Impediments to Financial Trade: Theory and Measurement*

Nicolae Gârleanu
UC Berkeley-Haas, NBER, and CEPR

Stavros Panageas
University of Chicago, Booth School of Business and NBER

Jianfeng Yu
University of Minnesota, Carlson School of Business

January 2015

Abstract

We propose a tractable model of an informationally inefficient market. We show the equivalence between our model and a substantially simpler model whereby investors face distortive investment taxes depending both on their identity and the asset class. We use this equivalence to assess existing approaches to inferring whether individual investors have informational advantages. We also develop a methodology of inferring the magnitude of the frictions (implicit taxes) that impede financial trade. We illustrate the methodology by using data on cross-country portfolio holdings and returns to quantify these frictions, and locate the directions in which financial trade seems to be especially impeded. We argue that our measure of frictions contains useful information for the sources of failure of frictionless models, and it helps in studying whether certain factors (such as the size of the financial sector) are associated with lower financial frictions.

Keywords: Financial frictions, Asset Pricing, Inefficient Markets, Performance Evaluation, Cross-border equity flows

JEL Classification: G12, G14, G15

*We are grateful for comments and suggestions from Eugene Fama, Motohiro Yogo, John Heaton, Helene Rey, Eric Van Wincoop, and seminar participants at Berkeley-Haas, Chicago Booth, University of British Columbia - Sauder, University of Maryland - Smith, University of Virginia-McIntire. Panageas acknowledges research support from the Fama-Miller Center for research in Finance.

1. Introduction

There is abundant evidence suggesting that (many) financial portfolios are under-diversified. One way to phrase this observation is to state that investors differ in their perception of value of many assets, and therefore tilt their portfolios away from a fully-diversified market portfolio.

We propose a tractable model of one channel that can generate such heterogeneity, namely asymmetric information. We show that such a model naturally creates a wedge in the valuation of the same asset by different investors, akin to raising investor- and asset-class- specific taxes. More broadly, treating financial frictions as equivalent to shadow taxes provides a simple, unified, and flexible way to model their distortionary effects.

We use this tax equivalence of frictions for two applications, the first theoretical and second empirical. First, we assess the value of the widely used "alpha" measures for inferring informational frictions, and in particular the informational advantage of certain investors. We find that there is no simple, one-to-one relation between informational advantage and alpha and discuss alternatives.

Second, using the broad equivalence between financial frictions and shadow tax rates, we develop a methodology to infer these shadow tax rates jointly from observed portfolio allocations and returns. We illustrate this methodology in the context of an example: we compute the matrix of (implicit) shadow taxes necessary to explain cross-country portfolio allocations. These shadow tax rates are expressed in economically meaningful units (percentage points of gross return) and allow us to create a "map" of financial frictions, i.e., they allow us to locate directions of financial trade that seem to be especially impeded. We analyze various factors that tend to correlate with the magnitude of these frictions. We argue that they contain important information for many applications, such as identifying sources of departure from frictionless models, determining the role of the financial industry in reducing financial frictions, etc.

Specifically, we consider a model featuring different locations, with a fraction κ of investors in every location being regular investors and the complement being "swindlers." Regular investors are endowed with common stocks that pay random location-dependent dividends at date one, while each swindler owns a "fraudulent" stock that pays nothing. Investors obtain signals on the type of a given stock (regular or fraudulent) in every location. Important, the quality of that signal depends on both the investor's and the firm's location.

Swindlers have a strong incentive to trade so as to equalize the price of their stock with the

prices of other stocks in their location. Moreover, the swindler can manipulate the earnings of her company, which deters short selling. A pooling equilibrium emerges with all common and fraudulent stocks in a given location trading at the same price. The failure rate f of an investor's signal to identify fraudulent stocks in a given location can be equivalently viewed as a tax rate when investing in that location: A proportion f of the stocks identified by the investor's signals as regular pay nothing. Indeed, we prove an equivalence between our model and a much simpler (Walrasian) economy where investors are faced with investor- and asset-specific capital taxation.

We utilize this simple framework for two purposes. First, we investigate the theoretical validity of common approaches to measuring informational (dis)advantages. We show that when markets are informationally inefficient, Jensen's alpha may fail to identify informational advantage; passive strategies may have alpha, and informed strategies may have negative alpha. We link these phenomena to the heterogeneity of informational inefficiency across markets. The limitation of Jensen's alpha as a performance measure is not merely a side effect of the CAPM being misspecified. In an informationally inefficient market it may be impossible to risk-adjust returns so that the intercept (alpha) of the postulated-asset pricing model maps into an investor's informational advantage. We show that the "style" analysis approach (originally proposed by W. Sharpe, and commonly adopted in practice), whereby the returns of an investment strategy are regressed on the returns of index strategies that *depend* on the type of asset classes in which the investment manager invests, would be a better alternative within our framework. However, style analysis requires a judicious choice of relevant asset classes, which can be a nontrivial task.

Second we develop a methodology to extract the model-implied shadow tax rates from information that is contained in *both* portfolio holdings and asset returns. We illustrate this methodology by inferring the shadow tax rates that are consistent with the patterns of international portfolio allocations and equilibrium returns in a cross section of OECD countries. We note that our theoretical model provides only one (informational frictions) of a number of possible ways to motivate the empirical exercise. This motivation is useful to fix ideas, but not crucial. More broadly, the implied shadow tax rates should be interpreted as a comprehensive measure of valuation discrepancies, which encompass all impediments to financial trade, both explicit and implicit.[1] The patterns of these shadow tax rates paint a detailed picture of the direction and magnitude of frictions.

[1]Examples of actual impediments would be actual capital taxes, while examples of implicit impediments would be portfolio constraints.

We document several patterns in the shadow tax rates. A first and noticeable pattern is that the majority of their variation is explained by destination-country fixed effects: Some countries seem to present foreign investors with higher implicit tax rates than others, no matter the origin of the investment. Origin-country fixed effects explain a smaller, but still non-trivial, amount of variation. Measures that span the residual, bilateral variation (i.e., the variation not spanned by either origin- or destination-country fixed effects) explain a very small part of the variation in implied tax rates. Indeed, any non-directional, bilateral variable (i.e., any variable that doesn't change value when we switch the identity of origin and destination country, such as geographical distance, common legal origin, common language or religion, etc.) is bound to *not* explain a substantial fraction of the variation in shadow tax rates. We also find that import shares into country j from country i have a positive but small correlation with our measured financial friction facing investors from i when investing in j, suggesting that iceberg-cost-related frictions in the goods markets[2] are not the primary driver of our measured frictions.

Consistent with common wisdom, we find that high shadow tax rates plague mostly investments directed towards less developed (economically and financially) economies. By contrast, shadow tax rates are quite small for financial trade between more developed economies. These statements follow both from an inspection of the data via a k-means cluster analysis (using estimated tax rates as a measure of distance), as well as a regression analysis. Smaller frictions correlate with a higher GDP per capita and a larger financial industry as a fraction of GDP. This suggests that the financial sector performs a useful function either because it helps lower frictions, or because it is a useful input for countries that have inherently lower frictions and need to process the flows from their increased appeal as investment destinations.

By performing several counterfactual exercises, we illustrate that our results are not merely driven by cross-sectional differences in home bias. The implication of this finding is that simply focusing on patterns of home bias may not be enough if one is to obtain an explanation of the frictions that impede financial trade.

In summary, this paper develops a new methodology to measure impediments to financial trade. Taking the view that financial frictions act as shadow taxes (a view consistent with, but not confined to, an asymmetric-information motivation), we utilize data on cross-country portfolios to obtain

[2]See, e.g., Obstfeld and Rogoff (2001) or Dumas and Uppal (2001).

and study patterns in implicit tax rates. These tax rates can be useful for a host of applications, such as providing guidance for the sources of frictions that seem most promising in explaining the data, quantifying the importance of some factors (such as the size of the financial industry) in reducing shadow tax rates, providing benchmarks for models that generate valuation wedges across different investors, etc. Additionally, and importantly, our measure of frictions is expressed in economically meaningful units and is easily interpretable.

1.1. Literature Review

The paper is related to various strands of the literature. At the theoretical level, the most closely related literature is that on multi-asset REE that focuses on explaining portfolio biases as due to asymmetric information.[3] A limitation of the asymmetric information literature is that investors tend to invest more heavily locally, but only if their superior local signal is positive. By contrast, in the data investors tend to have permanently higher allocations to their local asset. Our model is a hybrid of a rational expectations equilibrium and a standard adverse-selection model, which generates the constantly higher allocation to local stocks.[4] Perhaps the main innovation relative to the literature, however, is that we obtain a particularly tractable multi-asset model, which can be analyzed as simply as an elementary model with differential tax rates. We exploit this simplicity to obtain implications for the measurement of informational advantages and frictions that we believe to be new in the literature. In particular, our critique of "alpha" as a measure of informational advantage is quite distinct from existing criticisms of that measure in the literature.[5]

The idea of adding to the canonical model frictions that drive wedges in the valuation of different investors dates back to the origins of modern international finance — at least as far as Black (1974) and Stulz (1981). Cooper and Kaplanis (1986) noticed an identification problem in inferring implied

[3]Representative examples include Admati (1985), Gehrig (1993), and Brennan and Cao (1997). Relatedly, Van Nieuwerburgh and Veldkamp (2010) propose an approach relying on rational inattention.

[4]Hatchondo (2008) is closer to the setup of the current model. An important difference to his model is that we do not rely on noise trading, assuming instead the existence of strategic "swindlers." Furthermore, we can obtain the no-shorting outcome endogenously, although in the main body of the paper we impose short selling restrictions directly for simplicity.

[5]To give an example, our criticism of "alpha" as a measure of performance relies on the different pricing of stocks in the presence of heterogeneous extents of asymmetric information in general equilibrium, and the inexistence of an asset pricing model capable to price even the "passive" strategies. Hence our argument is quite distinct from earlier criticisms, which emphasized the difficulties in using an uninformed investor's information set to evaluate the ex-post performance of an ex-ante efficient portfolio. (See, e.g., Admati and Ross (1985), Dybvig and Ross (1985), and Mayers and Rice (1979) amongst many others). For instance, in the latter two papers SML analysis is valid if all information is security-specific, unlike in our framework.

frictions from observed portfolio holdings. Specifically, the requirement that portfolios add up to one, along with the market clearing conditions, make the mapping from portfolios to frictions non-unique. This identification problem does not arise in our setup for two reasons. First, we assume the presence of (country-specific) risk-free assets that can be traded without frictions. Second, our tax-equivalence result implies taxes that are distortionary, but redistributive (from regular investors to swindlers), unlike the iceberg costs in the literature.[6]

Finally, the paper is related to the various strands of empirical literature that analyze cross-country equity allocations.[7] Our main departure from this literature is methodological. For instance, it is common to regress portfolio allocations on various bilateral measures. However, portfolios alone are not sufficient to provide a picture of the underlying frictions (i.e., valuation wedges), which are the ultimate object of interest. We combine the information contained in the moments of asset returns, portfolios, and market capitalizations, along with optimality and general equilibrium conditions, to obtain a comprehensive measure of deviations from a frictionless benchmark.[8]

The paper is remotely related to the voluminous literatures on the home bias and the benefits of international diversification. We do not attempt to summarize either of these strands of the literature here.[9] The modern literature on home bias typically starts with primitive assumptions on endowments, information, etc., and then derives implications for such quantities as equilibrium portfolios, return correlations, exchange rate correlations, and expected returns. Multi-country setups are not easily computable except in very special cases;[10] more importantly, the goal of such models is to assess the quantitative power of one particular mechanism, without aspiring to match all conceivable moments in the data. The spirit of the exercise in this paper is different: rather than write elaborate models, we propose to revisit an (intentionally) basic, well-understood model, which

[6] In a related vein, Sharpe (1974) proposes to use holdings information to impute perceived expected returns. The idea in this paper is similar in some ways, in that the implied taxes essentially present different investors with different investment-opportunity sets. However, an aspect of our analysis is the joint usage of first order optimality conditions and general equilibrium conditions to infer frictions.

[7] This literature is quite large and we do not attempt to provide a full summary here. Indicatively we mention Faruqee et al. (2004), Portes and Rey (2005), Lane and Milesi-Ferretti (2005), Aviat and Coeurdacier (2007), Berkel (2007), Daude and Fratzscher (2008), and Bekaert and Wang (2009) among many others. A common theme in this literature is to estimate "gravity" equations for equity flows. Okawa and van Wincoop (2012) contains a thorough treatment of the theory of gravity equations in international finance and its limitations.

[8] Koijen and Yogo (2014) use institutional portfolio holdings to estimate an equilibrium demand system and the price impact of institutional investors. However, they parametrize directly the demands (through a logit specification) whereas we impute wedges from first order optimality conditions.

[9] See, e.g., Coeurdacier and Rey (2013) for a recent survey.

[10] See, e.g. Pavlova and Rigobon (2008) for a particularly tractable framework

we enrich with a very flexible structure of frictions. Subsequently we infer the frictions that can re-produce the observed allocations in the data. Such an approach has the advantage of providing, in economically rather than statistically meaningful units, measures of the strength of the sources of success or failure of the basic economic model. Our paper belongs to a growing body of research that performs similar exercises in other contexts.[11]

The paper is also remotely related to the literature that assesses the benefits of international diversification (typically in a partial rather than general equilibrium framework). In contrast to that literature, we do not provide estimates of the benefits of diversification for — say — a US investor. Instead we try to develop an entire geography of the severity of bilateral impediments to financial trade. We therefore provide a complement to this literature, by localizing and quantifying the financial frictions.

2. Model

2.1. Locations, preferences, and firm and investor types

Time is discrete and there are two dates, $t = 0$ and $t = 1$. All trading takes place at time $t = 0$, while at $t = 1$ all payments are made and contracts are settled. There is a set $\mathcal{L}$ of locations, and each investor is located in one location in the set $\mathcal{L}$. There is a continuum of investors in each location and we index a representative investor in a given location by $i \in \mathcal{L}$. Investors maximize expected utility of period-1 wealth, $E\left[U\left(W\right)\right]$, for some increasing and concave U.

Investors' time-zero endowments consist of shares in firms that are domiciled in their location. Investors in every location i are of two types, common investors and swindlers, while firms are of two types, regular and fraudulent. The number of shares in each firm is normalized to one, as are the measures of investors and firms at each location.

Common investors in location i are a fraction $\kappa \in (0, 1)$ of the population in that location. They are identically endowed with an equal-weighted portfolio of all regular firms in location i. All regular firms in location i produce the same random output D_i, and pay it out as a dividend. The total measure of regular firms is κ in each location.

Swindlers are a fraction $1-\kappa$ of the population in each location. Each swindler is endowed with

[11] See, e.g., the "growth accounting" framework of Chari et al. (2007), or the voluminous international trade literature, which we do not attempt to summarize here, on the identification of trade-friction parameters.

the share of one fraudulent firm. Fraudulent firms produce no output or dividend ($D_i = 0$).

For every firm in every location, there is a market for shares where any investor can submit a demand. Moreover, there exists a market for a riskless bond, available in zero net supply. The interest rate is denoted by r.

2.2. Signals

Each investor obtains a signal of the type — regular or fraudulent — of every firm in every location. The precision of these signals depends on the locations of the investor and the firm.

Specifically, investors in every location i obtain a signal $\iota^i_{jk} \in \{0, 1\}$ about every firm k in location j. (All investors in i obtain the same signal about any given firm.) This signal characterizes the firm as either regular ($\iota^i_{jk} = 1$) or fraudulent ($\iota^i_{jk} = 0$). The signal is imperfect. It correctly identifies every regular firm as such. However, it fails to identify all fraudulent firms: it correctly identifies a fraudulent with probability π_{ij} and misclassifies it as regular with probability $1 - \pi_{ij}$. We assume $\pi_{ii} = 1$, so that investors are fully informed about their local markets.

Given this setup, Bayes' rule implies that the probability that a firm in location j is fraudulent given that investor i's signal identifies it as regular is given by

$$f_{ij} \equiv \frac{(1 - \pi_{ij})(1 - \kappa)}{\kappa + (1 - \pi_{ij})(1 - \kappa)}. \tag{1}$$

The law of large numbers implies then that f_{ij} can also be interpreted as the fraction of fraudulent firms among all firms in location j identified by the signal of investor i as regular.

2.3. Budget constraints

Letting B^{ci} denote the amount that a common investor in location i invests in riskless bonds and dX^{ci}_{jk} a bivariate signed measure capturing the number of shares of firm k in location j that she buys, the time-one wealth of a common investor located in i is given by

$$W^{ci}_1 \equiv B^{ci}(1 + r) + \int_{j \in \mathcal{L}} \int_{k \in [0,1]} D_{jk} dX^{ci}_{jk}. \tag{2}$$

The first term on the right-hand side of (2) is the amount that the investor receives from her bond position in period 1, while the second term captures the portfolio-weighted dividends of all the

firms that the investor holds. The time-zero budget constraint of a common investor in location i is given by

$$B^{ci} + \int_{j\in\mathcal{L}} \int_{k\in[0,1]} P_{jk} dX^{ci}_{jk} = \frac{1}{\kappa} \int_{k\in[0,1]} P_{i,k}\rho_{(i,k)} dk, \tag{3}$$

where $\rho_{(i,k)}$ is an indicator function taking the value one if the firm k in location i is a regular firm and zero otherwise, and P_{jk} refers to the price of security k in location j. The left-hand side of (3) corresponds to the sum of the investor's bond and risky-security spending, while the right-hand side reflects the value of the (equal-weighted) portfolio of regular firms the investor is endowed with.

The budget constraint of a swindler owning firm l in location i is similar to (3), except that the value of the agent's endowment is given by P_{il}:

$$B^{sil} + \int_{j\in\mathcal{L}} \int_{k\in[0,1]} P_{jk} dX^{sil}_{jk} = P_{il}. \tag{4}$$

Note that, as before, the notation allows investors' portfolios to have atoms, which is further useful here because, in equilibrium, swindlers optimally hold a non-infinitesimal quantity of shares of their own firms. We denote the post-trade number of shares of fraudulent firm l in location i retained by the original owner by $S^{il} = dX^{sil}_{il}$.

Finally, the time-1 wealth of a swindler is

$$W^{sil}_1 \equiv B^{sil} + \int_{j\in\mathcal{L}} \int_{k\in[0,1]} D_{jk} dX^{sil}_{jk}. \tag{5}$$

2.4. Optimization problem

Common investors are price-takers. Taking a set of prices for risky assets as given for all firms in all locations and an interest rate, a common investor maximizes

$$\max_{B^{ci}, dX^{ci}_{jk}} E\left[U(W^{ci}_1)|\mathcal{F}_i, P_{jk}, r\right] \tag{6}$$

subject to (3) and a short-selling constraint: $dX^{ci}_{jk} \geq 0$. We note that even though we impose the short-selling restriction exogenously in the body of the text, we relax this assumption in the appendix. Specifically, we show that our results are identical when we allow short sales, but extend the setup to allow costly earnings manipulation by the swindler, which acts as an (out-of-equilibrium) deterrence mechanism to potential short sellers of fraudulent firms. We relegate the details to appendix A, and for the rest of the paper we simply exclude short sales.

The investor conditions on her own information set $\mathcal{F}_i$ (i.e., on her signals about every security), as well as on the prices of all securities in all markets.

The problem of the swindler is similar to the one of the common investor with the exception that the swindler takes into account the impact of her trading on the price of her stock.[12] Similar to a common investor, the swindler who owns firm l in location i solves

$$\max_{B^{sil}, dX_{jk}^{sil}} E\left[U(W_1^{sil})|\mathcal{F}_{il}, P_{jk}, r\right] \tag{7}$$

subject to the budget constraint (4) and $dX_{jk}^{sil} \geq 0$.

2.5. Equilibrium

An equilibrium is an interest rate r and a collection of prices $P_{i,k}$ for all risky assets, asset demands and bond holdings expressed by all investors in all locations, such that: 1) Markets for all securities clear: $\kappa \int_{i\in\mathcal{L}} dX_{jk}^{ci} + (1-\kappa)\int_{i\in\mathcal{L}} dX_{jk}^{sil} = 1$ for all (j,k); 2) Risky-asset and bond holdings, $\left\{X_{jk}^{ci}, B^{ci}\right\}$, are optimal for regular investors in all locations given prices and the investors' expectations; 3) Bond holdings B^{sil} and asset holdings for all securities X_{jk}^{sil} (including a swindler's own holdings of her own firm S^{il}) are optimal for swindlers given their expectations; 4) All investors update their beliefs about the type of stock k in location j by using all available information to them — prices, interest rate, and private signals — and Bayes' rule, whenever possible.

Our equilibrium concept contains elements of both a rational expectations equilibrium and a Bayes-Nash equilibrium. All investors make rational inferences about the type of each security based on their signals, the equilibrium prices, and the interest rate, by using Bayes' rule and taking the optimal actions of all other investors (regular and swindlers) in all locations as given. The continuum of regular investors are price takers in all markets.

Swindlers, however, are endowed with the shares of a fraudulent company and take into account the impact of their trades on the share price. In formulating a demand for their security, swindlers have to consider how different prices might affect the perceptions of other investors about the type of their security. As is standard, Bayes' rule disciplines investors' beliefs only for demand realizations that are observed in equilibrium. As is usual in a Bayes-Nash equilibrium, there is freedom in specifying how out-of-equilibrium prices affect investor posterior distributions of security types.

[12] In the extension presented in Appendix A, the swindler also has the ability to manipulate the dividends of her fraudulent stock.

We note that the distinction between regular investors who are price takers and swindlers who are strategic about the impact of their actions on the price of their firm is helpful for expediting the presentation of results, but not crucial. In Appendix A we show that our equilibrium obtains in the limit (as the number of traders approaches infinity) of a sequence of economies with finite numbers of traders — both regular and swindlers — who are strategic about their price impact, as in Kyle (1989).

By Walras' law, we need to normalize the price in one market. Since we abstract from consumption at time zero for parsimony, we normalize the price of the bond to be unity ($r = 0$).

2.6. Tax equivalence

While our economy is seemingly complex, its equilibrium outcomes coincide with those of a much simpler Walrasian economy featuring bilateral taxes. The intuition behind this result is quite straightforward: Agents optimally invest in all assets for which they have positive signals and in no others (the only exception is the swindler investing in her firm), but the signal is imperfect. The failure rate of the signal translates into a lower payoff relative to that obtained by a perfectly informed investor; the proportional loss can be thought of as a tax rate, depending on the identities of both the investor and the market. In addition, swindlers have strict incentives to invest in their own firms so as to render them indistinguishable from regular firms, which ensures a pooling equilibrium that justifies the behavior of the other investors. We record this result formally:

Proposition 1 *There exists an equilibrium of the original economy in which the prices of all assets in each location are equal. Furthermore, the prices P_j and positions $dX_j^i \equiv \int_k dX_{jk}^i$ taken by investors located in market i when investing in market j, excluding swindlers' positions in their own firms, are given as a solution to the problem:*[13]

$$dX_j^i \in \arg\max_{dX_j \geq 0} EU\left(\int_{j\in\mathcal{L}} ((1-f_{ij})D_j - P_j)\, dX_j + P_i\right) \tag{8}$$

$$\kappa = \int_{i\in\mathcal{L}} (1-f_{ij}) dX_j^i. \tag{9}$$

Equation (8) formalizes the decision problem of an investor facing taxes f_{ij}, as explained above. Equation (9) is the market clearing equation. The left-hand side, κ, is due to the fact that only

[13] In the interest of concision, we plugged in the investor's budget constraint in the objective.

κ of the firms are regular. We also note that the right-hand side depends on the "tax rates"; the reason is that a proportion f_{ij} of the investment dX_j^i is made into fraudulent firms, leaving only the remainder to purchase the share of regular firms.[14]

Proposition 1 makes the description of an equilibrium relatively easy, as we illustrate in the following section. In addition, it provides us with an intuitive language to describe the degree to which any investor is at a disadvantage when investing in any given market. We make considerable use of this concept in our empirical section, where the primary goal is to quantify these disadvantages — indeed, construed as taxes.

3. Informationally Inefficient Markets: Implications

In this section we exploit the equivalence formalized in Proposition 1 between informational frictions and taxes to obtain a number of theoretical implications. We first describe the solution to the model, but the main question we address is methodological: We investigate whether existing approaches that compare the returns obtained by an investor to those of certain passive (index) investment strategies identify correctly the investor's informational advantage. Both for simplicity and in order to compare our results to the Sharpe-Lintner-Mossin CAPM, in this section we assume that the dividends D_j are jointly normal (so that in the absence of frictions the CAPM would hold). Moreover, to obtain closed-form solutions for equilibrium prices, we also assume that investors have CARA utilities, $U(W) = -e^{-\gamma W}$. Even though we make these assumptions to ease the exposition in the text, the lemmas obtain irrespective of these two assumptions.

3.1. Equilibrium prices

We start from the optimality condition of an investor in location i faced with problem (8). We let $\lambda_{ij} \geq 0$ denote the Lagrange multiplier associated with $dX_j^i \geq 0$, and $p_{ij} := 1 - f_{ij}$ be the effective payoff to investing in assets of location j.[15] We also assume that $\int_{k\in\mathcal{L}} dk = 1$. Given the

[14] We note that in a pooling equilibrium the swindler submits an elastic demand for her own firm, i.e., absorbs the residual demand for her own firm at the price P_j, so that the market for fraudulent firms clears by construction.

[15] Note that p_{ij} is the probability that security j is regular given that the signal of investor i identifies it as such. Clearly, $p_{ij} \geq \kappa$.

CARA-normal setup, the first-order condition is

$$\gamma \text{cov}\left(p_{ij} D_j, \int_{k\in\mathcal{L}} p_{ik}\, D_k dX_k^i\right) = p_{ij} - P_j + \lambda_{ij}. \tag{10}$$

Dividing this equation by p_{ij} and summing over all agents i yields

$$\gamma \text{cov}\left(D_j, \kappa D^a\right) = 1 - P_j \int_{i\in\mathcal{L}} p_{ij}^{-1} di + \int_{i\in\mathcal{L}} p_{ij}^{-1} \lambda_{ij} di, \tag{11}$$

where we introduced the notation D^a for the aggregate dividend $D^a = \int_{j\in\mathcal{L}} D_j$. We note that, by (9) and Fubini's theorem,

$$\int_{i\in\mathcal{L}} \int_{k\in\mathcal{L}} p_{ik}\, D_k dX_k^i = \int_{k\in\mathcal{L}} D_k \int_{i\in\mathcal{L}} p_{ik} dX_k^i = \kappa D^a. \tag{12}$$

The price P_j is consequently expressed as

$$\begin{aligned} P_j &= \left(\int_{i\in\mathcal{L}} p_{ij}^{-1}\right)^{-1} \times \left(1 - \gamma \text{cov}\left(D_j, \kappa D^a\right) + \int_{i\in\mathcal{L}} \lambda_{ij} p_{ij}^{-1} di\right) \\ &= \left(\int_{i\in\mathcal{L}} p_{ij}\right) \times \left(1 - \gamma \text{cov}\left(D_j, \kappa D^a\right) + \int_{i\in\mathcal{L}} \lambda_{ij} p_{ij}^{-1} di\right) \times \frac{\left(\int_{i\in\mathcal{L}} p_{ij}^{-1}\right)^{-1}}{\left(\int_{i\in\mathcal{L}} p_{ij}\right)}, \end{aligned} \tag{13}$$

which provides a natural formula. The first term captures the average post-tax payoff to investors, the second the risk adjustment and the effect of the shorting constraint, while the third measures dispersion in p_{ij} across agents. A larger dispersion, i.e., a lower value of the third term, translates into a lower price as the effective price per unit of mean dividend paid by investor i, namely P_j/p_{ij} is convex in p_{ij}, and thus agents with low p_{ij} have a stronger weakening effect on the price than agents with high p_{ij} have in the opposite direction. Clearly, due to possible differences in the average and dispersion of p_{ij}, different asset classes may be priced differently even when containing the same amount of aggregate risk and being held in positive amounts by all agents ($\lambda_{ij} = 0$).

3.2. Alpha does not measure skill

A common approach to measuring the skill (i.e., the informational advantage) of a fund manager is to regress her historical returns on an asset pricing model (such as the Sharpe-Lintner-Mossin CAPM) or more generally some model of the stochastic discount factor. It is then quite common

to interpret the intercept (alpha) of such a regression as a measure of the manager's skill, after "having controlled for risk."

We show that, in an informationally inefficient market, even passive index strategies have alphas (either negative or positive), while informed investors may come across as having negative alpha. This shows that alphas are only an indication of model misspecification; the sign and size of alphas, however, do not generally map into a meaningful measure of skill. As we show shortly, when an asset pricing model fails due to the fact that different investors face essentially different returns within the same asset class, a conceptually more appropriate measure of informational advantage is a measure proposed by W. Sharpe and known as "style alpha" (we provide a formal description of this concept shortly). This measure does not employ a universal model of risk adjustment. Rather, it measures an investor's performance relative to the returns of the asset classes she invests in.

Since in our model the CAPM would hold in the absence of informational frictions, we start by computing the CAPM alpha of a passive index strategy in location j. From the perspective of an econometrician,[16] the return of an index strategy (i.e., an uninformed strategy, which utilizes no private signals) investing in location j is given by $M_j = \frac{\kappa D_j}{P_j}$ which has expected return $\frac{\kappa}{P_j}$. Similarly, the return on an index replicating the market portfolio is $M = \frac{\kappa D^a}{\int_{k\in\mathcal{L}} P_k dk}$. Using these observations, and recalling that the interest rate is normalized to zero, the alpha of buying the index in location j is given by

$$\begin{aligned}\alpha_j &= \frac{\kappa}{P_j} - 1 - \frac{cov\left(\frac{\kappa D_j}{P_j}, \left(\int_{k\in\mathcal{L}} P_k dk\right)^{-1} \kappa D^a\right)}{\left(\int_{k\in\mathcal{L}} P_k dk\right)^{-2} \kappa^2 \sigma_a^2} \left(\frac{\kappa}{\int_{k\in\mathcal{L}} P_k dk} - 1\right) \\ &= \left(\beta_j^D \frac{\int_{k\in\mathcal{L}} P_k dk}{P_j} - 1\right) + \frac{\kappa}{P_j}\left(1 - \beta_j^D\right), \end{aligned} \tag{14}$$

where β_j^D is the "cash-flow beta"

$$\beta_j^D = \frac{\text{cov}(D_j, D^a)}{\text{var}(D^a)}. \tag{15}$$

In the special case in which there is no asymmetric information ($p_{ij} = \kappa$) equations (13) and (14) imply the usual CAPM relation ($\alpha_j = 0$).[17] However, in general $\alpha_j \neq 0$, even for passive

[16] Note that we are studying the original, asymmetric-information economy, using the simplification provided by the tax economy for the computation of prices.

[17] To see this, notice that equation (13) implies that $P_j = \kappa - \gamma\kappa^2\beta_j^D\sigma_a^2$. Then it follows from (14) that $\alpha =$

strategies. To see this in the simplest possible case, consider a world with $\beta_j^D = 1$ for all j. Accordingly, $\alpha_j = \left(\frac{\int_{k\in\mathcal{L}} P_k dk}{P_j} - 1\right)$. In addition, equation (13) implies that some asset classes may still exhibit lower (or higher) than average prices, despite all assets having the same exposure to aggregate risk. For instance, lower overall quality of information in asset class j (low values of p_{ij} compared to other asset classes) will translate into a lower than average price for that class; since $\alpha_j = \left(\frac{\int_{k\in\mathcal{L}} P_k dk}{P_j} - 1\right)$ even an index investment in such a class will have positive alpha.[18]

If uninformed (passive) strategies command alphas, then alphas cannot be an accurate measure of an investor's skill. Indeed, continuing with the assumption that $\beta_j^D = 1$ for all j, the alpha resulting from a regression of the return that investor i obtains (when investing in location j) on the return of the market portfolio is given by

$$\alpha_{ij} = \frac{p_{ij}}{\kappa} \frac{\int_{k\in\mathcal{L}} P_k dk}{P_j} - 1. \tag{16}$$

Hence, even an investor who has an informational advantage $p_{ij} > \kappa$ might exhibit a negative alpha when that informational advantage happens to be in an asset class that is comparatively more expensive than the average asset class, i.e., $\int_{k\in\mathcal{L}} P_k dk < P_j$.

The fact that alpha does not measure skill is not simply a restatement of E. Fama's "joint hypothesis problem". Fama observed that positive alpha could either mean rejection of informational efficiency or rejection of the asset pricing model. We address a different issue: Taking as given the presence of informational inefficiency, we ask whether the resulting alphas map into a meaningful measure of skill. We obtain a negative answer.

Relatedly, we would like to remark that the failure of the CAPM is not a result of mis-specifying the model. Rather, the equilibrium outcome of binding shorting constraints — due to agents' differential information even conditional on prices — leads to a violation of the law of one price, and thus implies that no stochastic discount factor exists from the perspective of an uninformed investor. To make a formal statement, we introduce a derivative security, in zero net supply, that

$\left(\beta_j^D \frac{\kappa-\gamma\kappa^2\sigma_a^2}{P_j} - 1\right) + \frac{\kappa}{P_j}\left(1-\beta_j^D\right) = \frac{\kappa-\gamma\kappa^2\beta_j^D\sigma_a^2}{P_j} - 1 = 0.$

[18] An additional, less interesting observation is that mispricing is not related to the amount of risk. Consider for instance the case $p_{ij} = p$. All indexes, including the market, earn negative excess returns $\kappa - p$ (compared to an informed strategy) before the risk adjustments, but a high β_j^D index is benchmarked against a leveraged market index, thus one with even higher negative returns, and therefore has positive alpha. Conversely, low β_j^D is associated with negative alpha. We shut down this channel by assuming $\beta_k^D = 1$ for all k.

pays the aggregate dividend κD^a, and denote by P^a its price.[19] Adjusting the budget constraints accordingly, and thus the second argument to the covariance operator on left-hand side of equation (10), equations (10) and (11) continue to hold, possibly with different Lagrange multipliers. Equations (14) and (16) do not depend on the existence of this derivative security.

Lemma 2 *Suppose that a stochastic discount factor ξ exists such that*

$$P_j = E\left[\xi \kappa D_j\right].$$

Then ξ does not price the aggregate-dividend security:

$$P^a < \int_j P_j dj = E[\xi \kappa D^a]. \tag{17}$$

As is well known (see, for instance, Cochrane (2005)) the existence of a stochastic discount factor is equivalent to a beta representation for returns (with one factor). There consequently exists no such representation, i.e., no factor model that yields zero alphas.

3.3. Style alphas

An alternative and quite popular approach, routinely used in practice to infer skill, is the so-called style analysis, proposed by Sharpe (1992). According to this approach, the return of each manager is regressed on the passive returns of all possible asset classes. Moreover, to interpret the betas as portfolio weights, one additionally requires that the sum of the betas on the passive strategies add up to one (and are restricted to be positive). In our economy, style analysis would correctly identify investors who possess superior information, as shown by the following lemma.

Lemma 3 *Let $dw_j^i = \frac{P_j dX_j^i}{\int_{j\in\mathcal{L}} P_j dX_j^i}$ be the risky-portfolio weight of the investment in location j by an investor in location i. Consider the style regression of the gross return obtained on her risky portfolio by an investor in location i. The constant α_i^s in this regression is a positive multiple of the portfolio-weighted informational advantage of investor i across all market in which she invests:*

$$\alpha_i^s = \nu \int_{j\in\mathcal{L}} \left(\frac{p_{ij}}{\kappa} - 1\right) dw_j^i. \tag{18}$$

[19] We note that, in the absence of aggregate risk, such a security already exists in our model: the risk-free security.

The multiplicative constant ν does not depend on the agent i.

Of course, despite its theoretical appeal, a well-understood limitation of style analysis is that, with a short sample of data, one may be restricted in the number of asset classes to include in the regression.

3.4. Active asset management and indexing

Leaving aside issues of risk adjustment, return comparisons between passive and active strategies (even within the same asset class) can yield surprising conclusions in a world where the degree of asymmetric information is random.

A common observation in financial markets is that active asset managers underperform index investing after fees. In this section we show that when markets are informationally inefficient, comparing the average returns of active and indexing strategies in the *same* asset class may not be an appropriate measure of the relative attractiveness of the two strategies from the perspective of an investor.

Specifically, we consider the following modification of the setup. We assume that κ is not known to investors, but instead is drawn from a distribution $F(\kappa)$ independently of other uncertainty. Furthermore, we assume that investors in all locations have a choice. Investors in location i can either pay a cost to obtain the signals $\{\iota^i_{jk}\}_k$, or invest without any information. If the investors choose to get informed, they pay $\varphi_{ij}D_j$, at time 1, for every share purchased in market j, and obtain the return $\frac{p_{ij}D_j}{P_j}$ in that market.[20] If uninformed, they obtain $\frac{\kappa D_j}{P_j}$. The following obtains.

Lemma 4 *Suppose that agent i must pay $\varphi_{ij}D_j$ (at time 1) per share in market j to invest based on the signals $\iota^i_{j\cdot}$. Then for any $\{p_{ij}\}_{i,j}$ there exist (an open set of) values of φ_{ij} and distributions F for κ such that investors choose to pay the proportional costs φ_{ij} and get informed rather than invest passively (not knowing κ), while the expected return of a passive strategy exceeds that of an active strategy:*

$$\frac{E(\kappa)}{P_j} > \frac{\int_{k\in\mathcal{L}}(p_{ik}-\varphi_{ik})\,dX^i_k}{\int_{k\in\mathcal{L}}P_k dX^i_k} \text{ for all } i,j\in\mathcal{L}. \tag{19}$$

[20] The assumption that the observation costs are incurred at time 1 (rather than time 0) is convenient for the proof, but not essential for the economics of the result.

The result is intuitive. The cost φ_{ij} reduces the average payoff from investing in market j, but also eliminates the risk introduced by the randomness of κ, since equilibrium prices, and therefore informed returns, do not depend on κ. Risk-averse investors are willing to pay a premium to avoid that risk.

To summarize, if the market is subject to time-varying asymmetric information, then the average return differential between active and passive strategies may be negative, even though every investor chooses to invest actively. From the perspective of an investor, active investing offers insurance against fluctuations in asymmetric information. This insurance makes the investor willing to accept a lower average return when investing actively.[21]

4. Empirical Specification and Data

The previous section highlights some limitations of relying exclusively on return differentials to infer financial frictions. In this section we explore the idea of using the information in *both* returns and observed portfolio choices to infer the shadow tax rates (i.e., the differences in marginal valuations of the same asset by different investors). To illustrate our proposed method, we apply it to international data. As a result, for the remainder of this section, an investor i refers to the representative investor of country i and the asset class j refers to country $j's$ assets. To facilitate empirical analysis and comparisons with the literature we make the same assumptions as the Sharpe-Lintner-Mossin CAPM, (i.e., we assume that all returns are jointly normal or that investors have mean-variance preferences). We extend this basic framework to allow for frictions and different currencies.

Specifically, we start by assuming that — in the absence of frictions and multiple currencies — the representative investor's i optimal portfolio would be given by the familiar Markowitz portfolio

$$w_i = \frac{1}{\gamma}\Omega^{-1}\left(\mu - Re_{N\times 1}\right), \tag{20}$$

where w_i is a vector of portfolio holdings for investor i, γ is the (relative) risk aversion (assumed common across investors), μ is a set of expected gross returns, Ω is the covariance matrix of returns, R is the gross rate of interest and $e_{N\times 1}$ is an $N \times 1$ vector of ones, with N the number of countries.

We depart from conventional mean-variance analysis by allowing investors in different countries to differ with respect to their information sets. Our earlier finding is that such informational

[21] See, e.g., Moskowitz (2000), Sun et al. (2009), and Glode (2011) who find evidence of active funds outperformance in "down" markets.

differences act as taxes, which change equation (20) to

$$w_i = \frac{1}{\gamma}\widehat{P}_i^{-1}\Omega^{-1}\widehat{P}_i^{-1}\left(\widehat{P}_i\mu - RQ_i e_{N\times 1}\right), \tag{21}$$

where the matrix $\widehat{P}_i = \text{diag}(p_{i1}, .., p_{iN})$ is a diagonal $N \times N$ matrix capturing the after-tax gross return of investor i in country j and $Q_i \equiv \left(I_{N\times N} - \frac{1}{R}\text{diag}\left(l_{i1}, ...l_{iN}\right)\right)$ with $I_{N\times N}$ being the $N \times N$ identity matrix and $\text{diag}\left(l_{i1}, ...l_{iN}\right)$ denoting a diagonal matrix containing the Lagrange multipliers associated with the short-sale constraint $w_{ij} \geq 0$ for all j.

We allow the constant κ of our previous analysis (i.e., the shadow tax rate associated with passive, uninformed strategies) to vary by location, and let $K = \text{diag}\left(\kappa_1, .., \kappa_N\right)$ denote the corresponding diagonal matrix. We obtain the following result.

Lemma 5 *Let $\Sigma \equiv K\Omega K$ denote the covariance matrix of passive, index returns and similarly let $\mu^o = K\mu$ denote the vector of expected index returns. Then, with $\Pi_i \equiv K\widehat{P}_i^{-1}Q_i$,*

$$\Pi_i^{-1}w_i = \frac{1}{\gamma}\Sigma^{-1}\left(\mu^o - R\Pi_i e_{N\times 1}\right). \tag{22}$$

Equation (22) is cental for our purposes. It relates quantities that one can observe in principle (expected returns of index strategies, covariances of index strategies, and country portfolios) to the unobserved diagonal matrix Π_i, whose typical diagonal element is $\pi_{ij} = \kappa_j\left(1 - \frac{l_{ij}}{R}\right)/(1 - f_{ij})$. In words, π_{ij} captures the ratio of the gross return obtained by a passive strategy in country j relative to the shadow after-tax return obtained by the representative investor from country i investing in country j.

There are three practical issues that arise when using equation (22) for identifying Π_i. First, not all countries have the same currency. Second, the expected returns μ^o are notoriously hard to estimate from limited time series of data, and — more importantly — equation (22) ignores the restrictions posed on μ^o by market clearing. Third, the measurement of cross-country portfolios is very likely subject to measurement error as well, which could also impact the recovery of Π_i.

To address the first problem in the most straightforward way, we simply follow the literature on the international CAPM. Specifically, we drop the assumption of one currency and instead assume that there are $L + 1 \leq N$ currencies in the world, where $L + 1$ is the reference country's currency — the US dollar for our empirical analysis. $L + 1$ is allowed to be smaller than N to allow for currency unions. We assume that there is no informational advantages/distortions when investing in currencies.

Allowing for foreign-country denominated bonds starts with enlarging the variance-covariance matrix to include the dollar-denominated returns from investing in foreign-currency denominated bonds. Specifically, we assume that the vector of expected returns is now given by $\mu = \left(\mu^S, \mu^f\right)'$, where μ^S is the vector of dollar-denominated stock returns and μ^f is the vector of dollar-denominated returns from investing in foreign bonds. Similarly, the covariance matrix of (passive-strategy) returns is now given by

$$\Sigma = \begin{bmatrix} \Sigma_S & \Sigma_{Sf} \\ \Sigma_{Sf}' & \Sigma_f \end{bmatrix},$$

where Σ_S captures the variance-covariance matrix of index investments in stocks, Σ_{Sf} is the covariance matrix of returns between stocks and bonds, and Σ_f is the variance-covariance matrix from investing in foreign-currency denominated bonds.

Replicating standard arguments of the international CAPM, the presence of random exchange rates modifies equation (22) to[22]

$$\Pi_i^{-1} w_i = \frac{1}{\gamma} \Sigma_{S|f}^{-1} \left(\mu^{S,f} - R\Pi_i e_{N\times 1}\right) \tag{23}$$

where w_i is once again the stock portfolio, but the covariance matrix is now given by $\Sigma_{S|f} \equiv \Sigma_S - \beta' \Sigma_f \beta$ with $\beta' \equiv \Sigma_{Sf}\Sigma_f^{-1}$, and the expected return vector is $\mu^{S,f} = \mu^S - \beta'\left(\mu^f - Re_{L\times 1}\right)$. (Throughout we assume that Σ_S and Σ_f are invertible.)

The "multiple-currencies" equation (23) is essentially identical to the "single-currency" equation (22), except that stock returns should be understood as "currency-hedged", i.e., as the residuals from regressions of dollar-denominated stock returns on the (dollar denominated) returns of all foreign bonds. We note that our approach to dealing with multiple currencies could be extended to allow for different consumption baskets (see, e.g., Adler and Dumas (1983) or Cooper and Kaplanis (1994)). We prefer to keep the currency-aspects of the model as simple as possible for simplicity; however, when we present our empirical results, we revisit this issue and include controls to account for possible hedging effects arising from different consumption baskets.

Next we address the second problem, by imposing market clearing conditions, deriving the vector $\mu = \left(\mu^S, \mu^f\right)$, and then providing an expression for Π_i, $i = 1, \ldots, N$. The next lemma provides an intermediate step.

[22] See Sercu (1980) and the proof of Lemma 6.

Lemma 6 *Let* Π *denote an* $N \times N$ *matrix whose* N *columns contain the* N *diagonal elements of the matrices* Π_i, $i \in \{1, \ldots, N\}$. *Similarly, let* W *denote a matrix whose columns are given by the (stock) portfolios* w_i. *Let* η *denote the vector of the the wealth weights (as a fraction of aggregate world-wealth) of each country, and similarly let* m *denote a vector of the market capitalization of the stock market of each country (as a fraction of aggregate stock market capitalization).*

Then, up to terms of second or higher order in $\|\Pi - 1\|$, *we have*[23]

$$vec\,(W) = A\, vec\,(\Pi)\,, \tag{24}$$

where

$$A \equiv I_{N\times N} \otimes diag\,(m) + \frac{R}{\gamma}\left(\left(e_{N\times 1}\eta' - I_{N\times N}\right) \otimes \Sigma_{S|f}^{-1}\right). \tag{25}$$

The vector of equilibrium expected returns is given by

$$\mu^S - Re_{N\times 1} = \gamma\Sigma_{S|f}m + R\left(\Pi\, diag(\eta) - I_{N\times N}\right)e_{N\times 1} + \beta'\left(\mu^f - Re_{L\times 1}\right), \tag{26}$$

where

$$\mu^f - Re_{L\times 1} = \gamma\Sigma'_{Sf}m + (1-\gamma)\,\Sigma_f\widehat{\eta}, \tag{27}$$

and $\widehat{\eta}$ *is a vector whose typical element* j *is the wealth weight of investors who use currency* j *as their reference currency.*

Equation (24) provides an explicit relation between the portfolio weights W and the matrix of frictions Π. The matrix A, which controls the correspondence from frictions to portfolios, has two terms: The first term reflects market capitalization weights. Indeed, when $vec\,(\Pi) = e_{N^2\times 1}$, then only the first term survives, in the sense that $vec\,(W) = e_{N\times 1} \otimes m$. Alternatively phrased, when there are no frictions, then all countries should hold the world market portfolio — a well-known implication of the CAPM. The second term of (25) reflects the magnitude of equilibrium portfolio deviations from the CAPM. These deviations depend on risk aversion, the interest rate, the wealth weight of different countries, and most importantly the covariance properties of the returns of different countries — all quantities that influence the tradeoff between diversification and tilting one's portfolio to locations where the investor faces lower taxes. Importantly, expected returns on stocks and foreign-denominated bonds do not enter the matrix A.

[23] As part of our empirical exercise, we also computed the terms Π exactly using a numerical procedure. Because the values of $\Pi - 1$ are within a few basis points of zero, the exact and approximate solutions are essentially the same.

We conclude by addressing the third practical problem in recovering Π, namely the presence of measurement error. We start by noting that in the absence of measurement error in W, equation (24) would allow recovery of Π, since the matrix A is invertible. However, due to data limitations, it is likely that portfolios are measured with non-trivial error. Hence, to obtain more reliable estimates of Π, we also use the N equations (26), i.e., we use the information in expected returns to improve the recovery of Π. To operationalize this idea, we allow for observation errors U and express equations (24) and (26) as

$$Y = X \times vec\,(\Pi) + U, \tag{28}$$

where

$$Y = \begin{bmatrix} vec(W) \\ \frac{\mu^S - \beta'\left(\mu^f - Re_{N\times 1}\right) - \gamma \Sigma_{S|f} m}{R} \end{bmatrix}, \tag{29}$$

$$X = \begin{bmatrix} A \\ \eta' \otimes I_{N\times N} \end{bmatrix}, \tag{30}$$

and U is assumed to have covariance matrix Z, which we specify shortly. Equation (28) leads to the estimator

$$vec\,(\Pi) = \left(X' Z^{-1} X\right)^{-1} X' Z^{-1} Y. \tag{31}$$

The estimator (31) weights the information contained jointly in equations (24) and (26) to obtain the most efficient estimate of $vec\,(\Pi)$. To compute Π we need return data, market and wealth weights and information on portfolios. The next section describes our data sources, while the following section describes some details of the estimation procedure, in particular our specification of the covariance matrix Z.

4.1. Data

To obtain estimates of Π we need country-level stock market returns and stock market capitalizations, country-level wealth weights, exchange rates, and bilateral portfolio holdings. The monthly data for country-level stock return and stock market capitalizations come from Compustat Global and Compustat North America. We use these data to calculate the monthly total market capitalization of each country in US dollars. We only keep ordinary shares and we omit ADRs from our

sample. In addition, the corresponding exchange rates are also obtained from Compustat Global and Compustat North America. Due to data limitations for a large panel of countries, we focus our stock return sample on the period past the introduction of the Euro.[24] Our results are similar if stock return data are obtained from MSCI.

The bilateral equity holding data are from consolidated portfolio investment survey (CPIS) from the International Monetary Fund (IMF). This survey reports bilateral portfolio equity holdings for 77 origin countries and 241 target countries for the years 2001-11. The survey for year 2011 contains the latest available holdings data from the IMF at the time of our study.

Under the guidance of the IMF, national compilers collect the portfolio holdings data through national surveys addressed to end-investors and custodians. The national compilers try to minimize under- or double- counting, and to make CPIS data comparable across countries. For each country, the IMF produces the geographic breakdown of its residents' aggregate holdings of securities issued by non-residents, and all holdings are denominated in U.S. dollars. The CPIS data set has some drawbacks. For example, if a parent company in country A has a foreign subsidiary in country B, which holds a financial asset in country C, this holding is counted as country B's holding in country C, rather than country A's holdings in country C. Thus, just as with any survey data on cross-border holdings, the CPIS data set also suffers from the third-party holdings problem. In addition, the coverage is incomplete. For these reasons, it is important to allow for measurement error in portfolios (the next section contains further details). We revisit the issue of measurement error in Section 4.4., after presenting the results. There we explain in greater detail why portfolio measurement error (and incomplete coverage) is unlikely to have significant impact on our results.

These CPIS data have been used by many prior studies in economics and finance such as Faruqee et al. (2004), Lane and Milesi-Ferretti (2005), Aviat and Coeurdacier (2007), Berkel (2007), Daude and Fratzscher (2008), and Bekaert and Wang (2009). We focus on the most recent available portfolio holding data for the year 2011. Using holding data for earlier years yield similar results. Due to incomplete data on portfolio observations, coupled with shortage of continuous, reliable stock and exchange rate data on a multitude of countries, we choose to focus on OECD countries only. A handful of countries — Chile, Estonia, Slovakia, and Slovenia — joined the OECD only

[24] We take January 2003 as the starting date, i.e., 12 months after the physical introduction of the Euro in all Euro-countries in our sample (including Greece). We make this choice to mitigate the impact of any transient adjustments to the new currency. Earlier start dates do not impact the results.

recently and their portfolio data are mostly missing. We exclude these OECD countries, as well as Ireland, which has a negative holding of its own stock market and Luxembourg, which is a financial center. Financial centers are likely to act as pure intermediaries that are neither the true source nor the true destination of foreign investments.

We consider several possible measures of wealth weights. One is country-level GDP in 2011, while the others are calculated from stock-market size. These measures have a correlation above 97%, and thus our results are insensitive to the choice of proxy. We report results based on GDP as a proxy for wealth weights.

In the regressions, we use IMF data on the real GDP per capita adjusted for PPP in year 2011. The data on the size of the financial sector as a fraction of GDP and the size of real estate finance sector (value added as a fraction of GDP) are from the national accounts database in the OECD's website. We use the most recent available data, which are for the year 2008. The data on domestic private credit over GDP in year 2011 are obtained from the World Bank Financial Structure database (see, e.g., Beck et al. (2010)).

We also use bilateral (geographical) distance measures from CEPII. CEPII has calculated bilateral distances (in kms) for most countries across the world (225 countries in the current version of the database). As a measure of familiarity between two countries, we use categorical data on Facebook friendships between countries i and j. The data are available on the Facebook website.[25] For each country, the website lists (in order of importance) the five closest countries in terms of friendship connections. We assign the number six to all countries that do not rank in the top five. Finally, we use data on the fraction of imports from country j as a fraction of the GDP of country i. We use data for 2011 from the IMF.

4.2. Estimation

Having described the data that define the matrices X and Y in equation (31), we finally need to specify the covariance matrix Z of error terms. Given the very large number of parameters compared to the dimension of equation (31), we adopt a semi-parametric approach. Specifically,

[25]The web address is https://www.facebookstories.com/stories/1574/interactive-mapping-the-world-s-friendships.

we assume that Z is block diagonal, so that

$$Z = \begin{bmatrix} Z_1 & 0_{N^2 \times N} \\ 0_{N \times N^2} & \frac{1}{T}\widehat{\Sigma}_{S|f} \end{bmatrix}.$$

In words, we assume that the measurement error in the N^2 portfolio equations is uncorrelated with the measurement error in the N expected-return equations. For the error terms of the expected return equations, we simply use the empirical estimate $\frac{1}{T}\widehat{\Sigma}_{S|f}$. Finally, we specify Z_1 parametrically. Specifically, we assume that the error terms in the measurement portfolios are uncorrelated across countries, but are correlated within a country:

$$u_{ij} = w_{ij}\left(\sigma_\varepsilon \varepsilon_i + \sigma_\xi \xi_{ij}\right) \text{ with } cov\left(\varepsilon_i\right) = I_{N \times N} \text{ and } cov(\xi_{ij}) = I_{N^2 \times N^2}. \tag{32}$$

We make three observations about (32): i) the magnitude of each error term is proportional to the portfolio weight w_{ij} (i.e. the share of country $i's$ portfolio that is invested in country j), ii) Error terms are correlated within each country i due to the presence of the error term ε_i. iii) The portfolio-specific term ξ_{ij} captures an idiosyncratic error in the measurement of each portfolio. The error specification (32) is motivated by the data limitation that we only observe the equity portfolio of each country. As a result, differences in leverage across countries are likely to be over- or under-stating the portfolio of each country by the same proportion of w_{ij}. Another reason that leads to a similar source of measurement error is the difficulty of determining the free float of market capitalization in every country, which may be affected by cross-holdings, etc.

Having specified the structure of Z, the computation of Π can now be accomplished by specifying $\gamma, \sigma_\varepsilon$, and σ_ξ. For reasons that we explain later, the specification of γ is unimportant for the structure of the matrix Π (γ acts as a scaling parameter). Therefore we simply set $\gamma = 2$, and in Section 4.4. we show that our conclusions are unaffected as we vary γ. The other two parameters, σ_ε and σ_ξ, are estimated in an iterative way. Specifically, we start with an initial guess for these parameters, estimate an initial Π from equation (31), obtain residuals, estimate σ_ε and σ_ξ by using the moment equation implied by (32), re-estimate Π, and proceed till convergence. This procedure results in the values $\sigma_\varepsilon = 0.15$ and $\sigma_\xi = 0.16$.[26]

[26] A technical issue arises because some w_{ij} are equal to zero, so that the covariance matrix of errors is singular. To ensure a non-singular covariance matrix we evaluate (32) with $\max(0.03, w_{ij})$, rather than w_{ij}, which implies that the minimum standard deviation of the observation error of any given portfolio is bounded below by

4.3. Results

4.3.1. Basic properties of the implied tax rates

Throughout this section, we focus on the quantity $\tau_{ij} = \frac{\Pi_{ij}}{\Pi_{ii}} - 1$. This quantity has a straightforward interpretation as the valuation wedge (the shadow tax) between a foreigner from country j and the local investor from country i when investing in country i. Besides this intuitive interpretation, this quantity has the important advantage of being essentially unaffected by the presence of measurement error in expected returns (μ^S), and also robust to different assumptions on risk aversion (up to multiplicative re-scaling). We substantiate these last two claims in the next section, where we show that τ_{ij} is primarily impacted by measurement error in both the covariance and portfolio matrices. Due to the presence of measurement error, one should exercise caution in interpreting individual values of τ_{ij}; collectively, however, the values of τ_{ij} exhibit some interesting patterns, which we describe next.

Figure 1 reports τ_{ij} as a histogram. Most values of τ_{ij} are positive, although a few are negative — a manifestation of measurement error in individual τ_{ij}, due to measurement error in the covariance matrix and the portfolio holdings. (Section 4.4. shows that it is possible to ensure positive values by making assumptions on the structure of the covariance matrix of returns to reduce the impact of measurement error. We prefer to not impose any restrictions on the covariance matrix for our baseline results.)

The first hypothesis that we test is that the average value of τ_{ij} equals zero. Indeed, if no frictions were present, then we would expect both positive and negative values of τ_{ij} due to measurement error, but a mean value of zero. In Table 1 we regress all the values of τ_{ij} on a constant and reject the hypothesis that the average τ_{ij} is zero.

Figure 2 reports τ_{ij} as a color-coded table. Countries are aligned according to their WEO (World Economic Outlook) code, which implies that entries for the relatively more advanced economies appear generally (but not always) at the left and top of the figure. Rows refer to the destination country of portfolio investments, while columns refer to origin country. For instance, the entry

$\max(0.03, w_{ij})\sqrt{\sigma_\varepsilon^2 + \sigma_\xi^2} \approx 0.5\%$. Besides ensuring that the covariance matrix of error terms is invertible — so that expression (31) is meaningful — this minimum variance of the error term ensures that small portfolios don't exert undue influence on the estimation procedure. We note that this truncation affects only the assumed minimum standard deviation of the observation error in the matrix Z. We do not perform any truncation of the actual values of the portfolios in the vector Y.

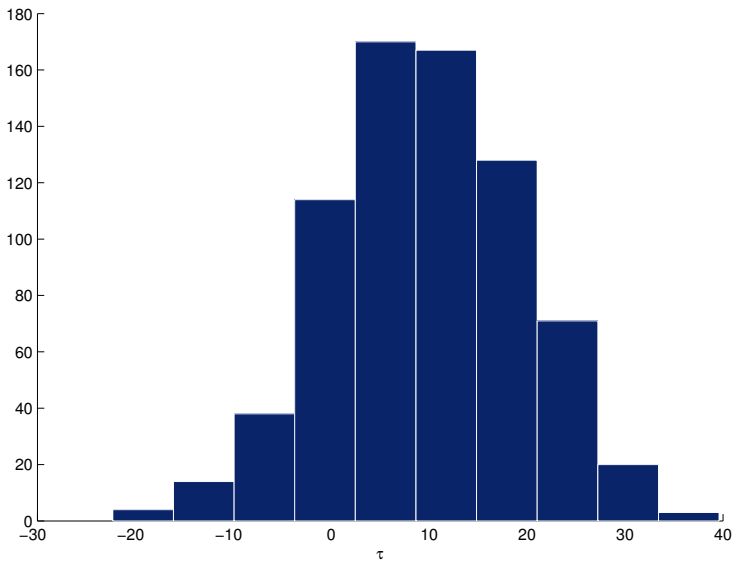


Figure 1: Histogram of implied tax rates τ_{ij} (in basis points per month).

in the second row, first column, is the implied tax faced by the representative US investor when investing in the UK (relative to the representative UK investor), while the first-row, second-column entry is the implied tax rate faced by the representative UK investor when investing in the US (relative to the representative US investor).

The similarity of colors across a given row shows that certain countries are lower-implied-tax destinations than others. Furthermore, the matrix is not symmetric around its diagonal, i.e., $\tau_{ij} \neq \tau_{ji}$. These facts suggests that certain countries seem to present foreign investors with high shadow tax rates, irrespective of the origin of the foreign investor. An additional implication is that destination-country fixed effects play a larger role in explaining the variation in implicit tax rates than symmetric, bi-directional variables whose value is not affected when origin and destination country are reversed (e.g., common institutions, culture, geographic distance, etc.). Table 1 confirms these visual impressions. Indeed, Table 1 shows that the bulk of the variation in implicit tax rates is due to destination-country fixed effects (73% of the variation). Origin-country fixed effects also account for a non-trivial part of the variation (approximately 25% on their own). The variation not explained by either type of fixed effect is small.

Another impression from Figure 2 is that frictions are rather small in the top left part of the table and become larger in the bottom and right part of the table. This suggests that the implied

		Origin																										
		USA	UK	Austria	Belgium	Denmark	France	Germany	Italy	Netherlands	Norway	Sweden	Switzerland	Canada	Japan	Finland	Greece	Iceland	Portugal	Spain	Turkey	Australia	New Zealand	Israel	Korea	Czech Rep.	Hungary	Poland
Destination	USA	0.00	0.20	3.96	6.92	3.97	8.35	3.47	11.60	2.23	2.59	2.06	-3.20	2.98	10.57	3.34	7.51	5.16	4.69	9.41	13.21	6.43	5.01	6.41	7.08	7.67	6.37	8.82
	UK	7.28	0.00	6.88	8.89	6.61	12.72	6.22	16.84	4.77	5.06	3.50	-3.63	7.57	15.14	6.40	8.61	7.83	4.59	13.43	18.07	10.91	9.60	11.51	10.86	9.90	9.30	12.90
	Austria	3.38	-5.05	0.00	3.30	0.53	9.54	2.17	14.19	-0.25	-1.80	-0.58	-10.88	-0.23	9.88	2.33	0.36	3.16	-2.20	9.70	16.51	5.90	6.73	7.13	7.95	-1.78	1.28	7.47
	Belgium	4.34	-4.58	1.44	0.00	1.13	8.20	0.79	12.93	-0.95	-1.21	-1.28	-11.14	1.84	11.52	1.46	4.16	4.13	-0.40	9.96	13.43	6.50	7.01	7.36	3.60	3.33	3.92	7.54
	Denmark	8.38	0.86	7.04	9.57	0.00	14.46	6.32	19.67	4.64	4.04	-1.15	-5.26	8.39	17.20	3.10	10.27	7.68	4.02	15.76	18.95	12.48	10.47	11.73	7.89	6.91	7.20	12.42
	France	-5.07	-11.36	-5.75	-3.28	-5.58	0.00	-7.27	5.43	-7.62	-7.29	-9.85	-17.21	-4.62	5.08	-6.34	-4.50	-3.15	-8.36	1.80	6.62	-0.42	-2.48	0.20	-2.68	-0.92	-2.90	1.34
	Germany	3.55	-1.92	2.58	5.63	3.11	8.98	0.00	13.69	1.04	1.14	-1.31	-8.12	4.23	12.47	1.65	4.25	5.45	1.91	10.32	14.17	7.37	5.37	7.24	5.82	7.43	6.30	10.26
	Italy	-8.67	-16.85	-10.47	-8.17	-9.95	-4.08	-11.65	0.00	-11.91	-11.44	-14.34	-22.30	-8.03	1.06	-11.21	-11.29	-7.60	-15.45	-3.50	2.40	-4.07	-6.46	-5.23	-6.46	-4.88	-7.59	-2.82
	Netherlands	5.59	-1.94	4.53	6.57	4.01	11.83	3.18	17.24	0.00	1.27	1.26	-7.60	4.06	14.94	4.64	7.08	6.35	4.48	13.02	17.99	9.56	9.58	10.35	7.39	8.95	7.57	12.62
	Norway	9.73	1.18	7.82	10.26	6.81	17.37	6.77	23.81	4.03	0.00	4.60	-6.12	3.84	20.85	6.95	12.15	10.75	10.35	19.63	26.12	12.24	13.05	12.89	12.40	9.58	10.52	17.08
	Sweden	11.67	4.99	10.43	13.45	7.35	16.57	8.46	21.35	9.06	8.97	0.00	-1.83	13.58	19.88	6.24	12.75	12.67	7.99	18.09	19.77	15.19	11.32	15.32	11.71	13.01	11.33	17.19
	Switzerland	13.52	8.11	12.94	14.87	12.94	18.64	11.89	22.67	11.49	11.49	9.16	0.00	13.52	21.28	12.17	15.22	14.76	11.68	19.88	23.34	16.39	15.32	16.22	16.38	16.53	14.73	18.69
	Canada	5.94	3.43	6.65	8.31	7.16	11.39	6.87	14.71	4.63	3.36	6.73	-0.57	0.00	12.50	6.70	8.46	7.66	8.89	12.79	16.95	8.02	10.54	9.59	9.16	8.56	8.76	11.22
	Japan	6.18	0.82	4.95	7.35	4.73	10.25	5.29	13.22	3.62	3.68	3.16	-1.74	4.60	0.00	7.03	5.48	4.90	3.68	10.90	10.35	6.69	5.91	8.11	6.16	4.65	5.56	8.68
	Finland	12.20	6.09	11.80	15.07	8.72	19.12	9.64	24.23	10.34	9.22	3.23	-1.21	12.56	24.15	0.00	16.47	14.32	10.32	20.96	25.62	16.39	14.21	16.31	14.23	17.18	15.27	19.92
	Greece	27.81	15.85	22.50	27.35	24.46	31.71	20.97	37.61	21.53	21.87	19.15	8.91	24.86	35.51	25.09	0.00	26.76	15.58	30.12	39.60	30.66	28.24	30.27	22.77	23.21	22.80	28.44
	Iceland	16.79	8.29	15.12	17.45	12.40	20.86	15.10	24.64	12.86	13.48	12.24	5.16	15.06	20.47	14.91	16.36	0.00	13.47	21.11	23.15	19.33	16.71	19.84	20.60	17.84	16.85	21.65
	Portugal	21.85	11.35	18.99	21.79	17.94	26.27	18.26	31.47	18.09	18.66	13.73	5.30	23.19	29.97	18.36	17.17	21.89	0.00	24.56	31.09	24.30	21.61	24.23	19.81	19.09	22.33	25.75
	Spain	-3.23	-11.38	-4.68	-1.17	-4.33	2.24	-5.73	7.16	-6.75	-5.74	-8.15	-16.17	-2.33	6.60	-4.41	-7.23	-2.31	-12.47	0.00	8.16	0.59	-1.11	1.07	-2.32	-2.16	-0.82	1.85
	Turkey	23.94	16.00	20.55	22.43	18.65	26.42	18.79	30.90	19.10	20.42	12.44	10.07	25.62	25.68	20.45	22.21	21.33	14.92	26.96	0.00	23.80	14.80	25.86	13.78	22.87	18.70	25.64
	Australia	1.62	-2.05	1.69	3.88	2.08	6.68	1.18	10.07	0.30	-0.55	-0.37	-6.86	0.07	7.33	1.10	4.10	3.40	0.70	7.25	9.74	0.00	-0.69	4.96	4.40	2.46	2.73	5.70
	New Zealand	9.08	6.09	9.98	13.10	9.80	14.43	8.76	17.72	9.29	8.79	5.62	1.92	11.77	15.60	8.75	11.98	10.89	8.30	15.25	14.57	9.43	0.00	12.38	11.51	12.98	9.59	13.92
	Israel	11.34	8.01	11.02	13.35	10.41	15.94	9.92	18.34	10.15	8.84	8.86	1.85	11.09	16.79	10.04	12.31	12.81	9.24	16.26	19.24	13.87	12.22	0.00	9.85	14.75	15.17	18.35
	Korea	24.09	18.53	22.92	24.28	20.52	28.52	21.42	33.26	20.35	20.61	17.33	13.13	23.17	30.63	21.14	20.85	25.99	18.44	29.02	28.87	26.97	24.54	24.47	0.00	22.29	22.42	25.03
	Czech Rep.	20.45	13.33	17.93	20.17	15.42	25.18	18.89	29.44	17.32	15.39	14.23	9.25	18.25	24.97	19.06	16.56	20.45	13.16	24.59	28.80	20.64	21.46	24.09	17.46	0.00	14.61	18.24
	Hungary	20.47	14.36	18.93	21.41	16.94	24.99	19.26	29.04	17.55	16.95	13.91	8.13	19.84	26.42	18.87	17.79	20.78	18.44	26.57	27.12	22.07	18.60	25.45	19.07	16.39	0.00	18.68
	Poland	6.80	-0.39	5.78	7.92	3.61	12.94	5.96	18.11	4.18	3.62	2.11	-6.68	6.05	15.04	5.78	3.70	8.67	2.40	13.24	17.99	9.46	7.74	14.05	3.87	1.32	0.21	0.00

Figure 2: Heatmap of implied tax rates (in basis points per month). Rows refer to destination countries. Columns refer to origin countries. Darker colors indicate higher implied tax rates.

tax rates affecting capital flows between developed countries are rather minimal (a few basis points per month). The frictions become more substantial when capital flows from more to less developed markets (but are not very large when capital flows from less to more developed economies).

Table 1: The four columns report results of regressions of implied tax rates on a constant and various combinations of destination- and origin-country dummy variables.

	(1) tax	(2) tax	(3) tax	(4) tax
constant	10.12***			
	(27.62)			
Observations	702	702	702	702
R^2	0.521	0.732	0.247	0.968
Dest. Dummies	No	Yes	No	Yes
Orig. Dummies	No	No	Yes	Yes

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Continuing with our basic data explorations, we perform a k-means cluster analysis, so as to identify group of countries that exhibit similar tax rates. The results are reported in Figure 3. Specifically, to measure the similarity between the destination countries i and i' we use a modified Euclidean distance between the tax-rate vectors $\tau_{i\cdot}$ and $\tau_{i'\cdot}$.[27] Roughly speaking, we consider countries i and i' as "similar" when international investors face similar shadow tax rates when investing in either country.

Figure 3 plots results of the cluster analysis. We performed our cluster analysis with two to eight clusters. Our analysis favored four clusters.[28] Figure 3 reports these clusters, along with silhouette scores for each country. These scores are meant to measure how well a country fits inside a cluster. For instance, a silhouette score of one would mean that a country is identical to the rest of the countries inside its cluster, a value of zero indicates that a country is not particularly well

[27]Specifically, our distance measure can be expressed as $\sum_{j \notin \{i,i'\}} (\tau_{ij} - \tau_{i'j})^2 + (\tau_{ii'} - \bar{\tau}_i)^2 + (\tau_{i'i} - \bar{\tau}_{i'})^2$, where $\bar{\tau}_i$ is the average tax rate faced by foreigners in country i (and similarly for i'). We use this modified Euclidean measure since $\tau_{ii} = \tau_{i'i'} = 0$, by construction, and it seems more meaningful to compare the tax rate faced by investor i' in country i to the average tax rate faced by foreign investors in country i, rather than the tax rate faced by investor i whose shadow tax rate is zero by construction. Excluding the last two terms in the empirical analysis has a very small impact on the results.

[28]The analysis favored 4 clusters in terms of two criteria frequently used in the literature, namely producing relatively homogenous and well-separated clusters (i.e. high average "silhouette" scores within each cluster) and in terms of finding the number of cluster beyond which the incremental drop in the sum of absolute distances becomes relatively small.

matched inside the cluster, while negative values would indicate that a country probably does not belong to that given cluster.[29]

Inspecting the clusters, we find that one of the clusters includes fairly developed financial markets (e.g., USA, Japan, Euro countries such as Germany, etc.), a second cluster contains France, Italy, and Spain, a third cluster contains mostly non-Euro countries at the periphery of the Eurozone (e.g., Scandinavian countries), and a fourth cluster contains countries that one would arguably associate with less developed financial markets (Portugal, Greece, Turkey etc.). The second and fourth clusters are well separated from each other and from the rest of the clusters (high silhouette scores). The first and third clusters are less well separated from each other with some of the countries having low silhouette scores (below 0.5). Even though certain countries could still be misclassified within a given cluster (after all the cluster analysis does not remove the error in the measurement of τ_{ij}), overall the clustering shows patterns that coincide with the way the investment community separates financial markets in terms of their financial and economic development, and their economic ties. This suggests that our measure of frictions plausibly captures what it is supposed to.

4.3.2. Factors that correlate with implied tax rates: Regression evidence

We next investigate what factors tend to be associated with high or low tax rates. The goal is not to provide an exhaustive list of factors that explain τ_{ij}, nor to dissect the independent role played by every possible factor that could be affecting τ_{ij}. Such a task would be impossible with the small number of countries at our disposal. The regressions that follow are merely meant to highlight that our implied tax rates are plausible measures of frictions, in that they correlate with the sort of variables that one would expect.

We start by following the long tradition of gravity equations in international finance and regress the implicit tax rates on the logarithm of geographical distance between two countries. We also use a second measure of connections between two countries i and j, namely a categorical measure of Facebook friendships between countries i and j.

To ensure that our tax rates do not simply capture real-exchange-rate hedging motives we

[29] Silhouette scores are a standard diagnostic used in cluster analysis. A silhouette score is defined as $s(i) = \frac{b(i)-a(i)}{\max(a(i),b(i))}$, where $a(i)$ is the average distance of element i to other elements inside its cluster, while $b(i)$ is the lowest average distance of i to any other cluster of which i is not a member.

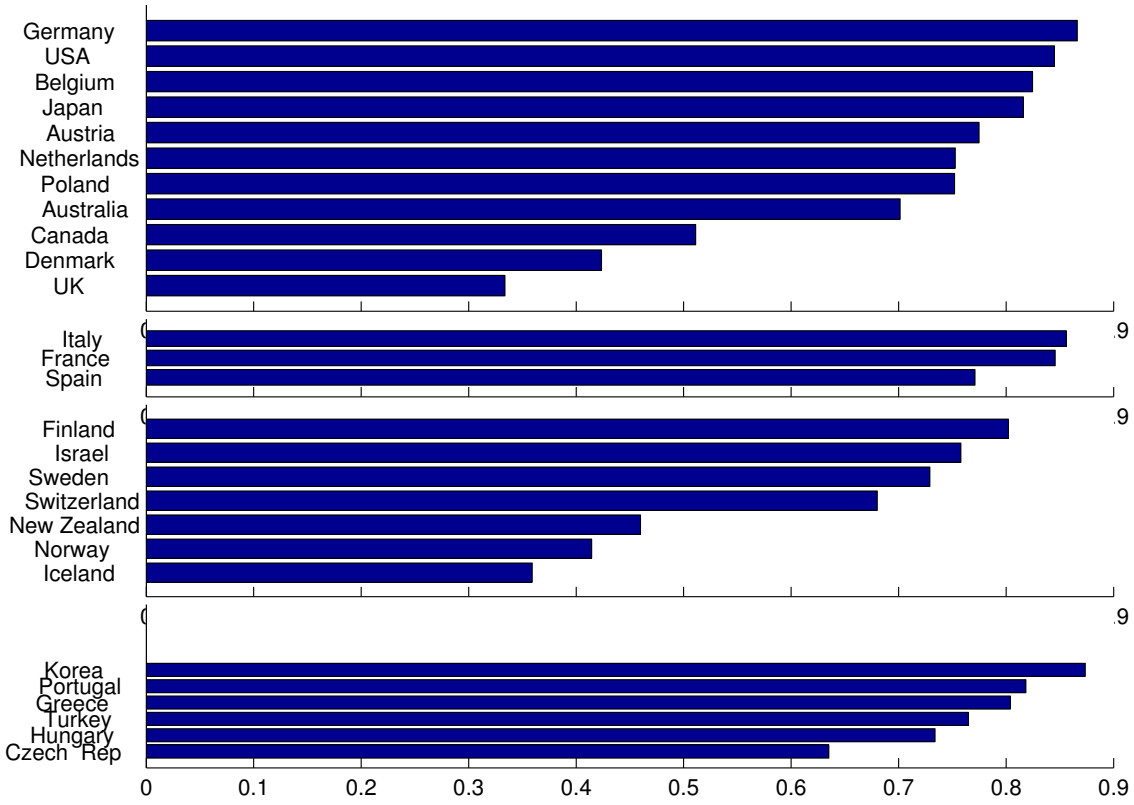


Figure 3: Results of cluster analysis along with sillhouette plots. Sillhouette values range from -1 to $+1$ and reflect how well a country fits into its cluster. A value of one reflects that a country fits perfectly in a cluster, a value of zero reflects that a country is not particularly well matched in the given cluster compared to some alternative cluster, and a value of negative one implies that a country probably does not belong in its cluster.

include a control for imports from country j as a fraction of country i's GDP. This measure is motivated by the literature on goods-market frictions,[30] which shows that "iceberg" transportation costs imply a wedge between the marginal utilities of foreign and domestic investors for the same good. These different marginal valuations translate into different marginal valuations for the firms that produce these goods, exactly as the taxes we assume here.[31] Obstfeld and Rogoff (2001)

[30]Dumas and Uppal (2001) develop a theoretical model with transportation costs; Obstfeld and Rogoff (2001) develop quantitative implications and show how a model with transportation costs can account for several puzzles of international macroeconomics.

[31]There is a formal link between the model in Obstfeld and Rogoff (2001) and our model. An elementary manipulation of the equations in Obstfeld and Rogoff (2001) implies that $D_H \frac{\partial U}{\partial C_H} = (1-\tau) D_H \frac{\partial U^*}{\partial C_H^*}$ where C_H is the

Table 2: The first four columns report regressions of implied tax rates on the logarithm of geographical distance between the countries (log_dist), a categorical measure of Facebook friendships between countries (fbook), and a measure of the share of goods imports from the origin country as a fraction of the GDP of the destination country (imp_share). Origin- and destination- country dummies are included in all except the last regression. Standard errors are heteroskedasticity-robust, clustered by recipient country.

	(1) tax	(2) tax	(3) tax	(4) tax	(5) tax	(6) tax	(7) tax	(8) tax
log_dist	0.775***			0.739***	1.138			0.296
	(5.92)			(4.81)	(1.54)			(0.36)
imp_share		-15.75		-1.233		-124.0*		-115.6*
		(-1.36)		(-0.14)		(-2.54)		(-2.16)
fbook			-0.230***	-0.0269			-0.934**	-0.108
			(-3.98)	(-0.31)			(-3.44)	(-0.24)
Observations	702	702	702	702	702	702	702	702
R^2	0.970	0.969	0.969	0.970	0.018	0.066	0.010	0.068
Dest. Dummies	Yes	Yes	Yes	Yes	No	No	No	No
Orig. Dummies	Yes	Yes	Yes	Yes	No	No	No	No

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

implies a direct and negative relationship between tax rates and import shares, in the sense that if the transportation cost from country j to i is large, then country i will import a smaller fraction of its imports from country j.

Table 2 shows that these bilateral measures have the sign that one would expect. (Tax rates are higher for relatively more distant countries, and taxes are lower for countries that have higher import shares and more Facebook friendships.) The statistical significance depends on whether origin and destination fixed effects are included. If they are, only distance is significant in a multivariate regression containing all variables. As one would expect from our earlier discussion, Table 2 shows that these bilateral variables have very limited ability to explain the variation in frictions (the R^2 is low). This is a consequence of the observation (in Table 1) that the bulk of the variation in tax rates is due to destination- (and to a lesser extent origin-) fixed effects.

In an attempt to understand the type of country characteristics that are correlated with the

consumption of the home good by a local, C_H^* is the consumption of the home good by a foreigner, D_H is the dividend of the home tree, and $(1-\tau)$ is the iceberg (transportation) cost. Accordingly, similar to this paper, the marginal valuation of locals and foreigners agrees on the valuation of the home tree only after applying a "tax rate" to the dividend of the home tree. Lane and Milesi-Ferretti (2005) generalize this model to multiple countries and show that import shares are decreasing in transportation costs.

severity of the frictions we turn our attention to Table 3. Table 3 reports results of regressions of implied tax rates on country characteristics. Viewing the financial industry as an input to lessen informational frictions, one would expect a larger financial industry to be correlated with a lower degree of frictions. Also motivated by Figure 3, one would expect GDP per capita to impact these frictions.

Table 3 provides evidence that both of these characteristics play a role. The regressions show that the value added in the financial industry of the destination country (as a fraction of its GDP) is associated with lower implied tax rates for the country. GDP per capita of the destination country also plays a significant role in lowering implied frictions, and so does the GDP per capita of the origin country. The last regression in Table 3 reports results when we include destination- and origin-country fixed effects along with interaction terms between origin-dummies and the size of the financial industry. We find that interaction terms between the GDP of two countries tend to play a role: Implied tax rates are lower when developed countries trade assets with each other, even after accounting for distance and destination- and origin-country dummies. In the last regression, we add interaction terms between the financial industry and the dummy variable of the origin country allowing for destination- and origin- fixed effects. The line labeled "D_originXFin_d" in the bottom part of the table reports the average value of these interaction terms. This value is negative, suggesting that a larger financial industry of the destination country is associated with lower tax rates for an origin country controlling for both origin- and destination- fixed effects. The p-value (line labeled "pval") performs a joint test that all the interaction terms between the financial industry and origin dummies are zero, and rejects that hypothesis.

The above results provide a flavor of the type of country characteristics that tend to be correlated with low frictions. Broadly speaking, we find that financial and economic development of the destination countries tend to be the factors that explain an important part of the variation in frictions — consistent with an informational interpretation.

4.3.3. Counterfactual experiments: The role of home bias

One could argue that the above results are due to some unobserved factor, which correlates with cross-country differences in home bias, and drives the differences in destination country fixed effects. To address this issue, we show next that the results of Table 3 are not simply due to cross-country

Table 3: The first three columns list results of regressions of implied tax rates (measured in basis points per month) on the destination (resp. origin) country's added value of the financial industry as a fraction of GDP (Fin_d, Fin_o), the logarithm of geographical distance between the countries (log_dist), goods import share (imp_share), GDP per capita of the destination, resp. origin country (GDP_d, GDP_o). The fourth regression includes origin country dummy variables. The next three regressions add destination- country fixed effects, and an interaction term of the GDPs per capita of the two countries (GDP_dXGDP_o). The last regression adds interaction terms between origin dummies and the size of the financial industry of the destination country. The third and second-last rows of the table report the average value of these interaction dummies along with the p-test value for the hypothesis that these interaction terms are jointly zero. We computed heteroskedasticity-robust standard errors and clustered three different ways (no clustering, by origin country, by destination country). We report the standard errors that were the most conservative for each regression, specifically: clustered by destination country for the first four regressions, clustered by origin country in the fifth and regular heteroskedasticity-robust standard errors in the last two regressions.

	(1) tax	(2) tax	(3) tax	(4) tax	(5) tax	(6) tax	(7) tax
Fin_d	-137.5*		-129.1*	-129.8*			
	(-2.53)		(-2.65)	(-2.61)			
Fin_o	10.17		18.29		25.75		
	(0.49)		(0.93)		(0.77)		
log_dist	0.957	0.300	0.859	0.915	0.955	0.683***	0.641***
	(1.41)	(0.43)	(1.33)	(0.96)	(1.23)	(5.16)	(4.36)
imp_share	-111.5*	-102.2*	-97.77*	-91.13*	-13.79	-1.618	-1.948
	(-2.47)	(-2.16)	(-2.42)	(-2.25)	(-0.66)	(-0.32)	(-0.39)
GDP_d		-0.0926	-0.0910*	-0.0916*			
		(-1.84)	(-2.05)	(-2.03)			
GDP_o		-0.0896***	-0.0891***		-0.0878**		
		(-19.04)	(-19.78)		(-2.90)		
GDP_dXGDP_o						-0.000569***	-0.000579***
						(-4.93)	(-5.21)
Observations	702	702	702	702	702	702	702
R^2	0.106	0.194	0.229	0.400	0.810	0.972	0.974
D_origXFin_d							-1.765
pval							0.000
Dest. Dummies	No	No	No	No	Yes	Yes	Yes
Orig. Dummies	No	No	No	Yes	No	Yes	Yes

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

differences in the extent of home bias (i.e., the diagonal elements of the portfolio matrix W).

Specifically, we consider a fictitious world in which a fraction δ_i of the population in every country i decides to hold only its local stock market — for an unmodeled reason (behavioral, hedging motives, etc.) — and not participate in any other market. Moreover, in this fictitious world, the rest of the population chooses to allocate its funds internationally, and subject to no frictions ($\tau_{ij} = 0$). We impose market clearing and solve for equilibrium portfolios in such a fictitious world. Importantly, we reverse-engineer the fraction δ_i in each country so that the *total* holdings of local stocks by local investors are exactly the same as in the data — i.e., we keep the diagonal elements of the new weight matrix, W^*, the same as W. However, we determine the off-diagonal elements of W according to what they would be in the frictionless world given by $\tau_{ij} = 0$.

With this modified matrix of portfolios and keeping everything else unchanged (moments of returns, assumptions on the observation errors, etc.), we infer the tax rates τ^*_{ij} that would result in such a counterfactual world. The left plot of Figure 4 provides a scatterplot of our implied tax rates τ_{ij} (using the actual matrix of portfolios W) against the tax rates τ^*_{ij} (implied by the counterfactual matrix W^*). The scatter diagram shows that the implied tax rates are substantially different. The middle and right diagram provide similar scatter plots for the estimated destination-country dummy variables and the estimated origin-country dummy variables. As the plot shows, there is some, but far from perfect, correlation between the destination-country dummy variables in the actual and counterfactual data. This implies that our tax rates are not mere reflections of cross-country differences in home bias.

To test formally whether the results in Table 3 are simply a manifestation of patterns of home bias, Table 4 tests the hypothesis $\tau_{ij} = \tau^*_{ij}$ by regressing $\Delta\tau_{ij} = \tau_{ij} - \tau^*_{ij}$ on τ^*_{ij} and the same regressors as in Table 3. Under the null hypothesis that $\tau_{ij} = \tau^*_{ij}$ (so that $\Delta\tau_{ij} = 0$), all the regressor-coefficients in Table 4 should not be significantly different from zero, including the coefficient on τ^*_{ij}. The reason is that any deviation of $\Delta\tau_{ij}$ from zero should be due to observation error in portfolios, return moments etc. Table 4 shows that the hypothesis $\Delta\tau_{ij} = 0$ can be rejected. The coefficient on τ^*_{ij} is different from zero in all specifications. Moreover, several of the variables in Table 3 remain significant after including τ^*_{ij} as a regressor. Comparing the magnitudes of the resulting coefficients, we find that the coefficient of the financial industry in column 2 is lower by about a third compared to the respective coefficient of Table 3, and remains significant. The import share, log distance,

Table 4: This table repeats some of the regressions of Table 3 except that the dependent variable is the difference between our implied tax rates and the tax rates that would result in a counterfactual world with unchanged portfolio allocations of domestic residents to domestic risky assets, but frictionless capital allocation for the fraction of capital that is allocated internationally. The counterfactual tax rate is included as one of the regressors ("tax2"). The rest of the variables are described in Table 3.

	(1) dtax	(2) dtax	(3) dtax	(4) dtax	(5) dtax	(6) dtax	(7) dtax
tax2	-0.688***	-0.719***	-0.645***	-0.661***	-0.654**	-0.397***	-0.386***
	(-7.78)	(-8.55)	(-5.31)	(-5.35)	(-3.46)	(-10.33)	(-9.80)
Fin_d		-91.69*		-79.85			
		(-2.00)		(-1.63)			
Fin_o		8.139		15.70	23.14		
		(0.46)		(0.94)	(0.65)		
log_dist		0.762	0.202	0.530	0.502	0.293**	0.235*
		(1.31)	(0.35)	(0.98)	(0.62)	(3.22)	(2.32)
imp_share		-69.21*	-76.06*	-74.80*	-18.72	0.0447	0.111
		(-2.03)	(-2.00)	(-2.17)	(-0.88)	(0.02)	(0.04)
GDP_d			0.0415	0.0363			
			(0.81)	(0.71)			
GDP_o			-0.102***	-0.101***	-0.101**		
			(-14.55)	(-14.44)	(-3.15)		
GDP_dXGDP_o						-0.0000345	-0.0000416
						(-0.36)	(-0.49)
Observations	702	702	702	702	702	702	702
R^2	0.633	0.655	0.691	0.697	0.910	0.992	0.993
D_origXFin_d							-10.63
pval							0.000
Dest. Dummies	No	No	No	No	Yes	Yes	Yes
Orig. Dummies	No	No	No	No	No	Yes	Yes

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

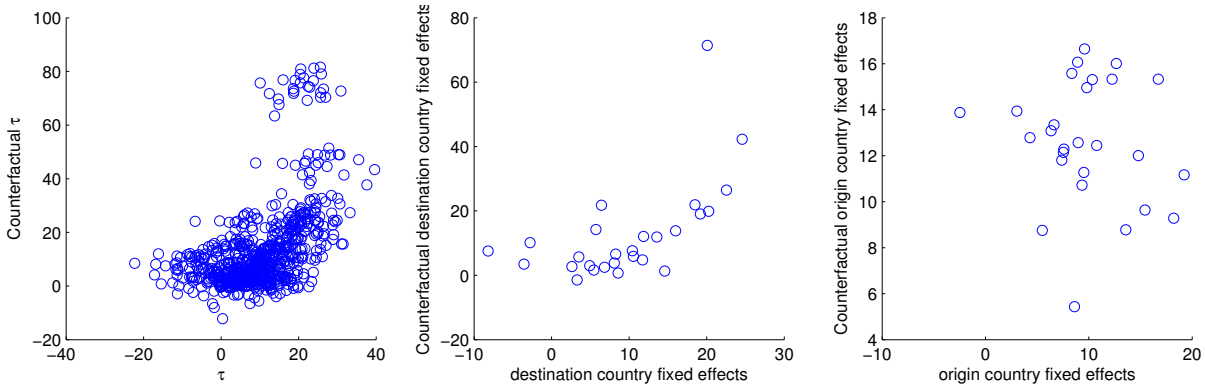


Figure 4: Scatter-plot of implied tax rates versus counter-factual implied tax rates. The counterfactual implied tax rates are computed by assuming that domestic investors' portfolio allocation to domestic risky asset is unchanged. However, the fraction of the portfolio that is invested internationally is invested as if there were no frictions. The left plot depicts all N^2 implied tax rates. The middle plot depicts the estimated destination-country fixed-effects for actual and counterfactual implied tax rates. The right plot depicts the estimated origin-country fixed-effects for actual and counterfactual implied tax rates.

and the GDP of the origin country also remain significant in several of the specifications. Hence, we can reject the hypothesis that $\tau_{ij} = \tau^*_{ij}$.

The fact that the off-diagonal elements of the portfolio matrix W play an important role for our results can be illustrated in an even simpler way. Figure 5 performs the following experiment: Keeping the diagonal elements of W fixed (i.e., the domestic allocations to the domestic asset), we perform a random reshuffling of the elements contained in the international portfolio of each country. We repeat this exercise 1000 times obtaining 1000 artificial matrices W, compute the resulting artificial values τ^*_{ij} (keeping all other inputs the same), and then regress the difference between the actual tax rates τ_{ij} and the counterfactual ones τ^*_{ij} on τ^*_{ij}, the size of the financial sector, the GDP of the destination country, log distance, and the import share. Figure 5 plots a histogram of the regression coefficients. The graph shows that the depicted coefficients are different from zero in almost all samples, and indeed statistically different from zero.[32] This implies that the off-diagonal elements of W matter for the results in Table 3.

[32]We computed the 0.025–0.975 coverage interval for all the depicted coefficients and zero was not in that interval for all of the depicted coefficients.

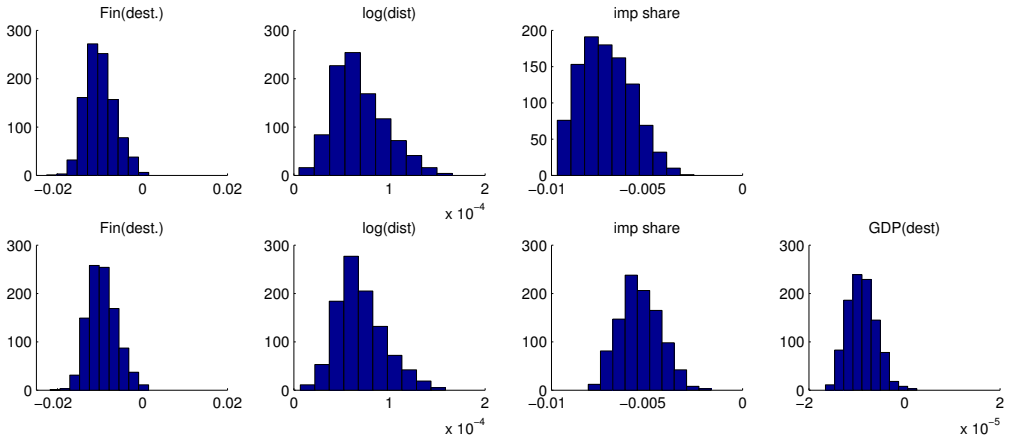


Figure 5: Histograms of regression coefficients from two regressions performed on randomly generated data. The artificial data samples are counterfactual shadow tax rates computed as follows: keeping the diagonal elements of the international portfolio of each country fixed, we re-shuffle the elements of the international portfolio holdings of each country in a random manner. We create 1000 artificial portfolio matrices and compute 1000 artificial tax rates. For each draw, we regress the difference between the tax rates we obtain from actual data (τ_{ij}) and the artificial tax rates corresponding to the random draw (τ^*_{ij}) on τ^*_{ij}, the size of the financial industry of both destination and origin country, log distance between the two countries, and the import shares as described in the text. The histograms for the coefficients on the destination country's industry, log distance, and import shares are given by the top three histograms. The lower histograms report results of the same regression, but including the GDP of destination and origin country as additional regressors.

4.3.4. The role of the financial industry

Table 3 suggests that the size of the financial industry of the destination country covaries inversely with the implied tax rates of the destination country. The table is silent about whether this is the result of direct causality (the financial industry helps reduce the frictions for incoming portfolio flows) or of reverse causality (countries that, for whatever reasons, have lower frictions attract more capital flows and thus need a larger financial industry to process the transactions). We investigate this issue in Table 5 by re-estimating some of the key regressions of Table 3 using an instrumental-variables approach. We use private domestic credit (as a fraction of GDP) and value added in the real-estate-finance sector as a fraction of GDP as instruments for the size of the financial sector. Arguably, since these two instruments capture the operation of the domestic finance sector, their size is less likely to be the mechanical result of international capital flows. Table 5 shows that when we use instrumental variables, the significance of our results remains the same.

We conclude with a remark on the role of the financial industry: whether lower shadow tax rates are caused by or cause a larger financial industry is of secondary importance for our purposes.

Table 5: Two-stage least squares instrumental variables regression. We use private domestic debt as a fraction of GDP and real estate finance as a fraction of GDP in the destination country as instruments for the size of the overall financial sector as a fraction of GDP in the destination country (Fin_d). The rest of the variables are described in table 3. Standard errors are heteroskedasticity robust, clustered by destination country. Results are essentially the same whether we estimate the regressions with two-stage least squares or GMM. In the latter case the test of overidentifying restrictions does not reject.

	(1) tax	(2) tax	(3) tax	(4) tax	(5) tax	(6) tax
Fin_d	-701.3*	-624.3*	-502.7*	-775.2*	-703.3*	-561.3*
	(-2.52)	(-2.30)	(-2.00)	(-2.38)	(-2.13)	(-1.97)
log_dist	4.472*	3.475	3.469	6.679	5.558	4.956
	(2.03)	(1.58)	(1.87)	(1.75)	(1.38)	(1.53)
imp_share		-89.52*			-69.23	
		(-2.34)			(-1.33)	
GDP_d			-0.0910*			-0.0919*
			(-2.19)			(-2.18)
Observations	702	702	702	702	702	702
Dest. Dummies	No	No	No	No	No	No
Orig. Dummies	No	No	No	Yes	Yes	Yes

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Either way, a larger financial industry seems useful: either because it directly lowers shadow tax rates, or because it is necessary to process the international transactions that result from lower shadow taxes, thus allowing a country to benefit from its inherently higher appeal.

4.4. Source of identification and robustness of implied tax rates

In this section we start by showing that the implied tax rates τ_{ij} are quite robust to the measurement of expected returns, which are notoriously hard to measure. Furthermore, different assumptions on risk aversion act essentially as a multiplicative scalar for the matrix τ_{ij}. After illustrating these statements, we provide an approximate, but intuitive expression for τ_{ij} that helps explain these statements, and more generally helps recognize the sources of identification of τ_{ij}.

We start with a graph that illustrates that our findings do not depend critically on assumptions or estimates of expected returns. Specifically, the left panel of Figure 6 shows a scatterplot of our implied tax rates plotted against the tax rate that would result if we used the constant interest rate R instead of our estimates of expected returns μ^S (i.e., the averages of historical returns).

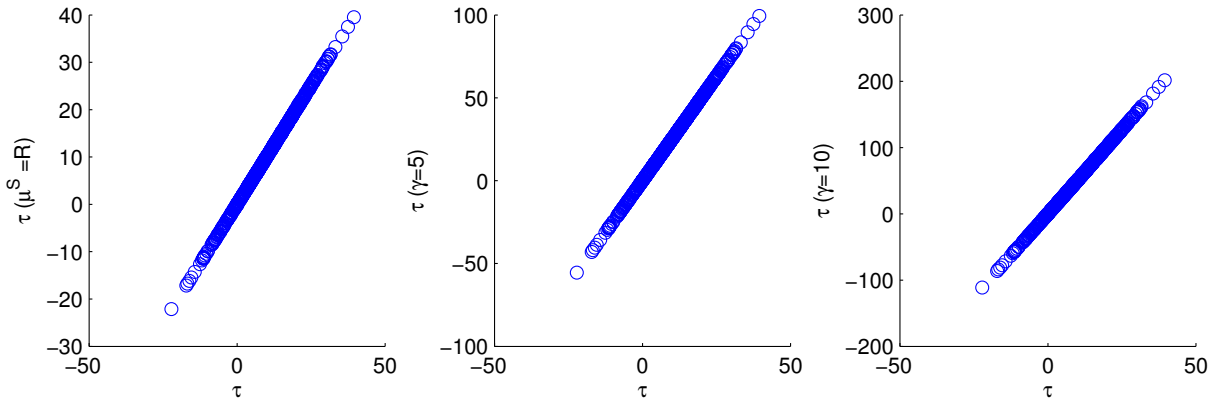


Figure 6: Left plot: Scatterplot of our implied tax rates (x-axis) versus implied tax rates imposing $\mu^S = R$ (y-axis). Middle and right plots: Scatterplots of our implied tax rates (x-axis) versus implied tax rates resulting when $\gamma = 5$, respectively $\gamma = 10$ (y=axis).

The scatterplot shows that the obtained estimates τ_{ij} remain essentially unchanged, showing that assumptions on expected returns do not affect τ_{ij}. (Indeed one would obtain essentially the same tax rates, as long as the assumption on gross returns is that they are not too far from one).

A related finding is the impact of risk aversion on τ_{ij}. The middle and right panels of Figure 6 illustrate the point. As risk aversion changes, the quantities τ_{ij} remain unaltered, up to scaling.

To understand these patterns, it is useful to perform a "back of the envelope" exercise by revisiting (22). Multiplying both sides by $\frac{\gamma\Sigma}{R}$ and evaluating the equation for two different investors i and j gives

$$\frac{\gamma\Sigma}{R}\left(\Pi_j^{-1}w_j - \Pi_i^{-1}w_i\right) = \left(\Pi_i - \Pi_j\right)e_{N\times 1}.$$

Since both Π_i^{-1} and Π_j^{-1} are close to one and $\frac{\gamma\Sigma}{R}$ is small, we approximate the left-hand side as $\frac{\gamma\Sigma}{R}\left(w_j - w_i\right)$. Focusing on element i in the above equation and approximating $\tau_{ij} \approx \pi_{ij} - \pi_{ii}$ gives

$$\tau_{ij} \approx \frac{\gamma}{R}\left[\Sigma\left(w_i - w_j\right)\right]_i, \tag{33}$$

where the notation $[x]_i$ refers to entry i in the vector x. Assuming momentarily that this basic approximation is accurate, it helps explain two things. First, it helps provide a reason why μ^S does not affect the computation of τ_{ij}. And, second, it shows that γ acts (approximately) as a multiplicative constant on all τ_{ij}.

A benefit of expression (33) is that it provides a more intuitive understanding of τ_{ij}. Specifically,

τ_{ij} can be understood as a difference in the marginal valuation of a given asset by foreign and local investors. Indeed, the right hand side of (33) is the difference between the covariance of asset i with (the local) investor's i overall portfolio and the respective covariance of asset i with investor j's portfolio. The difference between these covariances should be equal to the (shadow) difference in the expected returns perceived by the two investors.[33] In the special case where Σ is the identity matrix (possibly multiplied by a scalar), the implied tax rates reflect only the extent that a local investor invests in the local asset as compared to the respective allocation of foreign investors. If variances are unequal, and more importantly, if the covariances differ from zero, then the τ_{ij} will not only reflect properties of portfolios, but also properties of second moments of returns.

The approximation (33) also helps illustrate which type of measurement error affects our estimates of τ_{ij} primarily, namely measurement error in portfolios and covariances. In particular, if there is measurement error in estimated covariances, it is possible that the estimated $\tau_{ij} < 0$ for some countries, even if locals allocate more to their own country than foreigners. (We note parenthetically that if one were willing to parameterize the covariance matrix in certain ways, one could easily ensure that $\tau_{ij} > 0$.[34]) Additionally, equation (33) shows that mis-measurement of a country's portfolio only impacts our results to the extent that it affects the covariance of the country's overall portfolio with the covariance of individual assets. For instance, incomplete coverage of countries that are unlikely to change the covariance properties of an investor's overall portfolio does not materially affect our results.

The simplicity of expression (33) makes it tempting to examine whether one could infer τ_{ij} by using (33) directly, rather than using the more cumbersome expression (31). It turns out that (33) is only in partial agreement with our obtained tax rates: the correlation coefficient between the two sets of rates is 0.56. The correlation, however, becomes virtually one if, rather than using the actual portfolios w_i and w_j on the right hand side of (33), we use instead the predicted portfolio values $\widehat{w_i}$ and $\widehat{w_j}$ that result from $vec(\widehat{W}) = A\, vec(\widehat{\Pi})$, where $vec(\widehat{\Pi})$ is our estimate of (33).[35]

The reason why (33) and (31) give different result is simple: Via (31), we intentionally over-

[33] We use the word 'shadow' to allow for a Lagrange multiplier on the requirement $w_{ji} \geq 0$ that these two investors face when investing in asset class i.

[34] For instance if one postulated that the off-diagonal elements of Σ are all equal, and the diagonal elements are greater than the off-diagonals, then $w_{ii} > w_{ij}$ implies $\tau_{ij} > 0$. This would be an appropriate assumption in a one-factor world, where all countries had the same exposure to that factor.

[35] The correlation coefficient between our measure of τ_{ij} and the approximate expression $\frac{\gamma}{R}\left[\Sigma\left(\widehat{w_i} - \widehat{w_j}\right)\right]_i$ is essentially one.

identify the model so as to allow our estimation procedure to lower the weight given to portfolios that are likely to be the result of measurement error. Besides "filtering" out portfolio noise, however, expected returns do not affect τ_{ij}, as equation (33) shows. This (indirect) dependence of τ_{ij} on μ^S is the reason why our findings are numerically insensitive to different assumption on μ^S.

4.5. Interpretation and practical uses of the implied tax rates

An insight from the previous section is that the implied tax rates τ_{ij} capture valuation discrepancies between different investors. The theory section of this paper suggested one possible theoretical underpinning for the presence of such shadow tax rates. However, we are quite open to the possibility that these discrepancies may be the result of other frictions, or may even reflect behavioral misperceptions and fears related to unfamiliarity.

Regardless of the exact friction that leads to these valuation discrepancies, we believe that they have several practical uses.

The first and most obvious use is to provide a direct measure of "bottlenecks" of financial flows, i.e., help identify the directions of financial trade that seem particularly impeded. As such our measure of financial frictions can be used in various contexts both within and outside financial economics (say, as an alternative to distance in gravity equations in economics).

The second usage is as a way to diagnose the directions (and likely reasons) for failures of asset-pricing models. Most existing empirical asset-pricing approaches rely on comparing discrepancies between average and the expected returns implied by some model. Our implied taxes capture a different aspect, namely the discrepancy in the valuation of the same asset by different investors. This helps paint a complementary picture (and also suggest a direction for the likely failure of asset-pricing models). The reason is that most theories of frictions have implications not only for equilibrium expected returns, but also for valuation discrepancies between investors.

If one were to adopt a more behavioral view of our measured frictions, one could envisage a practical use for investment purposes. To give an extreme example, suppose that someone viewed these implied tax rates as resulting from irrational non-participation decisions, not informational disadvantages. Then the countries with the largest measured shadow tax rates would be good candidates for investing, since they are irrationally "cut off" from markets.

5. Conclusion

If markets are not frictionless, then frictions can act in a manner similar to distortionary taxes. We illustrate this analogy explicitly in the context of a model where frictions are caused by asymmetric information. More generally, financial frictions are similar to taxes in that they drive a wedge in the valuation of the same asset by different investors. This leads to different tilts in optimal investor portfolios.

Accordingly, the optimal holdings of investors contain valuable information in terms of inferring these implicit taxes. We employ a standard mean-variance framework (appropriately modified to allow for fictions) to infer a set of taxes capable to reproduce observed patters of portfolio holdings and equilibrium returns. These taxes provide a novel view of the direction and extent to which financial linkages seem to be particularly impeded: Consistent with commonly held views, tax rates are very small when developed countries trade with each other. However, these implicit taxes become non-trivial, when the capital flows are directed towards lesser developed financial markets. This helps us identify the location of financial "bottlenecks", i.e. directions where financial trade is particularly impeded. More importantly, it allows us to quantify (in units of expected returns) the relative importance of different directions of deviation from the frictionless benchmark.

A finding of our analysis is that tax rates manifest themselves mostly in the form of country-specific destination- and to a lesser extent origin-fixed effects. This means that certain countries are high-tax-rate destinations, no matter where the financial flows originate. This implies that bidirectional variables that are immune to the permutation of origin and destination country (e.g., geographical distance, common legal origin, common language or religion etc.) cannot explain but a very small fraction of the variation of the tax rates we observe in the data. The sort of variables that seem to perform well in terms of explaining the variation in tax rates are predominantly destination-country characteristics, indicatively the size of the financial industry of the destination country and the overall level of economic development.

References

Adler, M. and B. Dumas (1983). International portfolio choice and corporation finance: A synthesis. *Journal of Finance 38*, 925–984.

Admati, A. R. (1985). A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica 53*(3), 629–658.

Admati, A. R. and S. A. Ross (1985). Measuring investment performance in a rational expectations equilibrium model. *Journal of Business 58*(1), 1–26.

Aviat, A. and N. Coeurdacier (2007). The geography of trade in goods and asset holdings. *Journal of International Economics 71*, 22–51.

Beck, T., A. Demirgüç-Kunt, and R. Levine (2010). Financial institutions and markets across countries and over time: The updated financial development and structure database. *World Bank Economic Review 24*(1), 77–92.

Bekaert, G. and X. Wang (2009). Home bias revisited. Working Paper, Columbia Business School.

Berkel, B. (2007). Institutional determinants of international equity portfolios - a country-level analysis. *The B.E. Journal of Macroeconomics 7*(1), article 34.

Black, F. (1974). International capital market equilibrium with investment barriers. *Journal of Financial Economics 1*(4), 337 – 352.

Brennan, M. J. and H. Cao (1997). International portfolio investment flows. *Journal of Finance 52*, 1851–1880.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2007). Business cycle accounting. *Econometrica 75*(3), 781–836.

Cochrane, J. (2005). *Asset Pricing*. Princeton University Press, Princeton and Oxford.

Coeurdacier, N. and H. Rey (2013). Home bias in open economy financial macroeconomics. *Journal of Economic Literature 51*(1), 63 – 115.

Cooper, I. and E. Kaplanis (1986). *Costs to Crossborder Investment and International Equity Market Equilibrium.*, pp. 209 – 240. Cambridge University Press, Cambridge.

Cooper, I. and E. Kaplanis (1994). Home bias in equity portfolios, inflation hedging, and international capital market equilibrium. *Review of Financial Studies 7*(1), 45–60.

Daude, C. and M. Fratzscher (2008). The pecking order of cross-border investment. *Journal of International Economics 74*, 94–119.

Dumas, B. and R. Uppal (2001). Global diversfication, growth and welfare with imperfect markets for goods. *The Review of Financial Studies 14*(1), 227–305.

Dybvig, P. H. and S. A. Ross (1985). Differential information and performace measurement using a security market line. *Journal of Finance 40*(2), 383–399.

Faruqee, H., S. Li, and I. K. Yan (2004). The determinants of international portfolio holdings and home bias. Working Paper.

Gehrig, T. (1993). An information based explanation of the domestic bias in international equity investment. *Scandinavian Journal of Economics 95*, 97–109.

Glode, V. (2011). Why mutual funds "underperform". *Journal of Financial Economics 99*, 546–559.

Hatchondo, J. C. (2008). Asymmetric information and the lack of portfolio diversification. *International Economic Review 49*(4), 1297–1330.

Koijen, R. and M. Yogo (2014). An equilibrium model of institutional demand and asset prices. Working Paper, London Business School and Minneapolis FED.

Kyle, A. S. (1989). Informed speculation with imperfect competition. *The Review of Economic Studies 56*(3), 317–355.

Lane, P. R. and G. M. Milesi-Ferretti (2005). International investment patterns. *Review of Economics and Statistics 90*(3), 538–549.

Mayers, D. and E. Rice (1979). Measuring portfolio performance and the empirical content of asset pricing models. *Journal of Financial Economics 7*, 3–28.

Moskowitz, T. (2000). Mutual fund performance: An empirical decomposition into stock-picking talent, style,transactions costs, and expenses: Discussion. *The Journal of Finance 55*(4), 1695–1703.

Obstfeld, M. and K. Rogoff (2001). The six major puzzles in international macroeconomics: Is there a common cause? *NBER Macroeconomics Annual 15*, 339–390.

Okawa, Y. and E. van Wincoop (2012). Gravity in international finance. *Journal of International Economics 87*, 205–215.

Pavlova, A. and R. Rigobon (2008). The role of portfolio constraints in the international propagation of shocks. *Review of Economic Studies 75*, 1215–1256.

Portes, R. and H. Rey (2005). The determinants of cross-border equity flows. *Journal of International Economics 65*(2), 269–296.

Sercu, P. (1980). A generalization of the international asset pricing model. *Revue de l'Association Francaise de Finance 1*(1), 91–135.

Sharpe, W. F. (1974). Imputing expected security returns from portfolio composition. *The Journal of Financial and Quantitative Analysis 9*(3), pp. 463–472.

Sharpe, W. F. (1992). Asset allocation: Management style and performance measurement. *Journal of Portfolio Management 18*(2), 7–19.

Stulz, R. M. (1981). On the effects of barriers to international investment. *Journal of Finance 36*(4), 923 – 934.

Sun, Z., A. Wang, and L. Zheng (2009). Do active funds perform better in down markets? - new evidence from a cross-sectional study. Working paper, U C Irvine.

Van Nieuwerburgh, S. and L. Veldkamp (2010). Information acquisition and under-diversification. *Review of Economic Studies 77*(2), 779 – 805.

Appendix

A Strategic Agents and Dividend Manipulation

Here we build a model extension designed to capture two desired phenomena. First, we model a finite economy populated by agents who behave strategically and show that the equilibrium approaches the one in the benchmark (continuum) economy as the number of agents grows without bound. Second, we also obtain no shorting as an endogenous consequence of allowing swindlers to manipulate the cash flows from their firms, in a sense made precise below.

In order to illustrate the point as quickly and easily to convey as possible, we make a number of (dispensable) simplifying assumptions.

Consider two locations, $\mathcal{L} = \{1, 2\}$, each populated by N agents and hosting N firms. A number κN of the agents[36] in each location are common, the others being swindlers. Similarly, there are κN regular and $N - \kappa N$ fraudulent firms per location. Agents have CARA preferences with risk-aversion parameter γN, and each regular firm in location j has output and dividend D_j/N, where $\mathrm{E}[D_j] = 1$ and $\mathrm{Var}(D_j) = \sigma^2$. Let Ω denote the variance of (D_1, D_2). Dividends in different locations are assumed to be independent. Fraudulent firms have output equal to zero. All agents are endowed with an equal number of shares of regular firms in their locations. Each swindler also owns entirely a fraudulent firm.

The information structure is as in the main text, with some simplifications. As in the text, we maintain that $p_{ii} = 1$, and also impose symmetry, i.e., $p_{12} = p_{21} \equiv p$. We also impose that some quantities, such as the proportion of fraudulent firms mis-identified as regular equal their ex-ante averages. More precisely, agent 1 receives $p^{-1}\kappa N$ good signals for the firms in location 2, of which exactly κN correspond to the regular firms and the remainder to fraudulent firms. The set of $(p^{-1} - 1)\kappa N$ of mis-identified firms is chosen from a uniform distribution on the set of cardinality-$(p^{-1} - 1)\kappa N$ subsets of the set of fraudulent firms.

The action space for common investors consists of demand functions $X(\bar{P})$ that give the numbers of shares in the $2N$ securities that a given investor is willing to purchase given the $2N$-dimensional price vector $\bar{P}$. Swindlers must take an additional action, which is the amount L that they borrow

[36] The number of agents must be an integer, of course. We therefore adopt the convention that all necessary quantities are rounded in some reasonable fashion. Alternatively, we restrict κ to be rational and N to an appropriate (unbounded) set. The same for $p^{-1}\kappa N$.

and divert into the firm to increase its liquidation value.

Specifically, we assume that each swindler has the ability to borrow any amount L of her choosing at time 0, divert these funds into the firm, and report earnings equal to $L(1+r) = L$ in period 1. (Equivalently, we could assume that the swindler can take an action to produce earnings L by incurring a personal non-pecuniary cost of effort, which would have a value L in monetary terms.) Given the possibility of such a diversion, equation (5), giving the swindler's time-1 wealth, becomes

$$W_1^{sil} \equiv B^{sil} + \int_{j\in\mathcal{L}} \int_{k\in[0,1]} D_{jk} dX_{jk}^{sil} + L^{il}\left(S^{il} - 1\right). \tag{A.1}$$

We note that the difference to (5) is the term $L^{il}\left(S^{il}-1\right)$, which is intuitive. If $S^{il}-1<0$, i.e., if the swindler reduces her ownership of shares by being a net seller, then she has no incentive to perform earnings diversion, since she will recover only a fraction of the funds she diverted into the company. If, however, the swindler is a net buyer of her own security $\left(S^{il}-1>0\right)$, then the ability to manipulate earnings becomes infinitely valuable, since L^{il} can be chosen to be an arbitrarily large number. Intuitively, the swindler can report arbitrarily large profits at the expense of outside investors who hold negative positions (short sellers) in the fraudulent firm. This feature discourages any other agent from shorting: with non-zero probability all other agents know that the firm is fraudulent and don't buy any shares, so that any shorting results in $S^{il}>1$.

Given that we are considering a sequential game — the swindler's decision to manipulate is taken after the asset market clears — of incomplete information, we are looking for a *perfect Bayesian Nash* equilibrium. Loosely speaking, this concept requires that all actions — demands and manipulation decisions — be optimal given beliefs, while beliefs be updated according to Bayes' rule wherever possible. Note that, unlike in the main body of the paper, all agents take into account their potential impact on the price, and on the other agents' beliefs, when submitting their demand functions.

We concentrate on the sub-class of symmetric equilibria, in which all agents in a given market behave identically conditional on their type and signals, while the differences in behavior between market-1 and market-2 agents come down to index permutations in the natural way. Important, we are interested in the existence of pooling equilibria, in which all securities in a given location, and therefore in the entire economy by symmetry, have the same price.

Proposition 7 *A perfect Bayesian Nash equilibrium exists with the following properties.*

(i) The equilibrium is symmetric.

(ii) All security prices are equal.

(iii) Common investors and swindlers have the same portfolios in equilibrium, with the exception of the swindler's holding of her own firm.

(iv) There is no shorting.

(v) There is no dividend manipulation.

Furthermore, as N increases, equilibrium prices and aggregate holdings of agents in any location i of all assets in location j converge to the competitive-strategic equilibrium in Proposition 1.

B Proofs

Proof of Proposition 1. The proof proceeds in a number of steps. We start with an equilibrium in the simplified competitive tax economy, and use it to construct demands in the original economy. Second, we specify out-of-equilibrium beliefs in the original economy that support the equilibrium. In a third step, we verify that all agents, regular as well as swindlers, find it optimal to submit the demands specified given prices and their beliefs. Finally, we verify that markets clear.

Consider a solution to the simplified problem (8)–(9). The demands in the original economy are defined naturally based on this solution:

$$dX_{jk}^{ci} = (1 - f_{ij})\,\kappa^{-1} dX_j^i \mathbf{1}_{\iota_{jk}^i} \mathbf{1}_{(P_{jk}=P_j)} \tag{B.1}$$

$$dX_{jk}^{sil} = (1 - f_{ij})\,\kappa^{-1} dX_j^i \mathbf{1}_{\iota_{jk}^i} \mathbf{1}_{(P_{jk}=P_j)}, \quad (i,l) \neq (j,k) \tag{B.2}$$

$$dX_{il}^{sil} = \begin{cases} [0,\infty) & \text{if} \quad P_{il} = P_i \\ 0 & \text{if} \quad P_{il} \neq P_i \end{cases}. \tag{B.3}$$

The conjectured prices are $P_{jk} = P_j$ for all j and k. Note that dX_{jk}^{ci} and dX_{jk}^{sil} are themselves demand curves, i.e., functions of the prices $\{P_j\}_j$.

In words, all investors buy the same number of shares in each market as in the tax economy, but they split this position (equally) only among the firms about which they receive a good signal — note that the multiplicative factor $(1 - f_{ij})\,\kappa^{-1}$ equals the reciprocal of the probability that a

given signal is good. Another proviso is that the price equal the pooling price P_j; for any other price, the agents shun the asset. The only exception to this behavior is provided by the insiders of fraudulent firms, who submit an elastic demand at $P_{jl} = P_j$.

In equilibrium, only prices P_j are realized, and therefore prices are not informative. We postulate that all agents believe that any firm k in market j that has price $P_{jk} \neq P_j$ is fraudulent with probability one.

To see that dX^{ci}_{jk} is optimal, start by writing the expected utility for the agent as

$$\mathrm{E}\left[U\left(\int_j\int_k \left(D_{jk}-P_j\right)dX^{ci}_{jk}+P_i\right)|\iota^i\right]=\mathrm{E}\left[U\left(\int_j\int_k \left(\rho_{(jk)}D_j-P_j\right)dX^{ci}_{jk}+P_i\right)|\iota^i\right] \quad \text{(B.4)}$$

and note that, by Jensen's inequality, this utility is maximized by choosing dX^{ci}_{jk}, for fixed j, to be measurable with respect to ι^i_{jk} — in words the agent invests identically in all assets in market j in which she received the same signal. Furthermore, the portfolio of assets with higher signals ($\iota^i_{jk}=1$) strictly dominates the portfolio with low signals ($\iota^i_{jk}=0$). Let $d\hat{X}^{ci}_j$ denote the number of shares in each asset in market j in which the investor has a positive signal. Consequently, (B.4) is equal to

$$\begin{aligned}&\mathrm{E}\left[U\left(\int_j\int_k \left(\rho_{(jk)}D_j-P_j\right)\mathbf{1}_{\left(\iota^i_{jk}=1\right)}dkd\hat{X}^{ci}_j+P_i\right)|\iota^i\right] \quad \text{(B.5)}\\ &=\ \mathrm{E}\left[U\left(\int_j\left((1-f_{ij})D_j-P_j\right)Pr\left(\iota^i_{jk}=1\right)d\hat{X}^{ci}_j+P_i\right)\right].\end{aligned}$$

It follows that the optimal position is

$$d\hat{X}^{ci}_j=Pr\left(\iota^i_{jk}=1\right)^{-1}dX^{ci}_j=(1-f_{ij})^{-1}\kappa\, dX^{ci}_j. \quad \text{(B.6)}$$

Equation (B.1) is immediate.

The same argument holds for the choice that a swindler makes with respect to all assets but her own. When choosing the position in her own asset, the only consideration is the time-zero revenue $(1-dX^{sil}_{il})P_{il}$, since the asset pays zero. Given the other investors' demands, the insider must ensure that $P_{il}=P_i$. To that end she submits a demand that fails to clear the market at $P_{il}\neq P_i$, and is willing to take any position at $P_{il}=P_i$.

To see that markets clear at prices P_j, start from (9) and a consider a regular firm k in market

j. Since by assumption we have $\iota^i_{jk} = 1$, the total demand follows from adding (B.1) and (B.2) over all i, which gives

$$\kappa \int_i dX^{ci}_{jk} + (1-\kappa) \int_i dX^{sil}_{jk} = \int_i (1-f_{ij})\kappa^{-1} dX^i_j = 1 \tag{B.7}$$

by (9). The markets for fraudulent assets clear due to the elastic demands submitted by insiders. ■

Proof of Lemma 2. The proof rests on the observation that every agent is either indifferent toward buying a particular market j or bound by the shorting constraint. Also recall that the risk-free rate is $r = 0$. This means

$$P_j \geq (E^i[U'(W^i_1)p_{ij}D_j])/E^i[U'(W^i_1)] \geq (E^i[U'(W^i_1)\kappa D_j])/E^i[U'(W^i_1)] = E^i[\xi^i \kappa D_j],$$

where we defined $\xi^i \equiv U'(W^i_1)/E^i[U'(W^i_1)]$. The first inequality is strict whenever the Lagrange multiplier λ_{ij} is strictly positive, while the second whenever $p_{ij} > \kappa$. Since $p_{ij} > \kappa$ for at least one pair (i, j), choosing investor i accordingly and summing over j, we have

$$E[\xi \kappa D^a] = \int_j P_j dj > E^i[\xi^i \kappa D^a] = P^a, \tag{B.8}$$

where the last equality follows from the fact that all agents, including agent i, are marginal in the aggregate-index derivative. ■

Proof of Lemma 3. Let M_j be the return on index j, fix an investor i, let R^i_j be this investor's return in market j and R^i on the risky portion of her portfolio; use a bar over a random variable to indicate its mean. The goal is to compute the weights $d\hat{B}^i$ and α^i in the decomposition

$$R^i = \alpha^i + \int M_j d\hat{B}^i_j + \eta^i$$

that minimizes $Var(\eta^i)$ subject to $\int d\hat{B}^i_j = 1$.

The minimization problem is equivalent to minimizing the Langragian

$$var\left(\int M_j d\hat{B}^i_j\right) - 2cov\left(R^i, \int M_j d\hat{B}^i_j\right) - 2\lambda^i \int d\hat{B}^i_j, \tag{B.9}$$

with first-order condition

$$\int cov(M_j, M_k) d\hat{B}_j^i = cov(R^i, M_k) + \lambda^i \tag{B.10}$$

for all k.

Let $\mu_j = \kappa / P_j$ be the expected return on index j, so that

$$M_j = \mu_j D_j, \tag{B.11}$$

and define

$$dB_j^i = \frac{p_{ij}}{\kappa} dw_{j.}^i \tag{B.12}$$

Notice that $R^i = \int_j M_j \frac{p_{ij}}{\kappa} dw_j^i$. The first-order condition can be written as

$$\int cov(\mu_j D_j, \mu_k D_k) d\hat{B}_j^i = \int cov(\mu_j D_j, \mu_k D_k) dB_j^i + \lambda^i, \tag{B.13}$$

with solution of the form

$$d\hat{B}_j^i = dB_j^i + \lambda^i Y_j^i \, dj \tag{B.14}$$

for an appropriate Y^i, i.e., solving the linear system

$$\int cov(\mu_j D_j, \mu_k D_k) Y_j^i \, dj = 1. \tag{B.15}$$

Note that this system is independent of the agent, enabling us to write Y instead of Y^i.

Then the style alpha is given by

$$\begin{aligned} \alpha^i &= \int \mu_j \left(dB_j^i - d\hat{B}_j^i \right) \\ &= -\lambda^i \int \mu_j Y_j \, dj. \end{aligned} \tag{B.16}$$

We note that, in the special case $\beta_j^D = 1$ for all j, the integral can be calculated by dividing (B.15)

by μ_k and integrating against dk to obtain

$$\int \mu_k^{-1} dk = \int cov(D_j, D^a)\mu_j Y_j \, dj \tag{B.17}$$

$$= \sigma_a^2 \int \mu_j Y_j \, dj. \tag{B.18}$$

In general, the value of the integral does not depend on the agent.

Finally, to write an expression for λ^i, integrate (B.14):

$$-\lambda^i = \frac{\int dB_j^i - 1}{\int Y_j \, dj}. \tag{B.19}$$

Putting together (B.16) and (B.19), the style alpha equals

$$\alpha_i = \left(\int \frac{p_{ij}}{\kappa} dw_j^i - 1 \right) \frac{\int \mu_j Y_j \, dj}{\int Y_j \, dj}. \tag{B.20}$$

The last term is independent of i and, indeed, of the agent's skill, so that the formula would apply even if there were agents with differing skills in each locations.

We conclude by noting the simplification that obtains in the special case in which $P_j = P$ for all j:

$$\alpha_i = \int \frac{p_{ij} - \kappa}{P} dw_j^i. \tag{B.21}$$

■

Proof of Lemma 4. The utility of an uninformed investor choosing dX_j^i satisfies

$$\mathrm{E}\left[U\left(\int_j (\kappa D_j - P_j) \, dX_j^i + P_i \right) \right] \leq \mathrm{E}\left[U\left(\int_j (E[\kappa] D_j - P_j) \, dX_j^i + P_i \right) \right], \tag{B.22}$$

since the function inside the square bracket on the left-hand side is concave in κ. For non-trivial distributions of κ the inequality is strict.

On the other hand, an informed investor attains utility

$$\max_{dX_j^i \geq 0} \mathrm{E}\left[U\left(\left((p_{ij} - \varphi_{ij}) \, D_j - P_j \right) dX_j^i + P_i \right) \right]. \tag{B.23}$$

It follows that, for fixed κ, choosing φ_{ij} so that $\mathrm{E}[\kappa] - (p_{ij} - \varphi_{ij})$ is sufficiently close to zero, but positive, for all i and j, investors choose to pay for the information, yet

$$\frac{E\left[\kappa\right]}{P_j} > \frac{p_{ij} - \varphi_{ij}}{P_j}. \tag{B.24}$$

The result follows. ■

Proof of Lemma 5. We note that since $L_i w_i = 0$, it follows that $Q_i w_i = Q_i^{-1} w_i = w_i$. Letting $\widehat{\Pi}_i$ denote the (diagonal) matrix $\widehat{\Pi}_i = \widehat{P}_i^{-1} Q_i$, equation (21) can be written more compactly as

$$\begin{aligned}
\widehat{\Pi}_i^{-1} w_i &= Q_i^{-1} \widehat{P}_i w_i = \widehat{P}_i Q_i^{-1} w_i = \widehat{P}_i w_i = \frac{1}{\gamma} \Omega^{-1} \widehat{P}_i^{-1} \left[\widehat{P}_i \mu - R Q_i e_{N\times 1} \right] \\
&= \frac{1}{\gamma} \Omega^{-1} \left[\mu - R \widehat{\Pi}_i e_{N\times 1} \right],
\end{aligned}$$

where $Q_i^{-1} \widehat{P}_i = \widehat{P}_i Q_i^{-1}$, since both Q_i^{-1} and $\widehat{P}_i$ are diagonal matrices. Letting $\Pi_i = K \widehat{P}_i^{-1} Q_i$, and $\mu^o = K\mu$, we obtain (22). ■

Proof of Lemma 6. Imposing the same market clearing conditions as in Sercu (1980), and using (22) we obtain

$$m = \sum_{i=1..N} \eta_i \Pi_i^{-1} w_i = \frac{1}{\gamma} \Sigma_{S|f}^{-1} \left[\mu^{S,f} - R \Pi diag(\eta) e_{N\times 1} \right], \tag{B.25}$$

Combining (B.25) with $\mu^{S,f} = \mu^S - \beta' \left(\mu^f - R e_{N\times 1} \right)$ leads to (26). Similarly, adapting the arguments as in Sercu (1980) and imposing bond market clearing leads to (27). From equations (23) and (B.25), we have

$$\begin{aligned}
w_i &= \frac{1}{\gamma} \Pi_i \Sigma_{S|f}^{-1} \left[\mu^{S,f} - R \Pi_i e_{N\times 1} \right] \\
&= \Pi_i m + \frac{R}{\gamma} \left(\Pi_i - I_{N\times N} + I_{N\times N} \right) \Sigma_{S|f}^{-1} \left(\Pi - e_{N\times N} + e_{N\times N} \right) diag(\eta) e_{N\times 1} \\
&\quad - \frac{R}{\gamma} \left(\Pi_i - I_{N\times N} + I_{N\times N} \right) \Sigma_{S|f}^{-1} \left(\Pi_i - I_{N\times N} + I_{N\times N} \right) e_{N\times 1}.
\end{aligned}$$

Noting that $e_{N\times N} diag(\eta) e_{N\times 1} = e_{N\times 1}$, where $e_{N\times N}$ is an N-by-N matrix of ones, we obtain

$$w_i = \Pi_i m + \frac{R}{\gamma} \Sigma_{S|f}^{-1} \Pi diag(\eta) e_{N\times 1} - \frac{R}{\gamma} \Sigma_{S|f}^{-1} \Pi_i e_{N\times 1} + o(||\Pi - 1||). \tag{B.26}$$

Re-arranging (B.26) into vector form leads to (24) and (25). ■

Proof of Proposition 7.

To construct an equilibrium, we proceed in a number of steps. We first construct demand functions under the postulate, later verified, that all assets in the same location have the same price, shorting is prohibited, and there is no dividend manipulation. In a second step we extend the demand functions to cover all other price configurations, while in subsequent steps we address out-of-equilibrium beliefs, dividend manipulation, and, in a final step, shorting. We focus on a particular investor 1 and use symmetry throughout.

Step 1: Two-asset equilibrium. Suppose that P_j exists such that $P_{jk} = P_j$ for all firms jk. Consider the investment problem of any agent 1, common investor or swindler. We note that, given her signals on any location j, the agent (i) excludes from consideration all firms with bad signals — these have zero payoff — and (ii) invests equally in all the others, thus minimizing idiosyncratic risk.[37] At this stage, we also assume that the beliefs about the asset qualities are given by the signals, and are not updated based on the price. We address out-of-equilibrium beliefs in a later step.

This problem is a relatively standard, perfect-information one, featuring mean-variance investors facing differential taxes who invest strategically in multiple assets in the presence of shorting constraints. For the sake of completeness, we sketch proofs of both existence and convergence towards the competitive outcome as the number of agents grows to infinity.

Given that the asset payoffs in the two locations are independent and preferences are CARA, investments in the two assets do not interact. Market clearing, however, involves the demands of both agents, so we choose to write the problem in matrix form, even if all endogenous matrices are diagonal.

Let Π be diagonal with $\Pi_{jj} = p_{1j}$. Agent 1 faces a two-asset universe with expected payoffs $\text{diag}(\Pi)$ and variance-covariance matrix of payoffs $\Pi\Omega\Pi$. With X the portfolio choice of the agent, it is convenient to focus on the quantity $Y = \Pi X$. We are looking for an equilibrium in which all demands, as functions of the two prices, are piece-wise linear — in fact, linear truncated at zero. Relying on symmetry, we need to parameterize only the demand of agent 1, as

$$Y = (A - BP)^+ = Z\,(A - BP)\,, \tag{B.27}$$

[37] The agent is perfectly informed about location 1, and thus does not face any idiosyncratic risk, but equal weighting is, of course, still optimal, albeit only weakly.

where $P \in \mathbb{R}^2$ denotes the vector of prices per share in the two markets, and Z is a diagonal matrix with $Z_{jj} = 1$ if entry j of $A - BP$ is positive and $Z_{jj} = 0$ otherwise. Z is a function of P. A requirement of the equilibrium is that, taking all other demand schedules as given, an agent 1's optimal portfolio choice, denoted $\hat{X}$ (respectively $\hat{Y}$), and therefore P, is optimal subject to the restriction $\hat{X} \geq 0$. The type of equilibrium we are looking for requires that $\hat{Y} = Z(A - BP)$.

Consider price setting for a regular firm. We note that if an agent 1 demands X ($\in \mathbb{R}^2$) total shares in each location, the total demand for regular-firm shares is only $\Pi X = Y$. The reason is that, in market j, a proportion $1 - p_{1j}$ of the demand flows to fraudulent firms.

Given (B.27), it follows that the residual demand faced by an agent 1 for the aggregate regular asset in each location is also linear, given as

$$Y^{(r)} = A^{(r)} - B^{(r)}P, \tag{B.28}$$

with $A^{(r)} = (N-1)ZA + N\mathcal{R}ZA$ and $B^{(r)} = (N-1)ZB + N\mathcal{R}ZB\mathcal{R}$, where the matrix $\mathcal{R}$ implements the permutation $1 \leftrightarrow 2$ applied to all indices to capture the demand of agents in location 2; that is,

$$\mathcal{R} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

An agent 1 maximizes

$$\hat{Y}^\top \left(\mathbf{1} - \Pi^{-1}P\right) + \frac{\kappa}{N} e_1^\top P - \frac{\gamma N}{2} \hat{Y}^\top \Omega \hat{Y}, \tag{B.29}$$

where $e_1 = (1, 0)^\top$ is a vector that selects market 1, to capture the agent's endowment, and $\mathbf{1} = (1, 1)^\top$. Note that the agent's endowment represents a fraction $1/N$ of the total endowment in market 1, and his risk tolerance a fraction $1/N$ of the aggregate risk tolerance of agents in location 1. The agent takes into account that the price depends on her demand through the market-clearing condition

$$\hat{Y} + Y^{(r)}(P) = \kappa \mathbf{1}. \tag{B.30}$$

The logic of the argument is familiar. The agent can be thought of as choosing the quantity $\hat{Y}$

and via (B.30) the price vector P, taking the residual demand as given. Before we compute the optimal demand, we remark, based on (B.29), that (i) if $P_1 < 1$, then $\hat{Y}_1 > 0$, and (ii) $\hat{Y}_2 > 0$ if and only if $P_2 < p$. Since the market for asset 1 must clear, $Y_1 > 0$ in equilibrium, while Y_2 may or not be positive.

The first-order condition for the Lagrangean associated with (B.29) is

$$0 = 1 - \Pi^{-1}P - (D_{\hat{Y}}P)\Pi^{-1}\hat{Y} + (D_{\hat{Y}}P)e_1\frac{\kappa}{N} - \gamma N\Omega\hat{Y} + \lambda, \tag{B.31}$$

where λ is the Lagrange-multiplier vector attached to the no-shorting condition, while differentiating (B.30) with respect to $\hat{Y}$ gives

$$\begin{aligned} 0 &= 1 + D_P Y^{(r)} D_{\hat{Y}} P \\ &= 1 - B^{(r)} D_{\hat{Y}} P. \end{aligned} \tag{B.32}$$

Note that, via Z, $A^{(r)}$ and $B^{(r)}$ are actually functions of P. Since they are step functions, though, we may treat them as constants when evaluating first-order conditions. We verify later that the first-order approach generates an equilibrium in our context.

Putting (B.31) and (B.32) together provides a candidate demand schedule for agent 1:

$$\hat{Y} = \hat{Z}\left(\left(B^{(r)}\right)^{-1}\Pi^{-1} + \gamma N\Omega\right)^{-1}\left(\mathbf{1} - \Pi^{-1}P + \left(B^{(r)}\right)^{-1}e_1\frac{\kappa}{N}\right). \tag{B.33}$$

We'll define the coefficients of the linear demand of the agent 1 under consideration based on (B.33), but we first determine the value that the matrix Z, and therefore $\hat{Z}$, takes in equilibrium. Let P^* denote the equilibrium price and suppose that $P_1^* > p$, so that agent 1 is the only one investing in market 1. Then, from (B.33) market clearing implies

$$\kappa = N\hat{Y} = \left(N^{-1}\left(B_{11}^{(r)}\right)^{-1} + \gamma\sigma^2\right)^{-1}\left(1 - P_1^* + N^{-1}\left(B_{11}^{(r)}\right)^{-1}\kappa\right), \tag{B.34}$$

which gives upon rearrangement

$$P_1^* = 1 - \gamma\kappa\sigma^2. \tag{B.35}$$

Thus, if $p < 1-\gamma\kappa\sigma^2$ and the equilibrium satisfies (B.33), then agent 1 doesn't invest in market 2.

Suppose now that the agent invests in both markets. We use again market clearing from (B.33),

$$\kappa = \left(N^{-1}\left(B_{11}^{(r)}\right)^{-1} + \gamma\sigma^2\right)^{-1}\left(1 - P_1^* + N^{-1}\left(B_{11}^{(r)}\right)^{-1}\kappa\right) + \left(N^{-1}\left(B_{22}^{(r)}\right)^{-1}p^{-1} + \gamma\sigma^2\right)^{-1}\left(1 - p^{-1}P_2^*\right), \tag{B.36}$$

which, using $P_1^* = P_2^*$, leads to

$$P_1^* = 1 - \gamma\kappa\sigma^2 + \frac{\left(N^{-1}\left(B_{11}^{(r)}\right)^{-1} + \gamma\sigma^2\right)}{\left(N^{-1}\left(B_{22}^{(r)}\right)^{-1}p^{-1} + \gamma\sigma^2\right)}\left(1 - p^{-1}P_1^*\right), \tag{B.37}$$

further rewritten as

$$\left(p - P_1^*\right)\left(1 + a_0\right) = p - \left(1 - \gamma\kappa\sigma^2\right) \tag{B.38}$$

with

$$a_0 = \frac{\left(N^{-1}\left(B_{11}^{(r)}\right)^{-1} + \gamma\sigma^2\right)}{\left(N^{-1}\left(B_{22}^{(r)}\right)^{-1} + p\gamma\sigma^2\right)} > 0.$$

Thus $p > P_1^*$, implying active participation of agent 2 in market 1 if and only if $p > 1-\gamma\kappa\sigma^2$. We have therefore shown that $Z_{22}(P^*) = 1_{(p>1-\gamma\kappa\sigma^2)}$, should an equilibrium exist.

Let $Z^* \equiv Z(P^*)$. We define the demand of agent 1 under consideration based on the coefficients

$$\phi\left(B\right) = Z^*\left(\left(B^{(r)}\right)^{-1} + \gamma N\Omega\right)^{-1}\Pi^{-1} \tag{B.39}$$

$$\psi\left(B\right) = \phi\left(B\right)\Pi\left(\mathbf{1} + \left(B^{(r)}\right)^{-1}e_1\frac{\kappa}{N}\right). \tag{B.40}$$

Note that we used the known value of Z^* in this definition, as defined above.

Suppose that the mapping $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ admitted a non-zero, positive fixed point. Then let

$A = \psi(B)$ and compute the price from the market-clearing condition. It follows from equation (B.33) that $\hat{Y} = \psi(B) - \phi(B)P^*$ as long as $\hat{Z}(P^*) = Z^*$, i.e., as long as (i) $(\psi(B) - \phi(B)P^*)_{11} = (A - BP^*)_{11}$ is positive, and (ii) $(\psi(B) - \phi(B)P^*)_{22} = (A - BP^*)_{22}$ is positive if and only if $Z^*_{22} = 0$, thus $p > 1 - \gamma\kappa\sigma^2$.

Note, however, that $Z^*_{22} = 0$ implies $B_{22} = A_2 = 0$, and therefore $\hat{Z}_{22}(P^*) = 0$. Given symmetry and market clearing, it follows that $A_1 - B_{11}P^*_1 > 0$. If $Z^*_{22} = 1$ and both investors invest in both markets, then we saw above that market clearing leads to (B.38), which implies $P^*_1 < p$. From the definitions,

$$A - BP^* = \left(\left(B^{(r)}\right)^{-1} + \gamma N\Omega\right)^{-1}\left(\mathbf{1} + \left(B^{(r)}\right)^{-1} e_1 \frac{\kappa}{N} - \Pi^{-1}P^*\right). \tag{B.41}$$

Since $P^*_1 < p$, $\mathbf{1} - \Pi^{-1}P^* > 0$, and therefore $(A - BP^*)_{jj} > 0$, thus $\hat{Z}^*_{jj} > 0$, for $j \in \{1, 2\}$.

One can use Brouwer's theorem to show that ϕ has a strictly positive fixed point, as follows. It is convenient to concentrate on the mapping $\phi_N(NB) \equiv N\phi(B)$. To invoke this theorem, we restrict attention to $B_1 > 0$ (and $B_2 \geq 0$) and note that, as a consequence, the image of ϕ_N is bounded above (in the operator sense) — uniformly in N, in fact. We also see that $\delta > 0$ exists such that, if $NB_1 \geq \delta$, then $(\phi_N(NB))_1 > \delta$. Specifically, from (B.39) it follows that δ must obey

$$N\left((2N-1)^{-1}\delta^{-1} + \gamma N\sigma^2\right)^{-1} \geq \delta, \tag{B.42}$$

which holds, for instance, for

$$\delta = \frac{1}{2}\left(\gamma\sigma^2\right)^{-1}. \tag{B.43}$$

Thus, we have verified that the continuous mapping ϕ_N maps a compact set into itself, and therefore has a fixed point B characterized by $B_1 > \delta > 0$, $B_2 \geq 0$.

The last fact that must be established before concluding that we have an equilibrium is that the agents' portfolios are, indeed, optimal. We constructed them to satisfy first-order conditions, but the agents' objectives (B.29) are not concave in general. Given equilibrium residual demands, though, they are concave.

To make the matters clear, we rewrite (B.29) as

$$\hat{Y}^\top \mathbf{1} - \left(\hat{Y}^\top \Pi^{-1} - \frac{\kappa}{N} e_1^\top\right) P - \frac{\gamma N}{2} \hat{Y}^\top \Omega \hat{Y}, \tag{B.44}$$

Since P is not a constant, this function is not quadratic in $\hat{Y}$. The implicit function $P(\hat{Y})$, though, defined via (B.30), is (piece-wise linear and) convex. To obtain concavity, it is sufficient that $P(\hat{Y})$ be linear whenever $\hat{Y}^\top \Pi^{-1} - \frac{\kappa}{N} e_1^\top < 0$. Differently put, that for $\hat{Y} < \frac{\kappa}{N} \Pi e_1 = \frac{\kappa}{N} e_1$, both agents are long in both markets given equilibrium demands and $P(\hat{Y})$.

Obviously, there is no problem concerning asset 2 (since $e_{12} = 0$: agent 1 has no endowment of asset 2). For asset 1, consider first the case in which agents only participate in their home markets in equilibrium. Since we defined $A_2 = B_{22} = 0$, i.e., constant, in this case, there is no issue. In the other case, suppose that the agent increases $\hat{Y}_1$ until agents 2 drop out of the market, i.e., the price becomes $P_1 = p$. Each other agent 1, at this price, holds

$$\begin{aligned} A_1 - B_{11} p &= \left((B_{11}^{(r)})^{-1} + \gamma N \sigma^2\right)^{-1} \left(1 + (B_{11}^{(r)})^{-1} \frac{\kappa}{N} - p\right) && \text{(B.45)} \\ &< \left((B_{11}^{(r)})^{-1} + \gamma N \sigma^2\right)^{-1} \left((B_{11}^{(r)})^{-1} \frac{\kappa}{N} + \gamma \kappa \sigma^2\right) && \text{(B.46)} \\ &= \frac{\kappa}{N}, && \text{(B.47)} \end{aligned}$$

where we used $p > 1 - \gamma \kappa \sigma^2$. The residual demand is therefore linear until a point where $\hat{Y}_1 = \kappa - (N-1)(A_1 - B_{11} p) > \frac{\kappa}{N}$.

Let's turn now to the behavior of this equilibrium as N grows large. Since NB is bounded below by $\delta > 0$ (independently of N), (B.39) implies

$$\bar{B} = Z^* (\gamma \Omega)^{-1} \Pi^{-1}, \tag{B.48}$$

with $\bar{B}$ denoting the limit of NB.

We note that $\lim_{N \to \infty} B^{(r)} = \bar{B} + \mathcal{R} \bar{B} \mathcal{R}$. It follows, using (B.40), that the limit of NA is

$$\bar{A} = \bar{B} \Pi \mathbf{1} = Z^* (\gamma \Omega)^{-1} \mathbf{1} \tag{B.49}$$

and that the limit price per share is

$$\bar{P} = \left(\bar{B} + \mathcal{R}\bar{B}\mathcal{R}\right)^{-1}(\bar{A} + \mathcal{R}\bar{A} - \kappa\mathbf{1}). \tag{B.50}$$

The quantities $\bar{A}$ and $\bar{B}$ also correspond to the solution to (B.29) taking the price as given. The price $\bar{P}$ is therefore the competitive price in the two-asset, short-sale constrained equilibrium. The term $\left(\bar{B} + \mathcal{R}\bar{B}\mathcal{R}\right)^{-1}(\bar{A} + \mathcal{R}\bar{A})$ captures the aggregate expected payoff from an asset one invests in. The remaining term is the risk adjustment, with $\kappa\mathbf{1}$ the supply of the asset, and $\left(\bar{B} + \mathcal{R}\bar{B}\mathcal{R}\right)^{-1}$ accounting for the covariance between one asset the aggregate investor purchases and one unit of the total supply.

Step 2: The other demands. Returning to the finite-N case, we consider now the demand of a swindler in her own firm. Given that the only way to generate a positive demand in her firm — given the other agents' equilibrium strategies, described below — is to ensure that its price is P_1, the swindler submits a perfectly elastic demand at the price P_1 at which all other firms clear, as long as the demand is a quantity that does not exceed one. At all other prices, the swindler submits a demand $(1, \infty)$, i.e., stands ready to clear the market as long as she takes a gross position higher than one. This case can only obtain when another agent is willing to short at the respective price, and the optimal reaction of the swindler is to accommodate the shorter and manipulate dividends, as we describe below.

Formally, the demand of the swindler for her own firm l is

$$X^{sl}(P_{1l}) = \begin{cases} (-\infty, \infty) & \text{if} \quad \exists P_1 \ \& \ P_{1l} = P_1 \\ (1, \infty) & \text{if} \quad \nexists P_1 \text{ or } P_{1l} \neq P_1 \end{cases}. \tag{B.51}$$

We also note that, in equilibrium, the swindler never shorts her own firm, since the demand for it is lower than the demand for a regular firm, about which the signals are better.

All agents demand a zero amount of shares in firms in which they have bad signals. Local agents know precisely all regular firms. If the set of prices of regular firms in market 1 is not a singleton, then we specify the demand of agent 1 as $\hat{Y}_{1j} = \infty$ for all j such that $P_{1j} < \max_k P_{1k}$, and $\hat{Y}_{1j} = -\infty$ for all j such that $P_{1j} = \max_k P_{1k}$. If all positive-signal firms in market 2 do not have the same price, then agent 1 maximizes utility conditional on his out-of-equilibrium beliefs, stated below. The important feature to note is that no deviation by a swindler can prevent the

regular firms having the same price — this is ensured by the demands of the local common investors — and therefore the only relevant belief concerns the case in which the price P_{2j} of a fraudulent firm is not equal to that of at least κN other firms in location 2.[38]

Step 3: Out-of-equilibrium beliefs. Investor 1 knows all types in market 1. Suppose that she observes prices P_{2k}. Given $\kappa > \frac{1}{2}$, there are only two possibilities. First, at least κN firms of the $p^{-1}\kappa N$ ones about which the investor has positive signals have the same price (there is only one price level for which this statement is true). Then the agent assigns probability one that all firms with different price are fraudulent. Second, there is no such subset of firms. Then the agent believes that the prices are entirely uninformative.

As remarked above, given the prescribed strategies, no swindler can bring about the second case, and it is optimal for each swindler to induce the pooling outcome, as it carries zero cost and unilateral deviation exposes the firm as certainly fraudulent.

Finally, common agents do not have an incentive to adjust demand for the regular stocks in market 1 in the hope of signaling, given the swindlers' equilibrium demand (which ensures the same price for the fraudulent firm as for the regular ones).

Step 4: Dividend manipulation. The swindler's action also includes the amount of dividend manipulation she engages in, subsequent to asset-market clearing. If the swindler borrows the amount $F \geq 0$ that she diverts in the firm, then she makes profit

$$X^{sl}\left(F - P_{1l}\right) - F + P_{il} = P_{il}\left(1 - X^{sl}\right) + F\left(X^{sl} - 1\right). \tag{B.52}$$

Clearly, if $X^{sl} > 1$, the swindler benefits from borrowing an arbitrarily large amount L to divert in the firm, pushing its liquidating price arbitrarily high and making arbitrarily high profits. As long as there is no shorting, however, there is no manipulation.

Step 5: Shorting. Agent 1 is perfectly informed about assets in market 1, and does not short regular assets in market 1, as discussed in Step 1. She also does not short fraudulent assets in market 1 because agent 2 may also know that they are fraudulent, and therefore the swindler would be the only buyer, would end up being a net buyer of the asset and manipulate the dividends to an arbitrary extent. Agent 1 does not short an asset in market 2 that is fraudulent, for everyone

[38] To simplify exposition, we make the parametric assumption $\kappa > \frac{1}{2}$, which excludes the possibility that there are two or more disjoint sets of firms of size κN, which may have the same price.

knows that it is fraudulent. Finally, she may consider shorting an asset in market 2 about which she has a positive signal. If the equilibrium demand for market 2 by agent 1 in the shorting-constrained economy without manipulation is zero, then there is a positive probability that the asset is fraudulent and no one else invests in it. If this demand is actually strictly positive, then taking a negative position would be suboptimal even in the absence of the manipulation threat. ■