

Externalities as Arbitrage*

Benjamin Hébert (Stanford GSB)

March 19, 2018

PRELIMINARY AND INCOMPLETE;

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 apparent arbitrage opportunities imply that regulations are sub-optimal? No. I develop a 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, Valentin Haddad, Barney Hartman-Glaser, Bob Hodrick, Anton Korinek, Arvind Krishnamurthy, Hanno Lustig, Semyon Malamud, Tyler Muir, and Jesse Schreger, and seminar participants at the MIT Sloan Juniors Conference, Yale SOM, EPFL Lausanne, Stanford, UCLA, and UC Irvine, for helpful comments. I would particularly like to thank Wenxin Du for help understanding data sources, and Jonathan Wallen for outstanding research assistance. 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.” The returns of this portfolio across states of the world are a set externalities that would justify the observed pattern of arbitrage.

Using data on interest rates, foreign exchange spot and forward rates, and foreign exchange options, I construct an externality-mimicking portfolio. The weights in this portfolio are entirely a function of asset prices; no estimation is required. 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. As a result, in “bad times,” we would expect this portfolio to have negative returns, consistent with the idea that the planner would like to use regulations to encourage intermediaries to hold more wealth in these states.

To formally test this idea, I consider two definitions of “bad times.” First, intuitively, bad times can be defined as times in which the intermediaries have a high marginal utility of wealth. Using this definition, I show that it is sufficient to study the expected returns of the externality mimicking portfolio, and test if they are positive. Second, I define “bad times” using the stress test scenarios developed 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 that the expected return of the externality mimicking portfolio is generally negative, and that its returns in the stress tests conducted at the end of 2014 and 2015 are positive. This implies that the externalities that would justify current regulation are positive in bad times, which appears inconsistent with intuition and suggests that regulations are not having their desired effect. For a regulator, who presumably understands her own objective, the procedure developed in this paper could be used as a diagnostic tool to understand whether the regulations she imposes 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 inefficient 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 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 the main contribution of the 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 of 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] show 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 which 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, 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 becomes 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 em-

¹Bech 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.

phasizes the connections between covered interest parity violations and non-financial outcomes.

Unfortunately, there are only a small set of “arbitrageable” assets for which data are 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 unique² projection of the externalities 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

²The definition of the externalities references a probability measure, and the projection I employ is under this particular probability measure. Uniqueness here means unique given this measure.

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

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 R^2 of this predictive exercise suggests that the projection is reasonable. Next, I study the expected return of this portfolio, and the returns of the externality-mimicking portfolio in the “severely adverse” scenarios from the Federal Reserve’s stress tests.

I find that expected returns are generally negative, and for several of the stress tests that the returns are positive. This occurs because the portfolio is usually “long” yen, a currency that appreciates in the stress scenario and has low interest rates (and hence low expected returns). This inconsistency can be summarized by the observation that, in the stress test scenarios, the 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.

I explore the robustness of these conclusions in two ways. First, I construct an alternative externality mimicking portfolio using empirically estimated variances and covariances, and verify that these portfolio also experiences positive returns in certain stress tests. Second, I include in my analysis an arbitrage between an S&P 500 ETF and options on that ETF, akin to the classic “index-futures arbitrage” described by, for example MacKinlay and Ramaswamy [1988]. In theory, including such an arbitrage might be helpful, because the stock market is quite likely to be correlated with externalities motivating regulation. However, I show that, because the stock market is quite volatile relative to currencies, and this arbitrage is usually small, the asset in question has a low “Sharpe ratio due to arbitrage” and therefore is not a substantial part of the mimicking portfolio on most days.

Section 2 introduces the GEI framework, and describes its general efficiency prop-

erties. 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 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,

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

where $U^h(\{X_{j,s}^h\}_{j \in J_s}; s)$ 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 .

In each state $s \in S$, household $h \in H$ has an endowment of good $j \in J_s$ equal to $Y_{j,s}^h$. 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_a^h 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

$$I_s^h = T^h \mathbf{1}(s = s_0) + \sum_{j \in J_s} P_{j,s} Y_{j,s}^h + \sum_{a \in A} D_a^h (Z_{a,s} - Q_a \mathbf{1}(s = s_0)),$$

and the household's budget constraint is

⁴Separating contingent commodities into states and goods available in each state will give meaning to the financial structure described below.

$$\sum_{j \in J_s} P_{j,s} X_{j,s}^h \leq I_s^h.$$

The constraints on households' asset positions are summarized by

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

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

$$V^h(I_s^h, \{P_{j,s}\}_{j \in J_s}; s) = \max_{\{X_{j,s}^h \in \mathbb{R}^+\}_{j \in J_s}} U^h(\{X_{j,s}^h\}_{j \in J_s}; s)$$

subject to

$$\sum_{j \in J_s} P_{j,s} X_{j,s}^h \leq I_s^h.$$

The portfolio choice problem is

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

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 $a \in A \setminus A^h$, $D_a^h = 0$. To be precise, if $a \in A \setminus A^h$ and $D_a^h \neq 0$, 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 is 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, and hence unaffected by regulation, does not constrain the social planner. 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 $X_{j,s}^h$ and $X_{j,s}^i$, goods prices $P_{j,s}$, asset positions D_a^h and D_a^i , 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 $j \in J_s$,

$$\sum_{h \in H} (X_{j,s}^h - Y_{j,s}^h) = \sum_{i \in I} (X_{j,s}^i - Y_{j,s}^i)$$

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

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

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 every intermediary, and the price of asset $a \in A^h$ will be pinned down by competition between intermediaries. The equilibrium definition supposes that this will continue to be the 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. Such a policy is unlikely to be 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

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

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. I employ the follow definition of arbitrage:

Definition 2. For any security tradable by a household ($a \in \cup_{h \in H} A^h$), define the amount of arbitrage as

$$\chi_a = -Q_a + \sum_{s \in S_1} Z_{a,s} Q_s,$$

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

Using these definitions, I first point out that without regulation, the economy is

constrained inefficient.

Proposition 1. *Absent regulation, there is no arbitrage: $\chi_a = 0$ for all $a \in \cup_{h \in H} A^h$. 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. Every asset $a \in A^h$ and its portfolio of replicating securities in A^I are perfect substitutes from the perspective of intermediaries, and therefore must have the same prices. 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 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: $\chi_a \neq 0$ for some $a \in \cup_{h \in H} A^h$.*

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. The arbitrage is equal to shadow cost that the intermediary pays due to the regulation (along the lines of Garleanu and Pedersen [2011]).

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, by looking at the patterns of arbitrage across financial assets, regulators can assess whether the regulations they implement are having their desired effects.

3 Arbitrage and Wedges

I begin this section by defining the “wedges,” $\tau_{r,j,s}$, which are defined for each state $s \in S$ and good $j \in J_s$, using a “reference” probability measure π_s^r . 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. The reference measure π_s^r is an arbitrary full-support probability distribution over the states $s \in S_1$.⁵ Let $\pi_s^r \mu_{r,j,s}$ 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’s date

⁵In the applications I consider, this distribution will be either a risk-neutral measure or the physical measure.

zero budget constraint. Define the wedges $\tau_{r,j,s}$ using an orthogonal projection of the multipliers on to prices:

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

with $\sum_{j \in J_s} \tau_{r,j,s} P_{j,s} = 0$ for all $s \in S$.⁶ Note that the multipliers $\mu_{r,j,s}$, and the associated wedges $\tau_{r,j,s}$, are defined in the context of the reference distribution π_s^r , and if defined instead under an alternative reference distribution $\pi_s^{r'}$ would be rescaled, $\tau_{r',j,s} = \frac{\pi_s^r}{\pi_s^{r'}} \tau_{r,j,s}$.

The multipliers $\mu_{r,j,s}$ 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 $\mu_{r,j,s}$ are not proportional to prices (i.e. that the $\tau_{r,j,s}$ are not zero), pecuniary externalities exist in the equilibrium. A high value of the wedge $\tau_{r,j,s}$ indicates that the multiplier $\mu_{r,j,s}$ 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 $\tau_{r,j,s}$ by the multiplier κ ensures that the units of these wedges are in “dollars,” rather than units of social utility.

These wedges 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 consume good j in state s out of income 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 $X_{I,j,s}^h$ denote the marginal effect that income has on the demand of household h for good j in state s , holding prices constant. If the wedge-weighted difference of the income effects for the household and the intermediary,

$$\Delta_{r,s}^{h,i} = \sum_{j \in J_s} \tau_{r,j,s} (X_{I,j,s}^h - X_{I,j,s}^i),$$

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. In what follows, I will reference to these $\Delta_{r,s}^{h,i}$ as externalities.

Transferring income from intermediary i to household h in state s 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

⁶This definition of wedges is essentially the same as the one employed by Farhi and Werning [2016], adjusted for the difference between production and endowment economies.

the assets tradable 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 $a \in A^h$,

$$\sum_{s \in S} \pi_s^r \Delta_{r,s}^{h,i} Z_{a,s} = \sum_{s \in S} \left(\frac{V_{I,s}^i}{V_{I,s_0}^i} - \frac{V_{I,s}^h}{V_{I,s_0}^h} \right) Z_{a,s}, \quad (1)$$

where $V_{I,s}^h$ denotes the marginal value of income for household h in state s , and $V_{I,s}^i$ is that object for intermediary i .

Put another way, we can view the externalities $\Delta_{r,s}^{h,i}$ 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 reference measure π^r . 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 externalities. This result is stated below in Proposition 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 $X_{I,j,s}^i = X_{I,j,s}^{i'}$ 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 allocation 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 externalities 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 intermediary $i \in I$,*

$$\chi_a = \sum_{s \in S} \pi_s^r \Delta_{r,s}^{h,i} Z_{a,s}.$$

Proof. See the appendix, ... □

In words, an asset tradable 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 ($\Delta_{r,s}^{h,i} > 0$). 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 other words, under an optimal policy, differences in the risk-free rates available to households and intermediaries reflect inter-temporal externalities associated with those agents’ consumption-savings decisions.

In the next section, I invert this exercise: given an observed pattern of arbitrage, what can we say about $\Delta_{r,s}^{h,i}$? 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 prices for a set of securities A^* that she believes are tradable by some household h , 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 $a \in A^*$, 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 at a higher rate (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.

The overnight maturity of the fed funds-IOER arbitrage illustrates an important point. For simplicity, I have assumed that the only constraints on intermediaries are induced by regulation. When taking the model to data, it is important to consider only arbitrages caused by regulation, which I will interpret to mean only arbitrages with a relatively short time until profit is guaranteed and whose existence appears to depend on the regulatory regime. For example, the fed funds-IOER trade converges daily, and various short term interest rates were all essentially identical before the financial crisis. Similarly, covered interest parity violations exist for maturities of days, weeks, and months, and did not exist before the financial crisis. In contrast, the TIPS/treasury/asset swap arbitrages studied by Fleckenstein et al. [2014]

take several years to converge, and pre-dated the financial crisis. I infer that other frictions preventing arbitrage aside from regulation might explain some or all these TIPS/treasury/asset swap arbitrages, and therefore do not use these arbitrages in my empirical exercises.

Suppose that we observe a set arbitrages that are caused by regulation, and we assume that the equation of Proposition 3 holds: for each $a \in A^*$,

$$\chi_a = \sum_{s \in S} \pi_s^r \Delta_{r,s}^{h,i} Z_{a,s}.$$

If the set A^* 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 externalities $\Delta_{r,s}^{h,i}$ (given the reference measure π_s^r). 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.⁷ As a result, it will not be possible, empirically, to recover $\Delta_{r,s}^{h,i}$ for all states $s \in S$.

We can, however, attempt to project the externalities $\Delta_{r,s}^{h,i}$ 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]. Hansen and Richard [1987] study the projection of a stochastic discount factor onto the space of returns. Those authors show that this procedure is equivalent to minimizing the variance of a stochastic discount factor, subject to the constraint that the SDF price a set of assets.⁸ Hansen and Jagannathan [1991] then show that the portfolio whose returns are the projection of the stochastic discount factor is also the portfolio with the highest Sharpe ratio. The projection, variance, and Sharpe ratio in these papers are computed under the physical probability measure.

I will provide analogous results for the externalities $\Delta_{r,s}^{h,i}$, which can be viewed as a difference in stochastic discount factors. I will consider the projection of the externalities onto the space of returns of arbitrageable assets, under the reference probability measure. I will show that this projection has a lower variance than the externali-

⁷The existence of complete market in A^* is also inconsistent with the assumption of limited participation.

⁸This equivalence holds if the constraint that the SDF be positive does not bind.

ties, again under the reference measure.⁹ I call the portfolio whose returns are the projection of the externalities the “externality-mimicking portfolio.” I show that the externality-mimicking portfolio has a higher “Sharpe ratio due to arbitrage” than any other portfolio, where the “Sharpe ratio due to arbitrage” is defined as the difference between the Sharpe ratio on the arbitrageable assets (tradable by households) and the Sharpe ratio on the replicating portfolio (tradable only by intermediaries). Again, this Sharpe ratio is defined with respect to the reference probability measure.

I assume that the set A^* includes exactly one risk-free security. I will also work in the space of returns, rather than payoffs, and therefore assume that every arbitrageable asset and its portfolio of replicating securities have strictly positive prices. This is without loss of generality, given that there is a risk-free arbitrageable security, as one could always add some amount of the risk-free security to another other security to ensure that its price is positive, while still ensuring that a replicating portfolio exists. I define the gross risk-free interest rate for intermediaries, R^i , as

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

and define another gross 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 also define returns under intermediary-only prices,

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

I define π_s^h and $R_{a,s}^h$ along similar lines, although in my empirical exercises I have found it more convenient to work with the intermediaries’ risk-neutral measure and returns. I will use the notation $\Sigma_r^{A^*,i}$ and $\mu_r^{A^*,i}$ to refer to the variance-covariance matrix and mean vector of the returns $R_{a,s}^i$ for all assets in A^* except the risk-free

⁹Following Hansen and Richard [1987], and for simplicity, I use an L^2 projection, and hence minimize variance and maximize the Sharpe ratio. Other authors (e.g. Sandulescu et al. [2017]) have explored projections using L^p norms.

asset, under the reference measure. Again, I will define $\Sigma_r^{A^*,h}$ and $\mu_r^{A^*,h}$ as the same variance-covariance matrix and vector, but for the returns available to households, $R_{a,s}^h$. I assume there are no redundant assets, meaning that the matrices $\Sigma_r^{A^*,i}$ and $\Sigma_r^{A^*,h}$ have full rank. I will also use the notation $Q^{A^*,h}$ and $Q^{A^*,i}$ to refer to the vector of risky asset prices and replicating portfolio, respectively, for assets in A^* .

Given a portfolio θ of risky assets in A^* , the “Sharpe ratio for intermediaries-only” is

$$S_r^{A^*,i}(\theta) = \frac{\frac{\theta^T \mu_r^{A^*,i}}{R^i} - \theta^T Q^{A^*,i}}{(\theta^T \Sigma_r^{A^*,i} \theta)^{\frac{1}{2}}}.$$

and define the “Sharpe ratio for households” along similar lines, with h in place of i . Note that this definition of the Sharpe ratio is signed (there is no absolute value), and that it might be scaled by the inverse of the gross risk-free rate when compared to other definitions. Note also that there is no requirement that the portfolio weights θ sum to one (the units of θ are “dollars”, not percentages), but the definition of the Sharpe ratio is homogenous of degree zero.

I will next define the “Sharpe ratio due to arbitrage.” First, observe that, because prices in the intermediary-only market are not the same as the prices available to households, an asset allocation in dollars to arbitrageable assets and the same dollar asset allocation to the replicating portfolios are in fact claims to different cashflows. For example, if both intermediaries and households can buy stocks at \$1/share but intermediaries pay \$2/bond whereas households pay \$3/bond, an allocation of \$4 split equally between stocks and bonds means two shares and one bond for the intermediaries, but one share and one bond for the households. To define the Sharpe ratio due to arbitrage, I would like instead to compare portfolios that are claims to the same cashflows (e.g. one share and one bond), but perhaps have different prices. To that end, define the portfolio transformation $\tilde{\theta}(\theta)$ by

$$\tilde{\theta}_a(\theta) = \frac{Q_a}{\sum_{s \in S} Z_{a,s} Q_s} \theta_a.$$

This transformation converts an asset allocation in dollars at intermediary-only prices to an asset allocation in dollars at prices available to the household with the same cashflow claims.

I define the “Sharpe ratio due to arbitrage” as the difference between the Sharpe ratio for households on a transformed portfolio and the Sharpe ratio for intermediaries-

only,

$$\hat{S}_r^{A^*,h,i}(\theta) = S_r^{A^*,h}(\tilde{\theta}(\theta)) - S_r^{A^*,i}(\theta).$$

Lastly, because I am working the space of returns, it is convenient to scale the amount of arbitrage on an asset by the price of the replicating portfolio of derivatives. I define percentage arbitrage as

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

Recall also that χ_{RF} is the dollar (not percentage) arbitrage on the risk-free asset, and is equal to the expected value of the externalities $\Delta_{r,s}^{h,i}$. I will use the notation $\tilde{\chi}^{A^*}$ and $R_s^{A^*,i}$ to refer to the vectors of percentage arbitrages $\tilde{\chi}_a$ and returns $R_{a,s}^i$ for each risky asset in A^* .

Using all these definitions, I define the externality-mimicking portfolio and describe its properties. The externality-mimicking portfolio has weights on the risky assets equal to

$$\theta_r^{A^*,i} = (\Sigma_r^{A^*,i})^{-1}(\tilde{\chi}^{A^*} - \mu_r^{A^*,i} \chi_{RF}), \quad (2)$$

and a weight on the risk-free asset equal to

$$\theta_{r,RF}^{A^*,i} = -(\theta_r^{A^*,i})^T \mu_r^{A^*,i} + \frac{\chi_{RF}}{R^i}. \quad (3)$$

The following lemma summarizes the properties of the externality mimicking portfolio.

Lemma 1. *The externality-mimicking portfolio with $\theta^{h,r,i}$ and $\theta_{RF}^{h,r,i}$ portfolio weights has the following properties:*

1. *The externalities can be expressed as the return on the externality mimicking portfolio plus a zero-mean residual, uncorrelated with the returns of all arbitrageable assets $a \in A^*$:*

$$\Delta_{r,s}^{h,i} = R^i \theta_{r,RF}^{A^*,i} + (R_s^{A^*,i})^T \theta_r^{A^*,i} + \epsilon_{r,s}^{A^*,i},$$

$$\sum_{s \in S_1} \pi_s^r R_{a,s}^i \epsilon_{r,s}^{A^*,i} = 0 \quad \forall a \in A^*.$$

2. *The variance of the externalities under the reference measure, $\sum_{s \in S_1} \pi_s^r (\Delta_{r,s}^{h,i} -$*

$\chi_{RF})^2$, is weakly greater than the variance of the externality-mimicking portfolio, $(\theta_r^{A^*,i})^T \Sigma_r^{A^*,i} \theta_r^{A^*,i}$.

3. The Sharpe ratio due to arbitrage of the externality-mimicking portfolio, $\hat{S}_r^{A^*,h,i}(\theta_r^{A^*,i})$, is weakly greater than the Sharpe ratio due to arbitrage of any other portfolio of assets in A^* .

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

The externality-mimicking portfolio is defined in the context of the reference measure π^r . In my empirical exercises, I consider (proxies for) two reference measures: the physical (or actual) probability measure, π^p , and the risk-neutral measure that prices the assets available only to intermediaries, π^i . The two corresponding externalities, $\Delta_{p,s}^{h,i}$ and $\Delta_{i,s}^{h,i}$, which I will refer to as the “physical externality mimicking portfolio” and “risk-neutral externality mimicking portfolio,” respectively, are linked,

$$\pi_s^p \Delta_{p,s}^{h,i} = \pi_s^i \Delta_{i,s}^{h,i}.$$

This connection reflects the usual equivalence in asset pricing between state-dependent preferences and beliefs. Using the risk-neutral externality mimicking portfolio, as opposed to the physical measure portfolio, has a particular advantage, which is that all expected returns are equal to the risk-free rate, and hence observable. Moreover, options prices (which I presume are traded only by intermediaries) can reveal the risk-neutral variances and (for currencies) covariances between different assets. Thus, restricting the set of arbitrageable assets to a risk-free asset and currencies, no estimation is required when constructing the risk-neutral externality mimicking portfolio, only a transformation of various asset prices. Using the physical measure, in contrast, requires estimating both expected returns and a variance-covariance matrix.

The externality-mimicking portfolio is a reflection of what regulation is actually accomplishing. Its returns represent the best linear projection of the externalities on to the space of returns. Consider a state in which the externality mimicking portfolio had a negative 10% return. The best linear prediction of the externalities is negative 10%, implying that for every dollar the planner transfers from households to intermediaries in that state, the planner alleviates externalities that are equivalent, in utility terms, to ten cents at state zero (the initial state), scaled by the probability of the negative 10% return state occurring.

Given this meaning, we should expect that the externalities (and hence the mimicking portfolio returns) are negative in “bad” states of the world. That is, empirically, governments seem tempted to bailout intermediaries in bad times, not in good times. To test whether regulations are consistent with this intuition, we need to define what we mean by “bad times.” I will pursue two definitions, which result in two different, and empirically feasible, tests. The first definition is to define bad times as being bad for the intermediaries, which is to say that the intermediaries’ stochastic discount factor ($\frac{\pi_s^i}{R^i \pi_s^p}$) is high. The second definition involves studying a particular situation in which the regulator is concerned about externalities, and would like intermediaries to have more wealth. I argue that the “stress tests” conducted by the Federal Reserve fall into this category—presumably, the idea is to ensure that intermediaries have sufficient wealth in the stress scenario so as to avoid a bailout ex-post—and therefore study the returns of the externality mimicking portfolio in the stress scenario.

The first approach yields a very simple test. The covariance of the intermediaries’ stochastic discount factor and the risk-neutral externality-mimicking portfolio, under the physical measure, is

$$\begin{aligned} Cov^p\left(\frac{\pi_s^i}{R^i \pi_s^p}, \Delta_{i,s}^{h,i}\right) &= \frac{\chi_R}{R^i} - \frac{1}{R^i} \sum_{s \in S_1} \pi_s^p \Delta_{i,s}^{h,i} \\ &= -\frac{1}{R^i} \theta_i^{A^*,i} (\mu_p^{A^*,i} - R^i) + Cov^p\left(\frac{\pi_s^i}{R^i \pi_s^p}, \epsilon_{i,s}^{A^*,i}\right). \end{aligned}$$

In other words, if the externalities are negatively correlated with the intermediaries’ SDF, the expected excess return of the risk-neutral externality mimicking portfolio under the physical measure should be positive. The error term in this equation shows that this test would be biased if there are components of the SDF that are not spanned by the space of returns, and which are correlated with components of the externalities that are also unspanned.

The second approach uses the stress test to identify a particular state in which externalities should be negative. Let $s^* \in S_1$ be the state corresponding to the stress scenario. We have

$$\Delta_{r,s^*}^{h,i} = R^i \theta_{r,RF}^{A^*,i} + (R_{s^*}^{A^*,i})^T \theta_r^{A^*,i} + \epsilon_{r,s^*}^{A^*,i},$$

and hence we expect that the return of both of the externality-mimicking portfolios (physical and risk-neutral) should be negative in the stress scenario. This test would be biased if the unspanned component of the externalities was particularly large in

absolute value in the stress scenario.

To summarize, I will estimate both the risk-neutral and physical measure externality mimicking portfolios, and then conduct two tests. First, for the risk-neutral portfolio, I will estimate its expected return under the physical measure, and check if this is positive. Second, for both portfolios, I will examine the portfolios' returns in the stress tests, and check if they are negative. That is, we expect

$$\frac{1}{R^i} \theta_i^{A^*,i} (\mu_p^{A^*,i} - R^i) > 0 \quad (4)$$

and, for $r \in \{i, p\}$,

$$R^i \theta_{r,RF}^{A^*,i} + (R_{s^*}^{A^*,i})^T \theta_r^{A^*,i} < 0. \quad (5)$$

5 Empirical Estimates

In this section, I construct externality-mimicking portfolios, and perform the tests described in the previous section. I will begin by describing the arbitrages, then describe my data sources, and discuss some additional assumptions I employ. I next provide summary statistics of the data, and then describe the results of my tests.

5.1 The Arbitrages

My empirical approach is motivated by the model described in the preceding sections. A key assumption of the model is that, absent regulation, the arbitrages in question would not exist. Consequently, I will focus on very short maturity arbitrages, to avoid issues like the classic limits to arbitrage argument (Shleifer and Vishny [1997]) and debt overhang (Andersen et al. [2017]). This will exclude many of the arbitrages documented in the literature, including the classic “anomalies” of Lamont and Thaler [2003], the asset-swapped TIPS arbitrage of Fleckenstein et al. [2014], and the corporate bond-CDS basis (see, e.g., Garleanu and Pedersen [2011]). I will also look primarily at arbitrages that did not exist before the financial crisis, again to focus on arbitrages due to regulation. Here, I am implicitly taking the view that minimal regulation existed before the crisis, so if there was a significant arbitrage, it must not have been caused by regulation.

The model also requires a sharp dichotomy between assets that are tradable by households and assets that are traded by intermediaries. In practice, such a sharp

dichotomy is not likely to exist. I will describe the particular assumptions I am making about limited participation in the context of each of the three arbitrages I study.

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 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 that 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), the one-month constant-maturity treasury rate was 27bps, the AA non-financial one-month commercial paper rates was 37bps, the one-month OIS rate was 40bps, the interest on excess reserve rate was 50bps, and one-month LIBOR was 52bps. In the notation of the model, 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 second group of arbitrages are covered interest parity violations. I will assume that households can purchase euros with dollars, and invest their euros at the one-month OIS rate in euros (again, assuming the one-month euro OIS rate is a proxy for a risk-free, illiquid rate available to households). What households cannot do (but intermediaries can) is trade currency forwards.¹⁰ The intermediary can replicate the household's euro/OIS trade by investing dollars at the IOER rate, and selling

¹⁰Implicitly, I am either assuming that households also cannot trade currency futures, or that there is arbitrage between futures and forwards.

dollars for euros using a one-month currency forward. The households can invest in euros/OIS instead of investing their dollars at the one-month dollar OIS rate, 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/OIS vs. the bank investing at IOER and using the forward). These two alternatives are simply linear rotations of the payoffs $Z_{a,s}$ and the arbitrages χ_a , and hence will yield identical externality-mimicking portfolios. 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.

Implicitly, with these arbitrages, I am assuming that the default risk on the forward contract is negligible (or, to be more precise, that the pricing data reflects forward rates available to a risk-free counterparty). Du et al. [2017] argue, persuasively in my view, that this risk is negligible.

The third arbitrage I study, in robustness exercises, is an arbitrage between the SPDR S&P 500 ETF and options on that ETF, which trade on the CBOE under the ticker SPY. This arbitrage is closely related to the classic index-future arbitrage involving S&P 500 futures (e.g. Chung [1991], MacKinlay and Ramaswamy [1988], Miller et al. [1994]). I assume that the household can buy the ETF. The intermediary, to replicate the ETF, can buy a call on the ETF, sell a put on the ETF at the same strike, and invest enough cash at the IOER rate over the next month to cover the exercise price of the put/call. Regardless of whether the ETF ends up above or below the strike price, the intermediary will end up owning the ETF in one month.

These two strategies will generate identical payoffs, as long as there are no dividends over the course of the month (more precisely, that an ex-dividend date does not occur within the month). The ETF has ex-dividend dates quarterly, usually on the third Friday of March, June, September, and December.¹¹ Limiting my sample to avoid these dates generates a pure arbitrage. This illustrates one of the two main advantages the ETF-based arbitrage has over the traditional arbitrage. The stocks

¹¹The prospectus, available at <https://us.spdrs.com/library-content/public/SPDR.500%20TRUST.PROSPECTUS.pdf>, describes the details of how the ex-dividend dates are determined.

of the S&P 500 index pay dividends often, and hence most studies of index arbitrage assume either perfect foresight of dividends or use a dividend forecast. The second advantage relates to transactions costs and stale prices. The traditional index arbitrage involves buying and selling 500 stocks, generating substantial transactions costs and exacerbating the issue that prices might not be synchronized. Using the ETF, which is one of the most actively traded securities in the equity market and has a very small bid-offer, mitigates many of these issues. However, as I will discuss in the data section, synchronizing the options prices and the ETF price is still critically important, as in Van Binsbergen et al. [2012].

For this arbitrage, I am assuming that the costs associated with posting margin on the options are negligible. That is, the margin is sufficiently small, and the interest rate the intermediary receives on the posted margin sufficiently close to the IOER rate, that these costs are negligible. This assumption is also, implicitly, being applied to the margin required by counterparties in the OTC market for FX swaps.

To summarize, I assume that the household has access to several investments: a risk-free asset at the one-month dollar OIS, risk-free foreign-currency assets at the one-month OIS rate in those currencies, with dollar prices determined by spot exchange rates, and the SPDR S&P 500 ETF. Intermediaries can replicate each of these payoffs, perhaps at different prices, by using the IOER rate, currency forwards, and options on the ETF. Having described the three kinds of arbitrage I will be studying, I will next describe my data sources.

5.2 The Data

My main data sample begins on Jan 4th, 2011, and runs through September 15th, 2016. I include only days at least 23 non-weekend days before the start of next scheduled FOMC meeting, excluding days with an FOMC meeting. Because the FOMC holds eight scheduled meetings each year, roughly one quarter of all non-weekend days are included in the dataset. The dates of the FOMC meetings come from the Fed’s website.¹² The data on the interest on excess reserve rate is from the St. Louis Federal Reserve’s FRED database.

My data on spot and forward exchange rates, and all OIS rates, are Bloomberg. I use the London closing time for all of these instruments, following Du et al. [2017]. I

¹²<https://www.federalreserve.gov/monetarypolicy/fomccalendars.htm>

assume that these instruments follow the standard day count conventions for the FX derivatives market.¹³ When computing the risk-neutral variance-covariance matrix, I will also employ data on one-month at-the-money straddles in each of the possible currency pairs (e.g. USD-EUR, USD-JPY, EUR-JPY, etc...).

For the SPDR arbitrage, I use data on the ETF from the Trade and Quote database (TAQ) and data on the options from OptionMetrics, both via WRDS. Currently, both the ETF and options data correspond to 3:59pm eastern time data, and hence are not aligned with the CIP violations data.¹⁴ For the ETF, I use the first standard lot size trade reported in TAQ after 3:59pm. I restrict my options sample to options with maturities between 22 and 57 days that do not cross an ex-dividend date for the ETF, and on days in which the price of the ETF changed by no more than 0.5% between 3:45pm eastern time and 3:59pm eastern time. This last restriction reduces the likelihood of stale option pricing data, which is a concern with the OptionMetrics data (Van Binsbergen et al. [2012]).

For each option maturity, I consider option strikes with a volume of at least one hundred contracts in both the put and the call. I compute an implied bid financing rate from the bid on the call and the offer on the put, along with the ETF price, and compute an implied offer financing rate in the same fashion. I then compute, separately, the maximum bid rate and minimum offer rate across strikes. I choose an option maturity by first excluding options that settle during or after the next FOMC meeting, or that cross a dividend date, and then choosing the maturity closest to the maturity on the currency trade. If there are multiple maturities that are equally close to the maturity on the CIP trade, I choose the one with the smallest bid-offer of implied financing. After choosing a maturity, I use the average of the best bid and offer financing rate for that maturity. When computing the risk-neutral variance-covariance, I use the mean of the OptionMetrics implied variances from all of the liquid strikes at the particular maturity used to construct the arbitrage.¹⁵

I will typically report annualized arbitrages. I use an actual/360 convention, based on the FX derivatives convention (for CIP) and the maturity of the relevant options (for the SPDR arbitrage). When I construct the mimicking portfolio, I scale the arbitrage on the SPDR and the risk-neutral variance on the SPDR by the ratio of the

¹³See Jewitt [2015] or https://en.wikipedia.org/wiki/Foreign_exchange_date_conventions.

¹⁴I intend to remedy this using data directly from the CBOE in future versions of this paper.

¹⁵Conceptually, a VIX-style calculation of the risk-neutral variance is the more appropriate choice. I intend to implement this in future versions of the paper.

FX derivatives maturity and the option maturity, to compute a hypothetical trade in the SPDR arbitrage with the same maturity as the CIP trade.

Finally, I collect information on the “stress test” scenarios from the Federal Reserve’s website.¹⁶ The “severely adverse” scenario described in the tests shows, among other variables, the level of euro, yen, and pound, as well as the Dow Jones Industrial Average, at a quarterly frequency. I assume that the price change in the Dow Jones Industrial Average is the return on the SPDR (ignoring both the issue of dividends and the differences between the indices). I collect both the one and four-quarter returns on each of the assets I study, and in my analysis will pretend that these returns occur over a one-month horizon.

5.3 Additional Assumptions

To conduct the tests described in the previous section, several additional assumptions are required. To construct the risk-neutral externality mimicking portfolio, I require a full variance-covariance matrix under the risk-neutral measure. To estimate that portfolio’s expected excess returns under the physical measure, an estimate of expected excess returns is required. To construct the physical externality mimicking portfolio, both a physical measure variance-covariance matrix and measures of expected excess returns are required.

For expected excess returns, I make simple assumptions based on an impressionistic reading of empirical facts about stock returns and exchange rates. Specifically, motivated by Meese and Rogoff [1983] and the related literature, I assume that currencies are, over my one-month horizon, random walks. As a result, the expected excess return of using the IOER rate and a forward to purchase, say, one yen one month from now, is determined by the difference between the forward and spot exchange rate. For the stock market, I use an annualized equity premium of 5% (based on the time series average of estimates by Martin [2017]). For both the exchange rates and the stock market, time variation in the rates of expected return have been documented, although the impact of these predictor variables on expected returns over a one-month horizon is likely to be small. Moreover, because these returns are multiplied by the risk-free arbitrage, χ_R , in equation (2), and this number is very small (roughly .01% on average), these assumptions have almost no influence on the

¹⁶<https://www.federalreserve.gov/supervisionreg/dfa-stress-tests.htm>

portfolio weights.

I construct a physical measure variance-covariance matrix based on these assumptions. I first construct surprise returns, and then construct a daily time series of squared surprise returns and products of surprise returns for different currencies and the SPDR ETF. To estimate a daily variance-covariance matrix on each day, without using future information, I use an exponentially weighted moving average, using a decay factor of 0.94 (a procedure known as the “RiskMetrics” methodology, see for example Alexander [2008]). I initialize my variance and covariance estimates at the beginning of 2011 with the realized variance/covariance for 2010. I then scale my daily estimated variance-covariance matrix to the horizon of the one-month CIP violations.¹⁷

To construct a risk-neutral variance-covariance matrix, I rely primarily on option-implied volatilities. For currencies, I can extract a full variance-covariance matrix from the implied volatilities of at-the-money currency options, under the assumption of log-normality. For the SPDR ETF, I use the average of implied variances described previously. In both cases, I make a quantitatively insignificant adjustment, motivated by my theory, and use a discount rate associated with the IOER rate, instead of the more standard OIS rate. Unfortunately, I do not have data on risk-neutral covariances between currencies and the SPDR ETF. I construct these using the correlations implied by my physical measure estimates, along with the risk-neutral variances on the currencies and the ETF.

5.4 Summary Statistics

The first summary statistics I will present contain the sample means and standard deviations of the arbitrage associated with each currency, the SPDR ETF, and 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

¹⁷More sophisticated approaches that incorporate higher-frequency data might yield better results. See, for example, Ghysels et al. [2006].

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 these arbitrages in my sample. The arbitrages have horizons of roughly one month, but are scaled to annualized values. The table also shows the option-implied volatility and correlations of each currency (with respect to the US dollar), and the option-implied volatility of the SPDR ETF. Finally, the table reports the empirically estimated correlations between the currencies and the SPDR. Here, a positive correlation means appreciation relative to the dollar when the SPDR ETF has positive returns.

Table 1: Summary Statistics for Arbitrage

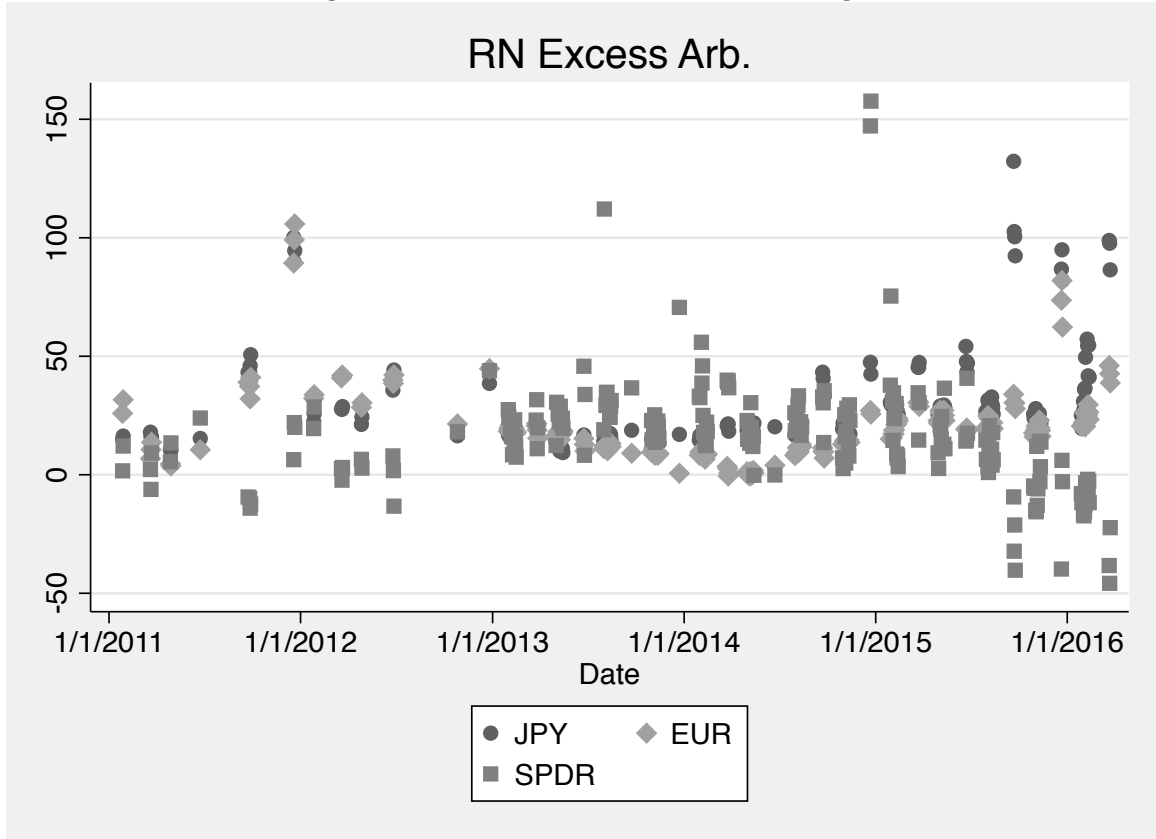
	Pounds	Euros	Yen	SPDR	FF
Arbitrage Mean (bps/year)	-5.2	8.2	16.4	2.7	-13.1
Arbitrage SD (bps/year)	8.1	18.7	22.4	22.9	2.6
OI Vol. (bps/year)	783	964	973	6168	-
OI Corr. with Pound/\$	1.00	0.59	0.24		-
OI Corr. with Euro/\$	0.59	1.00	0.29		-
Currency/\$ Corr. with SPDR (empirical)	0.21	0.20	-0.22		-
N	287	287	287	206	287

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. 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 (2) above will be roughly proportional to the amount of

“excess arbitrage,” $\tilde{\chi}_a - R^i \chi_R$, observed for each asset. However, the same is not true when considering equities; the SPDR ETF is much more volatile than any of the major currencies.

Figure 1 below shows the time series of the “risk-neutral” excess arbitrages, $\tilde{\chi}_a - R^i \chi_R$, for euro and yen, and for the SPDR ETF, in my sample. Using the risk-neutral excess arbitrage, as opposed to the physical measure excess arbitrage, eliminates the dependence on an estimate of expected returns. However, the arbitrage on the risk-free rate (χ_R) is sufficiently small that even large differences in expected returns have only small effects on this measure. 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 (when using only the yen and euro in the portfolio) closely track the excess arbitrages. In contrast, the excess arbitrage on the SPDR does not appear to be closely correlated with the excess arbitrage on the currencies, and is roughly of the same magnitude.

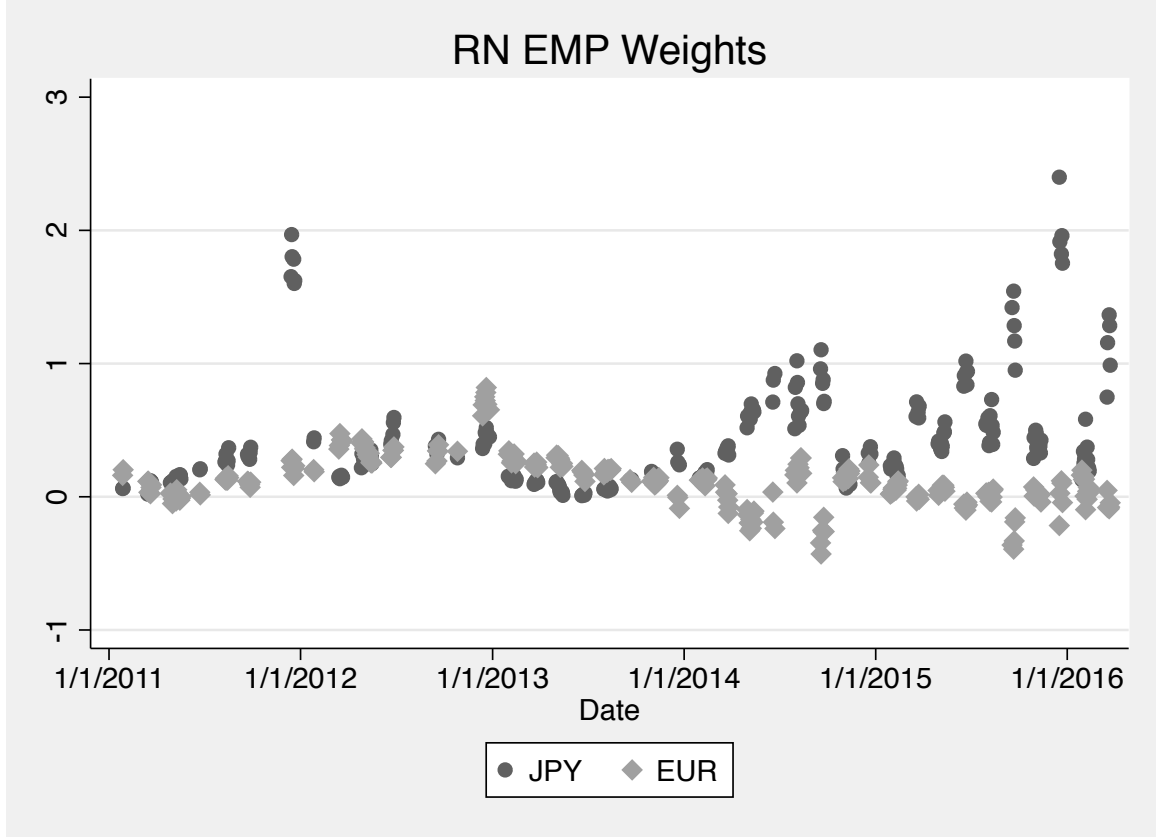
Figure 1: Time Series of Excess Arbitrage



5.5 Results

The externality-mimicking portfolio can be constructed at daily frequency. I begin by presenting portfolio weights, for a portfolio using only the euro and yen exchange rates and a risk-free asset.

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 compute

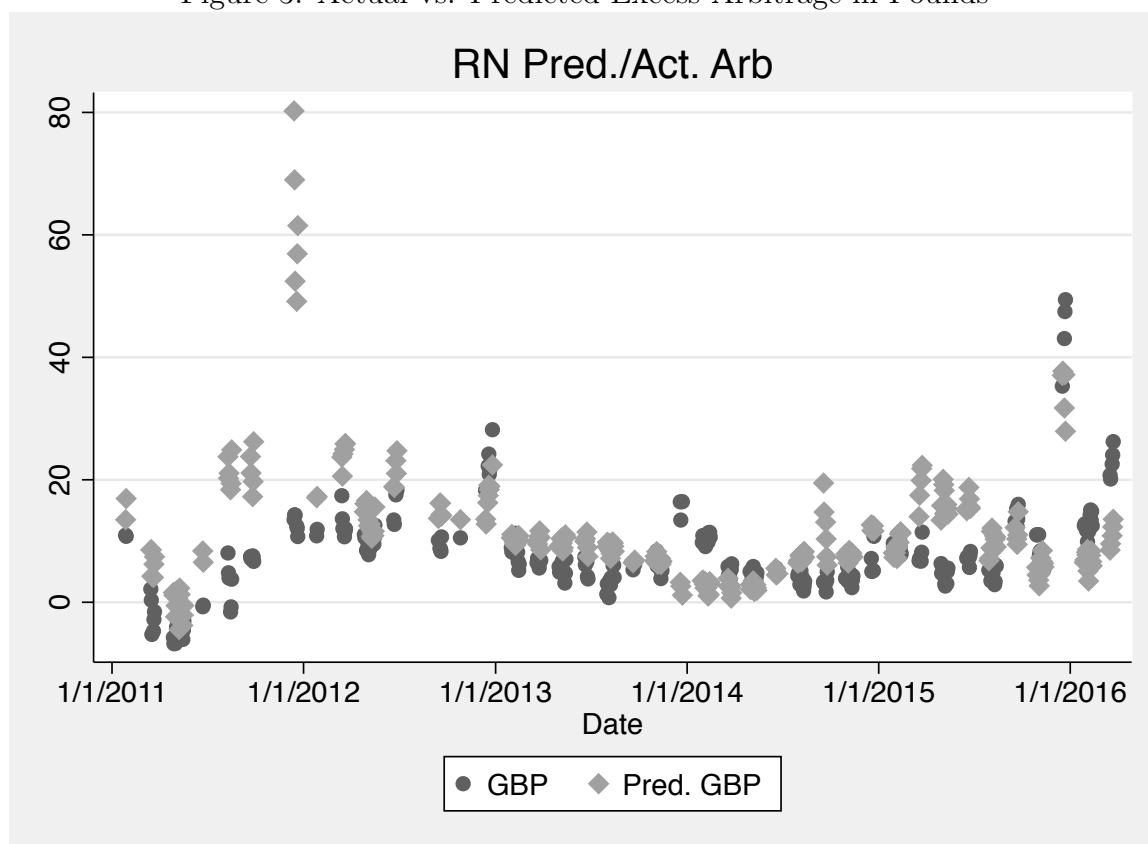
$$\tilde{\chi}_{GBP} - R^i \chi_R = \Sigma_{GBP}^i \theta^*,$$

where θ^* is the externality-mimicking portfolio (equation 2) and Σ_{GBP}^i 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 (lemma 1), that this is equivalent to computing the excess arbitrage under those externalities, if there is no covariance between

the pound excess arbitrage and the error term in the projection.

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_{GBP}^i . 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 track each other, except near the end of 2011.

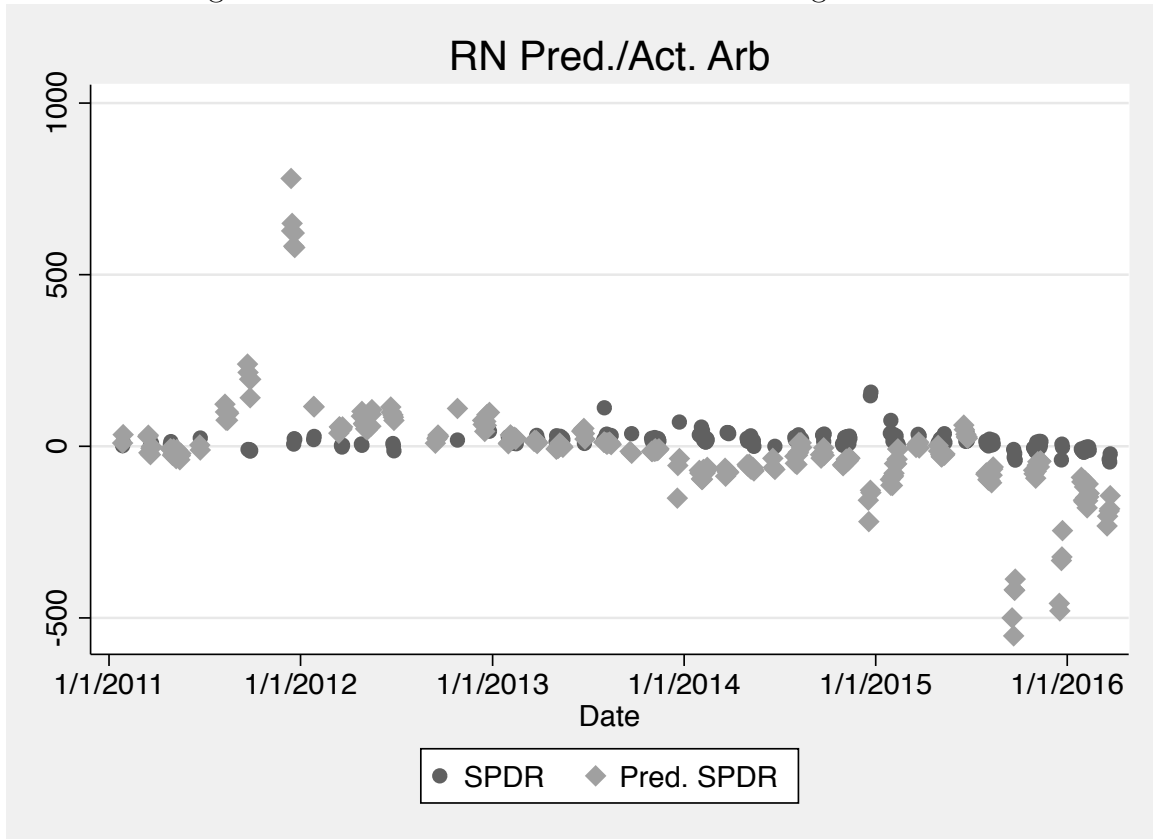
Figure 3: Actual vs. Predicted Excess Arbitrage in Pounds



However, the same cannot be said for the arbitrage based on the SPDR ETF. This arbitrage does not appear to be closely linked to the other arbitrages. I show these results in the figure below, and (because of this issue) will present results with

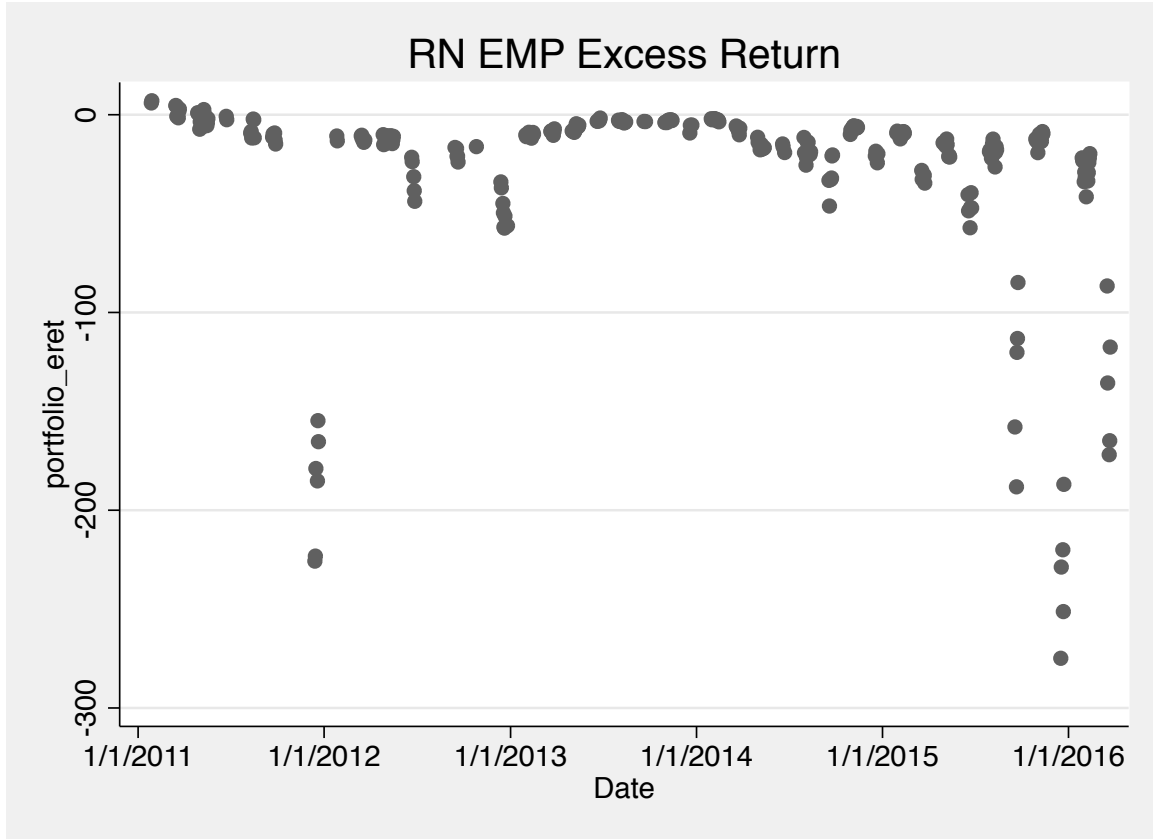
the SPDR arbitrage added to the portfolio later in this section.

Figure 4: Actual vs. Predicted Excess Arbitrage for SPDR



I next consider the expected return of this portfolio (the first test described in the previous section). Intuitively, because the portfolio is generally long yen and roughly flat euros, and both of these currencies have low interest rates and hence negative expected returns, the expected return on the portfolio is negative. This contradicts the intuition that the externalities should be negatively correlated with the stochastic discount factor.

Figure 5:



In the table below, I formally test whether the average expected return over my sample is greater than or equal to zero (a one-sided test). I show results for the full sample, only for dates for which the trade crosses a quarter-end, and only for dates for which the trade crosses a year-end.¹⁸ I also formally test whether the quarter-end dates are different from other dates, and whether the year-end dates are different from other quarter-end dates. Both Bech and Klee [2011], for fed funds vs. IOER, and Du et al. [2017], for CIP, have documented that the arbitrage spikes near quarter-ends, and many of the outlier days for the SPDR arbitrage are near quarter and year end as well. As the results below demonstrate, the problem of negative expected returns documented above is particularly acute at quarter and year-ends.

I now turn to the second test, using the stress tests. Once per year, the Federal Reserve describes a “severely adverse” scenario and requires banks to maintain various

¹⁸Formally, the trade “crosses quarter end” if the settlement dates of the spot and forward FX trades are before and after the end of some quarter, respectively. Crossing year end is defined the same way, but only for the fourth quarter.

Table 2: Risk-Neutral EMP Expected Returns

	N	Mean (bps)	Standard Deviation (bps)	Test	P-Value
Full Sample	287	-25.2	45.4	≥ 0	0.0000
Quarter-Ends	104	-51.0	67.6	≥ 0	0.0000
Year-Ends	26	-107.1	92.9	≥ 0	0.0000
QE - Full		-40.4		$= 0$	0.0000
YE - QE		74.9		$= 0$	0.0000

leverage and capital ratios in this scenario. In table 3 below, I report the returns of the yen, euro, and stocks¹⁹ 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. The stock return itself very negative in all of these scenarios.

Table 3: Stress Test “Severely Adverse” Scenarios

Stress Test Date	Euro		Stocks		Yen	
	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	-21.4	-72.4	2.9	-1.0
9/30/13	-15.4	-24.1	-30.8	-68.4	3.1	-1.1
9/30/14	-12.7	-14.4	-17.7	-84.7	7.9	6.7
12/31/15	-8.1	-15.0	-22.6	-70.8	2.8	5.2
12/31/16	-9.5	-12.7	-41.5	-68.8	3.3	7.8

Each of the stress test scenarios is associated with a particular date (listed in table 3) which is the date at which the scenario starts. For each date in my sample that is also within 180 calendar days of the stress test date, I report the returns of the risk-neutral externality-mimicking portfolio 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

¹⁹As mentioned above, the stress test scenarios report the level of the DJIA. I assume the change is due entirely to returns, and that the SPDR ETF will experience the same returns.

a single stress test per date, dropping only a handful of days that are far from any stress test date.

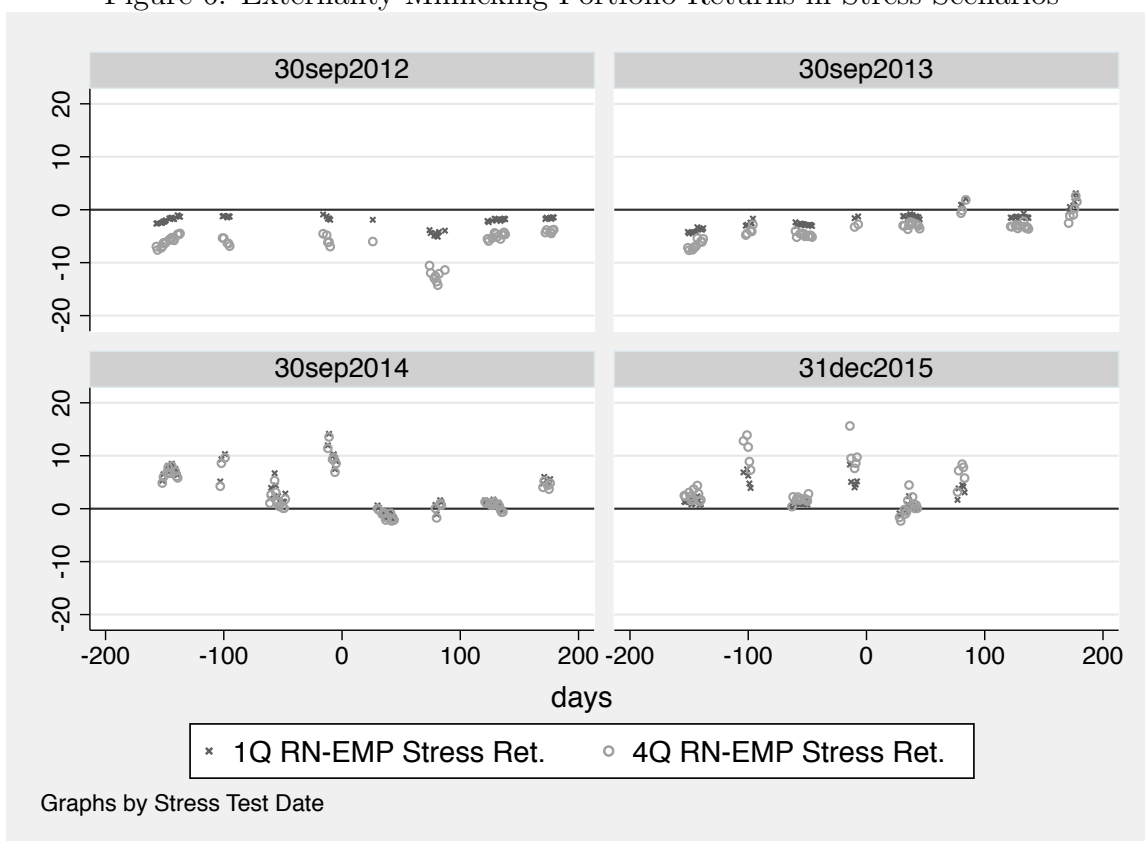
In figure 6 below, I report the returns of the externality mimicking portfolio for each of these days, under the associated stress scenario. The return of the externality-mimicking portfolio is “unit-less,” and has a particular meaning. A 10% positive return means that in the specified scenario, to rationalize 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, 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 6: Externality Mimicking Portfolio Returns in Stress Scenarios



I also show formal t-test results, averaging across stress test dates within a particular stress test, in the table below. The p-values correspond to a one-sided test that the mean is less than or equal to zero.

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

Table 4: Risk-Neutral Returns in Stress Scenario

Stress Test Date	N	Mean (1Q,%)	S.D. (1Q,%)	P-value (1Q)	Mean (4Q,%)	S.D. (4Q,%)	P-value (4Q)
9/30/12	51	-2.2	1.1	1.0000	-6.5	2.8	1.0000
9/30/13	60	-1.8	1.7	1.0000	-3.6	2.1	1.0000
9/30/14	61	3.5	3.9	0.0000	2.9	3.9	0.0000
12/31/15	49	2.0	2.2	0.0000	3.9	4.2	0.0000

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 in-

struments 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.

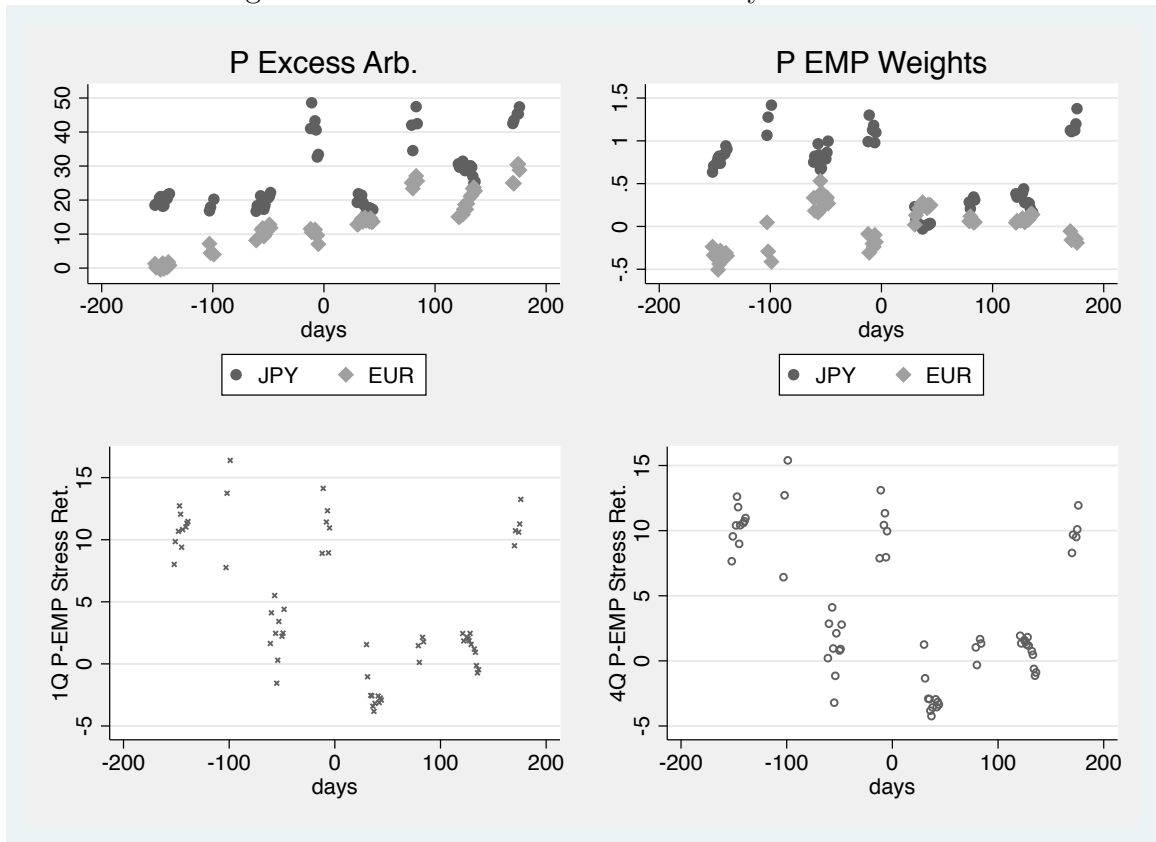
For robustness, I also present results for the physical measure externality-mimicking portfolio, and for a risk-neutral measure portfolio that includes the SPDR ETF. In the latter case, the SPDR ETF plays almost no role in the externality-mimicking portfolio, because the ETF is quite volatile relative to currencies, and the arbitrage is not much larger than for the CIP deviations, and hence the “Sharpe ratio due to arbitrage” is small.

First, I present the data for the 2014 stress test using the physical measure. The results are qualitatively similar to the results for the physical measure.

Table 5: Physical Returns in Stress Scenario

Stress Test Date	N	Mean (1Q,%)	S.D. (1Q,%)	P-value (1Q)	Mean (4Q,%)	S.D. (4Q,%)	P-value (4Q)
9/30/12	51	-3.3	1.4	1.0000	-10.0	4.0	1.0000
9/30/13	60	-1.8	2.4	1.0000	-3.8	2.9	1.0000
9/30/14	61	4.7	5.7	0.0000	3.9	5.6	0.0000
12/31/15	49	2.8	3.1	0.0000	5.3	5.8	0.0000

Figure 7: 2014 Stress Test Data for Physical Measure

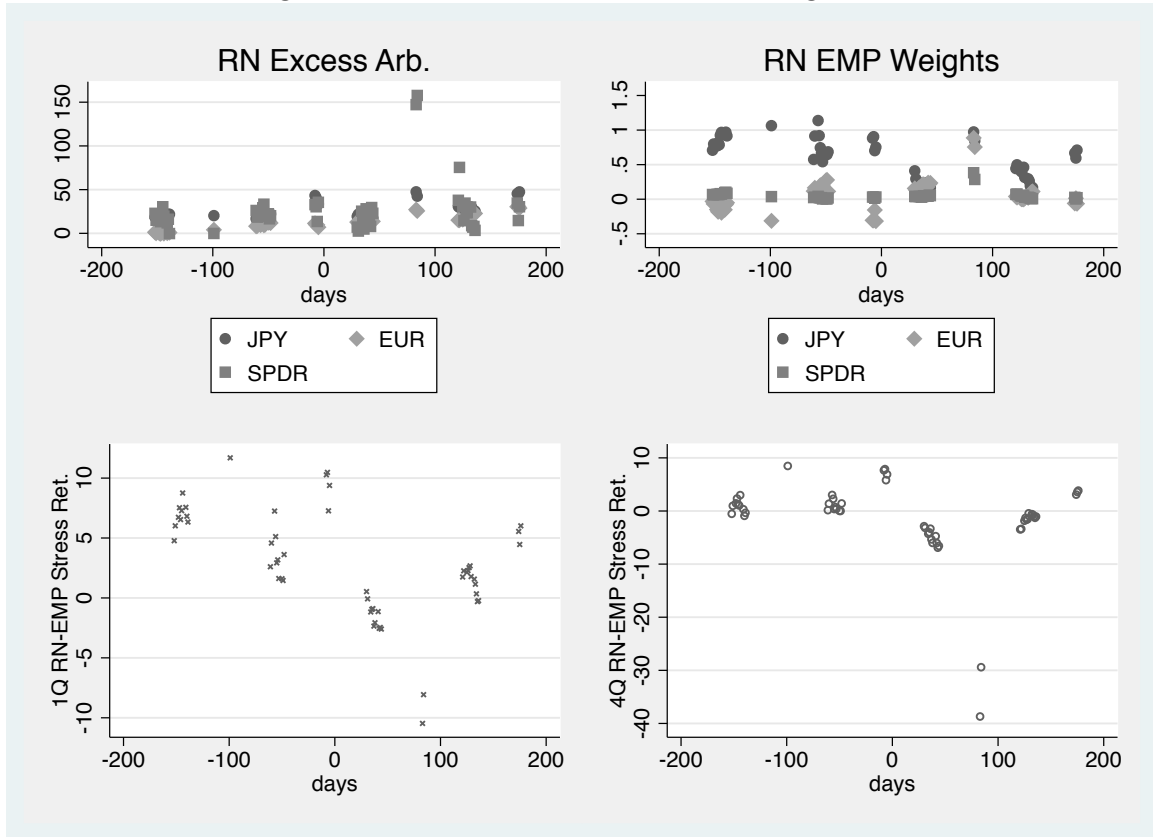


The same results, in table form, are also similar to the results for the risk-neutral measure. The physical measure estimates of the covariance matrix that I employ are more volatile than their risk-neutral counterparts, and hence introduce volatility into the portfolio weights and stress test returns.

Second, I present the data, again for the 2014 stress test, using the risk-neutral measure and including the SPDR arbitrage. As mentioned previously, for the most

part, due to its low Sharpe ratio due to arbitrage, the SPDR ETF has a small weight in the mimicking portfolio. The exception appears around day +100 (close to year end 2014), when the SPDR arbitrage is large. This part of the data looks good, in the sense that the portfolio experiences positive expected returns and negative stress test returns at this time. This illustrates again the idea that it is the absence of arbitrage that is problematic; if the SPDR arbitrage were large on most days, the regulations creating such an arbitrage could be easily rationalized from a macro-prudential perspective.

Figure 8: 2014 Stress Test Data including SPDR



Below, I display tables documenting the tests for the risk-neutral portfolio that includes the SPDR ETF arbitrage. Broadly speaking, the results are similar to the ones without the SPDR ETF, with the exception of the 2014 and 2015 four-quarter stress test returns. For these stress tests, the four-quarter returns for stocks are so negative that returns are likely to be negative with even a small positive weight on the stress test. As the graph above shows, the tables below hide some heterogeneity

Table 6: Risk-Neutral EMP Expected Returns with SPDR

	N	Mean (bps)	Standard Deviation (bps)	Test	P-Value
Full Sample	206	-10.1	60.0	≥ 0	0.0081
Quarter-Ends	55	-52.3	99.5	≥ 0	0.0001
Year-Ends	10	-152.2	183.0	≥ 0	0.0137
QE - Full		-57.5		$= 0$	0.0000
YE - QE		-122.2		$= 0$	0.0002

Table 7: Risk-Neutral Returns in Stress Scenario with SPDR

Stress Test Date	N	Mean (1Q,%)	S.D. (1Q,%)	P-value (1Q)	Mean (4Q,%)	S.D. (4Q,%)	P-value (4Q)
9/30/12	22	-2.1	0.7	1.0000	-5.6	1.7	1.0000
9/30/13	53	-2.0	2.0	1.0000	-6.5	3.7	1.0000
9/30/14	53	2.8	4.4	0.0000	-1.4	1.0	0.9104
12/31/15	44	0.5	1.5	0.0132	-0.8	2.3	0.9879

during the 2014 period. There are many days with positive four-quarter stress test returns, and a few days with very negative stress returns, corresponding to dates in which the SPDR ETF arbitrage is large.

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 Model Example and Extensions

6.1 Micro-Prudential Motives for Regulation

6.2 Borrowing Constraints and Price Rigidities

6.3 An Example

7 Conclusion

In this paper, I have analyzed the connection between externalities and arbitrage. In a general equilibrium model with incomplete markets and two classes of agents, I have shown that under an optimal policy, there is a close connection between the externalities the planner is attempting to correct with regulation and the arbitrage that regulation creates. Using this connection, I have developed a method of backing out a set of externalities that would rationalize a particular pattern of arbitrage across assets. This method constructs an externality-mimicking portfolio, whose returns are a projection of the externalities onto the space of returns. This portfolio is also the portfolio that maximizes what I call the “Sharpe ratio due to arbitrage.” I argue that these externalities should negatively covary with the SDF, and be particularly negative in “stress” scenarios. Using these intuitions, I develop two simple tests: does the externality-mimicking portfolio have positive expected returns, and does it have negative returns in the Federal Reserve’s stress tests? I show, in current data, that the answer to both these questions is no, implying that there is something inconsistent about current regulatory policy.

References

- Carol Alexander. Moving average models for volatility and correlation, and covariance matrices. *Handbook of finance*, 2008.
- 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.
- Falk Bräuning and Kovid Puria. Uncovering covered interest parity: the role of bank regulation and monetary policy. 2017.
- Y Peter Chung. A transactions data test of stock index futures market efficiency and index arbitrage profitability. *The Journal of Finance*, 46(5):1791–1809, 1991.
- 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.

- Eric Ghysels, Pedro Santa-Clara, and Rossen Valkanov. Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131(1-2):59–95, 2006.
- 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.
- Giles Jewitt. *FX derivatives trader school*. John Wiley & Sons, 2015.
- Owen A Lamont and Richard H Thaler. Anomalies: The law of one price in financial markets. *The Journal of Economic Perspectives*, 17(4):191–202, 2003.
- 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.
- A Craig MacKinlay and Krishna Ramaswamy. Index-futures arbitrage and the behavior of stock index futures prices. *The Review of Financial Studies*, 1(2):137–158, 1988.
- Semyon Malamud and Andreas Shrimpff. An intermediation-based model of exchange rates. Technical report, Working Paper, 2017.
- Ian Martin. What is the expected return on the market? *The Quarterly Journal of Economics*, 132(1):367–433, 2017.

Richard A Meese and Kenneth Rogoff. Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of international economics*, 14(1-2):3–24, 1983.

Merton H Miller, Jayaram Muthuswamy, and Robert E Whaley. Mean reversion of standard & poor’s 500 index basis changes: Arbitrage-induced or statistical illusion? *The Journal of Finance*, 49(2):479–513, 1994.

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.

Jules Van Binsbergen, Michael Brandt, and Ralph Koijen. On the timing and pricing of dividends. *American Economic Review*, 102(4):1596–1618, 2012.

A Additional Figures

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

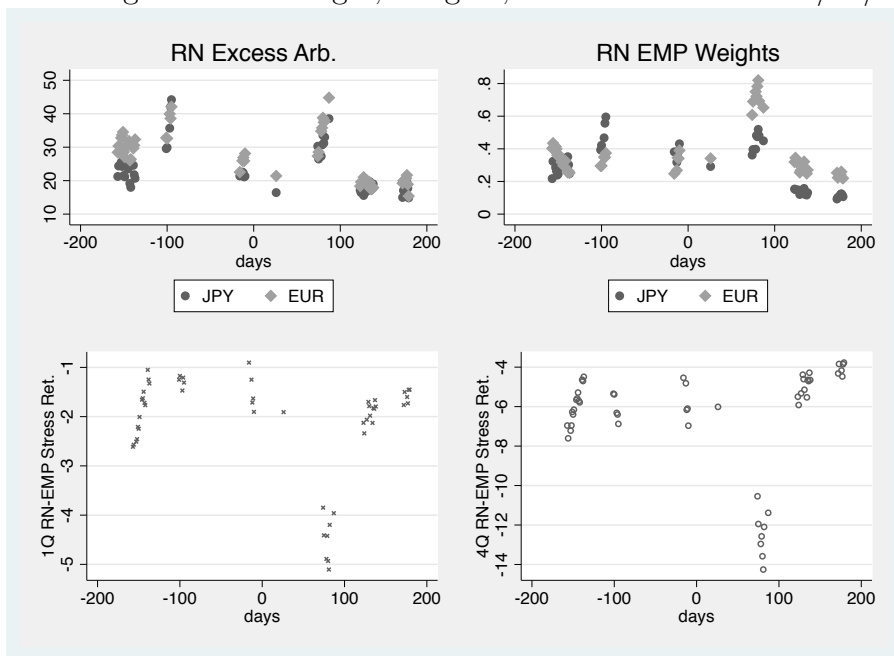


Figure 10: Arbitrages, Weights, and Returns for the 9/30/13 Stress Test

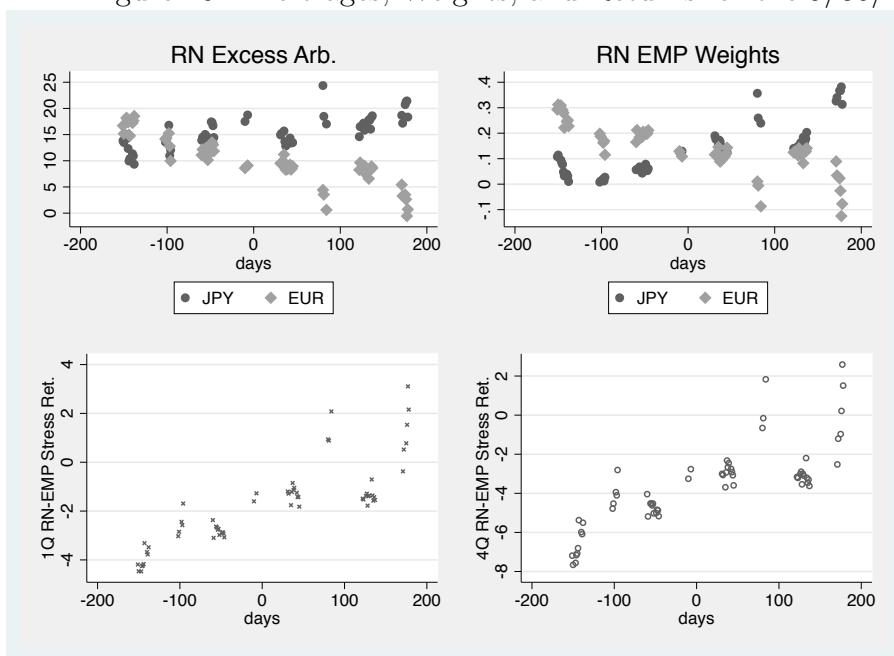


Figure 11: Arbitrages, Weights, and Returns for the 9/30/14 Stress Test

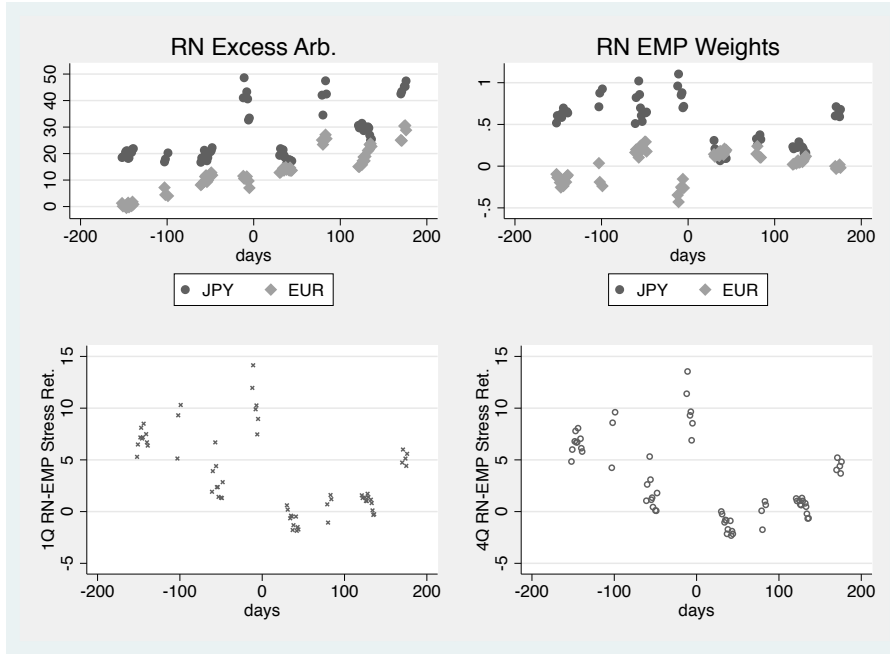


Figure 12: Arbitrages, Weights, and Returns for the 12/31/15 Stress Test

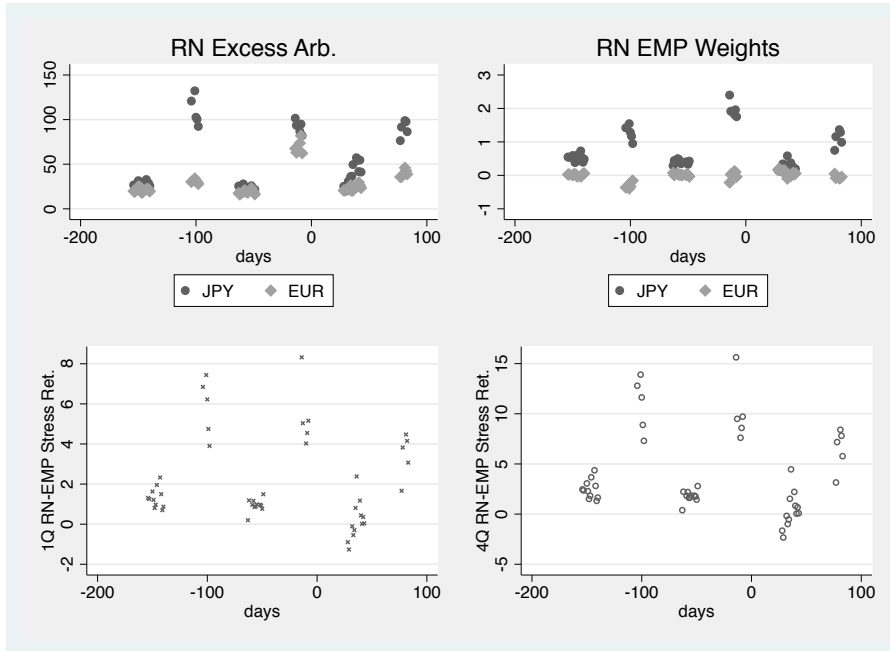
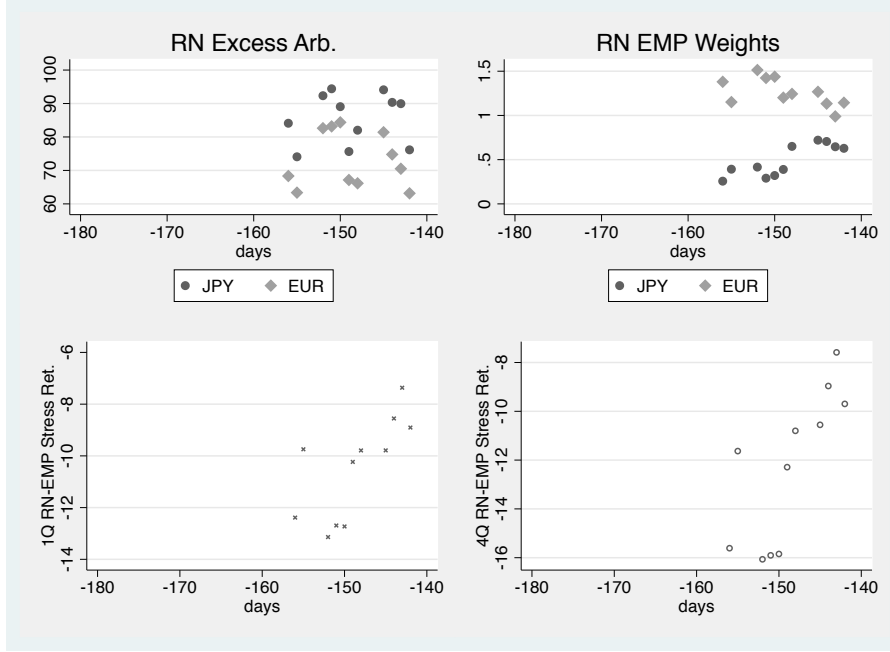


Figure 13: Arbitrages, Weights, and Returns for the 12/31/16 Stress Test



B Proofs

B.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

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

and

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

We can rewrite the private FOCs as

$$\begin{aligned} \kappa Q_a - \psi_a &= (\xi^h - \rho^h) \cdot \Phi_a^h + \sum_{s \in S} \mu_s X_{I,s}^h Z_{a,s} \\ &= \xi^i \cdot \Phi_a^i + \sum_{s \in S} \mu_s X_{I,s}^i Z_{a,s} \end{aligned}$$

We use the identities

$$S_s^i = X_{P,j,s}^i + X_{I,s}^i (X_s^i)^T$$

$$V_{P,s}^i = -V_{I,s}^i X_s^i$$

Hence, we can write

$$\begin{aligned} \sum_{i \in I} [-\lambda^i V_{I,s}^i (X_s^i + Y_s^i) - \mu_s \cdot S_s^i + (\mu_s \cdot X_{I,s}^i) X_s^i] + \\ \sum_{h \in H} [-\lambda^h V_{I,s}^h (X_s^h + Y_s^h) - \mu_s \cdot S_s^h + (\mu_s \cdot X_{I,s}^h) X_s^h] = 0. \end{aligned}$$

Using the Arrow securities,

$$\lambda^i V_{I,s}^i = \mu_s \cdot X_{I,s}^i + \psi_s$$

$$\lambda^h V_{I,s}^h = \mu_s \cdot X_{I,s}^h + \psi_s - \rho^h \cdot \Phi_s^h.$$

Therefore,

$$\sum_{i \in I} [-\lambda^i V_{I,s}^i Y_s^i - \mu_s \cdot S_s^i - \psi_s X_s^i] + \sum_{h \in H} [-\lambda^h V_{I,s}^h Y_s^h - \mu_s \cdot S_s^h - \psi_s X_s^h + (\rho^h \cdot \Phi_s^h) X_s^h] = 0.$$

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

$$\sum_{i \in I} [-\lambda^i V_{I,s}^i Y_s^i + \kappa \tau_s D(p_s) \cdot S_s^i - \psi_s X_s^i] + \sum_{h \in H} [-\lambda^h V_{I,s}^h Y_s^h + \kappa \tau_s D(p_s) \cdot S_s^h - \psi_s X_s^h + (\rho^h \cdot \Phi_s^h) X_s^h] = 0.$$

and

$$\begin{aligned} \lambda^i V_{I,s}^i &= \bar{\mu}_s - \kappa \tau_s D(P_s) \cdot X_{I,s}^i + \psi_s \\ \lambda^h V_{I,s}^h &= \bar{\mu}_s - \kappa \tau_s D(P_s) \cdot X_{I,s}^h + \psi_s - \rho^h \cdot \Phi_s^h. \end{aligned}$$

Using the Arrow securities and the private FOCs,

$$\begin{aligned} \kappa Q_s - \psi_s &= (\xi^h - \rho^h) \cdot \Phi_s^h + \bar{u}_s - \kappa \tau_s D(P_s) X_{I,s}^h \\ &= \xi^i \cdot \Phi_s^i + \bar{u}_s - \kappa \tau_s D(P_s) X_{I,s}^i \end{aligned}$$

and therefore

$$\kappa Q_s = \lambda^h V_{I,s}^h + \xi^h \Phi_s^h = \lambda^i V_{I,s}^i + \xi^i \Phi_s^i$$

Define

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

If an asset is tradable by household h ,

$$\kappa Q_a - \psi_a = - \sum_{s \in S} \kappa \tau_s D(P_s) X_{I,s}^h Z_{a,s} = \xi^i \cdot \Phi_a^i - \sum_{s \in S} \kappa \tau_s D(P_s) X_{I,s}^i Z_{a,s}$$

Therefore, with no regulation of the Arrow market,

$$\begin{aligned} \chi_a &= \sum_{s \in S} \tau_s D(P_s) (X_{I,s}^h - X_{I,s}^i) Z_{a,s} \\ &= \sum_{s \in S} \left(\frac{\lambda^i}{\kappa} V_{I,s}^i - \frac{\lambda^h}{\kappa} V_{I,s}^h \right) Z_{a,s} \end{aligned}$$

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

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

B.2 Proof of lemma ??