Structural Econometric Tests Of General Equilibrium Theory On Data From Large-Scale Experimental Financial Markets

Peter Bossaerts[‡], Charles Plott[§] and William Zame[¶]

This version: 8 March 2001

Preliminary!

[‡]California Institute of Technology and CEPR
[§]California Institute of Technology
[¶]UCLA

Structural Econometric Tests Of General Equilibrium Theory On Data From Large-Scale Experimental Financial Markets

Peter Bossaerts, Charles Plott and William Zame

Abstract: We develop structural econometric tests of asset pricing theory for application to data from experimental financial markets. The tests differ from those used in the analysis of field data because they verify the consistency between prices and allocations, as opposed to merely testing whether only prices satisfy equilibrium restrictions. Our tests also differ from standard field tests because the experimentor controls (knows) the very parameters that need to be estimated in the field. The tests build on an extension of general equilibrium theory where agents' choices need not be predicted perfectly, rendering the theory more realistic, without however destroying the main pricing results. The tests are implemented on two series of large-scale financial markets experiments. The first series is the certainty equivalent of the second one, with nonlinearity in payoffs replacing uncertainty. The goal of the tests is to resolve an apparent price-allocation paradox in the experiments, namely, why asset prices can satisfy theoretical restrictions (in our case, the CAPM) when allocations deviate substantially (portfolio separation fails). Most puzzling is the finding that allocations seem not to be any closer to CAPM predictions than when prices grossly violate CAPM restrictions. We find that our theory explains the paradox in both sets of experiments. That is, when end-of-period prices are such that the market portfolio is close to mean-variance efficient (the CAPM pricing prediction), the allocations and our theory explain why. When end-of-period prices make the market portfolio far from mean-variance optimal, our tests reject.

Structural Econometric Tests Of General Equilibrium Theory On Data From Large-Scale Experimental Financial Markets‖

Peter Bossaerts, Charles Plott and William Zame

1 Introduction

The purpose of this paper is to develop and implement structural econometric tests of general equilibrium asset pricing theory on data from large-scale experimental financial markets. These tests differ from standard tests on field data in two significant ways. First, the tests use cross-sectional portfolio holdings data, and therefore, verify the consistency between securities prices and allocations. In field studies, one generally only verifies whether prices (price dynamics) satisfy some necessary conditions for equilibrium. Second, unlike tests on field data, the joint distribution of asset returns is not treated as unknown, because it is one of the design parameters, and hence, known to the experimentor. This also implies that the sampling error in the estimation of the joint distribution of asset returns cannot be the source of the uncertainty on which the structural econometric tests are built.

Econometric tests of asset pricing theory on field data almost invariably focus on prices, ignoring the allocational predictions of the same theory. Tests of the Capital Asset Pricing Model (CAPM), for instance, have merely verified whether prices are such that the market portfolio (supply of risky securities) is mean-variance optimal. The allocational prediction of the CAPM, namely, that all investors should hold the same portfolio of risky securities (a result which has been referred to as portfolio separation), is rarely if ever verified. Yet economists are not directly interested in prices. They are mostly interested in allocations, and whether competitive markets manage to indeed generate optimal allocations (the first welfare theorem).

To be sure, the allocational predictions of standard asset pricing theory are extreme and a quick look at any evidence makes one reject them off hand. Investors do not hold risky securities in the same proportions, unlike predicted by the CAPM, for instance. One wonders what the meaning is of the multitude of tests of asset pricing theory that have been performed solely on prices if the allocations are plainly at odds with the theory and if, as is widely believed, prices depend crucially on allocations. This paper will provide a constructive answer.

It is our view that any realistic theory of general equilibrium ought to allow for discrepancies between the tight theoretical allocational predictions and the observed outcomes. There are a variety of reasons why allocations may

‖Financial support was provided by the National Science Foundation, the California Institute of Technology, and the R.G. Jenkins Family Fund. The paper benefited from comments during seminars at the Swedish School of Economics (Stockholm), the Norwegian Central Bank, UCLA and Caltech.

not be exactly as predicted. Determination and implementation of optimal portfolio holdings requires a subtantial amount of computational ability and trading skill. Investors may be satisfied with near-optimal holdings if there are no further incentives (an issue that many have brought up in the context of financial markets experiments, where potential earnings, while generous, are nevertheless limited). In addition, prices are observed at a relatively arbitrary point in time (e.g., the end of a trading period), and it is not clear whether any/many/all investors could have re-traded at those prices if they had to in order to obtain optimal holdings. Also, the preference assumptions needed to operationalize the theory (e.g., quadratic preferences) may only be correct as an approximation. Even if optimal for the true preferences, actual holdings may deviate from those computed on the basis of the hypothesized preferences. Finally, even if subjects are mean-variance optimizers, their risk aversion may change randomly over time, in which case the strict predictions of the theory (based on constant risk aversion) will fail.

To accommodate these discrepancies, we opted to introduce investor-specific perturbation terms at the demand level. These perturbations terms are meant to capture unmodeled features of the problem so as to render realism to the outcomes. Stochastic structure is imposed on the perturbation terms so as to generate falsifiable predictions. The specific stochastic structure used in the present paper is such that, through the law of large numbers, exact CAPM pricing is recovered when the number of investors (subjects) is large, even if no investor's portfolio holdings satisfies strict portfolio separation. Deviations from exact CAPM pricing with a small number of subjects are then to be explained as a standard small-sample effect.

The latter result is not only of interest as a basis for formal econometric tests of the CAPM. It also teaches that strict portfolio separation is not crucial for CAPM pricing, unlike widely believed. The market portfolio (supply of risky securities) can be mean-variance optimal even if nobody demands a fully optimal portfolio, i.e., even if everybody demands a portfolio below the frontier of mean-variance efficient portfolios.

Our theory explains why we had to resort to large-scale experiments before we observed clean evidence of CAPM pricing. In small-scale experiments, full equilibration rarely obtains, even if the tendency toward CAPM pricing is apparent. See [1]. Assuming that our theory is indeed correct, the large scale of our experiments at the same time facilitates econometric analysis, as correctness of the distributional properties of our tests relies on a sizeable number of subjects, and not on a long time series.

The demand perturbations in our theory provide a natural source on which to build econometric specification tests. This is fortunate, because tests of asset pricing theory have traditionally been based on the sampling error in the estimation of the joint distribution of asset returns. The latter is unknown in field research. It is a design parameter in experiments, and therefore, perfectly known. Moreover, by building specification tests on individual portfolio holding perturbations, one obtains more comprehensive tests of asset pricing theory: unlike tests on field data, ours will not only require that prices satisfy some necessary conditions for equilibrium, but also that they be consistent with the cross-sectional pattern of portfolio holdings.

Our tests are applied to large-scale financial markets experiments organized at Caltech in the period 1998-2000. Some of these experiments did not involve any uncertainty; payoffs on what would normally be interpreted as risky securities were nonlinear in holdings, effectively inducing the equivalent of risk aversion in subjects' preferences. The theory considers the two series of experiments to be equivalent both in terms of pricing and in allocations. Our econometric tests will establish whether this theoretical equivalence obtains. All experiments are discussed qualitatively and in more detail in [2, 3]. In this paper, we focus on econometric tests.

We analyze the pricing in these experiments with a standard asset pricing model, the Capital Asset Pricing Model (CAPM). This model assumes that investors like return but dislike risk. In theory, the CAPM captures pricing in markets with small risk fairly well (see, e.g., [7]). It has been objected, however, that the risk in experiments like ours is too small for any risk aversion to show up (see, e.g., [8]). Such an objection rests on an interpretation of what constitutes small risk that is outside the theory.[1] In addition, there are the facts. As [2] already documented, prices do consistently deviate from risk neutrality and appear to be explainable in terms of standard notions of risk aversion.[2] Because they are based on allocations as well as on prices, our structural econometric tests will verify whether prices are explained by allocations as predicted by equilibrium theory with standard risk aversion. If our tests reject, we must conclude that what looks like equilibrium pricing under standard risk aversion is in fact masking something very different. Either the principles of general equilibrium theory are violated (perhaps in favor of the popular view in finance that a few marginal, rational investors determine prices – see, e.g., [9]). Or the representation of preferences as expected utility with standard risk aversion may be wrong (maybe to be replaced with preferences that exhibit ambiguity aversion).

It is not the goal of our structural econometric tests to measure how frequently our theory is rejected on prices and allocations at the end of a trading period. We already know that it takes time for prices to move in line with the predicted equilibrium configuration and, primarily in the case of the uncertainty experiments, that markets easily move away from the predicted equilibrium configuration. See [2, 3]. Instead, the goal is to determine whether our version of general equilibrium theory (with the added demand perturbations) can explain an apparent price-allocation paradox, namely, when prices satisfy CAPM restrictions, allocations are vastly at odds with CAPM predictions and seem not to be any closer to CAPM predictions than when prices substantially violate CAPM restrictions. In particular, portfolio separation (all subjects hold risky securities in the same proportions) *always* fails. Our theory predicts CAPM pricing even if every subject violates portfolio separation.

If prices are close to CAPM predictions and our test fails to reject, one cannot anymore claim that the theory is not

[1] In particular, it rests upon the unproven proposition that expected utility theory only applies to lifetime wealth, and hence, that risk has to be evaluated relative to lifetime wealth only.

[2] Risk aversion has also been observed in a multitude of experiments on individual decision making (on willingness to bet) and auctions with uncertain payoffs.

supported because prices were not *exactly* as predicted by CAPM. The structural econometric tests would generate further confidence in our theory if, in addition, they generate estimates of subjects' risk aversions that are significantly positive.

The remainder of this paper is organized as follows. Section 2 discusses how to introduce realism in standard general equilibrium theory by adding perturbation terms to individuals' demands. Section 3 derives the structural econometric tests that this extended theory suggests. Section 4 presents the test results for the certainty-equivalent experiments. Section 5 presents them for the uncertainty experiments. Section 6 concludes.

2 The Theory

We are going to study simple financial markets where agents (subjects) are endowed with a number of risky and riskfree securities and cash. Subjects can trade in these securities for a pre-determined interval of time, after which the securities are liquidated. The liquidating dividend is drawn from a simple, commonly known distribution.

General equilibrium theory makes predictions about the eventual pricing and allocations in this setup. Because risk in our markets will be relatively small, the equilibrium can conveniently be taken to be that of the CAPM (Capital Asset Pricing Model), as if subjects' preferences traded off only mean against variance. See, e.g., [7]. The general equilibrium formulation of the CAPM that is convenient for our purposes may be described as follows. There are two dates, with uncertainty about the state of nature at the terminal date. A single good is available for consumption at the terminal date; there is no consumption at the initial date. $J+1$ assets (claims to state-dependent consumption at the terminal date) are traded at the initial date and yield claims at the terminal date. The first asset is riskless, yielding one unit of consumption independent of the state; all other assets are risky. Without loss, we assume asset payoffs are non-negative.[3]

The N subjects in the market are endowed only with assets; write $h_n^0 \in R_+$ for subject n's endowment of the riskless asset and $z_n^0 \in \mathbf{R}_+^J$ for subject n's endowment of the risky assets. Subjects have quadratic expected utility. We will work with the equivalent representation of this utility in terms of a trade-off between mean return against variance.[4] We assume that subjects hold common priors, so that they agree on the mean and variance of any portfolio. Write $D_j(s)$ for the return of the j-th risky asset in state s. Let μ be the vector of expected payoffs of risky assets and $\Delta = [\mathrm{cov}\,(D_j, D_k)]$ be the covariance matrix. A subject who holds h units of the riskless asset and the vector z

[3] In general, the number of assets may be smaller than the number of states, so markets may be incomplete. In all the experiments, markets were complete, however.

[4] The two representations are equivalent in the sense that the demand for risky securities of an agent with mean-variance preferences will be proportional to that of one with quadratic expected utility, with the constant of proportionality depending only on the risk aversion coefficient. We will prove this after introducing some more notation.

of risky assets will enjoy utility

$$u_n(h,z) = h + [z'\mu] - \frac{b_n}{2}[z'\Delta z] \tag{1}$$

Since there is no consumption at the initial date, we may normalize so that the price of the risky asset is 1; write p for the vector of prices of risky assets. Given prices p, consumer n's budget set consists of portfolios (h,z) that yield non-negative consumption in each state and satisfy the budget constraint

$$h + p \cdot z \leq h_n^0 + p \cdot z_n^0 \tag{2}$$

Assuming the portfolio choice yields consumption between 0 and $1/b_n$ in each state of the world (an assumption that holds in all our experiments), the first order conditions characterize optima, so investor n's demand for risky assets given prices p is

$$\tilde{z}_n(p) = \frac{1}{b_n}\Delta^{-1}(\mu - p) \tag{3}$$

In particular, all investors choose a linear combination of the riskless asset and the *same* portfolio of risky assets — a conclusion usually referred to as *portfolio separation*.[5]

As usual, an *equilibrium* consists of prices $\tilde{p}$ for assets and portfolio choices $\tilde{h}_n, \tilde{z}_n$ for each investor so that subjects optimize in their budget sets and markets clear. Given the nature of demands (3), market clearing requires that the portfolio $\Delta^{-1}(\mu - \tilde{p})$ be a multiple of the market portfolio $z^0 = \sum z_n^0$ (the supply of risky assets). Solving for equilibrium prices yields

$$\tilde{p} = \mu - \left(\sum_{n=1}^{N} \frac{1}{b_n}\right)^{-1} (\Delta z^0)$$

It is convenient to write $B = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{b_n}$ for the mean inverse risk aversion (we'll refer to it as risk tolerance) and $\bar{z} = \frac{1}{N}\sum_{n=1}^{N} z_n^0$ for the mean endowment (mean market portfolio); with this notation, the pricing formula becomes

$$\tilde{p} = \mu - B^{-1}\Delta\bar{z} \tag{4}$$

A little analysis shows that the pricing formula (4) implies that the mean market portfolio $\bar{z}$ (equivalently, the market portfolio z^0) is mean-variance optimal (ie., among all portfolios with the same variance, it has the highest

[5] We are ready to prove that the demand we have just derived is proportional to the demand of an agent with expected quadratic utility. Consider an expected quadratic utility index with risk aversion coefficient β_n:

$$E[u_n^q(h,z)] = h + [z'\mu] - \frac{\beta_n}{2}[z'\Delta z] - \frac{\beta_n}{2}(h + [z'\mu])^2.$$

The optimum demand for risky securities equals:

$$\tilde{z}_n(p) = \frac{1+\beta_n}{\beta_n}\Delta^{-1}(\mu - p),$$

which is the same as (3), with

$$b_n = \frac{\beta_n}{1+\beta_n}.$$

expected net return). This is not obvious, but is not surprising either. Investors demand mean-variance optimal portfolios, which can all be decomposed into the riskfree security and a single portfolio of risky securities. In equilibrium, aggregate demand for risky securities must equal supply, which is the market portfolio. Hence, in equilibrium, the market portfolio must be the single mean-variance optimal portfolio that is part of each subject's demand. That is, the market portfolio must be mean-variance efficient.

While standard, the foregoing implicitly makes extreme assumptions about subjects' abilities or even willigness to compute and implement optimal allocations. In addition, our representation of preferences may only be approximately true. Moreover, we inevitably test the model at a relatively arbitrary point in time (the end of a trading period). It is not clear whether subjects would all have been able to trade at those prices. While the latter is usually ignored in tests of asset pricing theory on field data (where all observations are treated as if they represent equilibrium outcomes), it must not and need not be in an experimental setting.

So, we expand the model such that actual portfolio choice is represented as a (possibly random) perturbation of the theoretical (CAPM optimal) portfolio choice. Formally, we suppose that actual demand z_n is the sum of theoretical demand $\tilde{z}_n$ and a perturbation ϵ_n:

$$z_n = \tilde{z}_n(p) + \epsilon_n \tag{5}$$

We interpret the perturbation terms as representing the effect of individual errors or incomplete information or as embodying unmodeled heterogeneity; we discuss these possibilities in greater detail below. We *do not* assume the perturbations are small. We only impose that perturbations are independent across n, with mean zero

$$E[\epsilon_n] = 0, \tag{6}$$

and finite variance V

$$E[\epsilon_n^2] = V > 0. \tag{7}$$

Direct computation shows that equilibrium prices will be:

$$p = \mu - \left(\sum_{n=1}^{N} \frac{1}{b_n}\right)^{-1} \Delta z^0 + \left(\frac{1}{N}\sum_{n=1}^{N} \frac{1}{b_n}\right)^{-1} \Delta \frac{1}{N}\sum_{n=1}^{N} \epsilon_n \tag{8}$$

Hence the difference between equilibrium prices p and CAPM equilibrium prices $\tilde{p}$ is

$$|p - \tilde{p}| = |\left(\frac{1}{N}\sum_{n=1}^{N} \frac{1}{b_n}\right)^{-1} \Delta \frac{1}{N}\sum_{n=1}^{N} \epsilon_n| \tag{9}$$

Now consider adding subjects, ensuring that the endowments average z^0 remains finite, risk tolerances $1/b_n$ have a finite non-zero mean, and perturbations ϵ_n continue to be independent with zero mean and finite variance. To emphasize the dependence on the number of subjects, write $\tilde{p}^N, p^N$ for equilibrium prices in the CAPM economy and perturbed economy with N investors. The Law of Large Numbers implies that, as $N \to \infty$

$$\begin{aligned} |p^N - \tilde{p}^N| &\to 0 \text{ a.s.,} && (10) \\ \frac{1}{N}\sum_{n=1}^{N}[z_n - \tilde{z}_n]^2 &\to V \text{ a.s..} && (11) \end{aligned}$$

That is, for large N equilibrium prices in the true economy will (with high probability) be close to equilibrium prices in the benchmark CAPM economy, but equilibrium portfolio choices in the true economy will *not* be close to equilibrium portfolio prices in the benchmark CAPM economy. (More specific assumptions about the distribution of perturbations lead, via the Central Limit Theorem, to estimates on the rates of convergence.)

Notice that our introduction of perturbation terms at the demand level is a deliberate modeling choice (we could have introduced them, e.g., at the utility level). It ensures that the pricing results (CAPM pricing) are not invalidated, at least as long as there is a large number of subjects.

3 The Econometrics

The model in the previous section provides a unified framework to think about structural tests on experimental data. Typically, we have data from several periods, each period being a replication of the same environment. We'd like to test the theory on the prices and allocations at the end of each period. We present an estimation and testing strategy within the framework of GMM. We elaborate how to estimate the weighting matrix that is needed for the GMM test to work. The latter is nonstandard.

3.1 A GMM Test

We first have to introduce a time index, to capture the fact that we will have price and allocation data in each of T periods. Use t for this time index: $t = 1, 2, ..., T$. This way, p_t denotes the (vector) of prices of risky securities in period t. Although we could choose differently, we'll take the end-of-period transaction prices. Likewise, $z_{n,t}$ denotes subject n's holdings of risky securities at the end of period t. In addition, $\tilde{z}_n(p_t)$ denotes what would have been subject n's optimal choice given transaction prices at the end of period t. The discrepancy is:

$$\epsilon_{n,t} = z_{n,t} - \tilde{z}_n(p_t).$$

As before, we assume:

$$E[\epsilon_{n,t}] = 0, \tag{12}$$

with $\epsilon_{n,t}$ mutually independent across n for a fixed t.

Because the condition in (12), imposes restrictions on moments, it is possible to construct a GMM (Generalized Method of Moments) test. The test is to be run on the cross-section of holdings at the end of a period t, given the price vector p_t. It will tell us whether prices and allocations are consistent with our version of general equilibrium theory. At the same time, it will generate an estimate of the average risk tolerance $(\sum_n 1/b_n)/N$.

The focus on a single cross-section may seem to be unfortunate, because one would need to estimate a number of parameters equal to the sample size: for each n, we would have to estimate the corresponding incidental parameter b_n. A clever choice of an $(N+1)$st paramater, however, makes all but the $N+1$st parameters unidentified. This really means that we do not need estimates of the b_ns to test our theory. Which is a plus: they would have been estimated with a large error anyway. Still, at one point we will need some information about the b_ns, namely, their cross-sectional variance. This we will estimate outside the GMM framework. The cross-sectional variance is needed to obtain a valid weighting matrix to be used in GMM. We will discuss this later on.

Let the single identified parameter be denoted b^N. It is defined implicitly to be:

$$\frac{1}{N}\sum_{n=1}^{N}\frac{b^N}{b_n}=1. \tag{13}$$

Now define $g_{n,t}$:

$$g_{n,t}=b^N\epsilon_{n,t}. \tag{14}$$

Notice:

$$E[g_{n,t}]=b^N E[\epsilon_{n,t}]=0. \tag{15}$$

While this is again a moment condition, it won't suffice as a basis for our GMM test, however. Since we want to run the test on the cross-section of holdings conditional on prices p_t, we would really like a restriction on $E[g_{n,t}|p_t]$. Unfortunately, while $E[\epsilon_{n,t}]=0$, the same does not hold for $E[\epsilon_{n,t}|p_t]$. This is because prices depend on $\epsilon_{n,t}$ (see 8), which means that $\epsilon_{n,t}$ will be correlated with prices. Nevertheless, as $N\to\infty$, prices converge to the perturbation-free CAPM pricing (see (10)) and therefore become independent of $\epsilon_{n,t}$. Therefore:

$$\lim_{N\to\infty}E[g_{n,t}|p_t]=0.$$

GMM tests rely on asymptotic (large-sample) arguments. Hence, we can ignore small-sample correlation between $g_{n,t}$ (and any transformation of $g_{n,t}$) and p_t. Because of this, we won't explicitly refer anymore to conditional moments (conditional on p_t).[6]

The restriction in (15) suggests the following GMM estimation and testing strategy. Define:

$$h_{N,t}=\frac{1}{N}\sum_{n=1}^{N}g_{n,t}. \tag{16}$$

[6] Strictly taken, we also need a minimum speed of convergence for the moment conditions, because we will eventually derive asymptotic

Estimate the parameters $\{b_n\}_{n=1}^N$ and b^N by minimizing the following quadratic form:

$$\min_{b1,\ldots,b_N,b^N}[\sqrt{N}h'_{N,t}]A^{-1}[\sqrt{N}h_{N,t}],$$

where the weighting matrix A is a consistent estimate of the asymptotic covariance matrix of $\sqrt{N}h_{N,t}$.

Notice, however, that the criterion function really does not depend on the first N parameters, because:

$$\begin{aligned} h_{N,t} &= \frac{1}{N}\sum_{n=1}^{N} g_{n,t} \\ &= \frac{1}{N}\sum_{n=1}^{N} b^N\epsilon_{n,t} \\ &= \frac{1}{N}\sum_{n=1}^{N} b^N z_{n,t} - \frac{1}{N}\sum_{n=1}^{N}\frac{b^N}{b_n}\Delta^{-1}(\mu - p_t) \\ &= \frac{1}{N}\sum_{n=1}^{N} b^N z_{n,t} - \Delta^{-1}(\mu - p_t). \end{aligned} \tag{17}$$

That is, the first N parameters are not identified. We already pointed out that this is a plus, the disadvantage being that we won't have estimates of the individual b_ns. Hence, the GMM estimation boils down to the following minimization:

$$\min_{b^N}[\sqrt{N}h'_{N,t}]A^{-1}[\sqrt{N}h_{N,t}], \tag{18}$$

The tricky part is to determine A. We shouldn't elaborate yet. Suffice it to say that the usual procedure cannot work. Traditionally, A is obtained as the sum of the sample covariance matrices of the $g_{n,t}$s when evaluated at the parameter values that minimizes the GMM criterion function for $A = I$ (the identity matrix). The problem is that

distributional results by multiplying sample moments with the square root of the sample size ($\sqrt{N}$). In particular, we could require

$$\lim_{N\to\infty}\sqrt{N}E[g_{n,t}|p_t] = 0.$$

Ignoring the distinction between mean independence and absence of correlation, this requirement is satisfied in the present case. From (8),

$$\begin{aligned} &\sqrt{N}E[g'_{n,t}p_t] \\ &= \sqrt{N}E[b^N\epsilon'_{n,t}(\mu - \tilde{p}^N)] + \sqrt{N}E[b^N\epsilon'_{n,t}b^N\Delta\frac{1}{N}\sum_{n=1}^{N}\epsilon_{n,t}] \\ &= 0 + \frac{1}{\sqrt{N}}E[(b^N)^2\epsilon'_{n,t}\Delta\epsilon_{n,t}] \\ &\to 0. \end{aligned}$$

We refrain from incorporating these complications in our analysis, because they would clutter the notation without adding any deep insight. Moreover, the arguments are standard. The interested reader should consult the literature on Pitman drift (see, e.g., [4], p. 154 ff.).

the sample covariance of each $g_{n,t}$ depends on its incidental parameter b_n, which are not identified, and hence, not estimable in GMM, as determined before.

Because the $\epsilon_{n,t}$ and, hence, the $g_{n,t}$ are independent across n for fixed t, and the $g_{n,t}$ are uniformly continuous in b^N for b^N in a compact interval, we obtain the usual convergence results. Let $\hat{b}^N$ denote the GMM estimate of b^N. First,

$$(\hat{b}^N - b^N) \to 0 \text{ a.s.}.$$

Because $1/B = \lim_{N\to\infty}(\sum_n 1/b_n)/N$, another way to state this is:

$$\hat{b}^N \to \frac{1}{B} \text{ a.s.}.$$

Second, at its minimum, the quadratic form in (18) will converge in distribution to a χ^2 random variable with degrees of freedom equal to J (the number of risky assets) minus 1. That is,

$$\min_{b^N}[\sqrt{N}h'_{N,t}]A^{-1}[\sqrt{N}h_{N,t}] \sim \chi^2_{J-1}$$

in large samples.

3.2 Obtaining The Weighting Matrix A

As mentioned before, one cannot obtain the weighting matrix A in the GMM criterion function in the usual way, which is to first estimate b^N as in (18) but substituting the identity matrix for A, and subsequently estimating A as the sum of the covariance matrices of the $g_{n,t}$s at the estimated b^N. The covariance matrices of the $g_{n,t}$s do not only depend on b^N, but on the incidental parameters $\{b_n\}_{n=1}^N$ as well. These are not identified in our GMM estimation, whether we set A equal to the identity matrix or not.

Let us first analyze the covariance matrix of $\sqrt{N}h_{N,t}$, in order to determine more precisely what we need to know about the incidental parameters to obtain A. We won't express the covariance matrix of $\sqrt{N}h_{N,t}$ in its simplest form. Instead, we will express it in such a way that errors in the estimation of the incidental parameters average out in large cross-sections (N large), effectively exploiting the typical structure of samples from our experiments, namely, a panel dataset, and obviating the need to have long time series (T large).

It will be important to distinguish between, on the one hand, the average covariance matrix of subject's holdings, namely,

$$\frac{1}{N}\sum_{n=1}^{N}\text{cov}(z_{n,t}) = \frac{1}{N}\sum_{n=1}^{N}E(z_{n,t} - E[z_{n,t}])(z_{n,t} - E[z_{n,t}])',$$

and, on the other hand, the average of covariance matrices computed relative to the grand mean of subjects' holdings (as opposed to each subject's own mean holding), to be denoted $\text{cov}(z_t)$:

$$\text{cov}(z_t) = \frac{1}{N}\sum_{n=1}^{N} E(z_{n,t} - \bar{z}_N)(z_{n,t} - \bar{z}_N)',$$

where

$$\bar{z}_N = \frac{1}{N}\sum_{n=1}^{N} E[z_{n,t}].$$

$\text{cov}(z_t)$ can be estimated consistently as the sample covariance matrix of holdings across individuals, and we will refer to it as such. The sample covariance matrix of holdings across individuals equals:

$$\frac{1}{N}\sum_{n=1}^{N}\left(z_{n,t} - \frac{1}{N}\sum_{n=1}^{N} z_{n,t}\right)'\left(z_{n,t} - \frac{1}{N}\sum_{n=1}^{N} z_{n,t}\right)'.$$

Close inspection of this expression reveals that the estimation error will decrease with the size of the cross-section ($N \to \infty$). Large time series are therefore not needed (T need not be large). Our experiments typically have large N, but low T.

We are ready to study the covariance matrix of $\sqrt{N}h_{N,t}$. Straightforward calculations lead to the following.

$$\begin{aligned}
&\text{cov}(\sqrt{N}h_{N,t}) \\
&= N\text{cov}(\frac{1}{N}\sum_{n=1}^{N} b^N \epsilon_{n,t}) \\
&= (b^N)^2 \frac{1}{N}\sum_{n=1}^{N} E[\epsilon_{n,t}\epsilon'_{n,t}] \\
&= (b^N)^2 \frac{1}{N}\sum_{n=1}^{N} E[(z_{n,t} - E[z_{n,t}])(z_{n,t} - E[z_{n,t}])'] \\
&= (b^N)^2 \frac{1}{N}\sum_{n=1}^{N} E[(z_{n,t} - \bar{z}_N + \bar{z}_N - E[z_{n,t}])(z_{n,t} - \bar{z}_N + \bar{z}_N - E[z_{n,t}])'] \\
&= (b^N)^2 \frac{1}{N}\sum_{n=1}^{N} E[(z_{n,t} - \bar{z}_N)(z_{n,t} - \bar{z}_N)'] + (b^N)^2 \frac{1}{N}\sum_{n=1}^{N}(\bar{z}_N - E[z_{n,t}])(\bar{z}_N - E[z_{n,t}])' \\
&\quad + (b^N)^2 \frac{1}{N}\sum_{n=1}^{N} E[(z_{n,t} - \bar{z}_N)](\bar{z}_N - E[z_{n,t}])' + (b^N)^2 \frac{1}{N}\sum_{n=1}^{N}(\bar{z}_N - E[z_{n,t}])E[(z_{n,t} - \bar{z}_N)'] \\
&= (b^N)^2 \,\text{cov}(z_t) - \frac{1}{2}(b^N)^2 \frac{1}{N}\sum_{n=1}^{N}(z_{n,t} - \bar{z}_N)(\bar{z}_N - E[z_{n,t}])' \\
&\quad + \frac{1}{2}(b^N)^2 \frac{1}{N}\sum_{n=1}^{N}\epsilon_{n,t}(\bar{z}_N - E[z_{n,t}])' - \frac{1}{2}(b^N)^2 \frac{1}{N}\sum_{n=1}^{N}(\bar{z}_N - E[z_{n,t}])(z_{n,t} - \bar{z}_N)'
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2}(b^N)^2\frac{1}{N}\sum_{n=1}^{N}(\bar{z}_N - E[z_{n,t}])\epsilon'_{n,t} + (b^N)^2\frac{1}{N}\sum_{n=1}^{N}E[(z_{n,t}-\bar{z}_N)](\bar{z}_N - E[z_{n,t}])' \\
&+ (b^N)^2\frac{1}{N}\sum_{n=1}^{N}(\bar{z}_N - E[z_{n,t}])E[(z_{n,t}-\bar{z}_N)'] \\
= \quad & (b^N)^2\operatorname{cov}(z_t) - \frac{1}{2}(b^N)^2\frac{1}{N}\sum_{n=1}^{N}E[(z_{n,t}-\bar{z}_N)](\frac{1}{b_n} - \frac{1}{N}\sum_{\nu=1}^{N}\frac{1}{b_\nu})(\mu - p_t)'\Delta^{-1} \\
&- \frac{1}{2}(b^N)^2\frac{1}{N}\sum_{n=1}^{N}\Delta^{-1}(\mu - p_t)(\frac{1}{b_n} - \frac{1}{N}\sum_{\nu=1}^{N}\frac{1}{b_\nu})E[(z_{n,t}-\bar{z}_N)'] \\
&+ \frac{1}{2}(b^N)^2\frac{1}{N}\sum_{n=1}^{N}\epsilon_{n,t}(\bar{z}_N - E[z_{n,t}])' + \frac{1}{2}(b^N)^2\frac{1}{N}\sum_{n=1}^{N}(\bar{z}_N - E[z_{n,t}])\epsilon'_{n,t}. \qquad (19)
\end{aligned}
$$

For N large, the last two terms can be ignored, by the law of large numbers. Estimation of the first term was discussed before. Estimation of the second and third terms can be based on

$$(b^N)^2\frac{1}{N}\sum_{n=1}^{N}(z_{n,t}-\bar{z}_N)(\frac{1}{b_n} - \frac{1}{N}\sum_{\nu=1}^{N}\frac{1}{b_\nu})(\mu - p_t)'\Delta^{-1} \qquad (20)$$

and

$$(b^N)^2\frac{1}{N}\sum_{n=1}^{N}\Delta^{-1}(\mu - p_t)(\frac{1}{b_n} - \frac{1}{N}\sum_{\nu=1}^{N}\frac{1}{b_\nu})(z_{n,t}-\bar{z}_N)', \qquad (21)$$

respectively. Because of linearity, any (zero-mean) error in the estimation of the risk tolerances ($1/b_n$s) will be averaged out in cross-section. That is, large N will help reduce the estimation error. At the same time, we need not rely on large T.

We are going to estimate the risk tolerances exploiting the panel data structure of our experimental data. Remember our basic moment restriction (see (12)):

$$E[\epsilon_{n,t}] = 0.$$

We added an independence restriction, however, which applied only to the cross-section ($\epsilon_{n,t}$s are mutually independent across n). Let us now impose some time-series structure on $\epsilon_{n,t}$ for a given n. We won't need independence. To make this explicit, factor the error into two components:

$$\epsilon_{n,t} = \alpha_n + \xi_{n,t}, \qquad (22)$$

where α_n denotes an individual's systematic mistake, assumed constant over time, and $\xi_{n,t}$ denotes the idiosyncratic component. For consistency with our basic assumption (12), we obviously require:

$$E[\alpha_n] = 0, \qquad (23)$$

$$E[\xi_{n,t}] = 0. \tag{24}$$

Assume that the α_ns and ξ_{nt} are mutually independent.

Consider now the history of subject n's holdings over time:

$$z_{n,t} = \tilde{z}_n(p_t) + \epsilon_{n,t}, \quad t = 1, 2, ..., T.$$

We can be more explicit:

$$z_{n,t} = \frac{1}{b_n}\Delta^{-1}(\mu - p_t) + \alpha_n + \xi_{n,t}, \ t = 1, 2, ..., T. \tag{25}$$

Now subtract the time-series mean from both the left-hand-side and the right-hand-side:

$$(z_{n,t} - E[z_{n,t}]) = \frac{1}{b_n}\left(\Delta^{-1}(\mu - p_t) - \Delta^{-1}(\mu - E[p_t])\right) + \xi_{n,t}, \ t = 1, 2, ..., T.$$

In words: a subject's holdings can be explained as the sum of a term proportional to a common component $\Delta^{-1}(\mu - p_t) - \Delta^{-1}(\mu - E[p_t])$ and an idiosyncratic term $\xi_{n,t}$. In the previous section, we proved that prices p_t would converge to error-free CAPM pricing as $N \to \infty$. Hence, the error $\xi_{n,t}$ and the price p_t will be asymptotically independent. This also implies that the error $\xi_{n,t}$ will be asymptotically independent of the common component $\Delta^{-1}(\mu - p_t) - \Delta^{-1}(\mu - E[p_t])$. In addition, the error $\xi_{n,t}$ is independent across n. Altogether, this means that mean-adjusted holdings will asymptotically exhibit a factor structure.

For large N, the covariance matrix of holdings of asset j of the N subjects, Σ_j^N, can thus be factored as follows:

$$\Sigma_j^N = \upsilon_j \gamma^N \gamma^{N'} + \Omega_j^N, \tag{26}$$

where Ω_j^N is a diagonal matrix,

$$\gamma^N = \begin{bmatrix} \frac{1}{b_1} \\ \frac{1}{b_2} \\ \cdots \\ \frac{1}{b_N} \end{bmatrix}, \tag{27}$$

and

$$\upsilon_j = (\Delta^{-1})_j \mathrm{cov}(p_t)(\Delta^{-1})'_j. \tag{28}$$

In the above, $(\Delta^{-1})_j$ refers to the jth row of Δ^{-1}. $\mathrm{cov}(p_t)$ denotes the time-series covariance matrix of prices p_t. We will discuss shortly how to estimate the latter.

Because of the factor structure, the vector of risk tolerances γ^N can readily be estimated from the eigenvector corresponding to the maximal eigenvalue of Σ_j^N. If one's eigenvalue decomposition algorithm produces orthonormal eigenvectors, then γ can be estimated from the eigenvector $l^{\max}$ corresponding to the maximal eigenvalue $\lambda^{\max}$, as follows:

$$\gamma = \sqrt{\frac{\lambda^{\max}}{v_j}} l^{\max}.$$

Closer inspection will reveal that we obtain J estimates of γ, one for each asset. We aggregate these estimates in the simplest possible way, by averaging across assets. The resulting estimates provide the input for the estimator of $\mathrm{cov}(\sqrt{N} h_{N,t})$.

In short, we obtain an estimate of the cross-sectional variance of the risk tolerances by means of an eigenvalue analysis on the covariance matrix of subjects' holdings over time. Several remarks are in order.

First, notice that, when $N > T$, the estimate of the covariance matrix Σ_j^N will be singular. Since we are only interested in the largest eigenvalue of Σ_j^N, singularity is not an impediment.

Second, the covariance matrix of $\sqrt{N} h_{N,t}$ depends also on the average risk tolerance b^N. There are two ways to obtain an estimate of b^N. Following tradition, one could obtain an estimate from minimization of the GMM criterion function with weighting matrix A set equal to the identity matrix. Alternatively, we could estimate b^N as the inverse of the average of the elements of γ. In the applications, we chose the latter procedure throughout. This generally produced a closer fit.[7]

Third, we need a good estimate of the covariance of prices $\mathrm{cov}(p_t)$. One's initial inclination would be to simply take the covariance of end-of-period transaction prices. Still, end-of-period prices are rather arbitrary, in the sense that it is highly unlikely that the majority of subjects could have traded at those prices. The covariance of end-of-period prices may not reflect the actual covariance that subjects faced during an experiment. It is important to use an estimate that reflects the latter. We are really interested in whether the cross-section of holdings can be squared with end-of-period prices, taking into account price uncertainty associated with execution of any trades that would move subjects closer to the ideal, error-free world of standard general equilibrium theory. Therefore, we opted to estimate $\mathrm{cov}(p_t)$ from the time-series covariance of *all* transaction price changes instead of just the end-of-period prices.

Fourth, our estimates of $\mathrm{cov}(\sqrt{N} h_{N,t})$, and hence, the GMM weighting matrix, need not be positive definite. It is not clear how to refine estimation in order to guarantee positive definiteness without destroying the feature that only large cross-sectional size (N) has to be relied on.

3.3 Discussion

It is important to understand what our χ^2 statistic is meant to capture. It tests whether deviations between end-of-period prices and CAPM pricing can be expected by chance in view of the randomness in allocations. In particular, the

[7] We also set equal to zero any elements of γ that were estimated to be negative.

tests ask whether observed mispricing is significant given the noise in subjects' holdings relative to the strict portfolio separation predicted by the CAPM.

If prices in a given period conform exactly to the CAPM (the market portfolio is exactly mean-variance efficient), our χ^2 statistic cannot reject. Inspection of the GMM criterion function (17) reveals why. If the market portfolio is mean-variance efficient, there exists a parameter b^N such that

$$\bar{z} \; (\equiv \frac{1}{N}\sum_{n=1}^{N} z_{n,t}) \; = \frac{1}{b^N}\Delta^{-1}(\mu - p_t)$$

exactly. The left-hand side gives the composition of the market portfolio, expressed as the vector of per-capita supplies of the risky securities (which always equal the average holdings of risky securities in our exchange economy). Equality with the right hand side for some choice of b^N merely states that the market portfolio is mean-variance optimal. If the equality holds, the GMM criterion function in (17) is identically zero.

Strict adherence to CAPM pricing almost never occurs. Instead, prices often do not conform exactly to the CAPM, but they are close. Our χ^2 statistic verifies whether the observed mispricing is likely given the size of typical deviations in subjects' holdings from strict portfolio separation over the entire experiment. If subjects' choices generally deviate substantially from strict CAPM demands, slight mispricing is expected, because subjects' mistakes (relative to CAPM holdings) unlikely average out exactly (which is what is required for mispricing to disappear). If subjects' choices generally follow mean-variance optimization closely, then even small mispricing may show up significantly. The novelty of our tests is that they use noise in allocations to determine whether mispricing is significant.

Note that the tests require the risk aversion coefficients (b_ns) to be constant over time for each subject. Despite this restriction, it should be repeated that the difference between a subject's actual choices and the mean-variance optimal choices implied by his/her b_n need not be mean zero over time. Cf. (22). In addition, there is no restriction on how random subjects' choices can be relative to an optimal demand with a fixed risk aversion b_n. Moreover, certain types of randomness in risk aversion are allowed. For instance, a subject's choice in period t may be based on a random risk tolerance $1/b_{n,t}$ when it is an independent (across subjects and over time) draw from a distribution with finite variance and mean $1/b_n$. If so, we can still write the subject's choice as in (25), with the error

$$\left(\frac{1}{b_{n,t}} - \frac{1}{b_n}\right)\Delta^{-1}(\mu - p_t)$$

absorbed in the shock $\xi_{n,t}$.

At this point, we are ready to summarize the fundamental differences between standard structural tests of asset pricing models on field data (e.g., [5]) and the structural tests we propose here for experimental data.

- Our tests verify an asset pricing model in its entirety: prices *and* allocations. Tests of field data verify only pricing restrictions.[8] They have so far ignored portfolio holdings.[9]

- The source of the sampling error in our tests differs from that on which field data tests are built.

  In field data tests, the payoff distribution is unknown and needs to be estimated in order to obtain the parameters that drive investors' choices, and hence, equilibrium prices. Observed prices are taken to be equilibrium prices, and the empiricist verifies whether these prices can be matched with an asset pricing model given his/her estimates of the parameters of the payoff distribution. The higher the uncertainty about those parameters, the more likely it is that an asset pricing model cannot be rejected (low power).

  In contrast, experimentalists cannot ignore the fact that they know the payoff distribution. The source of the modeling error must be found elsewhere. We identify it as the deviations between actual holdings and theoretical, error-free portfolio choices. The modeling error is not arbitrary: we impose the restrictions that are necessary for sub-optimal (perturbed) allocations still to generate correct pricing in the limit.

- Tests on field data rely on long time series ($T \to \infty$), because that is what it takes for the paramaters of the payoff distribution to be estimated accurately. Of course, this also requires strong restrictions on the stationarity (and memory) of the return processes.

  In contrast, our test builds on large cross-sections ($N \to \infty$). This is fortunate, because it is difficult and expensive to organize experiments with a long duration. It does require, however, that a large number of subjects be present. Typical experiments used to be arranged with only up to 20 subjects ($N = 20$). Our experiments, however, have been arranged with typically 40 to 60 subjects. The results from the experiments about to be discussed in the next section will illustrate that this is the sample size required to generate satisfactory power.

The last point can be put differently: the theory in Section 2 predicts that a large number of subjects are needed for traditional pricing results to obtain; the econometrics of the present section requires that there be a large number of subjects for power reasons.

4 Application I: Certainty-Equivalent Experiments

In this section, we implement our test on a set of experiments that were originally meant to be a benchmark for the asset market experiments that we discuss later on. Specifically, instead of endowing subjects with assets that return an uncertain payoff (with known probabilities), we allocated commodities whose payoff was certain but dependeded nonlinearly on the holdings. According to the theory, subjects in the uncertainty experiments are supposed to evaluate

[8] This may involve nonfinancial data, however, such as aggregate consumption data.

[9] But see, e.g., [6], for calibration exercises meant to explain investors' portfolio holdings.

portfolio choice based on a nonlinear expected utility index like (1). In the certainty experiments, subjects were paid according to the mapping in (1).

Of course, we did not give them the mapping itself – it would demand too much computation for subjects to figure out what the payoff on the portfolio choices would be that he or she is contemplating. Instead, we gave them a table from which they could readily read the payoff for all possible combinations of commodity holdings.

According to standard theory, these certainty experiments should be equivalent in all respects: prices, allocations. Therefore, we have been referring to them as *certainty-equivalent experiments*. We ran them because we wanted to verify whether these experiments also generated the perturbations in subject choice (relative to standard theory) that we perceived in the uncertainty experiments. See [3] for details.

The certainty-equivalent experiments provide an ideal background against which to evaluate the performance of our structural tests. Unlike in the uncertainty experiments, we know exactly what the traditional (CAPM) equilibrium price vector is. When we observe large, economically significant deviations between end-of-period transaction prices and equilibrium prices, we hope to reject (power). When these deviations are small and errors in portfolio choice sufficiently large, we hope to accept.

Likewise, the coefficients b_n are known for each subject in the certainty-equivalent experiment. (In the uncertainty experiments, the b_n are purely theoretical speculation, but it is hoped that individual portfolio choices can be explained by means of the b_ns.) We hope to recover at least the inverse of the average risk tolerance (b^N). Moreover, we are interested in determining the accuracy of the estimates. In particular, we would like the estimates to be significantly different from zero. This is important, because in the uncertainty experiments, we would like to be able to reject risk neutrality ($b^N = 0$).

4.1 Experimental Design

Specifics of the experiments were as follows.[10] There were three commodities, denoted A, B and Notes. Each note paid 100 francs (the experimental currency, to be converted into dollars at a pre-determined exchange rate) at the end of a period. The payoff on holdings of A and B depended on the exact combination in which they were held. In particular, letting z denote the vector of holdings of commodities A and B, the payoff $W(z)$ equals:

$$W(z) = [z \cdot \mu] - \frac{b_n}{2}[z \cdot \Delta z], \tag{29}$$

[10]Instructions and screens for the experiments we discuss here can be viewed at, e.g., http://eeps2.caltech.edu/market-000509. Use identification number:1 and password:a to login as a guest. You will not have a payoff but you will be able to see the forms used. If you wish to interact with the software in a different context you can go to http://eeps.caltech.edu, continue to the experiment and then demo links. This exercise will provide the reader with some understanding of how the software works.

where

$$\mu = \begin{bmatrix} 230 \\ 200 \end{bmatrix},$$

and

$$\Delta = \begin{bmatrix} 9867 & -1000 \\ -1000 & 1400 \end{bmatrix}.$$

μ and Δ are the mean and covariance of the payoffs on securities A and B in the uncertainty experiments. The parameter b_n varies across subjects. It is the equivalent of the risk aversion parameter in the uncertainty experiment. In return for the initial allocation of commodities, subjects were asked to pay a pre-determined loan amount at the end of each period. Subjects were also given cash.

In each of several consecutive periods the commodities were allocated unevenly across subjects. (For a given individual initial allocations remained the same across all periods.) The three markets were then opened for a fixed period of time. When markets closed the holdings of commodities and cash were recorded and earnings computed, after subtracting the loan amount. New allocations of the asset were distributed and a fresh period opened. Subjects whose earnings were sufficiently low were declared bankrupt[11] and were prevented from participating in subsequent periods. Earnings ranged from nothing (the bankrupt) to almost two hundred dollars. The remaining data and parameters for the experiments are displayed in Table 1.

CAPM equilibrium prices in the three certainty-equivalent experiments were as follows.

Experiment	p_A	p_B	p_{Notes}
000320	222	200	100
000509	221	199	100
000511	221	199	100

Figure 1 plots the evolution of transaction prices in the three experiments. In no period do prices exactly coincide with equilibrium prices (although they are really close in 000320). The question is whether the deviations are significant.

In addition, as documented more thoroughly in [3], allocations do not conform to CAPM predictions, even when prices are very close: portfolio separation fails uniformly. Figure 2 illustrates this. Plotted are subjects' holdings of A as a proportion of the wealth they allocated to A and B by the end of each period. Subjects should be holding the market portfolio of A and B (circles) when prices are in line with CAPM. They do not, and while prices over time move closer to CAPM (see Figure 1), allocations seem not to improve at all.

[11]Negative period earnings were possible because of the loan payment. Subjects were not guaranteed a positive payoff, only if they moved to a better position, away from the initial allocation. When doing nothing, a subject would hardly make money or even lose some.

The question then is whether the portfolio holdings are compatible with a version of the CAPM that allows subjects to deviate in specific ways from CAPM optimal demands and that would generate exact CAPM pricing as the number of subjects increases. Our test will shed light on this issue.

4.2 Tests

Table 2 displays the results of our structural econometric test for each period in each experiment. Listed are: (i) the χ^2 statistic and corresponding p value, (ii) the estimate of b^N and corresponding standard error (the actual values of b^N that were used in the experiments are given in footnotes).

It is most illuminating to discuss the results for experiment 000511 in detail. From Figure 1, we know that end-of-period prices deviated substantially from equilibrium values in periods 1 through 3. In principle, our structural econometric tests should reject. Otherwise, the power can be questioned. From periods 4 to 8, end-of-period prices move closer. But even in period 8, prices don't really exactly fit the theory. The question is whether the deviations we observe have not been caused by noise in subjects' choices, and hence, that these deviations are insignificant.

Table 2 displays values for the χ^2 statistic in periods 1 through 3 that reject the null of CAPM pricing (with allocation errors). In periods 4 through 8, we do not find rejections (at the 1% level). Also, the fit improves systematically. All this means that our test statistic "works": it rejects where we expect it to reject, suggesting power; at the same time, it does not find deviations from equilibrium prices of about 10 francs (periods 5 through 8) to be significant.

Experiment 000509 shows how price volatility enters in the structural econometric test. Inspection of Figure 1 reveals that prices are "close" to equilibrium predictions only from period 4 on. Yet our χ^2 statistic fails to reject already from period 2 on. This finding can be attributed to the effect of price volality on the GMM weighting matrix. Volatile prices are likely to generate large dispersions in asset holdings, which directly affects the estimator of the weighting matrix, through $\text{cov}(z_t)$ in (19). In addition, price volatility affects the estimation of the risk tolerances (see (28)). Covariance between the risk tolerances and subjects' holdings is also an ingredient of the estimator of the GMM weighting matrix; see (20) and (21).

The estimation results for 000320 suggest small-sample problems. Figure 1 indicates that prices were extremely close to theoretical predictions during the entire experiment. The χ^2-statistics in Table 2 reflect this, but are often in the far left tail of the χ_1^2 distribution. It is not difficult to conjecture a source of the distributional mis-specification: the GMM weighting matrix is likely to be estimated with substantial error. The χ_1^2 distributional assumption is based on asymptotics, when the weighting matrix is estimated with full accuracy. One can plausibly expect estimation error in the weighting matrix to affect the right tail of the distribution as well. That is, the sizes of our test statistic may be larger than the nominal values obtained from the χ_1^2 distribution.

Comparison between the estimates of b^N (inverse of average risk tolerances) in Table 2 and the corresponding theoretical values (in footnotes to the table) reveals that our procedure generally overestimates. This should be born in mind when interpreting the results for uncertainty experiments, where the theoretical value of b^N is not known. Another indication that our estimation procedure works, however, is that the estimates of b^N are generally significantly positive. There is only one exception: period 8 of 000509. Inspection of Figure 1 explains why. At the end of period 8 of 000509, $p_A = 230$ and $p_B = 200$. At these values, prices conform to a situation in which $b^N = 0$, i.e., the payoff function (29) is linear (which would be interpreted as risk neutrality in the equivalent economy with uncertainty). The estimation picks this up: b^N is insignificantly different from zero.

Finally, it should be pointed out that our test statistic not only correlates with the graphical evidence on pricing in Figure 1. It also correlates with the calculations in [3] of the average amount of money that subjects left on the table because they did not fully optimize.

5 Application II: Uncertainty Experiments

Let us now turn to our uncertainty experiments. We will follow the same plan as with the certainty-equivalent experiments: we first discuss experimental design, then we present the tests.

5.1 Experimental Design

The uncertainty experiments worked like the certainty-equivalent experiments, but subjects were given risky assets instead of commodities whose payoff depended nonlinearly on the holdings.[12]

In particular, there were two risky securities A, B in addition to the riskless security (Note), and limited cash. End-of-period payoffs (dividends) on the risky securities depended on a state variable. Three states were possible and one was drawn at random, with a commonly known distribution. The payoff matrix for all experiments that we discuss here remained the same, namely:

State	X	Y	Z
Security A	170	370	150
Security B	160	190	250
Note	100	100	100

Again, the number of subjects was on the order of 50. In each of several consecutive periods the assets were allocated unevenly across subjects. The three markets were then opened for a fixed period of time. When the period closed the state was drawn and announced, and earnings recorded. New allocations of the asset were distributed and a

[12] Instructions and screens for the experiments we discuss here can be viewed at, e.g., http://eeps2.caltech.edu/market-981116uclacit. Use identification number:1 and password:a to login as a guest.

fresh period opened. Subjects whose earnings were sufficiently low were declared bankrupt and were prevented from participating in subsequent periods. Earnings ranged from nothing (the bankrupt) to over two hundred dollars. The remaining data and parameters for the experiments are displayed in Table 3.

The evolution of transaction prices in the six experiments are plotted in Figure 3. In contrast with the certainty-equivalent experiments, theoretical equilibrium prices cannot be computed, absent knowledge of subjects' risk aversions. So, horizontal lines in Figure 3 do not indicate equilibrium values, but expected payoffs. Transaction prices generally deviate substantially from expected payoffs. When converted to Arrow-Debreu securities prices, these prices eventually signal something that would traditionally be interpreted unambiguously as they equilibrium effects of standard risk aversion: prices for the state with a lowest aggregate dividend are highest; those for the state with the highest aggregate divided are lowest. See the discussion in [2]. The question is: once we take portfolio choices into account, is this really to be interpreted as risk aversion in the expected utility sense? As mentioned before, our structural econometric tests will generate a negative answer.

While we do not know equilibrium prices in each of the uncertainty experiments, we can still determine to what extent prices are in line with the standard version of the CAPM. Indeed, CAPM provides a convenient way to measure how far prices are away from a configuration that is consistent with equilibrium. When prices satisfy the error-free version of the CAPM – equation (4) –, the market portfolio is mean-variance efficient, i.e., provide maximum expected return for its volatility. One can readily measure how far the market portfolio is from mean-variance efficiency by comparing the *Sharpe ratio* of the market portfolio with the maximum possible Sharpe ratio. The Sharpe ratio is defined as the ratio of the expected return in excess of the riskfree rate over the volatility.

Figure 4 displays the evolution of the difference between the Sharpe ratio of the market portfolio and the maximum possible Sharpe ratio. For the CAPM to hold, it must be that this difference is zero. There is a clear tendency for the Sharpe ratio difference to move in the direction required by the CAPM, even if one can certainly reject that the market is always at the CAPM equilibrium. Statistical modeling of the evolution of the Sharpe ratio difference confirms the visual impression: it is constantly being attracted by the CAPM.In particular, [2] reject at high significance levels that CAPM pricing comes about by chance (even if prices are a random walk, they may sometimes make the market portfolio mean-variance optimal, by chance).

As with the certainty-equivalent experiments, we are interested in discovering whether general equilibrium theory (modulo demand perturbations) can explain the pricing successes, i.e., the instances when prices are close to CAPM predictions. The question is motivated by the same paradox as in the certainty-equivalent experiments: when prices are close to CAPM, the predicted portfolio separation does not obtain; in fact, allocations appear to be no closer to equilibrium predictions than when prices violate CAPM restrictions. Figure 5 illustrates this. Plotted are subjects' holdings of A as a proportion of the wealth they allocated to A and B by the end of each period. Subjects should be holding the market portfolio of A and B (circles) when prices are in line with CAPM. They do not, and while prices over time move closer to CAPM (see Figure 4), allocations seem not to improve at all. (See [2, 3] for further evidence on the price-allocation paradox.)

Is this paradox to be explained by our version of general equilibrium theory, where subjects' demand are perturbed versions of fully optimal demands? Our structural econometric tests shed light on this.[13]

5.2 Tests

The results of our structural econometric tests are reported in Table 4. They test whether end-of-period prices and allocations can be explained in terms of our extension of standard general equilibrium theory that can accomodate deviations at the individual level between exact mean-variance optimal holdings and actual holdings.

We find a high correlation between the structural test and the visual evidence in Figure 4. That is, when end-of-period prices are close to CAPM predictions, the χ^2 statistic tends to be low. Which implies that what is visually close in Figure 4 is also close in an econometric sense. Table 5 aptly summarizes the evidence. It provides a two-way classification: significance of the χ^2 statistic (three categories: p-level below 1%, between 1% and 5% and above 5%) against distance from CAPM pricing at last transaction prices as measured by the difference between the market portfolio Sharpe ratio and the maximum Sharpe ratio (terciles, based on the relative distance within each experiment). The χ^2 statistic under the null that there is no relation between the two classifications equals 19.9; its p value is substantially below 1%.

The results provide strong support for our model, because they indicate that the paradox between pricing success and failure of portfolio separation is to be explained in terms the random perturbations that we introduced in standard general equilibrium theory.

[13] Two additional potential explanations that are based on institutional constraints in the experiments can readily be rejected. These are the following.

1. Presence of shortsale constraints. If we observe risk averse pricing it may be because risk neutral subjects (most subjects should be like that according to conventional wisdom) are shortsale constrained and therefore do not affect pricing anymore. We reject that possibility, because of the following. (i) We should see (the risk neutral) subjects only holding the security with the highest expected return, usually security A. Others (those who generated the CAPM pricing) hold a diversified portfolio. We don't observe this. See Figure 5. (ii) Our theory explains CAPM pricing because *everybody's* holdings equals optimal holdings (for a risk averse agent) plus a *mean zero* error. If a large number of subjects only hold a single security (the one with the highest expected return), the errors won't average out (risk neutral holdings are a systematic deviation from mean-variance optimal holdings). When we recover CAPM pricing, however, the econometric tests below will fail to reject that errors average out.

2. Bankruptcy rule. If subjects are really risk neutral, but take into account our bankruptcy rule (they get barred from further trading when negative two periods in a row, and thereby forego further opportunities to make money) then they should act in a risk averse way in early rounds (an implication of option theory). That is, pricing should reflect risk aversion in early rounds but this should disappear over time. We don't see that. See Figure 3.

When the χ^2 statistic is low, the estimates of b^N are small but always significantly different from zero. That is, we reject risk neutrality.

The noise in the estimation of the weighting matrix is apparent in the cases where standard errors could not be computed because the weighting matrix was not positive definite. While this does not make it impossible to compute the goodness-of-fit χ^2 statistics, it can be suspected to invalidate the distributional assumption (χ^2_1 distribution) underlying the tests. This suspicion is confirmed by the fairly high number of cases when the χ^2 statistic is in the far left tail of the χ^2_1 distribution. Of course, such problems should not come as a surprise: the distributional assumptions are based on large samples, whereas our samples had anywhere between 19 and 63 observations (N).

6 Conclusion

This paper is the first attempt to bring formal structural econometric testing to competitive financial markets experiments. We develop tests based on a modeling approach that allows for deviations in allocations from the tight predictions of general equilibrium theory. These deviations generate the stochastics behind the tests, in contrast with tests of asset pricing models on field data, where sampling error in the estimation of the payoff distribution provides the basis for statistical testing. The tests differ from field tests in one other respect: we test whether prices and allocations are consistent with theory, and not only whether prices alone satisfy some necesary conditions for equilibrium. That is, our tests could reject even if prices on their own are close to satisfying equilibrium restrictions (as prices often did in our experiments).

Our tests provide a metric to determine whether graphical evidence in favor of general equilibrium asset pricing theory is reliable. They are able to determine whether deviations from strict adherence to equilibrium pricing are econometrically significant.

More importantly, the tests verify our explanation for the price-allocation paradoxes (prices are close to equilibrium predictions yet allocations seem no closer than when prices violate equilibrium predictions). They answer the question whether deviations from strictly optimal holdings constitute noise that averages out across subjects, and hence, does not affect pricing in the limit.

Our modeling strategy explains the price-allocation paradox in both the certainty-equivalent experiments (where subjects were given the nonlinear payoff functions that would be expected utility indices in an uncertainty context) and the uncertainty experiments. In the latter, average risk aversion is tiny, but significant. The similarity between the statistical results from certainty-equivalent and uncertainty experiments confirms the theoretical speculation that the two sets of experiments are equivalent. Still, as already observed in [2, 3], and confirmed by the number of significant χ^2 statistics in this paper, markets are less likely to be in equilibrium when there is uncertainty. It remains to be explored why this is.

Altogether, we find that a general equilibrium model with standard risk aversion fits prices and allocations fairly well, modulo the demand perturbations we introduced. That is, there seems to be nothing wrong with the principles underlying general equilibrium theory (competitive-market equilibration). Nor do we have to resort to more general representations of attitudes towards risk and uncertainty, although our tests encompass cases where individual risk aversion changes randomly over time, which is nonstandard.

References

[1] Bossaerts, P. and C. Plott [2001], "The CAPM In Thin Experimental Financial Markets," *Journal of Economic Dynamics and Control* (forthcoming).

[2] Bossaerts, P. and C. Plott [1999], "Basic Principles of Asset Pricing Theory: Evidence from Large-Scale Experimental Financial Markets," Caltech working paper.

[3] Bossaerts, P., C. Plott and W. Zame, "Prices And Allocations In Financial Markets: Theory and Evidence," Caltech working paper.

[4] Gallant, A.R. [1987], *Nonlinear Statistical Models*, Wiley.

[5] Hansen, L. P. and K. J. Singleton [1982], "Generalized Instrumental Variables Estimation Of Nonlinear Rational Expectations Models," *Econometrica* 50, 1269-86.

[6] Heaton, J. and D. Lucas [2000], "Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk," *Journal of Finance* 55, 1163-1198.

[7] Judd, K. and S.M. Guu [2000], "Bifurcation Methods For Asset Market Equilibrium Analysis," Stanford University, Hoover Institute working paper.

[8] Rabin, M. [2000], "Risk Aversion and Expected-Utility Theory: A Calibration Theorem," *Econometrica* 68, 1281-92.

[9] Rubinstein, M. [1999], "Rational Markets: Yes or No?" Notes from a Debate, Berkeley Program in Finance.

Table 1: Experimental Design Data – Certainty Experiments.

Experiment	Subject Category (Number)	b_n [a]	Signup Reward (franc)	Endowments A	Endowments B	Endowments Notes	Cash (franc)	Loan Repayment (franc)	Exchange Rate \$/franc
000320	8	$2.3.10^{-3}$	0	6	2	0	400	1675	0.06
	9	$2.8.10^{-4}$	0	2	6	0	400	2020	0.09
	8	$1.5.10^{-4}$	0	6	2	0	400	2120	0.09
	8	$1.5.10^{-4}$	0	2	6	0	400	2015	0.09
000509	14	$2.3.10^{-3}$	0	2	8	0	400	2340	0.06
	13	$2.8.10^{-4}$	0	8	2	0	400	2480	0.06
	14	$1.5.10^{-4}$	0	2	8	0	400	2365	0.06
000511	15	$2.3.10^{-3}$	0	2	8	0	400	2340	0.06
	14	$2.8.10^{-4}$	0	8	2	0	400	2480	0.06
	15	$1.5.10^{-4}$	0	2	8	0	400	2365	0.06

[a] Coefficient in the payoff function (29) assigned to subject.

Table 2: Structural Econometric Tests – Certainty Experiments.

Experiment	Statistic	Periods							
		1	2	3	4	5	6	7	8
000320	χ^2_1	0.1	0.0	0.0	0.0	0.2	0.0	0.0	0.2
	p-level	.76	.94	.85	.87	.71	.86	.86	.61
	b^N $(*10^{-3})^a$	0.2	0.2	0.3	0.2	0.4	0.3	0.3	0.4
	s.e. $(*10^{-3})$	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
000509	χ^2_1	5.0	1.3	0.6	0.0	0.0	0.1	0.0	0.0
	p-level	.02	.17	.39	.89	.81	.79	.92	.95
	b^N $(*10^{-3})^b$	6.6	4.3	2.5	0.7	0.9	0.9	0.8	0.2
	s.e. $(*10^{-3})$	0.4	0.4	0.3	0.2	0.2	0.3	0.3	0.3
000511	χ^2_1	16.3	12.0	7.7	6.6	4.1	2.4	1.8	0.4
	p-level	.00	.00	.00	.01	.03	.08	.12	.57
	b^N $(*10^{-3})^c$	2.8	2.3	1.7	1.4	1.3	1.1	0.9	0.7
	s.e. $(*10^{-3})$	0.2	0.2	0.3	0.2	0.2	0.2	0.2	0.2

[a] Theoretical value: $0.2 * 10^{-3}$

[b] Theoretical value: $0.3 * 10^{-3}$.

[c] Theoretical value: $0.2 * 10^{-3}$.

Table 3: Experimental design data – uncertainty experiments.

Experiment	Subject Category (Number)	Signup Reward (franc)	Endowments A	B	Notes	Cash (franc)	Loan Repayment (franc)	Exchange Rate $/franc
981007	30	0	4	4	0	400	1900	0.03
981116	23	0	5	4	0	400	2000	0.03
	21	0	2	7	0	400	2000	0.03
990211	8	0	5	4	0	400	2000	0.03
	11	0	2	7	0	400	2000	0.03
990407	22	175	9	1	0	400	2500	0.03
	22	175	1	9	0	400	2400	0.04
991110	33	175	5	4	0	400	2200	0.04
	30	175	2	8	0	400	2310	0.04
991111	22	175	5	4	0	400	2200	0.04
	23	175	2	8	0	400	2310	0.04

Table 4: Structural Econometric Tests – Uncertainty Experiments.

Experiment	Statistic	Periods							
		1	2	3	4	5	6	7	8
981007	χ_1^2	11.6	0.7	33.9	11.0	5.3	3.2		
	p-level	.00	.33	.00	.00	.01	.04		
	b^N (*10^{-3})	0.7	0.6	1.4	1.1	1.2	1.1		
	s.e. (*10^{-3})	0.1	0.1	0.1	0.2	0.1	0.1		
981116	χ_1^2	15.6	0.5	0.6	2.4	1.4	18.4		
	p-level	.00	.42	.38	.08	.17	.00		
	b^N (*10^{-3})	1.6	1.1	0.8	1.0	1.8	2.3		
	s.e. (*10^{-3})	0.1	0.2	0.1	0.1	*[a]	0.1		
990211	χ_1^2	2.3	3.0	3.5	15.4	1.1	12.7	0.0	
	p-level	.08	.05	.04	.00	.22	.00	.93	
	b^N (*10^{-3})	1.3	1.4	1.9	5.5	1.1	2.4	-0.2	
	s.e. (*10^{-3})	0.0	*	*	*	*	0.0	0.2	
990407	χ_1^2	1.6	0.1	1.6	5.7	12.9	0.0	1.5	21.3
	p-level	.14	.73	.14	.01	.00	.87	.15	.00
	b^N (*10^{-3})	0.5	0.5	0.2	-0.2	-0.6	0.3	0.2	-1.0
	s.e. (*10^{-3})	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.1
991110	χ_1^2	32.4	13.5	7.1	1.9	0.7	1.8		
	p-level	.00	.00	.00	.11	.33	.12		
	b^N (*10^{-3})	2.9	2.4	1.9	1.2	1.1	1.5		
	s.e. (*10^{-3})	0.1	*	0.1	0.2	0.2	0.2		
991111	χ_1^2	0.9	0.2	24.6	10.4	5.2	5.7	6.3	4.1
	p-level	.27	.72	.00	.00	.01	.01	.01	.03
	b^N (*10^{-3})	0.4	0.3	1.9	1.4	1.0	0.9	1.5	1.3
	s.e. (*10^{-3})	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1

[a]Asterisk denotes that the weighting matrix was not positive definite, and hence, standard errors could not be computed.

Table 5: Two-Way Classification: Structural Econometric Tests Against Distance From CAPM Pricing – Uncertainty Experiments.

		Distance From CAPM Pricing[a]		
		Farthest 1/3	Medium	Closest 1/3
	$p > .05$	1	4	13
Structural Test[b]	$p \in (.01, .05]$	2	4	3
	$p <= .01$	8	6	0

[a]Measured as the difference between the Sharpe ratio of the market portfolio and the maximum Sharpe ratio at the end of each period; distances are compared per experiment. See Figure 4.

[b]p-level of χ^2 statistics in Table 4.

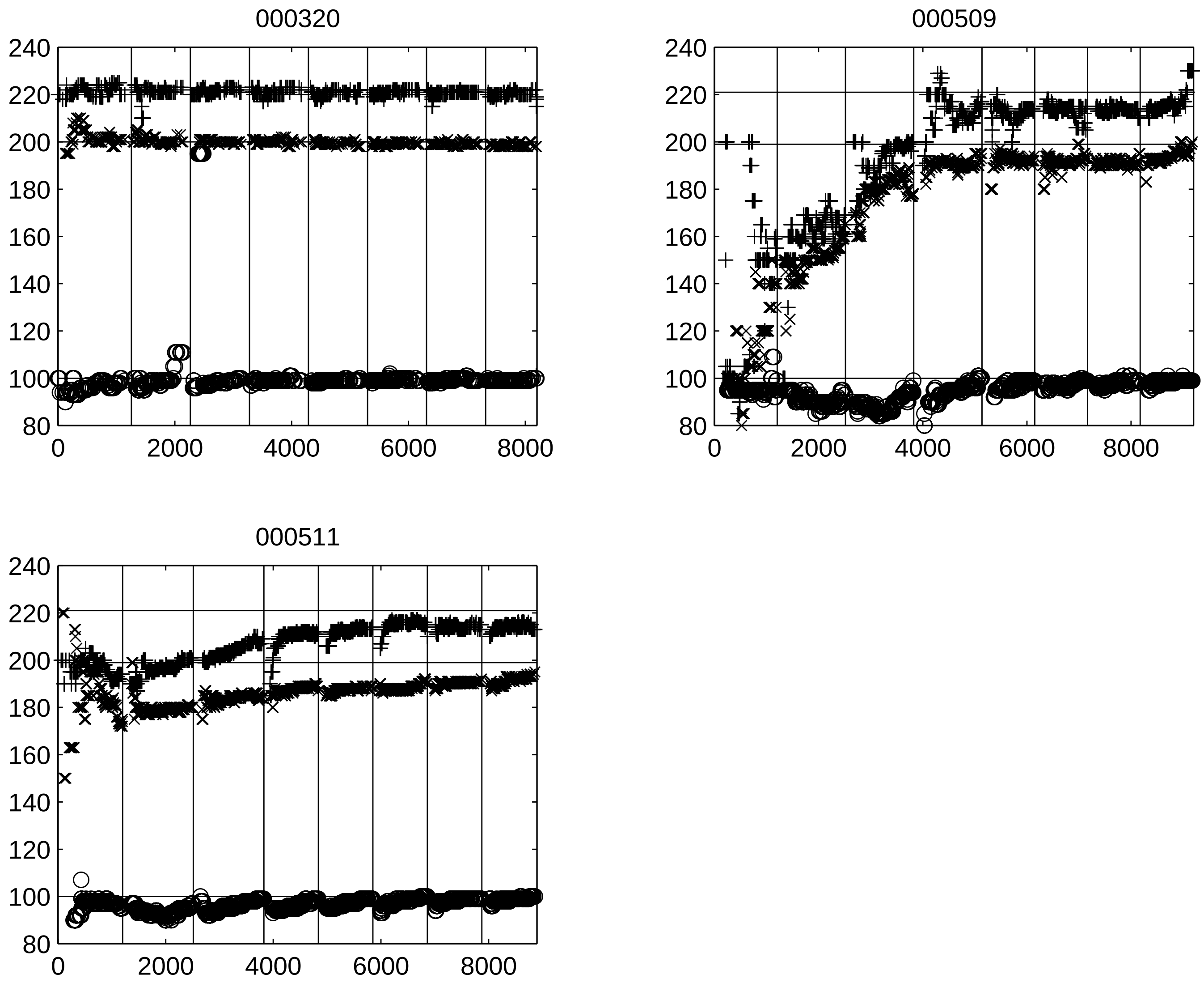

Figure 1: Evolution of the transaction prices in three certainty-equivalent experiments. Time (horizontal axis) is measured in seconds; Prices (vertical axis) in francs. Vertical bars delineate periods. Horizontal lines indicate theoretical equilibrium prices. Commodity A: +; commodity B: x; Notes: o.

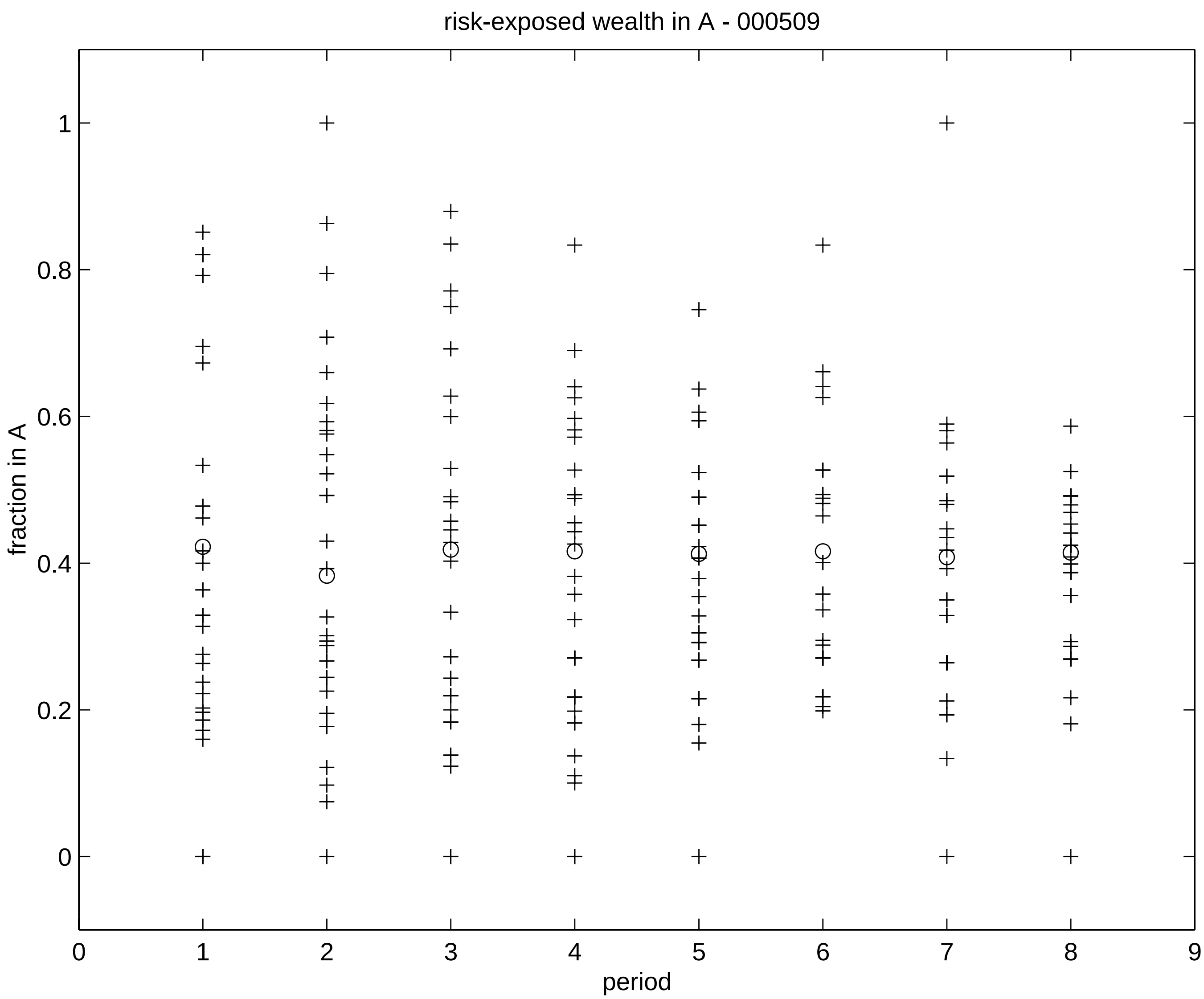

Figure 2: End-of-period holdings in 000509. Horizontal axis: periods. Vertical axis: proportion of A in subjects' total wealth allocated to A and B (pluses). Subjects that did not allocate any wealth to A and B are plotted as holding the market portfolio of A and B (the circles in the plot).

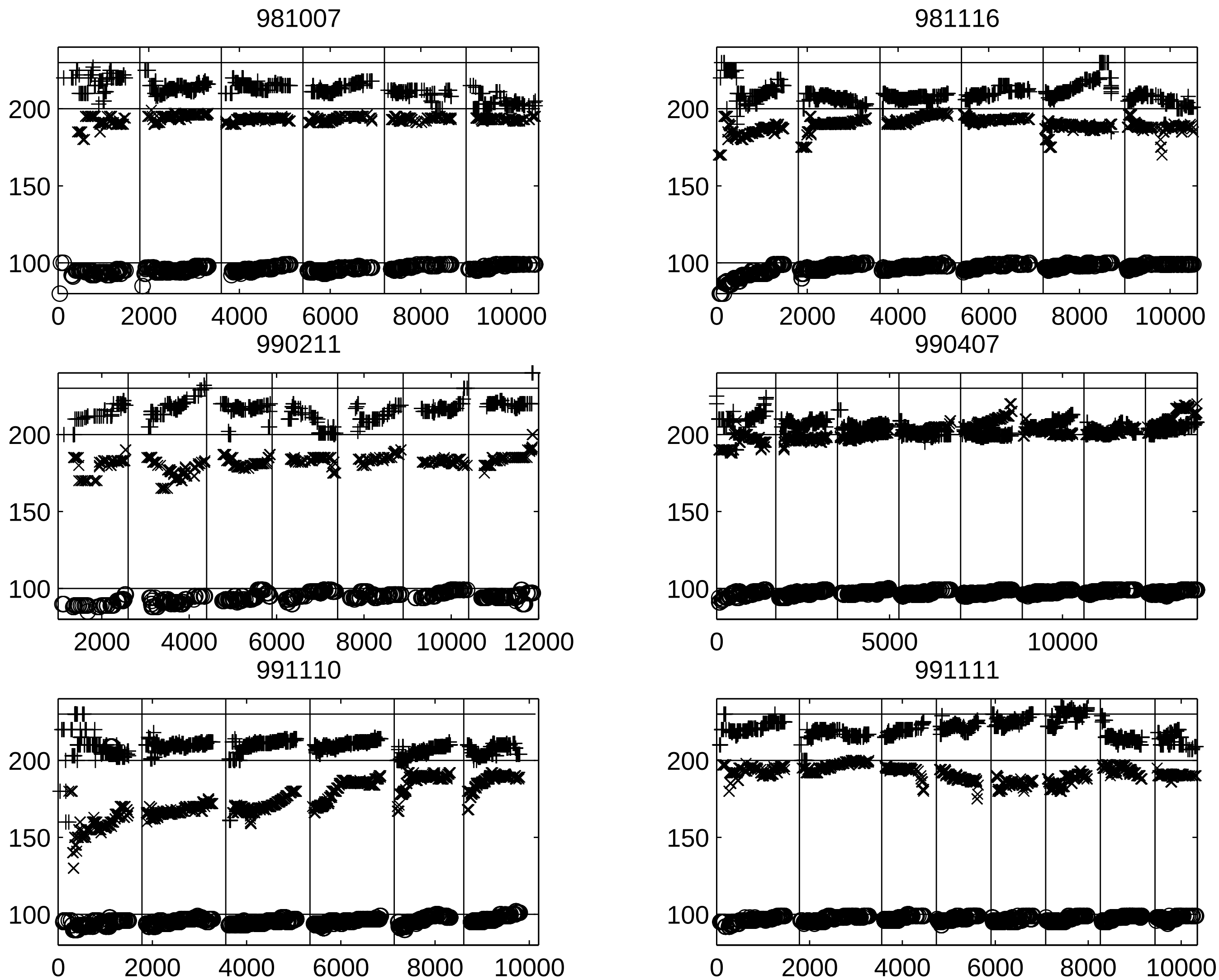

Figure 3: Evolution of the transaction prices in six uncertainty experiments. Time (horizontal axis) is measured in seconds; Prices (vertical axis) in francs. Vertical bars delineate periods. Horizontal lines indicate expected payoffs. Security A: +; Security B: x; Notes: o.

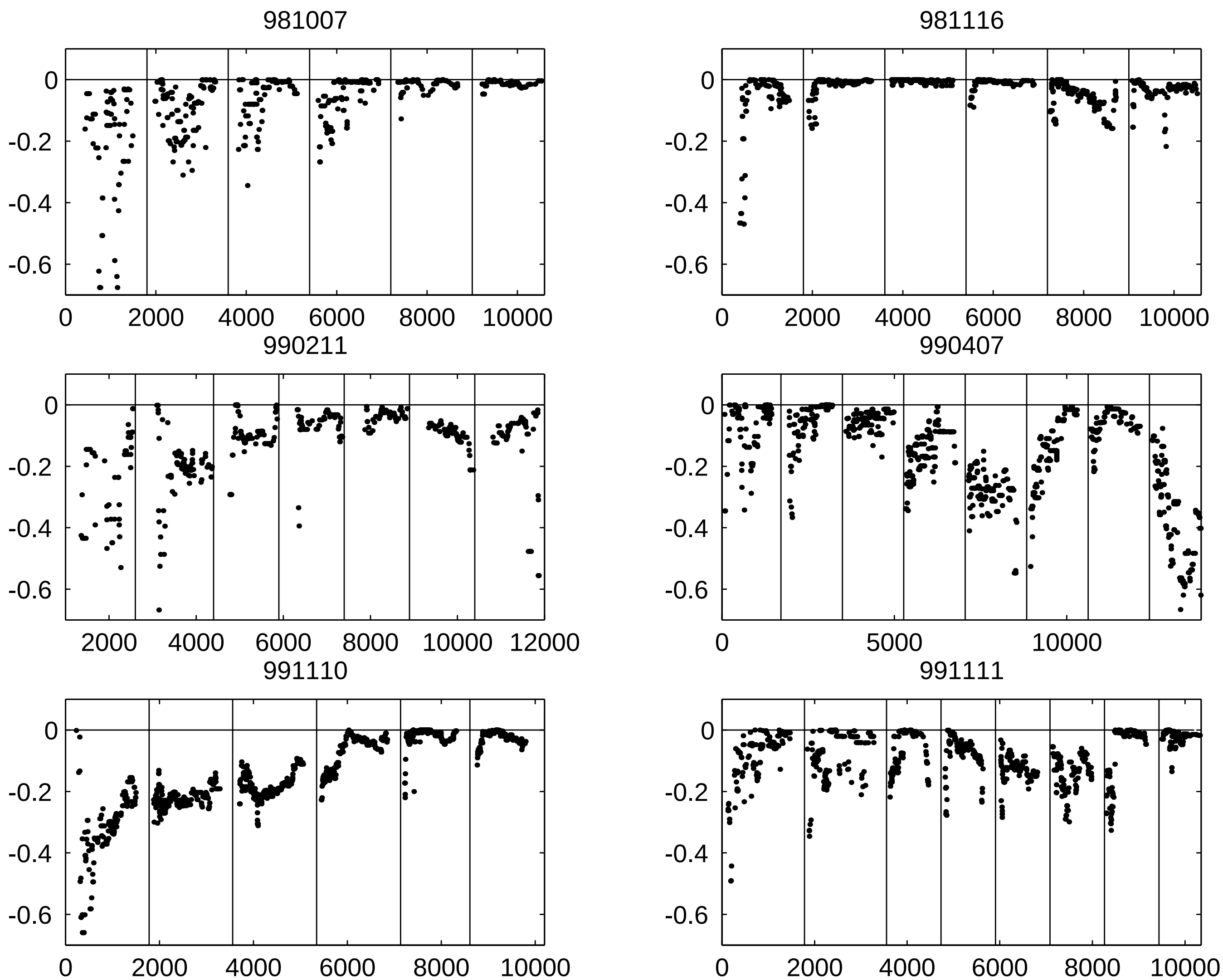

Figure 4: Evolution of the difference between the Sharpe ratio of the market and the maximum Sharpe ratio. transaction prices in six uncertainty experiments. Time (horizontal axis) is measured in seconds. Sharpe ratio differences are on the vertical axis. Vertical bars delineate periods. Horizontal line indicates the value that is needed for prices to be in line with the CAPM.

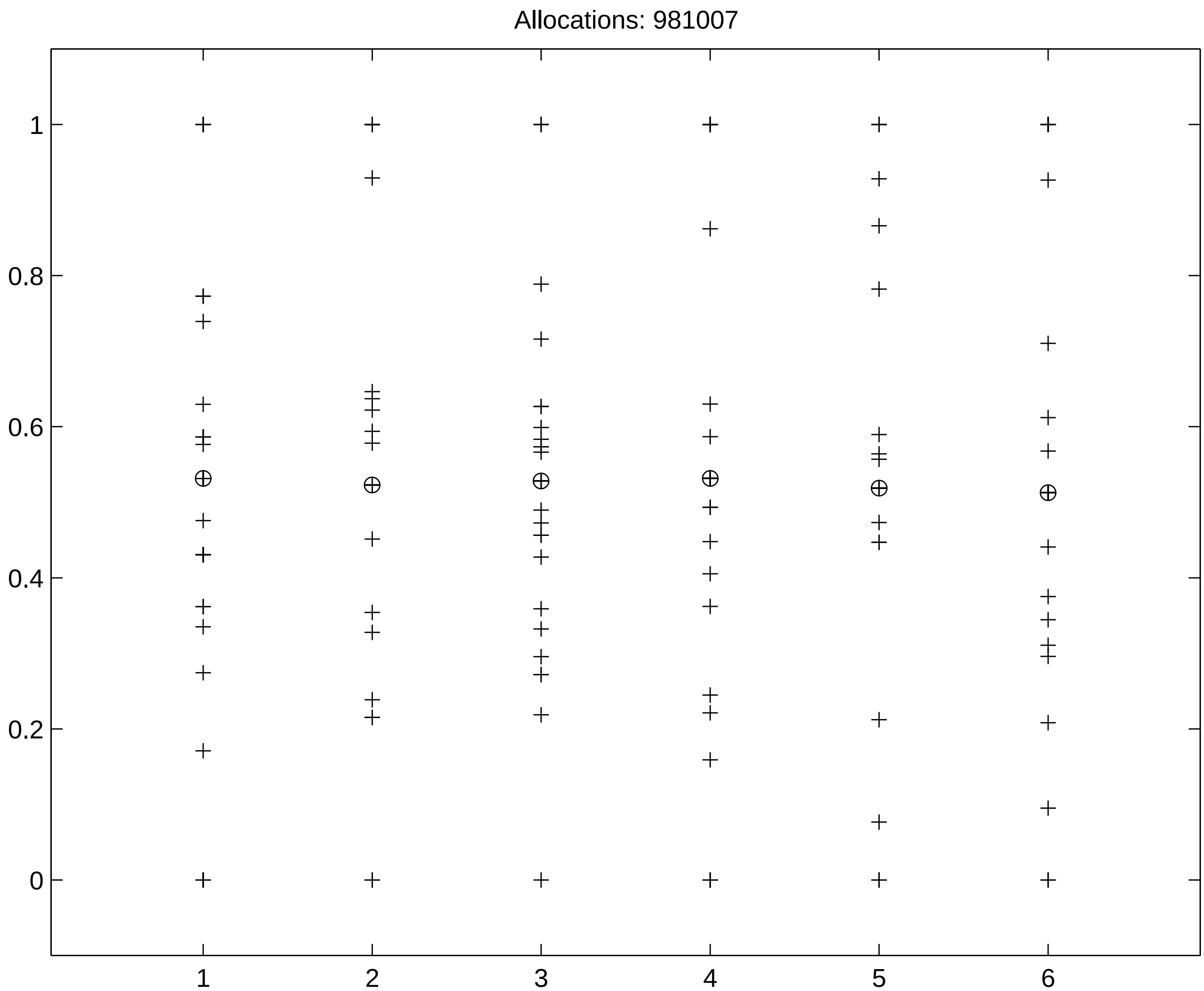

Figure 5: End-of-period holdings in 981007. Horizontal axis: periods. Vertical axis: proportion of subjects' risk-exposed wealth in security A (pluses). Subjects that did not allocate any wealth to risky securities are plotted as holding the market portfolio of risky securities (the circles in the plot).