# A Trilemma for Asset Demand Estimation*

Most recent version

William Fuchs† Satoshi Fukuda‡ Daniel Neuhann§

May 8, 2026

**Abstract**

We establish fundamental limits on demand analysis in asset markets. Given observational data, one cannot jointly maintain (i) no-arbitrage asset pricing, (ii) investor preferences over payoffs, and (iii) model-free identification of structural asset demand. These results follow from a decomposition of asset demand into fundamental preferences and a latent mapping from preferences to asset holdings. This mapping is fundamentally unobservable, varies with routine changes to expected asset payoffs, and links all asset-level demand functions into a non-separable system. As such, even perfectly exogenous supply shifters cannot produce the price variation required to estimate stable asset demand curves. We conclude that asset demand elasticities are contingent, model-specific constructs which reflect—rather than validate—a priori assumptions on investor behavior.

*We thank Dan Ackerberg, Yu An, Jonathan Berk, Olivier Darmouni, Aureo de Paula, Julie Fu, Zhiguo He, Andrey Ordin, and seminar participants at UCL, UT Austin, University of Zurich, 35th Annual Utah Winter Finance Conference, 7th Columbia Workshop in New Empirical Finance, 19th Jackson Hole Finance Group Conference, Tepper-LAEF Advances in Macrofinance Conference, and WashU Finance Conference for helpful comments. Xinran Zhang provided outstanding research assistance.

†University of Texas at Austin, FTG, and CEPR. william.fuchs@mccombs.utexas.edu, https://orcid.org/0000-0001-9936-6712

‡Santa Clara University. sfukuda@scu.edu.

§University of Texas at Austin and FTG. daniel.neuhann@mccombs.utexas.edu.

# 1 Introduction

How much do investors want to hold of a given asset, and how sensitive are their portfolio choices to the price? These questions lie at the heart of asset pricing. They determine how much prices move when central banks purchase bonds, when passive funds rebalance indices, or when financial intermediaries suffer shocks that force asset sales. An influential empirical literature has set out to answer them by estimating asset demand functions from data on portfolio holdings and prices.

A critical question for this literature is whether observed demand responses can be given a structural interpretation without relying on a fully specified equilibrium model of portfolio choice. If so, demand elasticities are invariant to the model used to estimate them and can be used to discriminate between different theories of investor behavior. If not, they are contingent, model-specific objects that reflect — rather than validate — a priori assumptions on investor behavior.

We provide a general theoretical analysis of this question. Our answer is that asset demand functions are not model-free empirical objects, and that structural modeling is unavoidable. Our results require only two foundational principles of asset pricing: that investors value assets for their payoffs, and that asset prices admit no arbitrage. As we discuss, these principles are difficult to discard without invalidating the basic premise of asset demand analysis.

Our results follow from a general decomposition of asset demand functions derived under preferences over payoffs and no arbitrage. Denote by $Y$ the *payoff matrix* summarizing investor beliefs over the state-contingent payoffs of assets in the choice set. Then the matrix of asset demand slopes (in which each element is the derivative of asset demand with respect to a specific asset price) satisfies

$$\mathcal{A}^i = (Y^+)^{\mathrm{T}} \mathcal{D}^i Y^+.$$

In this decomposition, $\mathcal{D}^i$ is the investor's *fundamental demand function* for state-contingent payoffs and $Y^+$ is the Moore-Penrose pseudo-inverse of $Y$. This has a natural interpretation: since preferences are defined over payoffs, not assets directly, $Y^+$ maps demand for payoffs into the associated asset quantities.

This simple decomposition reveals two main challenges. First, the demand function for any individual asset is *commingled with those of all other assets*: because investors care about state-contingent consumption, the optimal quantity of any asset generically depends on the payoffs of all other assets through $Y^+$. This means that one cannot analyze demand for any given asset in isolation. Second, $Y^+$ is *latent and unidentifiable* from past data. Since the payoff matrix reflects investor beliefs about future payoffs, including resale prices, no finite sample of realized returns can pin it down—one can always alter the payoff of an unrealized state, changing $Y^+$ while leaving every historical return intact.

These features of the decomposition have immediate implications for asset demand analysis. Since $Y^+$ is unobservable, fundamental demand $\mathcal{D}^i$ cannot be recovered from portfolio data: different combinations of preferences and latent mappings are always observationally equivalent. And since $Y^+$ shifts whenever beliefs over payoffs are revised, asset demand functions are not structural with respect to standard perturbations that occur during regular market functioning.

What do these considerations imply for the estimation of asset demand functions without structural models? A common approach in the literature is to estimate *individual* asset demand curves through quasi-exogenous variation in asset supply. For this approach to work, supply shocks must generate *ceteris paribus* variation in a single asset price, holding all other prices and payoffs fixed.

Unfortunately, this identification assumption is generically inconsistent with the principle of no arbitrage and equilibrium price determination. Under minimal conditions, an increase in the supply of a given asset reduces the marginal cost of a unit payoff in a given state (i.e., the *state price*) in all states where the asset pays off. But by no arbitrage, this must lead to a change in the prices of all other assets that pay off in overlapping states. Since such payoff overlap is generic for essentially all asset markets, we prove that the identification of individual asset demand curves from individual supply shocks is generically infeasible. Most strikingly, we show that supply shocks generically imply state price changes that differ *directionally* from those required to estimate a fixed demand curve.

This leaves the possibility of jointly estimating the entire $J \times J$ *system* of demand slopes, where $J$ is the number of assets, using multiple independent shocks to the price vector. This requires both at least $J$ linearly independent price changes and that $Y^+$ remains fixed across all experiments. The first is a standard rank condition that arises in many settings; the second is the binding constraint. Since our decomposition shows that belief revisions generically shift all demand functions, it is implausible that the econometrician observes multiple independent shocks to the same demand system. Moreover, standard shocks used in the literature—such as central bank interventions or index inclusions—directly shift the demand system by altering future payoffs.

One might hope that imposing weak statistical structure on $Y$—for instance, through a factor model for asset returns—is enough to make progress. To investigate this, we use random matrix theory to study the statistical properties of the Moore-Penrose inverse $Y^+$ for factor-structured payoff matrices. Our results show the inverse mapping is generically ill-conditioned: the *sign* of any given element of $Y^+$ is a coin flip in large economies, and two economies sharing identical factor structures but different idiosyncratic payoff realizations have sign-independent inverses. Controlling for factor exposures therefore provides no systematic correction for the misalignment between supply shocks and the price variation required for demand estimation. We confirm these predictions numerically through Monte Carlo simulations and empirically using payoff data from S&P 500 stocks.

We summarize our results as a trilemma: given observational data, one cannot jointly maintain (i) no-arbitrage asset pricing, (ii) investor preferences over payoffs, and (iii) model-free identification of structural asset demand functions.

Our findings suggest a critical role for structural models in asset demand analysis. Since $Y^+$ cannot be identified from data, two models that agree on all observable implications of the data can imply arbitrarily different demand elasticities. For example, Fuchs, Fukuda, and Neuhann (2025) show that the logit model of Koijen and Yogo (2019) can infer an elasticity below one even if the true elasticity is infinite. Estimated asset demand elasticities should therefore not be treated as credible calibration targets, and should be evaluated on the plausibility and robustness of the assumed mapping rather than empirical fit.

**Related literature.** Our paper relates to an important literature in finance and economics studying demand effects in financial markets. Early work in this area includes portfolio balance models (Tobin, 1969), and the price effects of index inclusions in equity markets (Shleifer, 1986; Harris and Gurel, 1986). More recently, this broad mechanism has found applications in unconventional monetary policy, foreign exchange markets, and fund flows in bond and equity markets.

This rightly influential literature shows that constraints on capital flows can have important effects on asset prices. However, it stops short of systematically establishing whether and when these price effects reveal structural aspects of investor and market behavior. This is important because critical aspects of asset price determination and policy transmission tightly depend on the price responsiveness of financial markets. We find that non-parametric approaches generically fail to identify asset demand elasticities because they are contaminated by cross-price effects. This means that implicit or explicit theoretical restrictions play a central role in determining the interpretation and policy relevance of the documented effects.

One consequence of our findings is that structural methods are important tools for understanding demand effects in asset markets, much like in many other settings (Berry and Haile, 2021). However, asset markets present particular challenges: investors form portfolios, marginal valuations depend on concurrent holdings of all other assets, the mapping from products to characteristics is latent, and choice is continuous. This fundamental non-separability of asset valuations under a latent mapping means that one cannot easily turn a decision problem with complementarities into, e.g., a discrete-choice problem over bundles. These differences clarify our relationship to recent work in industrial organization which estimates demand systems with complementarities (e.g., Iaria and Wang, 2020; Wang, 2024; Fosgerau, Monardo, and de Palma, 2024; Ershov, Laliberté, Marcoux, and Orr, 2024). These approaches typically study settings in which consumers make discrete choices over a limited number of bundles, or where substitution patterns are governed by exogenous functional-form parameters.

To overcome these challenges, structural models of asset demand must accurately account for the cross-asset linkages and price spillovers inherent to portfolio choice. Fuchs, Fukuda, and Neuhann (2025) show that the prominent logit approach in Koijen and Yogo (2019) can exhibit large biases in standard portfolio choice models with asymmetric substitution between assets. While our analysis in this paper focuses on contemporaneous cross-asset spillovers, similar issues would also arise in a dynamic setting where investors can trade securities referencing different states and dates, as these would also have to be priced by a common pricing kernel and governed by no arbitrage. This broader view helps connect our findings to those in Binsbergen, David, and Opp (2025) and He, Kondor, and Li (2025). Allen, Kastl, and Wittwer (2025) propose a structural model to estimate asset demand without reliance on price instruments. Consistent with our results, this approach requires a priori restrictions and uses data on bid schedules. Perhaps most closely related to this paper is Haddad, He, Huebner, Kondor, and Loualiche (2025), who aim to recover relative demand elasticities from supply shocks without a structural model. Our findings suggest that their approach must impose theoretical restrictions if the estimated elasticities are to have a structural interpretation.

# 2 Framework

## 2.1 Environment

We study a canonical portfolio choice model. A mass of potentially heterogeneous investors $I$ decide how much to consume at $t = 0$, and how to invest their savings to consume at $t = 1$. There are $J$ financial assets, each of which yields a random payoff at date 1. Uncertainty is represented by a set of $Z$ states of the world, one of which is realized at date 1. The payoff of asset $j$ in state $z$ is $y_j(z) \geq 0$. We denote by $Y \equiv (y_j(z))_{j,z}$ the $J \times Z$ matrix of cash flows. Since matrix $Y$ reflects investors' beliefs about state-contingent payoffs, it is unobserved by the econometrician. We denote by $\pi \equiv (\pi_z)_z$, where $\pi_z \in (0, 1)$ is the probability of state $z$.

We are agnostic about the determinants of asset payoffs, and assume investors take the payoff matrix as given. However, in general the payoffs of a given asset are the sum of a direct cash component (i.e., dividends) and its expected resale price (i.e., the expected state-contingent market price). As we will discuss in more detail later, this means that one cannot easily assume that $Y$ is a physical constant that remains fixed across time periods or economic regimes.

At time zero, each investor chooses a *portfolio* to maximize the expected utility of the state-contingent consumption across both dates. A portfolio is a vector of asset positions $a^i \equiv (a^i_j)^J_{j=1} \in \mathbb{R}^J$, where element $a^i_j$ is the investor's holdings of asset $j$. Investor $i$'s preferences are represented by a twice differentiable, strictly increasing and strictly concave von Neumann-Morgenstern utility function $u^i$.

Investors are competitive and take prices as given. The price of asset $j$ is $p_j$, and time-zero consumption is the numeraire (or *outside asset*) with price normalized to one. Investor $i$ is endowed with $e^i_j \geq 0$ units of asset $j$ and $e^i_0 \geq 0$ units of the numeraire, and non-traded consumption endowments $w^i_0 \geq 0$ and $w^i(z) \geq 0$ at date 0 and in state $z$, respectively. Denote by $e^i \equiv (e^i_j)_j$ and $w^i \equiv (w^i(z))_z$. Portfolio choice may be curtailed by exogenous constraints: the investor must choose a portfolio from the set of feasible portfolios $\Phi^i$, which we assume is a convex subset of $\mathbb{R}^J$.

Investor $i$'s *portfolio choice problem* can then be formally stated as:

$$\begin{aligned} \sup_{a^i \in \Phi^i} \quad & (1-\delta^i)u^i(c^i_0) + \delta^i \pi \cdot u^i(c^i) \\ \text{s.t} \quad & c^i_0 = e^i_0 - p \cdot (a^i - e^i) + w^i_0 \quad \text{and} \\ & c^i = Y^{\mathrm{T}} a^i + w^i. \end{aligned} \tag{PCP}$$

Investor *i's asset span* $\mathcal{S}^i$ is set of payoff profiles that can be achieved through some feasible portfolio. That is,

$$\mathcal{S}^i \equiv \{Y^{\mathrm{T}} a^i \in \mathbb{R}^Z \mid a^i \in \Phi^i\}. \tag{1}$$

The portfolio choice problem embeds the canonical notion of *preferences over payoffs*: investors value state-contingent consumption and demand assets *instru-*

*mentally* for the payoffs they provide, not because they provide direct utility. Our results are robust to including a direct utility from holdings but are more sharply stated without them—all we require is that investors have at least some preferences over payoffs. A solution to this problem is jointly determined by several parameters: (i) the utility function $u^i$ and rate of time preference $\delta^i$, (ii) initial wealth $w_0^i$ and state-contingent endowments $w^i$, which are demand shifters that shift state-specific marginal utility, (iii) portfolio constraints which determine the set of feasible portfolios $\Phi^i$, and (iv) the payoff matrix $Y$ and probability distribution $\pi$. We call the utility function, rate of time preference, demand shifters, and portfolio constraints *preference parameters*, which we denote by

$$\Theta^i \equiv \left(u^i, \delta^i, w_0^i, w^i, \Phi^i\right).$$

The optimal portfolio also depends on payoff matrix $Y$, which determines the mapping from asset positions to state-contingent payoffs, and the probability distribution $\pi$, which determines weights on states of the world. Since these objects do not pertain to investor preferences, we refer to these as *external parameters*.

**Asset Demand Functions.** A solution to problem (PCP) can be written in terms of $J$ Marshallian *asset demand functions* which map parameters and the asset price vector into portfolio holdings. That is, the asset demand functional of investor $i$ is

$$a^i(\cdot \mid \Theta^i, Y, \pi) : \mathbb{R}^J_{++} \to \mathbb{R}^J.$$

Standard portfolio choice theory shows that all asset demand functions generically depend on the entire vector of asset prices. That is, asset demand is *non-separable*.

In line with empirical practice, we will typically focus on identifying the $J \times J$ matrix of asset demand derivatives given a prevailing asset price vector $p$:

$$\mathcal{A}^i(\Theta^i, Y, \pi) \equiv -\frac{\partial a^i(p \mid \Theta^i, Y, \pi)}{\partial p^{\mathrm{T}}}.$$

This object characterizes asset demand within a neighborhood of price vector $p$.

In general, the econometrician observes neither preference parameters $\Theta^i$ nor external parameters $(Y, \pi)$. The demand identification problem thus is to infer combinations of these parameters which determine asset-level demand functions and are invariant to perturbations or counterfactuals of interest.

## 2.2 Consistent Pricing Systems and No Arbitrage

Before analyzing the demand identification problem in more detail, we establish the importance of a consistent pricing system for all possible portfolios of assets. This motivates our use of no arbitrage to structure the pricing system.

**Necessity of internally consistent prices.** A defining feature of portfolio choice is that investors can flexibly bundle and unbundle assets to achieve desired payoff processes. For example, two assets with state-contingent payoffs $[1, 1]$ and $[1, 0]$ can be combined into a portfolio with payoff $[0, 1]$, or indeed *any* payoff in $\mathbb{R}^2$. Given continuous choice over assets, investors can thus choose among a *continuum* of potential payoff vectors whose mapping into portfolios depends on the unobserved payoff matrix $Y$. To permit inference about preferences from portfolio holdings, the econometrician must therefore impose a priori assumptions on the pricing system which can be used to construct prices for all feasible payoff vectors.

**No arbitrage ensures consistent pricing.** The canonical approach to ensuring consistent pricing in financial markets is the principle of *no arbitrage*, which asserts that pricing system should not permit trading strategies which offer "something for nothing." In particular, this principle states that there should not exist feasible trading strategies which offer strictly positive payoff at some date while offering weakly positive payoffs in all other states and dates.

**Definition 1 (No Arbitrage)** *There is no arbitrage if there does not exist a portfolio $a^* \in \mathbb{R}^J$ such that $Y^{\mathrm{T}} a^* \geq 0$ and either (i) $p \cdot a^* \leq 0$ and $(Y^{\mathrm{T}} a^*)_z > 0$ for some $z$ or (ii) $p \cdot a^* < 0$.*

No arbitrage is a weak restriction which rules out the existence of profitable trading strategies that would be exploited by any investor with increasing preferences. Nevertheless, it is sufficient to ensure consistent pricing of *all* portfolios. In particular, the fundamental theorem of asset pricing shows that no arbitrage is equivalent to the existence of a vector of *state prices* $q \in \mathbb{R}^Z$ which serve as reference prices from which one can recover any asset price. State prices can be interpreted as the marginal cost of unit payoff in a given state of the world, so that asset prices are payoff-weighted sums of state prices. See Duffie (2001) for the proof.

**Theorem 0 (Fundamental Theorem of Asset Pricing)** *Let $\Phi^i = \mathbb{R}^J$. There is no arbitrage if and only if there exist state prices $q \in \mathbb{R}^Z_{++}$ such that asset prices satisfy*

$$p = Yq. \tag{2}$$

Under no arbitrage, the prices of all traded portfolios are thus linear combinations of state prices, with weights determined by state-contingent payoffs. This yields a consistent pricing system that links asset prices to the marginal cost of what they offer to investors, namely state-contingent payoffs. Because of its foundational role in modern asset pricing, we adopt this pricing system as well.

**Existence of Optimal Portfolios and Dimension Reduction.** No arbitrage provides two additional advantages for asset demand analysis. First, it allows researchers to combine individual asset positions into aggregated portfolios while ensuring that the demand for and price of the bundle is internally consistent. Such portfolio aggregation is a foundational tool. For example, Koijen and Yogo (2019) aim to summarize asset demand using a small number of asset characteristics. Second, a canonical result—recapitulated below—shows that no arbitrage ensures the *existence* of a solution to the portfolio choice problem. Naturally, existence is a prerequisite for demand analysis in asset markets. See Duffie (2001) for the proof.

**Proposition 0 (No arbitrage and the investor's problem)** *Let $\Phi^i = \mathbb{R}^J$. Then there is a solution to (PCP) if and only if there is no arbitrage.*

Taken together, the principle of no arbitrage thus serves to ensure the internal consistency of asset demand systems while imposing only weak assumptions. While some trading frictions could prevent no arbitrage from holding exactly, as long as the frictions do not completely rule out general equilibrium price adjustments, our arguments hold.

**Redundant assets.** No arbitrage pricing is particularly salient in the presence of redundant assets (i.e., when there are multiple portfolios that deliver identical cash flow processes). In such cases, an arbitrarily small change in the price of a redundant asset immediately triggers an arbitrage opportunity with discontinuous changes in demand functions (see Example 3 in Appendix D.1 for an illustration). If such arbitrages do persist on the equilibrium path, it is infeasible to identify asset-specific demand functions (i.e., the slope of asset quantities with respect to variation in a single price) for redundant assets from observational data. For the remainder, we therefore focus on the case without redundant assets.

**Assumption 1 (No redundant assets)** $Z \geq J$ *and* $\operatorname{rank}(Y) = J$.

## 2.3 Outline of the Argument

Our argument proceeds in three steps. *First* (Section 3), we derive the decomposition $\mathcal{A}^i = (Y^+)^\mathrm{T}\mathcal{D}^i Y^+$, which separates asset demand slopes into fundamental preferences over payoffs and a latent mapping $Y^+$ from payoff demand into portfolios. Since $Y^+$ is unobservable in principle, asset demand functions are not structural with respect to the belief revisions that occur during ordinary market functioning.

*Second* (Section 4.1), we ask whether the latent mapping can be bypassed through supply shocks and show that it cannot: no-arbitrage forces prices of all payoff-overlapping assets to move jointly, so supply shocks generically produce price variation that is misaligned with the ceteris paribus requirements of demand identification.

*Third* (Section 4.3), we ask whether imposing factor structure is sufficient to make progress and show that it is not: using random matrix theory, we establish that $Y^+$ is generically ill-conditioned, with individual elements having the wrong sign with probability approaching one-half in large economies. Section 5 collects these findings into the trilemma and draws implications for the interpretation of estimated demand elasticities.

# 3 The Problem of the Unobservable Mapping

To understand basic properties of asset demand function, we begin by establishing a general decomposition of asset demand functions into two components: fundamental demand over state-contingent consumption, and a *latent mapping* which determines the asset portfolio required to achieve target state-contingent consumption profile. The first component reflects the standard notion of demand that is common to all demand analysis, in that it reflects preference parameters and willingness to pay for consumption in different states of the world. The latent mapping is unique to the case of financial assets, in that it reflects beliefs over future payoffs which guide how different assets must be combined with each other.

Our main result is that the latent mapping is fundamentally *unobservable*. This has two implications: (i) fundamental demand can never be identified from data on portfolio choices, and (ii) asset demand can be given a structural interpretation (that is, being invariant to perturbations) only if the generalized inverse of the payoff matrix $Y$ remains fixed. Since payoff and forecast revisions occur across essentially all time horizons and financial markets, this suggests that asset demand is not a structural object in essentially all settings of interest.

## 3.1 Demand Decomposition and Non-Separability

To arrive at our decomposition, we must define an appropriate notion of demand functions for state-contingent consumption. We thus consider the following *con-*

*sumption choice problem* given a vector of state prices.

$$\max_{c^i \in \mathcal{C}^i} \quad (1-\delta^i)u(c_0^i) + \delta^i \pi \cdot u(c^i) \qquad \text{(CCP)}$$
$$\text{s.t} \quad c_0^i + q \cdot (c^i - w^i) \leq w_0^i + q \cdot Y^{\mathrm{T}} e^i.$$

A solution to problem (CCP) is a set of $Z$ state-contingent consumption functions $c^i$ which map the vector of state prices into a consumption profile. The budget constraint states that net purchases of consumption (over and above endowments) must be equal to the value of the investor's asset endowments. The set $\mathcal{C}^i$ encodes constraints on feasible consumption choices due to, e.g., incomplete markets or other portfolio constraints. By analogy with asset demand functions, we can therefore study the derivative of the Marshallian *consumption* demand curve with respect to *state prices*.

**Definition 2 (Consumption demand)** *The slope of consumption demand is a $Z \times Z$ matrix of Marshallian consumption demand slopes with respect to state prices,*

$$\mathcal{D}^i \equiv -\frac{\partial c^i}{\partial q^{\mathrm{T}}}$$

*which depends on preference parameters, state probabilities $\pi$ and state prices but is independent of the payoff matrix conditional on investor i's asset span $\mathcal{S}^i$ defined by (1).*

Consumption demand functions thus have a certain robustness property with respect to small perturbations of the payoff matrix $Y$. When these perturbations do not affect the set of attainable payoffs, they do not alter consumption plans. As we will see, this is *not* the case for asset demand functions, which do depend directly on perturbations of $Y$.

However, consumption demand functions are not observable because neither state prices nor the payoff process are directly observed by the econometrician. However, they can be linked to observable portfolios and asset prices using the portfolio choice problem and no arbitrage pricing. This follows directly from the chain rule. Specifically, the consumption process induced by portfolio is $c^i = Y^{\mathrm{T}} a^i + w^i$. The portfolio yielding a target consumption process $c^*$ thus is $a^*(c^*) = (Y^+)^{\mathrm{T}}(c^* - w^i)$, where $Y^+$ is the generalized Moore-Penrose inverse of $Y$. Differentiating asset demand with respect to asset prices therefore yields

$$\frac{\partial a^*(c^*)}{\partial p^{\mathrm{T}}} = (Y^+)^{\mathrm{T}} \frac{\partial c^*}{\partial p^{\mathrm{T}}} = (Y^+)^{\mathrm{T}} \frac{\partial c^*}{\partial q^{\mathrm{T}}} \frac{\partial q^{\mathrm{T}}}{\partial p^{\mathrm{T}}},$$

where the second equality follows from the chain rule. Next, observe that no arbitrage implies that asset prices $p$ are related to state prices through $p = Yq$, and thus state prices can be related to asset prices through $q = Y^+ p$. Applying this observation to $\frac{\partial q^{\mathrm{T}}}{\partial p^{\mathrm{T}}}$ then yields the following decomposition.

**Proposition 1 (Demand Decomposition)** *If asset prices satisfy no arbitrage, then*

$$\mathcal{A}^i = (Y^+)^{\mathrm{T}} \mathcal{D}^i Y^+. \tag{3}$$

*If preferences over state-contingent consumption are separable across states as in* (PCP) *and there are no binding portfolio constraints, then one can further decompose*

$$\mathcal{A}^i = (Y^+)^{\mathrm{T}} \Pi^{-1} \tilde{\mathcal{D}}^i Y^+, \tag{4}$$

*where* $\Pi$ *is a diagonal matrix of state probabilities and* $\tilde{\mathcal{D}}^i$ *depends only on preferences.*

The decomposition reveals that asset demand is generically *non-separable*: given a desired consumption process, the optimal position in any given asset is jointly determined by the payoff processes of *all* assets in the choice set. In particular, changing the payoff characteristics (state-contingent payoffs) of any given asset generically alters several elements of the inverse payoff matrix, triggering changes in the optimal demand for other assets.

A particularly stark illustration of this fact is that whether two assets are substitutes or complements depends on the attributes of other assets in the choice sets. To the best of our knowledge, this issue is distinct from other settings in industrial organization, which may consider flexible specifications in which two goods can be substitutes or complements, but these parameters are invariant to attributes of other goods. This however is a central feature of the portfolio choice.

We illustrate this with a simple 3 asset example, which satisfies the condition that any perturbations of the payoff structure do not change the asset span.

**Example 1 (Complementarity and substitutability depend on other assets)** *Consider a three-asset, three-state economy with the payoff matrix $Y$, where rows index assets:*

$$Y = \begin{bmatrix} 1 & \epsilon & 0 \\ \epsilon & 1 & 0 \\ \xi & \xi & 1 \end{bmatrix},$$

*where $\xi, \epsilon \in (0,1)$. Thus, we have complete markets. Assuming homogeneous consumption elasticities, the cross-elasticity between Assets 1 and 2 is:*

$$-\frac{\partial a_1}{\partial p_2} = \sum_z (Y^{-1})_{z,1}(Y^{-1})_{z,2} = \frac{\xi^2(1-\epsilon)^2 - 2\epsilon}{(1-\epsilon^2)^2}.$$

*For a given $\epsilon$ we can find the value of $\xi$ that sets the cross-elasticity to zero:*

$$\xi^* = \frac{\sqrt{2\epsilon}}{1-\epsilon}.$$

*Thus Assets 1 and 2 complements for $\xi < \xi^*$ and substitutes for $\xi > \xi^*$.*

Figure 1 illustrates the change in complementarity as a function of $\xi$. To understand the intuition for this example, consider first a very low value of $\xi$. In this case, Asset 3 plays an insignificant role in the payoff of the first two states and Assets 1 and 2 are good hedges for each other in states 1 or 2. This force is stronger when $\epsilon$ is smaller. Next, consider a very large value of $\xi$: Asset 3 now replicates the payoffs of Assets 1 and 2 in states 1 and 2 ever more closely, crowding out the hedging roles of both. When the price of Asset 2 rises, investors turn to Asset 3 as an alternative hedge for state 2, but in doing so they simultaneously acquire exposure to state 1. This reduces their demand for Asset 1 as well. The two assets thus become substitutes not through any direct relationship between them, but because they share a common "competitor." This force is stronger when $\epsilon$ is larger.

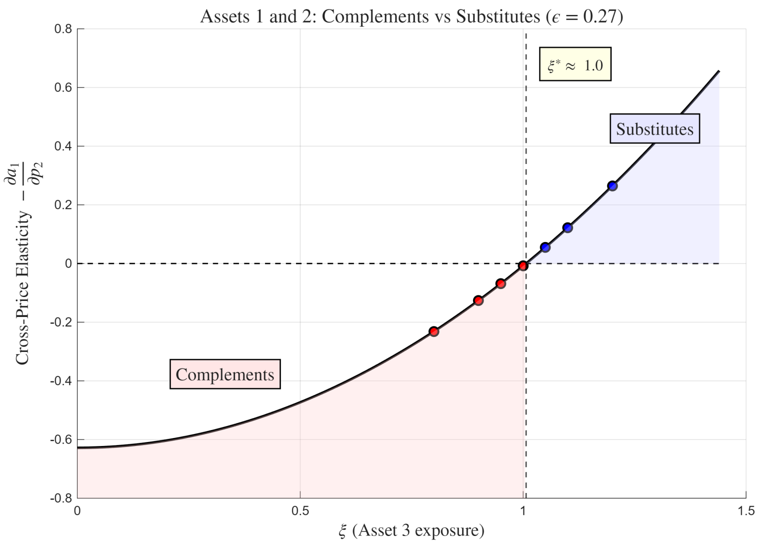


Figure 1: Switch in complementarity given $\xi$ in Example 1

## 3.2 Unobservable Mapping

We have shown that asset demand functions—as well as the degree of complementarity between any two assets—is determined by global properties of the inverse payoff matrix $Y^+$. We next show a fundamental constraint on asset demand analysis: the mapping from fundamental preferences to portfolio choices is unobservable because the generalized inverse payoff matrix $Y^+$ cannot be identified from any finite sample of observed payoffs *even if the payoff matrix is assumed to be stable.*

**Proposition 2 (Non-identification of the Latent Mapping $Y^+$)** *Consider any finite sample of realized asset payoffs $\mathcal{S}$. Then there exist arbitrarily many candidate payoff matrices $Y$ which are observationally equivalent given $\mathcal{S}$ but have different generalized inverses.*

Intuitively, the proposition holds because a researcher can always alter the payoff of a state of the world that has not yet been realized in the data, and altering payoffs in this state changes $Y^+$ but leaves all observed historical returns identical. This fundamental non-identification of $Y^+$ implies that asset demand estimation must reckon with two commingled identification problems: that of fundamental preferences (i.e., the standard identification problem that is common to all demand estimation exercises), and that of the latent mapping linking the primitive object of preferences (payoffs) to observed choices (asset positions).

An important implication of this is that fundamental demand can never be identified from observed portfolio choices. The reason is that observed choices and preferences pertain to different objects, namely assets and payoffs, and the mapping between the two is unobservable. As such, observed asset choices can always be rationalized by different combinations of preferences and the latent mapping.

**Corollary 1 (Non-recoverability of $\mathcal{D}^i$)** *For any $\mathcal{A}^i$, the consumption demand slope $\mathcal{D}^i$ is generically not uniquely determined absent knowledge of $Y^+$.*

**Proof.** The result follows because $Y^+$ cannot be identified from data (Proposition 2) and different inverses $Y^+$ generically induce different $\mathcal{D}^i$ through equation (3) even for the same $\mathcal{A}^i$. ■

## 3.3 Asset Demand is Not Structural under Weak Conditions

We now use our decomposition to provide conditions under which demand functions are *structural* in the sense of Hurwicz (1962) and Marschak (1953): invariant to relevant perturbations to the economic environment. To formalize this, we allow all model primitives to vary with a latent variable $\omega \in \Omega$ which we call the *economic environment*. We then say that a demand function is *structural* with respect to a class of perturbations if it is invariant across all $\omega$ within the class.

**Definition 3 (Structural Demand)** *Let $\mathcal{P} \subseteq \Omega \times \Omega$ be a class of perturbations, where each $(\omega, \omega') \in \mathcal{P}$ represents a transition from environment $\omega$ to $\omega'$. We say that a demand function $\mathcal{F}$ is structural with respect to $\mathcal{P}$ if:*

$$\mathcal{F}(\omega) = \mathcal{F}(\omega') \quad \textit{for all } (\omega, \omega') \in \mathcal{P}.$$

The structural properties of demand thus depend on the perturbations under consideration. While these are often application specific, we emphasize perturbations that are of particular interest to financial markets: *unobserved* revisions to payoff expectations $Y$ and state probabilities $\pi$ which occur in the regular course of financial market operations. If asset demand is not structural with respect to these perturbations, there can be no guarantee that demand is structural with respect to other interventions or perturbations either.

Our decomposition then implies two main results: consumption and demand functions cannot be jointly be structural with respect to perturbations that vary the payoff matrix, and asset demand functions can be structural only if consumption preferences are a specific function of payoff parameters themselves. This contradicts the fundamental dichotomy between investor preferences and asset characteristics which permits demand analysis in the first place.

**Proposition 3 (Non-structural demand)** *Consider a class of perturbations with unobservable changes to $Y$ or $\pi$. Then:*

1. *Asset demand function $\mathcal{A}^i$ and $\mathcal{D}^i$ cannot both be structural.*
2. *If asset demand functions are to be structural, then fundamental preferences must respond to any shock to future prices, dividends, or probabilities. In particular, we must have $\mathcal{D}^i = Y^T B^i Y$ for an arbitrary $B^i$, so that $\mathcal{A}^i = B^i$.*

That is, one cannot maintain the assumption that demand functions are structural unless one is also willing to make the assumption that probabilities and payoffs must remain fixed. This presents a sharp constraint on asset demand analysis because the assumption is (i) unverifiable, and (ii) it rules out that asset demand is structural with respect to interventions whose *goal* is to shift payoff expectations. This includes quantity-based policy experiments such as quantitative easing or foreign exchange interventions which aim to influence broader economic conditions. Such applications are of central interest to much of the literature.

# 4 The Problem of Identification

We now establish conditions under which asset demand can be identified from observational data on portfolio holdings and asset prices. We provide two main results. First, we show that the canonical approach to estimating demand curves—namely, exogenous shocks to the supply of a given asset—does not identify individual demand curves because they generically fail to produce ceteris paribus variation in a single asset price. Second, we ask whether multiple supply shocks can be combined to identify the $J \times J$ matrix of demand slopes $\mathcal{A}^i$. We show this to be possible if (i) the econometrician observes at least $J$ linearly independent, exogenous shocks to the price vector and associated portfolio responses and (ii) the unverifiable assumption that the latent mapping $Y^+$ remains fixed across all such experiments. While the requirement of $J$ shocks is a standard rank condition that also arises other settings with demand complementarities, the second is specific to financial markets, where the mapping between assets and characteristics is unobservable. Since forecast revisions are a defining function of financial market, this assumption is both strong and unverifiable in essentially all settings of interest. Furthermore, there is a natural tension: many experiments which shift prices today are likely to shift future prices as well, but this would also change $Y^+$, which contains resale prices.

## 4.1 Supply Shocks Produce Misaligned Price Variation

The canonical approach to estimating demand curves is to rely on supply shocks to provide suitably exogenous variation in a given price. With demand complementarities and long-lived assets, a central endogeneity concern is that the supply shock creates correlated changes in other asset prices (so-called *price spillovers*), thereby contaminating the price variation needed to identify a particular demand slope. We now show that this problem is generic under no arbitrage: except in the implausible knife-edge case where assets never pay off in overlapping states of the world, even perfectly exogenous supply shocks must always induce price spillovers.

**Ideal experiment.** We first define the price variation necessary to identify a particular asset demand slope. Given that asset demand exhibits demand complementarities and depends on expected future resale prices, we require that all other prices and all future payoffs must remain unchanged. To understand whether such variation is obtainable under even ideal conditions, we study the price changes induced by *perfectly exogenous supply shocks that leave all future payoffs unchanged.*

Under preferences over payoffs, it is useful to describe the ideal experiment in terms of state prices. The investor observes asset prices $p$ and payoff matrix $Y$. Under no arbitrage, prevailing asset prices imply the state price vector

$$q = Y^{+}p, \tag{5}$$

where $Y^{+}$ is the Moore-Penrose pseudo-inverse of $Y$.[1] A hypothetical *pure price shock* to asset $j$ thus induces a specific state price change which is fully determined by $Y^{+}$.

**Lemma 1 (State price changes in the ideal experiment)** *Let $v_j$ denote the unit vector in $\mathbb{R}^J$ with $1$ in the $j$-th position and zeros elsewhere. Then the changes in state prices given the exogenous variation in a single price $p_j$ are*

$$\Delta \mathbf{q}_j^{\text{ideal}} \equiv \frac{\partial q}{\partial p_j} = Y^{+}v_j.$$

The assertion follows immediately from equation (5). Identifying asset demand thus requires shocks which generate the state price variation $\Delta \mathbf{q}_j^{\text{ideal}}$ associated with the ideal experiment.

**Measurement using supply shocks.** Since pure price shocks are rarely observed, we now study whether even perfectly exogenous supply shocks can generate the variation required by the ideal experiment. To do so, we must describe how supply shocks affect state prices in a general class of models. Given the standard assumption of risk-averse preferences with decreasing marginal utility, we study settings

[1] If $Y$ is square, as when markets are complete, then $Y^{+} = Y^{-1}$ and there is a unique vector of state prices. If markets are incomplete ($J < Z$), then there exist multiple feasible state price vectors. As is standard, we select the minimum norm solution with pseudo-inverse $Y^{+} = Y^{\mathrm{T}}(YY^{\mathrm{T}})^{-1}$, which also arises endogenously in our demand decomposition.

in which a positive supply shock to asset $j$ must reduce *state prices* in all states where asset $j$ has a strictly positive payoff. We call this property *downward-sloping consumption demand*. Since our definition is written directly in terms of state prices, it must be understood purely in terms of fundamental preference parameters.

**Definition 4 (Downward-sloping consumption demand)** *Let $E \equiv (E_j)_{j=1}^{J} \in \mathbb{R}_{++}^{J}$ denote the vector of aggregate asset endowments. An economy has downward-sloping consumption demand if there exists a $Z \times Z$ matrix $V$ with strictly positive diagonal elements such that*

$$\Delta \mathbf{q}_j^{\text{supply}} \equiv \frac{\partial q}{\partial E_j} = -V y_j^{\mathrm{T}} \qquad \textit{for all assets } j,$$

*where $y_j^{\mathrm{T}}$ is the transpose of the $j$-th row $y_j \equiv (y_j(z))_{z=1}^{Z}$ of $Y$.*

In this definition, $V$ captures the marginal change in the market-wide pricing kernel, which is taken as given by each individual investors. That $V$ has strictly positive diagonal elements then captures our assumption that increases in the supply of state-contingent payoffs reduce the marginal price of these payoffs.

Definition 4 imposes no assumptions on $V$'s off-diagonal entries, which capture potential *direct* preference-based spillovers across state prices in response to a supply shock. The existence of such spillovers depends on the economic model. The canonical model with additively separable utility over consumption (as in Section 2) has zero off-diagonal elements. Example 2 below illustrates this with a representative investor. Non-separable models such as recursive utility (Epstein and Zin, 1989; Kreps and Porteus, 1978) or more general aggregators instead generally imply non-zero off-diagonal elements. Since spillovers are the main threat to identification, the identification challenge is generically *weaker* when there are no preference-based spillovers in state prices. To provide favorable conditions for identification, we thus assume that no such spillovers exist.[2]

[2] The only case in which non-diagonal $V$ can undo price spillovers occurs when the off-diagonal elements in $V$ exactly offset the cross-asset restrictions implied by no arbitrage. However, $V$ is determined by preferences and aggregate endowments while the no-arbitrage relation depends only on the payoff matrix $Y$. Hence there is no economic reason for such a mechanical offset to occur. More generally, Section 4.3 shows that, for large matrices, the sign of each element of $Y^+$ is close to a coin flip, with odds that depend only on the payoff matrix. Hence small perturbations to the payoff matrix can flip the sign of an element in $Y^+$ without meaningfully altering $V$.

**Assumption 2 (No Direct Spillovers Across State Prices)** *The marginal pricing kernel $V$ is a diagonal matrix. Hence there are no direct state price spillovers.*

**Example 2 ($V$ in an additive separable representative-agent model)** *In a standard representative-agent model with additive separable preferences over consumption, state prices relate to marginal utility over aggregate consumption,*

$$\frac{\partial q_z}{\partial E_j} = \frac{\delta}{1-\delta}\pi_z \frac{u''(C_z)}{u'(C_0)} y_j(z) < 0,$$

*where $C_0$ and $C_z$ are aggregate consumption at date 0 and in state $z$. Thus the marginal pricing kernel is a strictly positive diagonal matrix,*

$$V = -\frac{\delta}{1-\delta}\text{diag}\left(\pi_1 \frac{u''(C_1)}{u'(C_0)}, \ldots, \pi_z \frac{u''(C_z)}{u'(C_0)}, \ldots, \pi_Z \frac{u''(C_Z)}{u'(C_0)}\right).$$

**Supply Shocks Do Not Generate the Ideal Experiment.** We now show that supply shocks generically fail to produce the ideal experiment. We consider two definitions of alignment between supply shocks and the ideal experiment: (i) that induced state price changes are identical to those of the ideal experiment (up to a scalar multiple), and (ii) that the induced state price changes are of the same *sign*.

While the first condition is required to exactly identify a demand slope, the second captures the much weaker requirement that the supply shock should at least trigger *directionally consistent* changes in the cost of consumption. If this condition fails, there are state-contingent payoffs which should become more expensive in the ideal experiment but actually become cheaper upon a supply shock. Such errors can lead to large biases when estimating substitution patterns.

**Condition 1 (Identical variation)** *A supply shock to asset $j$ generates the ideal state price variation for asset $j$ if there exists some scalar $k_j$ such that $\Delta\mathbf{q}_j^{\text{ideal}} = k_j \Delta\mathbf{q}_j^{\text{supply}}$.*

**Condition 2 (Variation of the same sign)** *The supply shock generates state price variation of the same sign if $\Delta\mathbf{q}_j^{\text{ideal}}$ has the same sign as $\Delta\mathbf{q}_j^{\text{supply}}$ element by element.*

We can then state our main result of this section, which is that Conditions 1 and 2 are satisfied only under highly restrictive, non-generic conditions on the payoff matrix. In particular, for every state of the world there must exist a *unique* asset which offers a positive payoff in that state. That is, in order to satisfy the minimal requirement that the induced state price variation is of the same sign as in the ideal experiment, there must be no assets with overlapping payoffs.

**Definition 5 (Overlapping payoffs)** *Assets $j$ and $j'$ have overlapping payoffs if there exists at least one state of the world $z$ such that $y_j(z) > 0$ and $y_{j'}(z) > 0$.*

**Theorem 1 (Supply Shocks Induce Misaligned Price Variation)** *If Conditions 1 or 2 are satisfied, then $YY^{\mathrm{T}}$ is diagonal, and:*

*(i) If $YY^{\mathrm{T}}$ is diagonal, then there are no assets with overlapping payoffs.*

*(ii) If markets are complete, then $YY^{\mathrm{T}}$ is diagonal if and only if $Y$ is diagonal up to permutations.*

The conditions set out in Theorem 1 are unrealistic for almost all standard financial assets, as they require that there are no states of the world in which any given asset has positive payoffs while another asset also has positive payoffs. This is plainly violated for generic payoff distributions where overlap in payoffs (that is, concurrent non-zero dividends and/or resale values) is the norm, not the exception. It is therefore striking that, outside of these knife-edge restrictions, supply shocks do not even guarantee *directional* alignment with the ideal experiment. As such, supply shocks generically fail to identify structural asset-level demand slopes in essentially all settings of interest. In Appendix B, we also illustrate our findings using a simple example economy based on Fuchs, Fukuda, and Neuhann (2025).

**Asset-by-asset misalignment.** Theorem 1 shows that there must be misaligned price variation for at least one asset in the payoff menu (that is, at least one row of the payoff matrix). This allows the possibility that misalignment may not occur for *some* assets (although this cannot be verified without knowledge of the payoff matrix). The next proposition further strengthens our result by providing a weak condition for which misaligned price variation is guaranteed for *every asset.*

**Proposition 4** *If each column of $Y$ has at least two strictly positive elements, then each column of the Moore-Penrose inverse $Y^+$ contains at least one negative element: for each $j \in \{1, \ldots, J\}$, there exists at least one $z \in \{1, \ldots, Z\}$ such that $(Y^+)_{z,j} < 0$.*

## 4.2 Identification from Multiple Shocks Requires Fixed Payoffs

So far we have shown that even perfectly exogenous supply shocks generically produce price variation that is contaminated by cross-asset spillovers. While this means that one cannot identify specific asset demand curves from asset-level supply shocks, one might yet jointly identify the entire matrix $\mathcal{A}^i$ by combining sufficiently many independent supply shocks.

We now show that this is the case only if (i) the econometrician has access to at least $J$ exogenous supply shocks which provide linearly independent variation in the price vector—a standard rank condition—and (ii) that the unobservable mapping $Y^+$ remains fixed across all supply shocks. The second condition is a central challenge in financial markets, where the mapping from goods to characteristics is unobserved and subject to frequent revisions over essentially any time horizon.

**Setting.** We consider an idealized scenario in which the econometrician observes $K$ independent shocks to the price vector and interprets investors' portfolio responses under a set of maintained assumptions $\mathcal{M}$. We refer to each price shock as an *experiment*, and assume that they are linearly independent. Since there are $J$ assets, it is sufficient for our argument to consider the case where the number of experiments is weakly smaller than the number of assets: $K \leq J$. Let $\mathcal{O}_P$ denote the observed price changes, and $\mathcal{O}_{a^i}$ the observed portfolio response for investor $i$. We have the following standard definition of identification.

**Definition 6 (Identified Demand Functions)** *Two asset demand functions $a^i(\cdot \mid \Theta^i, Y, \pi)$ and $a^i(\cdot \mid \tilde{\Theta}^i, \tilde{Y}, \tilde{\pi})$ are observationally equivalent given data $(O_P, O_{a^i})$ and maintained* *assumptions $\mathcal{M}$ if both satisfy $\mathcal{M}$ and, for each observed price vector $p \in O_P$,*

$$a^i(p \mid \Theta^i, \Upsilon, \pi) = a^i(p \mid \tilde{\Theta}^i, \tilde{\Upsilon}, \tilde{\pi}) = O_{a^i}.$$

*Demand slope $\mathcal{A}^i$ is identified under maintained assumptions $\mathcal{M}$ if all observationally equivalent demand functions imply the same slope:*

$$\mathcal{A}^i(\Theta^i, \Upsilon, \pi) = \mathcal{A}^i(\tilde{\Theta}^i, \tilde{\Upsilon}, \tilde{\pi}).$$

We have already established that $\mathcal{D}^i$ is not identified because $\Upsilon^+$ is not observable. We now show that asset demand is identifiable only if the econometrician observes $J$ experiments *and* maintains the unverifiable assumption that the unobservable mapping $\Upsilon^+$ is fixed across all experiments.

**Proposition 5 (Necessary Conditions for Identification of $\mathcal{A}^i$)** *Let $(O_P, O_{a^i})$ denote a data set of $K$ observed price vectors and associated portfolio positions for investor $i$. Generically, $\mathcal{A}^i$ is identified only if the econometrician observes $K = J$ independent experiments and maintains the assumption that $\Upsilon^+$ is fixed across all experiments.*

Identification thus necessarily relies on the unverifiable assumption that the unobservable inverse payoff matrix is fixed across at least $J$ shocks to the price vector. This finding relates our work to Haddad, He, Huebner, Kondor, and Loualiche (2025), who aim to recover an asset elasticity matrix without a fully specified structural model by combining (i) cross-sectional variation in asset-level holdings, and (ii) time series shocks to the prices of certain factor portfolios. Proposition 5 shows that this approach identifies a well-defined and stable elasticity matrix only if the unobserved payoff matrix remains fixed over time. However, changes in expected payoffs are a natural byproduct of financial market activity, including risk sharing, investment, or price discovery. (Section 4.3 shows that even approximate stability in the payoff matrix does not ensure a stable *inverse* payoff matrix, which is what matters for the stability of asset demand functions.)

**Implications for Instrument Validity.** Propostion 5 also sharpens the conditions for instrument validity in asset demand estimation: in addition to providing suitably exogenous variation in prices *today*, the instrument must *not* alter future expected payoffs. This rules out any shock to current prices that simultaneously induces changes in expected asset payouts or resale prices, such as central bank asset purchases or index inclusion events. The former are often used precisely to influence expectations over future market conditions, whereas index inclusion is known to alter return comovements and the level of prices.

## 4.3 The Latent Mapping is Ill-conditioned and Unstable

We have shown that asset demand functions are neither structural nor identifiable unless the latent inverse payoff matrix is assumed to remain fixed. These issues arise because one can at best observe a subset of *realized* payoffs, but not the matrix of *expected* payoffs, although it is the latter which matters for asset demand.

A potential solution to this problem is to use statistical information on realized returns to impose structure on the payoff process, and to hope that this structure is sufficient to ensure a stable mapping from preferences to asset holdings. The predominant approach in the literature is to impose a factor structure on payoffs, whereby asset returns are driven by a relatively small number of common factors. We therefore use random matrix theory to analyze the asymptotic properties of factor-structured payoff processes.

We find that the inverse payoff matrix is ill conditioned: the *sign* of any given element of the inverse payoff matrix is a coin flip. Hence even well-behaved factor structures yield poorly behaved latent mappings that can flip signs even with small changes to the payoff process. Monte Carlo simulations show that our theoretical limit results hold even for small $J$ and $Z$. Appendix C shows that similar results hold in data from the S&P500. We conclude that imposing realistic statistical structure on the payoff matrix is *not* sufficient to ensure a well-behaved mapping from preferences to asset positions.

**Random matrix approach.** Because true payoffs are latent, we study random draws of $Y$ generated from a factor structure. This allows us to characterize, in probability, the expected sign structure of its pseudo-inverse. Specifically, let payoff matrix $Y \in \mathbb{R}^{J \times Z}$ with $J \leq Z$ be defined by the following single factor structure, where $y_{j,z}$ represents the payoff of asset $j$ in state $z$:

$$y_{j,z} = \alpha_j + \beta_j f_z + \varepsilon_{j,z} = \underbrace{\alpha_j + \beta_j \overline{f}}_{\equiv \gamma_j} + \beta_j (f_z - \overline{f}) + \varepsilon_{j,z}, \quad \text{where } \overline{f} \equiv \mathbb{E}[f_z].$$

The analysis extends to multi-factor processes: see Remark 2 in Appendix A.4. As before, let $Y^+$ denote the Moore-Penrose pseudo-inverse of $Y$.[3] We impose the following assumptions.

**Assumption 3 (Factor structure)** *$(\alpha_j, \beta_j)_j$ are i.i.d., independent of $(f_z)_z$ and $(\varepsilon_{j,z})_{j,z}$, with finite second moments.*

**Assumption 4 (Factor returns)** *$(f_z - \overline{f})_z$ are i.i.d. with bounded, continuous, and symmetric densities around 0, and $\sigma_f^2 \equiv \mathbb{V}[f_z] < \infty$.*

**Assumption 5 (Idiosyncratic shocks)** *$(\varepsilon_{j,z})_{j,z}$ are i.i.d. across $(j, z)$ with bounded, continuous, and symmetric densities around 0, and $\sigma_\varepsilon^2 \equiv \mathbb{V}[\varepsilon_{j,z}] > 0$. Factors and errors are mutually independent.*

**Population objects and sequential limits.** The properties of small random matrices are difficult to characterize with any generality. Theorem 2 thus considers the sequential limit $Z \to \infty$ followed by $J \to \infty$.[4] As $Z \to \infty$ the sample Gram matrix $G_Z \equiv \frac{1}{Z} YY^{\mathrm{T}}$ converges almost surely to the population second-moment matrix,

$$\Sigma = \gamma\gamma^{\mathrm{T}} + \sigma_f^2 \beta\beta^{\mathrm{T}} + \sigma_\varepsilon^2 I_J = \sigma_\varepsilon^2 I_J + UU^{\mathrm{T}}, \quad \text{where } U \equiv \begin{bmatrix} \gamma & \sigma_f \beta \end{bmatrix} \in \mathbb{R}^{J \times 2}.$$

[3] The rank of $Y$ equals $J$ almost surely under Assumption 5, since the set of $J \times Z$ matrices with $\mathrm{rank}(Y) < J$ has measure zero. Hence $Y^+ = Y^{\mathrm{T}}(YY^{\mathrm{T}})^{-1}$ a.s.

[4] The sequential limit $Z \to \infty$ followed by $J \to \infty$ is adopted for transparency of proof, not out of necessity. Our numerical simulations suggest that $Z$ and $J$ could be taken to infinity at the same time yet allowing for that would significantly complicate the proof.

The sign of $(Y^+)_{z,j}$ is asymptotically determined by the sign of the population quantity $(\Sigma^{-1}y_z)_j$, which can also be written as $v_j^\top \Sigma^{-1} y_z$ where $v_j \in \mathbb{R}^J$ is the $j$-th unit vector. We thus work directly with the following population objects:

1. *Individual sign probability.* For each fixed $(j, z)$, define

$$\pi(j,z) \equiv \lim_{Z\to\infty} P\left((Y^+)_{z,j} > 0\right).$$

2. *Sign-agreement frequency.* Let $Y$ and $\tilde{Y}$ be two payoff matrices generated by the same factor loadings $(\alpha_j, \beta_j)_j$ and factor realizations $(f_z)_z$ but independent idiosyncratic shocks $(\varepsilon_{j,z})$ and $(\tilde{\varepsilon}_{j,z})$. Define the sign agreement frequency

$$q(J,Z) \equiv \frac{1}{JZ}\sum_{j,z} \mathbf{1}\left(\text{sign}(Y^+)_{z,j} = \text{sign}(\tilde{Y}^+)_{z,j}\right),$$

where $\mathbf{1}(\cdot)$ denotes the indicator function, and its population limit

$$q(J) \equiv \operatorname*{plim}_{Z\to\infty} q(J,Z).$$

Theorem 2 characterizes the limits of these population objects as $J \to \infty$. We also use simulations to validate our results away from these limits.

**Theorem 2 (Sign Instability of $Y^+$)** *Under Assumptions 3–5, for almost every realization of $(\alpha_j, \beta_j)_j$, the following hold.*

*(i)* ***Individual coin flip.*** *For each fixed asset $j$ and state $z$,*

$$\lim_{J\to\infty} \pi(j,z) = \lim_{J\to\infty} P\left((\Sigma^{-1}y_z)_j > 0\right) = \frac{1}{2}.$$

*Moreover, the distribution of $(\Sigma^{-1}y_z)_j$ is continuous and centered at zero in the limit, so the positive and negative tails are mirror images of equal magnitude.*

*(ii)* ***Factor structure knowledge is insufficient.*** *Let $Y$ and $\tilde{Y}$ be two payoff matrices generated by the same factor loadings $(\alpha_j, \beta_j)_j$ and factor realizations $(f_z)_z$ but independent idiosyncratic shocks $(\varepsilon_{j,z})$ and $(\tilde{\varepsilon}_{j,z})$. The population sign-determining vari-*

*ables* $(\Sigma^{-1}y_z)_j$ *and* $(\Sigma^{-1}\tilde{y}_z)_j$ *are asymptotically independent for each fixed* $(z, j)$, *each with limiting sign probability* $\frac{1}{2}$. *Consequently,*

$$\lim_{J\to\infty} q(J) = \frac{1}{2}.$$

Knowledge of statistical properties of the return process is thus *not* sufficient to guarantee a well-behaved mapping from preferences to asset holdings. To the contrary, the latent mapping is generally ill-conditioned, with the sign of any given element being a coin flip. Hence the misalignment between supply shocks and the ideal experiment is pervasive for realistic payoff processes, and cannot be corrected for by controlling for factor exposures.

**Calibration and Numerical Exploration** Our theoretical results consider the limit $J, Z \to \infty$. We now study the behavior outside the limit using Monte Carlo simulations with payoff parameters that generate a share of idiosyncratic risk roughly consistent with the empirical data. Concretely, we assume:[5]

$$\alpha_j \sim \mathcal{U}[10, 20], \qquad f_z \sim \mathcal{N}(1, \sigma_f^2) \quad \text{with} \quad \sigma_f = \frac{1}{2},$$
$$\beta_j \sim \mathcal{U}[0.5, 1.5], \qquad \varepsilon_{j,z} \sim \mathcal{N}(0, \sigma_\varepsilon^2) \quad \text{with} \quad \sigma_\epsilon = 1.$$

Figure 2 shows that the theoretical prediction for $Z \to \infty$ and large $J$ can perform remarkably well even for moderate values of $Z$ and small $J$. The left panel depicts the proportion of times any individual element of the matrix $Y^+$ is positive.[6] Thus, given $Y > 0$, almost half the elements of $Y^+$ have the wrong sign. The right panel compares the signs of $(Y^+)_{z,j}$ and $(\tilde{Y}^+)_{z,j}$ where both $Y$ and $\tilde{Y}$ are generated from the same factor model and are thus indistinguishable in practice. We again observe that the signs coincide only 50% of the time, implying that there is no systematic way to correct for these sign errors.

[5]The high values of $\alpha_j (\sim \mathcal{U}[10, 20])$ effectively guarantee that all entries of $Y$ are positive. Note, however, that our theoretical results do not require that. Also, truncation of the normal distributions for $f$ and $\varepsilon$ (to force $Y$ to be always non-negative) do not qualitatively alter our results.

[6]For each $Z \in \{150, 300\}$, we vary the number of assets $J \in \{2, 4, \ldots, 100\}$. We took the average of 1000 runs (of the Monte Carlo simulations). Figure 2 also depicts the 95% confidence interval.

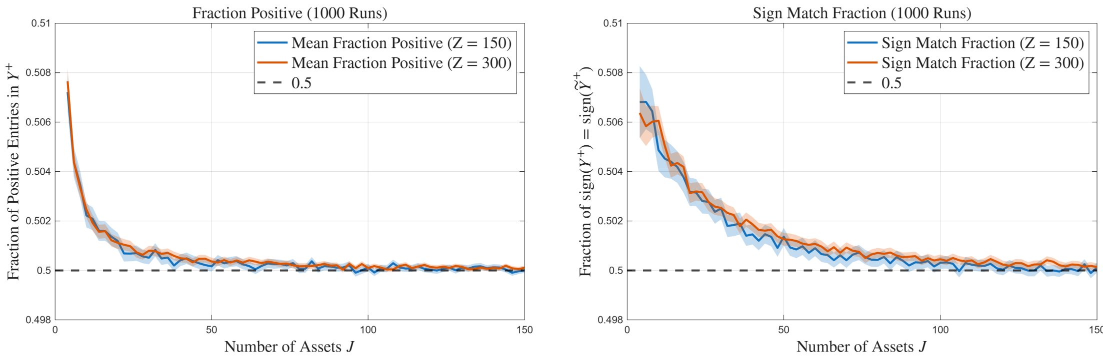


Figure 2: The Monte Carlo Simulation Results for Theorem 2. The left panel displays the empirical frequency of positive entries in $Y^+$, while the right panel shows the empirical frequency of sign matches between $Y^+$ and $\tilde{Y}^+$. Both panels report the results for $Z \in \{150, 300\}$ across various values of $J$, based on 1000 runs.

# 5 The Trilemma and its implications

We have established that the two principles of no arbitrage and preferences over payoffs sharply curtail the scope for non-parametric demand analysis in asset markets. We now summarize this result as a trilemma.

**Theorem 3 (Trilemma)** *Given observational data on portfolios, asset prices, and payoffs, one cannot jointly maintain (i) no-arbitrage asset pricing, (ii) investor preferences over payoffs, and (iii) model-free identification of structural asset demand functions.*

None of the stated conditions is easily discarded. No arbitrage is the prototypical internally-consistent pricing system, which ensures existence, consistency, and external validity of demand functions. Payoff-based asset valuation is the basic guiding principle of asset pricing. Since our definition of an asset is entirely generic, our results also apply equally to *portfolios* of primitive assets, which are themselves simply collections of payoffs. This leaves the assumption of constant payoffs. Unfortunately, this assumption is in tension with one of the basic functions of financial markets, which is price discovery—and thus revisions in expected payoffs—in response to news. It also cannot be directly verified in the data.

The trilemma is robust to small departures from its stated conditions. Relaxing no-arbitrage—for instance, by allowing for transaction costs or small arbitrage bands—replaces exact state-price equalities with inequalities but leaves the latent mapping $Y^+$ unobserved and ill-conditioned. Similarly, introducing non-pecuniary tastes or asset-level utility alters the fundamental demand object $\mathcal{D}^i$ but is orthogonal to the identification problem posed by $Y^+$: as long as investors retain any preference over state-contingent payoffs, the decomposition $A^i = (Y^+)^\mathrm{T}\mathcal{D}^iY^+$ remains operative and $Y^+$ remains latent. The cross-asset linkages and sign instability documented in Theorem 2 depend only on the geometry of the payoff matrix, not on the preference specification. A model in which non-pecuniary considerations entirely dominate pecuniary ones is not, in any meaningful sense, a model of asset demand. Hence neither perturbation restores model-free identification of structural asset demand.

Our analysis suggests a critical role for structural models in asset demand estimation. However, since $Y^+$ cannot be identified from data, structurally estimated demand functions will depend sensitively on the assumed payoff structure. For example, Fuchs, Fukuda, and Neuhann (2025) show that the logit asset demand model proposed by Koijen and Yogo (2019) can yield low estimated demand elasticities because of the substitution patterns it assumes. The estimated demand elasticities should therefore be evaluated based on the validity and plausibility of the assumed payoff structure, not on empirical fit, which provides no information about whether the assumed latent mapping is correct. Our decomposition provides a framework for understanding what those restrictions imply and where misspecification is likely to be most consequential.

# References


ALLEN, J., J. KASTL, AND M. WITTWER (2025): "Estimating Demand Systems with Bidding Data," Working Paper. [Cited on page 5.]

BERRY, S. T., AND P. A. HAILE (2021): "Foundations of Demand Estimation," in Handbook of Industrial Organization, vol. 4, pp. 1–62. [Cited on page 4.]

BINSBERGEN, J. H. V., B. DAVID, AND C. C. OPP (2025): "How (Not) to Identify Demand Elasticities in Dynamic Asset Markets," Working paper. [Cited on page 5.]

DUFFIE, D. (2001): Dynamic Asset Pricing Theory. Princeton University Press, third edn. [Cited on page 9.]

EPSTEIN, L. G., AND S. E. ZIN (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," Econometrica, 57(4), 937–969. [Cited on page 20.]

ERSHOV, D., J.-W. LALIBERTÉ, M. MARCOUX, AND S. ORR (2024): "Estimating Complementarity With Large Choice Sets: An Application to Mergers," Working Paper. [Cited on page 4.]

FOSGERAU, M., J. MONARDO, AND A. DE PALMA (2024): "The Inverse Product Differentiation Logit Model," American Economic Journal: Microeconomics, 16(4), 329–370. [Cited on page 4.]

FUCHS, W., S. FUKUDA, AND D. NEUHANN (2025): "Demand-System Asset Pricing: Theoretical Foundations," Working Paper. [Cited on pages 3, 5, 22, 30, and 45.]

HADDAD, V., Z. HE, P. HUEBNER, P. KONDOR, AND E. LOUALICHE (2025): "Causal Inference for Asset Pricing," Working Paper. [Cited on pages 5 and 24.]

HARRIS, L., AND E. GUREL (1986): "Price and Volume Effects Associated with Changes in the S&P 500 List: New Evidence for the Existence of Price Pressures," Journal of Finance, 41(4), 815–829. [Cited on page 4.]

HE, Z., P. KONDOR, AND J. S. LI (2025): "Demand Elasticity in Dynamic Asset Pricing," Working Paper. [Cited on page 5.]

HURWICZ, L. (1962): "On the Structural Form of Interdependent Systems," in Logic, Methodology and Philosophy of Science: Proceedings of the 1960 International Congress, ed. by E. Nagel, P. Suppes, and A. Tarski, pp. 232–239. Stanford University Press. [Cited on page 16.]

IARIA, A., AND A. WANG (2020): "Identification and Estimation of Demand for Bundles," Working Paper. [Cited on page 4.]

KOIJEN, R. S. J., AND M. YOGO (2019): "A Demand System Approach to Asset Pricing," Journal of Political Economy, 127(4), 1475–1515. [Cited on pages 3, 5, 9, and 30.]

KREPS, D. M., AND E. L. PORTEUS (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," Econometrica, 46(1), 185–200. [Cited on page 20.]

MARSCHAK, J. (1953): "Economic Measurements for Policy and Prediction," in Studies in Econometric Method, ed. by W. C. Hood, and T. C. Koopmans, no. 14 in Cowles Commission Monograph, pp. 1–26. John Wiley & Sons. [Cited on page 16.]

PLEMMONS, R. J., AND R. E. CLINE (1972): "The Generalized Inverse of a Nonnegative Matrix," Proceedings of the American Mathematical Society, 31(1), 46–50. [Cited on page 35.]

SHLEIFER, A. (1986): "Do Demand Curves for Stocks Slope Down?," Journal of Finance, 41(3), 579–590. [Cited on page 4.]

TOBIN, J. (1969): "A General Equilibrium Approach to Monetary Theory," Journal of Money, Credit and Banking, 1(1), 15–29. [Cited on page 4.]

WANG, A. (2024): "Demand with Complementarity: A Market-Share Approach," Working Paper. [Cited on page 4.]

# A Proofs of Propositions

## A.1 Section 3

**Proof of Proposition 1.** First, the consumption process induced by portfolio $a^i$ and payoff matrix $Y$ is $c^i = Y^{\mathrm{T}} a^i + w^i$. Multiplying $Y$ from the left, $Yc^i = YY^{\mathrm{T}} a^i + Yw^i$. Since $YY^{\mathrm{T}}$ is a $J \times J$ invertible matrix by Assumption 1, we have $a^i = (Y^+)^{\mathrm{T}} c^i - (Y^+)^{\mathrm{T}} w^i$. Since $w^i$ does not depend on $p$, differentiating this expression yields $\frac{\partial a^i}{\partial p^{\mathrm{T}}} = (Y^+)^{\mathrm{T}} \frac{\partial c^i}{\partial p^{\mathrm{T}}}$. By no arbitrage, $p = Yq$ and hence $q = Y^+ p$. Then, equation (3) follows from the chain rule.

Second, under the stated conditions, the first-order condition with respect to $c^i_z$ is:

$$\delta^i u'(c^i_z) = \lambda \pi_z^{-1} q_z,$$

where $\lambda$ is the Lagrange multiplier on the budget constraint, which depends on the state price vector, as $\lambda = (1 - \delta^i) u'(c_0)$ where $c_0$ is optimal consumption at time 0. Differentiating this first-order condition with respect to $q_{z'}$ yields:

$$\delta^i u''(c^i_z) \frac{\partial c^i_z}{\partial q_{z'}} = \pi_z^{-1} \left( \frac{\partial \lambda}{\partial q_{z'}} q_z + \lambda \mathbf{1}(z = z') \right),$$

that is,

$$\frac{\partial c^i_z}{\partial q_{z'}} = \pi_z^{-1} \frac{1}{\delta^i u''(c^i_z)} \left( \frac{\partial \lambda}{\partial q_{z'}} q_z + \lambda \mathbf{1}(z = z') \right).$$

Letting $\tilde{\mathcal{D}}^i_{z,z'} \equiv \frac{1}{\delta^i u''(c^i_z)} \left( \frac{\partial \lambda}{\partial q_{z'}} q_z + \lambda \mathbf{1}(z = z') \right)$ and $\Pi \equiv \mathrm{diag}(\pi_1, \ldots, \pi_Z)$, we have:

$$\mathcal{D}^i = \Pi^{-1} \tilde{\mathcal{D}}^i.$$

Substituting this equation into (3) yields equation (4). ■

**Proof of Proposition 2.** Let $Y \in \mathbb{R}^{J \times Z}$ be a payoff matrix, and let $(z_1, \ldots, z_T)$ be any sample of realized states. Let $y \in \mathbb{R}^J$ be any vector with $y \notin \{y(1), \ldots, y(Z)\}$, where $y(z)$ is the $z$-th column of $Y$. Define $\tilde{Y} = [Y \mid y] \in \mathbb{R}^{J \times (Z+1)}$, where $y$ corresponds to a state that does not realize in the sample.

With these in mind, first, we show that $\tilde{Y}$ is observationally equivalent to

$Y$: the realized return on every asset $j$ in every period $t$ is identical under $Y$ and $\tilde{Y}$. To see this, since the additional state corresponding to $y$ does not realize in the sample, the realized return on asset $j$ in period $t$ is $(Y)_{j,z_t}$ under both $Y$ and $\tilde{Y}$ for all $j \in \{1, \ldots, J\}$ and $t \in \{1, \ldots, T\}$.

Second, the Moore-Pensrose inverses of the two matrices differ: $\tilde{Y}^+ \neq Y^+$. On the one hand, the Moore-Penrose inverse $Y^+ \in \mathbb{R}^{Z \times J}$ is given by

$$Y^+ = Y^{\mathrm{T}}(YY^{\mathrm{T}})^{-1}.$$

On the other hand, the Moore-Penrose inverse $\tilde{Y}^+ \in \mathbb{R}^{(Z+1) \times J}$ is given by

$$\tilde{Y}^+ = \begin{bmatrix} Y^{\mathrm{T}}(YY^{\mathrm{T}} + yy^{\mathrm{T}})^{-1} \\ y^{\mathrm{T}}(YY^{\mathrm{T}} + yy^{\mathrm{T}})^{-1} \end{bmatrix}.$$

Thus, in addition to $Y^+ \neq \tilde{Y}^+$, the $Z \times J$ block of $\tilde{Y}^+$ is different from $Y^+$ as $(YY^{\mathrm{T}} + yy^{\mathrm{T}})^{-1} \neq (YY^{\mathrm{T}})^{-1}$. ■

**Proof of Proposition 3.** First, recall that $\mathcal{D}^i$ is the matrix of partial derivatives of optimal Marshallian consumption demand with respect to state prices, evaluated at given preferences and endowments. It is determined solely by the investor's preference ordering and budget set, and does not depend on $Y$ directly. Hence, for any intervention that holds preferences and endowments fixed, $\mathcal{D}^i$ is unchanged.

In contrast, by Proposition 1, $\mathcal{A}^i = (Y^+)^{\mathrm{T}}\mathcal{D}^i Y^+$. Let $\tilde{Y}$ be a perturbation of $Y$, i.e., $\tilde{Y}$ is a $J \times \tilde{Z}$ non-negative matrix with $\operatorname{rank}(Y) = J \leq \tilde{Z}$. Let $\tilde{\mathcal{A}}^i = (\tilde{Y}^+)^{\mathrm{T}}\mathcal{D}^i\tilde{Y}^+$. Then, $\tilde{\mathcal{A}}^i \neq \mathcal{A}^i$ with probability 1 (also, the set of $\tilde{Y}$ with $\tilde{\mathcal{A}}^i \neq \mathcal{A}^i$ is open and dense).

As Proposition 2 establishes that $Y$ is not identified from realized return data, $\mathcal{A}^i$ cannot be identified as a structural object, and that $\mathcal{A}^i$ alone does not suffice to predict asset demand responses to interventions without additional identifying assumptions on $Y^+$. This proves the first part, and the second part is a contraposition of the first part. ■

## A.2 Section 4.1

**Proof of Theorem 1.** First, we show that Condition 1 implies that $YY^{\mathrm{T}}$ is diagonal. Suppose $Y^{+} = -VY^{\mathrm{T}}K$ for some diagonal matrix $K \equiv \mathrm{diag}(k_1, \ldots, k_J)$. Operating $Y$ on both sides from the left,

$$I_J = -YVY^{\mathrm{T}}K.$$

If $k_j = 0$ for some $j$, then the $j$-th column of $K$ is the zero vector, and so is the $j$-th column of the right-hand side, which is impossible. Thus, $k_j \neq 0$ for all $j$. Then, $YVY^{\mathrm{T}}$ is a diagonal matrix:

$$\begin{cases} \sum_{z=1}^{Z} y_j(z) v_z y_{j'}(z) \neq 0 & \text{if } j = j' \\ \sum_{z=1}^{Z} y_j(z) v_z y_{j'}(z) = 0 & \text{if } j \neq j' \end{cases}.$$

Since $y_j(z), y_{j'}(z) \geq 0$, and $v_z > 0$, it follows that

$$\begin{cases} \sum_{z=1}^{Z} y_j(z) y_{j'}(z) \neq 0 & \text{if } j = j' \\ \sum_{z=1}^{Z} y_j(z) y_{j'}(z) = 0 & \text{if } j \neq j' \end{cases}.$$

Hence, $YY^{\mathrm{T}}$ is diagonal.

Second, we show that, more generally, Condition 2 implies that $YY^{\mathrm{T}}$ is diagonal. By Condition 2, the Moore-Penrose pseudo-inverse $Y^{+} = Y^{\mathrm{T}}(YY^{\mathrm{T}})^{-1}$ is non-negative. By Plemmons and Cline (1972, Theorem 1), the pseudo-inverse $Y^{+}$ is non-negative if and only if there exists a diagonal matrix with positive elements $D \equiv \mathrm{diag}(d_1, \ldots, d_Z)$ such that

$$Y^{+} = DY^{\mathrm{T}}. \tag{6}$$

Then, operating $Y$ from the left,

$$I_J = YDY^{\mathrm{T}}.$$

Then, extracting the $(j, k)$ element (with $j \neq k$) from each of both sides,

$$0 = \sum_{z=1}^{Z} y_j(z) d_z y_k(z).$$

Since $y_j(z) \geq 0$, $d_z > 0$, and $y_k(z) \geq 0$ for all $z \in \{1, \ldots, Z\}$, it follows that

$$y_j(z) y_k(z) = 0 \text{ for all } z \in \{1, \ldots, Z\}.$$

This implies that the $(j, k)$ element (with $j \neq k$) of $YY^{\mathrm{T}}$ is 0:

$$0 = \sum_{z=1}^{Z} y_j(z) y_k(z). \tag{7}$$

Thus, $YY^{\mathrm{T}}$ is a diagonal matrix.

Third, we show that, given that $YY^{\mathrm{T}}$ is diagonal, there are no assets with overlapping payoffs. Since $YY^{\mathrm{T}}$ is invertible, it is a diagonal matrix with positive elements. Equation (7) implies that, for any $z \in \{1, \ldots, Z\}$, there exists at most one $j \in \{1, \ldots, J\}$ such that $y_j(z) > 0$.

Fourth, we show that if markets are complete then $YY^{\mathrm{T}}$ is diagonal if and only if $Y$ has exactly one non-zero element in each row and in each column (so that $Y$ is a diagonal matrix up a re-ordering of rows or columns). If $YY^{\mathrm{T}}$ is diagonal, then its $(j, k)$ element is:

$$\begin{cases} \sum_{z=1}^{Z} y_j(z) y_j(z) > 0 & \text{if } j = k \\ \sum_{z=1}^{Z} y_j(z) y_k(z) = 0 & \text{if } j \neq k \end{cases}.$$

Hence, for each row $j$, there exists exactly one element $z$ such that $y_j(z) > 0$. Thus, $Y$ has $J$ non-zero elements. Since $Y$ is square and invertible, for each column $z$, there exists exactly one element $j$ such that $y_j(z) > 0$.

Conversely, if $Y$ has exactly one non-zero element in each row and in each column, then

$$\begin{cases} \sum_{z=1}^{Z} y_j(z) y_j(z) > 0 & \text{if } j = k \\ \sum_{z=1}^{Z} y_j(z) y_k(z) = 0 & \text{if } j \neq k \end{cases}.$$

Thus, $YY^{\mathrm{T}}$ is diagonal. ■

**Remark 1 (Proof of Theorem 1)** *Two remarks on the proof of Theorem 1 are in order. First, if $YY^{\mathrm{T}}$ is diagonal, then since $YY^{\mathrm{T}}$ is invertible under Assumption 1, $(YY^{\mathrm{T}})^{-1}$ is* *a diagonal matrix with positive entries. Since $Y$ is non-negative, so is $Y^{\mathrm{T}}$. Then, $Y^{+} = Y^{\mathrm{T}}(YY^{\mathrm{T}})^{-1}$ is non-negative.*

*Second, when each column of $Y$ is not a zero vector, i.e., for each $z \in \{1, \ldots, Z\}$, there exists at least one $j \in \{1, \ldots, J\}$ such that $Y_{j,z} = y_j(z) > 0$, it can be shown that the diagonal matrix $D$ in expression (6) is unique.*

**Proof of Proposition 4.** Let $y_j$ denote the $j$-th row of $Y$. Let $y_k^{+}$ denote the $k$-th column of $Y^{+}$. It follows from $YY^{+} = I_J$ that:

$$\sum_{z=1}^{Z} y_k(z) Y^{+}_{z,k} = 1 \quad \text{for all} \quad k \in \{1, \ldots, J\}; \tag{8}$$

$$\sum_{z=1}^{Z} y_j(z) Y^{+}_{z,k} = 0 \quad \text{if} \quad j \neq k. \tag{9}$$

Suppose to the contrary that there exists a column $k$ in $Y^{+}$ such that $y_k^{+} \geq 0$ element-by-element.

Consider the orthogonality condition (9) for some $j \neq k$. Since $Y$ is non-negative, $y_k \geq 0$. We assumed $y_k^{+} = (Y^{+}_{z,k})_z \geq 0$. Thus, if $y_j(z) > 0$ then $Y^{+}_{z,k} = 0$. This must hold for all $j \neq k$. Therefore, $y_k^{+}$ must be zero at any index $z$ where any other row of $Y$ is positive.

Now consider the normalization condition (8). For the sum to be strictly positive, there must exist at least one index $z^*$ such that:

$$y_k(z^*) > 0 \quad \text{and} \quad Y^{+}_{z^*,k} > 0. \tag{10}$$

However, we know that $Y^{+}_{z^*,k} > 0$ is only possible if $y_j(z^*) = 0$ for all $j \neq k$. Combining this with expression (10), we see that index $z^*$ represents a column in $Y$ where: the entry in row $k$ is positive: $y_k(z^*) > 0$; and the entries in all other rows $i$ are zero: $y_i(z^*) = 0$ for $i \neq k$. This implies that column $z^*$ of matrix $Y$ has exactly one strictly positive element, which is a contradiction to the assumption of the statement. ■

## A.3 Section 4.2

**Proof of Proposition 5.** Suppose for contradiction that $\mathcal{A}^i$ is identified but $Y^+$ is not assumed fixed. Then, there exist two observationally equivalent parameter vectors $(\Theta^i, Y, \pi)$ and $(\tilde{\Theta}^i, \tilde{Y}, \tilde{\pi})$ with $Y^+ \neq \tilde{Y}^+$. By Proposition 2, such pairs exist for any finite dataset. By the demand decomposition of Proposition 1,

$$\mathcal{A}^i(\Theta^i, Y, \pi) = (Y^+)^T \mathcal{D}^i Y^+ \neq (\tilde{Y}^+)^T \mathcal{D}^i \tilde{Y}^+ = \mathcal{A}^i(\tilde{\Theta}^i, \tilde{Y}, \tilde{\pi})$$

generically, contradicting observational equivalence. Hence $\mathcal{M}$ must include the assumption that $Y^+$ is fixed. The requirement of $K = J$ independent experiments then follows from the fact that $\mathcal{A}^i$ is a $J \times J$ matrix, and point-identification of all its elements requires $J$ linearly independent price vectors in $O_P$. ■

## A.4 Section 4.3: Proof of Theorem 2

The proof proceeds in four steps. The first step establishes the asymptotic limit (i.e., the population covariance matrix) $\Sigma$ of the Gram matrix $\frac{1}{Z}Y^{\mathrm{T}}Y$ as $Z \to \infty$. This allows the pseudo-inverse $Y^+ = Y^{\mathrm{T}}(YY^{\mathrm{T}})^{-1}$ to be approximated by $\frac{1}{Z}Y^{\mathrm{T}}\Sigma^{-1}$. Then, the population limits $\pi(j,z)$ and $q(J)$ equal expressions involving the population inverse $\Sigma^{-1}$, reducing both parts of the theorem to statements about $(\Sigma^{-1}y_z)_j$. The second step shows that each column of $Y^{\mathrm{T}}\Sigma^{-1}$ can be decomposed into the deterministic shift (i.e., $(\mu_j)_j$ in the main text) and the stochastic component centered around 0 (i.e., $(\nu_{z,j})_{z,j}$ in the main text). Then, we show the sense in which the deterministic shift is small compared to the stochastic component $(\nu_{z,j})_{z,j}$ when $J$ is large by applying the Woodbury identity to $\Sigma$. With these in mind, the third step establishes part(i). The proof shows that the sign of the $(z,j)$ element of $Y^+$, which is approximated by that of $\Sigma^{-1}Y^T$ (up to scaling) is asymptotically a fair coin flip. Finally, the fourth step establishes part (ii). The proof shows that the signs for two economies with a share factor structure but independent realizations of idiosyncratic shocks are asymptotically independent.

Hereafter, fix a realization of $(\alpha, \beta)$ for which all laws of large numbers used below hold. Probabilities are conditional on the loadings unless noted otherwise.

**Step 1.** In the first step, we replace the sample Gram matrix $G_Z \equiv \frac{1}{Z}YY^{\mathrm{T}}$ with the population covariance matrix $\Sigma$ by the law of large numbers. Namely, as $Z \to \infty$ with $J$ fixed, the sample covariance matrix $G_Z$ converges almost surely to the population second moment matrix $\Sigma$, where, as in the main text,

$$\Sigma = \gamma\gamma^{\mathrm{T}} + \sigma_f^2 \beta\beta^{\mathrm{T}} + \sigma_\varepsilon^2 I_J = \sigma_\varepsilon^2 I_J + UU^{\mathrm{T}} \quad \text{with} \quad U \equiv \begin{bmatrix} \gamma & \sigma_f \beta \end{bmatrix} \in \mathbb{R}^{J\times 2}$$

is a rank-two perturbation of a scaled identity matrix. Note that $\Sigma$ is positive definite so that it is invertible. This allows the pseudo-inverse $Y^+$ to be approximated by $\frac{1}{Z}Y^{\mathrm{T}}\Sigma^{-1} = \frac{1}{Z}\Sigma^{-1}Y^{\mathrm{T}}$. Lemma 2 below formally shows that the sign of $(Y^+)_{z,j}$ is determined by the sign of the variable $(\Sigma^{-1}y_z)_j$.

To that end, we decompose $(\Sigma^{-1}y_z)_j$ into the deterministic shift and the stochastic part symmetric around 0. Writing $y_z = \gamma + \beta(f_z - \bar{f}) + \varepsilon_z$ as in the main text, one can express

$$(\Sigma^{-1}y_z)_j = \underbrace{(\Sigma^{-1}\gamma)_j}_{\mu_j} + \underbrace{(f_z - \overline{f})(\Sigma^{-1}\beta)_j + (\Sigma^{-1}\varepsilon_z)_j}_{\nu_{z,j}}. \tag{11}$$

Conditional on the loadings $(\alpha, \beta)$, the term $\mu_j$ is a deterministic shift and the term $\nu_{z,j}$ is symmetric around zero.

Let $F_{\nu,j}$ be the CDF of $nu_{z,j}$ conditional on loadings $(\alpha, \beta)$. Then,

$$\begin{aligned} P((\Sigma^{-1}y_z)_j > 0) &= P(\mu_j + \nu_{z,j} > 0) \\ &= 1 - F_{\nu,j}(-\mu_j) = \frac{1}{2} + f_{\nu,j}(0)\mu_j + O(\mu_j^2), \end{aligned} \tag{12}$$

where the last equality follows from the Taylor approximation of $1 - F_{\nu,j}(\cdot)$ and $F_{\nu,j}(0) = \frac{1}{2}$ (which follows because $\nu_{z,j}$ is symmetric around zero).

With these in mind, we now establish Lemma 2, which guarantees that the replacement of $Y^+ = \frac{1}{Z}Y^{\mathrm{T}}G_Z^{-1}$ with $\Sigma^{-1}Y^{\mathrm{T}}$ does not change the limiting sign frequency: since $G_Z^{-1} \to \Sigma^{-1}$ with $\|G_Z^{-1} - \Sigma^{-1}\| = O(Z^{-1/2})$, the difference between the two matrices vanishes in operator norm, and any potential sign disagreement occurs only when an entry of $(\Sigma^{-1}y_z)_j$ lies in a vanishing neighborhood of zero. Since replacing $G_Z^{-1}$ by $\Sigma^{-1}$ changes each entry by at most $O(Z^{-1/2})$, a sign disagreement can occur only with probability $o(1)$. This ensures that the asymptotic sign frequency is unaffected by the finite-$Z$ approximation. Formally:

**Lemma 2 (Population replacement)** *Fix $J$ and $(\alpha, \beta)$. As $Z \to \infty$, the sample Gram matrix $G_Z \equiv \frac{1}{Z} Y Y^{\mathrm{T}}$ converges almost surely to $\Sigma$ and hence $G_Z^{-1} \to \Sigma^{-1}$ a.s. Moreover,*

$$\lim_{Z\to\infty} \max_{1\leq j\leq J} \left| \frac{1}{Z} \sum_{z=1}^{Z} \mathbf{1}\left( \left(\tfrac{1}{Z} y_z^{\mathrm{T}} G_Z^{-1}\right)_j > 0 \right) - P\left( (\Sigma^{-1} y_z)_j > 0 \right) \right| = 0 \quad a.s. \tag{13}$$

*Consequently, conditional on $(\alpha, \beta)$,*

$$\pi(j, z) = P\left( (\Sigma^{-1} y_z)_j > 0 \right), \tag{14}$$

$$q(J) = \frac{1}{J} \sum_{j=1}^{J} P\left( \mathrm{sign}\big((\Sigma^{-1} y_z)_j\big) = \mathrm{sign}\big((\Sigma^{-1} \tilde{y}_z)_j\big) \right), \tag{15}$$

*where $y_z$ and $\tilde{y}_z$ denote the z-th columns of $Y$ and $\tilde{Y}$.*

**Proof of Lemma 2.** By the law of large numbers, $G_Z \to \Sigma$ a.s. Hence, $G_Z$ is positive definite for large $Z$ and $\|G_Z^{-1} - \Sigma^{-1}\| \to 0$ a.s. Let

$$D_{z,j} \equiv \left( \frac{1}{Z} y_z^{\mathrm{T}} G_Z^{-1} \right)_j - \left( \frac{1}{Z} \Sigma^{-1} y_z \right)_j = \frac{1}{Z} y_z^{\mathrm{T}} (G_Z^{-1} - \Sigma^{-1}) v_j,$$

where $v_j$ is the unit vector in the $j$-th coordinate.

For any $\eta > 0$ and all large $Z$, $\|G_Z^{-1} - \Sigma^{-1}\|_F \leq \eta$ a.s., so $|D_{z,j}| \leq \frac{\eta}{Z} \|y_z\|$. A sign can flip only if $|(\frac{1}{Z} \Sigma^{-1} y_z^{\mathrm{T}})_j| \leq |D_{z,j}|$. Since $(\Sigma^{-1} y_z)_j = \mu_j + \nu_{z,j}$ has a continuous density at around 0 with value $f_{\nu,j}(0)$, we have:

$$P\left( \left|(\Sigma^{-1} y_z)_j\right| \leq \delta \right) \;\leq\; 2 f_{\nu,j}(0)\delta + o(\delta) \quad (\delta \downarrow 0).$$

Since the inequality holds uniformly across $j \in \{1, \ldots, J\}$, the sign disagreement probability vanishes uniformly across the entire cross-section $j \in \{1, \ldots, J\}$. Taking $\delta = \frac{\eta}{Z}\|y_z\|$ and averaging over $z$ (using $Z^{-1}\sum_z \|y_z\| \to \mathbb{E}\|y_z\|$ a.s.) shows the empirical fraction of sign disagreements is $O(\eta)$ a.s. Letting $\eta \downarrow 0$ proves (13). Then, the representations (14)–(15) follow immediately. ■

**Step 2.** The second step examines the structure of $\Sigma^{-1}$. The key insight is that $\Sigma = \sigma_\varepsilon^2 I_J + UU^\mathrm{T}$ is a rank-two perturbation of a scaled identity. Applying the Woodbury identity yields $\Sigma^{-1} = \sigma_\varepsilon^{-2} I_J - \sigma_\varepsilon^{-4} U A_J^{-1} U^\mathrm{T}$, where $A_J \equiv I_2 + \sigma_\varepsilon^{-2} U^\mathrm{T} U$ is a $2 \times 2$ matrix with $\|A_J^{-1}\| = O(J^{-1})$, because $U^\mathrm{T}U$ grows like $J$ as more assets are added. The correction term $UA_J^{-1}U^\mathrm{T}$ is therefore small relative to the leading $\sigma_\varepsilon^{-2} I_J$ term. This has the following consequences established in the lemma:

**Lemma 3 (Small deterministic shift)** *Under Assumptions 3–5, the following hold uniformly in $j$ as $J \to \infty$:*

$$(\Sigma^{-1}\gamma)_j = O(J^{-1}), \qquad (\Sigma^{-1}\beta)_j = O(J^{-1}), \qquad (\Sigma^{-2})_{j,j} \to \sigma_\varepsilon^{-4}.$$

*Consequently, the stochastic fluctuation $\nu_{z,j} \equiv (f_z - \bar{f})(\Sigma^{-1}\beta)_j + (\Sigma^{-1}\varepsilon_z)_j$ is symmetric around zero with a continuous density and variance satisfying $\sigma_{\nu,j}^2 \to \sigma_\varepsilon^{-2}$ uniformly in $j$.*

**Proof.** We apply the Woodbury identity to $\Sigma = \sigma_\varepsilon^2 I_J + UU^\mathrm{T}$:

$$\Sigma^{-1} = \sigma_\varepsilon^{-2}\left(I_J - U\left(I_2 + \sigma_\varepsilon^{-2}U^\mathrm{T}U\right)^{-1}\sigma_\varepsilon^{-2}U^\mathrm{T}\right). \tag{16}$$

Since the $2 \times 2$ matrix $U^\mathrm{T}U$ satisfies

$$U^\mathrm{T}U = J\begin{bmatrix} \mathbb{E}_J[\gamma^2] & \sigma_f \mathbb{E}_J[\gamma\beta] \\ \sigma_f \mathbb{E}_J[\gamma\beta] & \sigma_f^2 \mathbb{E}_J[\beta^2] \end{bmatrix},$$

where $\mathbb{E}_J$ denotes the empirical mean over $j \in \{1, \ldots, J\}$, each entry is of order $O(J)$. Thus, $\sigma_\varepsilon^{-2}U^\mathrm{T}U = O(J)$, which implies that the dominant term in the matrix $I_2 + \sigma_\varepsilon^{-2}U^\mathrm{T}U$ is the $O(J)$ contribution from $\sigma_\varepsilon^{-2}U^\mathrm{T}U$. Thus, when $J$ is large, the $2 \times 2$ matrix $\left(I_2 + \sigma_\varepsilon^{-2}U^\mathrm{T}U\right)^{-1}$ is of order $O(J^{-1}) = O(1) \cdot O(J^{-1}) \cdot O(1)$. Consequently, each entry in the matrix $U\left(I_2 + \sigma_\varepsilon^{-2}U^\mathrm{T}U\right)^{-1}\sigma_\varepsilon^{-2}U^\mathrm{T}$ is of order $O(J^{-1})$, meaning that $\Sigma^{-1}$ is asymptotically diagonal with off-diagonal entries that vanish at the same rate. Economically, the pseudo-inverse suppresses variation along the factor directions while leaving idiosyncratic risk largely unaffected.

Since we have established $(I_2 + \sigma_\varepsilon^{-2} U^{\mathrm{T}} U)^{-1} = O(J^{-1})$, substituting this back into Woodbury identity (16) and noting that $\gamma = U v_1$ yields

$$\Sigma^{-1}\gamma = \sigma_\varepsilon^{-2}\gamma - \sigma_\varepsilon^{-2} U \big(I_2 + \sigma_\varepsilon^{-2} U^{\mathrm{T}} U\big)^{-1} \sigma_\varepsilon^{-2} U^{\mathrm{T}} \gamma.$$

The first term $\sigma_\varepsilon^{-2}\gamma$ is $O(1)$ in each component. However, since $\gamma$ lies in the column space of $U$, we have $U^{\mathrm{T}}\gamma = U^{\mathrm{T}} U v_1$, which is $O(J)$. Thus, the second term equals

$$\sigma_\varepsilon^{-4} U \big(I_2 + \sigma_\varepsilon^{-2} U^{\mathrm{T}} U\big)^{-1} U^{\mathrm{T}} U v_1 = \sigma_\varepsilon^{-2} U \cdot O(J^{-1}) \cdot O(J) = \sigma_\varepsilon^{-2} U \cdot O(1).$$

Each component of this correction term is $O(1)$, and it precisely cancels the leading $O(1)$ term $\sigma_\varepsilon^{-2}\gamma$. What remains is a residual of order $O(J^{-1})$: each component $\mu_j = (\Sigma^{-1}\gamma)_j$ satisfies $|\mu_j| = O(J^{-1})$ uniformly in $j$. The same reasoning applies to $\beta$, giving $(\Sigma^{-1}\beta)_j = O(J^{-1})$.

Next, squaring expression (16) gives

$$\Sigma^{-2} = \sigma_\varepsilon^{-4}\Big(I_J - 2UA_J^{-1}\sigma_\varepsilon^{-2}U^{\mathrm{T}} + UA_J^{-1}\sigma_\varepsilon^{-4}(U^{\mathrm{T}}U)A_J^{-1}U^{\mathrm{T}}\Big), \quad \text{with} \quad A_J \equiv I_2 + \sigma_\varepsilon^{-2} U^{\mathrm{T}} U.$$

Since $A_J^{-1} = O(J^{-1})$ and $U^{\mathrm{T}}U = O(J)$, the corrections are $O(J^{-1})$. Thus,

$$(\Sigma^{-2})_{j,j} = \sigma_\varepsilon^{-4}\{1 + O(J^{-1})\} \to \sigma_\varepsilon^{-4} \quad \text{uniformly in} \quad j.$$

Substituting these orders into

$$\sigma_{\nu,j}^2 = \sigma_f^2 \big[(\Sigma^{-1}\beta)_j\big]^2 + \sigma_\varepsilon^2 (\Sigma^{-2})_{j,j}$$

yields $\sigma_{\nu,j}^2 = \sigma_\varepsilon^{-2} + O(J^{-1})$ uniformly in $j$. ■

Lemma 3 formalizes the intuition that as the cross-section expands, the factor-induced corrections to $\Sigma^{-1}$ become negligible: the pseudo-inverse behaves almost like a scaled identity, and the density at zero governing the linearization of the sign probability is determined primarily by the idiosyncratic variance $\sigma_\varepsilon^2$.

**Step 3: Individual coin flip, proof of part (i).** By Lemma 2, it suffices to analyze $P\left((\Sigma^{-1}y_z)_j > 0\right)$. By (12), we have:

$$\pi(j,z) = \frac{1}{2} + f_{\nu,j}(0)\mu_j + O(\mu_j^2). \tag{17}$$

By Lemma 3, $\mu_j \to 0$ uniformly as $J \to \infty$. Lemma 3 also implies that the density $f_{\nu,j}(0)$ bounded uniformly in $j$ as $J \to \infty$. Thus, $\lim_{J\to\infty} \pi(j,z) = \frac{1}{2}$ for each fixed $(j,z)$.[7] This establishes part (i).

**Step 4: Sign instability, part (ii).** By Lemma 2, $q(J)$ equals the average over $j$ of

$$P\left(\text{sign}((\Sigma^{-1}y_z)_j) = \text{sign}((\Sigma^{-1}\tilde{y}_z)_j)\right).$$

Since $Y$ and $\tilde{Y}$ share the same factor structure, both Gram matrices converge to the same $\Sigma$. Applying decomposition (11) to each matrix at a fixed $(z,j)$ gives

$$\begin{aligned}(\Sigma^{-1}y_z)_j &= \mu_j + (f_z - \bar{f})(\Sigma^{-1}\beta)_j + (\Sigma^{-1}\varepsilon_z)_j;\\ (\Sigma^{-1}\tilde{y}_z)_j &= \mu_j + (f_z - \bar{f})(\Sigma^{-1}\beta)_j + (\Sigma^{-1}\tilde{\varepsilon}_z)_j.\end{aligned}$$

Both variables share the deterministic shift $\mu_j$ and the common factor component $(f_z - \bar{f})(\Sigma^{-1}\beta)_j$. By Lemma 3, $(\Sigma^{-1}\beta)_j = O(J^{-1})$, and thus the common factor component vanishes as $J \to \infty$. The residual stochastic terms $(\Sigma^{-1}\varepsilon_z)_j$ and $(\Sigma^{-1}\tilde{\varepsilon}_z)_j$ are independent (since $\varepsilon \perp \tilde{\varepsilon}$) and each symmetric around zero. In the limit, the two sign-determining variables reduce to independent symmetric random variables $\xi$ and $\tilde{\xi}$. Hence,

$$\begin{aligned}P\left(\text{sign}((\Sigma^{-1}y_z)_j) = \text{sign}((\Sigma^{-1}\tilde{y}_z)_j)\right) &\to P(\xi > 0, \tilde{\xi} > 0) + P(\xi < 0, \tilde{\xi} < 0)\\ &= \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.\end{aligned}$$

[7] An earlier working paper shows that $\frac{1}{J}\sum_{j=1}^{J} f_{\nu,j}(0)\mu_j \to \frac{1}{J} f_\nu(0)\Theta_1$, where the constant $\Theta_1$ is determined by the model primitives.

Since this holds for each fixed $j$ and the indicators are bounded, the dominated convergence theorem gives $\lim_{J\to\infty} q(J) = \frac{1}{2}$. ■

To conclude the proof of Theorem 2, two remarks are in order. Remark 2 shows that a multi-factor extension is possible. Remark 3 reiterates the economic interpretation of Theorem 2.

**Remark 2 (Multi-factor extension)** *For a K-factor model with*

$$y_{j,z} = \gamma_j + \sum_{k=1}^{K} \beta_j^{(k)} (f_z^{(k)} - \bar{f}^{(k)}) + \varepsilon_{j,z},$$

*the population Gram matrix takes the form*

$$\Sigma = \sigma_\varepsilon^2 I_J + UU^{\mathrm{T}},$$

*where*

$$U \equiv \begin{bmatrix} \gamma & \sigma_{f,1}\beta^{(1)} & \ldots & \sigma_{f,K}\beta^{(K)} \end{bmatrix} \in \mathbb{R}^{J\times(K+1)}.$$

*Since $(I_{K+1} + \sigma_\varepsilon^{-2} U^{\mathrm{T}} U)^{-1} = O(J^{-1})$ for any fixed K, Lemma 3 carries over verbatim: both $(\Sigma^{-1}\gamma)_j$ and each $(\Sigma^{-1}\beta^{(k)})_j$ are $O(J^{-1})$. Hence both parts of Theorem 2 hold unchanged for any finite number of factors.*

**Remark 3 (Economic interpretation)** *Theorem 2 sharpens the identification impossibility established in Theorem 1 of the paper. Theorem 1 shows that supply shocks generically fail to generate the correct direction of state-price changes for at least one asset. Part (i) here quantifies how severe this directional failure is for each state-asset pair individually: the sign of the required correction is a coin flip in large economies. Part (ii) shows that this instability cannot be resolved by obtaining a second sample from an economy with the same factor structure: even sharing the same systematic risk, two economies require corrections of opposite sign approximately half the time. This rules out any procedure for controlling directional errors that relies solely on the factor structure of payoffs.*

# B Illustration in a General Equilibrium Model

We illustrate our findings using a simple example economy based on Fuchs, Fukuda, and Neuhann (2025). The decision problem is as in (PCP). For tractability, we assume that all investors are symmetric, face no portfolio constraints and have log utility.

**Payoffs.** There are two assets and two aggregate states of the world, both denoted by $g$ (green) and $r$ (red). The probability of state $z \in \{g, r\}$ is $\pi_z \in (0, 1)$. Table 1 depicts the payoff matrix. Parameter $\epsilon \in (0, 1)$ determines the complementarity between green and red assets. As $\epsilon \to 0$, green and red assets are perfect substitutes. As $\epsilon \to 1$, the green and red assets are Arrow securities paying exactly one unit in one state of the world.

| | State $g$ ($\pi_g$) | State $r$ ($\pi_r$) |
|---|---|---|
| Asset $g$ | $\frac{1}{2}(1+\epsilon)$ | $\frac{1}{2}(1-\epsilon)$ |
| Asset $r$ | $\frac{1}{2}(1-\epsilon)$ | $\frac{1}{2}(1+\epsilon)$ |

Table 1: Payoff matrix.

The aggregate endowments are given by$(e_0, e_g, e_r) = (1, 1 + s_g, 1)$, where $s_g$ is a supply shock to the green asset which we use to create price variation.

**Asset demand.** We will be interested in analyzing asset demand functions in a neighborhood around $s_g = 0$. We derive the demand functions $a_g$ and $a_r$ directly from the representative agent's portfolio choice problem:

$$\max_{a_g, a_r} (1-\delta)u(E_0 - p_g(a_g - E_g) - p_r(a_r - E_r)) + \delta\pi_g u(y_g(g)a_g + y_r(g)a_r) + \delta\pi_r u(y_g(r)a_g + y_r(r)a_r).$$

After substituting payoff matrix $Y$, the first-order conditions are:

$$(1-\delta)\frac{p_g}{E_0 - p_g(a_g - E_g) - p_r(a_r - E_r)} = \delta\pi_g \frac{1+\epsilon}{(1+\epsilon)a_g + (1-\epsilon)a_r} + \delta\pi_r \frac{1-\epsilon}{(1-\epsilon)a_g + (1+\epsilon)a_r}; \tag{18}$$

$$(1-\delta)\frac{p_r}{E_0 - p_g(a_g - E_g) - p_r(a_r - E_r)} = \delta\pi_g \frac{1-\epsilon}{(1+\epsilon)a_g + (1-\epsilon)a_r} + \delta\pi_r \frac{1+\epsilon}{(1-\epsilon)a_g + (1+\epsilon)a_r}. \tag{19}$$

Then, since $\pi_g = 1 - \pi_r$, the representative agent's demand functions are:

$$a_g(p_g, p_r) = \delta \frac{(E_0 + p_g E_g + p_r E_r)\left((1-\epsilon^2)p_g - ((1+\epsilon)^2 - 4\epsilon\pi_r)p_r\right)}{(p_g - p_r)^2 - (p_g + p_r)^2\epsilon^2};$$

$$a_r(p_g, p_r) = \delta \frac{(E_0 + p_g E_g + p_r E_r)\left((1-\epsilon^2)p_r - ((1-\epsilon)^2 + 4\epsilon\pi_r)p_g\right)}{(p_g - p_r)^2 - (p_g + p_r)^2\epsilon^2}.$$

$$a_g(p_g, p_r) = \delta \frac{(1 + p_g(1+s_g) + p_r)\left((1-\epsilon^2)p_g - ((1+\epsilon)^2 - 4\epsilon\rho)p_r\right)}{(p_g - p_r)^2 - (p_g + p_r)^2\epsilon^2}. \tag{20}$$

Varying only the green assets' price yields the standard own-price elasticity:

$$\mathcal{E}_g^{\text{ideal}} \equiv -\frac{\partial a_g(p_g, p_r)}{\partial p_g}\frac{p_g}{a_g}.$$

**Misalignment between ideal experiment and supply shock.** In the ideal experiment, the investor faces an exogenous increase in the price of the green asset $p_g$ while $p_r$ remains fixed. By Lemma 1, the induced change in state prices is

$$\Delta \mathbf{q}_g^{\text{ideal}} = \frac{\partial}{\partial p(g)}\begin{bmatrix} q(g) \\ q(r) \end{bmatrix} = \frac{1}{2\epsilon}\begin{bmatrix} 1+\epsilon \\ -(1-\epsilon) \end{bmatrix}. \tag{21}$$

A pure shock to $p(g)r$ thus *raises* the state price in state $g$, but *lowers* it in state $r$. This decrease in $q(r)$ is necessary to keep $p(r)$ unchanged.

Now consider how a supply shock $s_g$ affects equilibrium state prices. Given resource constraints $c(z) = y_g(z)(1+s_g) + y_r(z)$, equilibrium state prices are

$$q(g) = \pi_g \frac{\delta}{1-\delta} \cdot \frac{1}{1+\frac{1+\epsilon}{2}s_g} \qquad \text{and} \qquad q(r) = \pi_r \frac{\delta}{1-\delta} \cdot \frac{1}{1+\frac{1-\epsilon}{2}s_g}. \tag{22}$$

Differentiating yields

$$\Delta \mathbf{q}_g^{\text{supply}} = \frac{\partial}{\partial s_g}\begin{bmatrix} q(g) \\ q(r) \end{bmatrix} = -\frac{1-\delta}{2\delta}\begin{bmatrix} (1+\epsilon)\cdot \frac{q(g)^2}{\pi_g} \\ (1-\epsilon)\cdot \frac{q(r)^2}{\pi_r} \end{bmatrix} < 0. \tag{23}$$

In contrast to the ideal experiment, a positive supply shock to the green asset thus decreases *both* state prices whenever $\epsilon < 1$. The simple reason is that the green asset pays off in both states of the world. As such, the supply shock generates a state price change $\Delta q_g$ that is of the *wrong sign* compared to the ideal experiment. The only exception is when both assets are Arrow securities ($\epsilon = 1$).

**Implications for demand elasticties.** The misalignment between supply shock and ideal experiment can sharply bias observed behavior. In the ideal experiment, the investor is more willing to substitute away from green assets because the price of the red asset is unchanged. In the supply shock, substitution is tempered because the red asset is endogenously repriced. The resulting "elasticity" measure $\mathcal{E}_g^{\text{supply}}$ thus has an additional term which accounts the spillover to $p_r$:

$$\mathcal{E}_g^{\text{supply}} \equiv -\frac{\frac{da_g}{ds_g}}{\frac{dp_g}{ds_g}}\frac{p_g}{a_g} = \left(-\frac{\partial a_g}{\partial p_g} - \frac{\partial a_g}{\partial p_r}\frac{\frac{dp_r}{ds_g}}{\frac{dp_g}{ds_g}}\right)\frac{p_g}{a_g}.$$

Substituting for the equilibrium prices, these two measures are equal to:

$$\mathcal{E}_g^{\text{ideal}} = (1+(1-2\pi_r)\epsilon)\frac{(1-\epsilon)^2 + 4\epsilon\pi_r(1-\delta\epsilon) + 4\delta\epsilon^2\pi_r^2}{8\pi_r(1-\pi_r)\epsilon^2};$$
$$\mathcal{E}_g^{\text{supply}} = (1+(1-2\pi_r)\epsilon)\frac{2-\delta(1+(1-2\pi_r)\epsilon)}{(1+\epsilon)^2 - 4\epsilon\pi_r}.$$

We plot both measures in Figure 3. The two differ by order of magnitude for small $\epsilon$. In this range, the two assets are close substitutes. In the ideal experiment

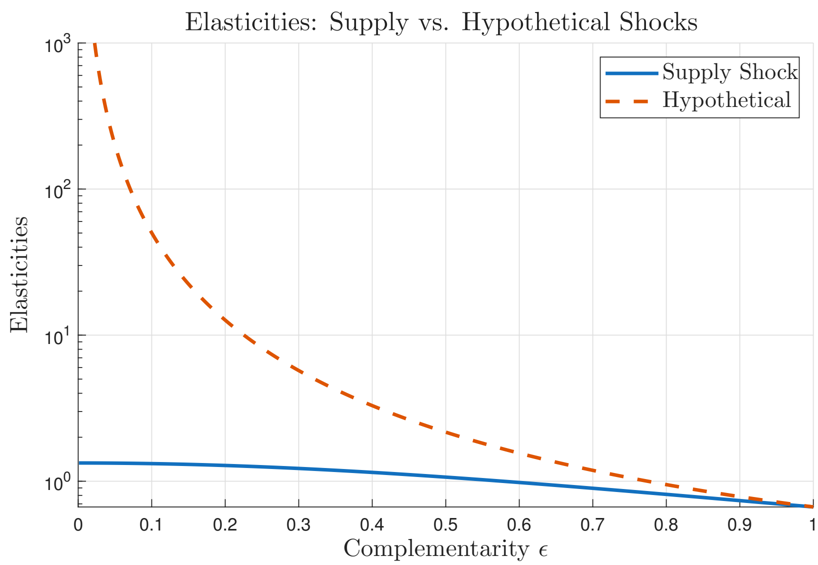


Figure 3: Ideal vs. supply-shock elasticities as a function of $\epsilon$ for $\delta = 2/3$ and $\pi_r = \pi_g = 1/2$. The ideal elasticity (solid line) diverges as $\epsilon \to 0$, while the supply-shock elasticity (dashed line) remains bounded. Both elasticities converge to $1 - \delta\pi_g = 2/3$ at the Arrow security limit $\varepsilon = 1$.

without price spillovers, this leads to very high demand elasticities with respect to a pure price shock. In the case of a supply shock, however, this very substitutability creates strong price spillovers that deter quantity changes on the equilibrium path. Hence, $\mathcal{E}_g^{\text{ideal}}$ diverges to infinity as $\epsilon \to 0$ while $\mathcal{E}_g^{\text{supply}}$ remains small. The only exception is when $\epsilon \to 1$ and the assets approach Arrow securities. In this case, there is no spillover across assets and thus no difference between the ideal experiment and the supply shock.

# C Empirical Illustration

To further gauge the empirical relevance of our results, we conduct a simple empirical exercise in which we use payoff data from the *S&P* 500 to assess the alignment between (subsets of) the payoff matrix and its inverse. The exercise is not intended to be exhaustive, but simply illustrates the immediate relevance of the issues we discuss. The sample consists of 428 stocks that remained in the *S&P* 500 from 2020 to 2024. Since the true payoff matrix is latent, we construct (subsets) of it by sampling realized payoffs.

The payoff for each stock is computed as the end-of-quarter price plus the sum of dividends paid during that quarter. We construct a $20 \times 20$ payoff matrix $Y$ by randomly selecting 20 stocks ($J$). The columns ($Z$) correspond to the 20 quarterly payoff observations from 2020Q1 to 2024Q4. This yields a $20 \times 20$ payoff matrix with weakly positive entries. We then invert this payoff matrix and compute the share of negative entries in $Y^+$ as well as the relative magnitude of the negative and positive entries (in terms of the median and the maximum).

We repeat this exercise ten times with replacement and report averages across all ten repetitions. Table 2 shows that our theoretical predictions hold remarkably well: the share of positive entries of $Y^+$ is approximately one half, and the negative entries are of equal magnitude. Taken together, the barriers to identification we document are generic and pervasive.

| **Metric (averaged over 10 iterations)** | **Value** |
|---|---|
| Percentage of positive entries in $Y^+$ | 50.58% |
| Ratio: (abs negative-entry median) / (positive-entry median) | 1.030 |
| Ratio: (absolute negative minimum) / (positive maximum) | 1.078 |

Table 2: Results of our empirical exercise averaged over 10 iterations.

# D Online Appendix

The Online Appendix is structured as follows. Appendix D.1 presents an example where redundant assets cause discontinuous demand. Appendix D.2 supplements Section 3 (specifically, Proposition 3). Appendix D.3 complements Section 4.1 by providing conditions under which $Y^+$ has the wrong sign for each state (Proposition 6), analogous to the asset-specific conditions in Proposition 4.

## D.1 Section 2.2

We present an example in which an asset demand function exhibits discontinuity in the presence of redundant assets.

**Example 3 (Discontinuous demand functions)** *Suppose there are two states of the world at date 1, and three assets. Given some $\epsilon \in (0,1)$, let a cash flow matrix $Y$ be given by*

$$\begin{bmatrix} \frac{1}{2}(1+\epsilon) & \frac{1}{2}(1-\epsilon) \\ \frac{1}{2}(1-\epsilon) & \frac{1}{2}(1+\epsilon) \\ 1 & 1 \end{bmatrix}.$$

*Now consider the demand functions for some investor $i$ with continuous utility function $u^i$.*

1. *Suppose $\Phi^i = \mathbb{R}^3$. The absence of unbounded arbitrage requires that $p_3 = p_1 + p_2$. Given this restriction on prices, well-defined demand functions exist for all three assets, with the investor taking weakly positive quantities in all three assets. Now suppose that, starting from an initial benchmark where no arbitrage pricing holds, $p_3$ increases slightly. Then, investor $i$'s problem (PCP) is no longer well-defined, and well-defined demand functions no longer exist.*

2. *Suppose instead that investor $i$ faces the short-sale constraint $a^i_j \geq -\chi$ for some $\chi > 0$. Given $p_3 = p_1 + p_2$, well-defined demand functions still exist for all three assets, with the investor taking weakly positive quantities in all three assets. Now suppose that $p_3$ increases slightly. Then it is optimal for the investor to jump to*

*a portfolio allocation where $a_3^i = -\chi$. This can trigger discontinuities in optimal demand.*

### D.2 Section 3

**Remark 4 (Proposition 3)** *Supposing that $\mathcal{D}^i$ is invertible, if $\text{row}(Y) \neq \text{row}(\tilde{Y})$ then $\mathcal{A}^i \neq \tilde{\mathcal{A}}^i$.*

*To see this, for ease of notation, we introduce matrices $M_Y \equiv (Y^+)^T \mathcal{D}^i Y^+$ and $M_{\tilde{Y}} \equiv (\tilde{Y}^+)^T \mathcal{D}^i \tilde{Y}^+$. We have $\mathcal{A}^i \neq \tilde{\mathcal{A}}^i$ if (and only if) $M_Y \neq M_{\tilde{Y}}$.*

*Since $(YY^\mathrm{T})^{-1}$ is an invertible $J \times J$ matrix, the range of $(Y^+)^\mathrm{T}$ satisfies:*

$$\text{Range}((Y^+)^\mathrm{T}) = \text{Range}(Y^\mathrm{T}(YY^\mathrm{T})^{-1}) = \text{Range}(Y^\mathrm{T}) = \text{row}(Y).$$

*Then, we consider the full product $M_Y$. Since $\mathcal{D}^i$ is assumed to be invertible and $Y^+$ has full rank $J$, the product $\mathcal{D}^i Y^+$ is a $J \times Z$ matrix with rank $J$. Thus, invoking the properties of the range of a matrix product,*

$$\text{Range}(M_Y) = \text{Range}((Y^+)^\mathrm{T}) = \text{row}(Y).$$

*Now, suppose toward a contradiction that $M_Y = M_{\tilde{Y}}$. If two matrices are equal, then: $\text{Range}(M_Y) = \text{Range}(M_{\tilde{Y}})$. Substituting our previous result:*

$$\text{row}(Y) = \text{row}(\tilde{Y}).$$

*This contradicts the initial hypothesis that $\text{row}(Y) \neq \text{row}(\tilde{Y})$. Hence, $M_Y \neq M_{\tilde{Y}}$, and consequently $\mathcal{A}^i \neq \tilde{\mathcal{A}}^i$. The proof also states if $\text{row}(Y^+) \neq \text{row}(\tilde{Y}^+)$ then $\mathcal{A}^i \neq \tilde{\mathcal{A}}^i$.*

### D.3 Section 4.1

We remark that we can also provide conditions under which $Y^+$ has a wrong sign for each state (i.e., row).

**Proposition 6** *Under the following two properties, each row of $Y^+$ contains at least one negative element: for each $z \in \{1, \ldots, Z\}$, there exists at least one $j \in \{1, \ldots, J\}$ such* *that* $(Y^+)_{z,j} < 0$.

*(i) Each row of $Y$ has at least two strictly positive elements.*

*(ii) Conical Independence: no column vector $y(z)$ of $Y$ can be written as a non-negative linear combination of the other column vectors of $Y$: for any $z \in \{1, \ldots, Z\}$, there exists no $(\alpha_{z'})_{z' \neq z} \in \mathbb{R}_+^{Z-1}$ such that*

$$y(z) = \sum_{z' \neq z} \alpha_{z'} y(z').$$

Before proving Proposition 6, we discuss its assumptions. Property (i) states that assets typically pay off in multiple states, ruling out only the knife-edge case of Arrow securities. Property (ii) is a weak linear independence requirement: it rules out perfectly redundant states whose payoffs can be exactly replicated by combinations of other states. In the special case in which $J = Z$, property (ii) is automatically satisfied because the assumption that $\text{rank}(Y) = J$ implies that the columns of $Y$ are linearly independent. These properties hold in virtually all realistic asset markets.

**Proof of Proposition 6.** Let $y(z)$ be the $z$-th column of $Y$. Let $y_k^+$ be the $k$-th row of $Y^+$. Suppose to the contradiction that there exists a row $k$ such that $y_k^+ \geq 0$ element-by-element.

Consider the projection matrix $P = Y^+ Y$. The entries are given by $P_{kz} = y_k^+ \cdot y(z)$. It follows from $y_k^+ \geq 0$ and $y(z) \geq 0$ that

$$P_{kz} \geq 0 \quad \text{for all} \quad z \in \{1, \ldots, Z\}.$$

The columns of $Y$ span the range of $Y$. The projection matrix $P$ acts as the identity on the row space of $Y^{\mathrm{T}}$, which implies $YP = Y$. Writing this column-wise for vector $y(z)$, for each $z \in \{1, \ldots, Z\}$, it follows from $y(z) = YP_{\cdot,z}$ that

$$y(z) = \sum_{k=1}^{Z} P_{kz} y(k), \quad \text{that is,} \quad (1 - P_{zz}) y(z) = \sum_{k \neq z} P_{kz} y(k).$$

Since $P$ is a projection matrix, $P_{zz} \leq 1$.

If $P_{zz} < 1$, then we have

$$y(z) = \sum_{k \neq z} \frac{P_{kz}}{1 - P_{zz}} y(k),$$

which is a contradiction to property (ii).

Thus, suppose that $P_{zz} = 1$. Then, $\sum_k P_{zk}^2 = P_{zz}$ implies $P_{zk} = 0$ for all $k \neq z$. This implies

$$P_{zk} = y_z^+ \cdot y(k) = 0 \quad \text{for all} \quad k \neq z.$$

Since $y_z^+ \geq 0$ and $y(k) \geq 0$, let

$$S = \{m \in \{1, \ldots, J\} \mid (y_z^+)_m > 0\}, \quad \text{where} \quad (y_z^+)_m = (Y^+)_{z,m}.$$

The set $S$ is not empty because $y_z^+ \cdot y(z) = P_{zz} = 1$. For all $k \neq z$, and for all $m \in S$, we must have $0 = y_m(k) (= Y_{m,k})$. Take any index $m \in S$. The row $m$ of matrix $Y$ has a value of 0 in every column $k \neq z$. Therefore, row $m$ contains at most one strictly positive element (potentially at column $z$). This contradicts property (i). ■