UNIL | Université de Lausanne
Département de finance

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

FINANCE RESEARCH SEMINAR SUPPORTED BY UNIGESTION

"Externalities as Arbitrage"

Prof. Benjamin HEBERT
(Stanford University, Graduate School of Business)

Abstract

Regulations on financial intermediaries can create apparent arbitrage opportunities. Intermediaries are unable to fully exploit these opportunities due to regulation, and other agents are unable to exploit them at all due to limited participation. Does the existence of arbitrage opportunities imply that regulations are sub-optimal? No. I develop of general equilibrium model, with financial intermediaries and limited participation by other agents, in which a constrained-efficient allocation can be implemented with asset prices featuring arbitrage opportunities. Absent regulation, there would be no arbitrage; however, allocations would be constrained-inefficient, due to pecuniary externalities and limited market participation. Optimal policy creates arbitrage opportunities whose pattern across states of the world reflects these externalities. From financial data alone, we can construct perceived externalities that would rationalize the pattern of arbitrage observed in the data. By examining these perceived externalities, and comparing them to the stated goals of regulators, as embodied in the scenarios of the stress tests, we can ask whether regulations are having their intended effect. The answer, in recent data, is no.

Friday, January 12, 2018, 10:30-12:00
Room 126, Extranef building at the University of Lausanne

Externalities as Arbitrage*

Benjamin Hébert (Stanford GSB)

December 29, 2017

PRELIMINARY AND INCOMPLETE; PLEASE DO NOT CITE

Abstract

Regulations on financial intermediaries can create apparent arbitrage opportunities. Intermediaries are unable to fully exploit these opportunities due to regulation, and other agents are unable to exploit them at all due to limited participation. Does the existence of arbitrage opportunities imply that regulations are sub-optimal? No. I develop of general equilibrium model, with financial intermediaries and limited participation by other agents, in which a constrained-efficient allocation can be implemented with asset prices featuring arbitrage opportunities. Absent regulation, there would be no arbitrage; however, allocations would be constrained-inefficient, due to pecuniary externalities and limited market participation. Optimal policy creates arbitrage opportunities whose pattern across states of the world reflects these externalities. From financial data alone, we can construct perceived externalities that would rationalize the pattern of arbitrage observed in the data. By examining these perceived externalities, and comparing them to the stated goals of regulators, as embodied in the scenarios of the stress tests, we can ask whether regulations are having their intended effect. The answer, in recent data, is no.

*Acknowledgements: I would like to thank Eduardo Dávila, Darrell Duffie, Emmanuel Farhi, Anton Korinek, Arvind Krishnamurthy, Hanno Lustig, and Jesse Schreger for helpful comments. I would particularly like to thank Wenxin Du for inspiring my work on this topic and for help understanding the data sources. All remaining errors are my own.

1 Introduction

Following the recent financial crisis, several apparent arbitrage opportunities have appeared in financial markets. These arbitrage opportunities, such as the gap between the federal funds rate and the interest on excess reserves (IOER) rate, or violations of covered interest rate parity, are notable in part because they have persisted for years after the peak of the financial crisis. Many authors have argued that the regulatory changes which occurred in response to the financial crisis have enabled these arbitrages to exist and persist.

If regulatory changes caused these arbitrage opportunities, does that imply that there is something wrong with the regulations? This paper addresses the question of the welfare implications of observing arbitrage. The paper first considers a general equilibrium model with incomplete markets and two types of agents, households and intermediaries. I point out that, just as a lack of arbitrage does not imply efficiency, the presence of arbitrage does not imply inefficiency. However, the underlying motivation for regulation in the model is incomplete markets and pecuniary externalities, and under an optimal policy the patterns of arbitrage across various assets are determined by whether those assets have payoffs in states with large or small externalities. That is, if regulation is working correctly, there should be a tight relationship between the externalities the social planner is correcting and the arbitrage on financial assets. Consequently, by observing asset prices, we can construct a set of "perceived externalities" that would justify the observed pattern of arbitrage.

The main contribution of the paper is to conduct this exercise. Because externalities can vary across states of the world, and markets are incomplete, there will necessarily be fewer assets than externalities, and hence it will not be possible to recover a unique set of perceived externalities. However, making an analogy to the projection of stochastic discount factors on to the space of returns, I show that it is possible to uniquely recover an "externality-mimicking portfolio." This returns of this portfolio across states of the world are a set externalities that would justify the observed pattern of arbitrage, and all other externalities that would also justify the observed pattern of arbitrage are more volatile than these returns.

Using data on interest rates, foreign exchange spot and forward rates, and foreign exchange options, I construct the externality-mimicking portfolio. The weights in this portfolio are entirely a function of asset prices; there are no econometrics involved.

The returns of this portfolio represent an estimate of the externalities the social planner perceives when considering transfers of wealth between the households and intermediaries in various states of the world. When the returns are positive, the planner perceives positive externalities when transferring wealth from intermediaries to households. I then consider the returns of this portfolio in the “stress tests” conducted by the federal reserve. I argue that these tests are statements about when the fed would like intermediaries to have more wealth, and as a result the returns of the externality-mimicking portfolio should be negative in the stress test. However, I find for the stress tests conducted at the end of 2014 and 2015 that the returns are positive, indicating an inconsistency between the stated goals of the fed and the actual effects of regulation. Along these liens, the procedure developed in this paper could be used as a tool by regulators to better understand whether the regulations they impose are having the desired effect.

This paper brings together and builds on several strands of literature. The theoretical framework builds on general equilibrium with incomplete markets (GEI) models of the sort studied by Geanakoplos and Polemarchakis [1986]. In particular, the definition of constrained inefficiency and the result that, absent regulation, the economy is constrained efficient follow from that paper. The model I develop specializes the standard GEI model in several respects. First, I assume that there are two classes of agents, households and intermediaries, who have different degrees of access to markets. Intermediaries have a complete market amongst themselves, and can also trade with any household. Households, on the other hand, cannot trade with each other, only via intermediaries, and face incomplete markets in their trades with intermediaries. These constraints, which I will refer to as limited participation constraints for the households, are very similar to standard incomplete markets constraints, except that they also allow a social planner to implement any feasible allocation entirely by regulating intermediaries. That is, the intermediaries can serve as a sort of “central point” for regulation, perhaps minimizing unmodeled costs of implementing any particular regulation. Second, I will assume that all intermediaries have the same homothetic utility functions. As a result, moving wealth between intermediaries will not change the total demand for goods, and hence will not generate pecuniary externalities. It follows that it is without loss of generality for a planner to implement an allocation without regulating trades between intermediaries. In this case, the intermediaries aggregate, and the planner is only concerned with total intermediary wealth. Put another way, the planner is only concerned with macro-prudential, as opposed to micro-prudential, regulation.

My emphasis on macro-prudential regulation, and the notation I employ, are shared with the work of Farhi and Werning [2016]. Furthermore, the connection between arbitrage and externalities depends on the asset market structure I impose, but not on the source of the externalities. Building on those authors' work, I show in an extension that using borrowing constraints or price rigidities, instead of or in addition to incomplete markets, would lead to the same relationship between arbitrage and externalities. A key difference between this paper and the work of Farhi and Werning [2016], and also the discussion of pecuniary externalities in Dávila and Korinek [2017], is my focus on an implementation of the constrained efficient allocation using borrowing constraints, rather than agent-state-good-specific taxes. Studying this implementation is both realistic, in the sense that regulation on banks takes this form, and it allows me to relate externalities and asset prices under the optimal policy, enabling the empirical exercises that are a focus of this paper.

Separating agents into multiple types, and enforcing limited participation for some types, is a common strategy in the literature on arbitrage (surveyed by Gromb and Vayanos [2010]). As in much of this literature, arbitrage arises in my model because of limited participation (for households) and constraints on trading (for intermediaries). My emphasis on welfare and optimal policy in the presence of arbitrage follows the spirit Gromb and Vayanos [2002]. My baseline model uses the pecuniary externalities and incomplete markets constrained inefficiency of Geanakoplos and Polemarchakis [1986], rather than the collateral constraint inefficiency of Gromb and Vayanos [2002], but the work of Farhi and Werning [2016] and Dávila and Korinek [2017] shows that this distinction is not essential. Again, the key difference between my paper and this literature is my emphasis on the connection between asset prices and externalities, and my attempt to use observed arbitrages to quantify these externalities.

However, a second difference concerns the choice of arbitrages to study. Going back to Shleifer and Vishny [1997], a central theme of this literature has been a focus on arbitrages, like the value of closed end funds relative to their constituent stocks, for which convergence is guaranteed only at a distant horizon, if ever. For these arbitrages, a central concern for any potential arbitrageur is that prices might move against them over short or medium horizons. Combined with certain other kinds of frictions, this "mark-to-market" risk might explain why the arbitrage can persist.

In contrast, my model and empirical exercises are focused exclusively on arbitrages that are guaranteed to converge over a short horizon. The interest on reserves/fed funds arbitrage documented by Bech and Klee [2011] converges daily, and the covered interest parity arbitrages documented by Du et al. [2017] can also be done at, for example, monthly frequencies.[1] As a result, the sort of mark-to-market risk central to many models of limits to arbitrage is absent from my model, and arbitrage exists only in the presence of regulation. This should not be taken as a rejection of those limits to arbitrage models, but rather as an attempt to focus on what we as economists can learn from the existence of arbitrages that are induced by regulation. Note also that, while CIP violations have been documented at a variety of horizons, but this does not imply that the underlying cause is the same at all horizons. For example, Andersen et al. [2017] argue that one-year horizon CIP violations are best explained by a debt-overhang type mechanism, but this mechanism because quantitatively small as the horizon (and hence default probability) shrinks.

There are also a number of papers that present limits-to-arbitrage type theories in a setting specific to covered interest parity violations (Amador et al. [2017], Ivashina et al. [2015], Liao et al. [2016]). These papers take borrowing constraints on intermediaries as given, rather than treating them as a policy instrument and considering optimal policy. The work of Ivashina et al. [2015] is particularly related, as it emphasizes the connections between covered interest parity violations and non-financial outcomes.

Unfortunately, there are only a small set of "arbitrageable" assets for which data is available. In the context of the model, an asset is arbitrageable if it can be traded by households, and if we (as economists observing the economy) can also find the price of a replicating portfolio of assets from the intermediary-only market. As a result of the lack of a complete set of arbitrageable assets, we cannot recover a unique set of externalities that would rationalize an observed pattern of arbitrage (the "perceived externalities"). However, we can construct a unique[2] projection of the externalities

[1]Bech and Klee [2011] describe the IOER-Fed Funds arbitrage as arising from market power by banks with respect to fed funds trades with the GSEs. Malamud and Shrimpf [2017] adopt a related perspective on CIP violations. Duffie and Krishnamurthy [2016] argue that a mix of the shadow costs of regulation and market power are responsible for the difference between IOER and various money market rates. Bräuning and Puria [2017] provide evidence on the significant impact of regulation. I adopt the view that these shadow costs are central in the federal funds market, whereas market power plays a larger role in, for example, bank deposit markets.

[2]The definition of the externalities references a probability measure, and the projection I employ

on to the space of returns. Formally, the exercise is analogous to the projection of a stochastic discount factor on the space of returns, as developed by Hansen and Richard [1987]. The procedure produces an "externality-mimicking portfolio," the returns of which are an estimate of the perceived externalities. Building on this analogy, I will also show that the standard deviation of any pattern of externalities consistent with the observed arbitrage is greater than the same standard deviation for this estimate. Moreover, the latter standard deviation is proportional to the "Sharpe ratio due to arbitrage" (a concept I will define) of the externality-mimicking portfolio, and this "Sharpe ratio due to arbitrage" is maximal for the externality-mimicking portfolio. These results are the analog of the bound of Hansen and Jagannathan [1991].

Armed with this empirical procedure, I construct externality-mimicking portfolios at daily frequency. The arbitrages I use to construct the portfolio are the fed funds/IOER arbitrage and the covered interest parity arbitrages for the dollar-euro and dollar-yen currency pairs. These arbitrages are constructed from daily data on interest rates and both spot and forward exchange rates.[3] The FF-IOER arbitrage serves as a sort of "risk-free arbitrage," meaning that it is the difference of two risk-free rates. The other two arbitrages are "risky arbitrages", meaning that they can be thought of as law-of-one-price violations on a risky payoff. To compute the weights in the externality mimicking portfolio, I require not only estimates of the arbitrage, but also a covariance matrix, under the intermediaries' risk-neutral measure, of the risky assets for which there is a law-of-one-price violation. Fortunately, because the assets in question are currency pairs, the entire risk-neutral covariance matrix can be extracted from options prices.

Having constructed the externality-mimicking portfolio, I then address two empirical questions. First, given that I am projecting the externalities on to a low-dimensional space of returns, there is some question about whether the portfolio return truly mimics the externalities. I perform a sort of "out-of-sample" test, predicting the arbitrage on the dollar-pound currency pair using the externality-mimicking portfolio weights and implied volatilities for the pound-euro and pound-yen currency pairs. The high R^2 of this predictive exercise suggests that the projection is reasonable. Second, I study the returns of the externality-mimicking portfolio in the "severely adverse" scenarios from the Federal Reserve's stress tests. These scenar-

is under this particular probability measure. Uniqueness here means unique given this measure.

[3] There is some dispute on the best way to do this (Du et al. [2017], Rime et al. [2017]).

ios are designed to reflect situations in which the Fed would like banks to have more wealth. Given that, we should expect (under my sign conventions) that the externality-mimicking portfolio has a negative return in this scenarios.

However, I find for several of the stress tests that the returns are close to zero, or even positive. This result occurs because, in the stress test scenarios, then yen appreciates and the euro depreciates relative to the dollar, but the dollar-yen and dollar-euro CIP violations have the same sign. This suggests an inconsistency in the regulatory regime. I speculate that this inconsistency arises from the joint effects of customer demand and leverage constraints, and argue that by constructing the externality mimicking portfolio and considering its returns, regulators can assess whether their regulations are having the intended effects.

Section 2 introduces the GEI framework, and describes its general efficiency properties. Section 3 relates the wedges in that framework to arbitrage on assets. Section 4 describes the projection used to construct the externality-mimicking portfolio. Section 5 presents my empirical results. Section 6 discusses extensions to the model, and section §7 concludes.

2 General Equilibrium with Intermediaries

In this section, I introduce financial intermediaries into an otherwise-standard incomplete markets, general equilibrium endowment economy. The notation and setup of the model follows Farhi and Werning [2016]. The model has two periods, time zero and one. At time one, a state $s \in S_1$ is determined. The state at time zero is s_0, and the set of all states is $S = S_1 \cup \{s_0\}$. The goods available in each states are denoted by the set J_s.[4] Households $h \in H$ maximize expected utility,

$$\sum_{s \in S} U^h(\{X^h_{j,s}\}_{j \in J_s}; s),$$

where $U^h(\{X^h_{j,s}\}_{j \in J_s}; s)$ is the utility of household h in state s, inclusive of the household's rate of time preference and the probability the household places on state s. I will assume non-satiation for at least one good in each state, implying that each household places non-zero probability on each state in S_1.

[4]Similar to Farhi and Werning [2016], separating contingent commodities into states and goods available in each state will give meaning to the financial structure described below.

In each state $s \in S$, household $h \in H$ has an endowment of good $j \in J_s$ equal to $Y^h_{j,s}$. In state s_0, the household might also receive a transfer. The set of securities available in the economy, A, has securities which offer payoffs $Z_{a,s}$ for security $a \in A$ in state $s \in S$. Let D^h_a denote the quantity of security a purchased or sold by household h, and let Q_a be the "ex-dividend" price at time zero (i.e. under the convention that $Z_{a,s_0} = 0$). In state s, the household's income, which it can use to purchase goods, is

$$I^h_s = T^h \mathbf{1}(s = s_0) + \sum_{j \in J_s} P_{j,s} Y^h_{j,s} + \sum_{a \in A} D^h_a (Z_{a,s} - Q_a \mathbf{1}(s = s_0)),$$

and the household's budget constraint is

$$\sum_{j \in J_s} P_{j,s} X^h_{j,s} \leq I^h_s.$$

The constraints on households' asset positions are summarized by

$$\Phi^h(\{D^h_a\}_{a \in A}) \leq \vec{0},$$

where Φ^h is a vector-valued function, convex in D^h. These constraints implement households limited participation in markets, in a manner that I will describe below.

Using this wealth and prices, the standard indirect utility function is

$$V^h(I^h_s, \{P_{j,s}\}_{j \in J_s}; s) = \max_{\{X^h_{j,s} \in \mathbb{R}^+\}_{j \in J_s}} U^h(\{X^h_{j,s}\}_{j \in J_s}; s)$$

subject to

$$\sum_{j \in J_s} P_{j,s} X^h_{j,s} \leq I^h_s.$$

The portfolio choice problem is

$$\max_{\{D^h_a \in \mathbb{R}\}_{a \in A}} \sum_{s \in S} V^h(I^h_s, \{P_{j,s}\}_{j \in J_s}; s)$$

subject to the budget constraint defined above and the asset allocation constraint.

Households are distinct from intermediaries, the other type of agent in the economy. I will use $i \in I$ to denote a particular intermediary. Intermediaries are like households (in the sense that all of the notation above applies, with some $i \in I$ in the place of an $h \in H$), except that they face different constraints on their portfolio choices, and they share a common set of utility functions. In particular, households are constrained to trade only with intermediaries, but intermediaries can trade with both households and other intermediaries. Intermediaries also share a common set of homothetic utility functions; that is, $U^i(\cdot; s) = U^{i'}(\cdot; s)$ for all $s \in S$ and $i, i' \in I$, and $U^i(\cdot; s)$ is homothetic for all $s \in S$. The assumption of a common, homothetic (but state-dependent) utility function ensures that redistributing wealth across intermediaries does not influence relative prices in goods markets, and hence is not particularly useful from a planner's perspective.

The constraint that households can trade only with intermediaries, but not each other, can be implemented using this notation in the following way. The set of assets, A, is a superset of the union of disjoint sets $\{A^h\}_{h \in H}$, denoting trades with household h. For a given household h, the function Φ^h implements the requirement that, for all $a \in A \setminus A^h$, $D^h_a = 0$. To be precise, if $a \in A \setminus A^h$ and $D^h_a \neq 0$, then there exists an element of $\Phi^h(D^h)$ strictly greater than zero. For simplicity, I assume there are no other constraints on household's portfolio choices, aside from these constraints.

The set of assets also includes assets that cannot be traded by any household. Define $A^I = A \setminus (\cup_{h \in H} A^h)$ as the set of securities tradable only by intermediaries. To simplify the exposition, I will assume this set includes a full set of Arrow securities, although nothing depends on this. I will say that household h is a limited participant if the span of $\{Z_{a,s}\}_{a \in A^h}$ is a strict subspace of the span of $\{Z_{a,s}\}_{a \in A}$, the latter of which is the space of all possible payoffs. The limited participation of households in this sense is crucial, in the model, for generating arbitrage. To be precise, the model will generate violations of the law of one price, by providing conditions under which a security $a \in A^h$, for some $h \in H$, will have a price that is different than the price of its replicating portfolio of Arrow securities in A^I. Law-of-one-price violations can also occur between the assets traded with two households (h and h'), although the paper will not emphasize these violations. There will not, in general, be arbitrage without law-of-one-price violations (i.e. getting something for nothing), due to the assumption of non-satiation.

For arbitrage between the asset market A^I and the asset market A^h to exist, intermediaries must face financial constraints. The approach of this paper, in contrast to the much of the existing literature on arbitrages, is to assume that the constraints faced by intermediaries are induced entirely by government policy. That is, I will assume that the Φ^i functions are the government's policy instrument; in contrast, the Φ^h functions are assumed to be exogenous. The assumption that the Φ^h are exogenous is without loss of generality. Because all trades are intermediated, and the government can constrain intermediaries, the government can effectively control all of the trades in the economy, and therefore implement any allocation that could be implemented with agent-specific taxes (as in Farhi and Werning [2016]).

The notion of equilibrium is standard:

Definition 1. An equilibrium is a collection of consumptions $X^h_{j,s}$ and $X^i_{j,s}$, goods prices $P_{j,s}$, asset positions D^h_a and D^i_a, transfers T^h and T^i, and asset prices Q_a such that:

1. Households and intermediaries maximize their utility over consumption and asset positions, given goods prices and asset prices, respecting the constraints that consumption be weakly positive and the constraints on their asset positions

2. Goods markets clear: for all $s \in S$ and $j \in J_s$,

$$\sum_{h \in H}(X^h_{j,s} - Y^h_{j,s}) = \sum_{i \in I}(X^i_{j,s} - Y^i_{j,s})$$

3. Asset markets clear: for all $a \in A$,

$$\sum_{h \in H} D^h_a + \sum_{i \in I} D^i_a = 0$$

4. The government's budget constraint balances,

$$\sum_{h \in H} T^h + \sum_{i \in I} T^i = 0$$

The definition of equilibrium presumes price-taking by households and intermediaries. Absent government constraints, each household h can trade with everyone intermediary, and the price of assets $a \in A^h$ will be pinned down by competition between intermediaries. The equilibrium definition supposes that this will continue to be case, even if the government places asymmetric constraints on intermediaries– for example, by granting a single intermediary a monopoly over trades with a particular household. In this case, it is as if the household had all of the bargaining power. In any case, such a policy is unlikely to optimal, and will never be the unique optimum.

I next describe a planner's problem for this economy. I assume that the planner is unable to redistribute resources ex-post (doing so would allow the planner to circumvent limited participation). Instead, in the spirit of Geanakoplos and Polemarchakis [1986], I will allow the planner to trade in asset markets on behalf of agents, trading for each agent only in markets she can participate in, to maximize a weighted sum of the agents' indirect utility functions. The planner solves

$$\max_{\{D_a^h\in\mathbb{R}\}_{a\in A,h\in H},\{D_a^i\in\mathbb{R}\}_{a\in A,i\in I},\{P_{j,s}\}_{s\in S,j\in J_s},\{T^i\}_{i\in I},\{T^h\}_{h\in H}} \sum_{h\in h}\lambda^h\sum_{s\in S}V^h(I_s^h,\{P_{j,s}\}_{j\in J_s};s)+ \sum_{i\in I}\lambda^i\sum_{s\in S}V^i(I_s^i,\{P_{j,s}\}_{j\in J_s};s),$$

subject to the household's limited participation constraints,

$$\Phi^h(\{D_a^h\}_{a\in A})\leq\vec{0}$$

for all $h\in H$, the definition of incomes I_s^h and I_s^i, market clearing in assets, the government's budget constraint, and goods market clearing for each state $s\in S$ and good $j\in J_s$,

$$\sum_{h\in H}(X_{j,s}^h(I_s^h,\{P_{j',s}\}_{j'\in J_s})-Y_{j,s}^h)=\sum_{i\in I}(X_{j,s}^i(I_s^i,\{P_{j',s}\}_{j'\in J_s})-Y_{j,s}^i).$$

Here, $X_{j,s}^h(I_s^h,\{P_{j',s}\}_{j'\in J_s})$ denotes the demand function for good j by agent h in state s. Note that the definition of the social planner's problem does not include constraints on the intermediaries trades, which, as discussed above, are instruments that can be used to implement the solution to the planning problem.

I begin by describing the equilibrium absent regulation. I will say (again following Geanakoplos and Polemarchakis [1986]) that the equilibrium is constrained inefficient if there is no set of Pareto weights (λ^h,λ^i) in the planner problem such that an equilibrium allocation in the economy coincides with a solution to the planning problem.

Proposition 1. *Absent regulation, there is no arbitrage. If there are at least two* *households who are limited participants, the allocation is generically constrained inefficient in the sense of Geanakoplos and Polemarchakis [1986].*

Proof. See the appendix, ... □

Absent regulation on intermediaries, there can be no arbitrage. If there are assets $a \in A^h$ and $a' \in A^I$ with identical payoffs ($Z_{a,s} = Z_{a',s}$ for all $s \in S$), they are perfect substitutes from the perspective of intermediaries, and therefore must have the same prices ($Q_a = Q_{a'}$). Constrained inefficiency is not surprising, either; although limited participation is not identical to incomplete markets, the same pecuniary externalities that generate constrained inefficiency in incomplete markets (see Geanakoplos and Polemarchakis [1986]) apply in the context of limited participation.

Proposition 1 establishes "generic" inefficiency. In this context, "generic" means that if there is an economy, with no regulation and at least two limited participants, that is constrained efficient (meaning the allocation coincides with a solution to a planning planning problem), there is a slightly perturbed version of the economy that will be constrained inefficient. The specific perturbation I use in the proof, following Geanakoplos and Polemarchakis [1986] and Farhi and Werning [2016], is a perturbation to the utility functions of the agents. In contrast, if there is a constrained inefficient economy, it will generically not be possible to perturb it to reach constrained efficiency. Speaking loosely, the set of economies that achieve constrained efficiency absent regulation form a "measure zero" subset of the set of all economies.

I next turn to implementations of solutions to the social planning problem. I will consider implementations of solutions to the planning problem that generate constrained efficiency through regulation on the trades of intermediaries. The next proposition shows that, generically, there exist regulations that simultaneously generate constrained efficiency and bring about arbitrage.

Proposition 2. *For a given set of strictly positive Pareto weights, there exist regulations $\{\Phi^i\}_{i \in I}$ and an equilibrium given those regulations such that the equilibrium allocation coincides with a solution to the planning problem with those Pareto weights. Generically, in the set of strictly positive Pareto weights, there is arbitrage in this equilibrium.*

Proof. See the appendix, ... □

Proposition 2 establishes that, for a given set of Pareto weights, it is possible for a planner to implement the constrained efficient allocation via regulation. This is not particularly surprising– all trade goes through intermediaries, and by regulating those intermediaries, the planner can achieve any allocation, and in particular any constrained efficient allocation. More surprising is the second part of the statement– that, generically, the equilibrium features arbitrage. This arbitrage reflects the desire of the planner to control the asset allocations of both households and intermediaries, using regulations on the trades of intermediaries.

To summarize, this section shows that an absence of arbitrage does not imply efficiency, and that efficiency does not imply an absence of arbitrage. Taken together, these results suggest that arbitrage and efficiency are simply unrelated. This is not the case; in the next section, I will show that, under the optimal regulatory regime, the "pattern of arbitrage" across various states of the nature is closely related to the notion of "wedges" that are typically used to analyze externalities. I will argue that, but looking at the patterns of arbitrage across financial assets, regulators can asses whether the regulations they implement are having their desired effects.

3 Arbitrage and Wedges

I begin this section by defining the "wedges," $\tau_{j,s}$, which are defined for each state $s \in S$ and good $j \in J_s$. These wedges represent the difference between the first-order conditions of the planner and of the agents– the latter do not take into account the effects that their asset allocation decisions have on goods prices, and these pecuniary externalities, due to limited participation, have welfare consequences. Let $\bar{\pi}_s$ denote an arbitrary full-support probability distribution over the states $s \in S_1$, let $\bar{\pi}_s \mu_{j,s}$ denote the multiplier, in the planner's problem, on the market clearing constraint of good $j \in J_s$ in state $s \in S_1$, and let κ denote the planner's multiplier on the government budget constraint. Define the wedges $\tau_{j,s}$ using an orthogonal projection of the multipliers on to prices:

$$\mu_{j,s} = \bar{\mu}_s P_{j,s} - \kappa \tau_{j,s} P_{j,s},$$

with $\sum_{j \in J_s} \tau_{j,s} P_{j,s} = 0$ for all $s \in S$.[5]

The multipliers $\mu_{j,s}$ represent the social marginal cost of demand for good j in state s, and the multiplier κ is the social marginal value of resources at time zero. To the extent that the multipliers $\mu_{j,s}$ are not proportional to prices (i.e. that the $\tau_{j,s}$ are not zero), pecuniary externalities exist in the equilibrium. A high value of the wedge $\tau_{j,s}$ indicates that the multiplier $\mu_{j,s}$ is low relative to the price $P_{j,s}$, which is to say that the social marginal cost of demand for the good is less than the price. Scaling the wedges $\tau_{j,s}$ by the multiplier κ ensures that the units of these wedges are in "dollars," rather than units of social utility.

These externalities can be compensated for by transferring income in state s to household h from an intermediary i, if household h has a different marginal propensity to demand good j in state s than the intermediary does. Transfers between households with differing marginal propensities to demand could also accomplish the same goal (transfers between intermediaries, who have the same homothetic utility function, would have no effect). Let $X^h_{I,j,s}$ denote the marginal effect that income has on the demand of household h for good j in state s, holding prices constant. Let $\bar{\pi}_s$ denote an arbitrary probability distribution with full support over the state $s \in S$. If the wedge-weighted difference of the income effects for the household and the intermediary,

$$\Delta^{h,i}_s = \sum_{j \in J_s} \tau_{j,s} P_{j,s} (X^h_{I,j,s} - X^i_{I,j,s}),$$

is positive, transferring income from intermediary i to household h in state s has a benefit, from the planner's perspective, because it alleviates pecuniary externalities. However, it might also have a cost, if the Pareto-weighted marginal utilities of income between intermediary i and household h in state s are not equalized. For transfers that are feasible (in the span of the assets tradeable by household h), under an optimal policy, these costs and benefits will exactly offset. That is, in a constrained efficient allocation, for all households $h \in H$, intermediaries $i \in I$, and assets $a \in A^h$,

$$\sum_{s \in S} \bar{\pi}_s \Delta^{h,i}_s Z_{a,s} = \sum_{s \in S} \left(\frac{V^i_{I,s}}{V^i_{I,s_0}} - \frac{V^h_{I,s}}{V^h_{I,s_0}}\right) Z_{a,s}, \tag{1}$$

where $V^h_{I,s}$ denotes the marginal value of income for household h in state s, and $V^i_{I,s}$

[5] This definition of wedges is the same as the one employed by Farhi and Werning [2016], adjusted for the differences between production and endowment economies.

is that object for intermediary i.

Put another way, we can view $\Delta_s^{h,i}$ as the difference of two stochastic discount factors– one for the household h and one for the intermediary. Each of these stochastic discount factors reflects the ratio of the agent's marginal utilities in state s and state s_0, adjusted for any differences between that agent's beliefs and the measure $\bar{\pi}$. Neither of these SDFs can be used to price all assets, because of the limited participation constraints (for households) and regulations (for intermediaries). However, for particular assets (the assets in A^h for household h, the Arrow securities for an intermediary) that are not affected by these constraints, the relevant SDF does in fact price the asset. Consequently, the amount of arbitrage between an asset $a \in A^h$ and its replicating portfolio of Arrow securities in the intermediary market will also be determined by the wedges. This result is stated below in 3.

I next elaborate on the importance of the assumption that the intermediaries have identical, homothetic preferences. A direct consequence of this assumption is that $X^i_{I,j,s} = X^{i'}_{I,j,s}$ for all $i, i' \in I$, $j \in J_s$, and $s \in S$. As a result, there is never any particular reason to transfer wealth across intermediaries. It follows that it is not necessary to regulate intermediaries' trades in the Arrow securities market, A^I. I will focus on implementations of the constrained efficient allocations without regulation of trade in the Arrow securities, primarily because these implementations feature meaningful prices for the Arrow securities. An alternative implementation that dictated all trades for intermediaries would not need to have asset prices at all, and hence the question of the existence of arbitrage would be ill-defined.

However, this assumption has economic content. It implies that the social planner is indifferent to the distribution of wealth across intermediaries. Put another way, the regulator has "macro-prudential" motives for regulation, but not "micro-prudential" motives. In section §6, I discuss how micro-prudential motives for regulation might change the interpretation of my results.

I now present the main result of this section: the relationship between the wedges (as summarized by the differences Δ) and arbitrage, in a constrained efficient allocation implemented without regulation of the Arrow securities market.

Proposition 3. *Consider a constrained efficient allocation implemented by regulations $\{\Phi^i\}_{i \in I}$ that do not regulate trade in the Arrow securities A^I. For any security* *tradeable by household h ($a \in A^h$), define the amount of arbitrage as*

$$\chi_a = -Q_a + \sum_{s \in S} Z_{a,s} Q_s,$$

where Q_a is the price of the asset and Q_s is the price of an Arrow security paying off in state s.

Then, for any intermediary $i \in I$,

$$\chi_a = \sum_{s \in S} \bar{\pi}_s \Delta_s^{h,i} Z_{a,s}.$$

Proof. See the appendix, ... □

In words, an asset tradeable by households will be cheap, relative to its Arrow-market replicating portfolio, if its payoffs occur mainly in states for which the planner would like to transfer wealth from intermediaries to households.

As a special case, consider a risk-free asset ($Z_{a,s} = 1$ for all $s \neq s_0$, recalling that $Z_{a,s_0} = 0$ by convention), and suppose that there are two goods, "houses" and "yachts," bought exclusively by households and intermediaries, respectively. If there are positive externalities in demand for houses in the future, and/or negative externalities in demand for yachts, ($\tau_{houses,s} < 0$ and $\tau_{yachts,s} > 0$), then it will alleviate externalities to transfer wealth to households in the future ($\Delta_s^{h,i} > 0$). As a result, in equilibrium, the risk-free asset available to households will have a lower price (higher interest rate) than the risk-free asset available to intermediaries.

In the next section, I invert this exercise: given an observed pattern of arbitrage, what can we say about $\Delta_s^{h,i}$? That is, presuming regulation is optimal, in which states must the regulator believe there are externalities that justify transferring wealth from intermediaries to households or vice versa?

4 The Externality-Mimicking Portfolio

Suppose a financial economist observes, in financial markets, a set of securities A^{arb} that she believes are tradable by households and for which the financial economist also observes a replicating portfolio of securities traded only by intermediaries (e.g. derivatives). I will call these assets "arbitrageable," meaning that, if the prices of these securities were not consistent with the prices of their replicating portfolio, there would be arbitrage. Many securities will not be arbitrageable, from the perspective of the financial economist, because she does not observe the price of the replicating portfolio of derivatives.

For each arbitrageable security $a \in A^{arb}$, by observing the price of the security and of the replicating portfolio of derivatives, the financial economist can compute the amount of arbitrage, χ_a. Examples of this sort of exercise include the covered interest parity violations documented by Du et al. [2017], the arbitrage between asset-swapped TIPS and treasury bonds documented by Fleckenstein et al. [2014], and the basis between corporate bonds and credit default swaps discussed in Garleanu and Pedersen [2011].

To varying degrees, each of these "arbitrages" is not exactly a textbook arbitrage. One can always imagine stories (e.g. jumps to default by derivatives counterparties correlated with the value of the derivative contract) that could justify small pricing deviations. I view these issues as quantitatively insignificant (in most of the cases cited above), and for the purposes of discussion will assume that each of these authors has in fact documented an arbitrage. The "cleanest" arbitrage of all is perhaps the ability of banks in the US to borrow overnight from the GSEs in the fed funds market and earn interest on excess reserves (Bech and Klee [2011]). Because this arbitrage has only an overnight maturity, and there is no counterparty risk, it is very difficult to come up with a story, aside from regulation, to explain why banks would not be willing to engage in the arbitrage.

Suppose we observed these arbitrages, and we assume that the equation of 3 holds: for each $a \in A^{arb}$,

$$\chi_a = \sum_{s \in S} \bar{\pi}_s \Delta_s^{h,i} Z_{a,s},$$

for any household $h \in H$ that can trade the security $a \in A^{arb}$. If the set A^{arb} contained a full set of Arrow securities (or was a complete asset market, more generally), there would be a one-to-one mapping between the arbitrage on these securities and the differences $\Delta_s^{h,i}$ (under a fixed probability distribution $\bar{\pi}_s$). Empirically, although there are multiple examples of assets with arbitrage, it would be a stretch to say that the financial economist observes a complete market in arbitrageable securities.[6] As a

[6] The existence of a full set of Arrow securities in A^{arb} is (at least in theory) consistent with the assumption of limited participation. To be in A^{arb}, some household must be able to trade the security, but these does not imply that every household can trade the security.

result, it will not be possible, empirically, to recover $\Delta_s^{h,i}$ for all states $s \in S$.

We can, however, attempt to project the differences $\Delta_s^{h,i}$ on to a lower-dimensional space, and require that this projection be consistent with the amount of arbitrage we observe in data. This procedure, which I will describe next, builds on ideas introduced by Hansen and Richard [1987]. There are essentially two ways to look at the procedure I employ. The first is that I am simply projecting the differences $\Delta_s^{h,i}$ on to the space of asset returns, perhaps after employing a non-linear transformation. The second is that I am attempting to make the risk-neutral measures for the intermediaries and for households as similar as possible. I will begin by describing this second perspective, and then demonstrate in lemma 1 that they are equivalent.

I will assume that the set A^{arb} includes a risk-free security, and I will define the risk-free interest rate for intermediaries, R^i, as

$$R^i = (\sum_{s \in S_1} \frac{V_{I,s}^i}{V_{I,s_0}^i})^{-1},$$

and define another risk-free interest rate R^h in similar fashion. If there is arbitrage on risk-free securities, these two interest rates will not be equal, and are both observable in data. Using these interest rates, I define a risk-neutral measure for intermediaries,

$$\pi_s^i = R^i \frac{V_{I,s}^i}{V_{I,s_0}^i},$$

and define π_s^h along similar lines.

Hansen and Richard [1987] study the problem of minimizing the variance of a stochastic discount factor, subject to the constraint that the SDF price a set of assets. Sandulescu et al. [2017] study a generalized version of this problem, minimizing some moment (not necessarily the variance) of the SDF. An SDF is, of course, also the Radon-Nikodym derivative between two measures– in these applications, a risk-neutral measure and the physical measure. The problem of Hansen and Richard [1987] (and of Sandulescu et al. [2017] more generally) can be recast as minimizing the value of a divergence between the risk-neutral measure and the physical measure, subject to the constraint that the risk-neutral measure correctly price assets. The particular divergence that corresponds to the variance minimization procedure of Hansen and

Richard [1987] is the "chi-squared" divergence,[7]

$$D_{\chi^2}(\pi||\pi') = \frac{1}{2} \sum_{s \in S_1} \pi_s^{'} (\frac{\pi_s}{\pi_s^{'}} - 1)^2.$$

These procedures have the virtue that they make the estimated risk-neutral measure as "close" as possible to the physical measure, under a definition of "close" defined by the divergence in question. In this context of this paper, I could apply this procedure to both the measure associated with intermediaries (π_s^i) and the measure associated with households (π_s^h), making each of them as close as possible to the physical measure. However, this requires an econometric exercise (estimating the moments of returns under the physical measure) and will not necessarily minimize the estimated amount of arbitrage. That is, by making the households' and intermediaries' measures as close as possible to the physical measure, we do not necessarily make them as close as possible to each other.

Instead, I will minimize the value of the chi-squared divergence between the households' measure, π^h, and the intermediaries' measure, π^i, subject to the constraint that the observed price of the asset for households is correct. That is, letting $\mathcal{P}(S)$ denote the simplex over time-one states, I solve

$$\hat{\pi}^h = \min_{\pi \in \mathcal{P}(S)} D_{\chi^2}(\pi||\pi^i)$$

subject to, for each $a \in A^{arb}$,

$$(R^h)^{-1} \sum_{s \in S_1} \pi_s Z_{a,s} + \chi_a = (R^i)^{-1} \sum_{s \in S_1} \pi_s^i Z_{a,s}.$$

This procedure has two virtues. First, it minimizes arbitrage, as measured by the divergence between the two risk-neutral probability measures. Second, we will see that the estimated measure $\hat{\pi}^h$ is entirely a function of the moments of asset payoffs under the risk-neutral measure of intermediaries.[8] Consequently, if we observe a

[7]The general definition of a divergence $D(\pi'||\pi)$ is a weakly positive scalar function of two probability measures that is zero if and only if $\pi' = \pi$. The more general procedure of Sandulescu et al. [2017] minimizes an α-divergence, which is a particular family of divergences that includes the χ^2 divergence.

[8]The procedure also has a disadvantage, which is that $\hat{\pi}^h$ might not have full support. For discussions of this issue, see Hansen and Richard [1987] and Hansen and Jagannathan [1991]. Using an alternative divergence, such as the Kullback-Leibler divergence, as in Sandulescu et al. [2017],

sufficiently rich set of asset prices traded only by intermediaries, we can construct $\hat{\pi}^h$ from those asset prices, without any econometrics. In the next section, I will use options prices on currency pairs to put this idea into practice.

One additional complication is that (at least in theory) there might not be a single household h^* that can trade all the assets in A^{arb}. I will ignore this issue, and assume that such a household does exist. Additionally, it is convenient to work in the space of returns, rather than payoffs. For this reason, I will assume that the payoff of each asset in A^{arb} is weakly positive in all states of the world.

As noted above, the differences $\Delta_s^{h^*,i}$ are defined with respect to a probability measure. I will work with the particular differences defined with respect to the intermediaries' risk-neutral probability measure, π^i, and call these differences the "risk-neutral externalities." From an estimated risk-neutral measure $\hat{\pi}^h$ and the risk-neutral measure π^i, we can construct an estimated set of risk-neutral externalities, $\hat{\Delta}_s^{h^*,i}$. Defining $\chi_R = \frac{1}{R^i} - \frac{1}{R^h}$ as the arbitrage on the risk-free asset, we can define the estimated risk-neutral externalities as (consistent with equation (1))

$$\hat{\Delta}_s^{h^*,i} = \chi_R + \frac{1}{R^h}(1 - \frac{\hat{\pi}_s^h}{\pi_s^i}). \tag{2}$$

An alternative approach would simply be to project (in the least squares sense) the risk-neutral externalities on to the space of returns. Define the excess returns for intermediaries as

$$R_{a,s}^i = R^i \frac{Z_{a,s}}{\sum_{s' \in S_1} \pi_{s'}^i Z_{a,s'}} - R^i,$$

and let $A^{arb} \setminus \{a_{RF}\}$ denote the set of arbitrageable assets excluding the risk-free assets. Let Σ^i be the covariance matrix of returns for risky assets in A^{arb} under the risk-neutral measure π^i. Given these definitions, the least squares projection is

$$\bar{\Delta}_s^{h^*,i} = \chi_R + R^i \sum_{a,a' \in A^{arb} \setminus \{a_{RF}\}} (\frac{\chi_a}{\sum_{s' \in S_1} \pi_{s'}^i Z_{a,s'}} - \chi_R)(\Sigma^i)_{a,a'}^{-1} R_{a',s}^i.$$

The following lemma shows that the least squares projection ($\bar{\Delta}_s^{h^*,i}$) is identical to the estimate of the risk-neutral externalities that minimizes the chi-squared divergence between the household and intermediaries' risk-neutral measures ($\hat{\Delta}_s^{h^*,i}$).

would remove this problem, but would break the equivalence between the divergence minimization and least squares projection.

Lemma 1. *Let $\hat{\pi}^h$ be the measure for that minimizes the chi-squared divergence between itself and the intermediaries' risk-neutral measure, subject to the constraint that $\hat{\pi}^h$ correctly price each asset in A^{arb} from the households' perspective:*

$$\hat{\pi}^h = \min_{\pi \in \mathcal{P}(S)} D_{\chi^2}(\pi || \pi^i)$$

subject to, for each $a \in A^{arb}$,

$$(R^h)^{-1} \sum_{s \in S_1} \pi_s Z_{a,s} + \chi_a = (R^i)^{-1} \sum_{s \in S_1} \pi^i_s Z_{a,s}.$$

If the measure $\hat{\pi}^h$ has full support, then the associated estimate of the risk-neutral externalities, $\hat{\Delta}^{h^*,i}_s = \chi_R + \frac{1}{R^h}(1 - \frac{\hat{\pi}^h}{\pi^i})$, is equal to the least-squares projection of the externalities on to the space of returns, $\bar{\Delta}^{h^*,i}_s$.

Proof. See the appendix, section C.2. □

This lemma can also be understood in terms of a portfolio problem. The portfolio weights

$$\theta^*_{a'} = \sum_{a \in A^{arb} \setminus \{a_{RF}\}} R^i \left(\frac{\chi_a}{\sum_{s' \in S_1} \pi^i_{s'} Z_{a,s'}} - \chi_R \right) (\Sigma^i)^{-1}_{a,a'}, \tag{3}$$

plus some amount of the risk-free asset, form a zero-cost portfolio whose returns are the unexpected component of $\bar{\Delta}^{h^*,i}_s$. I will call this portfolio the "risk-neutral externality-mimicking portfolio." The caveat in lemma 1, which I will ignore in what follows, applies to situations when the risk-neutral externality-mimicking portfolio has a gross return of less than negative one, breaking limited liability. This could occur if the portfolio is "short" one of the assets, but is very unlikely for the types of assets and time horizons that I study.

The risk-neutral externality-mimicking portfolio is the portfolio of assets that is the unique projection of the planner's perceived externalities on to the space of available payoffs, given the intermediaries' risk-neutral measure π^i. In other words, its return is the best linear approximation of the conditional expectation of $\Delta^{h^*,i}_s$ given the returns of the arbitrageable assets. In this sense, when its return is high, the planner must generally think that it would alleviate externalities to transfer wealth from intermediaries to households, and when its payoff is low, the planner must generally think it would alleviate externalities to transfer wealth to intermediaries from households.

Continuing the analogy with the procedure of Hansen and Richard [1987], I discuss a bound along the lines of Hansen and Jagannathan [1991]. In this analogy, as above, the intermediaries' measure π^i_s will play the role of the physical measure, and the households measure π^h_s will play the role of the risk-neutral measure. As in the latter paper, the standard deviation (under the measure π^i_s) of any "stochastic discount factor" $\frac{\pi^h_s}{R^h \pi^i_s}$, divided by its mean (also under π^i_s) must be greater than the corresponding statistic for the least-squares projection, $\frac{\hat{\pi}^h_s}{R^h \pi^i_s}$. Moreover, for the least squares projection, this statistic is also equal to the "arbitrage Sharpe ratio" of the portfolio θ^*_a. I define the "arbitrage Sharpe ratio" for any portfolio θ as

$$S^i(\theta) = \frac{|\sum_{a \in A^{arb}} \theta_a R^i (\frac{\chi_a}{\sum_{s' \in S_1} \pi^i_{s'} Z_{a,s'}} - \chi_R)|}{\sigma^i(\sum_{a \in A^{arb}} \theta_a R^i_{a,s})},$$

where $\sigma^i(\cdot)$ denotes the standard deviation under the measure π^i. The numerator is the expected excess return (under the intermediaries' risk-neutral measure) due to arbitrage available to intermediaries that purchase assets available to households. That is, it is not the entire expected excess return, as in the usual definition of the Sharpe ratio, but only the portion due to "excess arbitrage," meaning arbitrage in excess of the arbitrage available on the risk-free asset.

The analogy to the bound of Hansen and Jagannathan [1991] is described in the following lemma.

Lemma 2. *For any risk-neutral externalities $\Delta^{h,i}_s$ that satisfying the arbitrage pricing equation of 3, and any portfolio θ,*

$$R^h \sigma^i(\Delta^{h,i}_s) \geq R^h \sigma^i(\hat{\Delta}^{h,i}_s) = S^i(\theta^*) \geq S^i(\theta).$$

Proof. See the appendix, section C.3. □

This lemma provides an alternative perspective on the least squares projection procedure. One can view the projection procedure (or, equivalently, the divergence minimization procedure) as searching for a portfolio with the highest arbitrage Sharpe ratio. It also shows that the arbitrage Sharpe ratio of any portfolio places a lower bound on the volatility of the risk-neutral externalities, and the tightest possible bound is formed using the externality-mimicking portfolio.

The externality-mimicking portfolio is a reflection of what regulation is actually accomplishing. We can compare it to the stated objectives of regulators, and in particular to the "stress tests" conducted by the Federal Reserve and other central banks. These "stress test" scenarios presumably reflect situations in which the Federal Reserve would like intermediaries to have more wealth than the intermediaries would choose to have on their own, indicating some perception of externalities.

For example, the 2017 Stress Test Scenarios[9] involve, in the "severely adverse" scenario, a depreciation of the euro and pound, and an appreciation of the yen, relative to the US dollar. Du et al. [2017] document that CIP violations with a common pattern across these currencies: investing in euros, pounds, or yen bonds, and using a forward contract to swap back to dollars, offers a higher return than investing in dollars, at the beginning of 2017. This suggests something inconsistent: the externality-mimicking portfolio treats euro, pounds, and yen roughly equally, but the stress test scenario treats them differently. In this next section, I explore this idea in more detail.

5 Empirical Estimates

In this section, I estimate externality-mimicking portfolios and study the returns that the portfolios would experience in the "stress tests" conducted by the federal reserve. Constructing these portfolios requires, for each asset price that we as financial economists observe, taking a stand on whether that asset price comes from the intermediaries-only market or from a market in which both intermediaries and households can participate. Reality, of course, is more complicated; for example, some assets are rarely owned by individuals, but held by many money-market funds or mutual funds, and for our purposes might be thought of as tradable by households.

These abstract difficulties are often easier to resolve in the context of specific arbitrages. The empirical application I consider features two arbitrages: an arbitrage between household interest rates and the interest on excess reserves (IOER), and covered interest parity violations. The horizon of the arbitrage I study is one month.

The first arbitrage involves the yield difference of two (essentially) risk-free assets. A bank can earn interest on excess reserves held at the Federal Reserve. If there is no meeting of the FOMC within the next month, the bank is essentially guaranteed to

[9] https://www.federalreserve.gov/newsevents/pressreleases/files/bcreg20170203a5.pdf

earn one month's worth of interest at the current overnight rate. Of course, in rare circumstances, the Federal reserve might choose to change the interest on reserves between meetings, but such changes have low ex-ante likelihood and are unlikely to materially alter the expected interest rate.

A household, in contrast, cannot earn the rate of interest on excess reserves. As an alternative, it could invest in treasury bills, highly-rated commercial paper, repo agreements (via money market funds), bank deposits, or the like. There is a significant literature arguing that treasury bonds are special, relative to other bonds the would appear to be close substitutes (e.g. Fleckenstein et al. [2014]), perhaps due to their additional liquidity or some other kind of benefits. I will use 1-month OIS swap rates, which closely track the yields of one-month maturity highly-rated commercial paper in the US, as a proxy for a risk-free rate available to households that provides no liquidity benefits. These rates tend to be higher than the rates on one-month constant maturity treasuries, but lower than LIBOR rates (which may include credit risk). For example, on August 19th, 2016 (a little over one month before the next FOMC meeting), 1 month constant-maturity treasury rates were 27bps, AA non-financial one-month commercial paper rates were 37bp, one-month OIS swaps were 40bps, the interest on excess reserve rates were 50bps, and one month LIBOR was 52bps. In the notation of the previous section, the annualized household risk free return $(R^h)^{12}$ is 1.004 and the annualized intermediary risk-free return $(R^i)^{12}$ is 1.005.

The next arbitrage is a covered interest parity violation. I will assume that households can purchase euros with dollars, and invest their euros at the 1-month OIS swap rate in euros (the EONIA rate). What households cannot do (but intermediaries can) is trade currency forwards. The intermediary can replicate the household's euro/EONIA trade by investing dollars at the IOER rate, and selling dollars for euros using a 1-month currency forward. The households can invest in euros/EONIA instead of investing their dollars at the 1-month OIS swap rate in dollars, and hence the difference between these two strategies is also an asset in their payoff space.

That is, one could either define the arbitrage as being about zero-cost portfolios (the difference in these two strategies vs. the forward), or using positive-cost strategies (investing in euros/EONIA vs. the bank investing at IOER and using the forward). However, these two alternatives are simply linear rotations of the payoff Z_s and the arbitrages χ, and hence will yield identical estimates of $\bar{\Delta}_s^{h^*,i}$. The first strategy is analogous to the approach of Du et al. [2017], but the latter has the benefit that the portfolio problem can be implemented in the space of returns rather than payoffs. For this reason, in what follows, I will pursue the latter approach and consider arbitrages between positive-cost strategies.

In my empirical exercise, I will consider portfolios of the risk-free arbitrage and investments in euros and yen, and use an investments in pounds as a sort of "out-of-sample" test. To estimate the risk-neutral externality mimicking portfolio, it is sufficient to estimate the arbitrage on the risk-free assets and each of the risky assets, along with the covariance matrix of the risky assets under the intermediaries' risk-neutral measure. Estimates of the arbitrage can be constructed from spot and forward currency rates, the IOER rate, and OIS swap rates, as in Du et al. [2017]. Estimates of the risk-neutral covariance matrix between various currency pairs can be constructed from options on exchange rates. In particular, the availability of dollar-euro, dollar-pound, dollar-yen, euro-yen, euro-pound, and pound-yen options allows me to extract an entire risk-neutral variance-covariance matrix from options prices. I assume that the at-the-money FX options prices I observe come from the intermediary-only market (and hence the IOER rate is the appropriate discount rate), and that the distribution of future exchange rates is log-normal. The formula of Black [1976] applies, and using this formula I extract risk-neutral variances and covariances. More elaborate constructions, using out-of-the-money options as in the construction of the VIX index, are possible and would dispense with the need to assume log-normality, at the expense of additional complexity.

This procedure can be run at daily frequency, for all dates at least one month before a scheduled FOMC meeting. I will begin by presenting summary statistics for the magnitude of the arbitrage in question. The dataset consists of daily data from financial markets between January 2011 and September 2016, including only those dates that are at least one month (more than 23 non-weekend days) before the start of the next scheduled FOMC meeting and not meeting dates themselves. Because the FOMC holds eight scheduled meetings each year, roughly one quarter of all non-weekend days are included in the dataset.

The first summary statistics I will present contain the sample means and standard deviations of the arbitrage associated with each currency, as well as the risk-free arbitrage. Conceptually, these statistics correspond to the term

$$\tilde{\chi}_a = \frac{-Q_a + \sum_{s \in S} Z_{a,s} Q_s}{\sum_{s \in S} Z_{a,s} Q_s}.$$

For example, for euros, it represents the percentage difference in price, in dollars today, of purchasing a single euro one month in the future by buying the euro at spot today and saving at OIS (households, Q_a), and obtaining the same euro one month in the future by savings at the IOER rate and using a currency forward (intermediaries, $\sum_{s \in S} Z_{a,s} Q_s$). I also present the difference in dollar zero-coupon bonds prices, scaled by the interest rate for intermediaries ($R^i \chi_R$), which is a function of the difference between the dollar OIS rate and the IOER rate.

Table 1 below shows the mean and standard deviation of the magnitude of this arbitrage in my sample. The arbitrage has a horizon of one month, but is scaled to annualized values. The table also shows the option-implied volatility and correlations of each currency (with respect to the US dollar). These volatilities and correlation are extracted from the sample mean of the daily variance-covariance matrix constructed from options prices.

Table 1: Summary Statistics for Arbitrage

	Pounds	Euros	Yen	Risk-Free
Arbitrage Mean (bps/year)	-2.4	13.0	20.8	-13.2
Arbitrage SD (bps/year)	12.0	22.1	27.2	2.6
Option-Implied Currency/$ Vol. (bps/year)	842	969	967	-
Currency/$ Correlation with Pound/$	1.00	0.58	0.20	-
Currency/$ Correlation with Euro/$	0.58	1.00	0.28	-
N	395	395	395	395

Notes: Describe Table here

From table 1, we can observe several notable features of the data. First, to clarify the meaning of the signs, the negative sign of the "risk-free" arbitrage means that a "zero-coupon bond" for intermediaries costs less than the equivalent bond for households. In other words, intermediaries are able to earn a higher rate of interest than households (IOER vs. OIS). However, the positive sign on the euro and yen arbitrages implies that it is more expensive for intermediaries to use derivatives to purchase a euro (or yen) one month in the future in exchange for a dollar today than it is for intermediaries to use products also available to households. Despite being able to save at the IOER rate instead of the USD OIS rate, the magnitude of the covered interest parity (CIP) violation more than offsets this effect. Of course, this is not necessarily "bad news" for intermediaries– it simply means that inter. However, for pounds, these two effects roughly offset (on average). That is, the extra interest an intermediary can earn by saving at the IOER rate is roughly offset by the magnitude of the pound-dollar CIP violation. Finally, note that, at least for euros and yen, the covariance matrix of currency returns is reasonably close to a scaled version of the identity matrix. As a result, the portfolio weights constructed as in equation (3) above will be roughly proportional to the amount of "excess arbitrage," $\tilde{\chi}_a - R^i \chi_R$, observed for each asset.

Figure 1 below shows the time series of the excess arbitrages $\tilde{\chi}_a - R^i \chi_R$ for my sample (which, as mentioned above, is restricted to dates at least 23 non-weekend days before the next FOMC meeting). We can observe that the magnitude of arbitrage is quite volatile, and there is significant positive comovement between the Euro and Yen arbitrage (these facts are essentially the same as those documented in Du et al. [2017]). As noted above, the covariance matrix that is used to convert these excess arbitrages into portfolio weights is (at least on average) reasonably close to a scaled version of the identity matrix. This is apparent in figure 2 below, which shows that the portfolio weights closely track the excess arbitrages.

Figure 1: Time Series of Excess Arbitrage

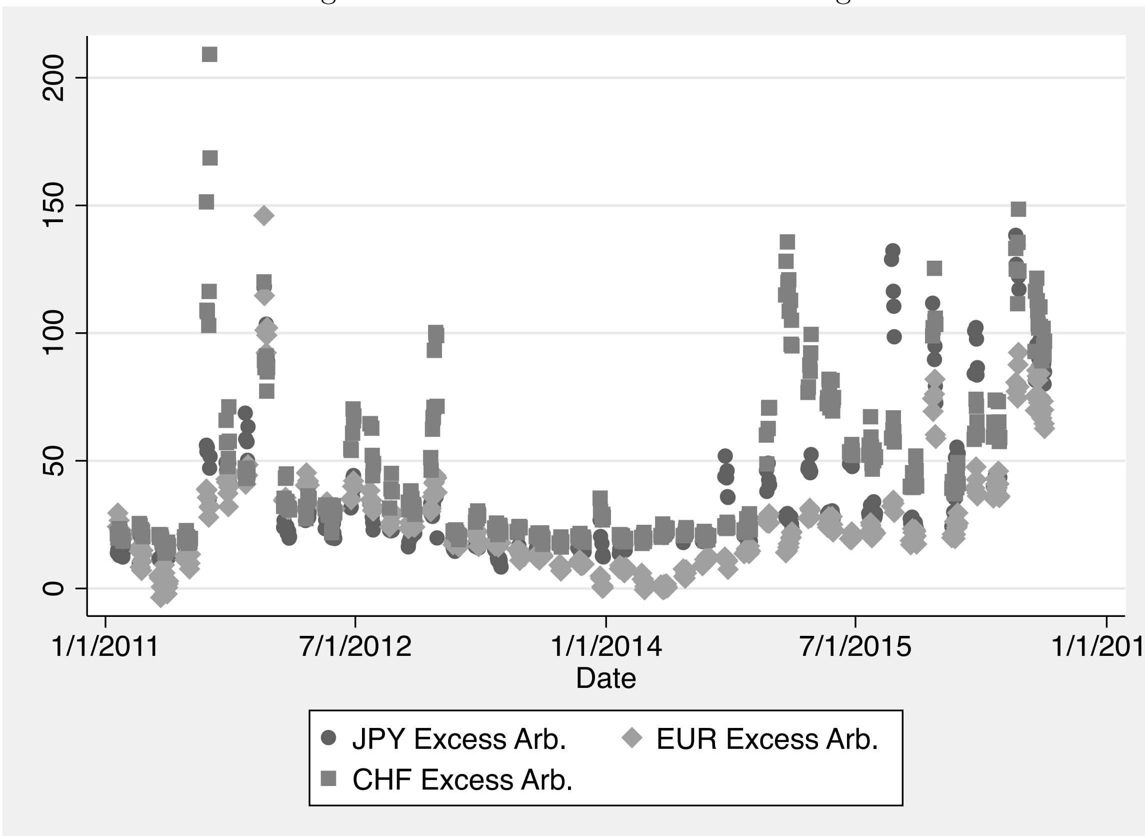
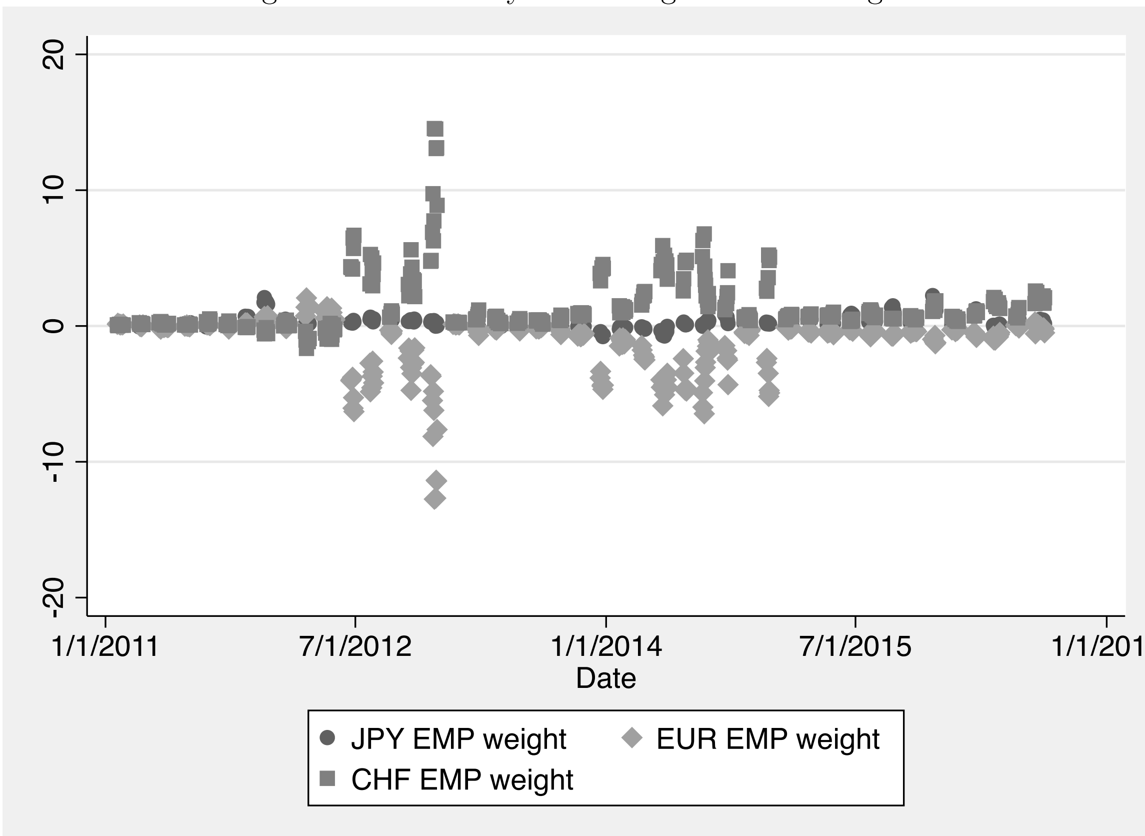


Figure 2: Externality-Mimicking Portfolio Weights



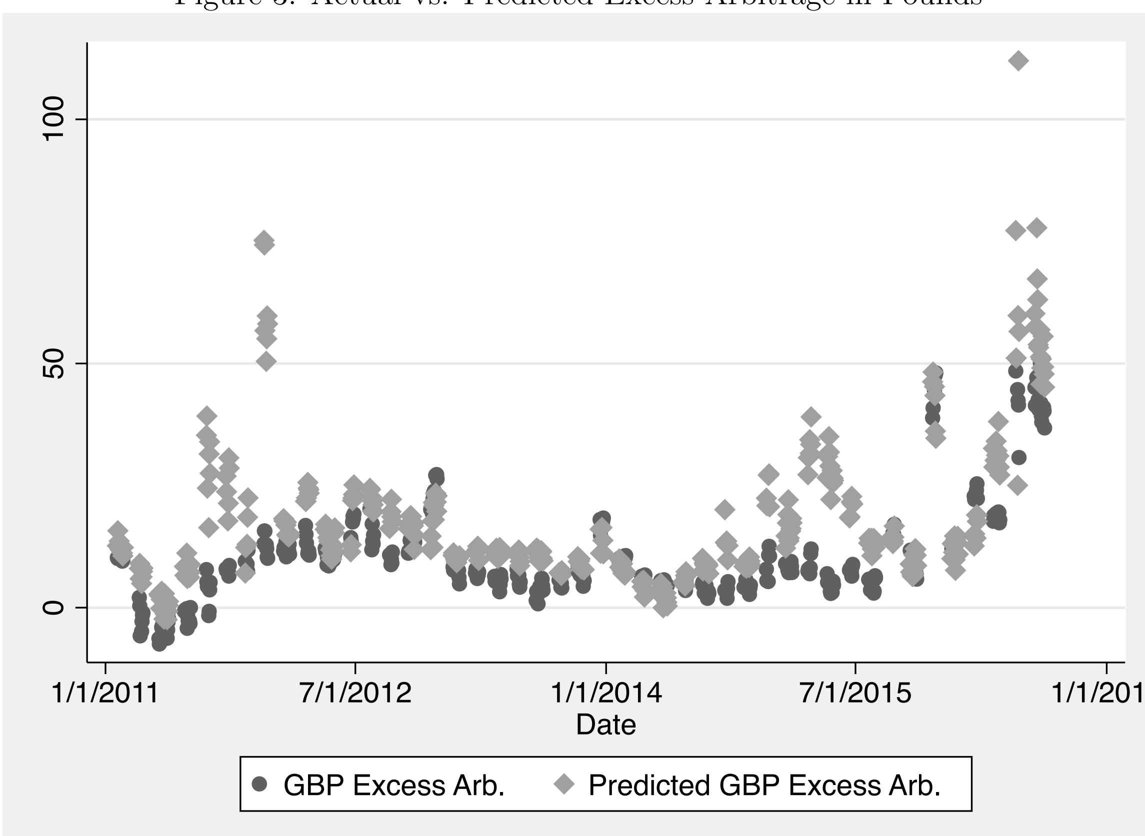
Having computed the portfolio weights of the externality-mimicking portfolio, I next consider the predictions that this portfolio has about other arbitrages. As mentioned above, I deliberately excluded pounds from the set of currencies used to form the externality-mimicking portfolio. This allows me to test whether the arbitrage predicted using the externality-mimicking portfolio is consistent with the arbitrage actually observed for the dollar-pound currency pair. Formally, I estimate

$$\tilde{\chi}_{GBP} - R^i \chi_R = \Sigma^i_{GBP} \theta^*,$$

where θ^* is the externality-mimicking portfolio (equation 3) and Σ^i_{GBP} is the covariance, under the intermediaries' risk-neutral measure, between the dollar-pound exchange rate and the assets used to form the externality-mimicking portfolio (the dollar-euro and dollar-yen exchange rates). One can observe, using the definition of the estimated risk-neutral externalities $\hat{\Delta}^{h^*,i}_s$, that this is equivalent to computing the excess arbitrage under those externalities.

Figure 3 displays the results graphically. The actual excess arbitrage in pounds is constructed from OIS rates in dollars and pounds, and the spot and forward dollar-pound exchange rates (using excess arbitrage eliminates the dependence on the IOER rate). The predicted excess arbitrage is constructed entirely from those same variables in euros and yen, along with options prices on all six possible currency pairs, which are used to both construct the externality mimicking portfolio (in the matrix Σ^i) and to construct the covariances Σ^i_{GBP}. In other words, the set of financial instruments used to construct the actual and predicted excess arbitrages do not overlap at all. Nevertheless, the predicted and actual excess arbitrages closely track each other. Regressing the actual excess arbitrage on the predicted excess arbitrage, without a constant, results in an R^2 of 74%. This suggests that the mimicking portfolio constructed from euros and yen tracks the unobservable risk-neutral externalities reasonably well (or at least that adding pounds does not improve things much).

Figure 3: Actual vs. Predicted Excess Arbitrage in Pounds



I now turn to the question of what the estimates of the externality-mimicking

Table 2: Stress Test "Severely Adverse" Scenarios

	Euro		Pound		Yen	
Stress Test Date	One-Quarter Return	Four-Quarter Return	One-Quarter Return	Four-Quarter Return	One-Quarter Return	Four-Quarter Return
9/30/12	-8.0	-16.7	0.0	0.0	2.9	-1.0
9/30/13	-15.4	-24.1	-13.4	-13.4	3.1	-1.1
9/30/14	-12.7	-14.4	-3.1	-3.6	7.9	6.7
12/31/15	-8.1	-15.0	-2.5	-3.9	2.8	5.2
12/31/16	-9.5	-12.7	-4.6	-9.2	3.3	7.8

Notes: Explain table here

portfolio imply about policy, and in particular whether the totality of the regulations on banks and other intermediaries is having an effect that is aligned with regulators' goals. This, of course, requires some notion of the regulators' perceived externalities; in other words, in which states would the regulators like intermediaries to have more wealth (negative $\Delta_s^{h,i}$) and in which states would the regulators like the households to have more wealth (positive $\Delta_s^{h,i}$). I will consider the "stress test scenarios" developed by the Federal Reserve to be a statement of this sort. Once per year, the Federal Reserve describes a "severely adverse" scenario and requires banks to maintain various leverage and capital ratios in this scenario. The scenario is specified in terms of asset prices, including the dollar-euro, dollar-yen, and dollar-pound exchange rates (the last of these will not play any role in the analysis). In table 2 below, I report the returns of the yen, euro, and pound in the stress test scenarios, at both the one quarter and four quarter horizons, for each stress test conducted. A general pattern emerges: recent stress tests have involved sizable euro depreciations relative to the dollar, and sizable yen appreciations. This pattern is consistent with the observation that, during my sample, stock market declines tend to coincide with euro depreciation and yen appreciation relative to the dollar, and that these sorts of correlations might influence how the Federal Reserve constructs the stress test scenarios.

I study the returns of the externality-mimicking portfolio in each of these scenarios. Unfortunately, the scenarios are specified on a quarterly basis, whereas my approach to the construction of the externality-mimicking portfolio has been to focus on monthly returns. The choice to emphasize monthly returns was driven by the need to observe a risk-free rate available to intermediaries (the IOER rate), and this rate is available at a daily frequency and (with very high probability) changes only on FOMC meetings dates. In what follows, I will present the returns of the one-month externality mimicking portfolio, pretending that the currency returns of the stress test scenario at either the one quarter or four quarter horizons in fact occur over a single month.

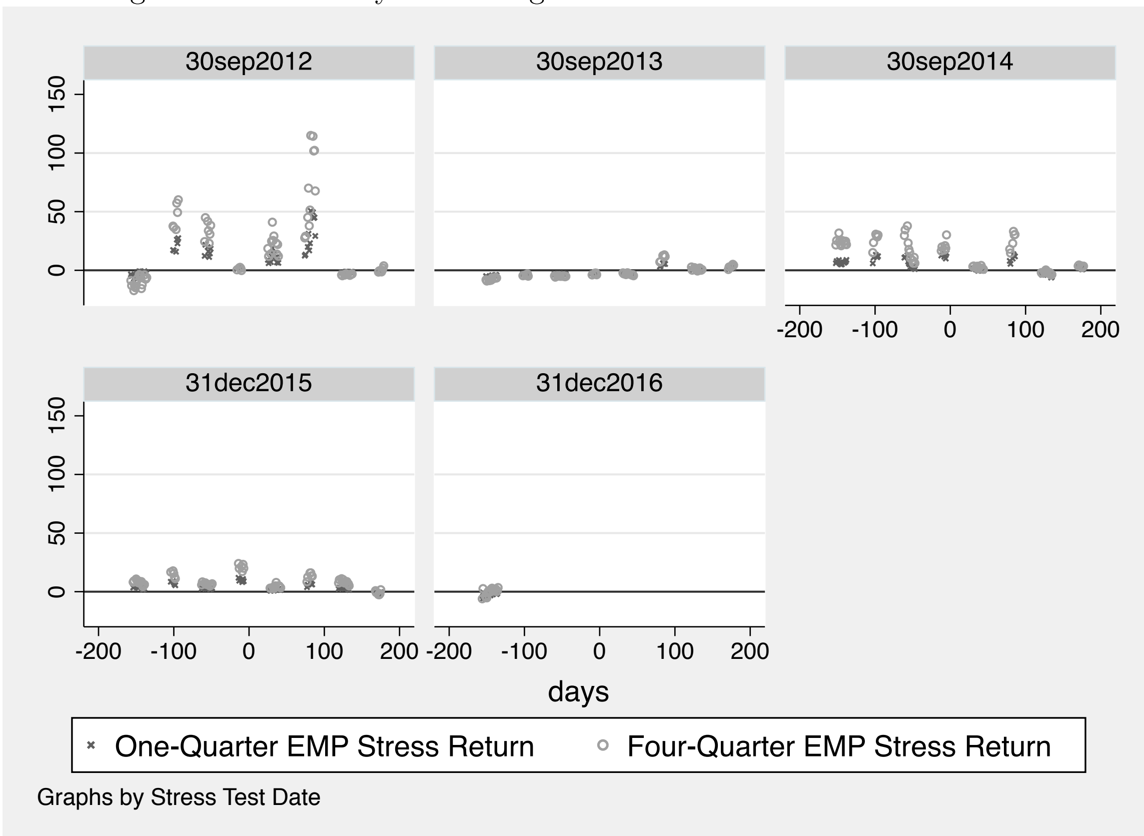
Each of the stress test scenarios is associated with a particular date (listed in table 2) which is the date at which the scenario starts. For each date in my sample (at least 23 non-weekend days before the next FOMC meeting) that is also within 180 calendar days of the stress test date, I report the returns of the externality-mimicking portfolio constructed from that day's interest rates, spot and forward exchange rates, and exchange rate options under the associated stress test scenario. Requiring that the relevant financial market data come from a day that is within 180 days of the stress test date effectively assigns almost all of the days in my sample to a single stress test date, dropping a handful of days that are far from any stress test date.

In figure 4 below, I report the returns of the externality mimicking portfolio for each of these days, under the associated stress scenario. The return of the externality-mimicking portfolio is "unit-less," and has a particular meaning. A 10% positive return means that in the specified scenario, to rationalize the observed arbitrages, the social planner must value wealth in the hands of households 10% more than it values wealth in the hands of intermediaries, due to externalities, on average. The "on average" caveat here refers to the projection of the true externalities onto the space of returns. There might be some states in which the portfolio returns 10% but the externalities are even larger, and other states in which the portfolio returns 10% but the externalities are smaller, or even go the opposite direction, but the average externality over these states (under the risk-neutral probabilities) will be 10%.

The purpose of the stress test, however, is to ensure that intermediaries have a sufficient level of wealth in a particular scenario. To the extent that the stress tests are effective, they must operate by inducing the intermediaries to hold different assets and issue different liabilities than they otherwise would have. Consequently, the intermediaries' counterparties (the households) must also hold different assets and issue different liabilities than they otherwise would have. In other words, if the regulations act to raise intermediaries' wealth in certain scenarios, they must lower the wealth of households in those scenarios (at least in an endowment economy). That is, the stress test scenarios should be taken as a statement when the regulator perceives negative externalities associated with transferring wealth from intermediaries to households (negative $\Delta_s^{h,i}$). Consequently, we would expect, if the regulations were having the desired effect, that the return on the externality mimicking portfolio in the stress test scenario would be sharply negative.

What I find in the data, however, is that this is not the case. In both 2014 and 2015, the arbitrage on yen was larger than the arbitrage on euros (figure 1), and as a result the externality mimicking portfolio placed more weight on yen than euros (figure 2). The yen appreciation in the stress test scenario was large enough, given this extra weight, to more than offset the euro depreciation, and as a result the return on the externality mimicking portfolio was positive. In other words, even though the Federal Reserve would like the intermediaries to have more wealth in the stress test scenario, the cumulative effect of all regulations (by the Fed and other entities) acted to encourage the banks to have less wealth in the stress test scenario. The situation was better for the 2012 and 2013 stress tests, and appears better at the end of the data for the 2016 test, mainly due to a higher weight on euros in the portfolio, which is caused by an increase in the magnitude of the euro CIP violation. Note the surprising logic of this statement: an increase in the magnitude of arbitrage can be understood as a sign that regulation is working more effectively. To aid the comparison between arbitrages, portfolio weights, and portfolio returns in the stress scenarios, section §A in the appendix shows each of these on the same axis (days relative to the stress test date), separately for each of the stress tests.

Figure 4: Externality Mimicking Portfolio Returns in Stress Scenarios



At first glance, the results for 2014 and 2015 seem like a contradiction. If the stress test requires intermediaries to have more wealth when the euro depreciates and the yen appreciates, shouldn't this have an effect on intermediaries' willingness to own euros and yen, and hence be reflected in market prices? How can we reconcile the fact that the stress test goes in "opposite directions" for euro and yen, but the arbitrages we observe go in the "same direction"? Although I cannot provide a definitive answer to this question, I will sketch a "story" that can explain these results. To explain the existence of arbitrage, it must be the case that households (and institutions like mutual funds that act on their behalf) are unable to execute the arbitrage, or face prohibitively high costs, and the evidence of Rime et al. [2017] supports this hypothesis. At the same time, there must be constraints on banks that raise the cost of conducting the arbitrage. Leverage constraints, such as the "supplementary leverage ratio," described in D'Hulster [2009], have been suggested by Du et al. [2017] as a relevant constraint, and are perhaps the only plausible interpretation of the IOER-

OIS basis in the United States. These constraints act to raise the cost of conducting the CIP arbitrage, regardless of "sign" of the arbitrage. Finally, as described by Du et al. [2017], the direction of the CIP arbitrage across currencies is predicted by the direction of the "carry trade" (the interest rate differential). Following Du et al. [2017], I interpret this as a sign of household demand to trade in particular directions, and banks being induced by arbitrage returns to take the other side of these trades.

Specifically, over the relevant time period, yen and euro interest rates were lower than dollar interest rates. As a result, households attempting to engage in the carry trade would wish to sell euros or yen and purchase dollars, and to save those dollars at prevailing interest rates. Intermediaries taking the other side of these trades would be, instead of saving dollars, purchasing euros or yen and saving at those countries' lower interest rates. To induce intermediaries make these trades, purchasing euros or yen and saving in those currencies must be a "good" deal from the intermediaries' perspective. This "good deal" manifests itself as arbitrage, with respect to currency forwards and the IOER rate, because of household's inability to trade in those instruments and intermediaries' leverage constraints. Because the sign of the interest rate differential is the same for dollar-euro and dollar-yen, households demand and the arbitrage induced by that demand move in the same direction. The stress test, to some degree, offsets this by encouraging banks to be "long" yen and "short" euros. However, this effect is dwarfed by demand from households (interpreted broadly) to engage in the carry trade, and the result is the arbitrage for banks often goes in the same direction, generating externality-mimicking portfolio weights that are not consistent with the stress test scenarios.

How might policy be altered so as to resolve this problem? A detailed exploration of this issue is beyond the scope of the paper, but the previous discussion will hint at one aspect of the problem. Demand from households will always induce intermediaries to take certain positions and not others, neglecting externalities (naturally). Differences in this demand across assets, differences in the elasticity of this demand across assets, and differences over time all imply that regulations that treat products symmetrically (without regards to these demands), such as leverage constraints, will have differential effects across assets and may not cause the desired reallocation of wealth across states of the world. In this view, exercises such as the stress test are a step in the right direction, but, as the results above demonstrate, not sufficient as currently implemented.

6 Extensions

6.1 Micro-Prudential Motives for Regulation

6.2 Borrowing Constraints and Price Rigidities

7 Conclusion

References

Manuel Amador, Javier Bianchi, Luigi Bocola, and Fabrizio Perri. Exchange rate policies at the zero lower bound. Technical report, National Bureau of Economic Research, 2017.

Leif Andersen, Darrell Duffie, and Yang Song. Funding value adjustments. Technical report, National Bureau of Economic Research, 2017.

Morten L Bech and Elizabeth Klee. The mechanics of a graceful exit: Interest on reserves and segmentation in the federal funds market. *Journal of Monetary Economics*, 58(5):415–431, 2011.

Fischer Black. The pricing of commodity contracts. *Journal of financial economics*, 3(1-2):167–179, 1976.

Falk Bräuning and Kovid Puria. Uncovering covered interest parity: the role of bank regulation and monetary policy. 2017.

Eduardo Dávila and Anton Korinek. Pecuniary externalities in economies with financial frictions. *The Review of Economic Studies*, page rdx010, 2017.

Katia D'Hulster. The leverage ratio: A new binding limit on banks. 2009.

Wenxin Du, Alexander Tepper, and Adrien Verdelhan. Deviations from covered interest rate parity. Technical report, National Bureau of Economic Research, 2017.

Darrell Duffie and Arvind Krishnamurthy. Passthrough efficiency in the fed's new monetary policy setting. 2016.

Emmanuel Farhi and Iván Werning. A theory of macroprudential policies in the presence of nominal rigidities. *Econometrica*, 84(5):1645–1704, 2016.

Matthias Fleckenstein, Francis A Longstaff, and Hanno Lustig. The tips-treasury bond puzzle. *the Journal of Finance*, 69(5):2151–2197, 2014.

Nicolae Garleanu and Lasse Heje Pedersen. Margin-based asset pricing and deviations from the law of one price. *The Review of Financial Studies*, 24(6):1980–2022, 2011.

John Geanakoplos and Heraklis M Polemarchakis. Existence, regularity, and constrained suboptimality of competitive allocations when the asset market is incomplete. *Uncertainty, information and communication: essays in honor of KJ Arrow*, 3:65–96, 1986.

Denis Gromb and Dimitri Vayanos. Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of financial Economics*, 66(2):361–407, 2002.

Denis Gromb and Dimitri Vayanos. Limits of arbitrage. *Annu. Rev. Financ. Econ.*, 2(1):251–275, 2010.

Lars Peter Hansen and Ravi Jagannathan. Implications of security market data for models of dynamic economies. *Journal of political economy*, 99(2):225–262, 1991.

Lars Peter Hansen and Scott F Richard. The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica: Journal of the Econometric Society*, pages 587–613, 1987.

Victoria Ivashina, David S Scharfstein, and Jeremy C Stein. Dollar funding and the lending behavior of global banks. *The Quarterly Journal of Economics*, 130(3): 1241–1281, 2015.

Gordon Y Liao et al. Credit migration and covered interest rate parity. *Project on Behavioral Finance and Financial Stability Working Paper Series*, (2016-07), 2016.

Semyon Malamud and Andreas Shrimpf. An intermediation-based model of exchange rates. Technical report, Working Paper, 2017.

Dagfinn Rime, Andreas Schrimpf, and Olav Syrstad. Segmented money markets and covered interest parity arbitrage. 2017.

Mirela Sandulescu, Fabio Trojani, and Andrea Vedolin. Model-free international sdfs in incomplete markets. 2017.

Andrei Shleifer and Robert W Vishny. The limits of arbitrage. *The Journal of Finance*, 52(1):35–55, 1997.

A Additional Figures

Figure 5: Arbitrages, Weights, and Returns for the 9/30/12 Stress Test

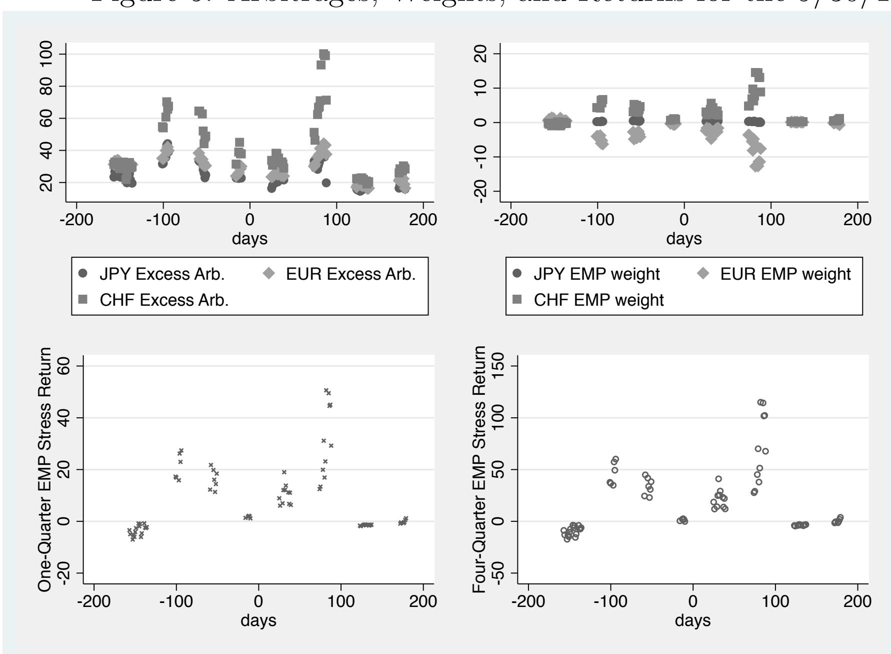


Figure 6: Arbitrages, Weights, and Returns for the 9/30/13 Stress Test

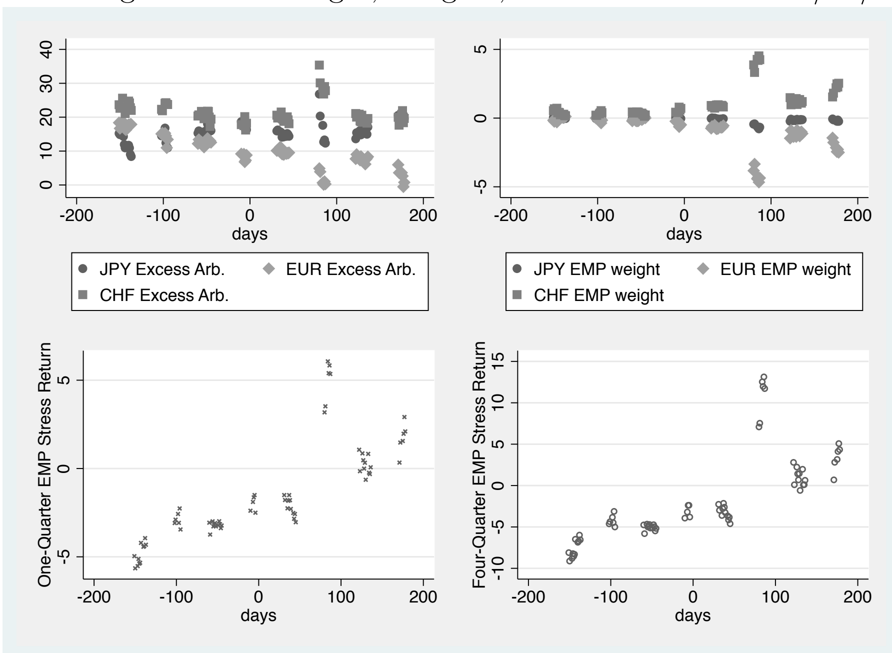


Figure 7: Arbitrages, Weights, and Returns for the 9/30/14 Stress Test

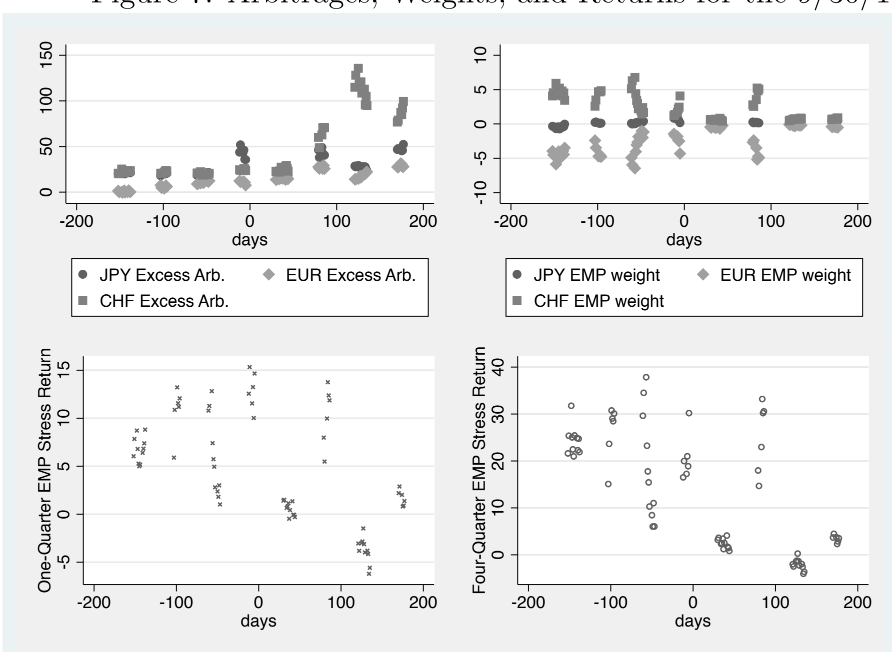


Figure 8: Arbitrages, Weights, and Returns for the 12/31/15 Stress Test

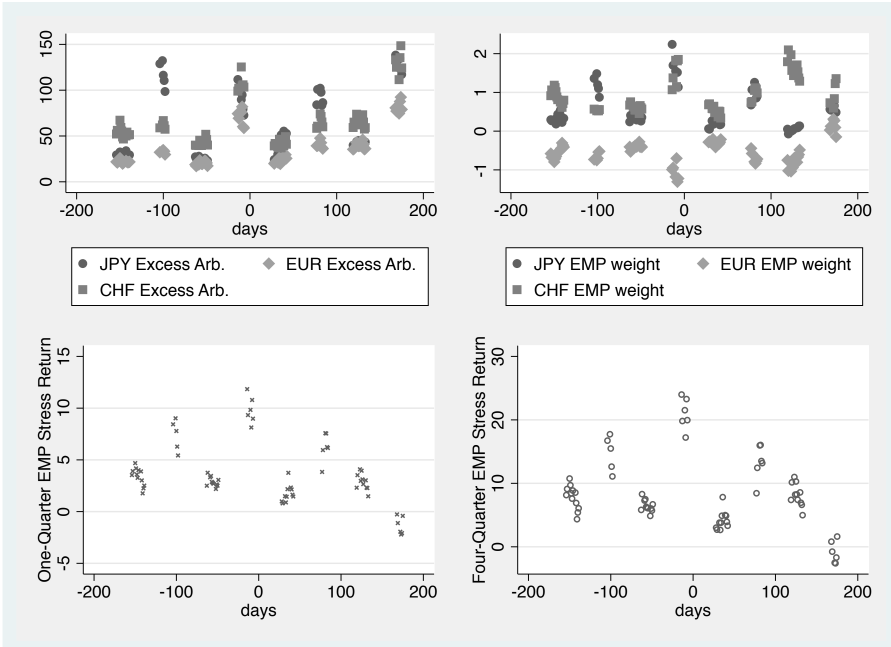
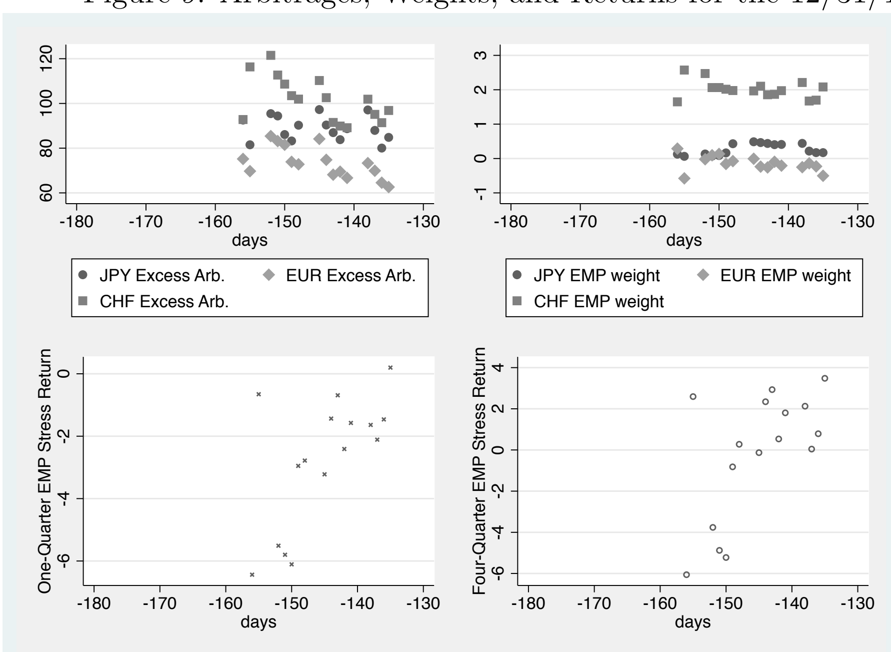


Figure 9: Arbitrages, Weights, and Returns for the 12/31/16 Stress Test



B Data Description

- All FX spot, forward, and options data, and all OIS rates, are London close from Bloomberg
- Arbitrage calculations consistent with Du et al. [2017]
- FX options volatilities rescaled to adjust for differences between IOER and OIS rates
- IOER rates from FRED
- FOMC dates from Federal Reserve website
- Sample period: 3 Jan 2011 to 15 Sep 2016

C Proofs

C.1 Wedges

The FOCs for the planner, for assets D_a^h and D_a^i, are

$$\rho^h \cdot \Phi_a^h + \sum_{s\in S} [\lambda^h V_{I,s}^h - \mu_s \cdot X_{I,s}^h] Z_{a,s} = \psi_a,$$

$$\sum_{s\in S} [\lambda^i V_{I,s}^i - \mu_s \cdot X_{I,s}^i] Z_{a,s} = \psi_a.$$

By comparison, the FOC the the agents are

$$\xi^h \cdot \Phi_a^h + \sum_{s\in S} \lambda^h V_{I,s}^h (Z_{a,s} - Q_a \mathbf{1}(s = s_0)) = 0,$$

and

$$\xi^i \cdot \Phi_a^i + \sum_{s\in S} \lambda^i V_{I,s}^i (Z_{a,s} - Q_a \mathbf{1}(s = s_0)) = 0$$

The FOC for goods prices in state s is

$$\sum_{i\in I} [\lambda^i V_{P,s}^i - \mu_s \cdot X_{P,s}^i + \lambda^i V_{I,s}^i Y_s^i] + \sum_{h\in H} [\lambda^h V_{P,s}^h - \mu_s X_{P,s}^h + \lambda^h V_{I,s}^h Y_s^h] = 0.$$

The FOC for transfers is

$$\kappa = \lambda^h V^h_{I,s_0}$$

and

$$\kappa = \lambda^i V^i_{I,s_0}.$$

We can rewrite the private FOCs as

$$\begin{aligned}\kappa Q_a - \psi_a &= (\xi^h - \rho^h) \cdot \Phi^h_a + \sum_{s\in S} \mu_s X^h_{I,s} Z_{a,s} \\ &= \xi^i \cdot \Phi^i_a + \sum_{s\in S} \mu_s X^i_{I,s} Z_{a,s}\end{aligned}$$

We use the identities

$$S^i_s = X^i_{P,j,s} + X^i_{I,s}(X^i_s)^T$$

$$V^i_{P,s} = -V^i_{I,s} X^i_s$$

Hence, we can write

$$\begin{aligned}&\sum_{i\in I}[-\lambda^i V^i_{I,s}(X^i_s + Y^i_s) - \mu_s \cdot S^i_s + (\mu_s \cdot X^i_{I,s})X^i_s]+ \\ &\sum_{h\in H}[-\lambda^h V^h_{I,s}(X^h_s + Y^h_s) - \mu_s \cdot S^h_s + (\mu_s \cdot X^h_{I,s})X^h_s] = 0.\end{aligned}$$

Using the Arrow securities,

$$\lambda^i V^i_{I,s} = \mu_s \cdot X^i_{I,s} + \psi_s$$

$$\lambda^h V^h_{I,s} = \mu_s \cdot X^h_{I,s} + \psi_s - \rho^h \cdot \Phi^h_s.$$

Therefore,

$$\begin{gathered}\sum_{i\in I}[-\lambda^i V^i_{I,s} Y^i_s - \mu_s \cdot S^i_s - \psi_s X^i_s]+ \\ \sum_{h\in H}[-\lambda^h V^h_{I,s} Y^h_s - \mu_s \cdot S^h_s - \psi_a X^h_s + (\rho^h \cdot \Phi^h_s)X^h_s] = 0.\end{gathered}$$

Define

$$\mu_s = \bar{u}_s P_s - \kappa \tau_s D(P_s),$$

so that

$$\tau_s \cdot P_s = 0.$$

We can rewrite the equations as

$$\sum_{i \in I} [-\lambda^i V^i_{I,s} Y^i_s + \kappa \tau_s D(p_s) \cdot S^i_s - \psi_s X^i_s] +$$
$$\sum_{h \in H} [-\lambda^h V^h_{I,s} Y^h_s + \kappa \tau_s D(p_s) \cdot S^h_s - \psi_s X^h_s + (\rho^h \cdot \Phi^h_s) X^h_s] = 0.$$

and

$$\lambda^i V^i_{I,s} = \bar{\mu}_s - \kappa \tau_s D(P_s) \cdot X^i_{I,s} + \psi_s$$
$$\lambda^h V^h_{I,s} = \bar{\mu}_s - \kappa \tau_s D(P_s) \cdot X^h_{I,s} + \psi_s - \rho^h \cdot \Phi^h_a.$$

Using the Arrow securities and the private FOCs,

$$\begin{aligned} \kappa Q_s - \psi_s &= (\xi^h - \rho^h) \cdot \Phi^h_s + \bar{u}_s - \kappa \tau_s D(P_s) X^h_{I,s} \\ &= \xi^i \cdot \Phi^i_s + \bar{u}_s - \kappa \tau_s D(P_s) X^i_{I,s} \end{aligned}$$

and therefore

$$\kappa Q_s = \lambda^h V^h_{I,s} + \xi^h \Phi^h_s = \lambda^i V^i_{I,s} + \xi^i \Phi^i_s$$

Define

$$\begin{aligned} \chi_a &= (-Q_a + \sum_{s \in S} Z_{a,s} Q_s) \\ &= \kappa^{-1} \xi^i \cdot (-\Phi^i_a + \sum_{s \in S} Z_{a,s} \Phi^i_s) \end{aligned}$$

If an asset is tradeable by household h,

$$\kappa Q_a - \psi_a = -\sum_{s \in S} \kappa \tau_s D(P_s) X^h_{I,s} Z_{a,s} = \xi^i \cdot \Phi^i_a - \sum_{s \in S} \kappa \tau_s D(P_s) X^i_{I,s} Z_{a,s}$$

Therefore, with no regulation of the Arrow market,

$$\begin{aligned}
\chi_a &= \sum_{s\in S} \tau_s D(P_s)(X^h_{I,s} - X^i_{I,s})Z_{a,s} \\
&= \sum_{s\in S} (\frac{\lambda^i}{\kappa} V^i_{I,s} - \frac{\lambda^h}{\kappa} V^h_{I,s})Z_{a,s}
\end{aligned}$$

$$\rho^h \cdot \Phi^h_a + \sum_{s\in S} [\lambda^h V^h_{I,s} - \mu_s \cdot X^h_{I,s}]Z_{a,s} = \psi_a,$$

$$\sum_{s\in S} [\lambda^i V^i_{I,s} - \mu_s \cdot X^i_{I,s}]Z_{a,s} = \psi_a.$$

C.2 Proof of lemma 1

The divergence problem is

$$\hat{\pi}^h = \min_{\pi\in\mathcal{P}(S)} D_{\chi^2}(\pi||\pi^i)$$

subject to, for each $a \in A^{arb}$,

$$(R^h)^{-1} \sum_{s\in S_1} \pi_s Z_{a,s} + \chi_a = (R^i)^{-1} \sum_{s\in S_1} \pi^i_s Z_{a,s}.$$

Because prices are positive, we can rescale the constraint. Defining Q^i_a as the price of the replicating portfolio of securities,

$$Q^i_a = (R^i)^{-1} \sum_{s\in S_1} \pi^i_s Z_{a,s},$$

and $R_{a,s}$ as the excess return (for all assets other than the risk-free asset),

$$R_{a,s} = \begin{cases} R^i & a = RF \\ \frac{Z_{a,s}}{Q^i_a} - R^i & a \neq RF \end{cases},$$

and

$$\chi_R = \frac{1}{R^i} - \frac{1}{R^h},$$

we can rewrite the constraint as

$$\sum_{s \in S_1} \pi_s R_{a,s} = \begin{cases} -R^h(\frac{\chi_a}{Q_a^i} - R^i \chi_R) & a \neq RF \\ R^i & a = RF. \end{cases}$$

Note that, for the risk-free asset, this constraint enforces that

$$\sum_{s \in S_1} \pi_s = 1$$

under the assumption that $R^i > 0$.

Defining the χ^2 divergence

$$D_{\chi^2}(\pi || \pi') = \frac{1}{2} \sum_{s \in S_1} \pi_s^{'} (\frac{\pi_s}{\pi_s^{'}} - 1)^2,$$

the Lagrangian is

$$\begin{aligned} \min_{\pi \in \mathcal{P}(S_1)} \max_{\theta \in \mathbb{R}^{|A^{arb}|}, \nu \in \mathbb{R}_+^{|S|}} \frac{1}{2} \sum_{s \in S_1} \pi_s^i (\frac{\pi_s}{\pi_s^i} - 1)^2 - \sum_{s \in S_1} \nu_s \pi_s \\ - \sum_{a \in A^{arb} \setminus \{a_{RF}\}} \theta_a [\sum_{s \in S_1} (\pi_s R_{a,s}) + R^h (\frac{\chi_a}{Q_a^i} - R^i \chi_R)] - \\ \theta_{RF} (\sum_{s \in S_1} \pi_s R^i - R^i) \end{aligned}$$

We can write the first-order condition for π_s

$$\frac{\hat{\pi}_s^h}{\pi_s^i} - 1 - \nu_s - \sum_{a \in A^{arb}} \theta_a R_{a,s} = 0.$$

where θ_a denotes the multiplier on the constraint. By complementary slackness, we must have either $\nu_s = 0$ or $\hat{\pi}_s^h = 0$, and hence will have

$$\nu_s = \max(-1 - \sum_{a \in A^{arb}} \theta_a R_{a,s}, 0)$$

and

$$\hat{\pi}_s^h = \pi_s^i \max(\sum_{a \in A^{arb}} \theta_a R_{a,s} + 1, 0).$$

Under the assumption that $\hat{\pi}^h$ has full support, $\nu_s = 0$ for all s. Plugging this into the Lagrangian, the problem solves

$$\max_{\theta \in \mathbb{R}^{|A^{arb}|}} \frac{1}{2} \sum_{s \in S_1} \pi^i_s (\sum_{a \in A^{arb}} \theta_a R_{a,s})^2 -$$
$$\sum_{a \in A^{arb} \setminus \{a_{RF}\}} \theta_a [\sum_{s \in S_1} \pi^i_s R_{a,s} (\sum_{a' \in A^{arb}} \theta_{a'} R_{a',s} + 1) + R^h (\frac{\chi_a}{Q^i_a} - R^i \chi_R)] -$$
$$\theta_{RF} (\sum_{s \in S_1} \pi^i_s R^i (\sum_{a' \in A^{arb}} \theta_{a'} R_{a',s} + 1) - R^i),$$

This expression simplifies (using $\sum_{s \in S_1} \pi^i_s R_{a,s} = 0$ for all $a \in A^{arb} \setminus \{a_{RF}\}$) to

$$\max_{\theta \in \mathbb{R}^{|A^{arb}|}} -\frac{1}{2} \sum_{s \in S_1} \pi^i_s (\sum_{a \in A^{arb} \setminus \{a_{RF}\}} \theta_a R_{a,s})^2 -$$
$$\sum_{a \in A^{arb} \setminus \{a_{RF}\}} \theta_a (R^h (\frac{\chi_a}{Q^i_a} - R^i \chi_R)) -$$
$$\frac{1}{2} (\theta_{RF} R^i)^2.$$

This expression is concave in θ, and the first-order condition for $\hat{\theta}_a$ (the optimum) is

$$\sum_{a' \in A^{arb}} \hat{\theta}_{a'} \Sigma^i_{a',a} = - R^h (\frac{\chi_a}{Q^i_a} - R^i \chi_R),$$

where $\Sigma^i_{a',a} = \sum_{s \in S_1} \pi^i_s R_{a,s} R_{a',s}$ for $a, a' \in A^{arb} \setminus \{a_{RF}\}$. Note that, trivially, $\hat{\theta}_{RF} = 0$. Solving, for $a \in A^{arb} \setminus \{a_{RF}\}$,

$$\hat{\theta}_{a'} = -R^h R^i \sum_{a \in A^{arb} \setminus \{a_{RF}\}} \frac{\chi_a}{\sum_{s' \in S_1} \pi^i_{s'} Z_{a,s'}} - \chi_R)(\Sigma^i)^{-1}_{a,a'}.$$

Therefore, using the definition of $\hat{\Delta}_s^{h,i}$ (equation (2)),

$$\begin{aligned}
\hat{\Delta}_s^{h^*,i} &= \chi_R + \frac{1}{R^h}(1 - \frac{\hat{\pi}_s^h}{\pi_s^i}) \\
&= \chi_R - \frac{1}{R^h} \sum_{a' \in A^{arb}} \hat{\theta}_{a'} R_{a',s} \\
&= \chi_R + R^i \sum_{a,a' \in A^{arb} \setminus \{a_{RF}\}} (\frac{\chi_a}{\sum_{s' \in S_1} \pi_{s'}^i Z_{a,s'}} - \chi_R)(\Sigma^i)^{-1}_{a,a'} R_{a,s} \\
&= \bar{\Delta}_s^{h,i},
\end{aligned}$$

which is the result.

C.3 Proof of lemma 2

First, observe by construction that the Lagrangian minimizes the variance over all SDFs with the correct expectation. Defining $m_s = \frac{\pi_s}{R^h \pi_s^i}$, the Lagrangian can be rewritten

$$\begin{aligned}
\min_{m \in \mathbb{R}^{|S|}} \max_{\theta \in \mathbb{R}^{|A^{arb}|}, \nu \in \mathbb{R}_+^{|S|}} & \frac{1}{2} R^h \sum_{s \in S_1} \pi_s^i (m_s - \frac{1}{R^h})^2 - \sum_{s \in S_1} \nu_s \pi_s^i m_s \\
& - \sum_{a \in A^{arb} \setminus \{a_{RF}\}} \theta_a [\sum_{s \in S_1} (\pi_s^i m_s R_{a,s}) + (\frac{\chi_a}{Q_a^i} - R^i \chi_R)] - \\
& R^i \theta_{RF} (\sum_{s \in S_1} \pi_s^i m_s - \frac{1}{R^h}).
\end{aligned}$$

Noting that the last constraint enforces

$$\sum_{s \in S_1} \pi_s^i m_s = \frac{1}{R^h},$$

implying that

$$\begin{aligned}
\sum_{s\in S_1}(\pi^i_s m_s R_{a,s}) + (\frac{\chi_a}{Q^i_a} - R^i\chi_R) &= \\
\sum_{s\in S_1}(\pi^i_s m_s \frac{Z_{a,s}}{Q^i_a}) - \frac{R^i}{R^h} + \frac{Q^i_a - Q_a}{Q^i_a} - 1 + \frac{R^i}{R^h} &= \\
\sum_{s\in S_1}(\pi^i_s m_s \frac{Z_{a,s}}{Q^i_a}) - \frac{Q_a}{Q^i_a} &= \\
\frac{Q^a}{Q^i_a}(\sum_{s\in S_1}(\pi^i_s m_s \frac{Z_{a,s}}{Q_a}) - 1)&.
\end{aligned}$$

Defining

$$R^h_{a,s} = \begin{cases} \frac{Z_{a,s}}{Q_a} - R^h & a \neq a_{RF} \\ R^h & a = a_{RF}, \end{cases}$$

we can rewrite the problem as

$$\begin{aligned}
\min_{m\in\mathbb{R}^{|S|}} \max_{\theta\in\mathbb{R}^{|A^{arb}|},\nu\in\mathbb{R}^{|S|}_+} \frac{1}{2}\frac{\sum_{s\in S_1}\pi^i_s(m_s - \frac{1}{R^h})^2}{\sum_{s\in S_1}\pi^i_s m_s} - \sum_{s\in S_1}\nu_s\pi^i_s m_s \\
- \sum_{a\in A^{arb}\setminus\{a_{RF}\}}\theta_a\frac{Q^a}{Q^i_a}[\sum_{s\in S_1}(\pi^i_s m_s R^h_{a,s})] \\
-\theta_{RF}\frac{R^i}{R^h}(\sum_{s\in S_1}(\pi^i_s m_s R^h) - 1).
\end{aligned}$$

It follows that m is the "SDF" satisfying the HJ bounds.

Enforcing the constraint and taking first-order conditions,

$$R^h m^*_s - 1 - \nu_s - \sum_{a\in A^{ARB}}\tilde{\theta}_a R^h_{a,s} = 0$$

where $\tilde{\theta}_a = \theta_a\frac{Q^a}{Q^i_a}$. Under the assumption that the positivity constraints do not bind, the Lagrangian is

$$\max_{\tilde{\theta}\in\mathbb{R}^{|A^{arb}|}} -\frac{1}{2R^h}\sum_{s\in S_1}\pi_s^i(\sum_{a\in A^{ARB}}\tilde{\theta}_a R_{a,s}^h)^2$$
$$-\frac{1}{R^h}\sum_{a\in A^{arb}\setminus\{a_{RF}\}}\tilde{\theta}_a\sum_{s\in S_1}\pi_s^i R_{a,s}^h$$

In other words, $\tilde{\theta}$ is the minimum second-moment portfolio among all portfolios with the same expected excess return, and hence is mean-variance efficient. It follows that any other portfolio must have a lower Sharpe ratio" under the measure π^i:

$$|\frac{\sum_{s\in S_1}\sum_{a\in A^{ARB}\setminus\{a_{RF}\}}\pi_s^i\tilde{\theta}_a R_{a,s}^h}{(\sum_{s\in S_1}\pi_s^i(\sum_{a\in A^{ARB}\setminus\{a_{RF}\}}\tilde{\theta}_a R_{a,s}^h)^2)^{\frac{1}{2}}}| \geq |\frac{\sum_{s\in S_1}\sum_{a\in A^{ARB}\setminus\{a_{RF}\}}\pi_s^i\theta_a R_{a,s}^h}{(\sum_{s\in S_1}\pi_s^i(\sum_{a\in A^{ARB}\setminus\{a_{RF}\}}\theta_a R_{a,s}^h)^2)^{\frac{1}{2}}}|.$$

Define, for any portfolio θ,

$$\theta'_a = \frac{1}{R^h}\frac{Q_a^i}{Q_a}\theta_a,$$

and note that θ_a^* (the EMP portfolio) is the corresponding portfolio for $\tilde{\theta}_a$. It follows that

$$\theta_a R_{a,s}^h = \frac{Q^a}{Q_a^i}\theta'_a(R^h)^2 - R^h\theta'_a\frac{Z_{a,s}}{Q_a^i}$$
$$= \theta'_a R^h(\frac{Q^a}{Q_a^i}R^h - R^i) - \theta'_a R^h R_{a,s}^i.$$

Therefore,

$$\sum_{s\in S_1}\pi_s^i\theta_a R_{a,s}^h = \theta'_a R^h(\frac{Q^a}{Q_a^i}R^h - R^i)$$
$$= \theta'_a R^h(-R^h\frac{\chi_a}{Q_a^i} + R^h - R^i)$$
$$= \theta'_a(R^h)^2(R^i\chi_R - \frac{\chi_a}{Q_a^i}).$$

Therefore,

$$|\frac{\sum_{s\in S_1}\sum_{a\in A^{ARB}\setminus\{a_{RF}\}}\pi_s^i\theta_a R_{a,s}^h}{(\sum_{s\in S_1}\pi_s^i(\sum_{a\in A^{ARB}\setminus\{a_{RF}\}}\theta_a R_{a,s}^h)^2)^{\frac{1}{2}}}| = R^h|\frac{\sum_{a\in A^{ARB}\setminus\{a_{RF}\}}\theta'_a(R^i\chi_R - \frac{\chi_a}{Q_a^i})}{(\sum_{s\in S_1}\pi_s^i(\sum_{a\in A^{ARB}\setminus\{a_{RF}\}}\theta'_a R_{a,s}^i)^2)^{\frac{1}{2}}}|$$

It follows that if there was a portfolio with a higher arbitrage Sharpe ratio than θ^*, there would be a portfolio with a higher Sharpe ratio under π^i than $\tilde{\theta}$, a contradiction.