HOW COMPUTATIONAL COMPLEXITY CAN RESTORE GENERAL EQUILIBRIUM IN MARKETS WITH INDIVISIBLE GOODS*

PETER BOSSAERTS
KONSTANTINOS IOANNIDIS
ROB WOODS
NITIN YADAV

January 2, 2026

Abstract

We study market equilibrium in settings with indivisible goods and tight budget constraints, where a traditional Walrasian Equilibrium (WE) may fail to exist. We introduce the Complexity Compensating Equilibrium (CCE), in which prices endogenously render the budget problem computationally difficult. Complexity induces heterogeneous demands even among agents with homogeneous preferences, as individuals allocate varying levels of cognitive effort. We define the equilibrium region as the set of price configurations that satisfy the necessary economic and computational conditions for equilibrium to exist. In this region, price configurations maximize the difficulty of the budget problem in addition to satisfying market clearing conditions. We evaluate the predictions of CCE through a controlled market experiment. We find that trading prices consistently force the budget problem to the equilibrium region. Further supporting and central to the CCE framework, the equilibrium bundles of goods generate markedly different utility levels across agents. This outcome contradicts a core feature of WE, namely, the equalization of utilities. In a setting where it exists, we reject WE on both prices and utilities, in favor of CCE.

*Peter Bossaerts, Konstantinos Ioannidis and Rob Woods are from the University of Cambridge, Nitin Yadav is from the University of Melbourne. The research reported here was generously funded by the Leverhulme Trust. The authors would like to thank Alec Smith for their valuable feedback on the paper, and conference participants at the Economic Science Association (Brno, 2025), 75th Allied Social Science Associations (San Francisco, 2025), 14th Society for Experimental Finance (Stavanger, 2024) and the 1st Network of the UK-Based Experimental and Behavioural Economists (Leicester, 2024) conferences, and to seminar attendees at the Royal Holloway University of London, the University of Exeter, the University of Vienna, the University of Pittsburgh, and Carnegie Mellon University for their comments.

I Introduction

A long-standing question in general equilibrium theory is whether and how markets equilibrate in the presence of indivisible goods. With homogeneous preferences and symmetric and tight individual budgets, Walrasian Equilibrium (WE) generically does not exist (Henry, 1970).To illustrate why, consider the simplest case of two agents and one unit of an indivisible good. Assume that both agents value the good at v and have budgets of b with $v > b$. For any price $p \leq b$, both agents demand the good, and for any price $p > b$, neither agent can afford the good. Thus, there is no price that clears the market.

Existence of general equilibrium in pure exchange economies involves three conditions: (i) budget feasibility (allocations satisfy budget constraints), (ii) optimality (utility is maximized given preferences and budgets), and (iii) market clearing (demand to equal supply in all markets). When indivisibilities stand in the way of general equilibrium, attempts to restore existence relax one of the typical assumptions about preferences or budgets, or relax one of the conditions for equilibrium. For instance, the economics literature has shown that general equilibrium can be restored if preferences exhibit a sufficient degree of substitutability or complementarity, if individual budgets are sufficiently heterogeneous, or if market clearing is allowed to be approximate. (See Section II for a more detailed literature review.)

Our contribution in this article is threefold. First, we propose a new equilibrium concept, the Complexity Compensating Equilibrium (CCE), which acknowledges the NP-hardness of the budget allocation problem when goods are indivisible, and hence, the need to spend substantial cognitive effort to find the utility maximizing choice. Second, we provide experimental evidence supporting generative predictions of CCE. Third, we pit our equilibrium against WE when the latter exists and provide experimental evidence in favor of CCE and against WE.

CCE builds on the intuition that when goods are indivisible, the agent is facing a computationally complex decision problem, namely, an NP-hard problem (Gilboa et al., 2021).[1] In a typical competitive market, the agent knows the budget as well as the payoffs (personal values) and prices of the goods. Even though this information is minimal, deciding which bundle is optimal is far from trivial. Our conjecture is that when the complexity of agents' individual budget problems is high, agents with homogeneous preferences are scattered in terms of their willingness or capacity to spend cognitive effort. As a result, they will select different consumption bundles, associated with different utility levels. This stratification allows markets to equilibrate.

To distinguish between instances of the budget problem with high and low complexity, we exploit the existence of a *phase transition* in the *decision version* of the budget problem. In the decision version, the goal is to determine whether utility can be improved by switching to another choice. If so, the instance is said to be *solvable*. The phase transition is the region where utility is increased so that the decision instances of the budget problem change from almost always solvable to almost never solvable. In the phase transition, the computations needed to solve an

[1] NP is an acronym that stands for "Nondeterministic Polynomial." Explanation of the origin of this term would require us to get into Turing machines. We believe that this is not necessary for the reader to understand the gist of our arguments, so we refrain from elaborating. The interested reader can consult standard textbooks, such as Arora and Barak (2009), or, perhaps more pertinent for economists, Bossaerts (2025).

instance are high on average, for both electronic computers and for humans.[2]

We argue that, to equilibrate, markets will select price configurations so that the optimal budget allocation reaches utility that locates the decision version in the phase transition. In this way, it is challenging for agents to determine whether they are at an optimum, let alone reach it. Many will stop searching early. This induces heterogeneous demands and hence assists in market equilibration.

We appeal to the existence of a second measure of computational complexity in NP-hard problems, which is the count of optimal solutions, technically referred to as the number of *witnesses*. Humans in particular are more likely to correctly assess whether they have reached an optimum when there are multiple solutions (Franco et al., 2021). To ensure maximum complexity, we therefore expect markets to select price configurations so that the budget problem has a minimal number of solutions. In this way, many agents may not find the optimum, ensuring that demands become heterogeneous and markets can equilibrate.

On top of these computational requirements for existence of CCE, we impose traditional market clearing conditions that are necessary given the scarcity of available goods. If, as will be the case in the experiment we run to verify our theory, there are only two units of three indivisible goods per three agents, then all goods pairs will have to be affordable, but nobody should be able to afford all three goods. Unique to CCE is the necessary condition that different levels of utility accrue to the various bundles that agents choose to hold. There cannot be more than one bundle that reaches maximal utility unless it takes equal cognitive effort to compute another optimum. Other bundles that are held in equilibrium must earn less utility because they require less cognitive effort. This will be in sharp contract with the predictions of WE (when it exists) as we shall point out later. The requirement that there should be a minimal number of optima squares nicely with the second computational condition for CCE we discussed earlier.

We limit our attention to verifying the necessary conditions of equilibrium. To prove existence or to compute an equilibrium (or equilibria if multiplicity obtains), we would have to make assumptions about human cognitive effort allocation in NP-hard problems, about which we have limited knowledge and hence, which we cannot reasonably control for. Our knowledge is limited, among other reasons because humans resort to many different algorithms (heuristics) to solve or approximate NP-hard problems, both over time and in cross-section. For the closely related 0-1 Knapsack Problem (KP) this was first shown in Meloso et al. (2009), confirmed in Murawski and Bossaerts (2016), and exploited in Bossaerts et al. (2024).[3] Moreover, effort allocation at times appears to defy economic intuition: as the chance to find the optimum decreases, the value (utility) reached increases (Murawski and Bossaerts, 2016; Bossaerts and Schultz, 2025; Bossaerts, 2025).

[2]Phase transitions exist for other NP-hard problems. For instance, Yadav et al. (2020) demonstrate existence of a phase transition in the 0-1 Knapsack Problem (KP), closely related to the budget problem. In the budget problem, cash can be used to exhaust one's budget, while in the KP, there is no cash; all items are indivisible. The relation is made precise in Section IV.

[3]In KP with 10 goods, Murawski and Bossaerts (2016) report that, individually, 20 participants visit only 3.6% of all feasible knapsacks, while collectively visiting 42.1%. The heterogeneity in search approaches suggests that communication should be beneficial. This has proven to be correct; purposely designed markets manage to substantially improve individual performance, thus spreading knowledge in society, as originally predicted in Hayek (1945). See Meloso et al. (2009) and Bossaerts et al. (2024) for details.

Nevertheless, to give the reader an idea of the nature of CCE equilibria, we introduce an example (see Section III) assuming that agents choose an algorithm from a family of algorithms referred to as the *Sahni-k family*. Although this family provides a good framework to elucidate how humans deal with the KP and the budget problem, it is far from all-encompassing. The class has been used to construct a metric of instance complexity that correlates with human performance (Meloso et al., 2009; Murawski and Bossaerts, 2016), the performance of nonhuman primates (Hong and Stauffer, 2023), and the informational efficiency of prices in financial markets (Meloso et al., 2009; Bossaerts et al., 2024).

It is insightful to consider the case where WE *does* exist (despite indivisibilities) and compare its predictions with those from CCE, both in terms of final holdings and in terms of prices. Under indivisibilities existence of WE typically requires there to be multiple optimal goods bundles, *all delivering the same level of utility.* The opposite happens in CCE: agents choose bundles with differing utility; which bundle they end up choosing and hence which utility they end up with depends on how much cognitive effort they are willing to apply. Thus, under CCE, we expect distinct categories of participants separated by the utility of their final holdings. In contrast, under WE, all participants should receive equal utility even if trading to different final holdings.

Since WE requires that the budget optimization problem feature multiple optima, in equilibrium the budget problem is computationally simpler. Indeed, as mentioned before, humans recognize optima more often when there are many. We therefore expect WE to be generally *outside the phase transition.* In our experiment, as an example, WE will indeed be far away from the phase transition, increasing the power to distinguish CCE and WE in the data.

To provide empirical evidence for CCE, we conducted a controlled laboratory market experiment with three indivisible goods and cash. As is standard in experimental research on general equilibrium theory (e.g., Bossaerts et al. (2007); Asparouhova et al. (2016)), the market institution we use is a continuous double-sided auction in which both consumers and sellers submit limit orders that remain in an open book unless they are cleared (traded) or they are canceled. Orders are cleared using price-time priority matching. The aggregate supply and demand are chosen so that there are two units of a good available per three consumers. Consumers are endowed with and value cash but can receive pre-announced payoffs for the first unit of each good they acquire in the marketplace, thereby increasing their take-home earnings if they can acquire it at a price below payoff. The sellers initially hold the entire supply of the goods but receive no payoff for them; only cash has value for them, so they want to sell as much as possible, at any price.

We implement two treatment variations. First, to ensure that our findings are robust to changes in the numerical values of the consumer problem we vary, within-participant, the goods payoffs. Each session consists of 16 trading periods split into two blocks of eight periods, and the payoff configuration is fixed within each block. The order of the two blocks is counterbalanced between sessions. Second, we vary, between-participants, the cash endowment ("income") of the consumers. In the low-income treatment WE does not exist, while it does in the high-income treatment. Between 16 and 20 participants participate in each session. We thus present experimental evidence from 180 participants and 2,825 participant-periods. In total, participants submitted 17,628 limit orders resulting in 4,121 trades.

In our setting, equilibrium requires, among others, that all goods are transferred from the sellers to the consumers since the former face zero marginal cost while the latter have positive utility for the first unit of a good. Equilibrium (whether CCE or WE) requires that prices are sufficiently high so nobody can afford all goods, implying that sellers will earn substantial surplus. But past experiments with a single good (plus cash) and a supply side with zero or low marginal costs have invariably produced prices that were biased towards zero, far below equilibrium predictions (Smith and Williams, 1982; Rasooly, 2022). If the findings in these simple partial-equilibrium experiments have any bearing on our general-equilibrium multi-market setting, then our design effectively stacks the deck against finding any evidence of equilibrium, whether CCE or WE.

We find overwhelming support for both the computational and economic requirements of CCE. When WE exists, it does not emerge; trade prices and utilities attained continue to be consistent with CCE. When prices move away from the region where they satisfy necessary computational and/or economic conditions for CCE, vector autoregression (VAR) analysis shows that they tend to revert. When the values (utility levels) of the equilibrium goods bundles are too close, VAR analysis demonstrates that they tend to diverge again, consistent with CCE. Importantly, our data show that sellers manage to extract massive economic rents, despite incurring zero marginal cost, and in contrast to the aforementioned evidence from partial-equilibrium experiments.

Past research in economics involving computational complexity has focused on the difficulty of computing equilibria; see, e.g., Rust (1996), or Judd (2001). Our analysis shifts the focus away from computational complexity at the market level and towards that of the agents trading in the markets. In an influential article, Hayek (1945) criticized the lack of appreciation among economists for the complexity individual agents face when solving budget (or production) problems even if they have to take market prices as given. However, Hayek's hypothesis was that market equilibrium would make individual agents' problems easier. We come to the opposite conclusion: in the face of homogeneous preferences, computational complexity at the agent level is needed for markets to equilibrate.

The remainder of the paper is organized as follows. We discuss related literature on equilibrium existence under indivisibilities in Section II. Section III provides a numerical example of a CCE based on stylized assumptions of effort allocation, while Section IV presents necessary computational and economic requirements for CCE that do not require us to understand how exactly humans solve the budget problem. Section V discusses our experimental design in full detail and identifies testable predictions for the necessary conditions of CCE. Results are analyzed in Section VI, and we provide a concluding discussion in Section VII.

II Literature on equilibrium existence with indivisible goods

Ever since Henry (1970), the literature has attempted to uncover conditions under which equilibrium existence can be restored. It is important to note that the literature often studies the corresponding allocation problem where a central planner allocates goods to agents. Under certain conditions, the second welfare theorem guaranties that an efficient allocation of the central planner problem can be supported as a competitive equilibrium. Although theoretically equivalent, our approach focuses on the market representation of the problem because most goods markets are organized without a central planner, and because the computational complexity the market participants are facing plays a central role in our approach.

One strand of literature focuses on the preferences of agents. Without giving up homogeneity of preferences, the equilibrium existence is restored if the indivisible goods are gross substitutes (Kelso Jr and Crawford, 1982; Gul and Stacchetti, 1999), net substitutes (Danilov et al., 2001; Baldwin and Klemperer, 2019; Baldwin et al., 2023), Δ-substitutes, i.e., pairs of bundles that can be obtained by replacing up to Δ goods from each other (Nguyen and Vohra, 2024), or can be split into two groups with goods substitutable within groups but complementary between groups (Sun and Yang, 2006; Rostek and Yoder, 2020). Choi et al. (2018) provides a model that relies on preference heterogeneity. They study the case where consumers have partial knowledge of the payoffs and need to costly search to fully uncover them, and show that sufficient heterogeneity in prior valuations leads to a unique market equilibrium.

Another strand of literature relaxes the assumption of homogeneous budgets. Budish (2011) shows that if budgets are unequal but arbitrarily close and the number of traders is large, equilibrium can be restored in the limit. Babaioff et al. (2021) analyze budget perturbations that allow equilibrium to exist with a small number of traders.

A third approach relaxes the assumption of exact market clearing. This approach builds on the notion of social-approximate equilibria in which excess demand is bounded, and crucially relies on the assumption of large number of agents (Dierker, 1971; Broome, 1972; Svensson, 1984). This assumption is typically combined with relaxing preference homogeneity (Nguyen and Vohra, 2024), relaxing budget homogeneity (Budish, 2011), or by allowing agents to be allocated a lottery over goods (Budish et al., 2013; Gul et al., 2024; Gul and Pesendorfer, 2025).

A further approach in the literature has been to define alternative equilibrium concepts. In the context of markets for votes, Casella et al. (2012) propose an ex-ante type of competitive equilibrium: agents submit probabilistic demands and market clearing happens only in expectation. They impose a rationing rule to clear the market ex-post. Florig and Rivera (2017) propose a rationing equilibrium which relies on fiat money having a positive price and consumers having knowledge of the demand-supply imbalance. In the context of competitive lending, Asparouhova (2006) studies equilibrium existence by building on the model of Rothschild and Stiglitz (1976), who define an equilibrium as a set of loan contracts such that further contracts can no longer be introduced without causing losses. By expanding this to profitable addition of *bundles* of contracts, Asparouhova (2006) clarifies equilibrium existence and Pareto sub-optimality. She provides experimental evidence that the original Rothschild-Stiglitz equilibrium emerges only if it is Pareto optimal.

Our approach is closer in spirit to the latter strand of literature as we also define a new equilibrium concept. Our equilibrium is relevant in that it exists despite homogeneous preferences and budgets. We do not require any substitutability of complementarity of goods, and only assume additive preferences. The assumption of homogeneous preferences and budgets may seem extreme, but it is meant to be in opposition to the mainstream literature, which has appealed to sufficient preference or budget heterogeneity to assure equilibrium existence.

III An illustrative numerical example

Consider an economy with identical agents, with additive utility in goods and cash, and unit demand for goods. The agents compete to obtain *three* possible goods: (i) a holiday to visit family in Spain, with a payoff of 1,200, (ii) an electric bike, with a payoff of 2,400,[4] and (iii) a professional camera kit, with a payoff of 3,000.[5] The payoffs can be interpreted as the monetary equivalent of the consumption utility of each good. Agents only value the first unit of a good they acquire. We denote the three goods as L, M, and H indicating low, medium, and high payoff respectively. The supply for each good is equal to *two* for every three agents, and they are indivisible. A market equilibrium in this economy would require one third of the agents to purchase goods $\{M, L\}$, one third to purchase goods $\{H, L\}$, and one third to purchase goods $\{H, M\}$.[6]

Traditional WE prices would be such that the agents are indifferent between all possible pairs of goods. The minimum budget for equilibrium to exist is 3,000.[7] If we assume that the agents have 3,200 in cash and a tick size of 50,[8] then the unique equilibrium prices would be 100 for the holiday (L), 1,300 for the bike (M) and 1,900 for the camera (H). For those prices, the agents obtain a profit (surplus) of 1,100 from each good and receive a total utility of 5,400 from the equilibrium pairs.

What has not been widely appreciated in the literature yet is that, with indivisibilities, the budget problem is difficult; in the language of computer science and complexity theory it is NP-hard (Gilboa et al., 2021). Our equilibrium concept, CCE, take this into account. We propose that agents use different algorithms in their attempt to solve their budget problems. These require different levels of cognitive effort. The more effort required, the closer an algorithm gets to the optimum. One can think of one of these algorithms as a *heuristic*. Some heuristics are better than others.

A family of approximation algorithms that has been found to explain to some degree how humans choose in the KP is the Sahni-k class, derived from the Sahni-Horowitz algorithm (Meloso et al., 2009; Murawski and Bossaerts, 2016). The simplest ($k = 0$), lowest-effort member of this class is the greedy algorithm, whereby goods are sorted in descending order of

[4]Note that the indivisibility of the bicycle can be resolved by introducing e-bike renting. Interestingly, however, the introduction of e-bike renting in reality has not solved the problem of indivisibilities, because rentals generally require subscriptions. The subscription contract reintroduces indivisibility.

[5]Like with the bicycle, rental contracts could also be introduced to overcome indivisibility in the case of the camera kit. There to, a subscription may be required however. But it is impossible to imagine a way to make the visit to family in Spain perfectly divisible. One cannot opt for a halfway trip, since that would deceive the purpose of the visit.

[6]To see this, remember that consumers are identical, and hence face the same budget constraint. If some consumers obtain a basket with only one good (say, H), and others baskets with two or three goods, H must be priced such that the former group cannot afford a second good. But since all agents have the same budget, the price of H must also be too high for the latter group. Equilibrium does not obtain because only one unit of H is demanded per three agents, while two units of H are supplied per three agents. Similarly, if some consumers buy nothing and the others all three goods, all the goods must be sufficiently expensive for the former group not to buy anything, meaning no one can afford the full basket. Therefore, for total demand to equal total supply, the number of agents holding each of the possible pairs of goods must be equal.

[7]Appendix A provides a formal derivation of the minimum budget for Walrasian equilibrium existence, and restrictions on the price range for each good when equilibrium is not unique.

[8]The *tick size* is the distance between allowable prices, usually starting from a single tick. Tick sizes are universally imposed in organized financial markets.

the ratio value/price, defined as the *density*, and the goods are purchased until exhausting one's budget.[9]

For $k = 1$, the algorithm is altered by cycling through all feasible initial allocations of one item, followed by application of the greedy algorithm on the remaining items, retaining eventually only the choice that maximizes total value. For $k = 2$, the algorithm cycles through all feasible pairs of items, followed by application of the greedy algorithm, again retaining only the highest-value choice. The algorithm is defined analogously for higher values of k. Clearly, *cognitive effort increases in* k. But *performance also increases in* k. With $k = 0$ the algorithm can reach the optimum only if the instance can be solved correctly by the greedy algorithm. With $k = n - 1$ where n is the total number of goods, the optimal is always found.

The following statistic offers an idea of how k affects the ability of humans to solve the KP: when comparing two instances where the k required to reach the optimum is increased by one, value attained (as a percentage of the optimal value) decreases as much as increasing the number of available items by *five* units. That is, with three items, going from $k = 2$ to $k = 3$ is equivalent to solving instances with eight items. See Bossaerts (2025).

For our example, we posit that agents sort into categories defined by k. Assuming agents have a budget of 3,200, this sorting could lead to the following CCE. Referring to the prices in Table Ia, the greedy algorithm would rank the holiday (L) first, followed by the bike (M), and the camera (H) would be ranked last. An agent using the greedy algorithm ($k = 0$) first selects L, and then proceeds with M. Their budget is insufficient for H, so they end up with the pair of $\{M, L\}$ goods. An agent using the algorithm with $k = 1$ preselects a good H which was not in the greedy solution, and then uses the greedy algorithm to select the second good. Given the ratios, that second good is L, so they end up with the pair of $\{H, L\}$ goods. Finally, an agent using the algorithm with $k = 2$ preselects H as the first good, and also preselects M as the second good. Their budget is insufficient for more goods, so they end up with the pair of $\{H, M\}$ goods. If one third of agents use the greedy algorithm ($k = 0$), one third uses the Sahni algorithm with $k = 1$, and one third uses the Sahni algorithm with $k = 2$, then the markets clear. A CCE is reached.

An important feature of our equilibrium, which also justifies the name, is that the agents who spend the highest effort find the optimal solution and gain the most (buying H and M); those who spend moderate effort find a good solution and gain less (buying H and L); and those who spend the least effort find the worst solution and gain the least (buying L and M). As can be seen from the last column of Table Ia, the agents are compensated for the cognitive effort of using a more sophisticated algorithm. The agent who uses $k = 2$ receives a utility of 5,600, the agent who uses $k = 1$ receives 5,200, and the agent who uses $k = 0$ receives 5,000. This is in sharp contrast with WE, which ignores the cognitive effort needed to make better choices and requires prices to ensure indifference between all equilibrium demand bundles. Thus, in CCE, larger cognitive effort is compensated for by higher utility. Notice also that agents applying maximal effort ($k = 2$) obtain a utility level that surpasses that of WE (5,600 *vs.* 5,400).

[9] More sophisticated versions exist, whereby, e.g., items are skipped when they violate the budget constraint as long as there are remaining items that do fit within the budget. It deserves emphasis that the greedy algorithm finds the optimal solution if all goods are infinitely divisible, but not if some goods are indivisible, or if only one good is infinitely divisible (as with cash remaining in one's budget in our example).

TABLE I Examples of CCE: (a) when WE exists, (b) and when WE does not exist

Market outcomes				Individual outcomes		
Good	Payoff	Price	Density	Sahni-k	Bundle	Utility
L	1,200	500	2.400	$k=0$	(M,L)	5,000
M	2,400	1,300	1.846	$k=1$	(H,L)	5,200
H	3,000	1,700	1.765	$k=2$	(H,M)	5,600

(a) Budget 3,200

Market outcomes				Individual outcomes		
Good	Payoff	Price	Density	Sahni-k	Bundle	Utility
L	1,200	400	3.000	$k=0$	(M,L)	4,800
M	2,400	1,200	2.000	$k=1$	(H,L)	5,000
H	3,000	1,600	1.875	$k=2$	(H,M)	5,400

(b) Budget 2,800

Now consider the case where the budget is 2,800, in which case WE does not exist. With tighter budgets, one would arguably expect prices for all goods to decrease. Consider such prices in Table Ib. The greedy algorithm still ranks good (L) first and good (H) last. If the agents use algorithms as before, then CCE continues to exist. We also note that while all pairs yield lower utility than before, there is still a cognitive premium so that the agents who demand the optimal bundle get the highest utility, of 5,400, and the agents who demand the second-best and third-best bundles receive lower utilities, of 5,000 and 4,800 respectively.

IV CCE

We now introduce necessary conditions for CCE which can be verified on empirical data. The idea behind CCE is that markets select price configurations that make the budget problem sufficiently complex so that agents, stratified in terms of cognitive effort, choose different consumption bundles, with only those who apply most effort reaching the optimum. Cognitive effort thereby induces demand heterogeneity.

Before introducing the computational requirements for CCE, we discuss computational complexity of the budget problem and show existence (and location) of a phase transition where the most difficult instances are expected to reside, and hence where we expect CCE to locate.

IV.A The budget problem with indivisible goods

Consider an economy with I indivisible goods, each fully characterized by a *price* p_i and a *payoff* ϕ_i. There are N agents in the economy, each endowed with a budget (income) of C. Agents exhibit linear utility in goods and cash, and demand at most one unit of each good. Agents intend to choose the payoff maximizing bundle among all the affordable bundles. Since all goods are indivisible, an agent's choice can be simplified to a series of binary choice variables x_i, which take the value of 1 if good i is purchased and 0 otherwise.

Formally, each agent faces the following optimization problem.

$$\max_{\substack{x_i \in \{0,1\},\\ i=1,2,\dots I}} \left\{ \sum_{i=1}^{I} x_i \phi_i + \left(C - \sum_{i=1}^{I} x_i p_i \right) \right\},$$

$$\text{subject to: } \sum_{i=1}^{I} x_i p_i \leq C.$$

The objective function can be rewritten as follows.

$$\max_{\substack{x_i \in \{0,1\},\\ i=1,2,\dots I}} \left\{ C + \sum_{i=1}^{I} x_i(\phi_i - p_i) \right\} = \max_{\substack{x_i \in \{0,1\},\\ i=1,2,\dots I}} \left\{ C + \sum_{i=1}^{I} x_i v_i \right\}. \tag{1}$$

The objective function is linear in the choice of goods, with *value coefficients* equal to $v_i = (\phi_i - p_i)$. We expect prices of chosen goods to be below goods payoffs ($p_i < \phi_i$), otherwise the agent prefers to hold cash. Therefore, the value coefficients are non-negative ($v_i \geq 0$). Written as in Equation 1, and ignoring the constant C, one can deduce that the budget problem is a special case of KP. See Gilboa et al. (2021) for a direct proof of NP-hardness.

The hardness of the budget problem originates in the difficulty of verifying whether one has found the optimum. With infinitely divisible goods, and provided utility satisfies minimal smoothness conditions, determining whether one has reached an (interior) optimum merely requires to check first and second order conditions, i.e., to check how the solution changes in a small neighborhood. With indivisibilities, an NP-hard *decision* problem has to be solved instead, namely, whether there exists a solution with target value at least as large as the one already attained plus one (utility) unit. If so, the instance is said to be *solvable*.

We are interested in distinguishing difficulty among different *instances* of the budget problem. In the numerical example of the previous section, we assumed that agents use algorithms in the Sahni-k family to solve the budget problem. Fixing the algorithms agents use is a powerful tool to obtain concrete analytical results, but it is too strong of an assumption for empirical analysis for the reason cited before: we lack knowledge of how cognitive effort is allocated in budget problems, or even whether this allocation is done optimally. In order to derive testable implications, we instead resort to generic complexity measures of instances of the budget problem which do not depend on the algorithms used by the agents.

For KP, an *ex ante* metric of difficulty is whether the decision version of the instance is in the phase transition or not. With *ex ante*, we mean that it depends on features of the instance that can be recognized without any attempt at solving it. With *phase transition*, we mean that there is a region where instances change from always-solvable to never-solvable.[10] Existence of such a phase transition for KP was first shown in Yadav et al. (2020).[11] There, it was also

[10]An example of a phase transition in nature is the dew point: a small temperature interval where evaporation and condensation offset each other; for a higher temperature, evaporation dominates; at lower temperatures, condensation dominates.

[11]Many other computationally hard problems exhibit phase transitions, such as the satisfiability (3SAT) problem (Mitchell et al., 1992), the graph coloring problem (Cheeseman, 1991), the number partition problem (Gent and Walsh, 1998), and the traveling salesman problem (Gent and Walsh, 1996).

shown that the instances requiring most effort – for humans and electronic computers alike – reside in the phase transition. See also Franco et al. (2021, 2024); Bossaerts (2025).

While the budget problem can be re-written as a KP [see Equation 1], it is not obvious whether such a phase transition exists for the subset of budget problems we are interested in, where we fix item (goods) payoffs (ϕ_i in our notation). The reason why is that payoffs are agents' utilities for the individual goods, which the market cannot change; to alter difficulty, the market can only change prices, p_i.[12] In terms of the classical KP, this means that value coefficients v_i and costs p_i are negatively correlated since $v_i = \phi_i - p_i$. The issue then is whether KP instances where values and costs are negatively correlated also feature a phase transition.

The phase transition of the *budget decision problem* is defined as a region in a two-dimensional space. The dimensions of this space depend on the payoffs and the prices of the available goods, the budget, and a target utility Υ. The first dimension is the *normalized capacity* (κ), defined as the ratio of the budget to the total price of all goods. Informally, normalized capacity measures budget feasibility. Formally, it is defined as follows:

$$\kappa = \frac{C}{\sum_i p_i}. \tag{2}$$

The second dimension is the *normalized profit* (π), defined as the ratio of the target value Υ and the total value from purchasing all goods. Informally, normalized profit captures how difficult it is to find an allocation that reaches the target value when disregarding the budget constraint. As such, normalized profit captures value infeasibility. Formally, it is defined as follows:

$$\pi = \frac{\Upsilon}{C + \sum_i(\phi_i - p_i)}. \tag{3}$$

In the decision version of our budget problem, we ask whether a target value Υ can be reached or exceeded. If yes, then the instance is solvable; otherwise it is not. When the instance is in the phase transition, i.e., the target value, payoffs and prices are such that κ and π are in the phase transition, then uncertainty about solvability is highest and effort required to determine solvability tends to be highest.

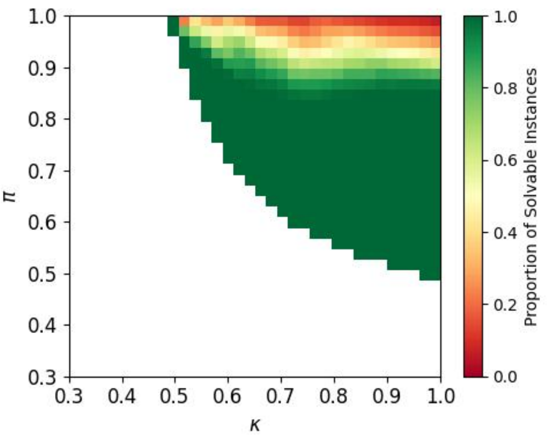
Figure Ia displays, for each combination of κ and π, the fraction of decision problem instances that are solvable. Budget and payoffs are fixed as in one of the treatments in our experiment.[13] Shown are all instances for which prices are "*rational*," in the sense that prices are below payoffs, and prices are ranked in the same order as payoffs. Prices are forced to be on a finite grid, with tick size equal to 0.05. This reflects the reality in financial markets (and in our experiment) that trade can happen only at discrete price levels.[14] For each price configuration, we test

[12] It is worth pointing out that fixing payoffs may render the budget problem less difficult, moving it outside the set of NP-hard problems. This does not make NP-hardness less relevant for the agents, who may face different payoffs in different settings, and hence, who may need to choose an appropriate algorithm (a way to solve budget problem instances) that takes into account all possible payoff structures they may encounter. But here we analyze the budget problem from the point of view of the market, who faces a particular payoff structure and needs to select prices that ensure maximal difficulty.

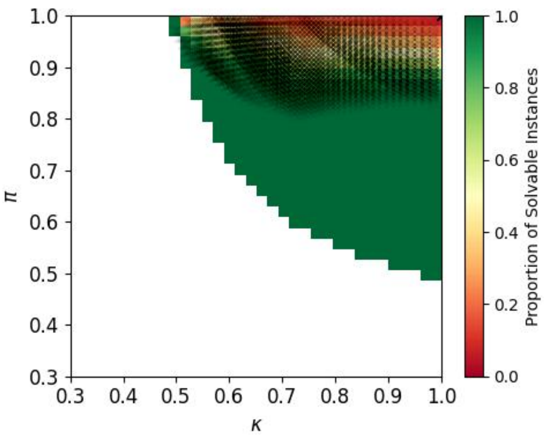
[13] Treatment "WE exists" with payoff configuration "C1." See Table II.

[14] The boundary of the region for which there exist instances reflects these constraints. For instance, at prices equal to payoffs, $\kappa \approx 0.48$ and $\sum_i p_i = \sum_i \phi_i$, π ($= C/(C + \sum_i \phi_i - \sum_i p_i)$)$= 1$. So there is only 1 possible value for π when $\kappa \approx 0.48$.

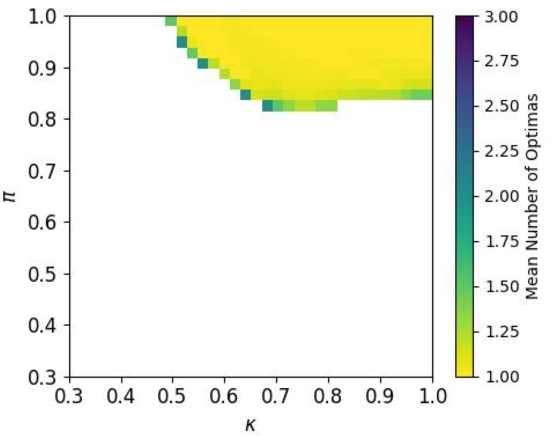
every feasible target.[15] We cut κ off at 1 (beyond which $\sum_i p_i > C$) for two reasons: (i) above $\kappa = 1$, the budget allocation problem is trivial since the budget constraint is not binding, (ii) for equilibrium to obtain it is necessary that prices are sufficiently high so purchasing all items is infeasible.[16] We also cut π at 1.



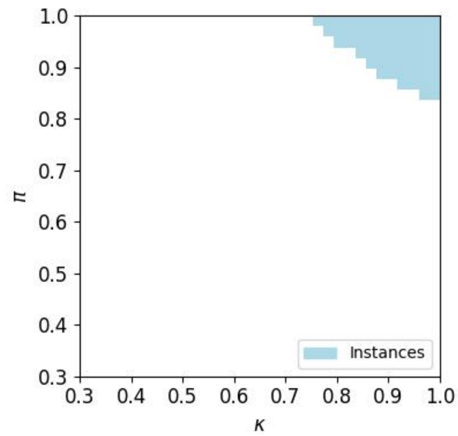
(a)



(b)



(c)



(d)

Notes: $\kappa = C/\sum_i p_i$; $\pi = \Upsilon/(C + \sum_i(\phi_i - p_i))$; C, ϕ_i as in Treatment "Walras – C1" in Table II; Υ and p_i varying. (a) Solvability of decision problem instances. (b) Location of optimization problem instances (black dots) against phase transition. (c) Location of optimization problem instances stratified by mean number of witnesses (optima); e.g., yellow locations contain instances for which the vast majority (more than 95%) have only one optimum. (d) Location of optimization problem instances only for price configurations that make all item pairs affordable, but not the triplet (necessary for equilibrium).

FIGURE I Location of budget problem instances in terms of normalized capacity (κ) and normalized profit (π)

[15] Feasible targets are defined as all multiples of the tick size, from a minimum equal to the budget (which can be reached without the need to trade) to a maximum equal to the total value of purchasing all assets at the lowest possible price.

[16] Unlike in past investigations of phase transitions in NP-hard problems, we do not randomly draw instances of a given κ and π. Instead, we plot the fraction of *all* instances that are solvable. In our case, the space of instances is bounded thanks to the finite price grid. There exist 3,400,320 instances in the treatment displayed in the figure. We thereby avoid biases caused by the way one would draw instances. See the discussion in Yadav et al. (2020).

Figure Ia reveals a distinct phase transition. Fixing κ while increasing the target value Υ – thereby increasing π and moving northwards in the plot – instances change from all-solvable (dark green) to non-solvable (dark red), with a small region where uncertainty about solvability is non-trivial (between light green and light red).

Normalized capacity and normalized profit metrics can be used to define a difficulty metric for the *optimization* version of the budget problem, as has been done for KP (Franco et al., 2021; Bossaerts, 2025). One does so by setting the target value in the definition of the normalized profit equal to the highest attainable value:

$$\pi = \frac{\Upsilon^*}{C + \sum_i (\phi_i - p_i)} = \frac{C + \sum_i x_i^*(\phi_i - p_i)}{C + \sum_i (\phi_i - p_i)} \tag{4}$$

where x_i^* denotes the optimal item allocation. The normalized capacity κ remains defined as before.

An instance of the *budget optimization problem* is then said to be *difficult* if its κ and π locate it in the phase transition of the corresponding budget decision problem. Figure Ib shows the locations of all possible budget optimization problems, mapped on top of the solvability of the decision version, i.e., overlayed on top of Panel (a) of the same figure. We observe that most instances of the budget optimization problem are located around the phase transition. For those below the phase transition, verifying that one has reached an optimum is easy because virtually all instances of the decision version in that region are solvable. In the phase transition, checking whether the optimum is reached is difficult because it is uncertain whether instances of the decision version at those locations are solvable, i.e., that the corresponding target values can be exceeded.

IV.B CCE: Computational Requirement 1

Our first computational requirement for CCE requires prices to locate the budget problem instance of an economy in the phase transition. Agents start with low target utility; in the experiment, they start with only cash equal to the budget C. Through trading, they will attempt to work their way up, increasing utility, and hence, increasing π. In CCE, they will have to go a long way, towards or even beyond the phase transition. As long as they are below the phase transition, it will be generically easy to find ways to improve the utility. Once entering the phase transition, the discovery of improvements is rendered far more difficult, causing those agents who prefer not to spend too much effort to stop in spite of further potential improvements. The result is heterogeneity in demands, and hence, possibility for the market to equilibrate.

IV.C CCE: Computational Requirement 2

Previous research has uncovered a second feature of instances of the KP decision problem that makes it harder for humans to correctly assess solvability, namely, the number of *witnesses* (goods bundles) that reach or exceed the target value (Franco et al., 2021). The fewer witnesses there are, the lower the accuracy in human responses. If there are only few optima to the budget optimization problem, then it will be more difficult to reach optimality. Markets can make the budget allocation problem harder by selecting prices for which there exists one unique optimum.[17]

It is worth recalling that, in WE, prices have to be such that there exist multiple optima for markets to equilibrate. Therefore, the *WE is not complex* when measuring complexity in terms of number of witnesses that support optimal choice.

In Figure Ic, we plot the mean number of optima of budget optimization problem instances in (κ, π) space. The same treatment is used as for Panels (a) and (b) in the figure. We observe a large overlap between the yellow region (where more than 95% of the instances have a single optimum) in Panel (c) and the phase transition region in Panel (a).

We predict that CCE will reside in the yellow region where most instances, if not all, feature unique optima. The scarcity of other solutions to the budget optimization problem thereby makes it less likely that humans discover the optimum, allowing those who spend most cognitive effort to acquire the best goods bundle, while others settle on inferior choices.

IV.D CCE: Economic Requirement 1

Computational complexity makes possible utility-ranked demands if it takes agents effort to find the optimum. This is the principle behind CCE. The other side is purely economic: demand has to equal supply. Let us now investigate the economic side.

We introduce additional notation. Let $S = (s_1, s_2, \cdots, s_I)$ denote the aggregate supply, with s_i denoting the supply for each good i. We assume $s_i < N$ for all goods, to ensure scarcity. Let $x^n = (x_1^n, x_2^n, \cdots, x_I^n)$ with $x_i^n \in \{0, 1\}$ denote the individual demand of agent n for good i. The aggregate demand is denoted as $X = \sum_n x^n$. An equilibrium bundle is a subset of the goods that will have to be held in equilibrium by some agent(s). Let $d^j = (d_1^j, d_2^j, \cdots, d_I^j)$ with $d_i^j \in \{0, 1\}$ denote equilibrium bundle j, and denote by $\mathcal{D}$ the set of all equilibrium bundles $j = 1, ..., J$, where J equals the number of equilibrium bundles. For readability, we will use symbols instead of binary numbers. For example, a bundle $d = (1, 0, 1)$ is equivalent to $d = \{L, H\}$ and $d = \{H, L\}$. Equilibrium requires markets to clear, such that $X = S$. Below we describe the properties that our equilibrium bundles need to satisfy.

CCE requires that the equilibrium bundles are utility ranked. Without loss of generality, we denote this ranking using an index j such that $U(d^j) < U(d^{j+1})$; in other words, bundle d^{j+1} yields higher utility than bundle d^j. The utility ranking ensures that cognitive effort to reach higher utility is compensated. Here, we use strict inequalities, though it could be that two different choices lead to the same utility level and with the same effort.[18] Sticking to strict inequalities, we have a simple necessary requirement for CCE to exist: prices have to be such that there exists a unique optimum. This can be tested in an experiment.

[17]Multiple optima can be allowed only if they can be reached by different algorithms with the same amount of effort (number of computations). The example of Section III assumes that agents choose one of the Sahni-k algorithms. There, the number of computations increases strictly in k. In addition, the algorithms are deterministic, implying that they will always estimate the same optimum for a given instance. Consequently, there does not exist an instance with two or more optima that can be reached with equal number of computations by two different algorithms.

[18]In Section III, this could not happen because of the nature of the algorithms agents use.

IV.E CCE: Economic Requirement 2

To ensure market clearing, we need a second requirement, namely, market clearing. Let y^j represent the mass of agents who choose equilibrium bundle d^j. We require that, for each agent n, there exists some equilibrium bundle d^j such that $x^n = d^j$. So, $y^j > 0$. Market clearing then imposes that the weighted aggregate demand (D) must satisfy

$$D \equiv N \times \sum_j y^j d^j = S.$$

Informally, this means that all agents only select an equilibrium bundle, in appropriate weights so that the weighted sum of the demanded bundles matches the aggregate supply.[19]

Choices have implications for prices, and hence, for the location of the budget problem in (κ, π) space. In particular, prices underlying κ and π have to be such that all equilibrium bundles are affordable, but not all the goods simultaneously. In Figure Id, we plot the location of budget optimization instances (based on π evaluated at the utility for the best bundle only) for price configurations for which all equilibrium bundles are affordable, but not the triplet (here, there are three goods). The same setting is used as for Panels (a), (b) and (c) of the figure.

For the reader to gain perspective, we note that there are 25,760 instances depicted in Panel (c) (where all optimization instances are plotted, regardless of whether they satisfy economic requirements), while in Panel (d), there are only 3,695 instances. Our second economic requirement alone reduces by 85% percent the number of potential budget optimization problem instances that the market can choose from!

IV.F Remark on CCE existence

We repeat here that we are not proving existence of CCE. We have only derived necessary conditions for CCE. Proof of existence would require one to make assumptions about the algorithms agents use to solve the budget problem, their (computational) costs, and how agents trade off cost against utility – if they do at all. Although evidence is emerging (Bossaerts, 2025; Bossaerts and Schultz, 2025), we still do not know enough to build a full theory of CCE and prove existence. But we have an idea where CCE could lie in (κ, π) space, and what equilibrium allocations look like. Consequently, to evaluate whether it is worthwhile investing more time into discovering how humans decide in the budget problem, we should at least determine to what extent markets locate the budget problem in the region that satisfy the above computational and economic requirements, and to what extent markets force agents to hold any of the possible equilibrium goods baskets.

[19] To illustrate, take our motivating example in Section III. There are three equilibrium bundles, namely, $d^1 = \{M, L\}$, $d^2 = \{H, L\}$, and $d^3 = \{H, M\}$. When the budget is 3,200, the bundles yield utilities equal to 5,000, 5,200, and 5,600 respectively. When the budget is 2,800, they yield utilities equal to 4,800, 5,000, and 5,400 respectively. Thus, they are always utility ranked as their corresponding utilities are increasing. Under the assumption that the agents use algorithms in the Sahni-k class, then the utilities correspond to solutions from k=2, k=1, and k=0 respectively. As such, utility rankings correspond to the ranking based on the difficulty of finding them. For market clearing, we require exactly one third of agents to select each demand bundle $\left(y^1 = y^2 = y^3 = \frac{1}{3}\right)$. In that case, the weighted sum of equilibrium demands is $D = 3 \times \left(\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right) = (2, 2, 2)$, which is equal to the aggregate supply of two units of each good for every three agents.

V Experimental design and predictions

The experiment is designed with the numerical example of Section III as inspiration. On the demand side, we induce preferences as in the example. The "goods" traded are assets: claims to a deterministic monetary payoff (hence, the assets are risk-free). On the supply side, we introduce sellers with zero marginal costs. We let markets determine prices and allocations. We organize exchange through a continuous double-sided open limit-order book system. We elaborate below.

Among agents, we distinguish between investors and sellers. We refer to assets instead of goods.[20]

V.A The market and the trading game

In each experimental session, unique participants trade in a market for 16 equal-duration periods. There are three assets available for trade with pre-determined payoffs to investors only. At the end of each period, assets are liquidated and participant period earnings are calculated based on final liquidated holdings. All periods are independent; one's final holdings and actions in one period do not influence any of the other periods. One quarter of the participants are sellers, and three quarters are investors. The sellers act as suppliers of assets. Each seller is endowed with two units of each of the three assets and no cash. They do not have a claim to the payoffs on the assets, so they are induced to sell the assets for as much cash as possible. The investors represent the demand side of the economy. Each investor is endowed with cash and no assets. The investors are only compensated for the *first unit* of each asset they hold at the end of the period, and for any remaining cash.

A noteworthy feature of our incentive scheme is that all assets are initially in the hands of sellers who have zero marginal cost, and all cash is in the hands of investors who all receive the same payoffs on the assets, and hence, who are homogeneous. This feature serves multiple purposes. First, it ensures that the decision of which asset to purchase is difficult for investors while keeping the complexity of the supply side minimal. Second, it maximizes trade, since it is in everyone's interest (i.e., Pareto-improving) that all assets move from sellers to investors. The large expected number of trades increases the power of our design.[21] Third, it creates price dynamics that stack the deck against our equilibrium. CCE requires high prices so that investors can only afford pairs of assets (see Economic Requirement 2). With divisible assets, it is known that inducing sellers with zero marginal cost causes prices to end up far below equilibrium levels (Smith and Williams, 1982; Rasooly, 2022). If such forces are at play for indivisible assets as well, then they must push prices away from CCE. These forces would destroy any evidence in favor of CCE.

The trading protocol is an online open-book continuous-time double-sided auction.[22] Traders

[20]In the Instructions (Appendix D), we refer to investors as consumers. Goods are referred to as assets, again to emphasize that participants trade claims for payouts.

[21]Contrast this with some of the designs in Plott and Sunder (1982, 1988): assuming risk neutrality, we do not expect a single trade, so in expectation we have no observations with which to verify whether prices are right. The same issue emerges in the design of Smith et al. (1988).

[22]The software, Flex-E-Markets, is used on a subscription basis. See quantahm.com. Most major stock exchanges, such as the New York Stock Exchange, the London Stock Exchange, Euronext, and NASDAQ, operate

submit limit orders, i.e., the maximum price they are willing to pay for buying one or more units of an asset or the minimum price they are willing to accept for selling one or more units of an asset. A trade occurs in two scenarios: (i) a buy order is submitted with a price higher than or equal to the lowest standing selling price, or (ii) a sell order is submitted with a price lower or equal to the highest standing buying price. If a trade occurs, the price is determined by the best standing order at the moment a trading order is submitted. In the first scenario, the trade is triggered by an investor submitting a buy order, so the standing sell order (ask) sets the price. In the second scenario, the trade is triggered by a seller submitting a sell order, so the standing buy order (bid) determines the price. If the buy and sell orders are for different quantities, the order with the higher number of units is split and the remaining units (after trade) are converted into a new order at the original order price and the new order is then subject to the same trade evaluation process. If trade does not occur, the order becomes a standing order and is anonymously displayed to all other traders, ready to trade with later orders. Standing orders can be canceled by the submitter prior to trading and are cleared at the end of each period.

The trading interface consists of three parts: the order book, the order form, and the holdings account. A snapshot of the trading interface can be found in Appendix E.

The order book consists of three panels, one for each of the assets. Each asset-panel updates in real time and stores all standing orders with their price and corresponding quantity. The standing orders are anonymized and participants can only identify their own standing orders. The orders are priority-ranked. Buy orders are shown in descending order and sell orders in ascending order. For orders with the same price, earlier orders receive priority and execute first, however quantities are combined in the interface to show the total standing volume at a given price. Participants also see the (anonymized) trade history of all trades that occurred for each asset in chronological order. Color coding ensures that the trader can determine whether the trade was triggered by an incoming buy order (blue) or an incoming sell order (red).

The order form is the only way participants can communicate with the market and its participants. No other communication between participants is allowed. Participants can submit buys and sells for any or all of the three assets. The submitted prices have to be positive and cannot exceed a maximum price.[23] A tick size of 0.05 is imposed. This tick size was chosen, among others, so that, in the two treatments where WE exists, WE prices are unique.

The holdings account displays for each participant their settled and available assets and cash. "Settled" holdings represent the participant's real time claims, which are set at the start of each period. Placing an order into the market does not affect these claims until the order is traded. When a trade occurs, the settled balance updates to reflect that trade (asset increase and cash decrease for a traded buy order, and asset decrease and cash increase for a traded sell order). At the end of the period, liquidation is based on the settled holdings at that time. "Available" holdings act as a constraint on participants. They equal the settled holdings minus any assets or cash committed to active orders. For example, if a seller has 2 units of H and places a sell order for 1 unit, that unit becomes committed and unavailable for other orders,

with this trading protocol. Settlement in Flex-E-Markets is immediate, however, and therefore part of the clearing mechanism, which puts additional computational constraints on the software.

[23]Maximum prices for all assets were set equal to 25% above the highest payoff of all three assets, rounded up to the nearest experimental currency unit.

leaving an available balance of 1. The seller can cancel the standing order to free the unit for a different price. Similarly, when an investor places a buy order, the cash offered is committed, reducing their available cash by the buy order amount. Available holdings adjust whenever orders are placed or canceled. Available holdings also incorporate shorting limits, though in our experiment short sales and cash borrowing are not permitted.

V.B Experimental treatments

Our experiment implements a mixed 2×2 design with four treatments. One treatment variable is manipulated between participants, and one is manipulated within participants. Table II presents a concise overview of all our experimental treatments.

TABLE II Overview of experimental treatments in terms of cash (budget) and payoffs for investors

Budget Treatment	Payoff Treatment	Cash C (Budget)	Asset Payoff (ϕ_i) L	M	H	Total sessions	Total participants
No Walras	C1	2.80	1.20	2.40	3.00	5	88
	C2	2.20	1.60	2.40	3.20	5	88
Walras	C1	3.20	1.20	2.40	3.00	5	92
	C2	2.60	1.60	2.40	3.20	5	92

In the first two sessions of the "Walras" treatment, the payoffs for all assets were one fourth of the values presented here; there, we used a conversion rate of experimental currency into pounds of 1:4. The payoffs in all other sessions were as displayed above, and a conversion rate of 1:1 was used.

Across sessions, and hence *across participants*, we vary the cash endowment to participants acting as investors. In one half of our sessions, referred to as "No Walras" sessions or Sessions S1 to S5, investors are endowed with cash that is insufficient for WE to exist. In the remaining sessions, called "Walras" sessions or Sessions S6 to S10, the cash endowment is sufficient for WE to exist.

Within participants, we vary the asset payoffs to the investors. We refer to the two payoff configurations as "C1" and "C2." In each session, one block of eight periods uses payoff configuration C1, and a second block uses payoff configuration C2. The order of the blocks is counterbalanced between sessions. We vary the asset payoffs as a robustness check to ensure that our results are not driven by a specific payoff parametrization. Note that while we refer to the assets as H, M and L here, in the experiment they were called A, B and C to prevent any biased expectations by the participants, and the letter assignment was shuffled between payoff configurations.

Within payoff configurations, we also vary roles (investor; seller), so that all participants are exposed to both sides of the economy. One in four periods, a participant is a seller; in the remaining periods, the participant acts as investor. The rotation is implemented to induce fairness and to ensure that investors understand the incentives for sellers and *vice versa*.

V.C Predictions

This section will motivate several testable predictions that come out of our theory. The predictions are presented informally, while formal statistical tests for the predictions are discussed along the results.

We first present predictions that need to be fulfilled to satisfy the economic requirements for CCE listed in Section IV. Subsequently, we deal with predictions that flow from the computational requirements for CCE. Finally, we add predictions that would obtain only for the "Walras" Treatment since they pertain to the WE.

V.C.1 Economic Predictions

Because sellers have zero marginal cost and investors assign positive value to the entire supply of available assets, all assets need to be sold. As such, the resulting re-allocation exhibits high efficiency. From the investor side, due to the ratio of potential aggregate supply and demand, market clearing would require an equal share of investors to purchase each of the three pairs of assets (see Economic Requirement 2).

Three predictions follow. They do not depend on the level of cash (budget) allocated to the investors, so obtain in both the Walras and No Walras treatments.

Prediction 1. *Efficiency is 100%: all asset units are sold to the investors.*

Prediction 2. *All investors end up holding pairs of assets.*

Prediction 3. *Exactly the same number of investors end up holding each pair of assets.*

As to prices, investors should be able to afford two but not three assets. Among others, this implies higher prices when budgets are higher (as in the Walras Treatment). Therefore, we formulate the following predictions.

Prediction 4. *Prices are sufficiently high so that any pair but not all three assets can be afforded.*

Prediction 5. *Prices are higher when investors have larger budgets.*

So far, our predictions are required for any generic equilibrium to hold. However, one more restriction on prices needs to be imposed. CCE requires that the equilibrium pairs entail different utility levels, so that cognitive effort expended to reach the optimum is compensated.

Prediction 6. *The three equilibrium asset pairs earn different levels of utility.*

V.C.2 Computational Predictions

We now present predictions regarding computational complexity. The first prediction is driven by the requirement that the computational complexity of the investors' budget problem needs to be high. As discussed earlier, we interpret this to mean that the market selects prices so that the budget optimization problem is located in the phase transition. This is a rewording of Computational Requirement 1.

Prediction 7. *Prices set the investor problem into the region of the phase transition.*

Since prices are determined endogenously as orders arrive in the market, they (prices) may occasionally move the budget problem out of the phase transition. There, the budget problem tends to be easy, and hence, with little effort all investors may find the optimum. Demands are no longer heterogeneous, so markets cannot equilibrate. Prices must push the budget back into the region of phase transition.

Prediction 8. *Prices mean-revert to locate the budget problem inside the phase transition.*

For humans, the chance of finding the optimum in allocation problems such as KP increases when there are multiple optima (multiple "witnesses"). This means that complexity, to humans at least, rises when there are fewer optima. We referred to this as Computational Requirement 2. It leads to the following prediction.

Prediction 9. *Prices are such that the budget optimization problem is in the region were most if not all instances feature a unique optimum.*

This prediction is obviously related to Prediction 6. When prices are such that the budget allocation problem has a unique optimum (Prediction 9), it must be that at least two of the equilibrium asset pairs generate different utility levels. Conversely, if utility levels of all equilibrium pairs are the same, contrary to Prediction 6, Prediction 9 cannot be right, unless nobody optimizes.

V.C.3 Walrasian Equilibrium

We also aim to test CCE against the traditional WE when it exists, as is the case in the "Walras" Treatment. In that case, the lower complexity of WE (it is outside the phase transition and it entails a budget optimization problem with three optima) favors its emergence. However, this assumes that the market has already settled on WE prices. Since there are more price configurations that are consistent with its necessary conditions, CCE may be obtained instead. We formulate a hypothesis that can be tested by investigating where prices appear to converge.

Prediction 10. *Asset prices converge to levels that are significantly different from WE prices.*

Another forceful way to reject WE is to reject that equilibrium pairs all gain the same level of utility. That is, WE is falsified if the data support Prediction 6.

V.D Implementation

We ran ten in-person experimental sessions in total. The sessions took place at the Cambridge Experimental and Behavioural Economics Group (CEBEG) Laboratory of the Judge Business School (Cambridge), between March and November of 2024. All participants were recruited from the CEBEG participant pool which is open to the general population of 18 years or older and able to attend in-person experiments. Ethical approval was obtained prior to data collection, which included the requirement of informed consent.[24]

[24]Ethics Approval UCAM-FOE-24-02 (University of Cambridge, 2024).

Experiment sessions lasted three hours. Participants were assigned randomly to a computer station. Before the experiment began, informed consent was obtained from all participants. We collected basic demographic data (age, gender, and field of study). Experiment instructions were read out by the experimenters and a demonstration of the trading interface was given, to ensure common knowledge of the rules of the experiment and the trading interface.[25] The instructions and demonstration of the trading interface lasted approximately one hour. The instructions included numerical examples of how period-earning and take-home pay were calculated. We ran two non-timed practice periods for participants to further familiarize themselves with the trading interface. We also ran four practice periods with the same duration as the experiment periods (namely, three minutes). When all clarifying questions were answered, a short break was introduced. The 16 experimental periods lasted about one hour, after which the participants were paid in cash, and the experiment concluded.

Roles were counterbalanced so that each participant was an investor in three periods and a seller in one period, while three out of four participants in any given period were investors. For the latter to obtain, we needed the number of participants for each session to be a multiple of four. We aimed to run each session with 20 participants, though some sessions ran with 16. Any participant in excess of 20 (or 16 for some sessions) was sent away with a show-up reward of £10. Participants were provided with an additional sheet of paper that they were instructed to use to record their role at the beginning of each period, to avoid confusion.

The participants knew their own role and their own endowment, but no information about other participants' roles and endowments was provided. This information structure is typically assumed in general equilibrium theory, and hence, a well-designed general equilibrium experiment should adhere to it. In principle, only trade prices are to be made available to participants, but the theory assumes that these are equilibrium prices, while the theory does not explain how they would come about. Therefore, general equilibrium experiments with continuous double-sided auctions usually allow access also to all standing (active) limit orders, i.e., the book of orders is made open. This approach has been successful to generate market equilibrium in past experiments, starting with Smith (1965).

Participants were informed at the beginning of the session that they would be paid for four randomly drawn periods out of the 16 experimental periods. The paid periods were drawn at the end of each experimental session, to ensure high incentives to perform well throughout all periods. The drawing ensured that all participants were paid thrice as an investor and once as a seller.

In total, we engaged 180 participants. They were on average 26 years old (median = 24, sd = 7.76, min = 18, max = 60), and evenly balanced across genders. Each participant was allowed to join only one session. They earned on average £35.50 (median = £36.00, sd = £2.75, min = £27.00, max = £42.00). An overview of demographic characteristics across treatments is provided in Table III.

[25] Instructions are reproduced in Appendix D.

TABLE III Participant demographics

Budget Treatment	Total sessions	Total participants	Earnings	Age	Gender			
					Male	Female	Other	Unknown
No Walras	5	88	£35.25	26	48	39	0	1
Walras	5	92	£35.75	26	42	45	3	2

For earnings and age, we report the mean. For gender, we report the number of participants identifying as male, female, other (non-binary or unlisted gender), or unknown (prefer not to say).

VI Results

We group our results into three subsections. The reader may want to inspect Appendix C to get an appreciation for the trading activity that emerged in our experiment. One of the features that stands out is the volatility of prices. We come back to this later.

VI.A Economic Requirements

We begin with Table IV, which presents the investor asset allocations at the close (end) of each period and the efficiency of asset sales.

TABLE IV Final allocations of assets per investor-period

Budget Treatment	Payoff Treatment	Efficiency	Triplet $\{H, M, L\}$	Pairs $\{M, L\}$	$\{H, L\}$	$\{H, M\}$	Singletons $\{L\}$	$\{M\}$	$\{H\}$	Zero Assets	Excess Units	Total
No Walras	C1	93.5%	32	94	127	143	21	32	26	24	29	528
	C2	89.9%	19	115	130	130	19	30	51	19	15	528
Walras	C1	89.4%	10	111	147	124	20	38	49	24	29	552
	C2	86.9%	13	143	119	98	13	27	98	17	24	552
Aggregate		89.9%	74	463	523	495	73	127	224	84	97	2,160

Efficiency is calculated as the percentage supply (asset units) that investors acquired from sellers by the end of a period. Unit of observation is "investor-period," that is, the final allocation of one investor at the end of one period. "Excess Units" refers to the number of investor-periods when the investor held at least one asset in excess of one unit (for which they were not compensated). Investor-periods when units are held in excess do not count towards the investor-periods of the triplet, pairs, or singletons. A more granular breakdown of final allocations including breakdown of excess demands is provided in Appendix B.

Prediction 1. Since sellers face zero marginal cost, they should be willing to sell assets at any price. Column 3 shows that roughly 90% of assets move from the sellers to the investors in all treatments. Turning this around, 10% of assets are mistakenly kept by sellers. Investors do occasionally make mistakes too: a small percentage of them (4.5% of investor-periods) end up buying more than one unit of the same asset, even though they gain nothing from the excess units (and pay a positive price for them). There is a reason not to consider over-buying as "mistakes," since it could reflect speculation, when investors purchase additional units in order to sell later for a profit.[26] In general, the data support Prediction 1.

Prediction 2. We observe that most investors end up with a pair of assets. This happens

[26] The idea that excess purchases reveal speculation is reinforced by the finding that almost an equal number of investor-periods (78, versus 97) managed to sell excess units before the end of the period. Therefore, we could claim that in total 78+97 = 175 attempts at speculation were tried, of which a bit less than half were successful. So, out of a total of 2,160 about 8% of investor-periods could be classified as involving speculation.

in approximately two thirds (1,481/2,160 or 68.6%) of the investor-periods. In equilibrium, this percentage should have been 100%, but the high volatility in prices occasionally allowed a small number of investors to purchase all three assets. Similarly, roughly one-fifth of investors bought only a single asset, and an even smaller fraction did not purchase any assets at all. Therefore, the evidence mostly supports Prediction 2.

Prediction 3. Among investors who purchased pairs of assets, the proportion choosing each possible pair is approximately the same. In total, investors obtained the pair $\{M, L\}$ in 463 investor-periods, $\{H, L\}$ in 523 investor-periods, and $\{H, M\}$ in 495 investor-periods. We formally test the equality of proportions by first estimating a multinomial logistic model with random intercepts at the investor and session level, and then testing whether the estimated proportions differ. We fail to reject the hypothesis that all three pairs are chosen with equal likelihood ($\chi^2(2) = 3.65,\ p = 0.1614,\ N = 1,481$). Thus, the data support Prediction 3.

We summarize the findings so far as follows.

Interim Conclusion 1. *The data support predictions 1 to 3.*

Prediction 4. Prices are expected to be sufficiently high so that investors cannot afford all three assets. Table V lists the frequencies with which participants could afford all three assets. We calculate the frequency as a fraction of total period duration that all three assets could be afforded. Duration is measured as either calendar time or trade time. Under the second method, one time unit corresponds to a trade. For prices, we take either the last traded prices[27] or the best standing sell price for the three assets.

We observe that participants could very rarely afford all three assets. This is even more acute when looking at the standing sell (ask) prices, which are expectedly higher than the trading prices.

TABLE V Fractions of time when asset triplet was affordable

Budget Treatment	Payoff Treatment	Last Traded Price Calendar Time	Last Traded Price Trade Time	Best Standing Sell Calendar Time	Best Standing Sell Trade Time
No Walras	C1	6.0%	6.8%	0.0%	0.1%
	C2	5.5%	6.8%	0.0%	0.0%
Walras	C1	1.3%	1.9%	0.0%	0.0%
	C2	4.6%	5.2%	0.1%	0.1%
Aggregate		4.3%	5.2%	0.0%	0.0%

Table presents fractions of time when the asset triplet $\{H, M, L\}$ was affordable. Time is measured as calendar time or as count of trades ("Trade Time"). Affordability is based on either transaction prices of most recent trades ("Last Traded Price") or from best sell offer in the book ("Best Standing Sell"). Affordability is only calculated from when all three assets have a price, that is, after each asset has traded at least once in the period.

However, Table VI shows that often prices were too high, so that one or more pairs became

[27] At any point in time, only one asset trades; prices for the non-traded assets were set equal to previous (last) trade prices, as is convention in financial economics.

unaffordable, contrary to our prediction. In the table, we only report results for affordability at traded prices. Time is measured as count of trades.

TABLE VI Fractions of time when one of the equilibrium asset pairs was not affordable

Budget Treatment	Payoff Treatment	Asset Pairs $\{M, L\}$	$\{H, L\}$	$\{H, M\}$	At Least One Pair
No Walras	C1	0.2%	3.5%	43.3%	43.3%
	C2	3.6%	26.2%	47.3%	49.8%
Walras	C1	0.1%	9.7%	69.7%	69.8%
	C2	3.6%	41.5%	69.4%	71.0%
Aggregate		1.8%	19.7%	57.1%	58.1%

Table presents affordability of pairs of assets. Affordability is defined as the fraction of time (measured as count of trades, i.e., trade time) in a period when the pair at the top of the column was not affordable, or when at least one pair was not affordable (rightmost column). Affordability is only calculated from when all three assets have a price, that is, after each asset has traded at least once in the period.

Prediction 5. In Appendix C we provide time series plots of the raw trade prices (Figure A1), and of the prices of pairs and the triplet of assets (Figure A2). Those graphs visualize the evolution of prices. The observation that prices are uniformly higher when the budget is higher (i.e., when WE exists) is plain obvious and does not need formal testing.

Interim Conclusion 2. *The data support Prediction 5. Support for Prediction 4 is qualified, however: while prices are sufficiently high so that all three assets could not be afforded together, between 1/2 and 2/3 of the time, not all pairs were affordable.*

Prediction 6. The last key economic prediction of our equilibrium is that the various equilibrium asset pairs ought to yield different utilities. To test this prediction, we estimate the long-run mean differences in total earnings from purchasing the pairs $\{H, M\}$ and $\{H, L\}$, as well as from purchasing $\{H, L\}$ and $\{M, L\}$.

We expect both of these differences to be significantly different from zero. We do not know whether the differences ought to be strictly *positive*. In the numerical example of Section III, they are. But we should allow for the possibility that, say, the first difference $(U_t(\{H, M\}) - U_t(\{H, L\}))$ is positive while the second one $(U_t(\{H, L\}) - U_t(\{M, L\}))$ is negative. This would happen if the pair $\{H, L\}$ generated lowest utility, while the pair $\{M, L\}$ generated medium utility. In that case, the expectation of the absolute value of the second difference in utility should be smaller than the expectation of the first difference.

The evolution of the two differences is estimated using a vector autoregression model (VAR). The unit of observation is a trade and one VAR is estimated per session–payoff treatment, concatenating the eight periods of the treatment.[28] Formally, we estimate the following VAR:

$$\begin{bmatrix} U_t(\{H,M\}) - U_t(\{H,L\}) \\ U_t(\{H,L\}) - U_t(\{M,L\}) \end{bmatrix} = \mu + \sum_{k=1}^{K} \Phi_k \begin{bmatrix} U_{t-k}(\{H,M\}) - U_{t-k}(\{H,L\}) \\ U_{t-k}(\{H,L\}) - U_{t-k}(\{M,L\}) \end{bmatrix} + \epsilon_t. \quad (5)$$

We then test whether the long-run expectations of the differences in values are significantly different from zero, i.e., whether elements of the vector $(I - \Phi_1 - \cdots - \Phi_K)^{-1}\mu$ are nonzero.[29] Inspection of the autocorrelations of the error terms invariably suggest that the most parsimonious VAR uses only one lag, i.e., $K = 1$ in Equation 5. Likewise, the estimates of the autoregression coefficient matrix (Φ_1) confirmed that the time series are stationary, i.e., that the earnings differences are mean-reverting. The results of the estimation are displayed in Table VII.

TABLE VII Estimated long-run earnings differences between asset pairs

Budget Treatment	Payoff Treatment	Session 1	2	3	4	5
No Walras	C1	71***	49***	73***	65***	84***
	C2	65***	59***	48***	67***	63***
Walras	C1	68***	43***	43***	66***	65***
	C2	45***	56***	62***	68***	67***

(a)

Budget treatment	Payoff treatment	Session 1	2	3	4	5
No Walras	C1	6	38***	47***	8	9*
	C2	14***	41***	57***	15***	-2*
Walras	C1	-19**	16**	15***	-3	7
	C2	13	37**	26***	-18***	-13***

(b)

(a) Earnings differences between $\{H,M\}$ and $\{H,L\}$, in 0.01. (b) Earnings differences between $\{H,L\}$ and $\{M,L\}$, in 0.01. Legend: ***: $p \leq 0.001$, **: $p \leq 0.01$, *: $p \leq 0.05$, all two-sided.

The results confirm that the long term expectations of earnings levels are generally significantly different across equilibrium assets pairs. The first difference $(U_t(\{H,M\}) - U_t(\{H,L\}))$ is uniformly positive and significant. That is, the rankings of the utility earned from buying $\{H,M\}$ and $\{H,L\}$ is as in the numerical example in Section III. As to the second difference $(U_t(\{H,L\}) - U_t(\{M,L\}))$, we observe that the long term expectation is significantly ($p \leq 0.05$) positive in the majority (11) of the 20 sessions, but it is significantly ($p \leq 0.05$) negative in 4 sessions. When negative, the magnitude (absolute value) of the long term expectation of the second difference is never more than 1/3 of that of the first difference. In the remaining

[28] We also estimated VARs for each period separately and then aggregated the results per session-treatment. This led to qualitatively the same inference, but because we computed standard errors from the cross-section of estimated long term means, we encountered a large loss in power.

[29] Standard errors of the estimated long term means are computed from the standard errors of the estimated coefficients using the Delta Method (see, e.g., Schervish (2012), Section 7.1.3).

5 sessions, the difference in the expected long term values of the pairs $\{H, L\}$ and $\{M, L\}$ is not significant. This is consistent with CCE only if reaching the utility level of these two pairs requires equal effort.

The utility differences of pairs $\{H, M\}$ and $\{H, L\}$ are also economically significant. For illustration, an estimated coefficient of 71 (Treatment: No Walras – C1) corresponds to earnings of £0.71 per period. The fact that higher compensation is provided for moving from the second-best to the first-best bundle, compared with the compensation for moving from the third-best to the second-best bundle, suggests that the cognitive effort cost associated with discovering better bundles is strictly convex.

TABLE VIII Average earnings implied from final period holdings, including cash

Budget Treatment	Payoff Treatment	Final Asset Holdings $\{M, L\}$	$\{H, L\}$	$\{H, M\}$
No Walras	C1	£4.19	£4.50	£5.50
	C2	£4.44	£4.90	£5.64
Walras	C1	£4.42	£4.49	£5.44
	C2	£4.57	£4.88	£5.64

Only participant-periods for which the participant held one of the three equilibrium pairs (shown at top of the columns) are included. In all cases (rows), the Kruskal-Wallis H test rejects equality of the distributions of earnings across asset pairs holdings ($p < 10^{-4}$).

Table VIII concludes the same in a different way. It reports the average earnings (in pounds, including cash) at the end of a period of investors who hold the asset pair shown at the top of each column. All investors started with the same cash, but spent part of it on acquiring the pair. In CCE, the average earnings across pairs should be significantly different. In the WE, when it exists, average earnings across pairs should be the same. We observe that average earnings for $\{H, M\}$ are the highest, followed by those for $\{H, L\}$, and lowest for $\{M, L\}$.

Because the distributions of earnings categorized by final asset pair holdings are distinctly non-gaussian, we resort to a non-parametric test to formally confirm the findings, namely, the Kruskal-Wallis H statistic. This tests whether the nature of final holdings has no effect on the earnings in the sense that there is no order among them. The test rejects when there is order: holding of one pair tends to give the lowest earnings, another gives the next lowest earnings, and so forth. Technically, it investigates stochastic dominance ordering. In all treatments (rows), we reject equality of earnings distributions at high significance levels ($p < 10^{-5}$). This is contrary to the WE, where earnings distributions should not exhibit stochastic dominance regardless of the pair a participant chooses to hold.[30]

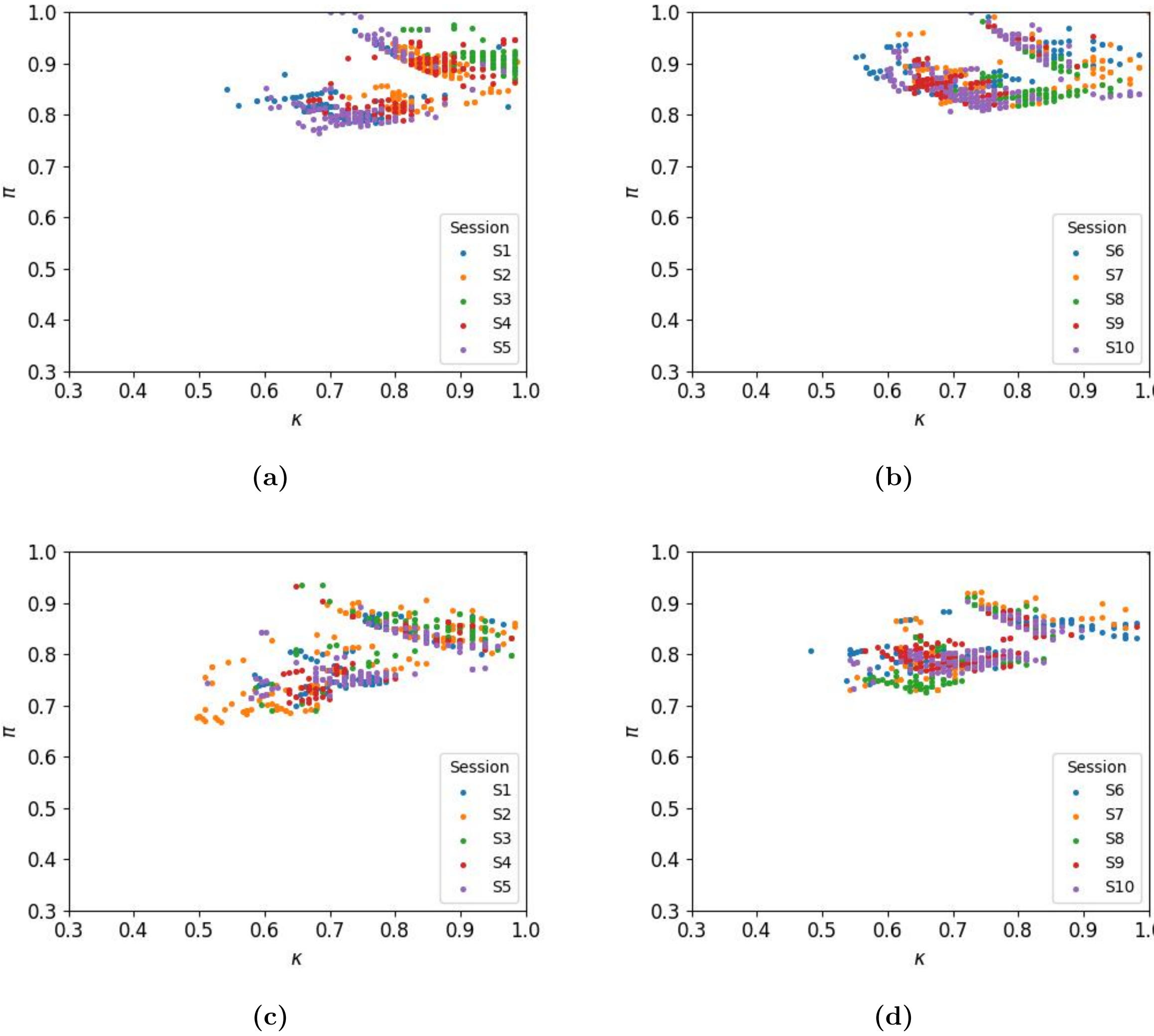
Interim Conclusion 3. *On balance, the data support Prediction 6.*

VI.B Computational Complexity

Prediction 7 and Prediction 8. The first Computational Requirement states that prices locate the budget problem inside the phase transition in (κ, π) space, and when not, that there

[30]Dunn's test with Bonferroni correction rejects pairwise equality across final asset holdings categories at $p < 0.01$, except when comparing $\{M, L\}$ and $\{H, L\}$ in treatment "Walras – C1."

is a strong tendency for the κ and π to mean-revert to the phase transition.



$\kappa = C/\sum_i p_i$; $\pi = \Upsilon^*/(C + \sum_i(\phi_i - p_i))$; C, ϕ_i, p_i, and Υ^* as in indicated sessions. Each dot corresponds to a trade. (a) Treatment "No Walras, C1" (Sessions S1 to S5). (b) Treatment "Walras, C1" (Sessions S6 to S10). (c) Treatment "No Walras, C2" (Sessions S1 to S5). (d) Treatment "Walras, C2" (Sessions S6 to S10).

FIGURE II Recorded locations of the budget problem in $\kappa \times \pi$ space.

For the reader to get perspective on the formal statistical analysis to follow, in Figure II we display the observed location of the budget problem in (κ, π) space after each trade.[31] The four subplots correspond to the four experimental treatments (see Table II). Sessions (S#) are color-coded. Realized κs and πs cover a large space, going beyond the phase transition (see Figure Ia). This is all right, as long as the evolution of κs and πs shows a strong tendency to mean-revert to the phase transition.

To formally validate mean-reversion to the phase transition region, we fit a VAR model to

[31] Since only one asset trades each time, we have to choose prices for the non-traded assets. As before, we take those to be the prices at which the assets last traded. Observations are only plotted after each asset has a last trade price, that is, after it has traded during the period.

the time series data on which Figure II is based. That is, we estimate the following model:

$$\begin{bmatrix} \kappa_t \\ \pi_t \end{bmatrix} = \mu + \sum_{k=1}^{K} \Phi_k \begin{bmatrix} \kappa_{t-k} \\ \pi_{t-k} \end{bmatrix} + \epsilon_t. \tag{6}$$

As we did with the differences in utilities from pairs of assets, we concatenate the data for all periods within a session–payoff treatment and estimate the parameters. We derive an estimate of the long term mean by computing the vector $(I - \Phi_1 - \cdots - \Phi_K)^{-1}\mu$. Inspection of the error autocorrelations suggests that the best parsimonious model uses only one lag of the observed time series. Estimates of the roots of the autoregression (not reported) prove strong mean reversion in the two time series, which confirms Prediction 8. The estimates of the long term means themselves are listed in Table IX.

TABLE IX Estimates of long term expectation of normalized cost κ and normalized profit π.

Budget Treatment	Payoff Treatment	(κ, π)				
		Session				
		1	2	3	4	5
No Walras	C1	(0.74, 0.84)	(0.85, 0.88)	(0.94, 0.93)	(0.82, 0.86)	(0.76, 0.84)
	C2	(0.76, 0.80)	(0.77, 0.82)	(0.85, 0.86)	(0.76, 0.80)	(0.77, 0.79)
Walras	C1	(0.76, 0.89)	(0.80, 0.89)	(0.82, 0.86)	(0.72, 0.88)	(0.74, 0.87)
	C2	(0.76, 0.80)	(0.78, 0.85)	(0.71, 0.80)	(0.72, 0.81)	(0.75, 0.81)

We subsequently compute estimates of the standard error of the long term mean of the VAR in Equation 6. From these, we calculate 95% confidence regions. Finally, we superimpose them on the phase transition region for each treatment separately. The phase transition regions are obtained analogously to the one depicted in Figure Ia.

Figure III shows substantial overlap between all 95% confidence ellipsoids and the phase transition. As such, the data provide convincing evidence that market forces push prices so that the budget problem is in the phase transition, making it most difficult for investors to determine whether they have reached the optimum. Prediction 7 is supported.

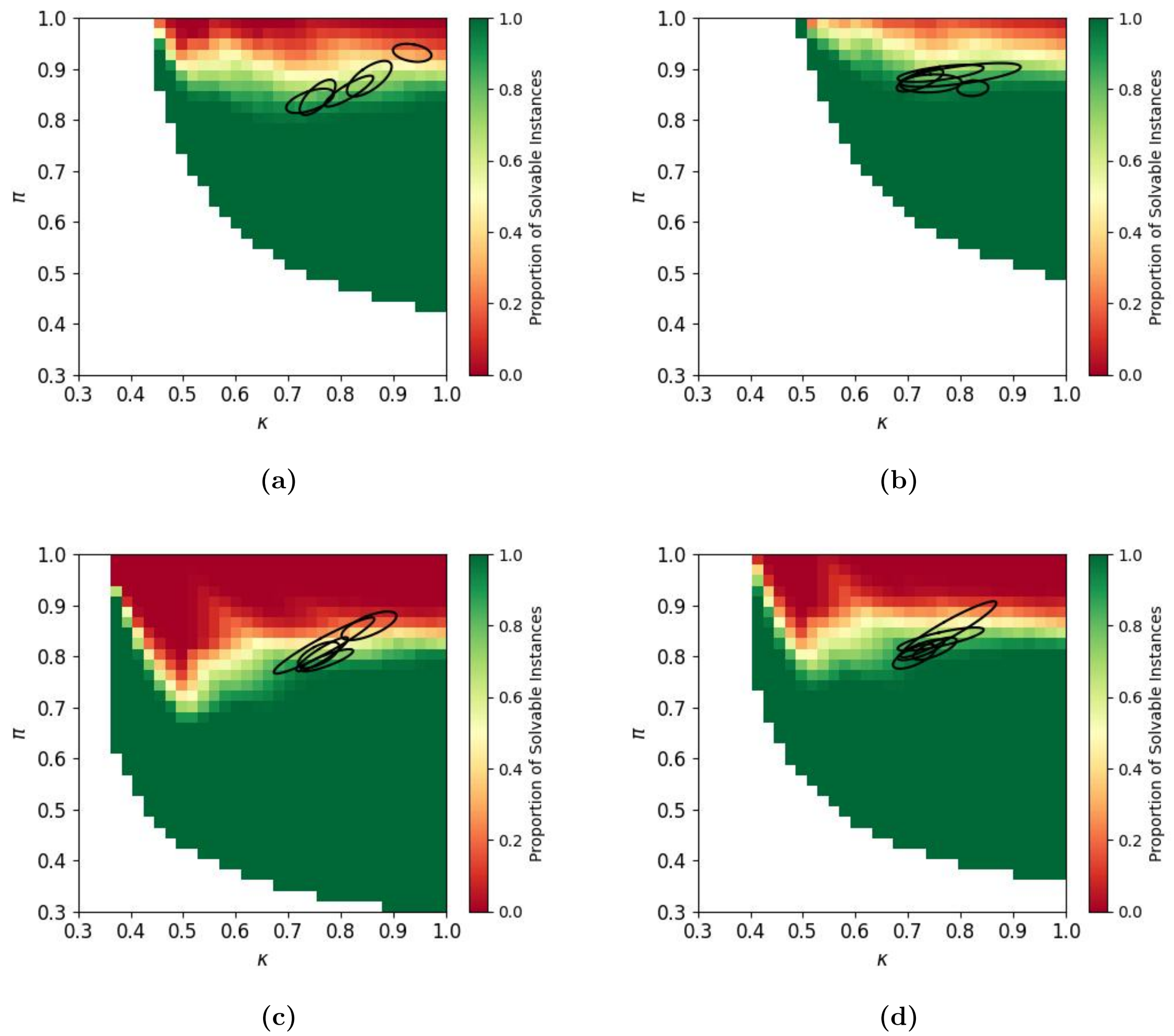
Prediction 9. We turn to the second Computational Requirement, which is that trade prices ensure that the budget problem has a unique optimum.

Figure IV superimposes the 95% confidence ellipsoids of the locations of the estimated long-term expectation of κs and πs on the color-coded map of mean number of optima (witnesses) of budget problems, constructed as in Figure Ic. We observe that the ellipsoids fully overlap with the yellow region, where 95% or more of the budget problems feature unique optima. Therefore, we have strong support for Prediction 9.

We take stock of the findings.

Interim Conclusion 4. *The data exhibit strong evidence in favor of Prediction 7, Prediction 8 and Prediction 9.*

Computational versus Economic Requirements. We observed strong evidence in favor of economic equilibrium in the sense that the vast majority, in about equal fractions, end up

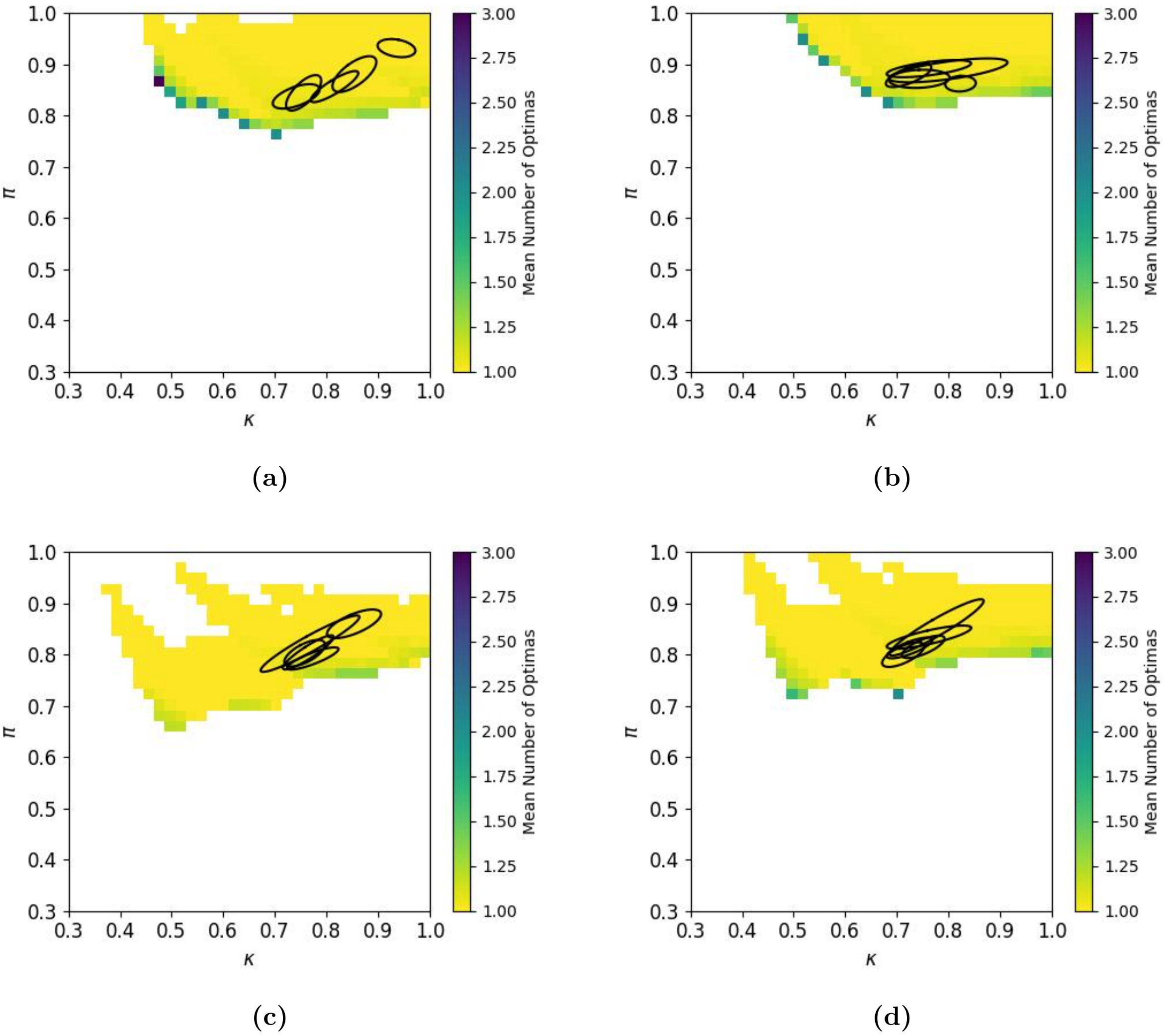


$\kappa = C/\sum_i p_i$; $\pi = \Upsilon/(C + \sum_i(\phi_i - p_i))$; C, ϕ_i as in indicated Treatment; p_i, and Υ varying; 95% confidence ellipsoids based on (6). (a) Treatment "No Walras, C1." (b) Treatment "Walras, C1." (c) Treatment "No Walras, C2." (d) Treatment "Walras, C2." Phase transitions are indicated in light green down to orange. The color-coded map in Panel (b) is the same as the one in Figure I.a.

FIGURE III 95% confidence ellipsoids of estimated asymptotic values for κ and π in each experimental session, arranged per treatment, relative to the phase transition map of the budget problem.

with one of the three equilibrium asset pairs. We also discovered that utilities accruing to these pairs were significantly and sizably different. All this is economic support for CCE. On the computational side, we confirmed that prices mostly forced the budget problem to reside in the phase transition and in the region where most instances have a unique optimum, and hence, in the region of highest difficulty. Yet prices often pushed the budget problem into the region where one or more equilibrium asset pairs are unaffordable: between 1/2 to 2/3 of the trades occurred at prices that made at least one equilibrium asset pair not affordable. That is, trade took places at prices that violated one of the key economic requirements, Prediction 4.

Figure V provides another way to visualize the issue. There, we plot the 95% confidence ellipsoids of κ and π on top of a map of the location of budget problems where all three equilibrium pairs are affordable, but not the triplet. The confidence ellipsoids tend to only



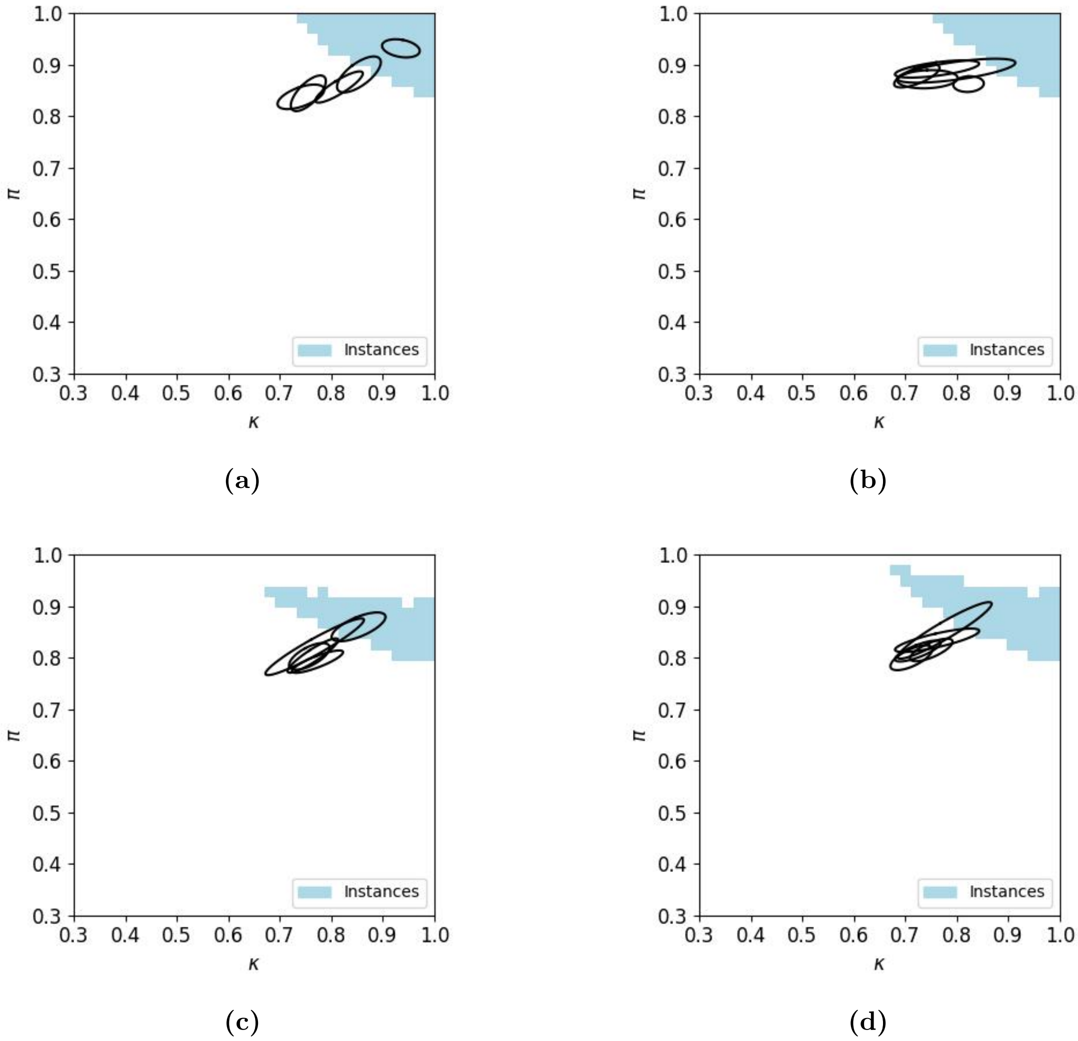
$\kappa = C/\sum_i p_i$; $\pi = \Upsilon^*/(C + \sum_i(\phi_i - p_i))$; C, ϕ_i as in indicated Treatment; p_i varying; Υ^* as resulting optimum; 95% confidence ellipsoids based on (6). (a) Treatment "No Walras, C1." (b) Treatment "Walras, C1." (c) Treatment "No Walras, C2." (d) Treatment "Walras, C2." Confidence ellipsoids overlay the mean number of witnesses. The colored regions in Panel (b) are the same as in Figure I.c.

FIGURE IV 95% confidence ellipsoids of estimated asymptotic values for κ and π for each experimental session, arranged per treatment, relative to mean number of witnesses (optima) in instances of budget problem.

partially overlap with the region where all asset pairs can be afforded, and sometimes not at all, suggesting that prices do not always tend to revert to levels that make all three pairs affordable.

We conjecture that markets ensure equilibrium by making at least one of the equilibrium pairs unaffordable during substantial periods of time. This is necessarily the most expensive pair. In our design, this happens to also be the pair that generates maximal utility. Not only is it the most difficult pair to recognize as generating maximal utility, but it is also often not available for sale within investors' budget constraints. Investors have to be attentive, track prices and submit orders when the best pair does become affordable. As a result, investors spend cognitive effort not only to find optimal choices given expected trade prices, but also to pay attention to order flow so that desired allocations can be implemented.

Interim Conclusion 5. *Markets appear to make it harder to attain maximum utility not only*



$\kappa = C/\sum_i p_i$; $\pi = \Upsilon^*/(C + \sum_i(\phi_i - p_i))$; C, ϕ_i as in indicated Treatment; p_i varying; Υ^* as resulting optimum; 95% confidence ellipsoids based on (6). (a) Treatment "No Walras, C1." (b) Treatment "Walras, C1." (c) Treatment "No Walras, C2." (d) Treatment "Walras, C2." The blue region in Panel (b) is the same as the one in Figure I.d.

FIGURE V 95% confidence ellipsoids of estimated asymptotic values for κ and π for each experimental session, arranged per treatment, relative to the region where prices are such that all asset pairs can be afforded (blue area).

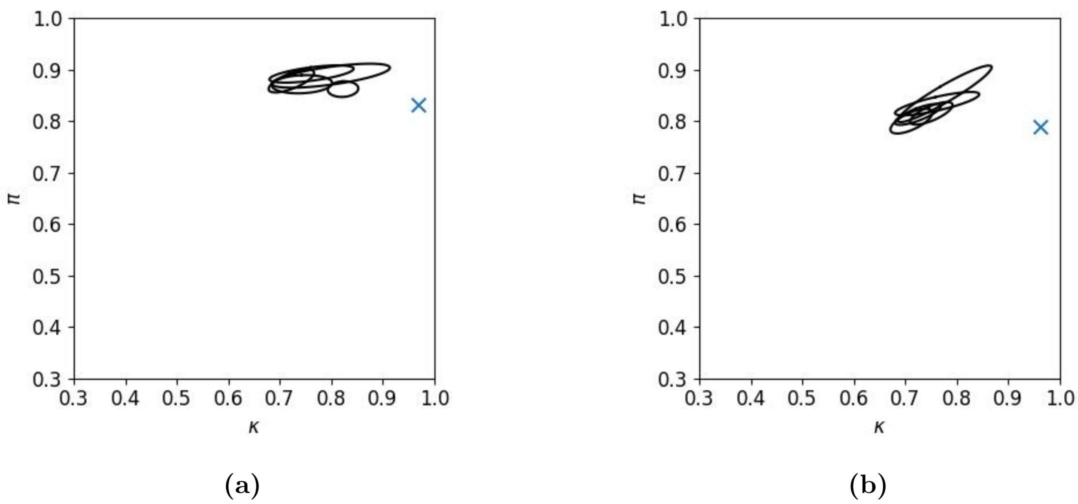
because of computational complexity but also by intermittently making the unique best asset pair unaffordable.

VI.C Walrasian Equilibrium

The final objective of our paper is to investigate the empirical support for WE in the Walras Treatment, where this equilibrium exists.

We already rejected WE because we verified Prediction 6: we found that earnings associated with the three equilibrium asset pairs are statistically and sizably different, unlike in WE, where they necessarily have to be equal because everyone is assumed to fully optimize, and hence, whatever choices are made in equilibrium must earn the same.

At the same time, WE implies that the budget problem investors face must have multiple



$\kappa = C/\sum_i p_i$; $\pi = \Upsilon^*/(C + \sum_i(\phi_i - p_i))$; C, ϕ_i as in indicated Treatment; p_i and Υ^* as in WE (blue cross); 95% confidence ellipsoids based on (6). (a) Treatment "Walras, C1." (b) Treatment "Walras, C2."

FIGURE VI 95% confidence ellipsoids of estimated asymptotic values for κ and π for each experimental session, per treatment, relative to location of WE (blue cross)

(three) optima. However, we discovered substantial overlap between the location in (κ, π) space of the budget problem instance that prices mean-revert to in the long run and the locations of budget problem instances that have mostly unique optima; see Figures IVb and IVd. That is, prices are such that the budget problem rarely if never exhibits more than one optimum.

In fact, we can locate exactly where the WE resides in (κ, π) space. See Figure VI. The figure reveals that the WE is never inside any of the 95% confidence ellipsoids constructed from the experimental sessions. This is conclusive evidence against WE.

We can also directly test whether the trade prices are consistent with WE. We remind the reader that WE prices are $(p_L, p_M, p_H) = (0.10, 1.30, 1.90)$ for payoff configuration C1, and $(p_L, p_M, p_H) = (0.10, 0.90, 1.70)$ for payoff configuration C2. Figure A1 in Appendix C displays the full time series of trade prices and compares them directly with the WE prices. For virtually all observations, the price levels of the lowest-value asset, L, are five to ten times higher than the equilibrium level of 0.10.

We re-emphasize that prices are high relative to expectations because of another reason: the sellers of the assets pay zero marginal cost for the goods. In past experiments, this has consistently caused low trade prices, even prices below WE (Smith and Williams, 1982; Rasooly, 2022). In contrast, prices in CCE *have to be* high, so that the budget problem becomes sufficiently complicated. This suggests that computational complexity more than offsets the (opposite) force caused by the large rents that accrue to the supply side.

Interim Conclusion 6. *When it exists, WE is overwhelmingly rejected.*

VII Overall Conclusions

When goods are indivisible, we find that markets set prices such that the budget problem becomes computationally complex. This complexity plays a central role in enabling market equilibration. When the budget problem is difficult, agents with homogeneous preferences select different bundles because discovering the optimal allocation requires non-trivial cognitive effort. If not everyone is able or willing to spend maximal effort to find the optimum budget allocation, demand becomes heterogeneous. This allows markets to clear. Those who expend more effort and identify better bundles are compensated with higher utility. The resulting Complexity Compensating Equilibrium (CCE) offers a framework in which equilibrium may exist even when WE does not.

In our experiment, computational complexity requires prices that are markedly different from those predicted by WE (provided the latter exists). Prices located the budget problem in the region where the most difficult instances of the budget problem are located, i.e., in the "phase transition," and in the region where most instances have a unique optimum. Prices were high, in spite of the downward pressure from suppliers who had zero marginal cost for selling the assets. The data further supported CCE in that the various equilibrium choices implied significantly different utility levels. When WE exists, all equilibrium bundles must earn the same, and the budget problem must have multiple optima.

We did discover an unexpected channel through which markets further raised the cognitive costs of reaching optimum utility levels, namely, price volatility. Prices moved sufficiently so that at times the best goods (assets) bundle was not affordable, forcing those that expected to be able to buy this bundle to be more attentive and wait till prices reverted to levels where they could afford it. The market was able to exploit this channel to equilibrate only because we happened to have chosen design parameters so that the optimal bundle consisted of the assets with the highest payoffs. Future work should look at the effects of eliminating this channel, by choosing parameters where the optimal solution contains assets with lesser payoffs.

Future work should also aim at understanding the role of computational complexity (and other features that raise cognitive effort) in the presence of preference (payoff) heterogeneity. In this paper, because of homogeneity, the market can only use prices to ensure equilibration. With heterogeneity, the market can also exploit payoff diversity. We would still propose that the market chooses prices that puts the budget problem in the phase transition. Because of payoff heterogeneity, however, the location of the phase transition changes. We conjecture that the phase transition will become more like that of the traditional KP. There, the phase transition consists of normalized profits (π) that are close to, and generally slightly above, normalized capacities (κ) (Yadav et al., 2020). This would then provide generic CCE price predictions that can be tested even in the field.

The broader implication of our findings is that cognitive constraints – often viewed as biases or limitations – may, in fact, be integral to the equilibration of decentralized markets. When indivisibilities make equilibration difficult or impossible because of preference homogeneity, the computational complexity of the budget problem – an NP-hard problem because of the indivisibilities – can generate sufficient demand heterogeneity for markets to equilibrate.

We close with a philosophical note. When they are repeatedly exposed to solve complex problems because markets exploit computational complexity in order to equilibrate, humans may lose confidence in their choices. A broader sense of unease may ensue. This leads us to conjecture that the rise in anxiety in modern societies (Goodwin et al., 2020) may have had its roots in the increasing reliance on market-based mechanisms to allocate resources. Indeed, markets have been used to facilitate transport (airline deregulation, railway privatization), medical care (healthcare marketplaces), education (higher education loan programs), and climate change (carbon markets), among others. Our findings raise the possibility that the computational complexity that would explain the success of the market mechanism in the real world may simultaneously have caused the growing push-back against market-based solutions of society's allocation problems.

References

Arora, S. and Barak, B. (2009). *Computational complexity: a modern approach*. Cambridge University Press.

Asparouhova, E. (2006). Competition in lending: Theory and experiments. *Review of Finance*, 10(2):189–219.

Asparouhova, E., Bossaerts, P., Roy, N., and Zame, W. (2016). 'lucas' in the laboratory. *Journal of Finance*, 71:2727–2780.

Babaioff, M., Nisan, N., and Talgam-Cohen, I. (2021). Competitive equilibrium with indivisible goods and generic budgets. *Mathematics of Operations Research*, 46(1):382–403.

Baldwin, E., Jagadeesan, R., Klemperer, P., and Teytelboym, A. (2023). The equilibrium existence duality. *Journal of Political Economy*, 131(6):1440–1476.

Baldwin, E. and Klemperer, P. (2019). Understanding preferences:"demand types", and the existence of equilibrium with indivisibilities. *Econometrica*, 87(3):867–932.

Bossaerts, P. (2025). Decision-making when computational complexity drives uncertainty. *Available at SSRN 5128250*.

Bossaerts, P., Bowman, E., Fattinger, F., Huang, H., Lee, M., Murawski, C., Suthakar, A., Tang, S., and Yadav, N. (2024). Resource allocation, computational complexity, and market design. *Journal of Behavioral and Experimental Finance*, 42:100906.

Bossaerts, P., Plott, C., and Zame, W. R. (2007). Prices and portfolio choices in financial markets: Theory, econometrics, experiments. *Econometrica*, 75(4):993–1038.

Bossaerts, P. and Schultz, W. (2025). The neuroeconomics of simple and complex choice. *Annual Reviews in Economics*, in print.

Broome, J. (1972). Approximate equilibrium in economies with indivisible commodities. *Journal of Economic Theory*, 5(2):224–249.

Budish, E. (2011). The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy*, 119(6):1061–1103.

Budish, E., Che, Y.-K., Kojima, F., and Milgrom, P. (2013). Designing random allocation mechanisms: Theory and applications. *American economic review*, 103(2):585–623.

Casella, A., Llorente-Saguer, A., and Palfrey, T. R. (2012). Competitive equilibrium in markets for votes. *Journal of Political Economy*, 120(4):593–658.

Cheeseman, P. (1991). Where the really hard problems are. In *International Joint Conference on Artificial Intelligence*.

Choi, M., Dai, A. Y., and Kim, K. (2018). Consumer search and price competition. *Econometrica*, 86(4):1257–1281.

Danilov, V., Koshevoy, G., and Murota, K. (2001). Discrete convexity and equilibria in economies with indivisible goods and money. *Mathematical Social Sciences*, 41(3):251–273.

Dierker, E. (1971). Equilibrium analysis of exchange economies with indivisible commodities. *Econometrica: Journal of the Econometric Society*, pages 997–1008.

Florig, M. and Rivera, J. (2017). Existence of a competitive equilibrium when all goods are indivisible. *Journal of Mathematical Economics*, 72:145–153.

Franco, J. P., Bossaerts, P., and Murawski, C. (2024). The neural dynamics associated with computational complexity. *PLOS Computational Biology*, 20(9):e1012447.

Franco, J. P., Yadav, N., Bossaerts, P., and Murawski, C. (2021). Generic properties of a computational task predict human effort and performance. *Journal of Mathematical Psychology*, 104:102592.

Gent, I. P. and Walsh, T. (1996). The tsp phase transition. *Artificial Intelligence*, 88(1-2):349–358.

Gent, I. P. and Walsh, T. (1998). Analysis of heuristics for number partitioning. *Computational Intelligence*, 14(3):430–451.

Gilboa, I., Postlewaite, A., and Schmeidler, D. (2021). The complexity of the consumer problem. *Research in Economics*, 75(1):96–103.

Goodwin, R. D., Weinberger, A. H., Kim, J. H., Wu, M., and Galea, S. (2020). Trends in anxiety among adults in the united states, 2008–2018: Rapid increases among young adults. *Journal of psychiatric research*, 130:441–446.

Gul, F. and Pesendorfer, W. (2025). Pseudo lindahl equilibrium as a collective choice rule. *Review of Economic Studies*, page rdaf043.

Gul, F., Pesendorfer, W., and Zhang, M. (2024). Efficient allocation of indivisible goods in pseudomarkets with constraints. *Journal of Political Economy*, 132(11):3708–3736.

Gul, F. and Stacchetti, E. (1999). Walrasian equilibrium with gross substitutes. *Journal of Economic theory*, 87(1):95–124.

Hayek, F. A. (1945). The use of knowledge in society. *The American Economic Review*, 35(4):519–530.

Henry, C. (1970). Indivisibilités dans une économie d'échanges. *Econometrica*, pages 542–558.

Hong, T. and Stauffer, W. R. (2023). Computational complexity drives sustained deliberation. *Nature Neuroscience*, pages 1–8.

Judd, K. L. (2001). Computation and economic theory: Introduction. *Economic Theory*, pages 1–6.

Kelso Jr, A. S. and Crawford, V. P. (1982). Job matching, coalition formation, and gross substitutes. *Econometrica: Journal of the Econometric Society*, pages 1483–1504.

Meloso, D., Copic, J., and Bossaerts, P. (2009). Promoting intellectual discovery: Patents versus markets. *Science*, 323(5919):1335–1339.

Mitchell, D., Selman, B., and Levesque, H. (1992). Hard and easy distributions of sat problems. In *Proceedings of the tenth national conference on Artificial intelligence*, pages 459–465.

Murawski, C. and Bossaerts, P. (2016). How humans solve complex problems: The case of the knapsack problem. *Scientific Reports*, 6:34851.

Nguyen, T. and Vohra, R. (2024). (near-) substitute preferences and equilibria with indivisibilities. *Journal of Political Economy*, 132(12):4122–4154.

Plott, C. R. and Sunder, S. (1982). Efficiency of experimental security markets with insider information: An application of rational-expectations models. *Journal of political economy*, 90(4):663–698.

Plott, C. R. and Sunder, S. (1988). Rational expectations and the aggregation of diverse information in laboratory security markets. *Econometrica: Journal of the Econometric Society*, pages 1085–1118.

Rasooly, I. (2022). Competitive equilibrium and the double auction. Technical report, Paris School of Economics.

Rostek, M. and Yoder, N. (2020). Matching with complementary contracts. *Econometrica*, 88(5):1793–1827.

Rothschild, M. and Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, pages 629–649.

Rust, J. P. (1996). Dealing with the complexity of economic calculations. *Available at SSRN 40780*.

Schervish, M. J. (2012). *Theory of statistics*. Springer Science & Business Media.

Smith, V. L. (1965). Experimental auction markets and the walrasian hypothesis. *The Journal of Political Economy*, 73(4):387–393.

Smith, V. L., Suchanek, G. L., and Williams, A. W. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica: Journal of the Econometric Society*, 56(5):1119–1151.

Smith, V. L. and Williams, A. W. (1982). The effects of rent asymmetries in experimental auction markets. *Journal of Economic Behavior & Organization*, 3(1):99–116.

Sun, N. and Yang, Z. (2006). Equilibria and indivisibilities: Gross substitutes and complements. *Econometrica*, 74(5):1385–1402.

Svensson, L.-G. (1984). Competitive equilibria with indivisible goods. *Zeitschrift für Nationalökonomie/Journal of Economics*, 44(4):373–386.

University of Cambridge (2024). Director of Research, Faculty of Economics, Ethics Approval. Reference number: UCAM-FoE-24-02, Approved on: 2 May 2024.

Yadav, N., Murawski, C., Sardina, S., and Bossaerts, P. (2020). Is hardness inherent in computational problems? Performance of human and electronic computers on random instances of the 0-1 knapsack problem. In *ECAI 2020*, pages 498–505. IOS Press.

Appendix A Calculations for Walrasian equilibrium existence

Here we provide a general formula for existence of Walrasian equilibrium in an economy with three assets and two agents as described in Section III. Denote with ϕ_L, ϕ_M, ϕ_H the payoffs of the three assets and p_L, p_M, p_H the corresponding prices. Denote the budget of each agent with C. By assumption, $\phi_L \leq \phi_M \leq \phi_H$. Assume unit demand and linear utility so that $U(x_1, x_2, x_3) = x_1(\phi_1 - p_1) + x_2(\phi_2 - p_2) + x_3(\phi_3 - p_3) + C$, where $x_i \in \{0, 1\}$.

All prices need to be positive and smaller than or equal to the payoff, otherwise no one would buy the assets. This gives us the following inequalities:

$$
\begin{aligned}
0 \leq p_L &\leq \phi_L \\
0 \leq p_M &\leq \phi_M \\
0 \leq p_H &\leq \phi_H
\end{aligned}
$$

All equilibrium bundles of assets need to lead to same utility. For markets to clear, the bundles need to be $\{L, M\}$, $\{L, H\}$ and $\{M, H\}$. Indifference between the equilibrium bundles would require:

$$
\begin{aligned}
U(\{L, M\}) = U(\{M, H\}) &\Rightarrow p_H = (\phi_H - \phi_L) + p_L \\
U(\{M, H\}) = U(\{H, L\}) &\Rightarrow p_M = (\phi_M - \phi_L) + p_L \\
U(\{H, L\}) = U(\{L, M\}) &\Rightarrow p_H = (\phi_H - \phi_M) + p_M
\end{aligned}
$$

Combining the equalities and inequalities, we get:

$$
\begin{aligned}
0 \leq p_L &\leq \phi_L \\
\phi_M - \phi_L \leq p_M &\leq \phi_M \\
\phi_H - \phi_L \leq p_H &\leq \phi_H
\end{aligned}
$$

The inequalities above provide a range of prices within which markets can clear. We additionally need to ensure that prices are such that all pairs of assets are affordable. To ensure affordability, we need:

$$
\begin{aligned}
p_L + p_M \leq C \Rightarrow p_L + (\phi_M - \phi_L) + p_L \leq C &\Rightarrow p_L \leq \frac{C - (\phi_M - \phi_L)}{2} \\
p_M + p_H \leq C \Rightarrow p_M + (\phi_H - \phi_M) + p_M \leq C &\Rightarrow p_M \leq \frac{C - (\phi_H - \phi_M)}{2} \\
p_H + p_L \leq C \Rightarrow p_H - (\phi_H - \phi_L) + p_H \leq C &\Rightarrow p_H \leq \frac{C + (\phi_H - \phi_L)}{2}
\end{aligned}
$$

Combining the last two sets of inequalities, we get the final restrictions on prices.

$$0 \leq p_L \leq \min\left\{\phi_L, \frac{C - (\phi_M - \phi_L)}{2}\right\}$$
$$\phi_M - \phi_L \leq p_M \leq \min\left\{\phi_M, \frac{C - (\phi_H - \phi_M)}{2}\right\}$$
$$\phi_H - \phi_L \leq p_H \leq \min\left\{\phi_H, \frac{C + (\phi_H - \phi_L)}{2}\right\}$$

The inequalities are satisfied if

$$C \geq (\phi_M - \phi_L)$$
$$C \geq (\phi_M - \phi_L) + (\phi_H - \phi_L)$$
$$C \geq \phi_H - \phi_L$$

Given that we assume $\phi_H \geq \phi_M \geq \phi_L$, a Walrasian equilibrium exists if

$$C \geq (\phi_M - \phi_L) + (\phi_H - \phi_L)$$

In the motivating example of Section III, payoffs are $(\phi_L, \phi_M, \phi_H) = (1200, 2400, 3000)$ and the minimum budget for equilibrium existence is $C = 3000$. For this budget, the price inequalities collapse to a unique price vector of $(p_L, p_M, p_H) = (0, 1200, 1800)$. For any budget $C > 3000$, a multiplicity of equilibria exists of the form $(p_L, p_M, p_H) = (p, 1200 + p, 1800 + p)$, where p can be any allowable number satisfying the inequality for asset L.

In our experimental sessions, we used two payoff configurations. Configuration C1 was identical to the one analyzed above, with quantities divided by 1,000. So, payoffs were $(\phi_L, \phi_M, \phi_H) = (1.20, 2.40, 3.00)$. There, the minimum budget for equilibrium existence is $C = 3.00$ and in that case equilibrium prices are $(p_L, p_M, p_H) = (0.00, 1.20, 1.80)$. For configuration C2, the payoffs were $(\phi_L, \phi_M, \phi_H) = (1.60, 2.40, 3.20)$. The minimum budget for equilibrium existence is $C = 2.40$ and in that case equilibrium prices are $(p_L, p_M, p_H) = (0.00, 0.80, 1.60)$.

Appendix B More details on asset choices in the experiment

In this appendix, we provide a more detailed breakdown of choices in our experiment.

We first look at the seller side of the markets. At the beginning of each period, each seller is endowed with two units of each asset, so six assets in total. Table A1 shows the distribution of sold and unsold units across treatments, using both seller-periods and the actual number of assets as the unit of observation. We see that the vast majority of sellers successfully sold all their assets in all periods (519/720 seller-periods), and generally almost all units were sold (3,882/4,320). Efficiency is calculated as the percentage of units sold as shown in the last three columns of the table.

TABLE A1 Distribution of sold and unsold asset units across treatments

Budget	Payoff	Asset units sold per seller-period								Number of Asset Units		
Treatment	Treatment	6	5	4	3	2	1	0	Total	Sold	Unsold	Total
Low	C1	143	13	10	6	2	2	-	176	987	69	1,056
	C2	120	28	16	6	1	5	-	176	949	107	1,056
High	C1	129	26	11	9	4	4	1	184	987	117	1,104
	C2	127	23	11	9	3	5	6	184	959	145	1,104
Aggregate		519	90	48	30	10	16	7	720	3,882	438	4,320

We move on to the investor side of the markets. Table A2a presents the full distribution of final asset holdings across all investor-periods. We observe that the majority (68.6%) of investors end up with a pair of assets. In rare occasions, investors manage to buy all three assets (3.4%). This leaves less supply for other investors, resulting in 19.6% of investors buying a single asset and 3.9% of investors not buying any asset at all.

Remarkably, in rare cases (4.5%) investors end up holding more than one unit of an asset. This can sometimes be by mistake, but it is also possible that some investors are buying excess units of an asset when the price is low, speculating that they will be able to sell later when the price is high (this was allowed). Since they did not successfully resell the excess units, those investors will not be paid a liquidating dividend for their additional units of holdings. Thus, such excess demand decreases welfare as it prevents other investors from increasing their utility.

For equilibrium to obtain, investors need to end up with pairs of assets in equal proportions. The upper panel shows a roughly similar proportion of investors purchasing each asset pair, suggesting that equilibrium allocations are reached. There is some mismatch between treatments with more investors purchasing asset pair $\{M, L\}$ in the high-budget treatment and more investors purchasing pair $\{H, M\}$ in the low-budget treatment. Table A2b shows the percentage of investor-periods purchasing each asset, aggregated over all possible bundles. For example in 335 out of 952 of investor-periods in the first treatment (column 3), investors purchased asset H (as a single asset or together with other assets). The proportions of investors purchasing each individual asset is statistically indistinguishable from equal proportions for all treatments.[32]

[32]Using a multinomial logistic model with random intercepts at the investor and session level and estimating it separately per treatment, we get the following results. For No Walras-C1 treatment, $\mathcal{X}^2(1) = 2.170, p = 0.338, N = 952$. For No Walras-C2 treatment, $\mathcal{X}^2(1) = 2.630, p = 0.268, N = 933$. For Walras-C1 treatment,

TABLE A2 Distribution of assets purchased per investor-period

Category	Assets	No Walras-C1		No Walras-C2		Walras-C1		Walras-C2		Aggregate	
		Count	%	Count	%	Count	%	Count	%	Count	%
Triplet	{H,M,L}	32	6.1%	19	3.6%	10	1.8%	13	2.4%	74	3.4%
Pairs	{M,L}	94	17.8%	115	21.8%	111	20.1%	143	25.9%	463	21.4%
	{H,L}	127	24.1%	130	24.6%	147	26.6%	119	21.6%	523	24.2%
	{H,M}	143	27.1%	130	24.6%	124	22.5%	98	17.8%	495	22.9%
	Subtotal	364	68.9%	375	71.0%	382	69.2%	360	65.2%	1481	68.6%
Singles	{L}	21	4.0%	19	3.6%	20	3.6%	13	2.4%	73	3.4%
	{M}	32	6.1%	30	5.7%	38	6.9%	27	4.9%	127	5.9%
	{H}	26	4.9%	51	9.7%	49	8.9%	98	17.8%	224	10.4%
	Subtotal	79	15.0%	100	18.9%	107	19.4%	138	25.0%	424	19.6%
Excess	{H,H}	1	0.2%	0	0.0%	0	0.0%	1	0.2%	2	0.1%
	{H,M,M}	0	0.0%	1	0.2%	0	0.0%	0	0.0%	1	0.0%
	{H,L,L}	6	1.1%	3	0.6%	4	0.7%	0	0.0%	13	0.6%
	{H,L,L,L}	0	0.0%	0	0.0%	3	0.5%	0	0.0%	3	0.1%
	{M,M}	4	0.8%	1	0.2%	9	1.6%	7	1.3%	21	1.0%
	{M,M,L}	0	0.0%	2	0.4%	3	0.5%	5	0.9%	10	0.5%
	{M,M,L,L}	0	0.0%	1	0.2%	0	0.0%	0	0.0%	1	0.0%
	{M,L,L}	11	2.1%	4	0.8%	3	0.5%	9	1.6%	27	1.3%
	{M,L,L,L}	2	0.4%	0	0.0%	2	0.4%	0	0.0%	4	0.2%
	{M,L,L,L,L}	1	0.2%	0	0.0%	1	0.2%	0	0.0%	2	0.1%
	{L,L}	3	0.6%	3	0.6%	2	0.4%	2	0.4%	10	0.5%
	{L,L,L}	0	0.0%	0	0.0%	1	0.2%	0	0.0%	1	0.0%
	{L,L,L,L}	1	0.2%	0	0.0%	0	0.0%	0	0.0%	1	0.0%
	{L,L,L,L,L}	0	0.0%	0	0.0%	1	0.2%	0	0.0%	1	0.0%
	Subtotal	29	5.5%	15	2.8%	29	5.3%	24	4.3%	97	4.5%
None	∅	24	4.5%	19	3.6%	24	4.3%	17	3.1%	84	3.9%
Total		528	100%	528	100%	552	100%	552	100%	2,160	100%

(a) Investor end-of-period holdings

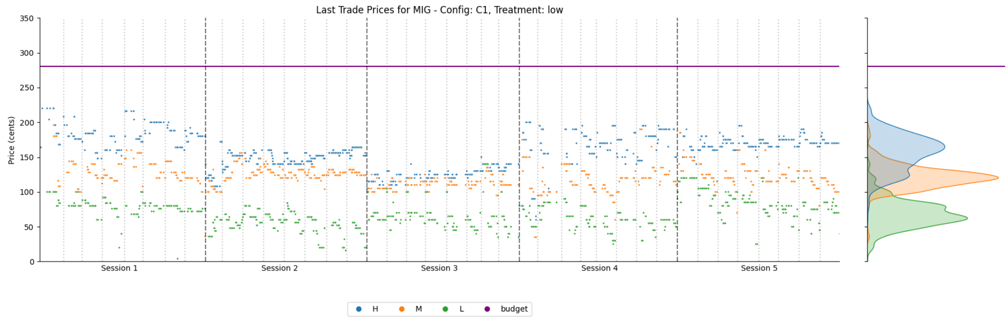
Category	Assets	No Walras-C1		No Walras-C2		Walras-C1		Walras-C2		Aggregate	
		Count	%	Count	%	Count	%	Count	%	Count	%
Holds asset	{H*}	335	35.2%	334	35.8%	337	35.6%	329	35.2%	1,335	35.4%
	{M*}	319	33.5%	303	32.5%	301	31.8%	302	32.3%	1,225	32.5%
	{L*}	298	31.3%	296	31.7%	308	32.6%	304	32.5%	1,206	32.0%
Total		952	100.0%	933	100.0%	946	100.0%	935	100.0%	3,766	100.0%

(b) Distribution of individual assets across investors

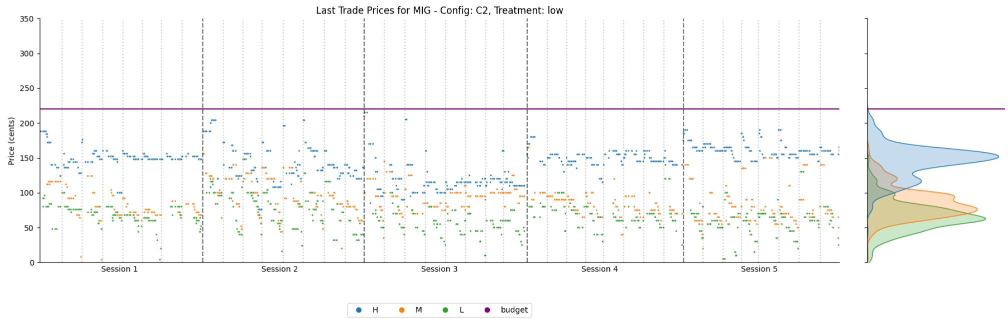
$\mathcal{X}^2(1) = 2.311, p = 0.315, N = 946$. For Walras-C2 treatment, $\mathcal{X}^2(1) = 1.452, p = 0.484, N = 935$. For aggregate data, $\mathcal{X}^2(1) = 7.728, p = 0.021, N = 3,766$.

Appendix C More details on asset prices in the experiment

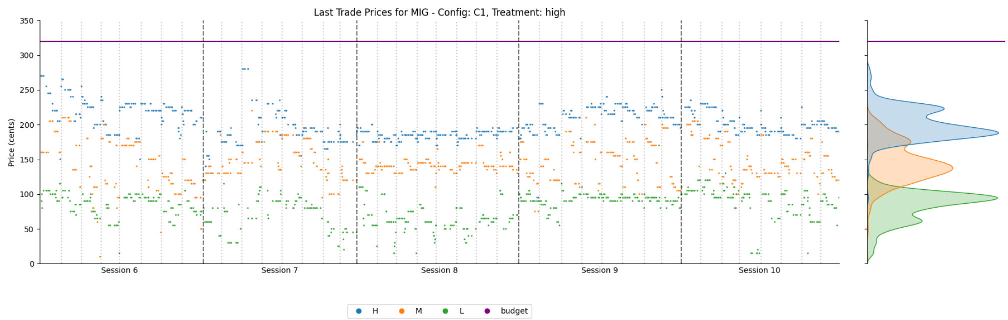
Here, we provide more details of the evolution of prices. We do so with two figures. Figure A1 plots the raw time series of the prices of the three assets for each treatment. We note the following observations: (i) prices are very volatile, increasing the complexity the investors are facing, (ii) prices are higher when investors have a high budget (i.e., when WE exists), (iii) prices of asset L are 5-10 times higher than £0.10, suggesting large deviations from Walrasian equilibrium. Figure A2 plots the prices of pairs and the triplet of assets for all treatments. The reader can easily verify that investors could rarely afford all assets together. This means that their budget constraint was binding and that they were solving a computationally hard decision problem.



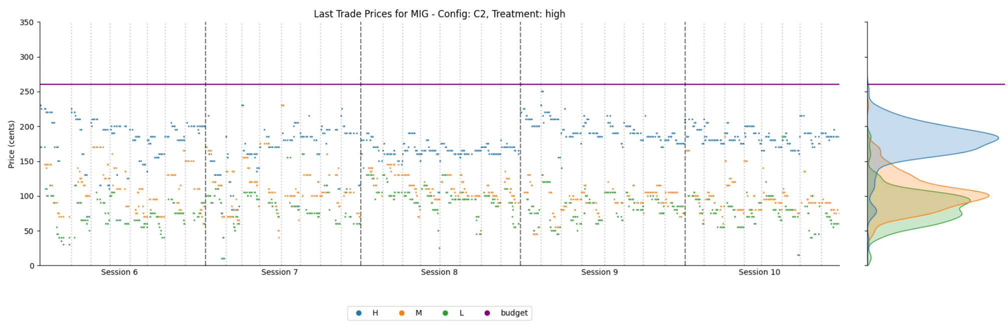
(a) No Walras, Payoff Configuration C1



(b) No Walras, Payoff Configuration C2



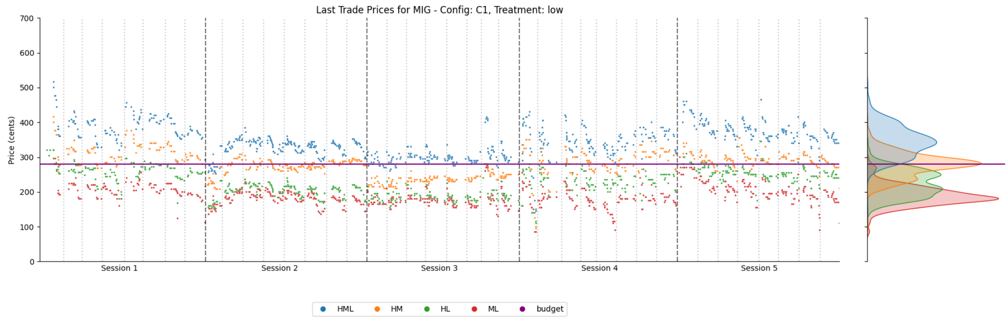
(c) Walras, Payoff Configuration C1



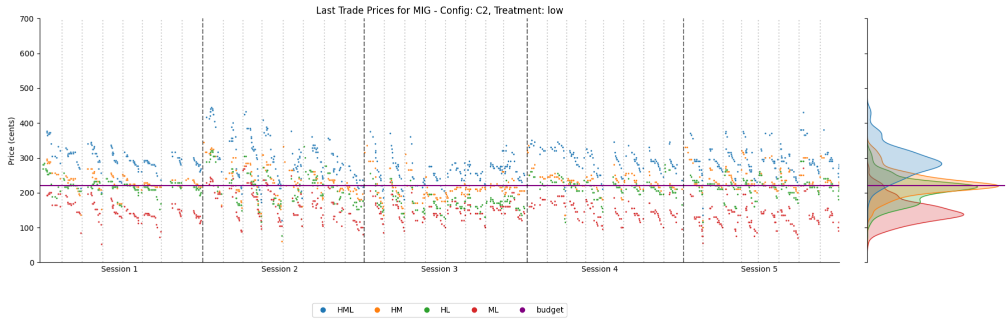
(d) Walras, Payoff Configuration C2

Notes: Dotted vertical lines separate experimental sessions. Purple horizontal line shows investor budget level. Right panel of each figure shows the histogram of prices. A price is shown whenever there is a trade. Only one asset traded; prices of other assets are obtained as last traded price.

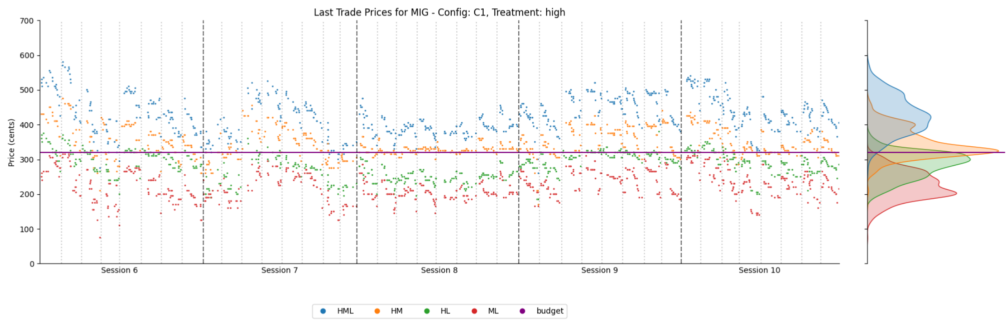
FIGURE A1 Time-series plot of asset trade prices, by treatment



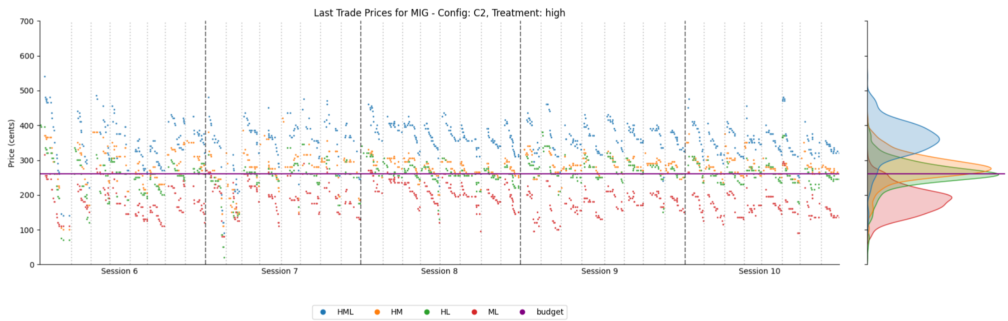
(a) No Walras, Payoff Configuration C1



(b) No Walras, Payoff Configuration C2



(c) Walras, Payoff Configuration C1



(d) Walras, Payoff Configuration C2

Notes: Dotted vertical lines separate experimental sessions. Purple horizontal line indicates investor budget level. Right panel of each figure shows histogram of prices. A price is shown whenever there is a trade. Only one asset traded; prices of other assets are obtained as last traded price.

FIGURE A2 Time-series plot of prices of various asset bundles, by treatment

Appendix D Experiment Participant Instructions

The following instructions were as given to participants in Session 10. Between sessions, the only change in the instructions was the Consumer First Asset Values (payoffs) table.

Markets with Indivisible Goods: INSTRUCTIONS

Summary
You will trade three assets and cash in an anonymous online marketplace with other participants. You will have one of two goals. You may start with cash but no assets, in which case your goal is to acquire **one** unit in as many assets as you can, for as little cash as possible. You may start with assets but no cash, in which case your goal is to sell **all** your assets, for as much cash as possible.

This game will be played multiple times. A replication will be referred to as a "period." Periods are identical, except that you may start with different initial endowments and the value of assets may change. Your take-home pay will depend on your earnings from four (4) randomly drawn periods out of all the periods we run.

Trading Game

In the trading game, in each period you will be assigned the role of either a **Seller** or a **Consumer**. You can find out which you are by looking at your holdings. This is important, so please note it each period.

A **Seller** will start with a positive quantity of the assets, but no cash. A Seller gets **NO** value from holding assets, but they do for cash, so they want to sell their assets for as much cash as possible. Any assets not sold by a Seller at the end of the period will be discarded for no value.

A **Consumer** starts with cash, but no assets. Assets provide significant value to the Consumers, so Consumers will likely want to use their cash to purchase assets. Importantly, Consumers only get value from the **FIRST** unit of each asset they hold, plus any remaining cash they have left over. A Consumer holding two or more units of an asset will not collect value for the subsequent units; only the first unit will count.

The three assets are called A, B, and C, and their **Consumer First Asset Values** can be found below:

Consumer First Asset Values	A	B	C
Periods 1-8 (and practice)	£1.20	£2.40	£3.00
Periods 9-16	£2.40	£3.20	£1.60

In the trading platform, currency will be expressed in experimental dollars, referred to with the symbol $. This is just the software that we use. $1.00 experimental dollar is equivalent to £1.00 British Pound.

Trading

Online Marketplace

In this experiment, you are asked to trade with other participants online through a platform called Flex-E-Markets. This will be opened and signed in on the computer in front of you. We will spend time demonstrating the platform soon, but for now we need to focus on the rules of trading. However, if you have any issues or are signed out during the experiment, raise your hand and an experimenter will sign you back in.

Trading Protocol

Trading takes place as follows. You submit limit orders: orders to buy a unit of an asset at a chosen price (or lower), or to sell a unit at a chosen price (or higher). Transactions take place from the moment a buy order with a higher (or equal) price crosses a sell order with a lower (or equal) price or the other way around. Buy orders are coloured blue; sell offers are coloured red.

<u>All trade occurs at the price specified by the best standing order</u>. In other words, if a trade occurs, the price of the earlier best order determines the price. Orders at a better price execute first. Given a price, orders arriving earlier execute first. Orders remain valid until you cancel them, or the marketplace closes.

We will now demonstrate the trading platform. After, you will be given sufficient time to practice submitting and cancelling orders. We will then complete the instructions, before running four full practice periods of the Trading Game.

Numerical Examples

Below we provide a few numerical examples to illustrate how your period earnings will be calculated.

1. Imagine you are a **Consumer**, and you start with $4.00 of cash. The Consumer First Asset Value of asset A is $1.20, the Consumer First Asset Value of asset B is $2.40, and the Consumer First Asset Value of asset C is $3.00. During the trading period,
    a. you bought one unit of asset A for $0.80 and one unit of asset B for $2.00. In this case, your final portfolio is worth $1.20 (from the first unit of asset A) + $2.40 (from the first unit of asset B) + $4.00 (your starting cash) - $0.80 (the cost to buy A) - $2.00 (the cost to buy B) = $4.80.
    b. you bought two units of asset A for $0.80 each and one unit of asset B for $2.00. In this case, your final portfolio is worth $1.20 (from the first unit of asset A) + $2.40 (from the first unit of asset B) + $4.00 (starting cash) - $0.80 (buying A) - $0.80 (buying A) - $2.00 (buying B) = $4.00. Note that the second unit of asset A has no value to you, but you still paid $0.80 for it (an $0.80 loss).
    c. you bought one unit of asset C for $3.25. In this case, your final portfolio is worth $3.00 (from the first unit of asset C), + $4.00 (starting cash) - $3.25 (buying C) = $3.75 (less than you started with)
2. Imagine you are a **Seller**, and you start with 2 units of asset A, 2 units of asset B, and 2 units of asset C. During the trading period,
    a. You sold two units of asset A for $0.75 and $0.85 respectively, two units of asset B for $2.10 and $1.90 respectively, and two units of asset C for $1.20 and $2.80 respectively. Your final payout would $0.75 + $0.85 (from the two units of asset A you sold) + $2.10 + $1.90 (from the two units of asset B you sold) + $1.20 + $2.80 (from the two units of asset C you sold) = $9.60.
    b. You sold two units of asset A for $0.80 each, two units of asset B for $2.00 each, and no units of asset C. Your final payout would $0.80 + $0.80 (from the two units of asset A you sold) + $2.00 + $2.00 (from the two units of asset B you sold) = $5.60. Note that the units of asset C you did not sell has no value for you.
    c. You sold no units at all. Your final payout would be $0.00 as the assets have no value to you.

Take-Home Pay

There will be **four** (4) **practice periods** and **sixteen** (16) **experiment periods**. Periods will last **3 minutes** each.

At the end of the experiment, we will randomly draw four (4) periods from the experiment periods and pay your earnings for those periods. This means it is in your best interest to perform well in all periods, because you do not know which will be paid. Your earnings in those four periods will be converted to British Pounds at the exchange rate of 1:1 ($1 converts to £1). This performance pay will be added to a base pay of £15. We expect take-home pay to be anywhere between £20 and £40.

GOOD LUCK!

(Version 23/05/2024)

Appendix E Flex-E-Markets Trading Interface

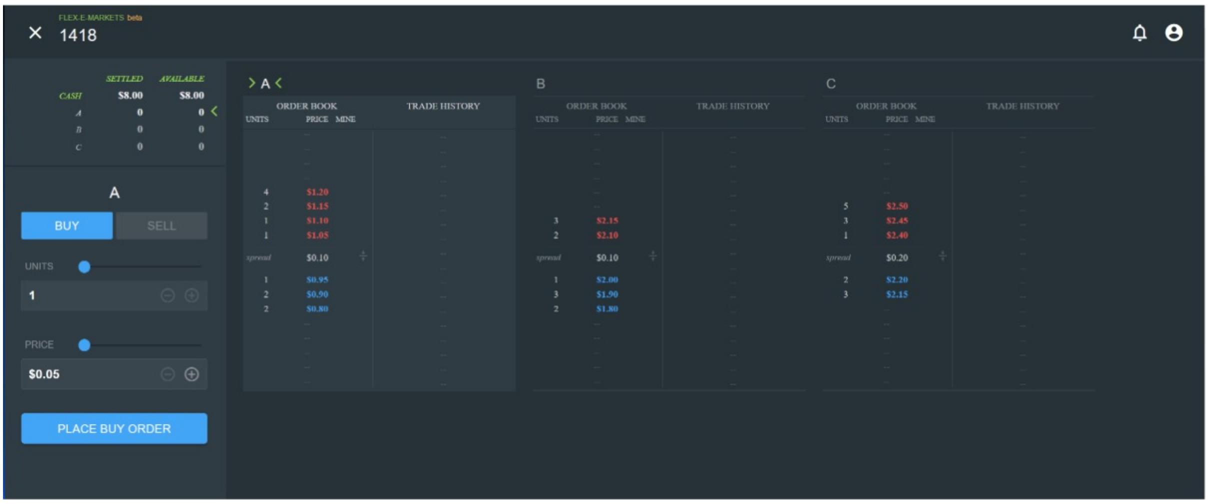


FIGURE A3 Flex-E-Markets Trading Interface