisfies the independent ceiling in Assumption 2. This uniformity does not cover sequences with ϕ tending to zero or infinity. The continuation results use the narrower domain in Appendix B.9, together with the strict branch and event separations specified below.

An additional support condition is needed when using the symmetric pair as a primitive representative of pair scope. For a compact parameter set $\mathcal{K}$, write $\mathcal{K}_P = \{(m, \phi) \in \mathcal{K} : 0 < \phi < \phi_U\}$. Whenever $\mathcal{K}_P$ is nonempty, the symmetric-pair continuation used for scope equality satisfies

$$\inf_{(m, \phi) \in \mathcal{K}_P} \left\{ 2 - m^2[10 + \phi - 8p_s(\phi)] \right\} > 0, \quad (\text{D.2})$$

where p_s is the lower symmetric break-even quote in (C.34). This is the uniform version of (B.20). It keeps both middle home positions exactly zero on the constructed symmetric continuation. Below we prove why this second-order support restriction also matters for pricing protection. It is not imposed on the empty or pooled continuation, or on the all-equilibrium exclusion result.

D.2 Eliminating the market-fund coordinate

The frictionless target expansion in (A.11) implies the composition statement used to motivate Section 3.2.1. Write $\pi^*(c) = e^*(c)/(\mathbf{1}'e^*(c))$. Since total exposure has no first-order change, normalization gives

$$\pi^*(c) - b_B = c \frac{m}{2(1-m^2)}(-1, 0, 1, 0)' + O(c^2). \quad (\text{D.3})$$

The corresponding exposure difference is $e^*(c) - Q^0 b_B = Kc g + O(c^2)$, with $g = (-2, 0, 2, 0)'$. This identifies the desired adjustment in the body's investor problem before incorporating fees and implementation constraints.

In Section 3.2's notation, $tc = \alpha$ and $Kc = \alpha/(1-m^2)$. Its demand lemma uses the exposure correction $r^* = Kc g$, rather than the change in portfolio weights in (D.3). Appendix C.1 gives the unit dictionary and the distance-form objective in the body's fee and holding variables.

Let $b = \mathbf{1}/4$, $d_T = b_T - b$, and $H_i = \iota_i - b$. For a type- i portfolio with only home access, write

$$q_{iT} = Kc u_{iT}, \quad x_i = Kc h_i, \quad y_i = \sum_T u_{iT} d_T + h_i H_i, \quad n_i = \sum_T u_{iT} + h_i.$$

The final exposure is $M_i b + Kc y_i$, where $M_i = q_{iB} + Kc n_i$ is total risky investment. Thus the actual market-fund position is $q_{iB} = M_i - Kc n_i$. The variable y_i is a zero-sum exposure adjustment, not an additional asset or fund. Let $v = b' \Sigma b$, put $z_i = Kc y_i$, and quote fees as $\delta_B = c^2 \beta$, $\delta_T = tc p_T$. Conditional on the adjustment, investor utility before any access charge equals

$$(1 - \delta_B) M_i - \frac{1}{2} v M_i^2 - M_i b' \Sigma z_i - \frac{1}{2} z_i' \Sigma z_i - \Theta c^2 \sum_T p_T u_{iT} + \delta_B Kc n_i. \quad (\text{D.4})$$

The last term accounts for market-fund fees saved when a home asset or a specialist replaces part of the common position. It is retained here even though it enters at an order beyond the exact quotient.

Lemma D.1 (Uniform portfolio normal form). *On every bounded set of nonnegative adjustment quantities, normalized specialist quotes, and β , the optimized market-fund position is positive for all sufficiently small c . Subtracting the optimized market-fund-only value and dividing by Θc^2 gives*

$$\Psi_i^c(u_i, h_i; p, \beta) = g' y_i - \frac{1}{2} \|y_i\|^2 - \sum_T p_T u_{iT} + \mathcal{E}_i^c(u_i, h_i; \beta), \quad g = (-2, 0, 2, 0)', \quad (\text{D.5})$$

where $\mathcal{E}_i^c = O(c)$ uniformly together with its first two quantity derivatives. The same assertion holds on each fixed raw portfolio support after differentiating its nonsingular linear demand system with respect to prices.

Proof. The maximizer over the common coordinate in (D.4) is

$$M_i^* = \frac{1 - \delta_B - b' \Sigma z_i}{v}. \quad (\text{D.6})$$

Uniformly on bounded adjustment sets, $M_i^* = Q^0 + O(c^2)$, whereas $Kcn_i = O(c)$, so $q_{iB} > 0$. Subtracting $(1 - \delta_B)^2/(2v)$ leaves, after normalization,

$$-\frac{(1 - \delta_B)b'\Sigma z_i}{v\Theta c^2} + \frac{(b'\Sigma z_i)^2}{2v\Theta c^2} - \frac{z_i'\Sigma z_i}{2\Theta c^2} - \sum_T p_T u_{iT} + \frac{\delta_B Kcn_i}{\Theta c^2}. \quad (\text{D.7})$$

At zero dispersion, $\Sigma_0 = (1 - m^2)I + m^2\mathbf{1}\mathbf{1}'$, $v_0 = w/4$, and $b'\Sigma_0 y_i = 0$. Writing $\nu = (1, 0, -1, 0)'$, direct differentiation gives $\Sigma'(0)b = (m/2)\nu$. Hence

$$b'\Sigma z_i = \frac{Km}{2}c^2\nu'y_i + O(c^3\|y_i\|), \quad z_i'\Sigma z_i = K^2c^2(1 - m^2)\|y_i\|^2 + O(c^3\|y_i\|^2).$$

The identities $Km/(2v_0\Theta) = 2$ and $K^2(1 - m^2)/\Theta = 1$ establish the displayed linear and quadratic terms. The fee-saving term is $c(\beta/t)n_i$, and the squared cross term is $O(c^2\|y_i\|^2)$. Every remaining coefficient is a smooth rational function of the primitives with denominators bounded away from zero. This proves the uniform derivative bounds. On a fixed support, the raw first-order conditions have the form $(A'\Sigma A)q = \mathbf{1} - \delta$. A linearly independent collection of fixed portfolio columns has a uniformly nonsingular Gram matrix; its price derivatives and the corresponding expansions follow by inversion. $\square$

The scale conversion is exact:

$$\frac{\delta_T \text{AUM}_T}{\Theta c^2} = p_T \left(\frac{1}{4} \sum_{i=0}^3 u_{iT} \right). \quad (\text{D.8})$$

Consequently the right-hand side of specialist break-even is precisely ϕ . The quotient does not equate a percentage fee to a fixed cost, or treat the primitive market fund as free. It removes a common value term and measures specialist fees and quantities in their respective units, tc and Kc .

D.3 Demand limits, ties, and profitable deviation witnesses

Lemma D.2 (Compact demand and the direction of convergence). *Retain B at a fee bounded by $C_B c^2$. Suppose the positive normalized prices in a fixed specialist menu lie in $[\epsilon, P]$, where $\epsilon > 0$. Every primitive investor optimum has no paid access, and its specialist and home quantities are $O(c)$, uniformly. Its market-fund quantity is $Q^0 + O(c)$. Every limit of normalized optimal implementations solves the exact investor problem. The assertion holds for a jointly selected incumbent allocation and for focal-upper optimal allocations under an entry or a lower-own-quote test.*

Conversely, a strictly profitable exact entry or lower-own-quote deviation can be approximated by primitive deviations that remain strictly profitable. This assertion allows an arbitrarily small additional reduction of the focal normalized quote; no assumption of unique product quantities is required.

Proof. Lemma A.10, applied to the option of using the cheap market fund alone, bounds the sum of the gross quadratic loss, all fee payments, and all access charges by Cc^2 , for every optimizer.

The access-cost floor excludes paid access. Uniform curvature gives $e_i - e^*(c) = O(c)$. The lower price bound ϵc then gives $q_{iT} = O(c)$ for every specialist. Projecting exposure onto the zero-sum subspace yields

$$x_i H_i = e_i - (\mathbf{1}' e_i) b - \sum_T q_{iT} d_T = O(c),$$

so $x_i = O(c)$. Since $\mathbf{1}' e_i = Q^0 + O(c)$, the claimed market-fund bound follows.

Exact optimizers are compact as well. If $y = \sum_T u_T d_T + h H_i$, comparison with the feasible zero adjustment gives

$$\frac{1}{2} \|y - g\|^2 + \sum_T p_T u_T \leq \frac{1}{2} \|g\|^2 = 4.$$

This bounds y , $\sum_T u_T \leq 4/\epsilon$, and then h . All such exact portfolios are feasible as primitive portfolios for small c , by (D.6). The uniform normal form and a comparison with each feasible exact portfolio show that every limit of primitive maximizers is an exact maximizer. Taking a common subsequence for the four investor types preserves a single compatible incumbent demand selection. Taking focal-upper selections instead proves the corresponding outer-limit assertion for deviation demand. In particular, on compact positive-quote sets, limiting primitive focal demand cannot exceed exact focal-upper demand.

For the converse, fix a profitable exact deviation at a normalized quote $\eta > 0$. Its optimizers form a compact set. Select an optimizer a^* that maximizes the focal product quantity on that set. Let $\epsilon_c \rightarrow 0$ bound the error of the normalized primitive objective relative to this exact objective, including the converging nonfocal prices. Choose $\gamma_c \downarrow 0$ with $\epsilon_c/\gamma_c \rightarrow 0$, and use the focal quote $\eta - \gamma_c > 0$. If a_c is any primitive optimizer at that quote, its optimality relative to a^* gives

$$f(a_c) + \gamma_c u_T(a_c) \geq f(a^*) + \gamma_c u_T(a^*) - 2\epsilon_c,$$

where f is the exact objective at η . Since $f(a_c) \leq f(a^*)$, every limit has focal quantity at least that of a^* . The first part of the proof makes the limit exact-optimal, so it is focal-maximizing. Applied type by type, this recovers the exact upper focal AUM and its strict revenue gain. In an own-quote test, $\eta < p_T$ already has a strict distance from the incumbent quote; the extra reduction remains admissible. This is why tied exact demand does not obstruct lifting a strictly profitable deviation. $\square$

The two directions serve different purposes. Outer convergence and upper focal demand transfer strict protection gaps from the exact problem to a candidate primitive configuration. Recovery after a vanishing price reduction transfers a strict profitable witness in the other direction, which is needed to exclude limits that are not exact equilibria.

D.4 The two quote tails and market-fund protection

Compact positive-quote convergence does not cover every primitive price. In particular, a triple with a fee of order c^2 can replace an order-one amount of B for its complementary home type. Its quantity need not be $O(c)$. The following bound treats that possibility explicitly.

Lemma D.3 (Global quote-tail bounds). *Consider a candidate configuration with*

$$\delta_B = \frac{F}{Q^0} + O(c^3), \quad \kappa c \leq \delta_T \leq Cc \quad \text{for each active specialist,}$$

where $\kappa > 0$. *In any omitted-product entry or specialist lower-own-quote test, retain all nonfocal incumbent fees. There is a common $P_* < \infty$ such that a focal fee $tc\eta$, $\eta > P_*$, attracts no positive demand. For every $\zeta > 0$ there are $\epsilon_* > 0$ and $c_* > 0$ such that every focal fee with $0 \leq \eta \leq \epsilon_*$ satisfies*

$$\delta_U \overline{\text{AUM}}_U \leq \left(\frac{1}{4} + \zeta\right) F \quad (0 < c < c_*). \quad (\text{D.9})$$

Both statements include all optimal implementations and all paid-access alternatives.

Proof. The cheap market-fund option and Lemma A.10 imply, under every tested quote, no paid access and $\|e_i - e^*(c)\| \leq Cc$. If product U is held positively, its first-order condition gives

$$\delta_U = 1 - b'_U \Sigma e_i = b'_U \Sigma(e^*(c) - e_i) \leq Cc.$$

The uniform positive lower bound on t proves the high-price claim.

For the other tail, the fee bound on every nonfocal specialist and the same utility comparison imply that each nonfocal specialist quantity is $O(c)$. The zero-sum exposure identity consequently gives

$$q_{iU} d_U + x_i H_i = O(c). \quad (\text{D.10})$$

Among the eight arc products, a nonnegative combination of d_U and H_i can equal zero nontrivially only when U is the triple complementary to type i . Indeed, cancellation requires $b_U - b$ to be a negative multiple of $\iota_i - b$, so all three non-home weights in b_U must coincide and its home weight must be zero. The unique such arc product is the complementary triple, with the identity

$$3b_U + \iota_i = 4b. \quad (\text{D.11})$$

For every other product-type pair, compactness of the two-generator nonnegative cone makes (D.10) imply $q_{iU} + x_i \leq Cc$, uniformly.

For the complementary type, if $3\delta_U > 4\delta_B$, a portfolio holding both U and its home asset positively can remove a positive multiple of $(3U, \text{home})$ and buy $4B$ instead. This preserves exposure and strictly lowers fees. At least one of the two holdings must therefore be zero in every optimum, and (D.10) again bounds both by Cc . In this case, and for every noncomplementary type, focal revenue is bounded by $C\eta c^2$.

It remains to consider $3\delta_U \leq 4\delta_B$, including the tie. Nonnegativity of all exposures and $e_{ij} = 1/w + O(c)$ for a non-home asset give $q_{iU} \leq 3/w + O(c)$. Only one type, of population mass $1/4$, can have this order-one demand. Consequently

$$\delta_U \overline{\text{AUM}}_U \leq \frac{4\delta_B}{3} \left(\frac{3}{4w} + O(c)\right) = \frac{\delta_B}{w} + O(c^3) = \frac{F}{4} + O(c^3).$$

This holds at the replication tie for the upper focal selection as well. Together the two cases bound the revenue by $\max\{C\eta c^2, F/4 + Cc^3\}$, after enlarging C . Because $F/c^2 = \phi\Theta$ has a uniform positive lower bound, first choosing ϵ_* and then c_* proves (D.9). $\square$

Lemma D.4 (Protection of the market-fund quote). *On the candidate menus $\{B\}$, $\{B, C\}$, and $\{B, L, R\}$, suppose specialist prices are in fixed positive order- c bands and B breaks even at a fee $\delta_B = F/Q^0 + O(c^3)$. Then every lower own quote by B earns strictly less than F , uniformly for sufficiently small c .*

Proof. At every $0 \leq \eta_B \leq \delta_B$, the cheap- B comparison excludes paid access and gives a positive market-fund position for each type. For $\{B, L, R\}$, the portfolio columns (b, b_L, b_R, ι_i) are linearly independent for every i . For $\{B, C\}$, the three columns are independent except for $i = 0$. That exception has the sole dependency $3b_C + \iota_0 = 4b$; because $3\delta_C > 4\eta_B$, no optimum can use both C and home positively. The remaining supports have independent columns.

The demand on each independent support is affine in η_B , with slopes uniformly bounded by the inverses of its Gram matrix. Strictly cheaper removal of the sole redundant cycle gives unique holdings also for type 0 on the pooled menu. The finitely many affine pieces agree at their support boundaries, because optimal holdings are unique there. Thus aggregate B demand is uniformly Lipschitz on the entire interval, with constant L_B . Write $A_B(\delta_B) \geq a_B > 0$. For every $\eta_B < \delta_B$,

$$\begin{aligned} F - \eta_B A_B(\eta_B) &\geq (\delta_B - \eta_B)\{A_B(\delta_B) - L_B \eta_B\} \\ &\geq \frac{a_B}{2}(\delta_B - \eta_B) > 0, \end{aligned}$$

when c is sufficiently small. This proves protection for the full lower-quote interval, including its endpoints and portfolio-support changes. $\square$

D.5 Continuing break-even configurations and their supports

A break-even configuration consists of prices and a common selection of investor optima at which every present fund covers F . It is not yet an equilibrium: all omitted-product and lower-own-quote tests still have to be checked. Keeping these steps separate is especially useful for the symmetric pair.

Lemma D.5 (Strict configuration continuation). *On the continuation domain of Appendix B.9, choose one of the following compact branch interiors:*

1. *the empty specialist branch, with ϕ in a compact subset of (ϕ_W, ∞) ;*
2. *the pooled branch, with ϕ in a compact subset of (ϕ_P, ϕ_W) ;*
3. *the symmetric pair branch, with ϕ in a compact subset of $(0, \phi_U)$, separated from $2/3$ and $\phi_+ = 45/64$, and satisfying (D.2);*
4. *either reflected lower unequal branch, with ϕ in a compact subset of (ϕ_-, ϕ_0) ;*

5. either reflected upper unequal branch, with ϕ in a compact subset of (ϕ_0, ϕ_+) .

Every such exact configuration has a primitive break-even continuation with

$$\begin{aligned}\delta_B &= \frac{F}{Q^0} + O(c^3), & \text{AUM}_B &= Q^0 + O(c), \\ \delta_T &= tc\{p_T(\phi) + O(c)\}, & \text{AUM}_T &= Kc\{Q_T(\phi) + O(c)\}.\end{aligned}\tag{D.12}$$

The constants and the maximum permitted c are uniform on each compact branch set. These configurations have globally optimal investor portfolios, a compatible incumbent selection, and no paid non-home access.

Proof. For a proposed raw support, solve

$$q_{i,S} = (A'_{i,S} \Sigma A_{i,S})^{-1} (\mathbf{1} - \delta_{i,S}).\tag{D.13}$$

A positive solution whose omitted-column marginal utilities are nonpositive satisfies the full conditional investor KKT system. Uniform curvature then gives global optimality conditional on home access. Lemma A.10 rules out all other access choices once $\delta_B = O(c^2)$. The raw support matrices used below have independent columns. Their inverses, and their derivatives on compact parameter sets, are uniformly bounded.

For the empty specialist menu, Proposition A.11 constructs the low cost-recovery root using exact market-fund/home demand. Its expansion gives $\delta_B = F/Q^0 + O(c^3)$. Any home-support changes in that calculation are handled by the positive-part formula, so smoothness across such a change is not needed.

On the pooled menu, types 0, 1 and 3 use B, C , while type 2 also uses its home asset. The positive quantities and inactive-home KKT margins are strict at first order on the stated compact interior. The normalized pooled revenue is

$$\mathcal{R}_C(z) = z \left(\frac{15}{2} - \frac{99}{8} z \right), \quad \mathcal{R}'_C(z) = \frac{15}{2} - \frac{99}{4} z > 0 \quad (0 < z < 10/33).$$

Writing $\delta_B = c^2 \beta$, the two rescaled break-even equations have limiting market-fund derivative $Q^0/\Theta > 0$ and the displayed specialist derivative. The implicit function theorem gives the continued root and its strict investor supports.

For the symmetric pair, impose $\delta_L = \delta_R = tcp$. Primitive reflection makes their revenues identical. Types 0, 1 and 3 use B, L, R with no home position. Type 2 uses B, L, R, ι_2 when $p < 1/3$, and B, ι_2 when $p > 1/3$. At the limiting symmetric quote, the two middle home KKT conditions are zero at first order. They must therefore be checked at the next order.

For that check, temporarily allow type 1's home holding to have either sign and solve the full (B, L, R, ι_1) system. If $v_i = \Sigma e_i$, its first-order conditions give

$$v_i = (1 - 4\delta_B + 2\delta_R, 1, 1 - 2\delta_L, 1 + 2\delta_L - 2\delta_R)', \quad x_1^{\text{free}} = (-1, 1, -1, 1)' \Sigma^{-1} v_i.\tag{D.14}$$

Here the superscript indicates the solution of a linear system in which only the sign of the home position is relaxed. Expanding (D.14) at equal specialist fees and $\delta_B = \phi\Theta c^2/Q^0 + O(c^3)$ gives

$$x_1^{\text{free}} = x_3^{\text{free}} = \frac{m^2(10 + \phi - 8p) - 2}{(1 - m^2)^2(1 + 3m^2)} c^2 + O(c^3). \quad (\text{D.15})$$

The expansion follows by inserting $\Sigma = \Sigma_0 + c\Sigma_1 + c^2\Sigma_2$ in (D.14) and using

$$\Sigma^{-1} = \Sigma_0^{-1} - c\Sigma_0^{-1}\Sigma_1\Sigma_0^{-1} + c^2\{\Sigma_0^{-1}\Sigma_1\Sigma_0^{-1}\Sigma_1\Sigma_0^{-1} - \Sigma_0^{-1}\Sigma_2\Sigma_0^{-1}\} + O(c^3).$$

The Schur complement for the home column is positive, so the inactive-home KKT margin has the sign of x_i^{free} . Condition (D.2) makes that margin strictly negative at order c^2 , uniformly. All other required support signs are strict at first order away from $p = 1/3$.

For each nearby p , the market-fund equation has its low root $\delta_B(c, p) = F/Q^0 + O(c^3)$, with $\partial_p \delta_B = O(c^3)$. It remains to solve one common specialist break-even equation. Its limiting revenue and derivative are

	$\mathcal{R}_P(p)$	$\mathcal{R}'_P(p)$
$p < 1/3$	$p(4 - 6p)$	$4 - 12p$
$p > 1/3$	$3p(1 - p)$	$3 - 6p$

(D.16)

They have a positive derivative on the stated lower-root interiors. An implicit function argument supplies the symmetric root. The support condition remains strict because the new normalized quote differs from $p_s(\phi)$ by $O(c)$. This use of a common-price path proves configuration existence; its derivative is not a substitute for the one-product lower-quote calculation below.

For the lower unequal orientation $\ell < r$, the supports are

i	positive raw columns	inactive menu columns
0	B, L, R	ι_0
1	B, L, R	ι_1
2	B, L, ι_2	R
3	B, L, R, ι_3	none.

Solving (D.13) and using the normal form gives

$$\frac{\text{AUM}_L}{Kc} = \frac{7}{2} - \frac{11}{2}\ell + r + O(c), \quad \frac{\text{AUM}_R}{Kc} = 3 + \ell - 4r + O(c), \quad \text{AUM}_B = Q^0 + O(c).$$

The nonzero quantities and the inactive-column inequalities are strict at first order on compact interiors of this branch. The determinant of the two limiting specialist revenue equations, in the parameterization (C.40), is

$$\det J_- = \frac{448s^3 - 456s^2 + 264s - 209}{12(8s^2 - 11)^2} < 0, \quad \frac{1}{8} \leq s \leq \frac{1}{2}. \quad (\text{D.17})$$

For a direct sign check, $448s^3 \leq 224s^2$ on this interval. The numerator is at most $-232s^2 + 264s - 209$, an increasing function there with terminal value -135 . The market-fund block is again $Q^0/\Theta > 0$, so the full rescaled Jacobian is nonsingular.

On the upper unequal orientation, types 0 and 1 use B, L, R , type 2 uses B, ι_2 , and type 3 uses B, L, R, ι_3 . The limiting specialist demands are $Q_L = 3 - 4\ell + r$, $Q_R = 3 + \ell - 4r$. On $\ell + r = 3/4$, the revenue Jacobian determinant is

$$\det J_+ = -\frac{5}{4}(8\ell - 3)^2 \neq 0, \quad 1/3 < \ell < 3/8. \quad (\text{D.18})$$

The implicit function theorem therefore applies to both unequal branches. Their strict first-order support signs persist. Reflection gives the other orientation without imposing symmetry on the prices. Compactness makes all of these constructions uniform and proves (D.12). $\square$

D.6 Protection against quotes close to the incumbent

Lemma D.6 (Local specialist pricing protection). *For each compact configuration family in Lemma D.5, there exist $a > 0$, $\epsilon_0 > 0$, and $c_0 > 0$ such that every active specialist T and every normalized own quote $p_T^c - \epsilon_0 \leq \eta < p_T^c$ satisfy*

$$\phi - \frac{tc\eta \overline{\text{AUM}}_T(\eta; \text{nonfocal incumbent fees})}{\Theta c^2} \geq a(p_T^c - \eta) > 0 \quad (0 < c < c_0). \quad (\text{D.19})$$

Proof. On the pooled and unequal branches, the incumbent investor supports are strict at first order. They remain fixed in a common normalized-price neighborhood. Demand there is smooth and uniquely implemented. The exact own-revenue slopes at the incumbent quote are

product and branch	own-revenue derivative
C	$15/2 - 99z/4$
lower unequal, cheaper	$7/2 + r - 11\ell$
lower unequal, expensive	$3 + \ell - 8r$
upper unequal, cheaper	$15/4 - 9\ell$
upper unequal, expensive	$9\ell - 3$.

(D.20)

They are strictly positive on the compact interiors specified above; see Lemma C.10. Shrinking the price neighborhood if necessary, their primitive counterparts are uniformly positive by the differentiated linear demand systems. Integrating the revenue derivative from the lower quote to the incumbent proves the result for these products.

For the symmetric pair, a first-order strict-support argument would be incomplete. At the incumbent, both middle homes are inactive with an order- c^2 margin. A lower quote can activate one of them after a normalized price movement of order c , within every fixed neighborhood of the incumbent. We check that change explicitly.

Hold δ_B, δ_R fixed and lower δ_L . Differentiating the full-support formula (D.14) gives, uniformly for small c ,

$$\frac{\partial x_1^{\text{free}}}{\partial \delta_L} = \frac{4}{1 - m^2} + O(c) > 0, \quad \frac{\partial x_3^{\text{free}}}{\partial \delta_L} = -\frac{4}{1 - m^2} + O(c) < 0. \quad (\text{D.21})$$

Both relaxed holdings are negative at the symmetric quote by (D.15) and the support condition. Lowering L 's quote makes type 1's relaxed home holding still more negative. Only type 3's home can enter. The sign test is valid for the constrained problem because its Schur complement is positive. These relaxed holdings are affine in the varied fee, so there is at most one such entry wall. The other investor supports are unchanged in a sufficiently small normalized-price neighborhood, uniformly away from $p = 1/3$. Thus the path has either both middle homes off or type 3's home alone on; it never reaches the cell with both middle homes positive. Reflection gives the identical conclusion for a lower quote by R .

For completeness, solving the linear systems on these two possible cells gives the following limiting derivatives at the symmetric reference quote:

	both middle homes off	one middle home on	
$p < 1/3$	$4 - 11p$	$4 - 12p$	
$p > 1/3$	$3 - 6p$	$3 - 7p$	(D.22)

For example, when $p > 1/3$, incumbent aggregate demand is $3(1 - p)$. Its own-price slope is -3 before the home-entry wall and -4 after it; differentiating fee times demand gives the second row. When $p < 1/3$, demand is $4 - 6p$, with own-price slopes -5 and -6 , respectively, giving the first row. All four revenue derivatives are uniformly positive on the specified compact protected intervals: $p < 1/3$ below the clientele event and $1/3 < p < 3/7$ above it. The primitive derivatives differ by $O(c)$ on each cell and vary continuously with the quote within that cell. They therefore remain positive after reducing the neighborhood and c_0 if needed. Demand is continuous at the home-entry wall, since the raw (B, L, R, ι_i) columns are independent and the entering home quantity is zero there. Integrating over the at most two cells proves (D.19), including quotes arbitrarily close to the incumbent. $\square$

D.7 Strict branch persistence under every price test

Proposition D.7 (Primitive persistence of protected exact branches). *Each compact exact branch family listed in Lemma D.5 has a primitive equilibrium continuation for all sufficiently small positive c , uniformly in its parameters. Its raw menu is $\{B\} \cup E$, with the prices and AUM in (D.12). The equilibrium satisfies all investor access and portfolio comparisons, exact active-fund break-even, every active-product lower-own-quote inequality, and every omitted one-product entry inequality over the full eight-product catalogue. The test uses the common incumbent and focal-upper demand conventions of Section 2. It imposes no restriction on entrant scope or price.*

Proof. Begin with the break-even configuration in Lemma D.5. Investor optimality and the access comparisons are already satisfied. Lemma D.4 protects every lower quote by B . It remains to check specialist entry and own-quote tests.

Choose a low normalized-price cutoff $\epsilon_* > 0$ from Lemma D.3, with $\zeta < 1/4$, and its high-price cutoff P_* . The same lemma protects all prices below ϵ_* by a revenue bound below $F/2$, and all prices above P_* have zero demand. The cutoffs are uniform on the compact branch family. On the remaining compact positive-price interval, exact omitted-entry revenues lie uniformly below ϕ by Lemma C.10. The outer-limit statement in Lemma D.2 transfers this strict inequality to the primitive configuration. More explicitly, failure would give a convergent sequence of tested prices and upper focal optimal portfolios; its exact limit would violate the strict exact revenue gap.

For an active specialist, first exclude a fixed neighborhood below its incumbent quote. On the remaining compact positive-price set, the same exact strict-gap lemma and outer-limit argument protect every lower quote. The remaining neighborhood, including price differences tending to zero with c , is protected by Lemma D.6. These regions exhaust the entire lower-quote interval. The arguments use all eight possible focal products and preserve the upper-demand convention at every tied deviation. Consequently the configuration passes all equilibrium tests. $\square$

D.8 Closure of all primitive equilibria

Proposition D.8 (Every primitive equilibrium limit is exact). *Maintain the general compact parameter and access-cost conditions at the start of this appendix. Let $c_n \downarrow 0$ and select any primitive equilibrium at each (m_n, ϕ_n, c_n) . After a subsequence, its non-market menu E is constant, its parameters converge, and*

$$p_{T,n} = \frac{\delta_{T,n}}{t(m_n)c_n}, \quad u_{iT,n} = \frac{q_{iT,n}}{K(m_n)c_n}, \quad h_{i,n} = \frac{x_{i,n}}{K(m_n)c_n}$$

converge to a compatible exact equilibrium allocation and fee vector at $\phi = \lim_n \phi_n$. Every product in E retains positive aggregate demand and a strictly positive price in the limit. In particular, its menu is one of the scopes classified in Proposition C.7.

Proof. Proposition 1 supplies the endogenous market fund in every primitive equilibrium under consideration. The bounds in Appendix B.9 and Lemma 4 imply, uniformly,

$$\text{AUM}_B = Q^0 + O(c_n), \quad \delta_B = \frac{F}{Q^0} + O(c_n^3), \quad 0 < \underline{p} \leq p_{T,n} \leq \bar{p} < \infty \quad (T \in E_n).$$

Active specialist AUM divided by Kc_n has uniform positive lower and finite upper bounds. All home positions divided by Kc_n are bounded. The lower bounds follow from exact break-even and the uniform upper bounds: either a vanishing fee or a vanishing AUM on that scale would make normalized revenue vanish, contradicting ϕ_n bounded away from zero. Thus these are consequences of primitive equilibrium, not extra restrictions selecting a class of equilibria.

There are only finitely many menus. Choose a subsequence with constant E , convergent parameters, and convergent normalized prices and incumbent quantities for all four types. Lemma D.2 makes each limiting portfolio exact-optimal and preserves their compatible joint selection. Dividing

each incumbent break-even equality by $\Theta(m_n)c_n^2$ gives

$$p_T \left(\frac{1}{4} \sum_i u_{iT} \right) = \phi.$$

The positive bounds ensure that no product of E disappears in the limit.

If the limiting configuration failed exact no-entry or lower-own-quote protection, the definition of failure would provide a focal product and a strictly profitable quote $\eta > 0$. For an own-quote failure the quote is strictly below its limiting incumbent fee. The focal-recovery part of Lemma D.2 supplies admissible primitive quotes with revenues converging to that strictly profitable exact revenue. They would earn more than F_n for all sufficiently large n , contradicting primitive equilibrium. This checks all eight possible entrant products and all exact lower-own-quote tests, with their focal-upper convention. The limit is therefore an exact equilibrium, to which the exhaustive exact classification applies. $\square$

The following proposition states the positive- c counterpart of Proposition 2, cited in its footnote. It separates the exclusion of other scopes from existence of the scopes in the classification. The existence claim retains the additional continuation conditions described above.

Proposition D.9 (Uniform correspondence of equilibrium scopes). *Let $\mathcal{S}_c(m, \phi)$ denote the set of non-market scopes that occur in primitive equilibria, and $\mathcal{S}_0(\phi)$ the exact scope correspondence in (C.53).*

On a compact general parameter set with positive distance from ϕ_P, ϕ_U, ϕ_W , there is a uniform $c_ > 0$ such that*

$$\mathcal{S}_c(m, \phi) \subseteq \mathcal{S}_0(\phi) \quad (0 < c < c_*). \quad (\text{D.23})$$

This all-equilibrium exclusion requires neither the narrower continuation domain nor (D.2).

If the compact set also lies in the continuation domain of Appendix B.9, satisfies (D.2) wherever pair scope belongs to $\mathcal{S}_0(\phi)$, and is separated from $2/3$ and $\phi_+ = 45/64$ on that pair region, then, after reducing c_ ,*

$$\mathcal{S}_c(m, \phi) = \mathcal{S}_0(\phi) \quad (0 < c < c_*). \quad (\text{D.24})$$

The equality concerns the set of scopes. It does not assert uniqueness of fees or portfolios, identify every primitive equilibrium branch, or locate an exact finite- c cutoff.

Proof. If the uniform inclusion failed, there would be primitive equilibria at a sequence $c_n \downarrow 0$, with parameter points in the compact set, whose scopes are not in $\mathcal{S}_0(\phi_n)$. Extract a constant menu and a convergent parameter subsequence. Proposition D.8 puts that menu in $\mathcal{S}_0(\phi_*)$. The exact scope table is locally constant away from its three scope boundaries. Positive separation from those boundaries therefore gives $\mathcal{S}_0(\phi_n) = \mathcal{S}_0(\phi_*)$ eventually, a contradiction. This proof does not construct an equilibrium and hence does not use a continuation support condition.

For the reverse inclusion, use one protected representative of each exact scope. If the empty menu belongs to the correspondence, then $\phi > \phi_W$, and the empty continuation applies. If C belongs, then $\phi_P < \phi < \phi_W$, and its pooled continuation applies. If the pair belongs, then $0 < \phi < \phi_U$, and the symmetric continuation applies with the additional support and event conditions. Each is a

primitive equilibrium by Proposition D.7. A finite compact cover of these branch regions gives a common positive bound on c . Where the exact table has overlapping scopes, the corresponding distinct primitive continuations both exist. Unequal pair branches provide additional fee vectors but are not needed to establish the reverse scope inclusion. Combining the two inclusions proves the equality. $\square$

D.9 Why the symmetric support condition cannot be omitted

The middle-home condition has a pricing implication even though the home position it controls is only order c^2 . That position determines which demand derivative an incumbent faces at a sufficiently nearby quote. The exact quotient alone does not retain this information.

Proposition D.10 (A failure of unqualified symmetric persistence). *The two numerical continuation bounds in Appendix B.9, without the additional symmetric-pair support condition, do not suffice to continue every protected exact symmetric pair as a symmetric primitive equilibrium. In particular, at*

$$m = \frac{3}{5}, \quad \phi = \frac{18}{25}, \quad p_s(\phi) = \frac{2}{5}, \quad (\text{D.25})$$

the nearby symmetric primitive break-even configuration has a profitable lower-own-quote deviation for all sufficiently small positive c . This statement does not exclude other primitive pair configurations outside the sufficient domain of (D.24).

Proof. The displayed parameters satisfy $m^2 < 44/69$, $\phi < 2(1 - m^2)/m^2$, and $2/3 < \phi < 36/49$. Thus the symmetric pair is strictly protected in the exact quotient and lies away from its clientele event. But (D.15) gives

$$x_1 = x_3 = \frac{425}{512}c^2 + O(c^3) > 0. \quad (\text{D.26})$$

The valid primitive supports are (B, L, R) for type 0, (B, L, R, ι_1) for type 1, (B, ι_2) for type 2, and (B, L, R, ι_3) for type 3. All other nontrivial support signs are strict at first order. Solving their linear systems and the two symmetric break-even equations gives a nearby symmetric configuration: the limiting common-price revenue derivative is $3 - 6p = 3/5 > 0$, so the corresponding implicit function theorem applies.

Now lower only L 's quote, holding the market-fund and R fees fixed. On the cell where both middle homes are positive, aggregate normalized own demand has derivative $-5 + O(c)$, and its incumbent level is $3(1 - p) + O(c)$. Hence the normalized own-revenue derivative is

$$3 - 8p + O(c) = -\frac{1}{5} + O(c) < 0. \quad (\text{D.27})$$

In contrast, the exact quotient's one-sided slope is $3 - 7p = 1/5 > 0$. The difference arises because the primitive deviation can initially keep both positive home positions.

Choose a sufficiently small fixed $\varepsilon > 0$ and lower the normalized L quote by εc . The home-price sensitivities in (D.21), together with (D.26), ensure that both home positions remain positive along this price change. The other strict supports remain unchanged. Integrating (D.27) shows that

normalized revenue rises by $(\varepsilon/5)c + O(c^2)$, hence raw revenue rises by $(\Theta\varepsilon/5)c^3 + O(c^4) > 0$. Since the incumbent already breaks even, this lower quote is profitable. Thus this symmetric configuration is not protected in the primitive economy. $\square$

This diagnostic also explains why a common-price break-even derivative and strict deviations at fixed positive distances from the incumbent are insufficient. The profitable price movement shrinks with c , and its profit is of a higher order than the revenue retained in the exact game. Condition (D.2) removes the relevant both-home-active cell from the local lower-quote path; the proof of Lemma D.6 checks the two remaining cells. No conclusion about every alternative pair branch is inferred from the failure of this particular symmetric continuation.

D.10 Events and the limits of the strict continuation statements

The exact game uses weak no-entry and lower-own-quote inequalities and therefore includes several boundary configurations. The strict primitive continuations above stay a positive distance from the events at which their proof margins vanish. The reasons are distinct:

- At ϕ_P , entry of a targeted product into the pooled menu reaches the operating-cost threshold.
- At $\phi_U = 36/49$, the symmetric pair's near-incumbent lower-quote margin vanishes.
- At $\phi_W = 25/22$, the pooled lower-root branch reaches its revenue maximum, which also determines entry into the empty exact menu.
- At $\phi = 2/3$, the symmetric pair changes its type-2 support; the lower-side own-revenue slope also vanishes there.
- At $\phi_- = 576/841$, the expensive product on the lower unequal branch has a lower-own-quote tangency. At $\phi_0 = 25/36$, the lower and upper unequal support cells join and an own-quote margin vanishes.
- At $\phi_+ = 45/64$, the two reflected unequal vectors merge into the symmetric vector; the unequal revenue Jacobian degenerates. The stated symmetric continuation retains separation from that event as well.
- At $m^2[10 + \phi - 8p_s(\phi)] = 2$, the next-order middle-home support test is zero. Its sign is not determined by the leading quotient, and the local primitive pricing proof used here does not cover the event.

The all-equilibrium upper inclusion only needs separation from the three scope boundaries ϕ_P, ϕ_U, ϕ_W , because it uses compact closure and the exact scope table. Scope equality additionally needs the conditions that construct a representative of every exact scope. The unequal branch endpoints are excluded when continuing those particular fee vectors, but they need not be excluded merely to represent pair scope by the protected symmetric branch. These statements neither assign the exact quotient's numerical cutoffs to a fixed positive c , nor assert a unique shifted primitive cutoff at a boundary event.

E Fees, performance, and investor portfolios

This appendix supplies the arguments for Propositions 3–5 and Corollary 1. Exact accounting and gross-return identities hold in the primitive model. The coefficient and holdings tables solve the normalized problem in Appendix C. Their positive- c expansions use the branch, event and support conditions of Appendix B.9 and the continuation results in Appendix D. The uniform performance result below instead concerns every primitive equilibrium under the general compact conditions; it does not extend the detailed holdings tables to that larger domain.

E.1 Fund sizes and fees

Let $A_T = \text{AUM}_T$, $Q_0 = 4/(1 + 3m^2)$, and use the scales $\alpha = tc$, $Kc = \alpha/(1 - m^2)$ and $F = \Theta(m)c^2\phi$ from (D.1). The exact equilibrium accounting is

$$\delta_T A_T = F, \quad \frac{\delta_T}{\delta_U} = \frac{A_U}{A_T}, \quad \sum_{T \in \mathcal{E}} \delta_T A_T = |\mathcal{E}|F. \quad (\text{E.1})$$

The last equality counts the market fund as well as specialists. It measures operating expenditure recovered through fees, not a welfare loss net of the service supplied by the funds.

Proposition B.9 and Lemma 4 give

$$A_B = Q_0 + O(c), \quad \delta_B = \frac{\Theta(m)\phi}{Q_0}c^2 + O(c^3). \quad (\text{E.2})$$

For every active specialist, both A_T/c and δ_T/c have positive lower and finite upper bounds on the stated compact sets. Consequently $A_B/A_T = \delta_T/\delta_B$ is bounded above and below by positive multiples of c^{-1} . These bounds support the size and fee hierarchy in Proposition 3 and give its stronger positive- c scope. Positive normalized cost bounded away from zero is needed for the lower bounds.

On a specified smooth continuation, write its limiting specialist price as p_T and aggregate holding as Q_T . Then

$$\delta_T = \alpha p_T + O(c^2), \quad A_T = KcQ_T + O(c^2), \quad \frac{A_B}{A_T} = \frac{4(1 - m^2)}{mcQ_T} + O(1). \quad (\text{E.3})$$

The coefficients are collected below. The quotes z , p_s , ℓ and r are the branch maps in Appendix C; in the unequal cases, $\ell < r$.

Organization and specialist	Normalized fee	Normalized AUM
Pooled C	z	$15/2 - 99z/8$
Symmetric pair, $p_s \leq 1/3$	p_s	$4 - 6p_s$, each
Symmetric pair, $p_s \geq 1/3$	p_s	$3(1 - p_s)$, each
Unequal pair, lower support	(ℓ, r)	$(7/2 - 11\ell/2 + r, 3 + \ell - 4r)$
Unequal pair, upper support	(ℓ, r)	$(5r, 5\ell)$

These are alternative equilibrium organizations, not products all offered in one menu. Every row satisfies $p_T Q_T = \phi$. For an unequal pair, $A_L/A_R = r/\ell > 1$ in the normalized problem; exact break-even supplies the corresponding inverse ratio at primitive fees.

Comparison across organizations. At the same ϕ in the reduced coexistence region, the pooled fund is cheaper and larger than either targeted fund. To verify this, let $R_C(x) = x(15/2 - 99x/8)$. Its increasing segment ends at $10/33$, and $R_C(z) = \phi$. For a symmetric pair quote $p \leq 10/33$,

$$R_C(p) - \phi = p \left(\frac{7}{2} - \frac{51}{8}p \right) > 0,$$

so $z < p$. For $p > 10/33$, the same conclusion follows from $z \leq 10/33$. Each unequal quote is at least $9/29 > 10/33$, giving the remaining cases. Equal revenue converts the fee ordering into the AUM ordering. A comparison between continued primitive branches retains the continuation conditions and compact separation; no comparison of total pair AUM is implied here.

E.2 Performance and its benchmark

Use the units of Appendix A.1: the investor objective has $\mu = \lambda = 1$, with $\lambda = \gamma\sigma^2$. Fees and alphas are measured in units of the original μ . Let R_T be the gross excess return of a unit of fund T , and use the gross market-fund return R_B as benchmark. Define

$$\beta_{TB} = \frac{\text{Cov}(R_T, R_B)}{\text{Var}(R_B)}, \quad a_T^g = 1 - \beta_{TB}, \quad a_T^n = a_T^g - \delta_T.$$

The superscripts distinguish gross and net fund returns against the same gross benchmark. Multiplying the alphas and fees by μ recovers return units before normalization.

Lemma E.1 (Exact gross alphas). *Put $D = 1 + 3m^2 - c^2/2$. For any unit-sum portfolio w ,*

$$a^g(w) = \frac{2mc(w_2 - w_0) + c^2(w_0 + w_2 - 1/2)}{D}. \quad (\text{E.4})$$

In particular,

$$a_B^g = 0, \quad a_L^g = a_R^g = \frac{mc}{D}, \quad a_C^g = \frac{2mc/3 - c^2/6}{D}. \quad (\text{E.5})$$

For every admissible $0 < c \leq m$, both specialist gross alphas are positive, and $a_L^g = a_R^g > a_C^g$.

Proof. Write the loading perturbations as $s = (1, 0, -1, 0)'$, so asset i 's loading is $m + cs_i$. The correlation matrix satisfies

$$(\Sigma b_B)_i = \frac{1 + 3m^2 + 2mcs_i - c^2 s_i^2}{4}, \quad b_B' \Sigma b_B = \frac{D}{4}.$$

Substituting $w'\Sigma b_B$ in the definition of beta proves (E.4). The specified fund weights give (E.5). Positivity and the ranking follow from $a_C^g = c(4m - c)/(6D) > 0$ and $a_L^g - a_C^g = c(2m + c)/(6D) > 0$. The formula also applies to each home asset and each other product in the maintained catalogue. $\square$

The pair products have identical gross mean, variance and market beta by reflection, regardless of their equilibrium fees. Thus an unequal pair has $a_L^n - a_R^n = \delta_R - \delta_L$: the cheaper, larger member has higher net alpha. Every gross fund has the same expected return, so a higher fee lowers its raw expected net return even when its market-benchmark net alpha is positive.

Fee shares in the normalized equilibria. Leading gross alphas are $2\alpha/3$ for C and α for either pair member. The exact normalized correspondence gives

$$0 < z \leq \frac{10}{33}, \quad 0 < p_s \leq \frac{3}{7}, \quad p_T \leq \frac{5}{12} \quad \text{on unequal pair branches.}$$

The maximum fee shares of gross alpha are therefore $5/11$, $3/7$ and $5/12$, respectively. All are strictly below one half. These numerical caps are reduced-model statements. The next result uses the strict gap to establish the weaker half bound in the primitive economy.

Corollary E.2 (Specialist fees below half of gross alpha). *Maintain the general compact conditions in Appendix D.1: m and positive ϕ lie in compact subsets of $(0, 1)$ and $(0, \infty)$, and paid non-home access costs have a common positive floor. There is a common $\bar{c} > 0$ such that, in every primitive equilibrium with $0 < c < \bar{c}$, each active specialist satisfies*

$$0 < \delta_T < \frac{1}{2}a_T^g, \quad a_T^n > \frac{1}{2}a_T^g > 0, \quad a_T^g A_T > 2F. \quad (\text{E.6})$$

Its net alpha is bounded above and below by positive multiples of c .

Proof. If the fee bound failed uniformly, choose equilibria with $c_n \downarrow 0$ and active specialists T_n satisfying $\delta_{T_n} \geq a_{T_n}^g/2$. By Proposition D.8, pass to a subsequence with constant menu and label, convergent parameters and positive normalized prices $p_T = \lim \delta_{T_n}/\alpha_n$. The limit is one of the complete normalized equilibria. In particular its specialist labels belong to $\{C, L, R\}$. By (E.5), $a_{T_n}^g/\alpha_n$ tends uniformly to ρ_T , with $\rho_C = 2/3$ and $\rho_L = \rho_R = 1$. The supposed failure implies $p_T/\rho_T \geq 1/2$, whereas the complete fee correspondence gives $p_T/\rho_T \leq 5/11 < 1/2$, a contradiction. All relevant gross alphas are positive and of order c uniformly for sufficiently small c . The net-alpha bound follows by subtraction, and multiplication by $A_T > 0$, together with $\delta_T A_T = F$, gives the last inequality. $\square$

This corollary strengthens the approximate performance statement in Proposition 4. It applies to all primitive equilibria under the established closure theorem's conditions. It neither proves a new continuation nor identifies every primitive investor's support. The stronger reduced bounds, including $a_T^g A_T/F \geq 11/5$, are not asserted as exact positive- c bounds.

Net-alpha rankings across organizations. In the reduced coexistence region, either pair fund also has higher net alpha than the pooled fund. In units of α , their difference is $1/3 - p + z$, positive immediately if $p \leq 1/3$. For a symmetric $p \in [1/3, 3/7]$,

$$R_C(p - 1/3) - \phi = \frac{-75p^2 + 102p - 31}{8} < 0.$$

The polynomial is increasing over this interval and equals $-13/98$ at its right endpoint. Since both $p - 1/3$ and z lie on the increasing segment of R_C , it follows that $p - 1/3 < z$. For unequal pairs, $p - 1/3 \leq 1/12$, whereas $R_C(1/12) = 69/128 < 576/841 = \phi_- \leq \phi$, so $z > 1/12$. These comparisons concern alternative organizations at the same primitives, with the usual conditions for continued branches.

Active share, size and alpha across fund types. This paragraph proves Corollary 1. Measure active share against the market fund's weights, $AS_T = \frac{1}{2} \sum_j |b_{T,j} - \frac{1}{4}|$. The pooled fund omits asset 0, a difference of $1/4$, and overweights each other asset by $1/12$, so $AS_C = 1/4$. Each pair fund differs from b_B by $1/4$ in every coordinate, so $AS_L = AS_R = 1/2$, and $AS_B = 0$. Lemma E.1 gives $a_B^g = 0 < a_C^g < a_L^g = a_R^g$ at every admissible $0 < c \leq m$. Active share alone does not order gross alphas across the catalogue: by (E.4), the pairs 01 and 30 also have active share $1/2$ but gross alpha $-mc/D$, and the triple 301 has active share $1/4$ and gross alpha $-(2mc/3 + c^2/6)/D$. The equilibrium products rank alike because every equilibrium specialist omits asset 0 and holds asset 2. At a common ϕ in the reduced coexistence region, the comparison across organizations in Appendix E.1 gives $z < p_T$ for every targeted quote, symmetric or unequal. Equal revenue then gives $A_C > A_T$, and the preceding paragraph gives the net-alpha ranking. Proposition 3 places the market fund above every specialist in size and below it in fee. For an unequal pair, $a_L^g = a_R^g$ and $a_L^n - a_R^n = \delta_R - \delta_L$, so the smaller, more expensive member has the lower net alpha. The cross-type size and net-alpha comparisons are reduced-problem statements; continued primitive branches retain the continuation conditions and compact separation of Appendix B.9.

Comparative statics within an organization. The cross-type ranking holds the primitives fixed. Changing a primitive instead moves each fee along its branch. On the pooled branch and the symmetric-pair branch, reduced demand is $Q(p) = a - bp$, with $(a, b) = (15/2, 99/8)$ for C , and $(4, 6)$ below and $(3, 3)$ above the pair event $p = 1/3$ for each pair member. Break-even $pQ(p) = \phi$ gives $p'(\phi) = 1/(a - 2bp) > 0$ on the protected lower root. Since $\delta = \alpha p$, with $\alpha = mc/(1 + 3m^2)$, and $\phi = F/(\Theta(m)c^2)$, a lower F at fixed m and c lowers the fee and, since $A = KcQ(p)$ with Q decreasing in p , raises the fund's assets, leaving its gross alpha unchanged. At fixed m and F ,

$$\frac{d\delta}{dc} = \frac{\alpha}{c}(p - 2\phi p'(\phi)) = -\frac{\alpha}{c} \frac{ap}{a - 2bp} < 0,$$

so greater dispersion also lowers the fee and raises assets, while the gross alpha of each fixed portfolio rises: $da_L^g/dc = m(1 + 3m^2 + c^2/2)/D^2 > 0$ and $da_C^g/dc = [(1 + 3m^2)(2m - c) + mc^2]/(3D^2) > 0$ for $0 < c \leq m$. Within an organization, size and gross alpha can therefore rise together. These

statements exclude branch endpoints and the pair event. On the upper unequal-pair branch, $\ell + r = 3/4$ and $\phi = 5\ell r$, and greater dispersion at fixed F lowers the cheaper fund's fee and raises the other's: $d(\alpha\ell)/dc = -(\alpha/c)\ell(\ell + r)/(r - \ell)$ and $d(\alpha r)/dc = (\alpha/c)r(\ell + r)/(r - \ell)$. No single statement therefore covers every specialist fee. These are reduced-branch comparisons; continued primitive branches retain the conditions of Appendix B.9.

Changing the benchmark. A costless index with the fund's own weights gives gross alpha zero and net alpha $-\delta_T$. The measure omits the investor's cost of constructing that index. The frictionless minimum-variance benchmark also gives every unit-sum gross portfolio alpha zero: interiority implies $\Sigma w^* = (w^{*\prime} \Sigma w^*) \mathbf{1}$, hence every such portfolio has beta one against it. Positive alpha against B is therefore a statement about a particular covariance comparison, not skill.

Using net B as the benchmark instead gives $a_T^g - \delta_T + \beta_{TB}\delta_B$; the difference is of order c^2 in the local region and must be retained in statistics of that order. In particular, the market fund's net alpha is $-\delta_B$ against gross B , and zero against net B .

Equal and AUM weights. When specialists operate, their equally weighted mean net alpha is of order c , while the industry AUM-weighted mean is of order c^2 . For the latter include all $N = |\mathcal{E}|$ active funds and define

$$\bar{a}_A^n = \frac{\sum_{T \in \mathcal{E}} A_T a_T^n}{\sum_{T \in \mathcal{E}} A_T}.$$

On the maintained smooth continuations,

$$\bar{a}_A^n = \frac{\Theta(m)c^2}{Q_0} \left(\sum_{T \neq B} \rho_T Q_T - N\phi \right) + O(c^3). \quad (\text{E.7})$$

The denominator is total managed risky exposure, $Q_0 + O(c)$. For pooling the bracket is $Q_C(2/3 - 2z) > 0$. For a pair it is $\phi(1/p_L + 1/p_R - 3) > 0$, using $p_T \leq 3/7$. With B alone it is $-\phi$. This statistic is an alpha per unit of managed exposure; it is not an investor-welfare measure.

E.3 Portfolio implementation

The relevant counterfactual. The value of adding a product is evaluated against the investor's optimum without it. Other positions must be reoptimized, including free home investment. The following precise rule applies on a fixed accessible instrument set, in particular in the maintained region without paid non-home access. All offered fund fees are positive; implementation ties are handled explicitly rather than interpreted as strict purchase rules.

Lemma E.3 (Adding a fund to the remaining menu). *Remove candidate fund T , hold other quotes fixed, and let e_i^{-T} be investor i 's optimal risky exposure from the remaining menu and accessible direct assets. Define*

$$v_{iT}^{-T} = 1 - b_T' \Sigma e_i^{-T}. \quad (\text{E.8})$$

If $v_{iT}^{-T} > \delta_T$, every optimum with the full menu uses T positively. If $v_{iT}^{-T} < \delta_T$, every optimum uses zero T . At equality, the optimum without T remains optimal; an alternative tied implementation may also use it.

Proof. The fixed-instrument objective is concave in nonnegative holdings and strictly concave in final exposure. At the optimum excluding T , the derivative in its newly available nonnegative holding is $1 - \delta_T - b'_T \Sigma e_i^{-T}$. A positive derivative is a feasible strict improvement. Any full-menu optimum omitting T would also solve the restricted problem. Strict concavity in exposure would give it exposure e_i^{-T} and the same positive derivative, a contradiction. A nonpositive derivative, together with the existing instruments' optimality conditions, supplies all full-menu KKT conditions and makes the excluded optimum full-menu optimal. If the derivative is strictly negative, the concavity inequality gives a strict value loss for any feasible candidate with positive T holding. At equality, any other optimum must have the same exposure and total fee bill, so positive T holdings can arise only through an equivalent implementation. $\square$

If $M_i^{-T} = \mathbf{1}' e_i^{-T} > 0$, write $w_i^{-T} = e_i^{-T} / M_i^{-T}$ and let $\bar{\delta}_i^{-T}$ be its fee bill per unit of risky exposure. Optimal scaling of the remaining holdings gives $1 - \bar{\delta}_i^{-T} = M_i^{-T} (w_i^{-T})' \Sigma w_i^{-T}$. Thus

$$v_{iT}^{-T} = 1 - \beta_{T, w_i^{-T}} (1 - \bar{\delta}_i^{-T}).$$

This is gross T 's alpha against the remaining portfolio's net return. It is not generally alpha against B , against B plus home alone, or against the unreoptimized remainder of the final portfolio. Applied to a held instrument at the full-menu optimum, the same scaling argument gives zero net alpha against that investor's own net portfolio. The benchmark is investor-specific.

In the reduced problem, the corresponding marginal value in units of α is $\rho_T - d'_T y_i^{-T}$. For type 2 with no specialists, $y_i^0 = (8/3)H_2$, giving $4/9$ for C and $1/3$ for either pair fund. With R still available at fee $p < 1/3$, removing L instead gives $u_R = 2(1 - 3p)$, $h_2 = 2(1 + p)$ and $v_{2L}^{-L}/\alpha = (1 - p)/2$. It exceeds p exactly when $p < 1/3$. At $p \geq 1/3$, the remaining optimum uses home alone for the correction and the marginal value is $1/3$.

The remaining menu also matters for middle investors. For type 1, removing L while leaving R at $p < 1$ gives $u_R = 6(1 - p)$, $h_1 = 2(1 - p)$ and $v_{1L}^{-L}/\alpha = (1 + p)/2$. A home asset with zero leading market alpha can therefore enter this reoptimized portfolio to offset the remaining fund's exposure. Its absence from the final symmetric portfolio does not make the no-specialist benchmark the relevant counterfactual for adding one of the two funds.

Holdings at symmetric prices. The next table gives specialist and home quantities in units of Kc . Every type also holds the order-one market-fund position. Under pooling, the specialist column is u_C ; under targeting it is the common holding in each of L and R .

Organization	Types 0, 1 and 3		Type 2	
	Specialist	Home	Specialist	Home
B only	0	0	0	8/3
$B + C$	$8 - 12z$	0	$6 - 27z/2$	$2 + 3z/2$
Pair, $p < 1/3$	$4(1 - p)$	0	$4 - 12p$	$8p$
Pair, $p \geq 1/3$	$4(1 - p)$	0	0	8/3

The formulas follow by substitution in the KKT system of (C.15). Pooling has $z \leq 10/33 < 4/9$, so every type buys C . The pair event $p = 1/3$ corresponds to $\phi = 2/3$; type 2's specialist holding is zero at that event. Zero home holdings of middle types at symmetric pair fees are exact in the reduced problem. Their exact persistence in the primitive economy requires the support condition in Appendix B.9.

Unequal pair prices. Orient prices so L is cheaper, $\ell < r$. On both unequal branches, types 0 and 1 have $(u_L, u_R, h_i) = (4(1 - \ell), 4(1 - r), 0)$. Type 3 has

$$(u_L, u_R, h_3) = (4(1 - 2\ell + r), 4(1 + \ell - 2r), 4(r - \ell)).$$

On the lower-support branch, type 2 has

$$(u_L, u_R, h_2) = (2(1 - 3\ell), 0, 2(1 + \ell)), \quad 9/29 \leq \ell \leq 1/3.$$

On the upper-support branch she has $(0, 0, 8/3)$. These are the cells in the proof of Proposition C.5; the other orientation follows by reflection. They show why the symmetric holdings table is not an all-pair clientele theorem: a middle investor can hold home, and type 2 can buy only the cheaper fund even when $\phi > 2/3$. At joins use the common boundary quantities; exact positive- c supports retain their branch-specific qualifications.

Direct exposure across organizations. Define the direct share as aggregate directly held risky exposure divided by aggregate total risky exposure. On the maintained continuations there is no paid non-home access. With h_i denoting normalized home holdings, its leading expression is

$$\mathcal{D}_{\mathcal{E}} = \frac{\sum_i x_{ii}/4}{\sum_i \mathbf{1}'e_i/4} = \frac{mc}{1 - m^2} \chi_{\mathcal{E}} + O(c^2), \quad \chi_{\mathcal{E}} = \frac{1}{16} \sum_i h_i. \quad (\text{E.9})$$

The denominator is $Q_0 + O(c)$, which explains the factor $1/16$. The coefficients are

Organization	$\chi\mathcal{E}$
B	$1/6$
$B + C$	$(4 + 3z)/32$
Symmetric pair, $p \leq 1/3$	$p/2$
Symmetric pair, $p \geq 1/3$	$1/6$
Unequal pair, lower support	$(1 - \ell + 2r)/8$
Unequal pair, upper support	$1/6 + (r - \ell)/4$

Both unequal coefficients exceed $1/6$ on their strict branches. For the lower one this follows from $r \geq 3/8$ and $\ell \leq 1/3$; for the upper one it follows from $r > \ell$.

In the symmetric coexistence interval $2/3 < \phi < \phi_U$, the pair coefficient is $1/6$, while the pooled coefficient is strictly smaller because $z < 4/9$. Hence the organization with three funds can leave more exposure invested directly than the one with two funds. Both are equilibrium organizations at the same primitives; B alone in this interval is only a reference implementation.

The crossing below $2/3$ can also be located analytically. Set

$$p_{\text{dir}} = \frac{53 - \sqrt{41}}{173}, \quad \phi_{\text{dir}} = 4p_{\text{dir}} - 6p_{\text{dir}}^2 \simeq 0.6421005.$$

Equality of the pooled and low-pair coefficients requires $z = (16p - 4)/3$. Substituting both break-even equations gives

$$R_C((16p - 4)/3) - (4p - 6p^2) = -2(173p^2 - 106p + 16).$$

Only the displayed root lies in $(1/4, 1/3)$; the other exceeds $1/3$. For $p \leq 1/4$, the pair coefficient is already below the pooled one. On the increasing revenue segment the polynomial's sign gives the comparison while the candidate $z = (16p - 4)/3$ is on that segment; once it exceeds $10/33$, it exceeds every pooled equilibrium quote directly. Thus, within coexistence, the pair has less direct exposure below ϕ_{dir} and more above it. Continued-branch rankings require separation from the crossing as well as the existing continuation conditions. These are comparisons across equilibria, not causal entry effects or welfare rankings.

E.4 Construction of the illustrations

Figures 6 and 7 illustrate the small-dispersion formulas at $m = 2/5$, $c = 1/20$, and $\phi = 1/2$. Both the pooled and symmetric targeted organizations are equilibria of the reduced problem at this normalized operating cost. The chosen m and ϕ also satisfy the symmetric-pair support condition in (B.20). The figures compare those alternative organizations. They do not report a calibrated economy or an exact equilibrium computation at positive c .

Use $Q_0 = 4/(1 + 3m^2)$, $\alpha = mc/(1 + 3m^2)$, $Kc = \alpha/(1 - m^2)$, and $F = \alpha Kc\phi$. The branch maps give $z = (10 - 2\sqrt{14})/33$ for pooling and $p = 1/6$ for either targeted fund. This pair fee is

below the diversifier's participation cutoff $1/3$. The plotted quantities are

	B	C	L or R
AUM	Q_0	$Kc(15/2 - 99z/8)$	$Kc(4 - 6p)$
fee	F/Q_0	αz	αp

The first figure divides AUM by $Q_0/100$ and fees by F/Q_0 . Thus the market fund has AUM 100 and fee 1 in each organization. Each plotted fee–AUM product recovers the common operating bill. The market-fund size is its leading approximation; organization-specific higher-order corrections are omitted consistently with (E.2)–(E.3).

The second figure measures alpha in units of α . Against gross B , the pooled fund's gross and net alphas are $2/3$ and $2/3 - z$, and either targeted fund's are 1 and $1 - p$. Against each fund's own costless portfolio replica, gross alpha is zero and net alpha is the negative of its fee: $-z$ or $-p$ in these units. These are the leading specialist quantities used in the body; the exact gross-return formulas appear in Lemma E.1.

F Numerical protocol and results

This appendix documents the computations reported in Section 5: what was solved and how, the parameter sources, the complete outcomes at the empirical anchor, the accuracy of the small-dispersion formulas, the access-cost sensitivity, and how the computed economy approaches the small-dispersion approximation. Every statement concerns the declared grids and the finite searches described here.

F.1 What is computed

Economy. Asset j has loading a_j on the common factor, with $a = (m + c, m, m - c, m)$, unit variance and covariance matrix $\Sigma = aa^\top + \text{diag}(1 - a_j^2)$. Expected excess returns and risk aversion are normalized, $\mu = \lambda = 1$. The four investors have equal mass $\frac{1}{4}$. Investor i holds asset i at no cost and pays $\hat{\tau} d(i, j)$ to hold asset $j \neq i$ directly, where the circle distance d is $\frac{1}{4}$ between neighbors and $\frac{1}{2}$ across the circle. Given fees, she chooses a set of directly held assets and nonnegative holdings of the offered funds and of those assets to maximize expected return minus half the variance, minus fees and access costs. Every fund has the operating cost $F = \phi \Theta(m)c^2$.

Equilibrium tests. A menu with fees is a computed equilibrium when every active fund recovers F and no deviation of Definition 1 is profitable: no active fund gains by lowering its own quote, and no inactive product earns more than F at any quote. Demand ties across access sets (values within 10^{-9}) are resolved in favor of the deviating product when an entrant is tested, and by the incumbent allocation otherwise. Implementation ties within an access set are not optimized.

Search. At every parameter point the computation covers all 512 subsets of the nine products through 288 representatives under the circle's reflection. For each menu it searches for break-even

fee vectors from 9 starting points (5 for menus of five or more products). A root is accepted at relative residual 2×10^{-6} ; revenue must equal F within 2×10^{-5} relative or 2×10^{-9} absolute. For every root it scans own lower quotes and entrant quotes on hybrid grids (8-point screens, 80-point dense grids, local refinement, special points at incumbent fees and a one-sided step of 10^{-8}). A root is *computed* if no deviation is profitable beyond the revenue tolerance, *unresolved* if the most profitable deviation is positive but within a small band (the larger of ten times the tolerance and 0.2 percent of F), and *rejected* otherwise. Investors' conditional portfolio problems are solved by nonnegative least squares with a ridge of 10^{-10} .

Bounded rescue. When a menu of Proposition 2 has no computed equilibrium at a point, and whenever another menu has one, the menu is solved again at three ridge values (10^{-8} , 10^{-10} , 10^{-12}), with twice as many starts, twice as dense quote grids, seeds from neighboring points at the same m and access cost, and a small-dispersion seed. An equilibrium outside the three organizations counts as confirmed if at least two of the three variants reproduce it. Production and rescue results are reported separately.

Units. Fees are in units of the expected excess return. Annual basis points are $10^4 \times 0.045 \times \hat{\delta}$ for a normalized fee $\hat{\delta}$; the same factor converts alphas. The premium does not enter the equilibrium.

Sample. The computations come from one locked archive (August 10, 2026). The anchor covers $m = .49$, access cost .05, $c \in \{.12, .15, .16\}$ and twelve values of ϕ (.05, .10, .12, .16, .40, .65, .72, .78, 1.00, 1.10, 1.18 and 1.5): 36 points and 10,368 menu problems, which represent 18,432 menus. The archive also contains refined transition brackets, canonical-menu sensitivity blocks at access costs .02 and .10, and the same search at $c \in \{.01, .03, .06, .10\}$, which Section F.4 uses only to compare the computed economy with the approximation. The archive reports no solver error. A separate archive contains exploratory computations with the same code at larger dispersion and at other average loadings, outside the range that the evidence supports. They are not part of the exercise reported here. Those exploratory runs also produce other organizations, equilibria without a market fund, and pooled funds larger than the market fund.

F.2 Parameters and sources

Covariance. The model's average pairwise correlation is $m^2 - c^2/6$. Pollet and Wilson (2010) report an average correlation of .237 among the 500 largest U.S. stocks, 1963–2006. Within each volatility quintile of the 25 volatility-and-correlation portfolios of Asness et al. (2020), Table 2, Panel A (1930–2015), the ratio $\sqrt{2}\text{SD}(\beta)/\text{mean}(\beta)$ of realized betas gives a value of c/m . The five quintiles imply c between .1215 and .1634 when the correlation moment is matched, with m between .489 and .491. The grid tests the representative rounded values .12, .15 and .16; it does not cover the whole interval, whose top lies above .16. Loadings are a covariance summary, not estimates of individual stocks' loadings in a four-asset economy.

Access cost. Under a local constant-absolute-risk-aversion approximation, the normalized access cost is $\hat{\tau} = (\gamma\sigma^2/\mu^2)(k/W)/L$, with relative risk aversion γ , return volatility σ , premium μ , recurring participation cost k relative to wealth W , and circle distance L of the access. The inputs of [Gálvez and Paz-Pardo \(2026\)](#) (risk aversion 6.83, a recurring cost of .0018 of income, a wealth-to-income ratio of 5.599 and volatility .157), with the premium of 4.5 percent, give $\hat{\tau} = .0267, .0535$ and $.1069$ for $L = 1, .5$ and $.25$. Setting the model’s market-portfolio volatility, rather than an asset’s volatility, equal to σ raises these values by the factor $1/.4273$ at $m = .49, c = .15$, to $.0626, .1251$ and $.2502$. The baseline .05 is therefore an illustrative anchor within these conventions, and .02 and .10 are sensitivity values; neither mapping identifies the access cost independently of the conventions. The source’s own premium is 4 percent; the substitution of 4.5 percent is mine.

Premium and comparisons. The 4.5 percent premium for annual units is close to the long-horizon premium of large asset managers in [Dahlquist and Ibert \(2024\)](#), Table 1 (median 4.40 percent), and to the one-year option-implied lower bound of [Martin \(2017\)](#), Table I (median 4.35 percent). Two further studies corroborate the range of risk aversion and participation costs without entering the mapping: [Choukhmane and de Silva \(2026\)](#) estimate risk aversion of 2.54 and a portfolio-adjustment cost of \$156 in their baseline specification, and [Bacher \(2025\)](#) reports risk aversion of about 5. Fee comparisons are external: [Khomyn et al. \(2024\)](#) report a median exchange-traded-fund expense ratio of 18 basis points (mean 25.35), and [Hitzemann et al. \(2022\)](#) relate mutual-fund fees to market beta. Neither calibrates the operating cost F , which is varied.

Maintained structure. Four assets on a circle, equal expected returns and variances, equal investor masses, the nine equal-weight products and a common operating cost are structural assumptions of the model, not calibrated quantities.

F.3 Results at the empirical anchor

The 36 anchor points have 49 computed equilibrium roots: 3 with the market fund alone, 24 pooled and 22 targeted. Every point has at least one, and 12 points have both pooled and targeted equilibria; Figure 8 shows the organizations at each point. Every product breaks even to within $3 \times 10^{-10}F$, so the cheaper fund is always the larger one. Every computed equilibrium contains the market fund, which is larger and cheaper than every specialist; the ratio of its assets to a specialist’s ranges from 4.1 to 25.3. Every specialist has positive alpha after fees against the market fund’s return before fees, from 5.7 to 20.4 basis points a year, and keeps at least 59 percent of its alpha before fees. No investor pays for direct access in any computed equilibrium.

Other menus and unresolved roots. The search found 1,798 break-even roots for other nonempty menus and rejected all of them: 1,180 because an entrant profits and 618 because an incumbent gains by lowering its fee. No break-even root was found for any other menu with six or more products. The closest rejected roots are those of the market fund with the targeted pair and the pooled fund at $\phi = .40$, where a fund gains 1.3 to 1.6 percent of F by lowering its fee. Three targeted roots with unequal pair fees, at $\phi = .72$ for each c , are unresolved: their most profitable

deviations earn between 0.006 and 0.07 percent of F , inside the band in which the protocol classifies neither way, and each coexists with a computed symmetric targeted equilibrium at its point. One computed targeted equilibrium, at $c = .12$ and $\phi = .65$, has pair fees that differ by 0.4 percent. The bounded rescue was applied to each of the 60 cases in which one of the three organizations has no computed equilibrium at a point; none of its 180 variants found one.

Transitions. Table F.1 reports the refined transition brackets. They lie above the cutoffs of Proposition 2 by .011 to .017 (lower), .027 (middle) and .060 to .082 (upper). At all three dispersions the upper limit of the targeted organization is set by a pooled entrant that covers its cost against the pair. Within the scanned range the refined search finds no cost at which a pair fund gains by cutting its own fee, the pricing limit that defines ϕ_U in the approximation.

Table F.1: **Transition brackets at the anchor** ($m = .49$, access cost .05). Normalized operating costs at which a targeted entrant stops covering its cost against the pooled fund (lower), at which the targeted organization stops being an equilibrium (middle), and at which a pooled entrant stops covering its cost against the market fund alone (upper). Midpoints of computed brackets of width below .01, from refined searches over the menus of the three organizations. The first column gives the cutoffs ϕ_P , ϕ_U and ϕ_W of Proposition 2.

	Approximation	$c = .12$	$c = .15$	$c = .16$
Lower	.140	.151	.157	.157
Middle	.735	.762	.762	.762
Upper	1.136	1.196	1.213	1.218

Portfolios. Pooled equilibria reproduce the holdings pattern of Proposition 5 in all 24 roots. In symmetric targeted equilibria, the middle investors hold small direct positions in their home assets, at most 0.38 percent of risky exposure, at every tested cost up to .65 for $c = .12$ and up to .40 for $c = .15$ and .16; the high-loading investor holds none. The condition of Appendix B.9 that keeps middle investors out of their home assets at equal pair fees holds at $m = .49$ only for costs above about .656. It is a sufficient condition: its failure permits such holdings but does not require them, and at $\phi = .65$ with $c \geq .15$ they are zero. The diversifier investor holds her home asset at every tested targeted cost except the lowest, .05, where she holds no direct position and the same fund portfolio as the high-loading investor. She buys neither pair fund at $\phi = .72$, nor at .65 for $c = .15$ and .16; in the approximation she stops buying the pair only above $\phi = 2/3$.

Accuracy. Table F.2 reports the errors of the small-dispersion formulas over the computed production roots of the pooled and targeted organizations, including the one targeted root with unequal fees. Errors are measured relative to the formula. The market-fund fee formula F/Q_0 uses the market fund's assets in the limit, $Q_0 = 4/(1 + 3m^2)$; a different approximation, F divided by the approximate assets of the computed organization, tracks the market fund's assets instead, whose errors the table also reports. At $\phi = 1.18$ the approximation has no pooled counterpart, so those

roots are excluded. In every pooled root the computed specialist fee lies below its formula and the computed specialist assets lie above theirs.

Table F.2: **Accuracy of the small-dispersion formulas at the anchor** ($m = .49$, access cost .05). Absolute error relative to the formula, median / maximum, in percent, over computed production roots. The last column also covers the three market-only roots.

c	Targeted organization		Pooled organization			All
	Specialist fee	Market-fund assets	Specialist fee	Specialist assets	Market-fund assets	Market-fund fee F/Q_0
.12	0.2 / 5.3	1.0 / 1.2	7.1 / 15.1	7.7 / 17.8	0.2 / 0.9	14.6 / 17.3
.15	0.2 / 10.6	1.7 / 2.3	8.9 / 18.3	9.8 / 22.5	0.4 / 1.5	19.2 / 22.5
.16	0.3 / 10.6	2.0 / 2.6	9.5 / 19.4	10.5 / 24.1	0.4 / 1.7	20.8 / 24.4

Access-cost sensitivity. The canonical-menu blocks recompute only the menus of the three organizations, at access costs .02 and .10, $c \in \{.15, .16\}$ and $\phi \in \{.10, .40, .72, 1.00, 1.18\}$, without rescue. They can show that an organization persists, not that others are absent. Each of the 20 points has a computed equilibrium, and in all 30 computed roots the market fund is larger and cheaper than every specialist. At .10 the organizations match those at .05 at every point; at .02 they match at seven of the ten points. The market fund alone is also computed at $(c, \phi) = (.15, 1.18)$ and $(.16, 1.00)$, and at $(.16, 1.18)$ it is the only organization found: no break-even root was found for the pooled menu there, a nonfinding in an unrescued block rather than a demonstrated disappearance of pooling. In each of these three market-only equilibria the two middle investors pay for direct access to the low-loading asset, the only paid access in the archive. Five targeted roots with unequal pair fees at $\phi = .72$ are unresolved, with margins below 0.08 percent of F .

F.4 Approach to the small-dispersion approximation

The archive also contains the full search at $c = .01, .03, .06$ and .10, with the same average loading, access cost and operating costs. These dispersions lie below the range that the evidence suggests and are not part of the empirical exercise; they show the computed economy approaching the small-dispersion approximation. At each of these 48 points the search finds at least one equilibrium, and every computed equilibrium belongs to the three organizations. At $c = .01, .03$ and .06 the computed organizations coincide with those of Proposition 2 at all twelve costs; at $c = .10$ they differ only at $\phi = 1.18$, as at the anchor. At $c = .01$ the transition brackets lie at .140, .746 and 1.142, against the cutoffs .140, .735 and 1.136, and the upper limit of the targeted organization falls in the same bracket as the cost at which a pair fund gains by cutting its own fee, as in the approximation. The median error of the specialist-fee formula at $c = .01$ is 0.1 percent for the targeted and 0.6 percent for the pooled organization. For the pooled organization it rises to 1.7, 3.5 and 5.9 percent at $c = .03, .06$ and .10, and to between 7.1 and 9.5 percent at the anchor.

F.5 Replication files

The replication files contain the locked archive with its input manifests and checksums, the configuration of every search, and the scripts that produce each table and figure in Section 5 and this appendix from the archive.

G A specialist operating-cost premium

This appendix proves Proposition 6 and the statements in Section 6.1. The market fund B costs $F_B > 0$ to operate. Every non-market product $T \in \mathcal{M} \setminus \{B\}$, whether active, a lower-quote copy of an active product or an omitted entrant, costs

$$F_T = F_S = \kappa_F F_B, \quad 1 \leq \kappa_F < \infty, \quad (\text{G.1})$$

with one premium κ_F for all eight products. Returns, access costs, standardized weights, the investor's problem and the timing are those of Section 2. The unit map (A.2) scales F_B and F_S by the same factor, so κ_F is unit-free. The normalized costs are

$$\phi_S = \frac{F_S}{\Theta(m)c^2}, \quad \phi_B = \frac{F_B}{\Theta(m)c^2} = \frac{\phi_S}{\kappa_F}, \quad (\text{G.2})$$

with Θ as in (16). Uniform statements take $\kappa_F \in [1, \bar{\kappa}_F]$ for a finite $\bar{\kappa}_F$.

Definition G.1 (Equilibrium with a specialist premium). *An equilibrium is defined as in Definition 1, with each product's own operating cost in its tests. Every active product T has positive AUM and satisfies $\delta_T \text{AUM}_T = F_T$, and no lower quote (10) earns more than F_T . Every omitted product U satisfies $\mathcal{R}_U \leq F_U$ in (11). The tie and focal-upper conventions of Definition A.5 are unchanged.*

Investors never pay operating costs directly, so the premium enters only through break-even, lower-quote and entry comparisons. Each comparison involves one product's cost: F_B when the product is B , and F_S otherwise. The proofs below list only the steps of the original arguments in which this distinction matters. Every other step is unchanged, with F read as the cost of the product being tested.

G.1 Exact accounting

Lemma G.2 (Cost-adjusted accounting). *In every equilibrium of Definition G.1, $\delta_T \text{AUM}_T = F_T$ for every active T , so $\delta_T/F_T = 1/\text{AUM}_T$. For active non-market products T and U , $\delta_T/\delta_U = \text{AUM}_U/\text{AUM}_T$. For every active specialist S ,*

$$\frac{\delta_S}{\delta_B} = \kappa_F \frac{\text{AUM}_B}{\text{AUM}_S}. \quad (\text{G.3})$$

With B and N_S specialists active, total fee revenue is $F_B + N_S \kappa_F F_B$.

Proof. Positive AUM and $F_T > 0$ give $\delta_T = F_T / \text{AUM}_T$. Dividing break-even equations gives the ratios, as in Lemma A.7. No limit or regularity condition is used. $\square$

The two ratios among and across tiers use the common premium. The cost-adjusted identity $\delta_T / F_T = 1 / \text{AUM}_T$ holds for each product separately.

G.2 Market-fund selection, rates and scale

Proposition G.3 (Selection with a specialist premium). *Fix $m \in (0, 1)$, $\tau > 0$ and $\bar{\kappa}_F \geq 1$. There are positive ceilings $\bar{c}$ and $\bar{F}$ such that, for every $0 < c < \bar{c}$, $0 < F_B < \bar{F}$ and $\kappa_F \in [1, \bar{\kappa}_F]$, every equilibrium contains B , held by every type with $q_{iB} \geq Q_0(m)/2$, and no investor pays for non-home access. The ceilings are independent of each other and may depend on $\bar{\kappa}_F$. They can be chosen uniformly when m ranges over a compact subset of $(0, 1)$ and non-home access costs have a common positive floor.*

Proof. Follow the proof of Proposition 1 in Appendix B.6, along equilibrium sequences with $c_n \rightarrow 0$, $F_{B,n} \rightarrow 0$ and $\kappa_{F,n} \in [1, \bar{\kappa}_F]$. Pass to a subsequence with a constant menu, convergent holdings and $\kappa_{F,n} \rightarrow \kappa_F^0 \in [1, \bar{\kappa}_F]$. The component lemmas change as follows.

Lemma B.5 is unchanged. Its contradictions use only market-fund entry at a fixed quote and the market fund's own lower quote, both compared with $F_{B,n}$.

In Lemma B.6, normalize prices by $F_{B,n}/Q_0$. Break-even of a non-market product gives

$$\frac{Q_0 \delta_{T,n}}{F_{B,n}} = \kappa_{F,n} \frac{Q_0}{A_{T,n}} \longrightarrow \beta_T := \frac{\kappa_F^0}{a_T} \quad (T \in \mathcal{O}), \quad \frac{Q_0 \delta_{T,n}}{F_{B,n}} \longrightarrow \infty \quad (T \text{ active, } A_T^0 = 0),$$

the second limit because $\kappa_{F,n} \geq 1$. This replaces (B.7); the market fund's own limit is unchanged. The lemma's conclusions use only these limits, the positivity of β_T , exact replication and the market fund's own quotes, so its proof is unchanged.

In Lemma B.7, identity (B.9) becomes

$$\frac{1}{4} \sum_{i=0}^3 \pi_i = \sum_{T \in \mathcal{O}} \beta_T a_T = \kappa_F^0 N \geq N. \quad (\text{G.4})$$

The case $N = 1$ is excluded by Lemma B.1, as before. For $N = 2$, every type has $\pi_i = 2\kappa_F^0 \geq 2 > 3/2$, so the quote coefficient $p = 3/2$ earns normalized revenue $3/2 > 1$ against B 's own cost. For $N \geq 3$, Lemma B.3 and (G.4) give $j\pi_{(j)}/4 \geq \kappa_F^0 N/H_4 \geq N/H_4$, and the coefficient $p = \frac{3}{4}\pi_{(j)}$ earns at least $9N/25 \geq 27/25 > 1$. Every margin is at least as large as in the original proof.

In Lemma B.8, the steps that use B 's lower quotes, including (B.10), involve only $F_{B,n}$ and are unchanged. The final accounting step becomes

$$\kappa_F^0 N = \frac{1}{4} \sum_{j=1}^{4-r} \pi_{(j)} \leq H_4 - H_r.$$

Since $\kappa_F^0 N \geq N \geq 1$, the cases $r = 2$ and $r = 3$ are impossible, and $r = 1$ forces $N = 1$, which Lemma B.1 excludes. Hence $r = 4$.

The assembly argument of Appendix B.6 applies verbatim. A counterexample sequence also carries $\kappa_{F,n}$, and compactness of $[1, \bar{\kappa}_F]$ supplies the convergent subsequence; this is the only step at which the ceilings may depend on $\bar{\kappa}_F$. Uniformity in m follows as in that appendix. $\square$

Proposition G.4 (Rates, necessary viability and scale). *After reducing the ceilings of Proposition G.3 if necessary, every equilibrium satisfies*

$$q_{iB} = Q_0 + O(c + F_B), \quad \text{AUM}_B = Q_0 + O(c + F_B), \quad \delta_B = \frac{F_B}{Q_0} + O(F_B(c + F_B)),$$

home holdings are $O(c + F_B)$, and every active non-market product satisfies

$$F_S \leq C_{\text{viab}} c^2, \quad \delta_T = O(c), \quad \text{AUM}_T = O(c). \quad (\text{G.5})$$

The constant C_{viab} does not involve κ_F . If, in addition, ϕ_S lies in a compact set $\Phi \subset (0, \infty)$, then for all sufficiently small c every active specialist has AUM_T/c and δ_T/c bounded above and away from zero. These bounds are uniform for $\kappa_F \in [1, \bar{\kappa}_F]$, m in a compact subset of $(0, 1)$ and a common positive access-cost floor.

Proof. In the proof of Proposition B.9, the ratio $\delta_T/\delta_B = \kappa_F \text{AUM}_B / \text{AUM}_T \geq \text{AUM}_B / \text{AUM}_T$ still tends to infinity for every active non-market product along a sequence tending to the origin. Specialist fees therefore eventually exceed $2\delta_B$. The cone bound, the residual inequality and the market fund's first-order condition then give the $O(c + F_B)$ bounds verbatim. For an active non-market product, break-even and the same fee and quantity bounds give

$$F_S = \delta_T \text{AUM}_T \leq K(c + F_B)^2 \leq 2Kc^2 + 2KF_B^2 \leq 2Kc^2 + 2KF_B F_S,$$

using $F_B \leq F_S$. Reducing $\bar{F}$ to at most $1/(4K)$ gives $F_S \leq 4Kc^2 =: C_{\text{viab}} c^2$. Hence also $F_B \leq 4Kc^2$, and the $O(c)$ bounds follow. The constant K comes from the original bounds and does not involve κ_F .

For the scale statement, follow the proof of Lemma 4. With $F_S = \Theta(m)c^2\phi_S$ and $F_B = F_S/\kappa_F \leq F_S$, both costs lie below the ceilings for all sufficiently small c . Break-even $\delta_T \text{AUM}_T = \Theta(m)c^2\phi_S$ together with $\text{AUM}_T, \delta_T \leq Cc$ gives the lower bounds, uniformly on the stated sets. $\square$

The two-sided bounds use the compact positive-normalized-cost condition. On the independent (c, F_B) rectangle alone, only the upper $O(c)$ bounds in (G.5) are available.

Proposition G.5 (Market-fund-only existence at fixed costs). *On the region (B.16) of Proposition B.10, with F read as F_B , a market-fund-only equilibrium exists for every $\kappa_F \geq 1$, with the same constants. Within the region of Propositions G.3 and G.4, every equilibrium is market-fund-only once $C_{\text{viab}} c^2 < \kappa_F F_B$.*

Proof. The construction in Proposition B.10, through Proposition A.11 and the protection of B 's lower quotes, uses only F_B . Its entry bounds show that every entrant's revenue is below F_B on

the region. Since $F_U = \kappa_F F_B \geq F_B$, entry protection holds a fortiori for every $\kappa_F \geq 1$. The last statement follows from (G.5). $\square$

Costs fixed in levels. The local results below hold ϕ_S and κ_F fixed as c falls, so both costs fall like c^2 . Two experiments with costs fixed in levels are covered. First, with $F_B > 0$ and a finite common κ_F fixed inside the cost rectangle of Propositions G.3 and G.4, $F_S = \kappa_F F_B$ is fixed and eventually exceeds $C_{\text{viab}} c^2$; Proposition G.5 then excludes all specialists for small c . Second, and more generally, specialists are excluded along any family that stays inside that domain with $F_S/c^2 \rightarrow \infty$. A premium fixed in levels while $F_B \rightarrow 0$, so that $\kappa_F \rightarrow \infty$, is not covered: the selection and rate constants above are not established uniformly over unbounded κ_F .

G.3 The reduced problem at the specialist cost

Proposition G.6 (Reduced classification at ϕ_S). *The normalized completion problem of Definitions C.1 and C.2, at $\phi = \phi_S$, is the reduced counterpart of the economy with a specialist premium. Consequently Proposition C.7 and the scope table (C.53) hold with ϕ replaced by ϕ_S . The cutoffs ϕ_P , ϕ_U and ϕ_W , the pooled, symmetric and unequal pair branches, and all events are unchanged in that coordinate.*

Proof. The exact investor problem (C.15) contains no operating cost. In Definition C.2, every break-even, lower-quote and entry comparison concerns a non-market product and uses one number. The market fund has no pricing equation there, because its fee is of order c^2 (Appendix C.1). Under the premium, (D.8) gives each specialist's normalized cost as $F_S/(\Theta c^2) = \phi_S$. The definition, and hence the exhaustive classification, is unchanged at ϕ_S . $\square$

Fixed primitives. At fixed (m, τ, c, F_B) , raising κ_F raises $\phi_S = \kappa_F F_B/(\Theta c^2)$. In the reduced correspondence this has the same effect on specialist viability, and hence on the set of organizations, as lowering c by the factor $\sqrt{\kappa_F}$; fees and holdings in original units do not scale in the same way. The scope set moves toward fewer specialists, but the overlap in (C.53) admits several organizations, so there is no unique transition path. The leading market fee F_B/Q_0 does not change; the exact fee F_B/AUM_B varies with the organization at finite c .

G.4 Positive dispersion: exclusion and continuation

Write $\mathcal{S}_c(m, \phi_S; \kappa_F)$ for the set of non-market scopes that occur in primitive equilibria and $\mathcal{S}_0$ for the exact scope correspondence (C.53).

Proposition G.7 (Primitive scopes with a specialist premium). *Let m lie in a compact subset of $(0, 1)$, ϕ_S in a compact subset of $(0, \infty)$ and $\kappa_F \in [1, \bar{\kappa}_F]$, with a common positive floor on paid non-home access costs.*

- (a) *On such a set with positive distance from ϕ_P , ϕ_U and ϕ_W , there is a uniform $c_* > 0$ such that $\mathcal{S}_c(m, \phi_S; \kappa_F) \subseteq \mathcal{S}_0(\phi_S)$ for $0 < c < c_*$. Every limit of normalized primitive equilibria is an exact equilibrium at ϕ_S .*

(b) Suppose the set also lies in the continuation domain of Appendix B.9 with ϕ read as ϕ_S , is separated from $2/3$ and $\phi_+ = 45/64$ on the pair region, and satisfies

$$\inf\{2 - m^2[10 + \phi_S/\kappa_F - 8p_s(\phi_S)]\} > 0 \quad (\text{G.6})$$

wherever the pair belongs to $\mathcal{S}_0(\phi_S)$. Then, after reducing c_* , $\mathcal{S}_c(m, \phi_S; \kappa_F) = \mathcal{S}_0(\phi_S)$. Each configuration of Lemma D.5, on its compact strict branch interior at ϕ_S , including the endpoint separations of the unequal branches, and with the symmetric-pair support condition read as (G.6), has a primitive break-even continuation with

$$\begin{aligned} \delta_B &= \frac{F_B}{Q_0} + O(c^3), & \text{AUM}_B &= Q_0 + O(c), \\ \delta_T &= tc\{p_T(\phi_S) + O(c)\}, & \text{AUM}_T &= Kc\{Q_T(\phi_S) + O(c)\}. \end{aligned}$$

Scope equality concerns the set of scopes. It does not continue every price branch at every event.

Proof. Every occurrence of F or ϕ in Appendix D is a specialist comparison, a test of B , or the market fee δ_B . Specialist comparisons use F_S and ϕ_S , tests of B use F_B , and on continued branches the market fee is $\delta_B = F_B/Q_0 + O(c^3)$. The changed steps are these.

Domain. $F_S = \Theta c^2 \phi_S$ and $F_B = F_S/\kappa_F$ lie below the ceilings of Proposition G.3 for all sufficiently small c , uniformly on the stated sets.

Normal form and demand limits. Lemma D.1 holds for every bounded $\beta = \delta_B/c^2$; with the premium, $\beta = \Theta \phi_S/(\kappa_F Q_0) + O(c)$. By (D.8), specialist break-even has right side ϕ_S . Lemma D.2 needs only $\delta_B \leq C_B c^2$. Its converse lifts a strictly profitable exact deviation: normalized revenue above ϕ_S gives primitive revenue above F_S .

Quote tails. In the proof of Lemma D.3, the complementary-triple bound is $\delta_B/(1+3m^2) + O(c^3) = F_B/4 + O(c^3)$. Since $F_B \leq F_S$ and $F_S/c^2 = \Theta \phi_S$ has a positive lower bound, choosing ϵ_* and then c_* gives (D.9) with the entrant's own cost F_S on the right, uniformly in $\kappa_F \geq 1$.

Market-fund protection. Lemma D.4 uses B 's own break-even and is unchanged.

Configurations. In Lemma D.5, the empty, pooled and unequal branches use B 's lower root with F_B , specialist break-even at ϕ_S , the market-fund Jacobian block and the specialist Jacobians, which are unchanged in ϕ_S coordinates. The first-order support signs do not involve the market fee. On the symmetric pair the middle homes are degenerate at first order, and their second-order coefficient (D.15) is derived with the market fee free. With $\delta_B = \Theta \phi_B c^2/Q_0 + O(c^3)$ it becomes

$$x_1^{\text{free}} = x_3^{\text{free}} = \frac{m^2(10 + \phi_B - 8p) - 2}{(1 - m^2)^2(1 + 3m^2)} c^2 + O(c^3), \quad \phi_B = \frac{\phi_S}{\kappa_F}, \quad (\text{G.7})$$

so (D.2) becomes (G.6). The common-price path and the implicit-function step are unchanged.

Local protection, persistence and closure. In Lemma D.6, the slopes (D.20) and (D.22) and the home-price derivatives do not involve δ_B at leading order. The symmetric pair uses the sign supplied by (G.6), and the margins are measured against ϕ_S . In Proposition D.7, the tail-region revenue is below half of the entrant's own cost F_S , and Lemma C.10 at ϕ_S covers the rest. Proposition D.8

uses Propositions G.3 and G.4 in place of Proposition 1 and its rates, and exact profitable witnesses against F_S .

Scopes. Part (a) is the compactness argument of Proposition D.9, which uses no support condition. Part (b) uses the continuations above in the reverse inclusion. $\square$

The support condition remains necessary. The rest of the proof of Proposition D.10, including the revenue slope (D.27), does not involve the market fee at leading order. At $m = 3/5$, $\phi_S = 18/25$ and $p_s = 2/5$, the numerator of (G.7) is

$$m^2 \left(10 + \frac{18}{25\kappa_F} - \frac{16}{5} \right) - 2 = \frac{56}{125} + \frac{162}{625\kappa_F} > 0,$$

and its denominator is positive. Both middle homes are therefore positive on the nearby symmetric configuration for every $\kappa_F \geq 1$; at $\kappa_F = 1$ the coefficient is $425/512$, as in (D.26). The diagnostic persists under any premium, so the support condition in Proposition G.7(b) cannot be omitted.

Continuation domain. Proposition G.7(b) keeps the conservative inherited domain of Appendix B.9 with ϕ read as ϕ_S : $M_{\text{cont}} \subset (0, \sqrt{44/69})$ and $\sup \Phi_{\text{cont}} < \inf_{m \in M_{\text{cont}}} 2(1 - m^2)/m^2$. It uses the support condition in the two-cost form (G.6), because (G.7) is the coefficient on which the proof relies. The inherited condition $m^2[10 + \phi_S - 8p_s(\phi_S)] < 2$ implies (G.6), since $\phi_S/\kappa_F \leq \phi_S$.

G.5 Fees, performance and portfolios

Corollary G.8 (Consequences of the premium).

- (a) *Sizes and fees.* On the compact sets of Proposition G.7(a), for all sufficiently small c , every active specialist satisfies $\text{AUM}_B / \text{AUM}_T \asymp c^{-1}$ and $\delta_T / \delta_B = \kappa_F \text{AUM}_B / \text{AUM}_T \asymp \kappa_F c^{-1}$; the market fund is larger and cheaper. On continued branches, $\text{AUM}_B / \text{AUM}_T = 4(1 - m^2)/(mcQ_T(\phi_S)) + O(1)$.
- (b) *Performance.* On the same sets, for all sufficiently small c , every active specialist satisfies

$$0 < \delta_T < \frac{1}{2}a_T^g, \quad a_T^n > \frac{1}{2}a_T^g > 0, \quad a_T^g \text{AUM}_T > 2F_S.$$

Gross alphas are those of Lemma E.1, and the market fund's net alpha is $-\delta_B = -F_B / \text{AUM}_B$. On continued branches the AUM-weighted net alpha of (E.7) becomes

$$\bar{a}_A^n = \frac{\Theta(m)c^2}{Q_0} \left(\sum_{T \neq B} \rho_T Q_T - \frac{\phi_S}{\kappa_F} - N_S \phi_S \right) + O(c^3),$$

with bracket $Q_C(2/3 - z(1 + 1/\kappa_F)) > 0$ for pooling, $\phi_S(1/p_L + 1/p_R - 2 - 1/\kappa_F) > 0$ for a pair, and $-\phi_S/\kappa_F$ for B alone.

- (c) *Portfolios.* In the small-dispersion approximation, the holdings table and direct-share coefficients of Appendix E.3 hold at ϕ_S . On the symmetric pair continuation of Proposition G.7(b), both

middle investors' home holdings are exactly zero. When B operates alone, the middle investors' home holdings are zero if and only if $\delta_B \leq \eta_1^H$, where $\eta_1^H = c^2/(2(1 + 3m^2))$; the inequality is weak at the exact zero-holding boundary. To leading order, $\phi_S/\kappa_F < 2(1 - m^2)/m^2$ gives zero middle homes for small c , and the reverse strict inequality gives positive ones.

Proof. (a) Combine Lemma G.2 with Proposition G.4. The expansion on continued branches uses Proposition G.7(b) and $Kc = mc/((1 - m^2)(1 + 3m^2))$. The comparison across organizations is the reduced calculation of Appendix E.1 at ϕ_S .

(b) The proof of Corollary E.2 uses closure, now Proposition G.7(a); the reduced fee bound $p_T/\rho_T \leq 5/11$, unchanged at ϕ_S by Proposition G.6; and $\delta_T \text{AUM}_T = F_T$, which now gives $a_T^g \text{AUM}_T > 2F_S$. Gross alphas depend only on portfolio weights and returns. For the weighted mean, $\sum_T A_T a_T^n = \sum_{T \neq B} A_T a_T^g - (F_B + N_S F_S)$, because the market fund's gross alpha is zero. Dividing by total managed exposure $Q_0 + O(c)$ and using the scales of Appendix E.1 gives the display. Break-even $zQ_C = \phi_S$ and $p_T Q_T = \phi_S$ give the brackets, which are positive because $z \leq 10/33$ and $p_T \leq 3/7$.

(c) The reduced tables involve only specialist quotes, which are functions of ϕ_S . Exact zero middle homes on the symmetric continuation follow from (G.7) and (G.6). For B alone, Lemma A.8 gives the threshold $\eta_1^H = 1 - v_B/s_1 = c^2/(2(1 + 3m^2))$, and $\delta_B/\eta_1^H \rightarrow m^2\phi_B/(2(1 - m^2))$ by Proposition G.4. $\square$

H Portfolio choice within mandates

This appendix proves Proposition 7 and the two specialist-redesign statements in Section 6.2. All results concern the extended small-dispersion problem defined below. None of them is a statement about the primitive economy at positive c .

H.1 Broad designs and the reduced design game

Designs and timing. An all-asset product may hold

$$b_B(s) = b_B + s(\iota_2 - \iota_0), \quad 0 \leq s \leq \bar{s} \leq \frac{1}{4}, \quad (\text{H.1})$$

with $b_B = 1/4$. Each s is a distinct product with operating cost F , and several broad designs may be active. The equal-weight design $s = 0$ is feasible but need not be selected. Specialists keep their standardized weights in this part, and all products share the cost F . A product is a design and a fee, posted publicly before investors choose access and portfolios and fixed for the period. Definition 1 applies with the catalogue enlarged to all broad designs: break-even, protection against a lower quote for the same design with focal-upper demand, and no profitable entry by any omitted design at any quote. There is no separate stage at which incumbents withdraw and redesign.

The full correction. By (D.3), the leading correction shifts the weight

$$s^* = \frac{mc}{2(1-m^2)} \quad (\text{H.2})$$

from asset 0 to asset 2. Write $\theta = s/s^*$ and $\bar{\theta} = \bar{s}/s^*$, and let $\theta^* = \min\{\bar{\theta}, 1\}$. The main results hold $\bar{\theta}$ fixed as c falls, so that the cap $\bar{s}$ in actual weights is proportional to c . A tilted broad fund can be replicated from standardized products:

$$b_B(s) = \left(\frac{1}{2} - 2s\right)b_{01} + 2s b_{12} + \frac{1}{2} b_{23}, \quad 0 \leq s \leq \frac{1}{4}. \quad (\text{H.3})$$

Its selection must therefore come from costs, not from the absence of a replicating combination.

Normal form. Let type i hold total broad exposure $M_i = \sum_k M_{ik}$ in designs $s_k = \theta_k s^*$ with fees δ_k , and let her specialist and home holdings be $q_{iT} = Kc u_{iT}$ and $x_i = Kc h_i$, with specialist fees $\delta_T = \alpha p_T$, as in Appendix C.1. Put

$$\lambda_{ik} = \frac{M_{ik}}{M_i}, \quad \theta_{\lambda_i} = \sum_k \lambda_{ik} \theta_k, \quad f_k = \frac{Q_0 \delta_k}{\Theta c^2}, \quad y_i = \sum_T u_{iT} d_T + h_i H_i.$$

Lemma H.1 (Normal form with several broad designs). *On every bounded set of tilts θ_k , normalized quantities, specialist quotes and broad fees f_k , the optimized total broad position is $Q_0 + O(c)$ for all sufficiently small c . Subtracting the value of the fee-free equal-weight position and dividing by Θc^2 gives, uniformly,*

$$\Psi_i = \frac{1}{2} \|g\|^2 - J_i + O(c), \quad J_i = \frac{1}{2} \|(1 - \theta_{\lambda_i})g - y_i\|^2 + \sum_T p_T u_{iT} + \sum_k \lambda_{ik} f_k. \quad (\text{H.4})$$

Proof. The broad holdings add $\sum_k M_{ik} s_k (\iota_2 - \iota_0)$ to exposure. The direction $\iota_2 - \iota_0 = g/2$ is a zero-sum adjustment. By (H.2) and (D.1), $Q_0 s^*/2 = Kc$, so the tilt equals $Kc \theta_{\lambda_i} g + O(c^2)$ when $M_i = Q_0 + O(c)$. The broad fee bill is $\bar{\delta}_i M_i$, with $\bar{\delta}_i = \sum_k \lambda_{ik} \delta_k = O(c^2)$. The computation in the proof of Lemma D.1 applies with the adjustment y_i replaced by $y_i + \theta_{\lambda_i} g$ and δ_B replaced by $\bar{\delta}_i$. The common coordinate is optimized as there; the fee enters through $-Q_0 \bar{\delta}_i$ at first order, and terms in $\bar{\delta}_i^2$ are $O(c^4)$. After division by Θc^2 , the linear and quadratic terms are $g' \tilde{y}_i - \frac{1}{2} \|\tilde{y}_i\|^2$, with $\tilde{y}_i = y_i + \theta_{\lambda_i} g$, and the broad fee term is $-\sum_k \lambda_{ik} f_k$. Completing the square, $g' \tilde{y}_i - \frac{1}{2} \|\tilde{y}_i\|^2 = \frac{1}{2} \|g\|^2 - \frac{1}{2} \|(1 - \theta_{\lambda_i})g - y_i\|^2$. $\square$

Definition H.2 (Extended small-dispersion problem and condition (R)). *Condition (R) states that broad designs carry each type's common position, so that $\sum_k \lambda_{ik} = 1$ with total broad position $Q_0 + O(c)$; that broad fees are of order c^2 , so the f_k are bounded; and that specialist fees are of order c . In the extended small-dispersion problem, type i chooses λ_i in the unit simplex over the offered broad designs and $u_i, h_i \geq 0$ to minimize J_i in (H.4).*

Condition (R) maintains for the broad tier what Proposition 1 and Proposition B.9 establish for the equal-weight market fund. It is not derived here for several broad designs.

Definition H.3 (Reduced design equilibrium). *At $\phi > 0$, a reduced design equilibrium consists of offered broad designs $\{(\theta_k, f_k)\}$ with $\theta_k \in [0, \bar{\theta}]$ and $f_k > 0$, a specialist menu $E \subseteq \mathcal{P}_8$ with fees $p_T > 0$, and one jointly compatible selection of minimizers (λ_i, u_i, h_i) of the problems in Definition H.2, such that:*

- (i) *every offered broad design has positive share $\Lambda_k = \frac{1}{4} \sum_i \lambda_{ik}$ and breaks even, $f_k \Lambda_k = \phi$;*
- (ii) *no lower quote for an offered design, and no omitted design $\theta \in [0, \bar{\theta}]$ at any quote $f' > 0$, earns $f' \bar{\Lambda}' > \phi$, where $\bar{\Lambda}'$ is the focal-upper share with all other fees fixed;*
- (iii) *the specialists satisfy the break-even, lower-quote and entry conditions of Definition C.2 against ϕ , with demand computed from the minimizers of J_i and focal-upper demand for deviations.*

Broad revenue $\delta_k \text{AUM}_k = \Theta c^2 f_k \Lambda_k$ and specialist revenue $\Theta c^2 p_T Q_T$ are both compared with $F = \Theta c^2 \phi$.

H.2 The value of tilt and conditional homogeneity

For a specialist menu E with positive fees, let

$$\gamma_i(y) = \min \left\{ \sum_{T \in E} p_T u_T : \sum_{T \in E} u_T d_T + h H_i = y, u, h \geq 0 \right\},$$

with $\gamma_i(y) = +\infty$ off the cone generated by these directions, and

$$V_i(\theta) = \min_{u, h \geq 0} \left\{ \frac{1}{2} \|(1 - \theta)g - y\|^2 + \sum_{T \in E} p_T u_T \right\}.$$

Lemma H.4 (Value of tilt). *For every type i , V_i is convex and continuously differentiable on $\mathbb{R}$, with $V'_i(\theta) < 0$ for $\theta < 1$, $V'_i(\theta) > 0$ for $\theta > 1$ and $V_i(1) = 0$. Hence $\theta = 1$ is its unique minimizer, and θ^* is its unique minimizer on $[0, \bar{\theta}]$. On $[0, \theta']$ with $\theta' < 1$, $V'_i \leq -\varepsilon$ for some $\varepsilon > 0$.*

Proof. The function γ_i is the value of a linear program in its right-hand side, hence polyhedral and sublinear, and $\gamma_i \geq 0$. Its Moreau envelope $\Omega_i(r) = \min_y \{ \frac{1}{2} \|r - y\|^2 + \gamma_i(y) \}$ is convex and continuously differentiable, with $\nabla \Omega_i(r) = r - y^*(r)$, where $y^*(r)$ is the unique proximal point. Since $V_i(\theta) = \Omega_i((1 - \theta)g)$, V_i is convex and continuously differentiable, with $V'_i(\theta) = -g'(r - y^*)$ at $r = (1 - \theta)g$. Optimality gives $r - y^* \in \partial \gamma_i(y^*)$, and every subgradient ζ of a sublinear function at y^* satisfies $\zeta' y^* = \gamma_i(y^*)$. Hence

$$(r - y^*)' r = \|r - y^*\|^2 + \gamma_i(y^*) \geq 0,$$

with equality only if $r = y^*$ and $\gamma_i(y^*) = 0$, that is, only if r lies on the free home ray $\{h H_i : h \geq 0\}$. For $r \neq 0$ this is impossible, because g is not proportional to any H_i . Thus, for $\theta \neq 1$, $g'(r - y^*) = (r - y^*)' r / (1 - \theta)$ has the sign of $1 - \theta$, which gives the signs of V'_i . At $\theta = 1$, $r = 0$, $y^* = 0$ and $V_i = 0$. The uniform bound follows from continuity of V'_i on a compact interval. $\square$

Lemma H.5 (Conditional homogeneity). *For a common broad tilt $\theta < 1$, the specialist problem at fees p equals $(1 - \theta)^2$ times the problem of Definition C.1 at fees $\hat{p} = p/(1 - \theta)$, with $u = (1 - \theta)\hat{u}$ and $h = (1 - \theta)\hat{h}$. Hence specialist equilibria at cost ϕ correspond one to one to equilibria of Definition C.2 at $\phi_{\text{eff}} = \phi/(1 - \theta)^2$, with fees and quantities equal to $1 - \theta$ times those of the equilibrium at ϕ_{eff} . At $\theta = 1$, no specialist has positive demand at any positive fee.*

Proof. Substituting $u = (1 - \theta)\hat{u}$, $h = (1 - \theta)\hat{h}$ and $p = (1 - \theta)\hat{p}$,

$$\frac{1}{2}\|(1 - \theta)(g - \hat{y})\|^2 + (1 - \theta)^2 \sum_T \hat{p}_T \hat{u}_T = (1 - \theta)^2 \left[\frac{1}{2}\|g - \hat{y}\|^2 + \sum_T \hat{p}_T \hat{u}_T \right],$$

which is $(1 - \theta)^2$ times (C.16). Optimizer sets, including focal-upper selections, correspond one to one. Aggregate demand scales by $1 - \theta$ and revenue $p_T Q_T$ by $(1 - \theta)^2$, so every comparison with ϕ is a comparison of the original problem with ϕ_{eff} . At $\theta = 1$ the target is zero, and $y^* = 0$ at any positive fees. $\square$

H.3 Selection of the broad design

Proposition H.6 (Broad-tier selection). *Under condition (R), with $\bar{\theta}$ fixed, every reduced design equilibrium has the following properties.*

- (a) *Exactly one broad design is offered.*
- (b) *That design is $\theta^* = \min\{\bar{\theta}, 1\}$.*
- (c) *Every type holds it entirely, $\lambda_i = 1$, at normalized fee $f = \phi$, that is, at a fee of about F/Q_0 .*
- (d) *If $\theta^* < 1$, the specialists form an equilibrium of Definition C.2 at $\phi_{\text{eff}} = \phi/(1 - \theta^*)^2$, with that equilibrium's fees and quantities scaled by $1 - \theta^*$. If $\theta^* = 1$, no specialist is offered.*

Conversely, (e) every configuration with properties (a)–(d) is a reduced design equilibrium. In particular, Proposition 7 holds.

Proof. Type i chooses λ_i to minimize $G_i(\lambda) = V_i(\theta_\lambda) + \sum_k \lambda_k f_k$, a convex problem. With a single broad design, a specialist deviation leaves λ unchanged. The steps below concern broad deviations with specialist fees fixed.

Step A: at most one broad design. Suppose that at least two designs are offered. Their shares sum to one, so $\Lambda_k < 1$ and $f_k = \phi/\Lambda_k > \phi$ for each of them. Let $f_{\min} = \min_k f_k > \phi$. An entrant, or a lower quote for an existing design, at θ^* with fee $f_E \in (\phi, f_{\min})$ is strictly preferred by every type to any mix that puts positive weight on the incumbents. Indeed $V_i(\theta^*) \leq V_i(\theta_\lambda)$ for every feasible θ_λ by Lemma H.4, and $f_E < \sum_k \lambda_k f_k$ whenever an incumbent receives positive weight. The entrant's share is therefore one and its revenue $f_E > \phi$, a contradiction. The preference is strict, so no tie rule is used. Exactly one design is offered, with share one and $f = \phi$. This proves (a) and (c).

Step B: the design is θ^ .* Suppose the design is $\theta_1 \neq \theta^*$ with $f_1 = \phi$. Choose an entrant tilt θ_E strictly between θ_1 and θ^* : take $\theta_E = \bar{\theta} = \theta^*$ when $\bar{\theta} < 1$, and $\theta_E = (\theta_1 + 1)/2$ when $\theta^* = 1$. By Lemma H.4, on the segment between θ_1 and θ_E every type satisfies $|V'_i| \geq \varepsilon_E > 0$, with the sign

pointing toward θ_E . For an entry fee $\phi + \Delta$ with $0 < \Delta < \varepsilon_E |\theta_E - \theta_1|$, moving a share μ into the entrant gives

$$\frac{d}{d\mu} G_i = V'_i(\theta_1 + \mu(\theta_E - \theta_1))(\theta_E - \theta_1) + \Delta < 0 \quad \text{on } [0, 1].$$

Every type moves her whole position, and the entrant's revenue is $\phi + \Delta > \phi$, a contradiction. This proves (b). A direct jump to $\theta = 1$ need not capture every type when $\theta^* = 1$, because $V'_i(1) = 0$ allows interior mixing; the intermediate entrant avoids this.

Step C: protection of (θ^, ϕ) .* A lower quote $f' < \phi$ for the same design earns at most $f' < \phi$. An omitted design $\theta \neq \theta^*$ at a fee $f' \geq \phi$ attracts no demand: for every $\mu > 0$,

$$G_i(\mu) \geq V_i(\theta^* + \mu(\theta - \theta^*)) + \phi > V_i(\theta^*) + \phi = G_i(0),$$

strictly by Lemma H.4. At a fee $f' < \phi$, it earns at most $f' < \phi$, because the broad tier has total share one. This covers every cheaper design with a different tilt.

Step D: specialists. With one broad design held by every type, Lemma H.5 turns the specialist conditions into those of Definition C.2 at ϕ_{eff} . This proves (d).

(e). For a configuration with (a)–(d), the broad design breaks even by (c), Step C verifies every broad deviation, and Step D verifies the specialist conditions. $\square$

Corollary H.7 (Scope and extinction). *Let $\theta^* < 1$. The equilibrium specialist scopes are $\mathcal{S}_0(\phi/(1 - \theta^*)^2)$ in (C.53), and each cutoff ϕ_X of Proposition C.7 becomes $\phi_X(1 - \theta^*)^2$. For $0 < \phi < \phi_W$, specialists operate in every equilibrium when $\theta^* < 1 - \sqrt{\phi/\phi_W}$, and the broad fund stands alone in every equilibrium when $\theta^* > 1 - \sqrt{\phi/\phi_W}$. At equality, both the broad fund alone and pooling at its capacity fee are equilibria. For $\phi > \phi_W$, the broad fund stands alone at every tilt.*

Proof. Apply Proposition H.6(d) and (C.53) at ϕ_{eff} , retaining the endpoint rows. The condition $\phi_{\text{eff}} \geq \phi_W$ is $\theta^* \geq 1 - \sqrt{\phi/\phi_W}$. $\square$

Corollary H.8 (Proportional and fixed caps). *With a cap proportional to c , Proposition H.6 applies at the fixed tilt $\theta^* = \min\{\bar{\theta}, 1\}$. With a fixed actual cap $\bar{s} > 0$, the full leading correction is feasible once $c \leq 2(1 - m^2)\bar{s}/m$. The extended reduced game then selects $\theta = 1$, that is $s = s^*(c)$, and has no specialist demand.*

Proof. With a fixed actual cap, $\bar{\theta} = 2(1 - m^2)\bar{s}/(mc) \geq 1$ once $c \leq 2(1 - m^2)\bar{s}/m$. Lemma H.1 involves the broad designs only through θ_λ and the fee bill, and Lemma H.4 holds on all of $\mathbb{R}$, so Proposition H.6 applies with $\theta^* = 1$ for every $\bar{\theta} \geq 1$. The last statement of Lemma H.5 gives zero specialist demand. $\square$

The fixed-cap statement is a statement about the extended reduced game. As $c \rightarrow 0$ the normalized design set $[0, \bar{\theta}(c)]$ becomes unbounded, outside the bounded sets of Lemma H.1. Transferring the conclusion to the primitive economy is not established here.

Replication. By (H.3), standardized products can replicate every tilted broad design. Under condition (R) this does not affect the leading order, because a specialist's fee per unit is of order

c and a broad fund's of order c^2 . At the primitive level, the corresponding comparison would be the cost comparison of Appendix B; it is not carried out here for tilted designs. As $s \rightarrow 0$, the replication reduces to the complementary pair, which the case $N = 2$ of Lemma B.7 already rules out.

H.4 Fees, sizes and alphas

Corollary H.9 (Fees, sizes and alphas with a tilted broad fund). *Let $\theta^* < 1$, and let $\hat{p}_T$ and $\hat{Q}_T$ denote the fees and demands of an equilibrium of Definition C.2 at ϕ_{eff} .*

- (a) *At leading order, specialist T 's fee and AUM are $(1 - \theta^*)\alpha\hat{p}_T$ and $(1 - \theta^*)Kc\hat{Q}_T$. Its revenue equals F , so equal revenue and inverse fee/size ratios hold among all active funds.*
- (b) *At fixed ϕ , on open smooth lower-root segments inside the protected pooled and symmetric-pair scope, each specialist's fee rises and its AUM falls as θ^* rises. Monotonicity across the symmetric pair's join at $\phi_{\text{eff}} = 2/3$ follows by continuity.*
- (c) *Against the equal-weight index, a specialist's gross alpha is the one in Lemma E.1, unchanged. Against the selected broad fund B' , with weights $b_B(\theta^*s^*)$, a portfolio w has gross alpha $(1 - \theta^*)a^g(w) + O(c^2)$, where $a^g(w)$ is the gross alpha against the index. A specialist's net alpha against B' , $(1 - \theta^*)\alpha(\rho_T - \hat{p}_T) + O(c^2)$, is positive.*
- (d) *The selected broad fund's gross alpha against the equal-weight index is exactly*

$$\frac{4mcs}{1 + 3m^2 - c^2/2} = \frac{2\theta^*m^2c^2}{(1 - m^2)(1 + 3m^2 - c^2/2)}, \quad s = \theta^*s^*.$$

Its leading term is $4s\alpha$, with $\alpha = mc/(1 + 3m^2)$. At leading order it exceeds the broad fund's fee F/Q_0 if and only if $\phi < 8\theta^$.*

Proof. (a) follows from Lemma H.5 and the scales $\delta_T = \alpha p_T$, $\text{AUM}_T = KcQ_T$, together with Lemma A.7.

(b) Let $R(\hat{p}) = \hat{p}\hat{Q}(\hat{p})$ be the revenue function on such a segment, with $R' > 0$ at the lower root and $\hat{Q}' \leq 0$. The fee at fixed ϕ is $\hat{f}(\theta) = (1 - \theta)\hat{p}(\phi/(1 - \theta)^2)$, and $d\hat{p}/d\phi_{\text{eff}} = 1/R'$. Hence

$$\hat{f}'(\theta) = -\hat{p} + 2\phi_{\text{eff}} \frac{d\hat{p}}{d\phi_{\text{eff}}} = \frac{2R - \hat{p}R'}{R'} = \frac{\hat{p}(\hat{Q} - \hat{p}\hat{Q}')}{R'} > 0.$$

AUM $(1 - \theta)\hat{Q}$ falls because both factors fall. At the symmetric pair's join $\phi_{\text{eff}} = 2/3$, the left derivative R' tends to zero, so the formula is used on each side and the fee is continuous across the join.

(c) Let Σ be the correlation matrix of Lemma E.1 and $D = 1 + 3m^2 - c^2/2$. For $B' = b_B + s(\iota_2 - \iota_0)$,

$$w'\Sigma(\iota_2 - \iota_0) = (1 - m^2 + c^2)(w_2 - w_0) - 2mc(w_1 + w_3), \quad b'_B\Sigma(\iota_2 - \iota_0) = -mc.$$

Since $s = O(c)$, $\text{Cov}(R_w, R_{B'}) = w' \Sigma b_B + s(1 - m^2)(w_2 - w_0) + O(c^2)$ and $\text{Var}(R_{B'}) = D/4 + O(c^2)$. Therefore

$$1 - \beta_{w,B'} = a^g(w) - \frac{4s(1 - m^2)(w_2 - w_0)}{D} + O(c^2) = (1 - \theta^*)a^g(w) + O(c^2),$$

because $4s(1 - m^2)/D = \theta^* \cdot 2mc/D$ and $a^g(w) = 2mc(w_2 - w_0)/D + O(c^2)$ by (E.4). For a specialist, $a^g = \alpha \rho_T + O(c^2)$ and the fee is $(1 - \theta^*)\alpha \hat{p}_T$ by (a). The fee share satisfies $\hat{p}_T/\rho_T \leq 5/11$ at ϕ_{eff} , so the net alpha is positive.

(d) Apply (E.4) with $w_2 - w_0 = 2s$ and $w_0 + w_2 = 1/2$, which gives $4mcs/D$ exactly. With $s = \theta^* s^*$ and (H.2) this is the displayed expression; its leading term is $4s\alpha$. The broad fee is $F/Q_0 = \Theta \phi c^2/Q_0$ at leading order, and the ratio of alpha to fee is $8\theta^*/\phi$. $\square$

H.5 Weight choice by specialists

Keep the broad fund at b_B . Each non-market mandate T may hold

$$w_{Tj}(t) = \frac{1 + t(\bar{a}_T - a_j^{\text{dir}})}{|T|} \quad (j \in T), \quad w_{Tj}(t) = 0 \quad (j \notin T),$$

where $a^{\text{dir}} = (1, 0, -1, 0)$, $\bar{a}_T$ is the mean of a^{dir} over T , and $|t| \leq \varepsilon < 1$. The parameter t here is unrelated to the scale t in (D.1), which this appendix does not use. Each weight choice is a separate product with operating cost F , and the commitment and entry conventions are those above. For the three-asset mandate $C = 123$, $w(t) = (0, \frac{1}{3} - \frac{t}{9}, \frac{1}{3} + \frac{2t}{9}, \frac{1}{3} - \frac{t}{9})$.

Proposition H.10 (Small discretion displaces the equal-weight pooled design). *Let the pooled configuration $(\{C\}, z)$ have $0 < z < 10/33$, so that $\phi = z(15/2 - 99z/8)$ by (C.31). For $t > 0$, let an entrant hold $w(t)$ at the quote $\eta(t) = z(1 - 13t/30)$, with C still offered at z . For all sufficiently small $t > 0$, no type holds C , and the entrant's normalized revenue exceeds ϕ . Hence the equal-weight pooled configuration fails the entry test for every positive ε .*

Proof. The entrant's active direction is $d_t = d_C + t h$, with $h = (0, -\frac{1}{9}, \frac{2}{9}, -\frac{1}{9})' = (H_2 - d_C)/3$. Consider the candidate in which types 0, 1 and 3 hold only the entrant, with quantity $q_3(t) = (g'd_t - \eta)/(d_t'd_t)$, and type 2 holds the entrant and her home asset, with quantities $q_2(t)$ and $x_2(t)$ from the two-column Gram system in d_t and H_2 . At $t = 0$ these reduce to the pooled holdings of Appendix E.3,

$$q_3(0) = 8 - 12z, \quad q_2(0) = 6 - \frac{27z}{2}, \quad x_2(0) = 2 + \frac{3z}{2},$$

all positive for $0 < z < 10/33$. The inactive home marginal values of types 0, 1 and 3 at $t = 0$ are $-3z$, $z - 2/3$ and $z - 2/3$, all negative, and remain negative for small t . At the candidate, the marginal value of the still-offered C less its fee z is

$$-\frac{t(117z + 240tz - 160t + 120)}{30(8t^2 + 9)} \quad \text{for types 0, 1, 3}, \quad \frac{3tz}{10(t - 3)} \quad \text{for type 2},$$

and both are negative for sufficiently small $t > 0$. The investor problem is convex, so these conditions make the candidate optimal. Optimal exposure is unique, so C is unheld in every optimum. The entrant's revenue $R(t) = \eta(t)[3q_3(t) + q_2(t)]/4$ satisfies $R(0) = \phi$ and $R'(0) = z(339z + 50)/40 > 0$, so $R(t) > \phi$ for sufficiently small $t > 0$. $\square$

Proposition H.11 (A flexible pooled mandate bundles the equal-fee pair). *Keep the broad fund at b_B , and suppose the three-asset mandate can reach*

$$C^* = w(3/4) = (0, \frac{1}{4}, \frac{1}{2}, \frac{1}{4}) = \frac{1}{2}(b_{12} + b_{23}).$$

Then no symmetric pair configuration $(\{L, R\}, (p, p))$ with $0 < p \leq 3/7$ passes the entry test. An entrant offering C^ at any quote $\eta < p$, with L and R still offered at p , has demand $2Q_P(\eta)$, where Q_P is given by (C.35). Its normalized revenue $2\eta Q_P(\eta)$ tends to $2\phi > \phi$ as $\eta \uparrow p$.*

Proof. (i) At the auxiliary symmetric fees (η, η) , $0 < \eta < p \leq 3/7$, the holdings table at symmetric prices in Appendix E.3 gives the reduced optimum. Types 0, 1 and 3 hold $u_L = u_R = 4(1 - \eta)$, type 2 holds $u_L = u_R = \max\{4(1 - 3\eta), 0\}$, and the middle investors hold no home asset. Thus every type holds equal amounts $u_i(\eta)$ of L and R .

(ii) Replace these holdings by $2u_i(\eta)$ units of C^* . Exposure and the fee bill are unchanged. Every portfolio that uses C^* alone is replicable in the (η, η) pair menu with the same exposure and fee bill, so the replaced portfolio is optimal in the menu with C^* alone at η .

(iii) In the actual menu, with L and R at p and C^* at η , the marginal values of L and R at that exposure are at most η , by the KKT conditions of the (η, η) optimum. Their fee is $p > \eta$, so their net marginal values are strictly negative, and they are unheld in every optimum. The candidate satisfies every KKT condition of this convex problem. The optimal exposure is unique, and d_{C^*} and H_i are linearly independent, so the entrant's demand is $\frac{1}{4} \sum_i 2u_i(\eta) = 2Q_P(\eta)$. Its normalized revenue $2\eta Q_P(\eta)$ tends to $2pQ_P(p) = 2\phi$ as $\eta \uparrow p$, by continuity of Q_P . $\square$

Proposition H.10 rules out the equal-weight pooled design at strict interior pooled fees; it does not characterize a redesigned pooled organization. Proposition H.11 concerns symmetric pair configurations only. Unequal pair branches and the equilibrium with flexible specialist mandates are not addressed. Reaching C^* moves each middle weight from one third to one quarter and the diversifier's weight from one third to one half.