Optimal Financial Transaction Taxes*

Eduardo Dávila[†]

August 2021

Abstract

This paper characterizes the optimal transaction tax in an equilibrium model of financial markets. If investors hold heterogeneous beliefs unrelated to their fundamental trading motives and the planner calculates welfare using any single belief, a positive tax is optimal, regardless of the magnitude of fundamental trading. Under some conditions, the optimal tax is independent of the planner's belief. The optimal tax can be implemented by adjusting its value until total volume equals fundamental volume. Knowledge of i) the share of non-fundamental trading volume and ii) the semi-elasticity of trading volume to tax changes is sufficient to quantify the optimal tax.

JEL Classification: G18, H21, D61

Keywords: transaction taxes, Tobin tax, belief disagreement, optimal corrective taxation, behavioral public economics

*The first version of this paper was circulated in May 2013. I am very grateful to Philippe Aghion, Adrien Auclert, Robert Barro, Roland Bénabou, Markus Brunnermeier, John Campbell, Raj Chetty, Peter Diamond, Emmanuel Farhi, Xavier Gabaix, Gita Gopinath, Robin Greenwood, Sam Hanson, Oliver Hart, Nathan Hendren, David Laibson, Sendhil Mullainathan, Adriano Rampini, Larry Samuelson, David Scharfstein, Florian Scheuer, Andrei Shleifer, Alp Simsek, Jeremy Stein, Adi Sunderam, Glen Weyl, Wei Xiong, and to seminar participants at many institutions and conferences. Johnny Xu and Ran Wang provided outstanding research assistance. Financial support from Rafael del Pino Foundation is gratefully acknowledged.

[†]Yale University and NBER. Email: eduardo.davila@yale.edu

1 Introduction

Whether to tax financial transactions or not remains an important open question for public economics that periodically gains broad relevance after periods of economic turmoil. For instance, the collapse of the Bretton Woods system motivated James Tobin's well-known 1972 speech — published as Tobin (1978) — endorsing a tax on international transactions. The 1987 crash encouraged Stiglitz (1989) and Summers and Summers (1989) to argue for implementing a transaction tax, while the 2008 financial crisis spurred further public debate on the issue, leading to a contested tax proposal by the European Commission. However, with the lack of formal normative studies of this topic, a financial transaction tax may still seem like "the perennial favorite answer in search of a question", per Cochrane (2013).

In this paper, I study the welfare implications of taxing financial transactions in an equilibrium model in which financial markets play two distinct roles. On the one hand, financial markets allow investors to conduct *fundamental trading*. Fundamental trading allows the transfer of risks towards those investors more willing to bear them. It also allows for trading on liquidity or life-cycle considerations, as well as trading for market-making or limited arbitrage purposes. On the other hand, financial markets also allow investors to engage in betting or gambling, which I refer to as *non-fundamental trading*.

I model non-fundamental trading by assuming that investors' trades are partly motivated by differences in beliefs about future payoffs, while the planner calculates welfare using a single belief. The discrepancy between the planner's belief and investors' beliefs implies that corrective policies, which can involve taxes or subsidies depending on the primitives of the economy, are generically optimal. Three main results emerge from the optimal taxation exercise.

First, the optimal transaction tax can be expressed as a function of investors' beliefs and equilibrium portfolio sensitivities to tax changes. Specifically, the optimal tax corresponds to one-half of the difference between a weighted average of buyers' beliefs and a weighted average of sellers' beliefs. Beliefs are the key determinant of the optimal tax because of the corrective nature of the policy. In general, optimal corrective policies are designed to correct marginal distortions, which in this case arise from investors' differences in beliefs.

Second, a simple condition involving the cross-sectional covariance between investors' beliefs and their status as net buyers or net sellers in the laissez-faire economy determines the sign of the optimal corrective policy. Importantly, when investors' beliefs are not related to their fundamental motives for trading, this condition implies that the optimal policy is a strictly positive tax. Therefore, as long as the planner is aware of the existence of some belief-driven trades, under the assumption that these trades are not related to other fundamental trading motives, a positive transaction tax is optimal.[1] Intuitively, the planner perceives that a reduction in trading, starting from the laissez-faire equilibrium, generates a first-order welfare gain for those investors who are optimistic buyers and pessimistic sellers.[2] When these are the majority of investors, their first-order welfare gains dominate the second-order welfare losses of those investors who share the planner's belief and the first-order welfare losses of optimistic sellers and pessimistic buyers. When instead the economy is populated by many optimistic sellers and pessimistic buyers, the optimal corrective policy is a trading subsidy.

Third, the optimal tax turns out to be independent of the belief used by the planner to calculate welfare under certain conditions. This surprising and appealing result relies on the fact that the traded assets are in fixed supply and that the planner does not seek to redistribute resources across investors. Intuitively, because welfare losses

[1] Although independence between beliefs and other trading motives is a plausible sufficient condition for an optimal positive tax, it is by no means necessary.

[2] Optimistic (pessimistic) investors are those who believe that the expected asset payoff is high (low) relative to other investors.

in this model arise from a distorted allocation of risk, only the dispersion in investors' beliefs — but not their absolute level — determines aggregate welfare and the optimal tax. Consequently, the planner does not need to know more than the investors to determine the optimal tax.

Because optimal tax characterizations are inherently local to the optimum, I study the convexity properties of the planner's problem. The planner's objective may fail to be quasi-concave, which implies that there may exist multiple locally optimal transaction tax rates. However, this phenomenon can only arise when the composition of marginal investors varies with the tax rate. I provide a natural sufficient condition under which the planner's problem is well-behaved and has a unique optimum. When investors exclusively trade for non-fundamental motives, the optimal policy is associated with an infinite tax that eliminates all trade. This result may help explain why some jurisdictions ban or heavily tax gambling activities.

Given the challenges associated with directly measuring investors' beliefs, I provide an implementation of the optimal tax policy that uses trading volume as an intermediate target. Under this implementation, a planner can simply adjust the tax rate until trading volume and the fundamental component of trading volume are equal. This alternative approach, which relies on a novel decomposition of trading volume into fundamental volume, non-fundamental volume, and the tax-induced volume reduction, shifts the planner's informational requirements from measuring investors' beliefs to finding an appropriate estimate of fundamental volume. Building on this volume decomposition, I also derive an approximation for the optimal tax that relies exclusively on two objects of the laissez-faire economy: the semi-elasticity of trading volume to tax changes and the share of non-fundamental trading volume.[3] This approximation — valid when the optimal tax is close to zero — does not impose any restrictions on investors' trading motives.

Next, after parameterizing the distribution of fundamental and non-fundamental trading motives, I provide explicit comparative statics results for the optimal tax with respect to primitives. Consistent with the main results, when fundamental and non-fundamental trading motives are jointly normally distributed and uncorrelated, the optimal tax is positive. Moreover, the optimal tax is increasing in the share of non-fundamental trading volume. Also, when the optimal tax is positive and finite, a mean-preserving spread of the distribution of investors' beliefs is associated with a higher optimal tax.

In the context of the parameterized model, a planner who seeks to determine the optimal tax rate only needs to know two high-level sufficient statistics. These are i) the semi-elasticity of trading volume to tax changes and ii) the share of non-fundamental trading volume. If the planner also knows the risk premium, it is possible to compute the aggregate marginal welfare gains associated with a tax change. I also describe how to find plausible empirical counterparts of the identified sufficient statistics using existing evidence. A calibration of the optimal tax that is consistent with empirically estimated volume semi-elasticities to tax changes and that features a 30% share of non-fundamental trading volume is associated with an optimal tax of 37bps (0.37%). I conduct a sensitivity analysis and provide a menu of optimal taxes for different values of the volume semi-elasticity as well as the share of non-fundamental trading volume. For instance, when the share of non-fundamental trading volume is 10% or 60%, the model predicts optimal taxes of 10bps (0.1%) or 105bps (1.05%), respectively. I also discuss two issues of key practical importance. First, I describe the connection between the optimal corrective tax and tax revenues. Second, I illustrate how the possibility of tax evasion driven by imperfect tax enforcement — in which trading activity migrates to a different (untaxed) trading venue — affects the determination of the optimal tax.

Finally, I establish the robustness of the results and discuss their limitations. First, I characterize the optimal

[3] Given existing estimates of trading volume elasticities to tax changes, this approximation implies an optimal tax of the same order of magnitude of the share of non-fundamental trading volume, when expressed in basis points. That is, a 10%, 20%, or 40% share of non-fundamental trading volume is associated (approximately) with an optimal tax of 10bps, 20bps, or 40bps.

tax for more general specifications of beliefs and utility and show that the optimal tax formula characterized in Proposition 1 remains approximately valid in the general case when investors' stochastic discount factors are approximately constant. This result validates the characterization of the optimal tax as an approximation to any specification of beliefs and preferences. Next, I briefly describe in the paper how the results extend to environments with short-sale constraints, pre-existing trading costs, imperfect tax enforcement, multiple traded assets, production, and dynamics. The Appendix formally includes these and other extensions. Finally, I discuss several channels that are not explicitly explored in this paper but that are relevant when considering the desirability of financial transaction taxes, in particular the role of asymmetric information.

This paper belongs to the literature that follows Tobin's proposal of introducing transaction taxes to improve the societal performance of financial markets. Although Tobin's speech largely focused on foreign exchange markets, it has become customary to refer to any tax on financial transactions as a "Tobin tax". Stiglitz (1989) and Summers and Summers (1989) verbally advocate for a financial transaction tax, with Ross (1989) taking the opposite view. Roll (1989) and Schwert and Seguin (1993) contrast the costs and benefits of such a proposal. Umlauf (1993), Campbell and Froot (1994), several chapters in ul Haq, Kaul and Grunberg (1996), and Jones and Seguin (1997) are representative samples of empirical work in the area. See McCulloch and Pacillo (2011) and Burman et al. (2016) for recent surveys and Colliard and Hoffmann (2017) and Cai et al. (2017) for evidence on the recently introduced transaction taxes in France and China, respectively.[4]

The theory in this paper differs substantially from that in Tobin (1978). Tobin postulates that prices are excessively volatile and that a transaction tax is a good instrument to reduce price volatility. This paper shows instead that transaction taxes are a robust instrument to reduce trading volume but that their effect on asset prices is a priori indeterminate. The normative results in this paper rely on the fact that a reduction in trading volume improves the allocation of risk in the economy from the planner's perspective.[5]

This paper is most directly related to the growing literature that evaluates welfare under belief disagreements in financial markets. Weyl (2007) is the first to study the efficiency of arbitrage in an economy in which some investors have mistaken beliefs. Gilboa, Samuelson and Schmeidler (2014) and Gayer et al. (2014) introduce refinements of the Pareto criterion that identify negative-sum betting situations. No-betting Pareto requires that there exists a single belief that, if held by all agents, implies that all agents are better off by trading. Unanimity Pareto requires that all agents perceive to be better off by trading using each agent's belief. The welfare criterion proposed by Brunnermeier, Simsek and Xiong (2014) assesses efficiency by using all possible convex combinations of investors' beliefs in the economy. These papers seek to identify outcomes related to zero-sum speculation, but do not discuss policy measures to limit trading, which is the raison d'être of this paper. In the same spirit, Posner and Weyl (2013) advocate for financial regulation grounded on price-theoretic analysis, which is precisely my goal with this paper.[6] Blume et al. (2013) propose a criterion in which a planner evaluates welfare under the worst-case scenario among a set of belief assignments. They quantitatively analyze several restrictions on trading but do not characterize optimal policies. Heyerdahl-Larsen and Walden (2014) propose a criterion in which the planner does not have to take a stand on which belief to use, within a reasonable set, to assess efficiency. I relate my results to

[4]Since this paper originally circulated, a number of other papers have further studied financial transaction taxes. See, among others, Coelho (2014), Vives (2017), Dávila and Parlatore (2021), Berentsen, Huber and Marchesiani (2015), Dang and Morath (2015), Alvarez and Atkeson (2018), Biais and Rochet (2020), and Dieler et al. (2020).

[5]Financial market interventions may be optimal in other environments — see, among others, Scheuer (2013) or Dávila and Korinek (2018). However, these theories do not imply that transaction taxes of the form studied in this paper are optimal or even desirable — see the Appendix for an elaboration of this point in a model with pecuniary externalities and incomplete markets.

[6]A growing literature exploits market design tools to study normative issues in market microstructure. See, for instance, Budish, Cramton and Shim (2015) or Baldauf and Mollner (2014).

these criteria when appropriate.

Many papers explore the positive implications of speculative trading due to belief disagreements, following Harrison and Kreps (1978). Scheinkman and Xiong (2003) analyze the positive implications of a transaction tax in a model with belief disagreements, but they do not draw normative conclusions. Panageas (2005) and Simsek (2013) study implications for production and risk-sharing of speculative trading motives. See Xiong (2013) and Simsek (2021) for recent surveys of this line of work. Since some trades are not driven by fundamental considerations, this paper also relates to the literature on noise trading that follows Grossman and Stiglitz (1980). However, the standard noise trading formulation makes it hard to understand how noise traders react to taxes and how to evaluate their welfare. By using heterogeneous beliefs to model non-fundamental trading, this paper sidesteps these concerns.

Given the additive nature of corrective taxes (see e.g., Sandmo (1975); Kopczuk (2003)), the normative conclusions that emerge from explicitly incorporating dispersed information and learning operate in parallel to the results of this paper. This is an active area of research. In recent work, Dávila and Parlatore (2021) characterize the conditions under which transaction costs/taxes do not affect information aggregation, even though they discourage the endogenous acquisition of information. This is a different margin through which transaction taxes may have an independent effect on welfare. Along the same lines, Vives (2017) examines an environment in which a positive transaction tax is welfare improving by correcting investors' information acquisition choices.

The literature on transaction costs is formally related to this paper, since a transaction tax is similar to a transaction cost from a positive point of view. This literature studies the positive effects of transaction costs on portfolio choices and equilibrium variables like prices and volume. I refer the reader to Vayanos and Wang (2012) for a recent comprehensive survey. While those papers focus on the positive implications of exogenously given transaction costs/taxes, in this paper I study the welfare effects of a transaction tax and its optimal determination. I explicitly relate the positive results of the paper to this work in the text when appropriate.

Finally, this paper contributes to the growing literature on behavioral welfare economics, recently synthesized in Mullainathan, Schwartzstein and Congdon (2012). This paper is related to Gruber and Koszegi (2001) and O'Donoghue and Rabin (2006), who characterize optimal corrective taxation when agents fail to optimize because of self-control or limited foresight. Within this literature, the work by Sandroni and Squintani (2007) and Spinnewijn (2015), who study optimal corrective policies when agents have distorted beliefs, is closely related. While those papers respectively study optimal policies in insurance markets and frictional labor markets, this paper derives new insights in the context of financial market trading. Farhi and Gabaix (2015) have recently studied optimal taxation with behavioral agents, while Campbell (2016) advocates for incorporating behavioral insights into optimal policy prescriptions.

Section 2 introduces the model and Section 3 studies its positive predictions. Section 4 conducts the normative analysis, presenting the main results. Section 5 provides explicit comparative statics for the optimal tax and explores the quantitative implications of the model. Section 6 discusses the robustness and limitations of the results and Section 7 concludes. The Appendix includes proofs and derivations, as well as additional extensions.

2 Model

In the absence of transaction taxes, the baseline environment of this paper resembles Lintner (1969), who relaxes the CAPM by allowing for heterogeneous beliefs among investors.

Investors There are two dates $t = 1, 2$ and there is a unit measure of investors. Investors (investors' types) are indexed by i and distributed according to a continuous probability distribution with c.d.f. $F(\cdot)$ such that $\int dF(i) = 1$.

Investors choose their portfolio optimally at date 1 and consume at date 2. They maximize expected utility with preferences that feature constant absolute risk aversion. Therefore, each investor maximizes

$$\mathbb{E}_i \left[U_i \left(W_{2i} \right) \right] \quad \text{with} \quad U_i \left(W_{2i} \right) = -e^{-A_i W_{2i}}, \tag{1}$$

where Equation (1) already imposes that investors consume all terminal wealth, that is, $C_{2i} = W_{2i}$. The parameter $A_i > 0$, which represents the coefficient of absolute risk aversion $A_i \equiv -\frac{U_i''(\cdot)}{U_i'(\cdot)}$, can vary across the distribution of investors. The expectation in Equation (1) is indexed by i because investors hold heterogeneous beliefs, as described below.

Market structure and beliefs There is a risk-free asset in elastic supply that offers a gross interest rate normalized to 1. There is a single risky asset in exogenously fixed supply $Q \geq 0$. The price of the risky asset at date 1 is denoted by P_1 and is quoted in terms of an underlying good (dollar), which acts as numeraire. To simplify the exposition, and without loss of generality, I assume that the fundamentals of the economy are such that the equilibrium price of the risky asset is always strictly positive, that is, $P_1 > 0$.[7] The initial holdings of the risky asset at date 1, given by X_{0i}, are arbitrary across the distribution of investors. Investors' initial holdings of the risky asset must add up to the total asset supply Q, therefore $\int X_{0i} dF(i) = Q$. Investors face no constraints when choosing portfolios: they can borrow and short sell freely.

The risky asset yields a dividend D at date 2, which is normally distributed with an unspecified mean and a variance $\mathbb{V}ar[D]$. An investor i believes that D is normally distributed with a mean $\mathbb{E}_i[D]$ and a variance $\mathbb{V}ar[D]$, that is,

$$D \sim_i N \left(\mathbb{E}_i [D], \mathbb{V}ar[D] \right).$$

For now, the distribution of mean beliefs $\mathbb{E}_i[D]$ across the population of investors, which is a key primitive of the model, is arbitrary.[8] Nothing prevents investors from having correct beliefs; those investors can represent market makers or (limited) arbitrageurs. Investors do not learn from each other or from the price, and agree to disagree in the Aumann (1976) sense.

Two arguments justify the assumption of investors who disagree about the mean — not other moments — of the distribution of payoffs. First, it is commonly argued that second moments are easier to learn. In particular, with Brownian uncertainty, second moments can be learned instantly. Second, as formalized in Section 6, in a precise approximate sense, only the mean of investors' beliefs enters explicitly in the optimal tax formula.

Hedging needs Every investor has a stochastic endowment at date 2, denoted by M_{2i}, which is normally distributed and potentially correlated with D. This endowment captures the fundamental risks associated with the normal economic activity of the investor. The quantity of endowment risk that an investor i faces is captured by the covariance $\mathbb{C}ov[M_{2i}, D]$, which is known to all investors. For now, the sign and magnitude of investors' hedging needs are arbitrary across the distribution of investors. Without loss of generality, I assume that $\mathbb{E}[M_{2i}] - \frac{A_i}{2}\mathbb{V}ar[M_{2i}] = 0$ and normalize investors' initial dollar endowment to zero.

[7] The Appendix provides a sufficient condition that guarantees that P_1 is strictly positive in equilibrium.

[8] A common prior model in which investors receive a purely uninformative signal (noise), but pay attention to it, maps one-to-one to the environment in this paper. Alternatively, investors could neglect the informational content of prices, as in the cursed equilibrium model of Eyster and Rabin (2005). In general, belief disagreements among investors can be interpreted as modeling departures from full rationality in information processing.

Trading motives Summing up, there are four motives to trade in this model:

(i) Different hedging needs: captured by $\mathbb{C}ov\left[M_{2i}, D\right]$ (fundamental)

(ii) Different risk aversion: captured by A_i (fundamental)

(iii) Different initial asset holdings: captured by X_{0i} (fundamental)

(iv) Different beliefs: captured by $\mathbb{E}_i\left[D\right]$ (non-fundamental)

The first three correspond to fundamental motives for trading: sharing risks among investors, transferring risks to those more willing to bear them, or trading for life cycle or liquidity needs. Trading on different beliefs is the single source of non-fundamental trading in the model. All four trading motives can equally determine the positive properties of the model: the assumed welfare criterion makes the last trading motive non-fundamental. Having multiple sources of fundamental trading, while not necessary, is important to show that all fundamental trading motives enter symmetrically in optimal tax formulas. I assume throughout that all four cross-sectional distributions have bounded moments and that the cross-sectional dispersion of risk aversion coefficients is small.

At times, to sharpen several results, I impose the following symmetry assumption on the cross-sectional joint distribution of primitives. I explicitly state when Assumption [S] is used in the paper.[9]

Assumption. [S] **(Symmetry)** *Investors have identical preferences: $A_i = A$, $\forall i$. The cross-sectional distribution of the following linear combination of investors' mean beliefs, hedging needs, and initial asset holdings is symmetric:* $\mathbb{E}_i\left[D\right] - A\mathbb{C}ov\left[M_{2i}, D\right] - A\mathbb{V}ar\left[D\right] X_{0i}$.

Assumption [S] simplifies the solution of the model by making the equilibrium price independent of the tax rate, which allows for sharper characterizations. As it will become clear in Section 5, this assumption does not restrict the levels of fundamental trading, non-fundamental trading, or the cross-sectional correlation between fundamental and non-fundamental trading motives.

Policy instrument: a linear financial transaction tax This paper follows the Ramsey approach of solving for an optimal policy under a restricted set of instruments. The single policy instrument available to the planner is an anonymous linear financial transaction tax τ paid per dollar traded in the risky asset. A change in the net asset holdings of the risky asset of $|X_{1i} - X_{0i}|$ shares at a price P_1 faces a total tax in dollars terms, due at the time the transaction occurs, for both buyers and sellers, of

$$\tau P_1 \left|\Delta X_{1i}\right|, \tag{2}$$

where $|\Delta X_{1i}| \equiv |X_{1i} - X_{0i}|$. The total tax revenue generated by the transaction is thus $2\tau P_1 \left|\Delta X_{1i}\right|$. The tax rate τ can in principle take any value on the extended real line $\overline{\mathbb{R}} = [-\infty, \infty]$. Consequently, investors may face negative taxes, i.e., subsidies.

In the Appendix, I discuss in detail how investors' portfolio decisions change when facing a subsidy instead of a positive tax. I also formally show there that trading subsidies can be implemented when paid on the net change of asset holdings over a given period, but cannot be implemented when paid on every purchase or sale.

Linearity, anonymity, and enforcement I restrict the analysis to linear taxes with the intention of being realistic. The conventional justification for the use of linear (as opposed to non-linear) taxes in this environment is that linear taxes are the most robust to sophisticated trading schemes. For example, a constant tax per trade creates incentives to submit a single large order. Alternatively, quadratic taxes create incentives to split orders

[9] A probability distribution is said to be symmetric if and only if there exists a value μ such that $f(\mu - x) = f(\mu + x)$, $\forall x$, where $f(\cdot)$ denotes the p.d.f. of the distribution.

into infinitesimal pieces. These concerns, which are shared with other non-linear tax schemes, are particularly relevant for financial transaction taxes, given the high degree of sophistication of many players in financial markets and the negligible costs of splitting orders given modern information technology.

I assume that transaction taxes must apply across-the-board to all market participants and cannot be conditioned on individual characteristics, which implies that the planner's problem is a second-best problem. A planner with the ability to distinguish good trades from bad trades could achieve the first-best outcome by taxing investors perceived to engage in welfare-reducing trades on an individual basis: this is a highly implausible scenario.

Furthermore, I assume that investors cannot avoid paying transaction taxes, either by trading secretly or by moving to a different exchange. This behavior is optimal when the penalties associated with evasion are sufficiently large, provided the taxable event is appropriately defined. I discuss the implications of imperfect tax enforcement for the optimal tax policy in Section 5 and in the Appendix.

Revenue rebates and welfare aggregation Lastly, since this paper focuses on the corrective (Pigouvian) effects of transaction taxes and not on the ability of this tax to raise fiscal revenue, I assume that tax proceeds are rebated lump-sum to investors.[10] All results in the paper are derived under an arbitrary rule for tax rebates across investors. Formally, an investor indexed by i receives a rebate T_{1i} that simply must satisfy budget balance on the aggregate: $\int T_{1i} dF(i) = \int \tau P_1 |\Delta X_{1i}| \, dF(i)$. When needed, I consider either a uniform rebate rule, in which all investors receive the same transfer, or an individually targeted rebate rule, in which each investor i receives a transfer equal to the tax liability paid by that investor. The uniform rebate rule respects the anonymity assumption, while the individually targeted rule mutes the redistributional effects of the policy and is useful for theoretical purposes. Since investors are small, they never internalize the impact of their actions on the rebate they receive. It is important that tax revenue is rebated and not wasted.[11]

Finally, until I revisit this issue in Section 6, I assume that the planner seeks to maximize the sum of investors' certainty equivalents, which is a conventional approach in normative problems. Assuming that the planner has access to lump-sum transfers to redistribute wealth across investors ex-ante, separating efficiency from distributional considerations, yields identical results. The Appendix includes a detailed discussion of this welfare aggregation approach and how it relates to other approaches.

Investors' budget/wealth accumulation constraint The consumption/wealth of a given investor i at date 2 consists of the stochastic endowment M_{2i}, the stochastic payoff of the risky asset $X_{1i}D$, and the return on the investment in the risk-free asset. This includes the proceeds from the net purchase/sale of the risky asset $(X_{0i} - X_{1i}) P_1$, the total tax liability $-\tau P_1 |\Delta X_{1i}|$, and the lump-sum transfer T_{1i}. It can be expressed as

$$W_{2i} = M_{2i} + X_{1i}D + X_{0i}P_1 - X_{1i}P_1 - \tau P_1 |\Delta X_{1i}| + T_{1i}. \tag{3}$$

[10] Broadly defined, there are two types of taxes: those levied with the aim of raising revenue and those levied with the aim of correcting distortions. This paper exclusively studies corrective taxation. Sandmo (1975) shows that corrective taxes and optimal revenue raising taxes are additive; see also Kopczuk (2003). This paper does not consider the additional benefits of corrective taxes generated by "double-dividend" arguments. Those arguments, surveyed by Goulder (1995) in the context of environmental taxation, apply directly to transaction taxes. Biais and Rochet (2020) have recently studied the desirability of transaction taxes as a revenue-raising instrument.

[11] Public debates surrounding transaction taxes often discuss how to spend tax revenues. Barring political economy considerations, it should be clear that the problem of how to spend tax revenues is orthogonal to the problem of characterizing the optimal tax. However, given their importance for public debates, I illustrate the behavior of tax revenue in the context of the quantitative assessment of the model in Section 5.

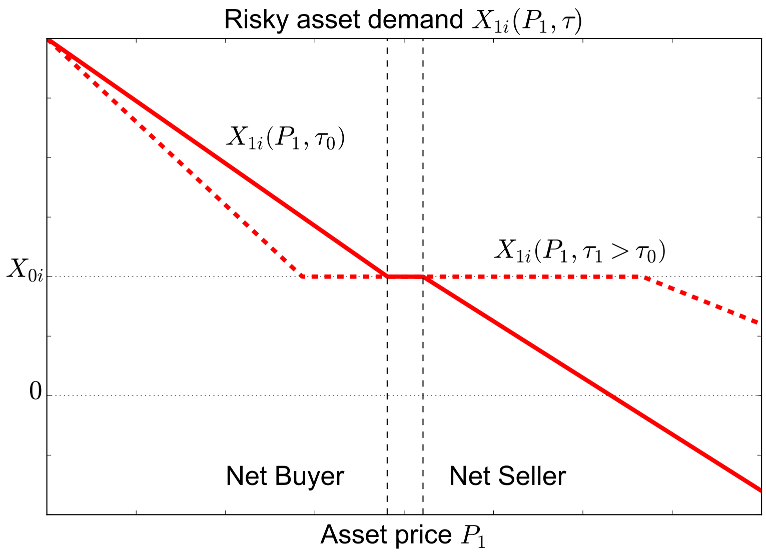


Figure 1: Risky asset demand

Note: Figure 1 illustrates the optimal portfolio demand $X_{1i}(P_1)$, characterized in Equation (5), for an investor i as a function of the asset price P_1 for two tax rates, τ_0 and τ_1, such that $\tau_1 > \tau_0 > 0$. A linear transaction tax $\tau > 0$ is reflected as a higher price $P_1(1+\tau)$ paid by buyers, a lower price $P_1(1-\tau)$ received by sellers, and an inaction region that is increasing in τ, all else equal.

Definition. (Equilibrium) A competitive equilibrium with taxes is defined as a portfolio allocation X_{1i}, $\forall i$, a price P_1, and a set of lump-sum transfers T_{1i}, $\forall i$, such that: i) given the price P_1, each investor i finds the allocation X_{1i} optimal by maximizing expected utility subject to the corresponding budget/wealth accumulation constraint, respectively introduced in Equations (1) and (3); ii) the price P_1 is such that the market for the risky asset clears, that is, $\int \Delta X_{1i} dF(i) = 0$; and iii) tax revenues are rebated lump-sum to investors, so that $\int T_{1i} dF(i) = \int \tau P_1 |\Delta X_{1i}| dF(i)$.

3 Equilibrium

I initially solve for investors' optimal portfolio decisions. Subsequently, I characterize the equilibrium price and allocations.

Investors' problem In this model, investors effectively choose their risky asset demand to maximize the certainty equivalent of their expected future wealth. Formally, investor i's risky asset demand is given by the solution to the following mean-variance problem:

$$\max_{X_{1i}} \left[\mathbb{E}_i[D] - A_i \mathbb{C}ov[M_{2i}, D] - P_1\right] X_{1i} + P_1 X_{0i} - \tau P_1 |\Delta X_{1i}| - \frac{A_i}{2} \mathbb{V}ar[D] (X_{1i})^2 . \tag{4}$$

As formally shown in the Appendix, the problem solved by investors is well-behaved. Given a price P_1 and a tax rate $\tau > 0$, investor i's optimal net asset demand $\Delta X_{1i}(P_1) = X_{1i}(P_1) - X_{0i}$ is given by

$$\Delta X_{1i}(P_1) = \begin{cases} \Delta X_{1i}^{+}(P_1) = \frac{\mathbb{E}_i[D] - A_i \mathbb{C}ov[M_{2i}, D] - P_1(1+\tau)}{A_i \mathbb{V}ar[D]} - X_{0i}, & \text{if } \Delta X_{1i}^{+}(P_1) > 0 \quad \text{Buying} \\ 0\,, & \text{if } \Delta X_{1i}^{+}(P_1) \leq 0,\ \Delta X_{1i}^{-}(P_1) \geq 0 \quad \text{No Trade} \\ \Delta X_{1i}^{-}(P_1) = \frac{\mathbb{E}_i[D] - A_i \mathbb{C}ov[M_{2i}, D] - P_1(1-\tau)}{A_i \mathbb{V}ar[D]} - X_{0i}, & \text{if } \Delta X_{1i}^{-}(P_1) < 0 \quad \text{Selling.} \end{cases} \tag{5}$$

Figure 1 illustrates the optimal portfolio demand $X_{1i}(P_1)$ for an investor i as a function of the asset price P_1. The presence of linear transaction taxes modifies the optimal portfolio allocation along two dimensions. First, a transaction tax is reflected as a higher price $P_1(1+\tau)$ paid by buyers and a lower price $P_1(1-\tau)$ received by sellers. Hence, for a given price P_1, a higher tax reduces the net demand of both buyers and sellers at the intensive margin.

Second, a linear tax implies that some investors decide not to trade altogether, creating an inaction region. If the initial holdings of the risky asset X_{0i} are not too far from the optimal allocation without taxes $\frac{\mathbb{E}_i[D]-A_i\mathbb{C}ov[M_{2i},D]-P_1}{A_i\mathbb{V}ar[D]}$, an investor decides not to trade. Only when $\tau=0$ the no-trade region ceases to exist. The envelope theorem, which plays an important role when deriving the optimal tax results, is also key to generating the inaction region, as originally shown in Constantinides (1986). Intuitively, an investor with initial asset holdings close to his optimum experiences a second-order gain from a marginal trade but suffers a first-order loss when a linear tax is present, making no-trade optimal.[12]

Equilibrium characterization Given the optimal portfolio allocation characterized in Equation (5) and the market clearing condition $\int \Delta X_{1i}(P_1)\,dF(i)=0$, the equilibrium price of the risky asset satisfies the following implicit equation for P_1:

$$P_1=\frac{\int_{i\in\mathcal{T}(P_1)}\left(\frac{\mathbb{E}_i[D]}{\mathcal{A}_i}-A\left(\mathbb{C}ov\left[M_{2i},D\right]+\mathbb{V}ar\left[D\right]X_{0i}\right)\right)dF(i)}{1+\tau\left(\int_{i\in\mathcal{B}(P_1)}\frac{1}{\mathcal{A}_i}dF(i)-\int_{i\in\mathcal{S}(P_1)}\frac{1}{\mathcal{A}_i}dF(i)\right)},\tag{6}$$

where $A\equiv\left(\int_{i\in\mathcal{T}(P_1)}\frac{1}{A_i}dF(i)\right)^{-1}$ is the harmonic mean of risk aversion coefficients for active investors and $\mathcal{A}_i\equiv\frac{A_i}{A}$ is the quotient between the risk aversion coefficient of investor i and the harmonic mean.[13] The notation $i\in\mathcal{T}(P_1)$ indicates that the domain of integration is the set of investors who actively trade in equilibrium at a price P_1. Analogously, $\mathcal{B}(P_1)$ and $\mathcal{S}(P_1)$ respectively denote the sets of buyers and sellers at a given price P_1. Equation (5) determines the identity of the investors in each of the sets. Because the sets $\mathcal{T}(P_1)$, $\mathcal{B}(P_1)$, and $\mathcal{S}(P_1)$, as well as A and $\mathcal{A}_i$, depend on the equilibrium price, Equation (6) provides an implicit characterization of P_1. Intuitively, only marginal investors directly determine the equilibrium price. As shown in Lemma 1 below, Equation (6) has a unique solution for P_1 whenever there is trade in equilibrium.

The numerator of the equilibrium price in Equation (6) has two components. The first term is a weighted average of the expected payoff of the risky asset. The second term is a risk premium, determined by the product of price and quantity of risk. The price of risk is given by the harmonic mean of risk aversion coefficients A. The quantity of risk consists of two terms. The first one is the sum of covariances of the risky asset with the endowments $\int_{i\in\mathcal{T}(P_1)}\mathbb{C}ov\left[M_{2i},D\right]dF(i)$. The second one is the product of the variance of the risky asset $\mathbb{V}ar\left[D\right]$ with the number of shares initially held by investors $\int_{i\in\mathcal{T}(P_1)}X_{0i}dF(i)$.

Trading volume is another relevant equilibrium object. I denote trading volume, measured in shares of the risky asset and expressed as a function of the tax rate, by $\mathcal{V}(\tau)$. Trading volume formally corresponds to

$$\mathcal{V}(\tau)=\int_{i\in\mathcal{B}(\tau)}\Delta X_{1i}(\tau)\,dF(i),\tag{7}$$

where $\Delta X_{1i}(\tau)$ denotes equilibrium net trades for a given tax rate τ and where only the net trades of buyers are considered, to avoid double counting. At times, it is useful to compute asset turnover, which expresses trading volume as a function of the total number of shares Q. Formally, turnover is given by $\Xi(\tau)=\frac{\mathcal{V}(\tau)}{Q}$.

Lemma 1 synthesizes the main positive results of the model. Lemma 1 shows that the model is well-behaved and that a transaction tax is a robust instrument to reduce trading volume. More broadly, Lemma 1 implies that

[12] If taxes were quadratic, the marginal welfare loss induced by the tax around the optimum would also be second-order, eliminating the inaction region.

[13] It should be clear from their definitions that $\mathcal{A}_i$ and A are also functions of P_1 and τ through the set of active investors.

theories in which transaction taxes are desirable must rely on a mechanism through which reducing trading volume is welfare improving.[14]

Lemma 1. (Competitive equilibrium with taxes)

a) [Existence/Uniqueness] An equilibrium always exists for a given τ. The equilibrium is (essentially) unique.

b) [Volume response] Trading volume is decreasing in τ.

c) [Price response] The asset price P_1 increases (decreases) with τ if

$$\int_{i\in\mathcal{B}(P_1)} \frac{1}{\mathcal{A}_i} dF(i) \leq (\geq) \int_{i\in\mathcal{S}(P_1)} \frac{1}{\mathcal{A}_i} dF(i) \,. \tag{8}$$

Under Assumption [S], the asset price P_1 is invariant to the level of the transaction tax.

Lemma 1 shows that an equilibrium always exists and that it is essentially unique. Equilibrium existence is effectively guaranteed because risky asset demands are everywhere downward sloping. Equilibrium portfolio allocations and trading volume are uniquely pinned down in any equilibrium. The equilibrium price P_1 is also uniquely pinned down in any equilibrium with positive trading volume. Every no-trade equilibrium is associated with a range of prices consistent with such an equilibrium. There are also many sets of individual lump-sum transfers consistent with any unique equilibrium allocation. Therefore, because of both dimensions of indeterminacy, I refer to the equilibrium as essentially unique.

Trading volume always goes down when transaction taxes increase. Even though a change in the transaction tax can change the asset price and indirectly induce some sellers to sell more or some buyers to buy more, this effect is never strong enough to overcome the direct effect of the tax, which always discourages trading.

The condition that determines the sign of $\frac{dP_1}{d\tau}$ in Equation (8) corresponds to the difference between the aggregate buying and selling price elasticities. When this term is positive, increasing τ reduces the buying pressure by more than the selling pressure, reducing the equilibrium price, and vice versa — see the Appendix for a simulation of the model that illustrates these effects. When the difference between aggregate buying and selling elasticities is zero, the equilibrium price is independent of the tax.[15] In particular, for the symmetric benchmark in which Assumption [S] holds, aggregate buying and selling price elasticities are everywhere identical, implying that the equilibrium price is invariant to the tax rate.

4 Normative analysis

After solving for the equilibrium allocations and the equilibrium price for a given tax, I first introduce the welfare criterion used by the planner to compute social welfare and then characterize the optimal tax policy.

[14] Existing empirical evidence is consistent with the prediction that trading volume decreases after an increase in transaction taxes/costs, although tax evasion may be at times a confounding factor. The empirical evidence regarding the effect of transaction taxes/costs on prices is mixed. Some studies find an increase in price volatility, but others find no significant change or even a reduction. Asset prices usually fall at impact following a tax increase, but seem to recover over time. See the review articles by Campbell and Froot (1994), Habermeier and Kirilenko (2003), McCulloch and Pacillo (2011), Burman et al. (2016), and the recent work on the European Transaction Tax by Colliard and Hoffmann (2017) and Coelho (2014).

[15] In this model, an increase in the asset price is automatically associated with a reduction in the risk premium, and vice versa. Since the asset price can increase, decrease, or remain constant in this model, it is straightforward to conclude that price volatility can increase, decrease, or remain constant in a dynamic extension of this model.

4.1 Welfare criterion

In order to aggregate individual preferences, I assume that the planner maximizes the sum of investors' certainty equivalents.[16] However, to conduct any normative analysis in this paper, one must also take a stand on how to evaluate social welfare when investors hold heterogeneous beliefs, which is a controversial issue.[17]

In this paper, the planner computes social welfare as follows. In any dimension in which investors' beliefs agree, I assume that the planner shares the investors' beliefs. Whenever investors disagree, I assume that the planner computes investors' welfare (certainty equivalents) using a single belief. Hence, when many investors disagree, the belief used by the planner will be necessarily different from the beliefs held by most investors. Given a planner's belief, I follow a two-step approach. First, I characterize the optimal tax policy for a given planner's belief. Subsequently, I identify the conditions under which the optimal policy does not depend on the planner's belief. In those cases, only the consistency requirement that investors' welfare be computed using a single common belief is relevant.

The two-step normative approach used in this paper can be applied more generally. In every normative problem with belief heterogeneity among investors it is possible to first characterize the solution to a planning problem for a given planner's belief and then seek to find conditions under which the optimal policy is independent of the planner's belief. When an optimal policy independent of the planner's belief cannot be found, the results simply characterize the optimal paternalistic policy.

This approach is paternalistic, because it ignores investors' subjective beliefs when finding the optimal policy. However, when the optimal policy is independent of the specific belief chosen by the planner, conventional criticisms of paternalistic policies on the grounds that the planner must be better informed than the individuals in the economy do not apply. Although overruling investors' beliefs when computing welfare creates a mechanical rationale for intervention, the welfare impact of belief distortions, the sign and magnitude of the optimal intervention, as well as the informational requirements needed to implement the optimal policy are far from obvious, as shown in this paper.

Two arguments support the welfare criterion adopted in this paper. First, since there is a single distribution of payoffs, but different investors hold different beliefs about such a distribution, all of them (but one) must be wrong. In that case, it may be reasonable to argue that a planner need not respect investors' beliefs when they are almost surely incorrect. Alternatively, a veil of ignorance interpretation can also support the welfare criterion used in this paper. If investors acknowledge that they may wrongly hold different beliefs when trading, they would be willing to implement ex-ante a tax policy that corrects their trading behavior.[18]

After presenting the main results of the paper in Proposition 1, I explain how the welfare criterion introduced in this paper relates to those proposed by Brunnermeier, Simsek and Xiong (2014) and Gilboa, Samuelson and Schmeidler (2014) in Section 4.3 and in the Appendix.

[16] If the planner has access to ex-ante transfers, the sum of investors' certainty equivalents is the appropriate measure of welfare for any set of welfare weights. If the planner does not have access to ex-ante transfers, maximizing the sum of investors' certainty equivalents can be interpreted as selecting an equal-weighted set of "generalized social marginal welfare weights" (Saez and Stantcheva, 2016), which in turn can be mapped to traditional social welfare weights, as described in the Appendix.

[17] In addition to the work discussed in the literature review, see Kreps (2012), Cochrane (2014), and Duffie (2014) for some reflections on this topic. Duffie (2014), in particular, poses both philosophical/axiomatic challenges and a practical challenge to policy treatments of speculative trading motivated by differences in beliefs. This paper provides an explicit solution to the practical challenge raised in that paper, which questions the ability of enforcement agencies to set policies when some trades are belief-motivated while other trades arise from welfare enhancing activities.

[18] The Appendix includes a formal discussion of how altruistic investors in this model always prefer an optimal tax of the same sign as the optimal tax chosen by the planner.

4.2 Optimal transaction tax

After introducing the welfare criterion used by the planner, I characterize the properties of the optimal tax policy. The planner's objective is a function of the investors' certainty equivalents from the planner's perspective. The certainty equivalent of investor i from the planner's perspective, denoted by $V_i^p(\tau)$, corresponds to

$$V_i^p(\tau) \equiv \left(\mathbb{E}_p[D] - A_i \mathbb{C}ov[M_{2i}, D] - P_1(\tau)\right) X_{1i}(\tau) + P_1(\tau) X_{0i} - \frac{A_i}{2} \mathbb{V}ar[D] \left(X_{1i}(\tau)\right)^2 + \tilde{T}_{1i}(\tau), \tag{9}$$

where $X_{1i}(\tau)$ and $P_1(\tau)$ represent equilibrium outcomes that are in general functions of τ and $\tilde{T}_{1i}(\tau) = T_{1i}(\tau) - \tau P_1(\tau) |\Delta X_{1i}(\tau)|$ denotes the net transfer received by investor i. The assumed welfare criterion implies that the expectation of the payoff of the risky asset used to calculate $V_i^p(\tau)$ does not have an individual subscript i because it is computed using the planner's belief, denoted by $\mathbb{E}_p[D]$.[19]

Social welfare, denoted by $V^p(\tau)$, corresponds to the sum of investors' certainty equivalents and is formally given by

$$V^p(\tau) = \int V_i^p(\tau) \, dF(i).$$

The optimal tax corresponds to $\tau^* = \arg\max_\tau V^p(\tau)$, where τ must lie in the extended real line $[-\infty, +\infty]$. Before characterizing the optimal tax, it is worth finding the marginal welfare impact of a tax change on investor i and on the aggregate from the planner's perspective.

Lemma 2, which formally introduces both, is helpful to understand the form of the optimal tax policy, characterized in Proposition 1 and illustrated in Figures 2 and 3 below.

Lemma 2. (Marginal welfare impact of tax changes)

a) [Individual welfare impact] The individual marginal welfare impact of a change in the tax rate from the planner's perspective is given by

$$\frac{dV_i^p}{d\tau} = \left[\mathbb{E}_p[D] - \mathbb{E}_i[D] + \operatorname{sgn}\left(\Delta X_{1i}(\tau)\right) P_1(\tau) \tau\right] \frac{dX_{1i}(\tau)}{d\tau} - \Delta X_{1i}(\tau) \frac{dP_1(\tau)}{d\tau} + \frac{d\tilde{T}_{1i}(\tau)}{d\tau}, \tag{10}$$

where $\operatorname{sgn}(\cdot)$ *denotes the sign function.*

b) [Aggregate welfare impact] The aggregate marginal welfare impact of a change in the tax rate from the planner's perspective is given by

$$\frac{dV^p}{d\tau} = \int_{i \in \mathcal{T}(\tau)} \left[-\mathbb{E}_i[D] + \operatorname{sgn}\left(\Delta X_{1i}(\tau)\right) P_1(\tau) \tau\right] \frac{dX_{1i}(\tau)}{d\tau} dF(i), \tag{11}$$

where $\mathcal{T}(\tau)$ *denotes the set of active investors for a given tax rate* τ.

The individual marginal welfare impact of a tax change, which is expressed in dollars, features three terms. The first term in Equation (10) captures the impact of a tax change on an investor's portfolio allocation. A change in the allocation $\frac{dX_{1i}}{d\tau}$ only affects welfare through the wedges in investors' portfolio demands perceived by the planner. A first wedge arises when $\mathbb{E}_p[D] \neq \mathbb{E}_i[D]$, through the difference in beliefs between the planner and an investor i. If the planner computed welfare respecting investors' beliefs, the envelope theorem would guarantee that this wedge is exactly zero. A second wedge arises because investors face the transaction tax at the margin. The second term in Equation (10) captures distributive pecuniary effects — using the terminology of Dávila and Korinek (2018). If P_1 increases with τ, the buyers (sellers) of the risky asset are worse (better) off, since the terms-of-trade of their transaction have worsened (improved). The opposite occurs when P_1 decreases with τ. The third term in Equation (10) simply accounts for the change in investor i's tax rebate net of the tax liability.

[19]The planner and all investors agree on the second moments of the distribution of asset payoffs. The Appendix includes an extension of the model in which investors disagree about the second moments of the distribution of payoffs.

Equation (11) shows that three elements cancel out after aggregating the individual welfare effects. First, and crucially for the results in this paper, the planner's belief $\mathbb{E}_p[D]$ drops out after the aggregation step. Intuitively, a planner with a very high (low) $\mathbb{E}_p[D]$ may find it desirable for all investors to hold more (less) shares of the risky asset. However, this is not possible in equilibrium: market clearing implies that the portfolio changes induced by a tax change must add up to zero on the aggregate; formally, $\int \frac{dX_{1i}(\tau)}{d\tau} dF(i) = 0$. Second, the distributive pecuniary effects cancel out, as in any competitive model. Finally, because all tax revenues are rebated to investors, the net government transfers also add up to zero. Consequently, the aggregate marginal welfare impact of a tax change will simply depend on the distribution of investors' beliefs and on the way in which taxes impact investors' portfolio allocations.

Proposition 1 introduces the main results of the paper. I first present Proposition 1 and then elaborate on each of its results below.

Proposition 1. (Optimal financial transaction tax)

a) [Optimal tax formula] The optimal financial transaction tax τ^* satisfies

$$\tau^* = \frac{\Omega_{\mathcal{B}(\tau^*)} - \Omega_{\mathcal{S}(\tau^*)}}{2}, \tag{12}$$

where $\Omega_{\mathcal{B}(\tau)}$ is a weighted average of buyers' expected returns, given by

$$\Omega_{\mathcal{B}(\tau)} \equiv \int_{i\in\mathcal{B}(\tau)} \omega_i^{\mathcal{B}}(\tau) \frac{\mathbb{E}_i[D]}{P_1(\tau)} dF(i), \quad \text{with} \quad \omega_i^{\mathcal{B}}(\tau) \equiv \frac{\frac{dX_{1i}(\tau)}{d\tau}}{\int_{i\in\mathcal{B}(\tau)} \frac{dX_{1i}(\tau)}{d\tau} dF(i)}, \tag{13}$$

and $\Omega_{\mathcal{S}(\tau)}$ is a weighted average of sellers' expected returns, analogously defined.

b) [Sign of the optimal tax] A positive tax is optimal when optimistic investors are net buyers and pessimistic investors are net sellers in the laissez-faire economy. Formally,

$$\text{if} \quad \left.\frac{dV^p}{d\tau}\right|_{\tau=0} = \mathbb{C}ov_F\left(\mathbb{E}_i[D], -\left.\frac{dX_{1i}}{d\tau}\right|_{\tau=0}\right) > 0, \text{ then } \tau^* > 0, \tag{14}$$

where $\mathbb{C}ov_F(\cdot,\cdot)$ denotes a cross-sectional covariance. As long as some investors have heterogeneous beliefs and fundamental and non-fundamental trading motives are independently distributed across the population of investors, this condition is endogenously satisfied, implying that the optimal corrective policy is a strictly positive tax.[20]

c) [Irrelevance of planner's belief] The optimal financial transaction tax does not depend on the belief used by the planner to calculate welfare.

Optimal tax formula Proposition 1a) shows that the optimal tax formula can be written exclusively as a function of investors' beliefs, $\mathbb{E}_i[D]/P_1$, and portfolio sensitivities, $\frac{dX_{1i}}{d\tau}$. Because the equilibrium price, portfolio sensitivities, and identity of the active investors are endogenous to the level of the tax, Equation (12) only provides an implicit representation for τ^*. This is a standard feature of optimal taxation exercises. Below, I provide conditions under which the optimal tax formula has a unique solution.

The corrective (Pigouvian) nature of the tax explains why investors' beliefs and portfolio sensitivities are the relevant variables that determine the optimal tax. Pigouvian logic suggests that corrective taxes must be set to target marginal distortions, which in this particular case arise from investors' differences in mean beliefs about asset payoffs. Ideally, the planner would like to target each individual belief distortion with an investor-specific tax.[21] However, because the planner employs a second-best policy instrument — a single linear tax — the portfolio

[20] Formally, fundamental and non-fundamental trading motives are independently distributed across the population of investors when the distributions of $\mathbb{C}ov[M_{2i}, D]$, A_i, X_{0i}, and $\mathbb{E}_i[D]$ are independent.

[21] See the Appendix for a characterization of the first-best policy with unrestricted instruments, which calls for investor-specific corrective policies.

sensitivities $\frac{dX_{1i}}{d\tau}$ determine the weights given to individual beliefs in the optimal tax formula. The planner gives more weight to the distortions of the most tax-sensitive investors.[22] Note that the weights assigned to buyers $\omega_i^{\mathcal{B}}$ and sellers $\omega_i^{\mathcal{S}}$ add up to one respectively and that investors who do not trade do not affect the optimal tax at the margin.

When Assumption [S] holds, the optimal tax satisfies the simpler condition

$$\tau^* = \frac{\mathbb{E}_{\mathcal{B}(\tau^*)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right] - \mathbb{E}_{\mathcal{S}(\tau^*)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]}{2}, \tag{15}$$

where $\mathbb{E}_{\mathcal{B}(\tau^*)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]$ and $\mathbb{E}_{\mathcal{S}(\tau^*)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]$ respectively denote the cross-sectional average of expected returns of buyers and sellers at the optimal tax rate τ^*. In this case, portfolio sensitivities drop out of the optimal tax formula, providing a tractable benchmark in which the optimal tax is exclusively a function of the average beliefs of active buyers and sellers.

If all investors agree about the expected payoff of the risky asset, so that $\mathbb{E}_i[D]$ is constant, the optimal tax is $\tau^* = 0$. Equations (12) and (15) suggest that an increase in the dispersion of beliefs across investors, by widening the gap between buyers' and sellers' expected returns, calls for a higher optimal transaction tax. In Section 5, I explicitly link the value of the optimal tax to primitives of the distribution of fundamental and non-fundamental trading motives.

Sign of the optimal tax Proposition 1b) shows that the optimal policy corresponds to a strictly positive tax when, in the laissez-faire economy, optimistic investors (those with a high $\mathbb{E}_i[D]$) are on average net buyers (those for which $-\frac{dX_{1i}}{d\tau}\big|_{\tau=0} > 0$) of the risky asset, while pessimistic investors are on average net sellers. If all trading is driven by disagreement, Equation (14) trivially holds — optimists buy and pessimists sell. However, because investors may also trade due to fundamental motives, it is possible for an optimistic investor to be a net seller in equilibrium and vice versa. When Assumption [S] holds, Equation (14) simplifies to the more intuitive condition for a positive tax:

$$\text{if} \quad \mathbb{E}_{\mathcal{B}(\tau=0)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right] > \mathbb{E}_{\mathcal{S}(\tau=0)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right], \text{ then } \tau^* > 0,$$

which highlights that identifying the difference in beliefs between buyers and sellers in the laissez-faire economy is sufficient to establish the sign of the tax.

Proposition 1b) not only establishes a necessary condition for the optimal tax to be positive, but it also provides a natural sufficient condition for Equation (14) to be satisfied. Hence, as long as some investors hold heterogeneous beliefs, and if the distribution of beliefs across investors is independent of the distribution of fundamental trading motives (risk aversion, hedging needs, and initial positions), a strictly positive tax is optimal.

Alternatively, one could argue on empirical grounds that Equation (14) holds. The evidence accumulated in the behavioral finance literature, surveyed in Barberis and Thaler (2003) and Hong and Stein (2007), suggests that investors' beliefs drive a non-negligible share of purchases/sales. Intuitively, in expectation, an optimistic (pessimistic) investor is more likely to be a buyer (seller) in equilibrium. Hence, unless the pattern of fundamental trading substantially counteracts this force, it is natural to expect the covariance in Equation (14) to be positive. Independence between fundamental and non-fundamental trading motives is a sufficient condition for an optimal positive tax, but it is not necessary.

This result puts fundamental and non-fundamental trading motives on different grounds when setting the optimal tax. The mere presence of non-fundamental trading motives unrelated to fundamental trading motives

[22] The presence of demand/portfolio sensitivities in optimal corrective tax formulas goes back to Diamond (1973), who analyzes corrective taxation with uniform taxes in a model of consumption externalities. See Dávila and Walther (2021) for a systematic treatment of corrective regulation with imperfect instruments.

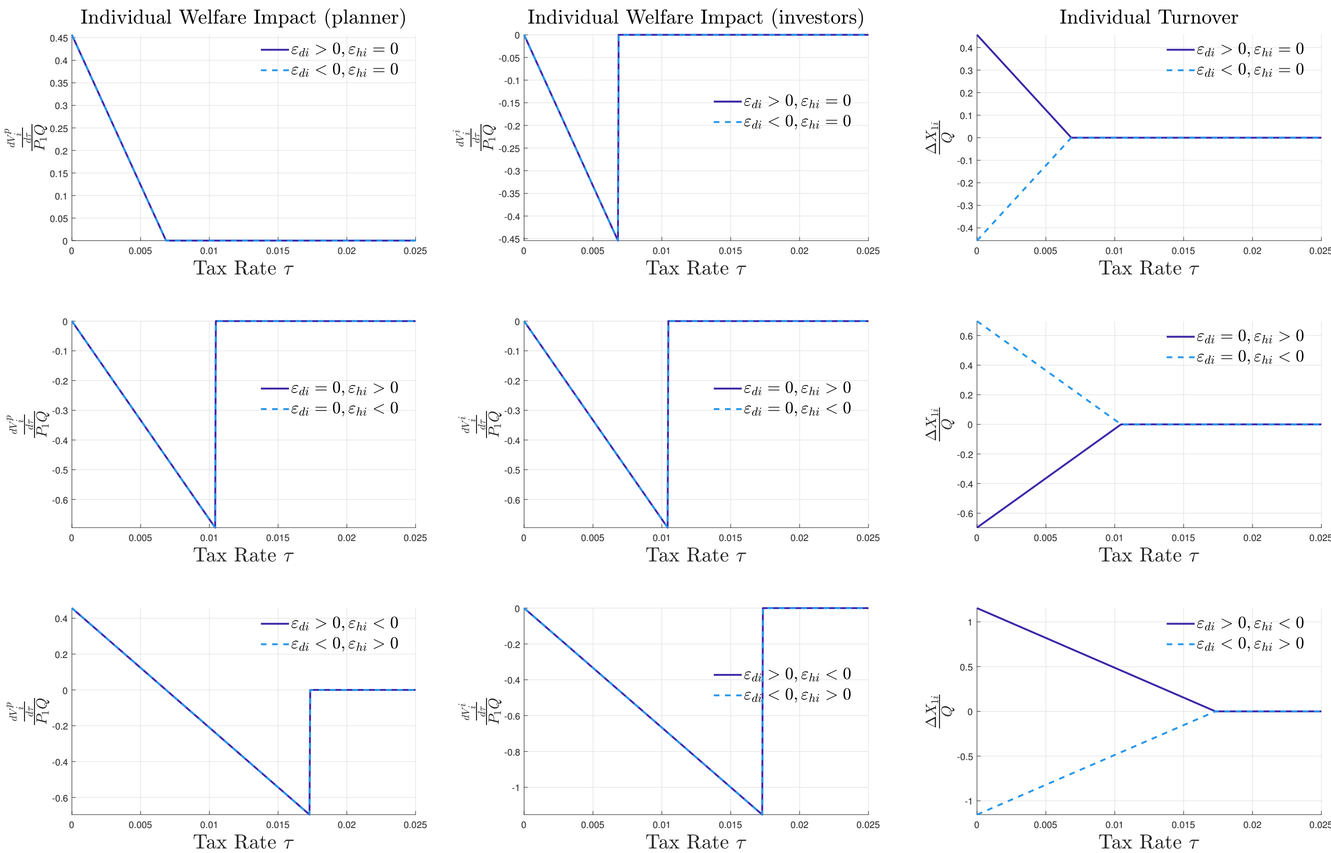


Figure 2: Individual marginal welfare impact (planner's and investors' perspectives)

Note: Figure 2 shows the normalized individual marginal welfare impact of a tax change for specific investors from the perspective of the planner (left plots) and from the perspective of each individual investor (middle plots), for different values of τ. These respectively correspond to Equations (10) and (31). All plots in Figure 2 satisfy Assumption [G], described on page 22, which explicitly defines ε_{di} and ε_{hi}. The right plots show individual asset turnover for specific investors for different values of τ. The top row plots show a buyer ($\varepsilon_{di} > 0$ and $\varepsilon_{hi} = 0$) and a seller ($\varepsilon_{di} > 0$ and $\varepsilon_{hi} = 0$) who only trade for non-fundamental reasons. The middle row plots show a buyer ($\varepsilon_{di} = 0$ and $\varepsilon_{hi} < 0$) and a seller ($\varepsilon_{di} = 0$ and $\varepsilon_{hi} > 0$) who only trade for fundamental reasons. The bottom plots show a buyer ($\varepsilon_{di} > 0$ and $\varepsilon_{hi} < 0$) who buys for non-fundamental and fundamental reasons and a seller ($\varepsilon_{di} < 0$ and $\varepsilon_{hi} > 0$) who sells for non-fundamental and fundamental reasons. The values of ε_{di} and ε_{hi} used correspond to one standard deviation of the distributions of $\frac{\varepsilon_{di}}{P_1}$ and $\frac{\varepsilon_{hi}}{P_1}$, respectively.

All plots use the baseline calibration from Section 5, that is, $\delta^{NF} = 0.3$ (share of non-fundamental trading volume), $\varepsilon_{\tau}^{\log \mathcal{V}}\big|_{\tau=0} = 100$ (laissez-faire semi-elasticity of volume to taxes), and $\Pi = 1.5\%$ (quarterly risk premium). The plots in Figure 2 assume that i) the planner's mean belief $\mathbb{E}_p[D]$ equals the average mean belief μ_d, and ii) an individually targeted rebate rule is implemented. See Figures A.8 and A.10b in the Appendix for an illustration of how the results change after relaxing both assumptions.

implies that it is optimal to have a positive tax, regardless of the relative magnitude of both types of trading motives. That is, a positive tax is optimal in that situation even when most trades are driven by fundamental motives.[23] Intuitively, the planner perceives that a reduction in trading, starting from the laissez-faire equilibrium, generates a first-order welfare gain for those investors who are optimistic buyers and pessimistic sellers. When these are the majority of investors (Equation (14) holds), their first-order welfare gains dominate the second-order welfare losses of those investors who share the planner's belief and the first-order welfare losses of optimistic sellers and pessimistic buyers.

The first column in Figure 2 illustrates this logic. The left three plots in Figure 2 illustrate the (normalized) marginal welfare gain/loss $\frac{dV_i^p}{d\tau}$ for three different types of investors from the planner's perspective. The top left plot shows the welfare impact on a set of investors who trade purely due to non-fundamental motives (belief differences). The planner perceives that increasing the tax rate at $\tau = 0$ is welfare improving for these investors. The middle left plot shows the welfare impact on a set of investors who trade purely for fundamental motives. The planner perceives that increasing the tax rate at $\tau = 0$ does not affect the welfare of these investors up to a first order. The bottom left plot shows the welfare impact on a set of investors who are buyers (sellers) for fundamental motives but who are also optimistic (pessimistic). Similarly to the first case, the planner perceives a positive but smaller gain from increasing the tax rate around $\tau = 0$. In all three cases, the marginal welfare gains from taxation decrease in τ whenever the investors actively trade. Figure 3, which shows the aggregate marginal welfare impact of a tax change, aggregates these effects in dollar terms and illustrates how the planner determines the optimal tax.

The middle column in Figure 2 shows instead the normalized individual marginal welfare impact of a tax change for the same investors when computed using their own beliefs. Leaving aside relative price changes and net transfers, these plots clearly illustrate that every investor perceives himself to be worse off using his own belief to compute his own welfare when facing a positive tax.

In certain circumstances, a trading subsidy rather than a transaction tax could be optimal. If many optimists happen to be sellers of the risky asset in the laissez-faire equilibrium, instead of buyers, the optimal policy may be a subsidy. An example of this trading pattern involves workers who are overoptimistic about their own company's performance and who fail to sufficiently hedge their labor income risk. They are natural sellers of the risky asset, as hedgers, but they sell too little of it. In that case, a transaction tax, by pushing them towards no-trade, has a negative first-order welfare effect. When Equation (14) holds, this phenomenon is not too prevalent among investors.

Irrelevance of planner's belief Proposition 1c) establishes that the optimal tax and, implicitly, its sign, are independent of the belief used by the planner to calculate welfare. This is a surprising and appealing result because, even though the planner does not respect investors' beliefs when assessing welfare, aggregate welfare assessments and the optimal tax policy do not depend on the planner's belief, but only on the consistency condition that there exists a single common payoff distribution.

This result holds for three reasons. First, the marginal impact of a tax change on a given investor certainty equivalent from the planner's perspective is a linear function of $\mathbb{E}_p[D] \frac{dX_{1i}}{d\tau}$. This is an exact property of CARA/mean-variance preferences that holds approximately for general utility specifications — see Proposition 5. Intuitively, the expected payoff captures the first-order welfare gain/loss of changing a portfolio allocation.

Second, the risky asset is in fixed supply, which implies that if one investor holds more shares of the risky asset, some other investor must be holding fewer shares. Formally, it is essential that $\int \frac{dX_{1i}}{d\tau} dF(i) = 0$. In that case,

[23] Proposition 2 provides an explicit decomposition of total trading volume into fundamental and non-fundamental trading volume.

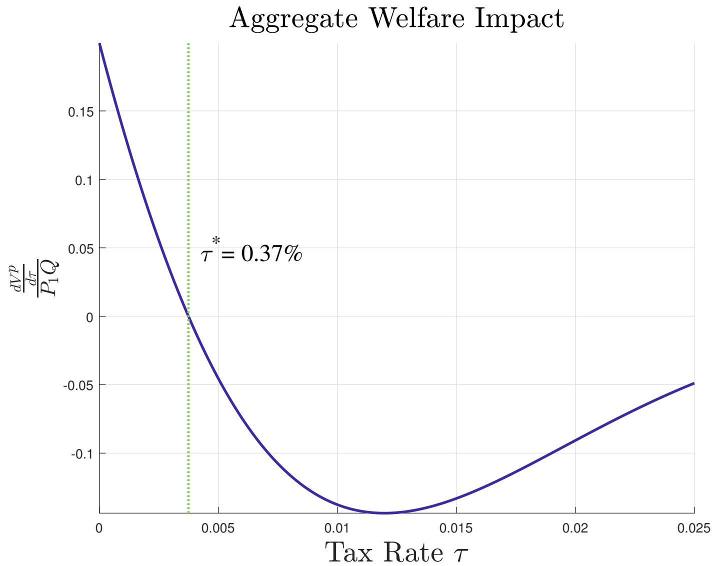


Figure 3: Aggregate marginal welfare impact

Note: Figure 3 shows the normalized aggregate marginal welfare impact of a tax change, $\frac{\frac{dV^p}{d\tau}}{P_1 Q}$, defined in Equation (11), for different values of τ. Figure 3 satisfies Assumption [G], described on page 22, and uses the baseline calibration from Section 5, that is, $\delta^{NF} = 0.3$ (share of non-fundamental trading volume), $\left.\varepsilon_{\tau}^{\log \mathcal{V}}\right|_{\tau=0} = 100$ (laissez-faire semi-elasticity of volume to taxes), and $\Pi = 1.5\%$ (quarterly risk premium). The optimal tax, represented by a vertical dotted line, is $\tau^* = 0.37\%$.

only relative asset holdings matter for welfare. Intuitively, the key economic outcome of this model corresponds to the allocation of risk among investors, which is determined by the dispersion on investors' beliefs, but not by the average belief.

Third, the planner does not use the transaction tax with the purpose of redistributing resources across investors, giving equal weight to the welfare gains/losses across investors in dollar terms. Intuitively, this allows the planner to add up certainty equivalents. When combined, these three features imply that $\frac{dV^p}{d\tau}$ is a function of $\mathbb{E}_p [D] \int \frac{dX_{1i}}{d\tau} dF(i)$, which is exactly 0, via market clearing.

Finally, note that these three arguments, and consequently the result of the irrelevance of the planner's belief, do not depend on the exact form of the policy (an anonymous linear financial transaction tax). As shown in the Appendix, if the planner has investor-specific taxes, different taxes on buyers and sellers, or other policy instruments (e.g., a short-sale or a borrowing constraint), these can also be optimally determined independently of the belief used by the planner to calculate welfare.

4.3 Welfare criteria comparison/Non-convexity/Pure betting

Welfare criteria comparison It is useful to compare the welfare criterion used in this paper with the welfare criteria of Brunnermeier, Simsek and Xiong (2014) and Gilboa, Samuelson and Schmeidler (2014), respectively referred to as BSX and GSS in this subsection.

The belief-neutral social welfare criterion proposed by BSX is the closest to the one used in this paper. Their criterion compares two allocations by aggregating investors' welfare using a set of social welfare weights and requiring that the preferred allocation is so for a planner who computes investors' welfare using every convex combination of agents' beliefs. The optimal tax characterized in Proposition 1 selects the best competitive equilibria with taxes according to their belief-neutral social welfare criterion for a specific set of welfare weights.[24] Because the optimal tax is independent of the belief selected by the planner, the restriction that the planner's belief must be in the convex hull of investors' beliefs is automatically satisfied. In fact, as highlighted in Proposition 1c), the optimal tax maximizes social welfare for *any* belief chosen by the planner, not only those in the convex hull of agents' beliefs.[25] Because the policy instrument considered in this paper is a second-best instrument, even though the optimal tax maximizes a belief-neutral social welfare criterion, this does not necessarily correspond with a belief-neutral Pareto efficient allocation. The planner would need additional instruments to implement such an allocation.

The conceptual differences between the no-betting Pareto criterion proposed by GSS and the criterion used in this paper are more significant. Importantly, the criterion in GSS refines the traditional Pareto criterion. For an allocation to no-betting dominate another one, it must be that all investors prefer the former allocation to the latter and that there exists a single belief that, if held by all agents, makes the former allocation be preferred by all agents. Because the welfare criterion used in this paper computes welfare without respecting investors' beliefs, the allocation implemented by the optimal tax will typically fail to no-betting Pareto dominate the laissez-faire allocation, since at least some investors will perceive to be worse off under the optimal tax policy. However, at times, the allocation implemented by the optimal tax policy may no-betting Pareto dominate the no-trade allocation. The Appendix includes several informative examples that illustrate scenarios in which the no-betting Pareto criterion can and cannot rank i) the no-trade allocation, ii) the competitive equilibrium allocation under the optimal tax, and iii) the laissez-faire competitive equilibrium allocation.

Failure of quasi-concavity of planner's objective Although Equation (12) must hold at the tax level that maximizes the planner's objective function, without further restrictions on the distribution of trading motives, the planner's problem may have multiple local optima. I formally show that the planner's objective function is quasi-concave at tax rates in which investors only adjust their trading behavior on the intensive margin. The planner's objective function can only fail to be concave when the composition of investors who actively trade varies with the tax rate. I provide a sufficient condition under which the planner's objective function is quasi-concave, implying that there exists a uniquely optimal tax. I summarize these results in the following Lemma.

Lemma 3. (Failure of quasi-concavity of planner's objective)

a) The planner's objective function may fail to be quasi-concave. Non-concavities can only arise if the composition of active investors varies in response to tax changes.

b) A sufficient condition for the planner's objective to be quasi-concave is that i) Assumption [S] is satisfied, and ii) investors' beliefs satisfy

$$\frac{\partial \mathbb{E}_{\mathcal{B}(\tau)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]}{\partial \tau} - \frac{\partial \mathbb{E}_{\mathcal{S}(\tau)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]}{\partial \tau} < 2, \quad \forall \tau \tag{16}$$

where $\mathbb{E}_{\mathcal{B}(\tau)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]$ *and* $\mathbb{E}_{\mathcal{S}(\tau)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]$ *respectively denote the average expected returns of buyers and sellers for a given tax rate* τ*.*

[24] As explained in the Appendix, maximizing investors' certainty equivalents is identical to using generalized social welfare weights, which in turn can be mapped to conventional social welfare weights, such as those used in the criterion of BSX.

[25] Note that the welfare criterion proposed by BSX can be extended to any set of reasonable beliefs (not just convex combinations of beliefs). They explicitly state that in endowment-economies (as in this paper) the reasonable beliefs can be extended to include any belief. The optimal tax characterized in Proposition 1 still selects the best competitive equilibria with taxes under this extended criterion.

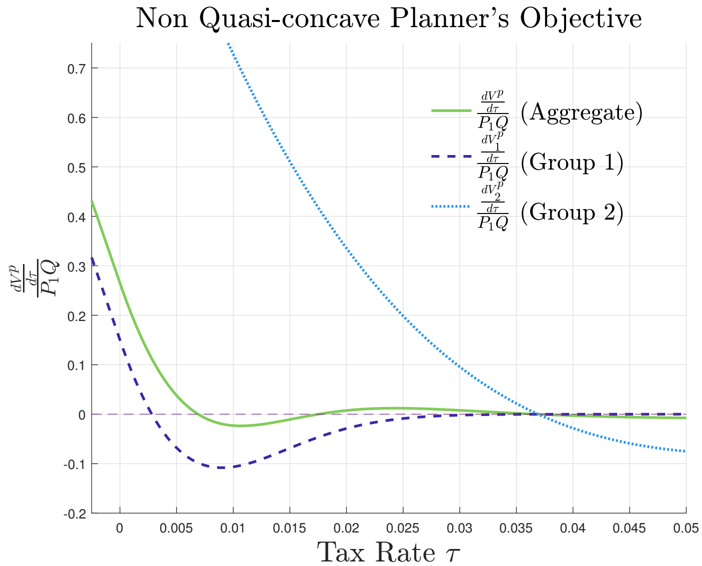


Figure 4: Failure of quasi-concavity of planner's objective

Note: Figure 4 shows the normalized aggregate welfare impact of a tax change from the planner's perspective in a scenario in which the planner's objective fails to be quasi-concave. This figure corresponds to an environment in which there are two groups of investors with the same degree of risk aversion and the same initial asset holdings. 90% of investors belong to group 1, while the remaining 10% belong to group 2. Group 1 investors have turnover of 1/4 and a share of non-fundamental trading volume of 0.3. Group 2 investors have turnover of 1 and a share of non-fundamental trading volume of 0.65. The risk premium is 1.5%. In this example, the planner's objective has three critical points, with two local maxima at $\tau = 0.67\%$ (the globally optimal tax) and $\tau = 3.66\%$, and local minimum at $\tau = 1.72\%$.

Since the set of marginal investors varies with the tax rate, the desirability of varying the tax rate may significantly change depending on which investors remain active. The sufficient condition for quasi-concavity is intuitive: it requires that the difference between buyers' and sellers' expected returns does not grow too fast when the tax increases, which implies that the change in the marginal benefit of increasing the tax level cannot become too large as the tax increases.

Beyond its technical interest, the fact that the planner's objective may have multiple local optima is of economic significance. Figure 4 illustrates this possibility by considering an example in which there are two groups of investors. The first group of investors has low turnover and a low share of non-fundamental trading volume. The second group of investors has high turnover and a high share of non-fundamental trading volume. In this case, the planner perceives a welfare gain for both groups when initially increasing the tax level starting from $\tau = 0$, but soon finds it costly to further increase the tax rate without discouraging the fundamental trades of the first group of investors. However, for a sufficiently high tax level, most of the investors of the first group have stopped trading and cease to enter the planner's marginal welfare assessment. At that point, the planner only finds large welfare gains from reducing the non-fundamental trades of the second group of investors at the margin, generating a second local optimum.

Pure betting Finally, I formally show that when investors exclusively trade for non-fundamental motives, the optimal policy is associated with an infinite tax, eliminating trade altogether.

Lemma 4. (No-trade is optimal if all trade is belief-motivated) *If investors exclusively trade on belief differences, that is, $A_i = A$ and $X_{0i} = X_0$, $\forall i$, and $\mathbb{C}ov[M_{2i}, D]$ is identical for all investors, $\tau^* = \infty$ is optimal and there is no-trade in equilibrium under the optimal tax policy.*

The pure betting case is an interesting benchmark conceptually and in practice. Conceptually, Lemma 4 shows that whenever τ^* is finite, there must be at least some investors with fundamental trading motives. In practice, this result can be used to justify why, in many jurisdictions, activities that are clearly identified as relying exclusively on differences in beliefs, including casino-style gambling or horse races, are heavily taxed or even banned completely, as implied by Lemma 4. Lemma 4 connects with Proposition 2 in Gilboa, Samuelson and Schmeidler (2014), which shows that when all trading is due to belief differences, as in Lemma 4, it cannot be that allocations that involve trading no-betting Pareto dominate the no-trade allocation.

4.4 Trading volume implementation

As described above, the distribution of beliefs is the key determinant of the optimal tax. However, empirically recovering credible measures of investors' beliefs is a challenging task. To avoid the direct measurement of beliefs, I now introduce an alternative approach that implements the optimal policy using trading volume as an intermediate target.[26] After decomposing total trading volume into fundamental volume, non-fundamental volume, and a tax-induced volume reduction, Proposition 2 shows that the optimal tax can be implemented when total trading volume and the fundamental component of trading volume are equal, or equivalently, when the non-fundamental component of volume and the tax-induced volume reduction are equal.

Proposition 2. (Trading volume implementation)

a) [Trading volume decomposition] Trading volume in dollars, $P_1\mathcal{V}(\tau)$, where $\mathcal{V}(\tau)$ is defined in Equation (7), can be decomposed as follows:

$$\underbrace{P_1\mathcal{V}(\tau)}_{\textit{Total volume}} = \underbrace{\Theta_F(\tau)}_{\textit{Fundamental volume}} + \underbrace{\Theta_{NF}(\tau)}_{\textit{Non-fundamental volume}} - \underbrace{\Theta_\tau(\tau)}_{\textit{Tax-induced volume reduction}},$$

where $\Theta_F(\tau)$, $\Theta_{NF}(\tau)$, and $\Theta_\tau(\tau)$ are defined in the Appendix for the general case. Under Assumption [S], they correspond to

$$\Theta_F(\tau) = \frac{1}{2}\left|\frac{dX_{1i}}{d\tau}\right| A\left(\int_{i\in\mathcal{S}(\tau)}\left(\mathbb{C}ov\left[M_{2i},D\right]+\mathbb{V}ar\left[D\right]X_{0i}\right)dF(i) - \int_{i\in\mathcal{B}(\tau)}\left(\mathbb{C}ov\left[M_{2i},D\right]+\mathbb{V}ar\left[D\right]X_{0i}\right)dF(i)\right)$$

$$\Theta_{NF}(\tau) = \frac{1}{2}\left|\frac{dX_{1i}}{d\tau}\right|\left(\int_{i\in\mathcal{B}(\tau)}\mathbb{E}_i\left[D\right]dF(i) - \int_{i\in\mathcal{S}(\tau)}\mathbb{E}_i\left[D\right]dF(i)\right)$$

$$\Theta_\tau(\tau) = \tau P_1\left|\frac{dX_{1i}}{d\tau}\right|\int_{i\in\mathcal{B}(\tau)}dF(i),$$

where $\left|\frac{dX_{1i}}{d\tau}\right| = \frac{P_1}{A\mathbb{V}ar[D]}$, $\forall i$, is invariant to the tax level and constant across investors.

b) [Alternative optimal policy implementation] The planner can implement the optimal corrective policy by adjusting the tax rate until total trading volume equals the fundamental component of volume or, equivalently, until the non-fundamental component of volume equals the tax-induced volume reduction. Formally,

$$\tau^* \text{ is optimal} \iff P_1\mathcal{V}(\tau^*) = \Theta_F(\tau^*) \iff \Theta_{NF}(\tau^*) = \Theta_\tau(\tau^*).$$

c) [Small-tax approximation] For values of the optimal tax close to zero, knowledge of two variables from the laissez-faire economy is sufficient to approximate the optimal tax. These variables are i) the share of non-fundamental

[26] I use the intermediate target nomenclature by analogy to the literature on optimal monetary policy. In this model, equilibrium trading volume becomes an intermediate target to implement optimal portfolio allocations.

trading volume and ii) the semi-elasticity of trading volume to the tax rate. Formally, τ^* must satisfy

$$\tau^* \approx \frac{\overbrace{\frac{\Theta_{NF}(0)}{\Theta_F(0)+\Theta_{NF}(0)}}^{\text{Non-fundamental volume share}}}{\underbrace{-\left.\frac{d\log\mathcal{V}}{d\tau}\right|_{\tau=0}}_{\text{Volume semi-elasticity}}}. \tag{17}$$

Proposition 2a) provides a novel decomposition of trading volume into three components. The first component of trading volume is a function of investors' initial asset holdings, risk aversion, and hedging needs. I refer to this component as fundamental volume. The second component of trading volume is a function of investors' beliefs. I refer to this component as non-fundamental volume. The third component of trading volume is a function of the tax rate. I refer to this component as the tax-induced volume reduction. Note that when $\tau = 0$, this last component is zero, and all volume can be attributed to fundamental and non-fundamental components. Note also that the fundamental and non-fundamental components of volume can be understood as hypothetical volumes that would occur if only one or the other trading motive were present, for given sets $\mathcal{S}(\tau)$ and $\mathcal{B}(\tau)$.[27] Hypothetical volume decompositions are useful since they can be measured or approximated with the appropriate data. This decomposition of trading volume allows us to develop alternative implementations of the optimal policy.

Proposition 2b) shows that if the planner can credibly predict the amounts of fundamental or non-fundamental trading volume, he can adjust the optimal tax until total trading volume equals the fundamental component of volume or until the non-fundamental component of volume equals the tax-induced volume reduction. This new approach is appealing because it shifts the informational requirements for the planner from recovering investors' beliefs to constructing a model that predicts the appropriate amount of fundamental volume. Importantly, it is not true that the optimal tax is such that the non-fundamental component of volume is fully eliminated — see, for instance, Figure 6a for an illustration. In fact, since generically $\Theta_\tau(\tau) > 0$ when $\tau > 0$, the non-fundamental component of volume will be strictly positive at the optimal tax.

It is not at all obvious that the alternative implementation presented in Proposition 2 is possible. Pigouvian logic (see Dávila and Walther (2021) for a systematic study of corrective/Pigouvian regulation) suggests that the optimal corrective tax should be a function of marginal distortions, in this case given by investors' beliefs, and policy elasticities, in this case given by the impact of the tax on equilibrium trading. It turns out that, in this particular environment, total trading volume is related in equilibrium to investors' beliefs and to the impact of the tax on equilibrium trading via the trading volume decomposition proposed in Proposition 2, which allows for a different implementation of the optimal tax based on trading volume. Since the trading volume decomposition relies on the assumption of CARA/mean-variance preferences, Proposition 2 is more special and requires more structure relative to Proposition 1, which I show in Section 6 to be approximately valid very generally.

Finally, Proposition 2c) provides a new alternative implementation that exploits the definition of trading volume. The upshot of this new approximation is that it provides a simple and easily implementable characterization of the optimal tax based exclusively on information from the laissez-faire economy: the semi-elasticity of trading volume to a tax change and the share of non-fundamental trading volume. Intuitively, Equation (17) equalizes the reduction in trading volume caused by a tax change of size τ with the share of non-fundamental trading volume: this insight is far from obvious, since trading that occurs between investors with only fundamental motives is also distorted by the tax. In practical terms, an economy in which a 20bps

[27] Simsek (2013) also presents a different but related decomposition of trading volume on risk-sharing and speculative components for a different purpose (to detect when financial innovation increases portfolio risks).

tax increase reduces trading volume by 20%, implying a semi-elasticity of $\frac{-20\%}{20(\%)^2} = -100$, and whose share of non-fundamental trading volume is 30% features an approximately optimal tax of 30bps. Section 5, which further explores the quantitative implications of the model, also shows that the non-fundamental volume share and the volume semi-elasticity are sufficient to find the optimal tax in an environment with fully specified trading motives.

5 Quantitative implications

The results derived in Sections 3 and 4 are valid for any distribution of trading motives. In this section, I parameterize the cross-sectional distribution of trading motives with a dual goal. The first goal is to derive comparative statics results on primitives to understand how changes in the composition of trading motives affect the optimal tax. The second goal is to explore the quantitative implications of the model. Initially, I show that knowledge of two high-level variables is sufficient to set the optimal tax. Next, I show how to compute estimates of the optimal tax using the best existing empirical counterparts of the identified variables.

Formally, I assume that investors' beliefs and hedging needs are jointly normally distributed, as described in Assumption [G].

Assumption. [G] **(Gaussian trading motives)** *Investors' beliefs and hedging needs are jointly distributed across the population of investors according to*

$$\mathbb{E}_i [D] \sim \mu_d + \varepsilon_{di}$$
$$A\mathbb{C}ov [M_{2i}, D] \sim \mu_h + \varepsilon_{hi},$$

where $\mu_d \geq 0$ and $\mu_h = 0$. The random variables ε_{hi} and ε_{di} are jointly normally distributed as follows:[28]

$$\begin{pmatrix} \varepsilon_{di} \\ \varepsilon_{hi} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_d^2 & \rho\sigma_d\sigma_h \\ \rho\sigma_d\sigma_h & \sigma_h^2 \end{pmatrix}\right), \tag{18}$$

where $\rho \in [-1, 1]$ and $\sigma_d, \sigma_h \geq 0$. Investors have identical preferences $A_i = A$, $\forall i$, and hold identical initial asset positions $X_{0i} = X_0$, $\forall i$.

The cross-sectional dispersion of investors' mean beliefs σ_d^2 and that of hedging needs σ_h^2 respectively parameterize the relative importance of non-fundamental and fundamental trading. The share of non-fundamental trading volume, which I denote by $\delta^{NF} = \frac{\Theta_{NF}(\tau)}{\Theta_F(\tau)+\Theta_{NF}(\tau)}$, where $\Theta_{NF}(\tau)$ and $\Theta_F(\tau)$ are defined in Proposition 2, turns out to be a key object of interest. As shown in the Appendix, the share of non-fundamental trading volume δ^{NF}, which is constant for any tax rate in this model, can be expressed in terms of primitives as follows:

$$\delta^{NF} = \frac{\sigma_d^2 - \rho\sigma_d\sigma_h}{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h} = \frac{\frac{\sigma_d}{\sigma_h} - \rho}{\frac{\sigma_h}{\sigma_d} - \rho + \frac{\sigma_d}{\sigma_h} - \rho}. \tag{19}$$

Note that δ^{NF} is exclusively a function of the ratio $\frac{\sigma_d}{\sigma_h}$ — equivalently $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}$ — and the correlation between trading motives across the population of investors — see Figure A.3 in the Appendix for an illustration of the different combinations of $\frac{\sigma_d}{\sigma_h}$ and ρ that generate the same value of δ^{NF}. When $\rho = 0$ and $\delta^{NF} = 0$, investors have identical beliefs and all trade is fundamental. When $\rho = 0$ and $\delta^{NF} = 1$, all trade is driven by investors' beliefs and hence non-fundamental. The parameter ρ determines the correlation between both motives to trade across the population. A negative (positive) value of ρ implies that optimistic investors are also more likely to be buyers

[28]Note that $\mathbb{V}ar [D]$ denotes the variance of the payoff of the risky asset while σ_d^2 corresponds to the cross-sectional dispersion of investors' beliefs about the expected payoff of the risky asset.

(sellers) for fundamental motives. Making investors' preferences identical and assuming that they have identical asset holdings of the risky asset eliminate other motives for trading, although both assumptions can be relaxed without impact on the insights. Finally, note that when $\rho > 0$, δ^{NF} can take values outside of the interval $[0, 1]$, which prevents us from interpreting δ^{NF} as a share in those extreme cases. As shown below, the optimal tax is infinite or negative in those cases.

5.1 Theoretical results

Note that Assumption [G] implies that Assumption [S] is satisfied, since the random variable $\mathbb{E}_i[D] - A\mathbb{C}ov[M_{2i}, D] - A\mathbb{V}ar[D]X_0$ is normally distributed and consequently symmetric. Therefore, the equilibrium price does not vary with the value of τ and can be expressed as a function of primitives. Formally, it corresponds to

$$P_1 = \mu_d - A\mathbb{V}ar[D]Q. \tag{20}$$

The Appendix provides explicit characterizations of equilibrium asset allocations, trading volume and turnover, as well as the share of buyers, sellers, and inactive investors. The Appendix also includes explicit characterizations of fundamental and non-fundamental trading volume, as well as the tax-induced volume reduction. The planner's objective is quasi-concave in this case, and the optimal tax satisfies a non-linear equation involving the inverse Mills ratio of the normal distribution. The following results emerge under Assumption [G].

Proposition 3. (Optimal tax and comparative statics)

a) As long as some investors have heterogeneous beliefs ($\sigma_d > 0$) and investors' beliefs and hedging needs are not positively correlated ($\rho \leq 0$), it is optimal to set a strictly positive tax.

b) When positive and finite, the optimal tax is increasing in the ratio of non-fundamental trading to fundamental trading $\frac{\sigma_d}{\sigma_h}$ for any correlation level ρ. Consequently, a mean-preserving spread of investors' beliefs is associated with a higher optimal tax.

c) If the share of non-fundamental trading volume $\delta^{NF} \geq 1$, then $\tau^* = \infty$; if $0 \leq \delta^{NF} < 1$, then $\tau^* \in [0, \infty)$; if $\delta^{NF} < 0$, then $\tau^* < 0$.

The result for the case in which $\rho = 0$ in Proposition 3a) is a particular case of the general result in Proposition 1b), which guarantees the optimality of a positive tax when fundamental and non-fundamental motives for trade are independent of each other and there exists some non-fundamental trading. With Gaussian trading motives, assuming that fundamental and non-fundamental trading motives are negatively correlated further increases the rationale for taxation, since it implies that optimistic (pessimistic) investors are also more likely to be buyers (sellers) for fundamental motives. Figure A.3 in the Appendix illustrates that in many instances in which $\rho > 0$, the optimal tax can still be positive and finite.

Proposition 3b) shows that the optimal tax increases with the share of non-fundamental trading volume in the more relevant region in which the tax is positive and finite. Consequently, a mean-preserving spread of investors' beliefs is associated with a higher optimal tax. Intuitively, an increase in belief dispersion makes optimistic (pessimistic) investors more likely to be buyers (sellers), increasing non-fundamental trading volume and the motive to tax by the planner.

Finally, Proposition 3c) shows that knowing the value of δ^{NF}, which is exclusively a function of $\frac{\sigma_d}{\sigma_h}$ and ρ, is sufficient to fully determine the sign of the optimal tax. When $\delta^{NF} \in [0, 1)$, the optimal tax is non-negative and finite. Outside of this interval, the optimal policy features either a subsidy or an infinite tax. As described above, these extreme scenarios can only arise when the correlation between trading motives ρ is sufficiently large. Next, as the last step before quantifying the model, I show that additional information besides δ^{NF} is needed to find the magnitude of the optimal tax.

5.2 Quantitative assessment

In this section, I characterize which variables ought to be measured to implement an optimal tax and provide a discussion of the empirical counterparts of such variables and what they imply for the optimal tax. Given that the model studied in this paper is stylized in a number of dimensions, it is important to highlight that any quantitative conclusion of this section should be taken with caution. Hence, it is perfectly valid to interpret the quantitative assessment of the model as a numerical illustration built on sensible numbers.

Optimal tax identification A significant challenge for any theory of optimal taxation is to clearly identify the informational requirements that a planner would need to actually implement the optimal tax and to measure the welfare consequences of a change in the tax rate. Proposition 4 shows that a small number of high-level variables are sufficient to answer both questions in this case. Interestingly, the same two variables that locally approximate the optimal tax in Proposition 2c) turn out to be the variables needed to find the globally optimal tax under Assumption [G].

Proposition 4. (Optimal tax identification/Sufficient statistics)

a) [Optimal tax] Knowledge of two variables is sufficient to determine the optimal tax. These variables are i) the share of non-fundamental trading volume, δ^{NF}, defined in Equation (19), and ii) the semi-elasticity of trading volume to the tax rate, given by $\varepsilon_{\tau}^{\log \mathcal{V}} \equiv \frac{d \log \mathcal{V}}{d\tau}$.[29]

b) [Marginal welfare impact of a tax change] Knowledge of three variables is sufficient to determine the normalized aggregate marginal welfare impact of a tax change, $\frac{\frac{dV^p}{d\tau}}{P_1 Q}$. These variables are i) the share of non-fundamental trading volume, δ^{NF}, defined in Equation (19), ii) the semi-elasticity of trading volume to the optimal tax, given by $\varepsilon_{\tau}^{\log \mathcal{V}} \equiv \frac{d \log \mathcal{V}}{d\tau}$, and iii) the risk premium, given by $\Pi \equiv \frac{A \mathbb{V}ar[D] Q}{P_1}$.

Proposition 4a) shows that in addition to knowing the share of non-fundamental trading volume δ^{NF}, which Proposition 3 determined to be sufficient to pin down the sign of the optimal tax, a planner must also know the sensitivity of total trading volume to a change in the tax rate to fully determine the magnitude of the optimal tax. As discussed below, when finding empirical counterparts, this elasticity can be directly estimated from tax policy changes.

Proposition 4b) shows that in order to assess the marginal welfare impact of a tax change (normalized in terms of the total capitalization of the risky asset), a planner needs to also take a stance on the risk premium, which is invariant to the tax rate in this model. Intuitively, the risk premium contains information on investors' willingness to pay for the ability to share risks. The Appendix shows that a planner with separate knowledge of $\frac{\sigma_d}{\sigma_h}$ and ρ, in addition to the three variables identified in Proposition 4b), can also fully recover the distribution of marginal welfare impacts for each individual, $\frac{dV_i^p}{d\tau}$.

Proposition 4 is the most relevant result from the perspective of implementing the optimal tax characterized in this paper. Importantly, by showing that only two high-level variables are needed to find the optimal tax, it avoids the need to specify the parameters of the model. For instance, while it may be hard to separately estimate $\frac{\sigma_d}{\sigma_h}$ and ρ, it may be easier to find different approaches that can help discipline δ^{NF}, as discussed below. In this context, it is also important that the sufficient statistics identified — the risk premium, the volume semi-elasticity

[29] As implied by Equation (53) in the Appendix, it is sufficient to know this semi-elasticity for any given value of τ. The quantitative results in this section are based on empirical counterparts of $\varepsilon_{\tau}^{\log \mathcal{V}}\big|_{\tau=0}$.

to tax changes, and the share of non-fundamental trading volume — are scale-invariant variables. The use of scale-invariant variables sidesteps common concerns associated with CARA calibrations — see, e.g., Campbell (2017) — and allows us to conjecture that the quantitative insights should remain valid, at least in approximate form, in more general quantitative models that match the relevant sufficient statistics.

Optimal tax calibration Proposition 4 shows that finding an empirical counterpart of the optimal tax exclusively requires measures of $\varepsilon_{\tau}^{\log \mathcal{V}}$ and δ^{NF}, and that an estimate of the risk premium is necessary to compute welfare gains. Next, I describe how to find the plausible empirical counterparts of these objects, given the existing evidence. I continue to set the gross risk-free rate to 1 in the calibration — the results are virtually indistinguishable for reasonable values of the risk-free rate.

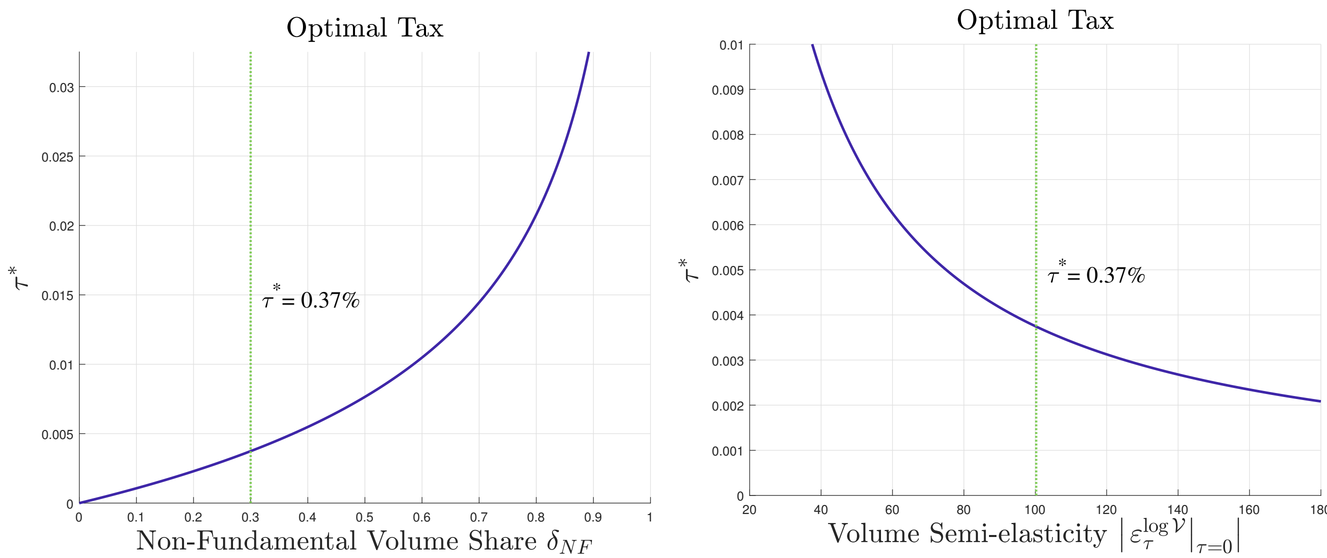
The evidence in Colliard and Hoffmann (2017), who precisely estimate the necessary volume semi-elasticity to tax changes using the recent implementation of a financial transaction tax in France, is best suited to discipline the choice of $\varepsilon_{\tau}^{\log \mathcal{V}}$. They find, starting from $\tau = 0$, that a 20bps tax increase (0.2%) persistently reduced trading volume for stocks by 20%, which corresponds to a semi-elasticity $\varepsilon_{\tau}^{\log \mathcal{V}}\big|_{\tau=0} = \frac{d \log \mathcal{V}}{d\tau}\Big|_{\tau=0} = \frac{-20\%}{0.2\%} = -100.$[30] As shown in the Appendix, in this model there is a tight relation among the semi-elasticity of trading volume, the risk premium, and the amount of asset turnover. I use this relation to choose the frequency at which to calibrate the model. By choosing a quarterly calibration, the model is able to jointly match a standard quarterly risk premium $\Pi = {}^{6\%}\!/_{4} = 1.5\%$; the turnover ratio of domestic shares for US stocks, which I compute to be $\Xi(0) = 33\%$ of total asset float in a quarter, using information from the World Federation of Exchanges database between 1990 and 2018; and the volume semi-elasticity $\varepsilon_{\tau}^{\log \mathcal{V}}\big|_{\tau=0} = -100$.

Providing an appropriate estimate of the share of non-fundamental trading volume δ^{NF} is certainly more challenging. For instance, Hong and Stein (2007) argue that "the bulk of volume must come from differences in beliefs that lead traders to disagree about the value of a stock." The Appendix includes an estimation procedure for δ^{NF} in this model that uses information on individual investors' portfolio choices and hedging needs. Importantly, this procedure does not use any information on investors' beliefs. I show that this procedure yields an unbiased estimate of δ^{NF} when $\rho = 0$ and explicitly characterize the potential bias when $\rho \neq 0$.

The recent work by Koijen and Yogo (2019) maps closely to the estimation procedure described in the Appendix and seems best suited to shed light on the value of δ^{NF}. They seek to explain investors' portfolio holdings using a rich characteristics-based model of investors' asset demands, which can be interpreted as modeling investors' hedging/fundamental trading motives. Their flexible approach is able to explain 40% of the variation in investors' portfolio holdings, leaving 60% of investors' portfolio holdings unexplained. Since we cannot guarantee that their explanatory variables include all possible hedging motives, 60% can be interpreted as an upper bound for δ^{NF}. Conservatively, I adopt $\delta^{NF} = 0.3$ as the reference value for the share of non-fundamental trading volume. This choice corresponds to imposing a uniform prior for δ^{NF} on the interval $[0\%, 60\%]$. This choice implies that a non-negligible share of trading is non-fundamental, while erring on the side of attributing most trades to fundamental motives. A reader who perceives that different values of $\varepsilon_{\tau}^{\log \mathcal{V}}\big|_{\tau=0}$ and δ^{NF} are more plausible, or may be more appropriate in alternative contexts, can refer to the sensitivity analysis in Figure 5.

Quantitative results Under the parameterization just described, Figures 5 and 6a illustrate the magnitudes implied by the model for the optimal tax and also provide a sensitivity analysis.

[30] I adopt as reference the average estimate in Colliard and Hoffmann (2017) for stocks outside of the Euronext's SLP program, although they find a range of semi-elasticities for different investors and market structures, with volume reductions between 10% to 40%. Alternatively, using data from the Swedish experience in the 80's, Umlauf (1993) finds that a 1% tax increase is associated with a decline in turnover of more than 60%, which corresponds to a volume semi-elasticity of −60.



(a) Share of non-fundamental trading volume

(b) Volume semi-elasticity

Figure 5: Optimal tax (sensitivity analysis)

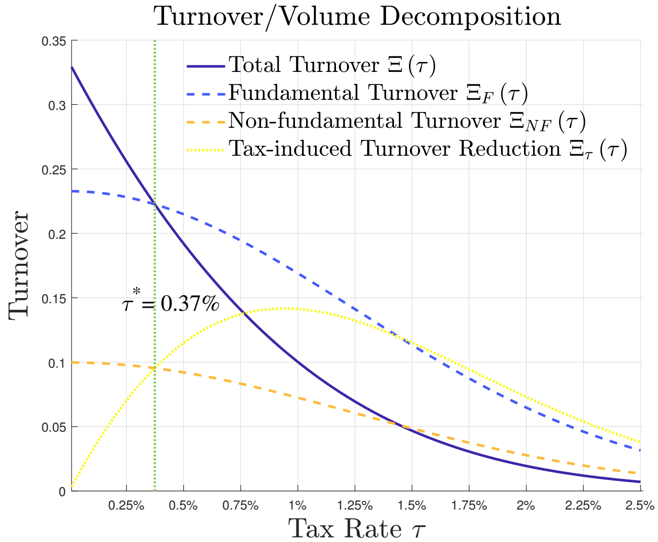
Note: Figure 5a shows the optimal tax τ^* as a function of the share of non-fundamental trading volume δ^{NF}, defined in Equation (19). Figure 5b shows the optimal tax τ^* as a function of the magnitude of volume semi-elasticity $\varepsilon_{\tau}^{\log \mathcal{V}} = \frac{d \log \mathcal{V}}{d\tau}$, evaluated at $\tau = 0$. The set of reference parameters are $\delta^{NF} = 0.3$ (used in the right plot) and $\varepsilon_{\tau}^{\log \mathcal{V}}\big|_{\tau=0} = -100$ (used in the left plot). In both figures, the optimal tax for the reference parameters, represented by a vertical dotted line, is $\tau^* = 0.37\%$.

Figure 5a illustrates how the optimal tax varies as a function of the share of non-fundamental trading volume δ^{NF}, for the reference level of the volume semi-elasticity to tax changes, $\varepsilon_{\tau}^{\log \mathcal{V}}\big|_{\tau=0} = -100$. For a given volume semi-elasticity, the optimal tax is increasing in δ^{NF}. Consistent with the theoretical results, when $\delta^{NF} = 0$, the optimal tax is also 0. For values of δ^{NF} close to 0, the optimal tax increases almost linearly with δ^{NF}. When δ^{NF} approaches 1, the optimal tax tends sharply to ∞. For reference, when $\delta^{NF} = 0.1$, $\tau^* = 0.11\%$, and when $\delta^{NF} = 0.6$, $\tau^* = 1.05\%$.

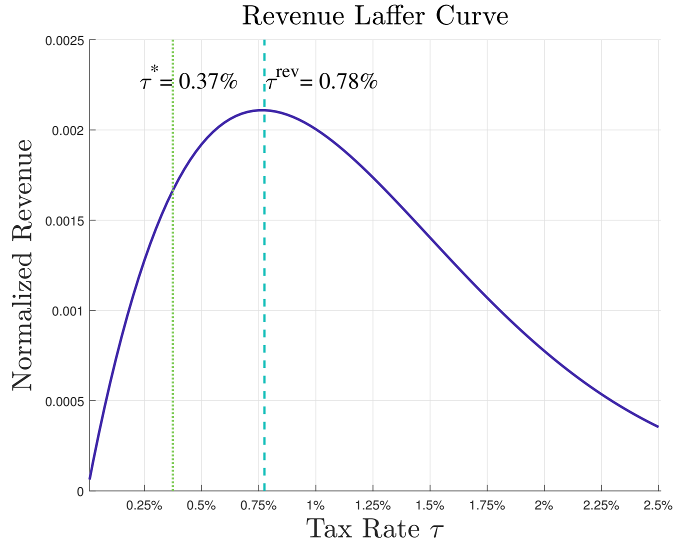
Figure 5b illustrates how the optimal tax varies as a function of the volume semi-elasticity to tax changes evaluated at $\tau = 0$, for the reference level of the share of non-fundamental trading volume, $\delta^{NF} = 0.3$. For a given share of non-fundamental trading volume, the optimal tax is decreasing in $\left|\varepsilon_{\tau}^{\log \mathcal{V}}\big|_{\tau=0}\right|$. Intuitively, when the magnitude of the semi-elasticity is high (low), a small (large) transaction tax is needed to eliminate the same amount of non-fundamental trading volume. For reference, when $\varepsilon_{\tau}^{\log \mathcal{V}}\big|_{\tau=0} = -50$, $\tau^* = 0.75\%$, and when $\varepsilon_{\tau}^{\log \mathcal{V}}\big|_{\tau=0} = -150$, $\tau^* = 0.23\%$.[31] Note that the optimal tax is convex both in the level of δ^{NF} and $\left|\varepsilon_{\tau}^{\log \mathcal{V}}\big|_{\tau=0}\right|$, which suggests that uncertainty about the level of both determinants of the optimal tax may call for higher optimal taxes.

While Figure 5 illustrates the results of Propositions 1 and 3, Figure 6a graphically illustrates how to make use of the results of Proposition 2 in practice. It shows the differential behavior of the three components of trading volume to changes in the tax rate for the baseline calibration. The reduction in fundamental and non-fundamental volume, which is monotonic, is driven by extensive margin changes in the composition of active investors. Meanwhile, the tax-induced reduction component of volume grows rapidly at first before it starts decreasing monotonically, due to

[31] Allowing investors to trade dynamically at different frequencies or introducing technological trading costs may change the positive predictions of the model. However, after calibrating to match the observed semi-elasticity of volume to tax changes, one would expect to find comparable values for the optimal tax for a given share of non-fundamental trading volume.



(a) Trading volume implementation



(b) Tax revenue

Figure 6: Trading volume and tax revenue

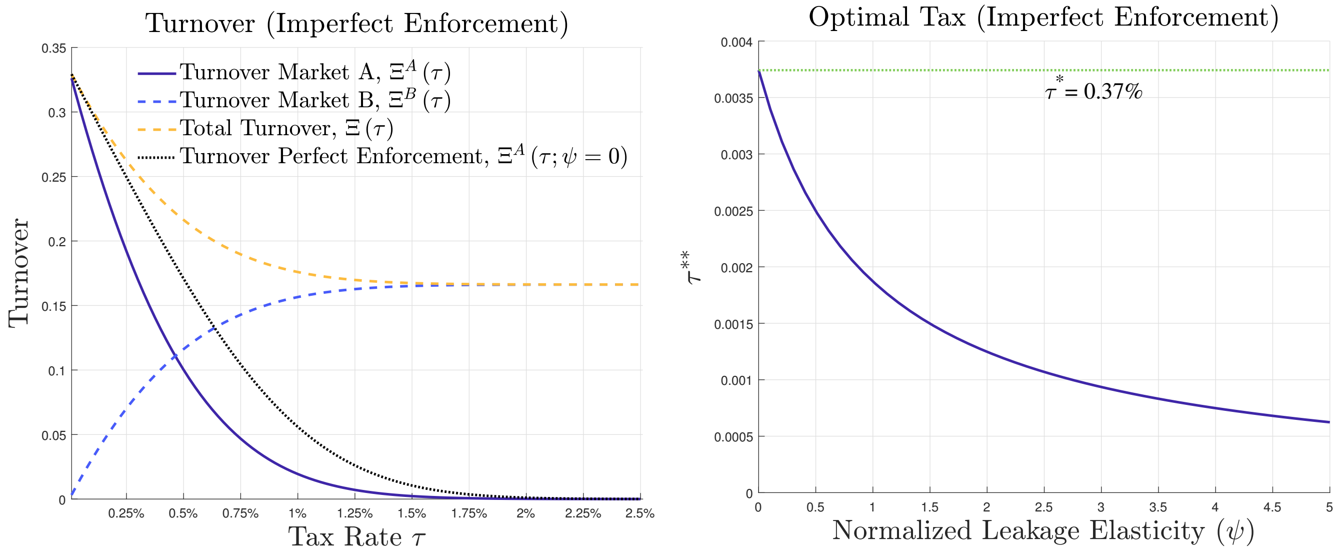
Note: Figure 6a illustrates the volume decomposition established in Proposition 2 when $\varepsilon_\tau^{\log \mathcal{V}}\Big|_{\tau=0} = -100$ and $\delta^{NF} = 0.3$. It shows total turnover, given by $\Xi(\tau) = \frac{\mathcal{V}(\tau)}{Q}$, fundamental turnover, given by $\Xi_F(\tau) = \frac{\Theta_F(\tau)}{P_1 Q}$, non-fundamental turnover, given by $\Xi_{NF}(\tau) = \frac{\Theta_{NF}(\tau)}{P_1 Q}$, and the tax induced turnover reduction $\Xi_\tau(\tau) = \frac{\Theta_\tau(\tau)}{P_1 Q}$. The optimal tax, represented by a vertical dotted line, is $\tau^* = 0.37\%$. Note that this line goes through to the points in which $\Xi(\tau)$ and $\Xi_F(\tau)$, as well as $\Xi_{NF}(\tau)$ and $\Xi_\tau(\tau)$, intersect, which is consistent with Proposition 2b). Figure 6b shows total tax revenue per dollar of market capitalization of the risky asset, given by $2\tau\frac{\mathcal{V}(\tau)}{Q} = 2\tau\Xi(\tau)$. The optimal corrective tax is $\tau^* = 0.37\%$, while the revenue maximizing tax is $\tau^{\text{rev}} = 0.78\%$.

the overall trading reduction on the extensive and intensive margins. Total and fundamental volume intersect at the optimal tax rate of $\tau^* = 0.37\%$. Consistent with Proposition 2, non-fundamental volume and the tax-induced volume reduction intersect as well at the same tax rate. As explained in Section 4.4, note that the non-fundamental component of volume remains strictly positive at the optimal tax.

We can also verify the validity of Proposition 2c) in this particular calibration. Given a volume semi-elasticity $\varepsilon_\tau^{\log \mathcal{V}}\Big|_{\tau=0} = -100$, and a share of non-fundamental trading volume $\delta^{NF} = 0.3$, the (approximate) optimal tax rate given in Equation (17) becomes $\tau^* \approx \frac{0.3}{100} = 0.3\%$, close to the exact value found. Given a volume semi-elasticity of -100, 2c) associates the percentage of non-fundamental trading to the optimal tax, when expressed in basis points. That is, a 20%, 40%, or 60% share of non-fundamental trading volume is approximately associated with an optimal tax of 20, 40, or 60bps, respectively.

Finally, it is also possible to compute the aggregate marginal welfare gain associated with a given tax change relying exclusively on the observables identified in Proposition 4. Starting from the laissez-faire equilibrium, I show in the Appendix that $\frac{\frac{dV^P}{d\tau}}{P_1 Q}\Big|_{\tau=0} = \frac{1}{\Pi}\frac{1}{\left|\varepsilon_\tau^{\log \mathcal{V}}\big|_{\tau=0}\right|}\delta^{NF}$, which for the reference calibration of the model corresponds to $\frac{1}{1.5\%} \cdot \frac{1}{100} \cdot 0.3 = 0.2$. This result implies that introducing a 10 basis points transaction tax is associated with a quarterly welfare gain of approximately 2 basis points of the capitalization of the risky asset. Taking the value of the US stock market as reference, of roughly 30 trillion dollars, introducing a 10bps tax is associated with a welfare gain of approximately 6 billion per quarter, or 24 billion per year, which roughly corresponds to 10 basis points of US GDP.

Tax revenue In practice, policymakers are partly interested in financial transaction taxes for their ability to raise revenue. However, given that the theory developed in this paper is of a corrective nature, I have not discussed



(a) Turnover with imperfect enforcement

(b) Optimal tax with imperfect enforcement

Figure 7: Imperfect enforcement

Note: Figure 7a illustrates the behavior of asset turnover in the extension of the model with imperfect tax enforcement when $\psi = 1$. The solid dark blue line shows turnover in market A, the one directly affected by the tax. The dashed light blue line shows turnover in market B, which starts at 0 when $\tau = 0$. The orange dashed line shows total turnover, given by the sum of turnover in both markets $\Xi(\tau) = \Xi^A(\tau) + \Xi^B(\tau)$. The dotted black shows, for reference, total turnover (in market A) in the baseline model, for which $\psi = 0$. As expected, turnover in market A under imperfect enforcement is lower than under perfect enforcement, but total volume (adding up both markets) is higher. Figure 7b shows the optimal tax under imperfect enforcement τ^{**}, defined in Equation (21), as a function of the normalized leakage elasticity ψ.

so far the behavior of tax revenue. Figure 6b illustrates how total tax revenue per dollar of market capitalization of the risky asset, given by $2\tau\Xi(\tau)$, varies with the tax rate for the baseline calibration. As expected, tax revenue takes the form of a "Laffer curve," initially increasing with the tax rate, reaching a maximum, and then decreasing. It should be evident that the revenue maximizing tax can in principle be higher or lower than the optimal corrective tax. This occurs because the revenue maximizing tax is exclusively a function of volume/turnover (regardless of whether such volume is fundamental or non-fundamental), while the key determinant of the optimal corrective tax is the share of non-fundamental trading volume. In this particular calibration, the revenue maximizing tax $\tau^{\text{rev}} = 0.78\%$ is higher than the optimal corrective tax $\tau^* = 0.37\%$.

Imperfect tax enforcement In practice, a major concern with imposing financial transaction taxes is that trading might migrate to a different country or to a different market within the same country — see, for instance, Cai et al. (2017) for recent direct evidence on this phenomenon. In the Appendix, I present a tractable extension of the model that allows investors to circumvent the financial transaction tax by trading the risky asset in a different market, at a cost. Formally, I allow investors to trade the same asset in two different markets, market A and market B. Trading in market A is subject to a transaction tax, as studied in this paper, while trading in market B is not. To prevent investors from avoiding the tax altogether by shifting all their trading to market B, I assume that trades in market B are subject to a quadratic adjustment cost $\frac{\nu}{2}\left(\Delta X_{1i}^B\right)^2$. This environment parsimoniously captures the notion that investors have incentives to avoid transaction taxes, but that this may be costly.

In this environment, I show that the optimal tax under imperfect enforcement can be characterized as follows:

$$\tau^{**} = \frac{1}{1+\psi} \frac{\mathbb{E}_{\mathcal{B}(\tau^{**})}\left[\frac{\mathbb{E}_i[D]}{P_1}\right] - \mathbb{E}_{\mathcal{S}(\tau^{**})}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]}{2}, \tag{21}$$

where $\psi = \frac{A \mathbb{V}ar[D]}{\nu}$ indexes how easy it is for trading activity to migrate/leak from the regulated market A to the unregulated market B.[32] Using the terminology of Dávila and Walther (2021), I refer to ψ as a normalized leakage elasticity, which captures how much activity moves to the unregulated market. Heuristically, Equation (21) implies that, for a given difference in beliefs between buyers and sellers, the optimal tax is lower when tax evasion is easier (ψ is higher). Formally, I show in the Appendix that, for given primitives that pin down the risk premium and laissez-faire turnover, the optimal tax when trading activity can move untaxed to a different market is decreasing in the value of the normalized leakage elasticity ψ.[33]

Figure 7a illustrates the behavior of asset turnover as a function of the tax rate for both market A and market B. It also shows total turnover and, for reference, turnover in the baseline model, in which $\psi = 0$. As expected, turnover in market A under imperfect enforcement is lower than under perfect enforcement, but total volume (adding up both markets) is higher. Figure 7b illustrates how the optimal tax varies as a function of how easy it is for trading activity to migrate to an unregulated market. For reference, as shown in the Appendix, when $\psi = 1$ — which implies that half of the reduction in turnover in market A migrates to market B — the optimal tax is $\tau^{**} = 0.185\%$, exactly half of the optimal tax with perfect enforcement. Finally, note that if the planner had an additional motive to keep trading activity with a given jurisdiction, this would push for further lowering the level of the optimal tax.

6 Robustness and limitations

In this section, I first show that the characterization of the optimal tax in Proposition 1 remains valid as an approximation in more general models. Next, based on results derived in the Appendix, I summarize the results of several extensions. Finally, I discuss several mechanisms that are not explicitly considered in the paper but that may be important to better understand the desirability of financial transaction taxes.

6.1 General utility and arbitrary beliefs

Here, I consider an environment in which investors also face a consumption/savings decision, have general utility specifications, and disagree about probability assessments in an arbitrary way. Investors' beliefs are now modeled as a change of measure with respect to the planner's probability measure, which (jointly) determines the realization of all random variables — asset payoffs and endowments — in the model. The beliefs of investor i about date 2 uncertainty are represented by a Radon-Nikodym derivative Z_i, which is absolutely continuous with respect to the planner's probability measure and satisfies $\mathbb{E}_p[Z_i] = 1$. This random variable Z_i flexibly captures any discrepancy between the probability assessments made by the planner and those made by investors. Formally, for a given random variable X, I use the notation $\mathbb{E}_i[X] = \mathbb{E}_p[Z_i X]$.

In this case, investors maximize

$$\max_{C_{1i}, C_{2i}, X_{1i}, Y_{1i}} U_i(C_{1i}) + \beta_i \mathbb{E}_i[U_i(C_{2i})],$$

[32] As formally shown in the Appendix, When $\nu \to 0$, or equivalently $\psi \to \infty$, avoiding the tax is costless, so all trades migrate to market B whenever $\tau > 0$. When $\nu \to \infty$, or equivalently $\psi \to 0$, no trades migrate to market B, nesting the baseline model.

[33] This is consistent with the general analysis of corrective regulation with imperfect instruments of Dávila and Walther (2021), who show that the optimal corrective policy is sub-Pigouvian when regulated and unregulated decisions are gross substitutes, as it is here with trading in markets A and B.

where $U_i(\cdot)$ satisfies standard regularity conditions, subject to the following budget constraints

$$C_{1i} + X_{1i}P_1 + Y_{1i} = M_{1i} + X_{0i}P_1 - \tau P_1 |\Delta X_{1i}| + T_{1i}$$
$$C_{2i} = M_{2i} + X_{1i}D + RY_{1i},$$

where Y_{1i} denotes the amount invested in the risk-free asset and R denotes the risk-free rate, now an equilibrium object. As in the baseline model, both D and M_{2i} are random and potentially correlated. I consider the same competitive equilibrium definition as in Section 3, now augmented to include market clearing of the risk-free asset.

In this model, when active, investors' optimal portfolio decisions satisfy the following pair of Euler equations

$$1 = R\mathbb{E}_i [m_i] \tag{22}$$
$$P_1 (1 + \tau \operatorname{sgn}(\Delta X_{1i})) = \mathbb{E}_i [m_i D], \tag{23}$$

where $m_i = \frac{\beta_i U'(C_{2i})}{U'(C_{1i})}$ denotes investor i's stochastic discount factor. Again, some investors may decide not to trade the risky asset at all when their optimal asset holdings are close to their initial asset holdings. In that case, $\Delta X_{1i} = 0$, and Equation (23), which is the counterpart of Equation (5), does not hold.

As in the baseline model, the planner computes social welfare under a single belief and aggregates the money-metric utility changes across all investors. In the Appendix, I characterize the exact optimal tax in this general environment. In Proposition 5, I show that the optimal tax characterization in this general model, approximated when investors' stochastic discount factors are roughly constant, turns out to be identical to the optimal tax characterization in the baseline model.

Proposition 5. (Optimal tax approximation) *In the limit in which investors' stochastic discount factors are approximately constant, formally, when $m_i \to \overline{m} \in \mathbb{R}_+$, the optimal financial transaction tax τ^* satisfies*

$$\tau^* \approx \frac{\Omega_{\mathcal{B}(\tau^*)} - \Omega_{\mathcal{S}(\tau^*)}}{2},$$

where $\Omega_{\mathcal{B}(\tau^*)}$ and $\Omega_{\mathcal{S}(\tau^*)}$ are described in Equation (13). This expression, which is also independent of the planner's beliefs, is identical to the one in Proposition 1a).

When investors' stochastic discount factors are approximately constant, Proposition 5 shows that the optimal tax in the general case collapses to the optimal tax in the baseline model. This approximation corresponds to a scenario in which the risks faced by investors are not too large in comparison to their risk-bearing capacity.[34] Proposition 5 allows us to interpret the results of Proposition 1 as an approximation to more general models. Note that only the first moment of the distribution of beliefs about the payoff of the risky asset appears explicitly in the approximated optimal tax formula, which motivates the sustained assumption in the paper restricting belief differences to the first moment of the distribution of the payoffs of the risky asset.

Finally, the Appendix includes several additional results. In addition to a detailed explanation of the optimal tax characterization in this general environment, the Appendix also includes a simulation of the non-linear model that aims to match the same high-level variables identified in Section 5.2, comparing and relating its quantitative findings to those in that section.

[34] This result is related to the classic Arrow-Pratt approximation (Arrow (1971); Pratt (1964)), which shows that the solution to the CARA-Normal portfolio problem approximates the solution to any portfolio problem for small gambles, but it is not identical, since Proposition 5 directly approximates the optimal tax formula, while the standard approximation is done over investors' optimality conditions.

6.2 Extensions

The Appendix of this paper includes multiple extensions. These show that the characterization of the optimal tax formula remains valid identically or suitably modified in more general environments.

First, I introduce the possibility that investors face short-sale or borrowing constraints that limit their portfolio decisions. I show that the optimal tax formula from Proposition 1a) remains valid in that case, and show in a simulated version of the model that the optimal tax becomes lower when investors face short-sale constraints. Second, I show that the optimal tax formula from Proposition 1a) remains valid when there are pre-existing trading costs, as long as these are compensation for the use of economic resources, not economic rents. Perhaps counter-intuitively, when pre-existing trading costs reduce the share of fundamental trading, the optimal transaction tax can be increasing in the level of trading costs and vice versa. Third, I show that the sign of the optimal tax is independent of whether tax enforcement is perfect or imperfect. However, I show that the magnitude of the optimal tax is decreasing in the investors' ability to avoid paying taxes. Fourth, in an environment with multiple risky assets, the optimal tax becomes a weighted average of the optimal tax for each asset, with higher weights given to those assets whose volume is more sensitive to tax changes. This result follows from the second-best Pigouvian nature of the policy. Fifth, I show that investor-specific taxes are needed to implement the first-best outcome. Sixth, I provide a formula for the upper bound of welfare losses induced by a marginal tax change when all trading is fundamental. Seventh, I derive an optimal tax characterization in the case in which investors and the planner disagree about second moments.

Finally, in a q-theory production economy, I show that a transaction tax may generate additional first-order gains/losses if the planner's belief differs from the average belief of investors. Intuitively, social welfare in a production economy also depends on the level of aggregate risk and investment. For instance, if a marginal tax increase reduces (increases) investment at the margin when investors are too optimistic (pessimistic) relative to the planner, a positive tax is welfare improving, and vice versa.

While, in principle, the optimal tax formula in a production economy depends on the belief used by the planner, there are three scenarios in which this is not the case and the optimal tax characterization from Proposition 1 remains valid. First, when a transaction tax has no impact on the asset price, i.e., $\frac{dP_1}{d\tau} = 0$, investment remains unchanged, and Proposition 1 applies without modification. Second, if the planner uses the average belief of investors in the economy to calculate welfare, there is no additional rationale for taxation due to production, so Proposition 1 remains valid. Finally, using an additional policy instrument to regulate aggregate investment would be optimal in this environment, allowing the planner to set the optimal transaction tax as in Proposition 1.[35]

6.3 Additional considerations

Before concluding, it is worth highlighting several channels that have not been explored explicitly in this paper, but that are relevant when considering the desirability of financial transaction taxes.

First, and perhaps more critically, by assuming that investors do not learn from each other or from the price, and agree to disagree, this paper has abstracted from the roles played by financial markets in aggregating information, enabling information acquisition, and guiding production via information diffusion, potentially with feedback effects (Bond, Edmans and Goldstein, 2012). Transaction taxes may impact these roles separately, with different positive

[35] An earlier version of this paper studied how dynamic trading affects the determination of the optimal tax. Consistent with Tobin's insight, a transaction tax affects more forward-looking investors who buy and sell at high frequencies, since the anticipation of future taxes reduces their incentives to trade. However, buy-and-hold investors are barely sensitive to a transaction tax. Portfolio sensitivities endogenously capture both possibilities.

and normative implications. On the positive side, Dávila and Parlatore (2021) provide a systematic study of how transaction costs/taxes impact price informativeness, characterizing the conditions under which transaction costs/taxes do not affect information aggregation, even though they discourage the endogenous acquisition of information. Vives (2017) examines an environment in which a positive transaction tax is welfare improving by correcting investors' information acquisition choices. While these papers provide a first step towards understanding how transaction taxes affect trading in financial markets and welfare, a systematic normative analysis of the desirability of transaction taxes when investors have dispersed information is still missing.[36] Relatedly, whether the volume decomposition and optimal tax implementation presented in Proposition 2 remain valid in models with dispersed information also remains an open question.

Second, given its static nature, this paper abstracts from dynamic considerations that may potentially play an important role, particularly quantitatively, when setting an optimal transaction tax. In particular, this paper does not explicitly account for fluctuations of investors' beliefs (perhaps via learning), risk attitudes, and risk exposures, which ultimately generate persistent trading over time. Relatedly, there is scope to further understand the impact of transaction taxes on wealth dynamics and how wealth dynamics affect the determination of the optimal tax.

Finally, this paper has assumed a Walrasian/centralized trading protocol. Understanding the normative implications of taxing financial transactions in models in which markets are decentralized or when some investors have market power remains a fruitful avenue for further research.

7 Conclusion

This paper studies the welfare implications of taxing financial transactions in an equilibrium model in which financial market trading is driven by both fundamental and non-fundamental motives. While a transaction tax is a blunt instrument that distorts both fundamental and non-fundamental trading, the welfare implications of reducing each kind of trading are different. As long as a fraction of investors hold heterogeneous beliefs that are unrelated to their fundamental motives to trade, a planner who weights investors equally and computes social welfare using a single belief will find a strictly positive tax optimal. Interestingly, the optimal tax may be independent of the belief used by the planner to calculate welfare.

The optimal transaction tax can be expressed as a function of investors' beliefs and portfolio sensitivities. Alternatively, the planner can determine the optimal tax rate by directly equating the level of total trading volume to the level of fundamental trading volume. Knowledge of two variables, the share of non-fundamental trading volume and the semi-elasticity of trading volume to the tax rate, is sufficient to compute the magnitude of the optimal tax, as shown and illustrated in this paper by using the best existing estimates of both variables. Even though this paper has finally provided a robust and an internally consistent corrective theory of financial transaction taxes, there is significant scope to better understand whether transaction taxes are a desirable policy instrument in practice.

[36] Any normative theory of this form must first characterize how transaction taxes affect the amount of information in the economy and then determine whether more or less information is desirable.

References

Alvarez, Fernando, and Andrew Atkeson. 2018. "The Risk of Becoming Risk Averse: A Model of Asset Pricing and Trade Volumes." *Working Paper.*

Arrow, Kenneth J. 1971. "The Theory of Risk Aversion." In *Essays in the Theory of Risk-Bearing.* 90–120. Markham Publishing Co. Chicago.

Auerbach, Alan J., and James R. Hines Jr. 2002. "Taxation and Economic Efficiency." *Handbook of Public Economics*, 3: 1347–1421.

Aumann, Robert J. 1976. "Agreeing to Disagree." *The Annals of Statistics*, 4(6): 1236–1239.

Baldauf, Markus, and Joshua Mollner. 2014. "High-Frequency Trade and Market Performance." *Working Paper.*

Barberis, Nicholas, and Richard Thaler. 2003. "A Survey of Behavioral Finance." *Handbook of the Economics of Finance*, 1(B): 1053–1128.

Berentsen, Aleksander, Samuel Huber, and Alessandro Marchesiani. 2015. "The Societal Benefit of a Financial Transaction Tax." *Working Paper*, , (176).

Biais, Bruno, and Jean-Charles Rochet. 2020. "Taxing Financial Transactions." *HEC Working Paper.*

Blume, Lawrence E., Timothy Cogley, David A. Easley, Thomas J. Sargent, and Viktor Tsyrennikov. 2013. "Welfare, Paternalism and Market Incompleteness." *Working Paper.*

Bond, Philip, Alex Edmans, and Itay Goldstein. 2012. "The Real Effects of Financial Markets." *Annual Review of Financial Economics*, 4(1): 339–360.

Brunnermeier, Markus K., Alp Simsek, and Wei Xiong. 2014. "A Welfare Criterion for Models with Distorted Beliefs." *The Quarterly Journal of Economics*, 129(4): 1753–1797.

Budish, Eric, Peter Cramton, and John Shim. 2015. "The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response." *The Quarterly Journal of Economics*, 130(4): 1547–1621.

Burman, Leonard E., William G. Gale, Sarah Gault, Bryan Kim, Jim Nunns, and Steve Rosenthal. 2016. "Financial Transaction Taxes in Theory and Practice." *National Tax Journal*, 69(1): 171–216.

Cai, Jinghan, Jibao He, Wenxi Jiang, and Wei Xiong. 2017. "The whack-a-mole game: Tobin tax and trading frenzy." *Working Paper.*

Campbell, John Y. 2016. "Restoring Rational Choice: The Challenge of Consumer Financial Regulation." *American Economic Review*, 106(5): 1–30.

Campbell, John Y. 2017. *Financial Decisions and Markets: A Course in Asset Pricing.* Princeton University Press.

Campbell, John Y., and Kenneth A. Froot. 1994. "International Experiences with Securities Transaction Taxes." In *The Internationalization of Equity Markets.* 277–308. University of Chicago Press.

Cochrane, John. 2013. "Finance: Function Matters, Not Size." *Journal of Economic Perspectives*, 27(2): 29–50.

Cochrane, John. 2014. "Challenges for Cost-Benefit Analysis of Financial Regulation." *The Journal of Legal Studies*, 43(S2): S63–S105.

Coelho, Maria. 2014. "Dodging Robin Hood: Responses to France and Italy's Financial Transaction Taxes." *UC Berkeley Working Paper.*

Colliard, Jean-Edouard, and Peter Hoffmann. 2017. "Financial Transaction Taxes, Market Composition, and Liquidity." *The Journal of Finance*, 72(6): 2685–2716.

Constantinides, George M. 1986. "Capital Market Equilibrium with Transaction Costs." *Journal of Political Economy*, 94(4): 842–862.

Dang, Tri Vi, and Florian Morath. 2015. "The Taxation of Bilateral Trade with Endogenous Information." *Working Paper.*

Dávila, Eduardo, and Ansgar Walther. 2021. "Corrective Regulation with Imperfect Instruments." *Working Paper.*

Dávila, Eduardo, and Anton Korinek. 2018. "Pecuniary Externalities in Economies with Financial Frictions." *The Review of Economic Studies*, 85(1): 352–395.

Dávila, Eduardo, and Cecilia Parlatore. 2021. "Trading Costs and Informational Efficiency." *The Journal of Finance*, 76(3): 1471–1539.

Diamond, Peter A. 1973. "Consumption Externalities and Imperfect Corrective Pricing." *The Bell Journal of Economics and Management Science*, 4(2): 526–538.

Dieler, Tobias, Swarnava Biswas, Giacomo Calzolari, and Fabio Castiglionesi. 2020. "Asset Trade, Real Investment and a Tilting Financial transaction tax." *Working Paper.*

Duffie, Darrell. 2014. "Challenges to a Policy Treatment of Speculative Trading Motivated by Differences in Beliefs." *The Journal of Legal Studies*, 43(S2): S173–S182.

Eyster, Erik, and Matthew Rabin. 2005. "Cursed Equilibrium." *Econometrica*, 73(5): 1623–1672.

Farhi, Emmanuel, and Xavier Gabaix. 2015. "Optimal Taxation with Behavioral Agents." *Working Paper.*

Fortuin, Cees M., Pieter W. Kasteleyn, and Jean Ginibre. 1971. "Correlation Inequalities on some Partially Ordered Sets." *Communications in Mathematical Physics*, 22(2): 89–103.

Gayer, Gabrielle, Itzhak Gilboa, Larry Samuelson, and David Schmeidler. 2014. "Pareto Efficiency with Different Beliefs." *The Journal of Legal Studies*, 43(S2): S151–S171.

Gilboa, Itzhak, Larry Samuelson, and David Schmeidler. 2014. "No-Betting-Pareto Dominance." *Econometrica*, 82(4): 1405–1442.

Glosten, Lawrence R., and Paul R. Milgrom. 1985. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics*, 14(1): 71–100.

Goulder, Lawrence H. 1995. "Environmental Taxation and the Double Dividend: A Reader's Guide." *International Tax and Public Finance*, 2: 157–183.

Greene, William H. 2003. *Econometric Analysis.* Prentice Hall.

Grossman, Sanford J., and Joseph E. Stiglitz. 1980. "On the Impossibility of Informationally Efficient Markets." *American Economic Review*, 70(3): 393–408.

Gruber, Jonathan, and Botond Koszegi. 2001. "Is Addiction "rational"? Theory and Evidence." *The Quarterly Journal of Economics*, 116(4): 1261–1303.

Habermeier, Karl, and Andrei A. Kirilenko. 2003. "Securities Transaction Taxes and Financial Markets." *IMF Economic Review*, 50: 165–180.

Harberger, Arnold C. 1964. "The Measurement of Waste." *American Economic Review*, 54(3): 58–76.

Harrison, J. Michael, and David M. Kreps. 1978. "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations." *The Quarterly Journal of Economics*, 92(2): 323–336.

Heyerdahl-Larsen, Christian, and Johan Walden. 2014. "Efficiency and Distortions in a Production Economy with Heterogeneous Beliefs." *Working Paper.*

Hong, Harrison, and Jeremy C. Stein. 2007. "Disagreement and the Stock Market." *The Journal of Economic Perspectives*, 21(2): 109–128.

Jones, Charles M., and Paul J. Seguin. 1997. "Transaction Costs and Price Volatility: Evidence from Commission Deregulation." *American Economic Review*, 87(4): 728–737.

Koijen, Ralph SJ, and Motohiro Yogo. 2019. "A demand system approach to asset pricing." *Journal of Political Economy*, 127(4): 1475–1515.

Kopczuk, Wojciech. 2003. "A Note on Optimal Taxation in the Presence of Externalities." *Economics Letters*, 80(1): 81–86.

Kreps, David M. 2012. *Microeconomic Foundations I: Choice and Competitive Markets.* Princeton University Press.

Lintner, John. 1969. "The Aggregation of Investor's Diverse Judgments and Preferences in Purely Competitive Security Markets." *Journal of Financial and Quantitative Analysis*, 4(4): 347–400.

McCulloch, Neil, and Grazia Pacillo. 2011. "The Tobin Tax - A Review of the Evidence." *IDS Research Reports*, 2011: 1–77.

Miller, Edward M. 1977. "Risk, uncertainty, and divergence of opinion." *The Journal of Finance*, 32(4): 1151–1168.

Mullainathan, Sendhil, Joshua Schwartzstein, and William Congdon. 2012. "A Reduced-Form Approach to Behavioral Public Finance." *Annual Review of Economics*, 4: 511–540.

O'Donoghue, Ted, and Matthew Rabin. 2006. "Optimal Sin Taxes." *Journal of Public Economics*, 90(10): 1825–1849.

Panageas, Stavros. 2005. "The Neoclassical Theory of Investment in Speculative Markets." *Working Paper.*

Posner, Eric, and E. Glen Weyl. 2013. "Benefit-Cost Analysis for Financial Regulation." *American Economic Review*, 103(3): 393–397.

Pratt, John W. 1964. "Risk Aversion in the Small and in the Large." *Econometrica*, 32(1): 122–136.

Roll, Richard. 1989. "Price Volatility, International Market Links, and Their Implications for Regulatory Policies." *Journal of Financial Services Research*, 3(2): 211–246.

Ross, Stephen A. 1989. "Using Tax Policy to Curb Speculative Short-Term Trading." *Journal of Financial Services Research*, 3: 117–120.

Saez, Emmanuel, and Stefanie Stantcheva. 2016. "Generalized social marginal welfare weights for optimal tax theory." *American Economic Review*, 106(1): 24–45.

Sandmo, Agnar. 1975. "Optimal Taxation in the Presence of Externalities." *The Swedish Journal of Economics*, 77(1): 86–98.

Sandmo, Agnar. 1985. "The Effects of Taxation on Savings and Risk Taking." *Handbook of Public Economics*, 1: 265–311.

Sandroni, Alvaro, and Francesco Squintani. 2007. "Overconfidence, Insurance, and Paternalism." *American Economic Review*, 97(5): 1994–2004.

Scheinkman, Jose A., and Wei Xiong. 2003. "Overconfidence and Speculative Bubbles." *Journal of Political Economy*, 111(6): 1183–1220.

Scheuer, Florian. 2013. "Optimal Asset Taxes in Financial Markets with Aggregate Uncertainty." *Review of Economic Dynamics*, 16(3): 405–420.

Schwert, G. William, and Paul J. Seguin. 1993. "Securities Transaction Taxes: An Overview of Costs, Benefits and Unresolved Questions." *Financial Analysts Journal*, 49(5): 27–35.

Simsek, Alp. 2013. "Speculation and Risk Sharing with New Financial Assets." *The Quarterly Journal of Economics*, 128(3): 1365–1396.

Simsek, Alp. 2021. "The Macroeconomics of Financial Speculation." *NBER Working Paper.*

Spinnewijn, Johannes. 2015. "Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs." *Journal of the European Economic Association*, 13(1): 130–167.

Stiglitz, Joseph E. 1989. "Using Tax Policy to Curb Speculative Short-Term Trading." *Journal of Financial Services Research*, 3(2): 101–115.

Summers, Lawrence H., and Victoria P. Summers. 1989. "When Financial Markets Work Too Well: A Cautious Case for a Securities Transactions Tax." *Journal of Financial Services Research*, 3: 261–286.

Tobin, James. 1978. "A Proposal for International Monetary Reform." *Eastern Economic Journal*, 4(3): 153–159.

ul Haq, Mahbub, Isabelle Kaul, and Inge Grunberg, ed. 1996. *The Tobin Tax: Coping with Financial Volatility.* Oxford University Press.

Umlauf, Steven R. 1993. "Transaction Taxes and the Behavior of the Swedish Stock Sarket." *Journal of Financial Economics*, 33(2): 227–240.

Vayanos, Dimitri, and Jiang Wang. 2012. "Liquidity and Asset Returns under Asymmetric Information and Imperfect Competition." *Review of Financial Studies*, 25(5): 1339–1365.

Vives, Xavier. 2017. "Endogenous Public Information and Welfare in Market Games." *The Review of Economic Studies*, 84(2): 935–963.

Weyl, E. Glen. 2007. "Is Arbitrage Socially Beneficial?" *Princeton University Working Paper.*

Weyl, E. Glen. 2019. "Price Theory." *Journal of Economic Literature*, 57(2): 329–384.

Xiong, Wei. 2013. "Bubbles, Crises, and Heterogeneous Beliefs." *NBER Working Paper no. 18905.*

Appendix

A Proofs and derivations: Section 3

Properties of investors' portfolio problem Given a price P_1 and a tax τ, investors solve $\max_{X_{1i}} J(X_{1i})$, where $J(X_{1i})$ denotes the objective function of investors, introduced in Equation (4) in the text, and reproduced here:

$$J(X_{1i}) = \left[\mathbb{E}_i[D] - A_i \mathbb{C}ov[M_{2i}, D] - P_1\right] X_{1i} + P_1 X_{0i} - \tau P_1 |\Delta X_{1i}| + T_{1i} - \frac{A_i}{2} \mathbb{V}ar[D](X_{1i})^2. \tag{24}$$

The first and second order conditions in the regions in which the problem is differentiable are respectively given by

$$J'(X_{1i}) = \mathbb{E}_i[D] - A_i \mathbb{C}ov[M_{2i}, D] - P_1 - \tau P_1 \operatorname{sgn}(\Delta X_{1i}) - A_i \mathbb{V}ar[D] X_{1i} = 0, \tag{25}$$

$$J''(X_{1i}) = -A_i \mathbb{V}ar[D] < 0. \tag{26}$$

When the tax rate is strictly positive, $\lim_{X_{1i} \to X_{0i}^-} J'(X_{1i}) > \lim_{X_{1i} \to X_{0i}^+} J'(X_{1i})$, so the transaction tax generates a concave kink for investors' objective function at $X_{1i} = X_{0i}$. The existence of a concave kink combined with the fact that $J''(\cdot) < 0$, $\lim_{X_{1i} \to +\infty} J'(X_{1i}) = -\infty$, and $\lim_{X_{1i} \to -\infty} J'(X_{1i}) = +\infty$ jointly imply that the solution to the investors' problem is unique, and that it can be reached either at an interior optimum or at the kink. Equation (5) provides a full characterization of the solution. When taxes are positive, for a given price P_1, an individual investor i decides not to trade when

$$\left| \frac{\mathbb{E}_i[D] - A_i \mathbb{C}ov[M_{2i}, D] - A_i \mathbb{V}ar[D] X_{0i} - P_1}{P_1} \right| \leq \tau.$$

When the tax rate is negative (a subsidy), $\lim_{X_{1i} \to X_{0i}^-} J'(X_{1i}) < \lim_{X_{1i} \to X_{0i}^+} J'(X_{1i})$, so the transaction subsidy generates a convex kink for investors' objective function at $X_{1i} = X_{0i}$. The existence of a convex kink combined with the fact that $J''(\cdot) < 0$, $\lim_{X_{1i} \to +\infty} J'(X_{1i}) = -\infty$, and $\lim_{X_{1i} \to -\infty} J'(X_{1i}) = +\infty$ jointly imply that the solution to the investors' problem is reached at an interior optimum. See Figure A.7 for a graphical illustration of both cases.

Lemma 1. (Competitive equilibrium with taxes)

Proof. a) [Existence/Uniqueness] For given set of primitives and a tax rate τ, let us define an aggregate excess demand function $Z(P_1) \equiv \int_{i \in \mathcal{T}(P_1)} \Delta X_{1i}(P_1) \, dF(i)$, where individual net demands $\Delta X_{1i}(P_1)$ are determined by Equation (5) and $\mathcal{T}(P_1)$ denotes the set of investors with non-zero net trading demands for a given price P_1. A price $P_1^\star$ is part of an equilibrium if $Z(P_1^\star) = 0$, which guarantees that market clearing is satisfied. The continuity of $Z(P_1)$ follows trivially. It is equally straightforward to show that $\lim_{P_1 \to \infty} Z(P_1) = -\infty$ and $\lim_{P_1 \to -\infty} Z(P_1) = \infty$. These three properties are sufficient to establish that an equilibrium always exist, applying the Intermediate Value Theorem.

To establish uniqueness, we must study the properties of $Z'(P_1)$, which can be explicitly computed as follows

$$Z'(P_1) = \int_{i \in \mathcal{T}(P_1)} \frac{\partial X_{1i}(P_1)}{\partial P_1} dF(i) = -\int_{i \in \mathcal{T}(P_1)} \frac{1 + \operatorname{sgn}(\Delta X_{1i}) \tau}{A_i \mathbb{V}ar[D]} dF(i) \leq 0,$$

where the first equality follows from Leibniz's rule. Because the distribution of investors is continuous, $Z(P_1)$ is differentiable.[37] Note that $Z'(P_1)$ is strictly negative when the region $\mathcal{T}(P_1)$ is non-empty. This is sufficient to conclude that if there exists a price $P_1^\star$ that i) satisfies $Z(P_1^\star) = 0$ and ii) is such that the set of active investors has positive measure, the equilibrium must be unique, because $Z'(P_1^\star) < 0$ at that point and $Z'(P_1) \leq 0$ everywhere else. However, a price $P_1^\star$ that satisfies $Z(P_1^\star) = 0$ but that implies that the set of investors who actively trade has zero measure can also exist. In that case, there is generically a range of prices that are consistent with no-trade.

Therefore, trading volume is always pinned down, although there is an indeterminacy in the set of possible asset prices in no-trade equilibria. In that sense, the equilibrium is essentially unique.

[37] When the distribution of investors $F(\cdot)$ is continuous, $P_1(\tau)$, $X_{1i}(\tau)$, and $V(\tau)$ are continuously differentiable whenever trading volume is positive in equilibrium. All the economic insights from this paper remain valid when the distribution of investors can have mass points, as shown in earlier versions of this paper. Assuming a continuous probability distribution simplifies all formal characterizations by preserving differentiability.

b) [Volume response] The change in trading volume is given by $\frac{d\mathcal{V}}{d\tau} = \int_{i\in\mathcal{B}(P_1)} \frac{dX_{1i}}{d\tau} dF(i)$. It follows that $\frac{dX_{1i}}{d\tau} = \frac{\partial X_{1i}}{\partial\tau} + \frac{\partial X_{1i}}{\partial P_1}\frac{dP_1}{d\tau}$ can be expressed as

$$\frac{dX_{1i}}{d\tau} = \frac{\partial X_{1i}}{\partial \tau} \underbrace{\left[1 - (\operatorname{sgn}(\Delta X_{1i}) + \tau) \frac{\int_{i\in\mathcal{T}(P_1)} \frac{\operatorname{sgn}(\Delta X_{1i})}{A_i} dF(i)}{\int_{i\in\mathcal{T}(P_1)} \frac{1+\operatorname{sgn}(\Delta X_{1i})\tau}{A_i} dF(i)}\right]}_{\equiv \varepsilon_i}, \tag{27}$$

where $\frac{\partial X_{1i}}{\partial\tau} = \frac{-P_1 \operatorname{sgn}(\Delta X_{1i})}{A_i \mathbb{V}ar[D]}$, $\frac{\partial X_{1i}}{\partial P_1} = \frac{-(1+\operatorname{sgn}(\Delta X_{1i})\tau)}{A_i\mathbb{V}ar[D]}$, and it is straightforward to show that $\varepsilon_i > 0$ for both buyers and sellers. Equation (27) implies that $\frac{dX_{1i}}{d\tau} < 0$ for buyers, while $\frac{dX_{1i}}{d\tau} > 0$ for sellers, implying that trading volume decreases with τ.

c) [Price response] The price P_1 is continuous and differentiable in τ when the distribution of investors is continuous. Using again Leibniz's rule, the derivative $\frac{dP_1}{d\tau}$ can be expressed as

$$\frac{dP_1}{d\tau} = \frac{\int_{i\in\mathcal{T}(P_1)} \frac{\partial X_{1i}}{\partial\tau} dF(i)}{-\int_{i\in\mathcal{T}(P_1)} \frac{\partial X_{1i}}{\partial P_1} dF(i)} = \frac{-\left(\int_{i\in\mathcal{B}(P_1)} \frac{P_1}{A_i\mathbb{V}ar[D]} dF(i) - \int_{i\in\mathcal{S}(P_1)} \frac{P_1}{A_i\mathbb{V}ar[D]} dF(i)\right)}{\int_{i\in\mathcal{T}(P_1)} \frac{1+\operatorname{sgn}(\Delta X_{1i})\tau}{A_i\mathbb{V}ar[D]} dF(i)}. \tag{28}$$

It follows that $\frac{dP_1}{d\tau} < 0$ if $\int_{i\in\mathcal{B}(P_1)} \frac{1}{A_i} dF(i) > \int_{i\in\mathcal{S}(P_1)} \frac{1}{A_i} dF(i)$ and vice versa. Under Assumption [S], which implies that $\frac{1}{A_i}$ is constant and that the share of buyers equals the share of sellers, the numerator of Equation (28) is zero, implying that $\frac{dP_1}{d\tau} = 0$. □

A sufficient (but not necessary) condition for P_1 to be strictly positive is that the expected dividend of every investor is large enough when compared to his risk-bearing capacity, that is:

$$\mathbb{E}_i[D] > A_i(\mathbb{C}ov[M_{2i}, D] + \mathbb{V}ar[D]Q), \ \forall i.$$

Note also that one must assume that $\mathbb{V}ar[M_{2i}]$ is sufficiently large to guarantee that the variance-covariance matrix of the joint distribution of M_{2i} and D is positive semi-definite. If we allowed P_1 to take negative values, Equation (2) would become $\tau|P_1||\Delta X_{1i}|$.

B Proofs and derivations: Section 4

To simplify the exposition, I suppress the explicit dependence on τ of many variables and sets at times, e.g., X_{1i}, P_1, $\mathcal{B}$, and $\mathcal{S}$, instead of $X_{1i}(\tau)$, $P_1(\tau)$, $\mathcal{B}(\tau)$, and $\mathcal{S}(\tau)$.

Lemma 2. (Marginal welfare impact of tax changes)

Proof. a) [Individual welfare impact] The derivative of the planner's objective function is given by $\frac{dV^p}{d\tau} = \int \frac{dV_i^p}{d\tau} dF(i)$, where investor i's certainty equivalent from the planner's perspective $V_i^p(\tau)$ is defined in Equation (9). Consequently, $\frac{dV_i^p}{d\tau}$ corresponds to

$$\frac{dV_i^p}{d\tau} = [\mathbb{E}_p[D] - \mathbb{E}_i[D] + \operatorname{sgn}(\Delta X_{1i}) P_1\tau] \frac{dX_{1i}}{d\tau} - \Delta X_{1i}\frac{dP_1}{d\tau} + \frac{d\tilde{T}_{1i}}{d\tau}. \tag{29}$$

The derivation of Equation (29) uses the envelope theorem for the choice of X_{1i} and for the extensive margin choice between trading and no trading, which are both made optimally. Note that $\frac{dV_i^p}{d\tau} = \frac{d\tilde{T}_{1i}}{d\tau}$ for investors who do not trade at the margin, because $\frac{dX_{1i}}{d\tau} = 0$ and $\Delta X_{1i} = 0$. Hence, a marginal tax change has no effect on the welfare of those investors who decide not to trade, besides potential tax rebates.

b) [Aggregate welfare impact] We can aggregate across investors to express the change in social welfare as follows

$$\frac{dV^p}{d\tau} = \int_{i\in\mathcal{T}(\tau)} [-\mathbb{E}_i[D] + \operatorname{sgn}(\Delta X_{1i}) P_1\tau] \frac{dX_{1i}}{d\tau} dF(i), \tag{30}$$

where Equation (30) follows from market clearing, which implies $\int \Delta X_{1i} dF(i) = 0$ and $\int \frac{dX_{1i}}{d\tau} dF(i) = 0$, and from the assumption that tax revenues are rebated to investors, which implies that $\int \frac{d\tilde{T}_{1i}}{d\tau} dF(i) = 0$. □

Note that the marginal welfare impact of a tax change from the perspective of investor i is given by

$$\frac{dV_i^i}{d\tau} = \operatorname{sgn}(\Delta X_{1i}) P_1\tau \frac{dX_{1i}}{d\tau} - \Delta X_{1i}\frac{dP_1}{d\tau} + \frac{d\tilde{T}_{1i}}{d\tau}. \tag{31}$$

Proposition 1. (Optimal financial transaction tax)

Proof. a) [Optimal tax formula] Starting from Equation (30), it follows that the optimal transaction tax τ^* must satisfy the following expression

$$\tau^* = \frac{\int_{i\in\mathcal{T}(\tau)} \frac{\mathbb{E}_i[D]}{P_1}\frac{dX_{1i}}{d\tau} dF(i)}{\int_{i\in\mathcal{T}(\tau)} \operatorname{sgn}(\Delta X_{1i})\frac{dX_{1i}}{d\tau} dF(i)} = \frac{1}{2}\frac{\int_{i\in\mathcal{T}(\tau)} \frac{\mathbb{E}_i[D]}{P_1}\frac{dX_{1i}}{d\tau} dF(i)}{\int_{i\in\mathcal{B}(\tau)} \frac{dX_{1i}}{d\tau} dF(i)}$$

$$= \frac{1}{2}\left[\overbrace{\int_{i\in\mathcal{B}(\tau)} \frac{\mathbb{E}_i[D]}{P_1} \underbrace{\frac{\frac{dX_{1i}}{d\tau}}{\int_{i\in\mathcal{B}(\tau)}\frac{dX_{1i}}{d\tau}dF(i)}}_{\omega_i^{\mathcal{B}}} dF(i)}^{\Omega_{\mathcal{B}}} - \overbrace{\int_{i\in\mathcal{S}(\tau)} \frac{\mathbb{E}_i[D]}{P_1} \underbrace{\frac{\frac{dX_{1i}}{d\tau}}{\int_{i\in\mathcal{S}(\tau)}\frac{dX_{1i}}{d\tau}dF(i)}}_{\omega_i^{\mathcal{S}}} dF(i)}^{\Omega_{\mathcal{S}}}\right].$$

This derivation exploits the fact that $\int_{i\in\mathcal{B}(\tau)} \frac{dX_{1i}}{d\tau} dF(i) = -\int_{i\in\mathcal{S}(\tau)} \frac{dX_{1i}}{d\tau} dF(i)$, as well as the fact that $\int_{i\in\mathcal{T}(\tau)} \frac{\mathbb{E}_i[D]}{P_1}\frac{dX_{1i}}{d\tau} dF(i) = \int_{i\in\mathcal{B}(\tau)} \frac{\mathbb{E}_i[D]}{P_1}\frac{dX_{1i}}{d\tau} dF(i) + \int_{i\in\mathcal{S}(\tau)} \frac{\mathbb{E}_i[D]}{P_1}\frac{dX_{1i}}{d\tau} dF(i)$.

b) [Sign of the optimal tax] Given the properties of the planner's problem, established below, it is sufficient to show that $\left.\frac{dV^p}{d\tau}\right|_{\tau=0} > 0$ to guarantee that the optimal policy is a positive tax. We can express $\left.\frac{dV^p}{d\tau}\right|_{\tau=0}$ as follows

$$\left.\frac{dV^p}{d\tau}\right|_{\tau=0} = -\int_{i\in\mathcal{T}(0)} \mathbb{E}_i[D] \left.\frac{dX_{1i}}{d\tau}\right|_{\tau=0} dF(i) = -\mathbb{C}ov_F\left(\mathbb{E}_i[D], \left.\frac{dX_{1i}}{d\tau}\right|_{\tau=0}\right)$$

$$= \frac{P_1}{\mathbb{V}ar[D]}\left[\mathbb{C}ov_F\left(\mathbb{E}_i[D], \frac{\mathbb{I}\left[\left.\Delta X_{1i}\right|_{\tau=0} > 0\right]}{A_i}\right)\varepsilon_B - \mathbb{C}ov_F\left(\mathbb{E}_i[D], \frac{\mathbb{I}\left[\left.\Delta X_{1i}\right|_{\tau=0} < 0\right]}{A_i}\right)\varepsilon_S\right],$$

where the sub-index F denotes cross-sectional moments. Hence, $\left.\frac{dV^p}{d\tau}\right|_{\tau=0}$ is positive if $\mathbb{C}ov_F\left(\mathbb{E}_i[D], \frac{\mathbb{I}\left[\left.\Delta X_{1i}\right|_{\tau=0}>0\right]}{A_i}\right) > 0$, since that result directly implies that $\mathbb{C}ov_F\left(\mathbb{E}_i[D], \frac{\mathbb{I}\left[\left.\Delta X_{1i}\right|_{\tau=0}<0\right]}{A_i}\right) < 0$. Under the assumption that all cross-sectional distributions are independent, we can decompose equilibrium net trading volume as

$$\Delta X_{1i} = \underbrace{\frac{\mathbb{E}_i[D] - \mathbb{E}_F[\mathbb{E}_i[D]] + A\mathbb{E}_F[\mathbb{C}ov[M_{2i}, D]] + A\mathbb{V}ar[D]Q}{A_i\mathbb{V}ar[D]}}_{\equiv Z_1} \underbrace{-\frac{\mathbb{C}ov[M_{2i}, D]}{\mathbb{V}ar[D]} - X_{0i}}_{\equiv Z_2}, \tag{32}$$

where Z_1 and Z_2 are defined in Equation (32) and $A \equiv \left(\mathbb{E}_F\left[\frac{1}{A_i}\right]\right)^{-1}$. For a low cross-sectional dispersion of risk tolerances/risk aversion coefficients, that is, when $\mathbb{V}ar\left[\frac{1}{A_i}\right] \approx 0$, the sign of the covariance of interest is identical to sign of $\mathbb{C}ov_F(Z_1, g(Z_1 + Z_2))$, where Z_1 and Z_2, given their definition above, are independent random variables, and $g(\cdot)$ is an increasing function. It then follows directly from the FKG inequality (Fortuin, Kasteleyn and Ginibre, 1971) that $\mathbb{C}ov_F(Z_1, g(Z_1 + Z_2))$ is positive, which allows us to conclude that $\left.\frac{dV^p}{d\tau}\right|_{\tau=0} > 0$ when fundamental and non-fundamental motives to trade are independently distributed across the population.

c) [Irrelevance of the planner's belief] The claim follows directly from Equation (30). The fact that the risky asset is in fixed supply, which implies that $\int \frac{dX_{1i}}{d\tau} dF(i) = 0$, combined with the fact that investors' welfare (measured as certainty equivalents) is linear in $\mathbb{E}_p[D] X_{1i}$, are necessary for the irrelevance result to hold. ☐

Lemma 3. (Failure of quasi-concavity of planner's objective)

Proof. The planner's objective function $V^p(\tau)$ is continuous as long as the distribution of investors is also continuous. Hence, the Extreme Value Theorem guarantees that there exists an optimal τ^*. The first order condition of the planner's problem is given by Equation (30).

Establishing the uniqueness of the optimum and its properties requires the study of $\frac{d^2V^p}{d\tau^2}$. I show that the planner's objective function is concave (has a negative second derivative) on the intensive margin, although changes in the composition of marginal investors on the extensive margin cause non-concavities.[38] Formally, the second order condition of the planner's

[38] Note that quadratic taxes, often used as a tractable approximation to linear taxes, do not generate extensive margin adjustments, since it is generically optimal for all investors to trade. Consequently, quadratic taxes cannot generate failures of quasi-concavity of the planner's objective.

problem is given by

$$
\begin{aligned}
\frac{d^2 V^p}{d\tau^2} = & \frac{d(P_1\tau)}{d\tau}\int_{i\in\mathcal{T}(\tau)} \operatorname{sgn}(\Delta X_{1i})\frac{dX_{1i}}{d\tau}dF(i) + \int_{i\in\mathcal{T}(\tau)}\left[-\mathbb{E}_i[D] + \operatorname{sgn}(\Delta X_{1i})P_1\tau\right]\frac{d^2X_{1i}}{d\tau^2}dF(i) \\
& - \int_{\tilde{\mathcal{B}}(\tau)} \underbrace{\left[-\mathbb{E}_i[D] + P_1\tau\right]\frac{dX_{1i}}{d\tau}}_{\frac{dV_i^p}{d\tau}} dF(i) - \int_{\tilde{\mathcal{S}}(\tau)} \underbrace{\left[-\mathbb{E}_i[D] - P_1\tau\right]\frac{dX_{1i}}{d\tau}}_{\frac{dV_i^p}{d\tau}} dF(i), \qquad (33)
\end{aligned}
$$

where $\tilde{\mathcal{B}}(\tau)$ and $\tilde{\mathcal{S}}(\tau)$ correspond to the set of buyers and sellers who are indifferent between trading and not trading, and are defined by the following surfaces

$$
\begin{aligned}
\tilde{\mathcal{B}}(\tau) &= \{i : \mathbb{E}_i[D] - A_i\mathbb{C}ov[M_{2i}, D] - P_1 - \tau P_1 - A_i\mathbb{V}ar[D]X_{1i} = 0\} \\
\tilde{\mathcal{S}}(\tau) &= \{i : \mathbb{E}_i[D] - A_i\mathbb{C}ov[M_{2i}, D] - P_1 + \tau P_1 - A_i\mathbb{V}ar[D]X_{1i} = 0\},
\end{aligned}
$$

when $\tau \geq 0$ (they are empty sets when $\tau < 0$) and $\int_{\tilde{\mathcal{B}}(\tau)}$ and $\int_{\tilde{\mathcal{S}}(\tau)}$ denote line/surface-integrals. It is possible to show that $\frac{d^2X_{1i}}{d\tau^2} = 2\frac{dX_{1i}}{d\tau}\frac{dP_1}{d\tau}\frac{1}{P_1} + e.m.$, where $e.m.$ denotes extensive margin terms that involve changes in the composition of investors, similar to the last two terms in Equation (33). The first term in Equation (33) is always negative. The sign of the second term is ambiguous, but it is always equal to zero at an interior optimum, leaving aside the extensive margin effects. The final two terms capture extensive margin effects, and can take on any sign. The last two terms can in general be written as $\int_{i\in\tilde{\mathcal{T}}(\tau)} \frac{dV_i^p}{d\tau}dF(i) \gtreqless 0$, where $\tilde{\mathcal{T}}(\tau)$ denotes the set of investors indifferent between trading and not trading. This is sufficient to show part a) of the Lemma.

It follows from Equation (33) that, at any interior optimum without extensive margin effects,

$$
\left.\frac{d^2V^p}{d\tau^2}\right|_{\tau=\tau^*,\, e.m.=0} = \frac{d(P_1\tau)}{d\tau}\int_{i\in\mathcal{T}(\tau)}\operatorname{sgn}(\Delta X_{1i})\frac{dX_{1i}}{d\tau}dF(i)\Bigg|_{\tau=\tau^*} \leq 0,
$$

because $\frac{d(P_1\tau)}{d\tau} = P_1\left(1 - \frac{\tau\int\frac{\operatorname{sgn}(\Delta X_{1i})}{A_i}dF(i)}{\int\frac{1+\operatorname{sgn}(\Delta X_{1i})\tau}{A_i}dF(i)}\right) > 0$. Because $\frac{dV^p}{d\tau}$ is differentiable given that the distribution of investors is continuous and the measure of active investors is non-zero, this result implies that, when there are no extensive margin effects (or when they are small), any interior optimum must be a maximum. If extensive margin effects are large, there could potentially be multiple interior optima as illustrated in Figure 4 in the text. Because there are no extensive margin changes when $\tau < 0$, the planner's objective function is concave in that region, implying that $\left.\frac{dV^p}{d\tau}\right|_{\tau=0} > 0$ is a sufficient condition for $\tau^* > 0$.

Note that it is possible to normalize the aggregate marginal impact of a tax change by the number of active investors' as follows

$$
\frac{\frac{dV^p}{d\tau}}{\int_{i\in\mathcal{T}(\tau)}dF(i)} = \mathbb{E}_{\mathcal{T}(\tau)}\left[\left[-\mathbb{E}_i[D] + \operatorname{sgn}(\Delta X_{1i})P_1\tau\right]\frac{dX_{1i}}{d\tau}\right].
$$

Given that $\int_{i\in\mathcal{T}(\tau)}dF(i)$ takes positive values, a sufficient condition for quasi-concavity is that $\mathbb{E}_{\mathcal{T}(\tau)}\left[\left[-\mathbb{E}_i[D] + \operatorname{sgn}(\Delta X_{1i})P_1\tau\right]\frac{dX_{1i}}{d\tau}\right]$ is decreasing in τ. Under Assumption [S], it follows that

$$
\frac{\frac{dV^p}{d\tau}}{\int_{i\in\mathcal{T}(\tau)}dF(i)} = \frac{P_1}{A\mathbb{V}ar[D]}2P_1\left(\frac{\mathbb{E}_{\mathcal{B}(\tau)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right] - \mathbb{E}_{\mathcal{S}(\tau)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]}{2} - \tau\right),
$$

so a sufficient condition for the planner's objective to be quasi-concave is that

$$
\frac{\partial\mathbb{E}_{\mathcal{B}(\tau)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]}{\partial\tau} - \frac{\partial\mathbb{E}_{\mathcal{S}(\tau)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]}{\partial\tau} < 2, \qquad (34)
$$

which establishes part b) of the Lemma 3. Since P_1 is exclusively a function of primitives under Assumption [S], and the regions of buyers and sellers are purely a function of primitives once P_1 is determined, Equation (34) provides an explicit restriction on the set of primitives of the model. □

Lemma 4. (No-trade is optimal if all trade is belief-motivated)

Proof. The planner's objective, $V^p = \int V_i^p dF(i)$, can be expressed under no-trade as follows

$$V^p_{\text{no-trade}} = \left(\mathbb{E}_p[D] - A\mathbb{C}ov[M_2, D]\right)Q - \frac{A}{2}\mathbb{V}ar[D](Q)^2.$$

The planner's objective for any other allocations involving $X_{1i} \neq Q$ is given by

$$V^p_{\text{trade}} = \left(\mathbb{E}_p[D] - A\mathbb{C}ov[M_2, D]\right)Q - \frac{A}{2}\mathbb{V}ar[D]\int (X_{1i})^2 dF(i).$$

Since $\int X_{1i}dF(i) = Q$, a straight application of Jensen's inequality immediately implies that $V^p_{\text{no-trade}} > V^p_{\text{trade}}$ whenever $X_{1i} \neq Q$ for at least a single investor. $\square$

Proposition 2. (Trading volume implementation)

Proof. a) [Trading volume decomposition] Trading volume (in dollars) is defined by

$$P_1\mathcal{V}(\tau) \equiv P_1\int_{i\in\mathcal{B}(\tau)} \Delta X_{1i}dF(i) = \frac{1}{2}\left(\int_{i\in\mathcal{B}(\tau)} P_1\Delta X_{1i}dF(i) - \int_{i\in\mathcal{S}(\tau)} P_1\Delta X_{1i}dF(i)\right).$$

We can express the individual net trade (in dollars) as

$$P_1\Delta X_{1i} = \frac{P_1}{A_i\mathbb{V}ar[D]}\left(\mathbb{E}_i[D] - A_i\mathbb{C}ov[M_{2i}, D] - P_1(1 + \text{sgn}(\Delta X_{1i})\tau) - A_i\mathbb{V}ar[D]X_{0i}\right),$$

which allows us to write trading volume as

$$\begin{aligned}
P_1\mathcal{V}(\tau) &= -\frac{1}{2}\left[\int_{i\in\mathcal{T}(\tau)}\left(\frac{\partial X_{1i}}{\partial\tau}\left(\mathbb{E}_i[D] - A_i\mathbb{C}ov[M_{2i}, D] - P_1(1 + \text{sgn}(\Delta X_{1i})\tau) - A_i\mathbb{V}ar[D]X_{0i}\right)\right)dF(i)\right] \\
&= -\frac{1}{2}\left[\begin{array}{c}\int_{i\in\mathcal{T}(\tau)}\left(\frac{dX_{1i}}{d\tau}\left(\mathbb{E}_i[D] - A_i\mathbb{C}ov[M_{2i}, D] - P_1\,\text{sgn}(\Delta X_{1i})\tau - A_i\mathbb{V}ar[D]X_{0i}\right)\right)dF(i) \\ +\frac{dP_1}{d\tau}\int_{i\in\mathcal{T}(\tau)}\left(-\frac{\partial X_{1i}}{\partial P_1}\right)A_i\mathbb{V}ar[D]\Delta X_{1i}dF(i)\end{array}\right] \\
&= -\frac{1}{2}\left[\int_{i\in\mathcal{T}(\tau)}\left(\frac{dX_{1i}}{d\tau}\left(\mathbb{E}_i[D] - A_i\mathbb{C}ov[M_{2i}, D] - P_1\,\text{sgn}(\Delta X_{1i})\tau - A_i\mathbb{V}ar[D]X_{0i}\right)\right)dF(i)\right] - \frac{d\log P_1}{d\tau}\tau P_1\mathcal{V}(\tau),
\end{aligned}$$

using the fact that

$$-\int_{i\in\mathcal{T}(\tau)}\frac{\partial X_{1i}}{\partial P_1}A_i\mathbb{V}ar[D]\Delta X_{1i}dF(i) = \int_{i\in\mathcal{T}(\tau)}(1 + \text{sgn}(\Delta X_{1i})\tau)\Delta X_{1i}dF(i) = 2\tau\mathcal{V}(\tau).$$

Therefore, we define $\kappa(P_1,\tau) \equiv \frac{1}{1+\tau\frac{d\log P_1}{d\tau}}$ with $\kappa(P_1, 0) = 1$, and express trading volume as

$$\begin{aligned}
P_1\mathcal{V}(\tau) &= \frac{\kappa(P_1,\tau)}{2}\int_{i\in\mathcal{T}(\tau)}\left(\left(-\frac{dX_{1i}}{d\tau}\right)\left(\mathbb{E}_i[D] - A_i\mathbb{C}ov[M_{2i}, D] - P_1\,\text{sgn}(\Delta X_{1i})\tau - A_i\mathbb{V}ar[D]X_{0i}\right)\right)dF(i) \\
&= \Theta_F(\tau) + \Theta_{NF}(\tau) - \Theta_\tau(\tau),
\end{aligned}$$

where each of the elements is given by

$$\begin{aligned}
\Theta_F(\tau) &\equiv \frac{\kappa(P_1,\tau)}{2}\int_{i\in\mathcal{T}(\tau)}\left(-\frac{dX_{1i}}{d\tau}\right)\left(-A_i\mathbb{C}ov[M_{2i}, D] - A_i\mathbb{V}ar[D]X_{0i}\right)dF(i) \\
\Theta_{NF}(\tau) &\equiv \frac{\kappa(P_1,\tau)}{2}\int_{i\in\mathcal{T}(\tau)}\left(-\frac{dX_{1i}}{d\tau}\right)\mathbb{E}_i[D]dF(i) \\
\Theta_\tau(\tau) &\equiv \frac{\kappa(P_1,\tau)}{2}\tau P_1\int_{i\in\mathcal{T}(\tau)}\text{sgn}(\Delta X_{1i})\left(-\frac{dX_{1i}}{d\tau}\right)dF(i).
\end{aligned}$$

When Assumption [S] holds, $\frac{dX_{1i}}{d\tau}$ is constant across investors and $\kappa(P_1,\tau) = 1$, justifying the expressions in the text.

b) [Optimal policy implementation] Note that the optimality condition for the planner characterized in Proposition 1 can be expressed as

$$\int_{i\in\mathcal{T}(\tau)}\frac{dX_{1i}}{d\tau}\mathbb{E}_i[D]dF(i) = \tau P_1\int_{i\in\mathcal{T}(\tau)}\text{sgn}(\Delta X_{1i})\frac{dX_{1i}}{d\tau}dF(i),$$

which is satisfied when $\Theta_{NF}(\tau^*) = \Theta_\tau(\tau^*)$ or, alternatively, when $P_1\mathcal{V}(\tau^*) = \Theta_F(\tau^*)$.

c) [Small tax approximation] First, note that $\frac{d\mathcal{V}}{d\tau} = \frac{1}{2}\int_{i\in\mathcal{T}(\tau)} \operatorname{sgn}(\Delta X_{1i}) \frac{dX_{1i}}{d\tau} dF(i) = \int_{i\in\mathcal{B}(\tau)} \frac{dX_{1i}}{d\tau} dF(i)$. Consequently, it is possible to express $\Theta_\tau(\tau)$ as follows

$$\Theta_\tau(\tau) = -\kappa(P_1,\tau)\,\tau P_1 \frac{d\mathcal{V}}{d\tau}.$$

Since, at the optimum, $\Theta_{NF}(\tau^*) = \Theta_\tau(\tau^*)$, we can write

$$\frac{\Theta_{NF}(\tau^*)}{P_1(\tau^*)\,\mathcal{V}(\tau^*)} = -\tau^*\kappa(P_1,\tau^*)\left.\frac{d\log\mathcal{V}}{d\tau}\right|_{\tau^*} \Rightarrow \tau^* = \frac{\frac{\Theta_{NF}(\tau^*)}{\Theta_F(\tau^*)+\Theta_{NF}(\tau^*)-\Theta_\tau(\tau^*)}}{-\kappa(P_1,\tau^*)\left.\frac{d\log\mathcal{V}}{d\tau}\right|_{\tau^*}}.$$

Using the fact that $\kappa(P_1,0) = 1$ and $\Theta_\tau(0) = 0$, this expression can be approximated around $\tau^* \approx 0$ as follows

$$\tau^* \approx \frac{\frac{\Theta_{NF}(0)}{\Theta_F(0)+\Theta_{NF}(0)}}{-\left.\frac{d\log\mathcal{V}}{d\tau}\right|_{\tau=0}},$$

which corresponds to Equation (17) in the paper. $\square$

Online Appendix

C Proofs and derivations: Section 5

C.1 Theoretical results

Equilibrium price and portfolio allocations The variance-covariance matrix, given by Equation (18) in the text, is positive semi-definite when $\sigma_d^2\sigma_h^2 > (\sigma_{dh})^2$, where σ_{dh} defines the covariance between both random variables, given by $\sigma_{dh} = \rho\sigma_d\sigma_h$. This restriction is implied by the fact that covariance matrices must be positive semi-definite or, equivalently, by the fact that the correlation coefficient is bounded, that is, $\rho \in [-1, 1]$.

Because the joint cross-sectional distribution of beliefs and hedging needs is symmetric, the equilibrium price can be expressed as follows

$$\begin{aligned} P_1 &= \frac{\int_{i\in\mathcal{T}} \left(\mathbb{E}_i[D] - A\left(\mathbb{C}ov[M_{2i}, D] + \mathbb{V}ar[D]\,Q\right)\right) dF(i)}{\int_{i\in\mathcal{T}} dF(i)} \\ &= \mu_d - \mu_h - A\mathbb{V}ar[D]\,Q, \end{aligned} \tag{35}$$

which corresponds to Equation (20) in the text, since $\mu_h = 0$. Note that the equilibrium price P_1 is fully characterized as a function of primitives and is independent of τ. Therefore, we can express equilibrium net trades (when non-zero) as

$$\begin{aligned} \Delta X_{1i} &= \frac{\mathbb{E}_i[D] - A\mathbb{C}ov[M_{2i}, D] - P_1\left(1 + \operatorname{sgn}(\Delta X_{1i})\,\tau\right) - A\mathbb{V}ar[D]\,X_{0i}}{A\mathbb{V}ar[D]} \\ &= \frac{\left(\mathbb{E}_i[D] - \mathbb{E}_{\mathcal{T}}[\mathbb{E}_i[D]]\right) - \left(A\mathbb{C}ov[M_{2i}, D] - \mathbb{E}_{\mathcal{T}}[A\mathbb{C}ov[M_{2i}, D]]\right) - \operatorname{sgn}(\Delta X_{1i})\,\tau P_1}{A\mathbb{V}ar[D]} \\ &= \frac{\varepsilon_{di} - \varepsilon_{hi} - \operatorname{sgn}(\Delta X_{1i})\,\tau P_1}{A\mathbb{V}ar[D]}, \end{aligned} \tag{36}$$

where P_1 is already determined as a function of primitives in Equation (35) and $\mathbb{E}_{\mathcal{T}}[\mathbb{E}_i[D]] = \frac{\int_{i\in\mathcal{T}} \mathbb{E}_i[D] dF(i)}{\int_{i\in\mathcal{T}} dF(i)}$ and $\mathbb{E}_{\mathcal{T}}[A\mathbb{C}ov[M_{2i}, D]] = \frac{\int_{i\in\mathcal{T}} A\mathbb{C}ov[M_{2i},D] dF(i)}{\int_{i\in\mathcal{T}} dF(i)}$. I refer to ΔX_{1i}^+ and ΔX_{1i}^- as latent net buying/selling positions, since ΔX_{1i}^+ only describes actual trades when positive and ΔX_{1i}^- when negative. In particular, following the formulation in Equation (5) in the text, we can write the distribution of equilibrium latent net trades in the population as follows

$$\Delta X_{1i}^+ = \frac{\varepsilon_{di} - \varepsilon_{hi} - \tau P_1}{A\mathbb{V}ar[D]} \sim N\left(\frac{-\tau P_1}{A\mathbb{V}ar[D]}, \frac{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}{(A\mathbb{V}ar[D])^2}\right) \tag{37}$$

$$\Delta X_{1i}^- = \frac{\varepsilon_{di} - \varepsilon_{hi} + \tau P_1}{A\mathbb{V}ar[D]} \sim N\left(\frac{\tau P_1}{A\mathbb{V}ar[D]}, \frac{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}{(A\mathbb{V}ar[D])^2}\right), \tag{38}$$

where P_1 is already determined in Equation (35) and where we use the fact that $\mathbb{C}ov\left[\frac{\varepsilon_{di}}{A\mathbb{V}ar[D]}, \frac{-\varepsilon_{hi}}{A\mathbb{V}ar[D]}\right] = \frac{-\sigma_{dh}}{(A\mathbb{V}ar[D])^2}$.

Note that $\mathbb{V}ar[\varepsilon_{di} - \varepsilon_{hi}] = \sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h$ can take different values depending on the correlation between both motives for trading. Specifically

$$\mathbb{V}ar[\varepsilon_{di} - \varepsilon_{hi}] = \begin{cases} (\sigma_d + \sigma_h)^2, & \text{if } \rho = -1 \\ \sigma_d^2 + \sigma_h^2, & \text{if } \rho = 0 \\ (\sigma_d - \sigma_h)^2, & \text{if } \rho = 1. \end{cases}$$

Therefore, when $\rho \to -1$, the dispersion of equilibrium allocations is maximal. Note also that, when $\rho \to 1$, if $\sigma_d = \sigma_h$, there is no trade in equilibrium. In terms of (latent) individual turnover, we can express $\frac{\Delta X_{1i}^+}{Q}$ and $\frac{\Delta X_{1i}^-}{Q}$ as follows

$$\frac{\Delta X_{1i}^+}{Q} \sim N\left(\frac{\tau}{\Pi}, 2\pi\left(\Xi(0)\right)^2\right) \quad \text{and} \quad \frac{\Delta X_{1i}^-}{Q} \sim N\left(\frac{-\tau}{\Pi}, 2\pi\left(\Xi(0)\right)^2\right), \tag{39}$$

where $\Pi = \frac{A\mathbb{V}ar[D]Q}{P_1} = \frac{\mu_d}{P_1} - 1$ denotes the risk premium and

$$\Xi(0) = \frac{1}{\sqrt{2\pi}} \frac{1}{\Pi} \frac{\sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}}{P_1}$$

denotes laissez-faire aggregate turnover, derived in Equation (52) below.

Tax revenue Note that that total revenue is given by

$$\mathcal{R} = \tau P_1 \int |\Delta X_{1i}| \, dF(i) = 2\tau P_1 \mathcal{V}(\tau),$$

which can be expressed, when normalizing by the total market capitalization of the risky asset, as follows:

$$\frac{\mathcal{R}}{P_1 Q} = 2\tau \frac{\mathcal{V}(\tau)}{Q} = 2\tau \Xi(\tau).$$

Trading volume and shares of buyers and sellers Because there is a continuum of investors, a Law of Large Numbers guarantees that the level of trading volume in equilibrium in this model, given by $\mathcal{V}(\tau)$, is deterministic. Exploiting the properties of truncated normal distributions, stated in Section C.3, we can express the two elements that will determine $\mathcal{V}(\tau)$ as follows:

$$\begin{aligned}
\mathbb{P}[\Delta X_{1i} > 0] &= 1 - \Phi(\alpha_+) \\
\mathbb{E}[\Delta X_{1i} | \Delta X_{1i} > 0] &= \frac{1}{A \mathbb{V}ar[D]} \left(-\tau P_1 + \sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h} \lambda(\alpha_+) \right),
\end{aligned}$$

where α_+ is defined in Equation (40). Throughout the Appendix, $\phi(\cdot)$ and $\Phi(\cdot)$ respectively denote the p.d.f. and c.d.f. of the standard normal distribution. Note that the share of buyers (and sellers, by symmetry) corresponds to $\mathbb{P}[\Delta X_{1i} > 0] = \int_{i \in \mathcal{B}(\tau)} dF(i)$.

Trading volume expressed in dollars, which follows from the definition of $\mathcal{V}(\tau)$ in Equation (7), can be expressed as a function of the tax rate and primitives as follows

$$\begin{aligned}
P_1 \mathcal{V}(\tau) &= P_1 \int_{i \in \mathcal{B}} \Delta X_{1i} dF(i) = P_1 \mathbb{P}[\Delta X_{1i} > 0] \, \mathbb{E}[\Delta X_{1i} | \Delta X_{1i} > 0] \\
&= \frac{P_1}{A \mathbb{V}ar[D]} (1 - \Phi(\alpha_+)) \left(-\tau P_1 + \sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h} \lambda(\alpha_+) \right) \\
&= \frac{P_1}{A \mathbb{V}ar[D]} \left(-\tau P_1 (1 - \Phi(\alpha_+)) + \sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h} \phi(\alpha_+) \right),
\end{aligned}$$

where $\lambda(\alpha_+) = \frac{\phi(\alpha_+)}{1 - \Phi(\alpha_+)}$ corresponds to the inverse Mills ratio, whose properties are described in Section C.3, and α_+ and α are defined as follows

$$\alpha_+ = \max\{\alpha, 0\} \quad \text{where} \quad \alpha = \frac{\tau P_1}{\sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}}. \tag{40}$$

We can use the fact that $\phi'(\alpha_+) = -\alpha_+ \phi(\alpha_+)$ and the definition of α_+ to compute the response of trading volume to a tax change, given by

$$\begin{aligned}
\frac{d\mathcal{V}}{d\tau} &= \frac{1}{A \mathbb{V}ar[D]} \left(-P_1 (1 - \Phi(\alpha_+)) + \tau P_1 \phi(\alpha_+) \frac{d\alpha_+}{d\tau} + \sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h} \phi'(\alpha_+) \frac{d\alpha_+}{d\tau} \right) \\
&= \frac{1}{A \mathbb{V}ar[D]} \left(-P_1 (1 - \Phi(\alpha_+)) + \left(\tau P_1 - \alpha_+ \sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h} \right) \phi(\alpha_+) \frac{d\alpha_+}{d\tau} \right) \\
&= -\frac{P_1}{A \mathbb{V}ar[D]} (1 - \Phi(\alpha_+)),
\end{aligned}$$

which takes strictly negative values. Note that we can express $\varepsilon_\tau^{\log \mathcal{V}} = \frac{d \log \mathcal{V}}{d\tau} = \frac{d\mathcal{V}}{d\tau} \frac{1}{\mathcal{V}}$ as follows

$$\begin{aligned}
\frac{d\mathcal{V}}{d\tau} \frac{1}{\mathcal{V}} &= \frac{-\frac{P_1}{A \mathbb{V}ar[D]} (1 - \Phi(\alpha_+))}{\frac{1}{A \mathbb{V}ar[D]} (1 - \Phi(\alpha_+)) \left(-\tau P_1 + \sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h} \lambda(\alpha_+) \right)} \\
&= \frac{-1}{-\tau + \frac{\sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}}{P_1} \lambda(\alpha_+)},
\end{aligned}$$

which implies that

$$\varepsilon_\tau^{\log \mathcal{V}} \Big|_{\tau=0} = \frac{d \log \mathcal{V}}{d\tau} \Big|_{\tau=0} = \frac{-P_1}{\sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h} \lambda(0)},$$

where $\lambda(0) = \sqrt{\frac{2}{\pi}}$, as shown in Section C.3.

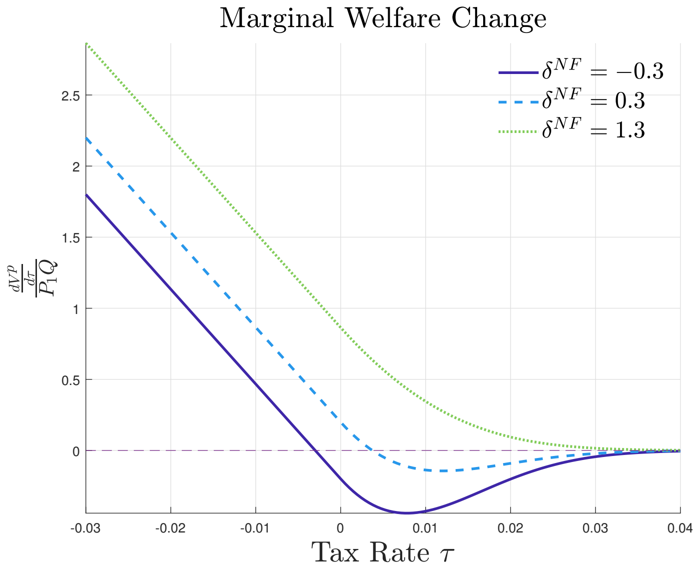


Figure A.1: Marginal welfare change

Note: Figure A.1 shows the normalized aggregate marginal welfare impact of a tax change from the planner's perspective, $\frac{\frac{dV^P}{d\tau}}{P_1 Q}$, defined in Equation (42), for the following values of the share of non-fundamental trading volume δ^{NF}: $\delta^{NF} = \{-0.3, 0.3, 1.3\}$. The optimal tax in each case is respectively given by $\tau^* = -0.30\%$, $\tau^* = 0.37\%$, and $\tau^* = \infty$.

Marginal welfare impact We can express the aggregate marginal welfare change $\frac{dV^p}{d\tau}$ as follows

$$\begin{aligned}\frac{dV^p}{d\tau} &= \int_{i\in\mathcal{T}(\tau)} \left[-\mathbb{E}_i[D] + \operatorname{sgn}(\Delta X_{1i}) P_1\tau\right] \frac{dX_{1i}}{d\tau} dF(i)\\
&= \frac{P_1}{A\mathbb{V}ar[D]}\left(\int_{i\in\mathcal{B}(\tau)} \varepsilon_{di} dF(i) - \int_{i\in\mathcal{S}(\tau)} \varepsilon_{di} dF(i) - \tau P_1 2\int_{i\in\mathcal{B}(\tau)} dF(i)\right)\\
&= \frac{P_1}{A\mathbb{V}ar[D]}\left(2\int_{i\in\mathcal{B}(\tau)} dF(i)\right)\left(\frac{\mathbb{E}_{\mathcal{B}(\tau)}[\mathbb{E}_i[D]] - \mathbb{E}_{\mathcal{S}(\tau)}[\mathbb{E}_i[D]]}{2} - \tau P_1\right)\\
&= \frac{P_1}{A\mathbb{V}ar[D]}\left(2\int_{i\in\mathcal{B}(\tau)} dF(i)\right)\left(\frac{\sigma_d(\sigma_d - \rho\sigma_h)}{\sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}}\lambda(\alpha_+) - \tau P_1\right)\\
&= \frac{P_1}{A\mathbb{V}ar[D]}\sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}\, 2\left(1 - \Phi(\alpha_+)\right)\left(\delta^{NF}\lambda(\alpha_+) - \alpha\right),\end{aligned}$$

where the definitions of α_+ and α are given in Equation (40) and where we can define δ^{NF}, shown below to correspond to the share of non-fundamental trading volume, as follows

$$\delta^{NF} = \frac{\sigma_d(\sigma_d - \rho\sigma_h)}{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h} = \frac{\frac{\sigma_d}{\sigma_h}\left(\frac{\sigma_d}{\sigma_h} - \rho\right)}{1 + \left(\frac{\sigma_d}{\sigma_h}\right)^2 - 2\rho\frac{\sigma_d}{\sigma_h}}. \tag{41}$$

Note that $\mathbb{E}_{\mathcal{B}(\tau)}[\mathbb{E}_i[D]] - \mathbb{E}_{\mathcal{S}(\tau)}[\mathbb{E}_i[D]]$ is explicitly computed in Equation (45) below. Since $\int_{i\in\mathcal{B}(\tau)} dF(i)$ is strictly positive for any value of τ, and α is a positive linear function of τ, it is sufficient to study the properties of $\delta^{NF}\lambda(\alpha_+) - \alpha$ to determine whether there is a uniquely optimal tax.

When normalized by the total value of the risky asset, we can express the marginal welfare impact of a tax change as follows

$$\begin{aligned}\frac{\frac{dV^p}{d\tau}}{P_1 Q} &= \underbrace{\frac{P_1}{A\mathbb{V}ar[D]Q}}_{1/\Pi}\underbrace{\frac{\sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}}{P_1}}_{\sqrt{2\pi}\Xi(0)\Pi} 2\left(1 - \Phi(\alpha_+)\right)\left(\delta^{NF}\lambda(\alpha_+) - \alpha\right)\\
&= \sqrt{2\pi}\,\Xi(0)\, 2\left(1 - \Phi(\alpha_+)\right)\left(\delta^{NF}\lambda(\alpha_+) - \alpha\right),\end{aligned} \tag{42}$$

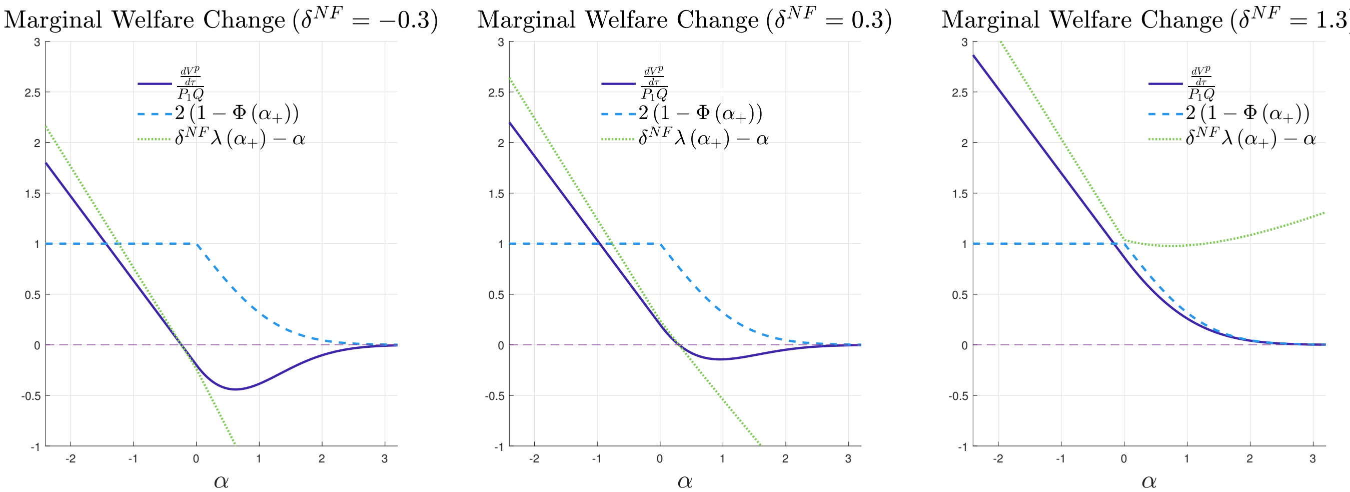


Figure A.2: Marginal welfare change (decomposition)

Note: Figure A.2 shows the normalized aggregate marginal welfare impact of a price change from the planner's perspective, $\frac{\frac{dV^P}{d\tau}}{P_1 Q}$, and its components, as defined in Equation (42), for the following values of the share of non-fundamental trading volume δ^{NF}: $\delta^{NF} = \{-0.3, 0.3, 1.3\}$. The optimal tax in each case is respectively given by $\tau^* = -0.30\%$, $\tau^* = 0.37\%$, and $\tau^* = \infty$.

where $\Xi(0)$ denotes laissez-faire turnover, characterized in Equation (52) below. Note that the normalized aggregate welfare impact at zero can be expressed exclusively as a function of Π, $\varepsilon_\tau^{\log \mathcal{V}}\big|_{\tau=0}$, and δ^{NF}, since

$$\left.\frac{\frac{dV^p}{d\tau}}{P_1 Q}\right|_{\tau=0} = \Xi(0)\sqrt{2\pi}\delta^{NF}\lambda(0) = \Xi(0)\,2\delta^{NF} = \frac{1}{\Pi}\frac{1}{\left|\varepsilon_\tau^{\log \mathcal{V}}\big|_{\tau=0}\right|}\delta^{NF},$$

where we use the fact that $\varepsilon_\tau^{\log \mathcal{V}}\big|_{\tau=0} = -\frac{1}{2}\frac{1}{\Pi\Xi(0)}$.

There are three possible scenarios to consider regarding the sign of the optimal tax, as illustrated in Figure A.1. First, when $\delta^{NF} < 0$, there is a uniquely optimal negative tax (a subsidy), given by $\alpha^* = \delta^{NF}\lambda(0)$. Second, when $0 \leq \delta^{NF} < 1$, there is a uniquely optimal finite and positive tax, which solves $\delta^{NF}\lambda(\alpha^*) = \alpha^*$. It is easy to verify that the function $\delta^{NF}\lambda(\alpha_+) - \alpha$ has a single root in that case. Third, when $\delta^{NF} \geq 1$, it is the case that $\frac{dV^p}{d\tau} > 0$, $\forall\tau$, so the optimal tax is $\tau^* = +\infty$. This follows from the fact that $\lambda(\alpha) > \alpha$, $\forall\alpha$, established in Section C.3. Formally,

$$\text{if} \quad \begin{cases} \delta^{NF} < 0 & \Rightarrow \tau^* < 0 \\ 0 \leq \delta^{NF} < 1 & \Rightarrow 0 \leq \tau^* < \infty \\ \delta^{NF} \geq 1 & \Rightarrow \tau^* = \infty. \end{cases}$$

The second order condition of the planner's problem can be written, for $\tau \neq 0$, as follows

$$\begin{aligned} \frac{d^2 V^p}{d\tau^2} &= \frac{P_1}{A\mathbb{V}ar[D]}\sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}2\left(\left(\delta^{NF}\phi'(\alpha_+) + \alpha\phi(\alpha_+)\right)\frac{d\alpha_+}{d\tau} - (1-\Phi(\alpha_+))\frac{d\alpha}{d\tau}\right) \\ &= \frac{P_1}{A\mathbb{V}ar[D]}\sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}2\left(\left(\left(\delta^{NF} - 1\right)\phi'(\alpha_+)\right)\frac{d\alpha_+}{d\tau} - (1-\Phi(\alpha_+))\frac{d\alpha}{d\tau}\right). \end{aligned}$$

The first term in the parentheses corresponds to the extensive margin effects described after introducing Equation (33). This term is zero whenever $\alpha < 0$ (equivalently, $\tau < 0$), so the planner's problem is always concave in that region. However, in general, the sign of this term is ambiguous. When $0 \leq \delta^{NF} < 1$, the first term is positive, since $\phi'(\alpha_+) \leq 0$ and $\frac{d\alpha_+}{d\tau} \geq 0$. When $\delta^{NF} > 1$, the first term is negative. The second term is always negative, as already shown after introducing Equation (33). Therefore, in the relevant case in which $0 \leq \delta^{NF} < 1$, the planner's problem is not necessarily concave. However, the characterization of $\frac{dV^p}{d\tau}$ immediately implies that the planner's problem is quasi-concave whenever $\delta^{NF} < 1$ for any set of parameters, so whenever the optimal tax is finite, it is unique.

It is also valuable to understand how different combinations of $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}$ (or equivalently $\frac{\sigma_d}{\sigma_h}$) and ρ map to δ^{NF}. Figure A.3 shows a contour plot with the combinations of $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}$ and ρ that generate the same value of δ^{NF}. The left black dashed

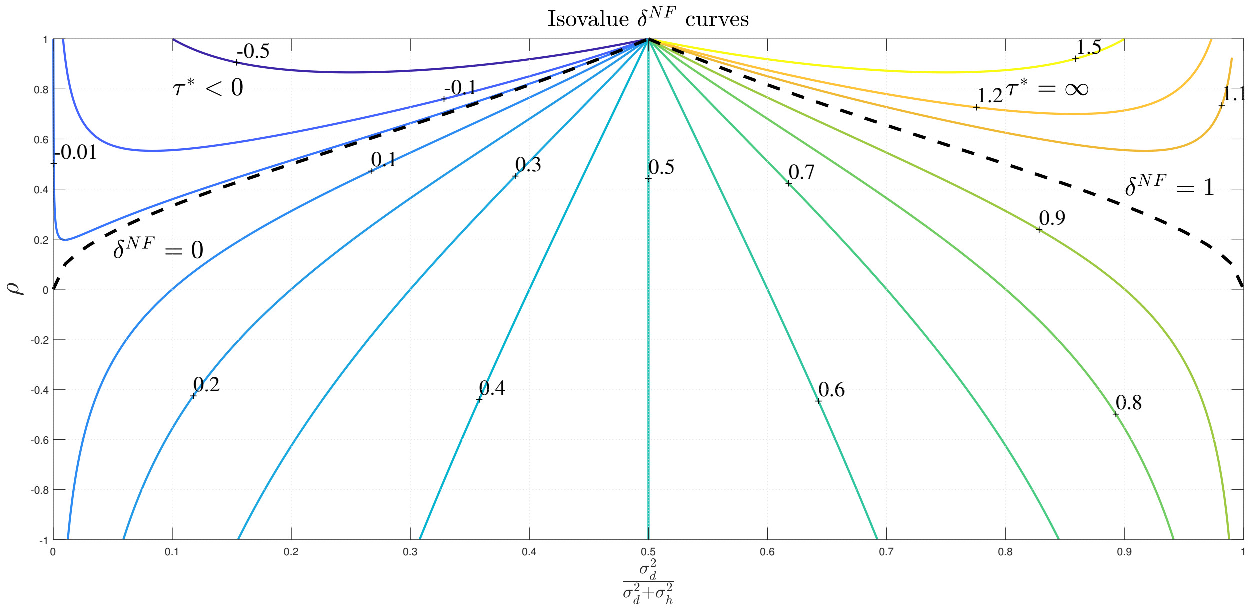


Figure A.3: Isovalue δ^{NF} curves for combinations of $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}$ and ρ.

Note: Figure A.3 shows the different combination of $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2} \in [0,1]$ and $\rho \in [-1,1]$ that are associated with the same value of the share of non-fundamental trading volume, δ^{NF}, defined in Equation (19), and consequently with the same optimal tax τ^*. The left black dashed line delimits the area in which δ^{NF} takes negative values (associated with a negative optimal tax). The right black dashed line delimits the area in which δ^{NF} takes values higher than one (associated with an infinite optimal tax). Figure A.4 below shows how δ^{NF} varies as a function of $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}$ for different values of δ.

line delimits the area in which δ^{NF} takes negative values (associated with a negative optimal tax). The right black dashed line delimits the area in which δ^{NF} takes values higher than one (associated with an infinite optimal tax). Figure A.4 below shows how δ^{NF} varies as a function of $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}$ for different values of δ.

For extreme choices of ρ, the share of non-fundamental trading volume δ^{NF} takes the following values

$$\delta^{NF} = \begin{cases} \frac{\sigma_d^2+\sigma_d\sigma_h}{\sigma_d^2+\sigma_h^2+2\sigma_d\sigma_h}, & \text{if } \rho = -1 \\ \frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}, & \text{if } \rho = 0 \\ \frac{\sigma_d^2-\sigma_d\sigma_h}{\sigma_d^2+\sigma_h^2-2\sigma_d\sigma_h}, & \text{if } \rho = 1. \end{cases}$$

Note that, when $\rho \leq 0$, the optimal tax is finite and non-negative for any value of $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}$.

Note that from the perspective of a planner who respects investors' beliefs, one can express the aggregate marginal impact of a tax change as follows

$$\frac{\frac{dV^{p,i}}{d\tau}}{P_1 Q} = -\frac{\tau}{\Pi}\int_{i\in\mathcal{T}(\tau)} dF(i) = -\frac{\tau}{\Pi}2\left(1-\Phi\left(\alpha_+\right)\right).$$

As discussed in the text, social welfare in this case always decreases with the level of the tax.

Finally, note that it is also possible to characterize the individual marginal welfare impact of a tax change for specific investors from the perspective of the planner. Formally, we can write $\frac{dV_i^p}{d\tau}$ as follows

$$\frac{dV_i^p}{d\tau} = \left[\mathbb{E}_p[D] - \mathbb{E}_i[D] + \operatorname{sgn}\left(\Delta X_{1i}\right)P_1\tau\right]\frac{dX_{1i}}{d\tau} - \Delta X_{1i}\frac{dP_1}{d\tau} + \frac{d\tilde{T}_{1i}}{d\tau}.$$

After normalizing for the total market capitalization of the risky asset, and since $\frac{dP_1}{d\tau} = 0$, we can further express $\frac{\frac{dV_i^p}{d\tau}}{P_1 Q}$ as follows

$$\frac{\frac{dV_i^p}{d\tau}}{P_1 Q} = \left[\frac{\mathbb{E}_p[D] - \mu_d + \overbrace{\mu_d - \mathbb{E}_i[D]}^{-\varepsilon_{di}}}{P_1} + \operatorname{sgn}\left(\Delta X_{1i}\right)\tau\right]\frac{\frac{dX_{1i}}{d\tau}}{Q} + \frac{\frac{d\tilde{T}_{1i}}{d\tau}}{P_1 Q}, \tag{43}$$

where $\frac{\frac{dX_{1i}}{d\tau}}{Q} = \frac{-\operatorname{sgn}(\Delta X_{1i})}{\Pi}$ and $\frac{\frac{d\tilde{T}_{1i}}{d\tau}}{P_1 Q}$ can take the value of zero if the planner fully rebates the tax liability to each individual investor or can be something different, as described in Section E.4, which characterizes the uniform rebate rule case. After the proof of Proposition 4, I fully describe the necessary informational requirements that characterize the distribution of $\frac{\frac{dV_i^p}{d\tau}}{P_1 Q}$. The counterpart of Equation (43) when welfare is computed from the perspective of an investor i is given by

$$\frac{dV_i^i}{d\tau} = \operatorname{sgn}(\Delta X_{1i})\,\tau \frac{\frac{dX_{1i}}{d\tau}}{Q} + \frac{\frac{d\tilde{T}_{1i}}{d\tau}}{P_1 Q}. \tag{44}$$

Optimal tax Because Assumption [S] is satisfied, the optimal tax formula can be written as in Equation (15), that is,

$$\tau^* = \frac{\mathbb{E}_{\mathcal{B}(\tau^*)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right] - \mathbb{E}_{\mathcal{S}(\tau^*)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]}{2}.$$

The average belief of buyers and sellers corresponds to

$$\begin{aligned}
\mathbb{E}_{\mathcal{B}}\left[\mathbb{E}_i\left[D\right]\right] &= \mu_d + \mathbb{E}\left[\varepsilon_{di}\middle|\,\Delta X_{1i} > 0\right] = \mu_d + \mathbb{E}\left[\varepsilon_{di}\middle|\,\varepsilon_{di} - \varepsilon_{hi} - \max\left\{\tau P_1, 0\right\} > 0\right] \\
\mathbb{E}_{\mathcal{S}}\left[\mathbb{E}_i\left[D\right]\right] &= \mu_d + \mathbb{E}\left[\varepsilon_{di}\middle|\,\Delta X_{1i} < 0\right] = \mu_d + \mathbb{E}\left[\varepsilon_{di}\middle|\,\varepsilon_{di} - \varepsilon_{hi} + \max\left\{\tau P_1, 0\right\} < 0\right].
\end{aligned}$$

Note that to compute those average beliefs, it is useful to write the joint distribution of the following relevant random variables as follows

$$\begin{pmatrix} \varepsilon_{di} \\ \varepsilon_{di} - \varepsilon_{hi} - \max\left\{\tau P_1, 0\right\} \\ \varepsilon_{di} - \varepsilon_{hi} + \max\left\{\tau P_1, 0\right\} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ -\max\left\{\tau P_1, 0\right\} \\ \max\left\{\tau P_1, 0\right\} \end{pmatrix}, \begin{pmatrix} \sigma_d^2 & \sigma_d^2 - \rho\sigma_d\sigma_h & \sigma_d^2 - \rho\sigma_d\sigma_h \\ & \sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h & \sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h \\ & & \sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h \end{pmatrix}\right),$$

using the fact that $\mathbb{C}ov\left[\varepsilon_{di}, \varepsilon_{di} - \varepsilon_{hi}\right] = \sigma_d^2 - \rho\sigma_d\sigma_h = \sigma_d\left(\sigma_d - \rho\sigma_h\right)$.

The correlation coefficient between ε_{di} and $\varepsilon_{di} - \varepsilon_{hi} - \max\left\{\tau P_1, 0\right\}$, or between ε_{di} and $\varepsilon_{di} - \varepsilon_{hi} + \max\left\{\tau P_1, 0\right\}$, is denoted by $\rho^{\mathcal{BS}}$ and given by

$$\rho^{\mathcal{BS}} = \frac{\sigma_d^2 - \rho\sigma_d\sigma_h}{\sigma_d\sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}} = \frac{\sigma_d - \rho\sigma_h}{\sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}}.$$

Exploiting the properties of truncated normal distributions,

$$\begin{aligned}
\mathbb{E}\left[\varepsilon_{di}\middle|\,\varepsilon_{di} - \varepsilon_{hi} - \max\left\{\tau P_1, 0\right\} > 0\right] &= \mathbb{E}\left[\varepsilon_{di}\right] + \rho^{\mathcal{BS}}\sigma_d\lambda\left(\alpha_+\right) = \rho^{\mathcal{BS}}\sigma_d\frac{\phi\left(\alpha_+\right)}{1 - \Phi\left(\alpha_+\right)} > 0 \\
\mathbb{E}\left[\varepsilon_{di}\middle|\,\varepsilon_{di} - \varepsilon_{hi} + \max\left\{\tau P_1, 0\right\} < 0\right] &= \mathbb{E}\left[\varepsilon_{di}\right] + \rho^{\mathcal{BS}}\sigma_d\lambda_-\left(-\alpha_+\right) = -\rho^{\mathcal{BS}}\sigma_d\frac{\phi\left(-\alpha_+\right)}{\Phi\left(-\alpha_+\right)} < 0,
\end{aligned}$$

where $\lambda_-\left(\cdot\right)$ is defined in Section C.3 and α_+ is defined in Equation (40). When $\rho^{\mathcal{BS}} > 0$, equivalently, $\delta^{NF} > 0$, the case associated with a positive tax, $\mathbb{E}_{\mathcal{B}}\left[\mathbb{E}_i\left[D\right]\right]$ is increasing in τ while $\mathbb{E}_{\mathcal{S}}\left[\mathbb{E}_i\left[D\right]\right]$ is decreasing in τ, that is, the average belief of marginal buyers increases with the tax rate while the average belief of sellers decreases. The opposite occurs when $\rho < 0^{\mathcal{BS}}$, equivalently, $\delta^{NF} < 0$.

Combining all these results, we can express the numerator of the optimal tax formula in Equation (15) as

$$\mathbb{E}_{\mathcal{B}(\tau)}\left[\mathbb{E}_i\left[D\right]\right] - \mathbb{E}_{\mathcal{S}(\tau)}\left[\mathbb{E}_i\left[D\right]\right] = 2\rho^{\mathcal{BS}}\sigma_d\lambda\left(\alpha_+\right), \tag{45}$$

which follows from the fact that

$$\frac{\phi\left(\alpha_+\right)}{1 - \Phi\left(\alpha_+\right)} + \frac{\phi\left(-\alpha_+\right)}{\Phi\left(-\alpha_+\right)} = 2\frac{\phi\left(\alpha_+\right)}{1 - \Phi\left(\alpha_+\right)} = 2\lambda\left(\alpha_+\right).$$

We can therefore write τ^* as

$$\begin{aligned}
\tau^* &= \frac{\mathbb{E}_{\mathcal{B}(\tau^*)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right] - \mathbb{E}_{\mathcal{S}(\tau^*)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]}{2} \\
&= \frac{\sigma_d\left(\sigma_d - \rho\sigma_h\right)}{\sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}}\lambda\left(\max\left\{\frac{\tau^* P_1}{\sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}}, 0\right\}\right)\frac{1}{P_1}.
\end{aligned}$$

We can rearrange this expression to find that

$$\frac{\tau^* P_1}{\sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}} = \frac{\sigma_d\left(\sigma_d - \rho\sigma_h\right)}{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}\lambda\left(\max\left\{\frac{\tau^* P_1}{\sqrt{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}}, 0\right\}\right),$$

which allows us to define $\alpha^* \equiv \frac{\tau^* P_1}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}$, implying that the fixed point that characterizes the optimal tax can be expressed as

$$\alpha^* = \delta^{NF}\lambda\left(\max\left\{\alpha^*, 0\right\}\right), \tag{46}$$

where δ^{NF} is defined in Equation (41) above. Note that α^* is exclusively a function of $\frac{\sigma_d}{\sigma_h}$ and ρ through δ^{NF}. Once a solution for α^* is found, τ^* simply corresponds to $\tau^* = \alpha^* \frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}$.

Trading volume implementation Under the new parametric assumption, it is possible to find explicit expressions for the trading volume decomposition. Before doing so, it is useful to compute $\mathbb{E}_{\mathcal{B}}\left[A\mathbb{C}ov\left[M_{2i}, D\right]\right]$ and $\mathbb{E}_{\mathcal{S}}\left[A\mathbb{C}ov\left[M_{2i}, D\right]\right]$, given by

$$\begin{aligned}
\mathbb{E}_{\mathcal{B}}\left[A\mathbb{C}ov\left[M_{2i}, D\right]\right] &= \mathbb{E}_{\mathcal{B}}\left[\varepsilon_{hi}\right] = \mathbb{E}\left[\varepsilon_{hi}\middle|\,\Delta X_{1i} > 0\right] = \mathbb{E}\left[\varepsilon_{hi}\middle|\,\varepsilon_{di} - \varepsilon_{hi} - \max\left\{\tau P_1, 0\right\} > 0\right] \\
&= \rho_h^{\mathcal{BS}}\sigma_h\lambda\left(\alpha_+\right) = \frac{\sigma_h\left(\rho\sigma_d - \sigma_h\right)}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}\frac{\phi\left(\alpha_+\right)}{1-\Phi\left(\alpha_+\right)} \\
\mathbb{E}_{\mathcal{S}}\left[A\mathbb{C}ov\left[M_{2i}, D\right]\right] &= \mathbb{E}_{\mathcal{S}}\left[\varepsilon_{hi}\right] = \mathbb{E}\left[\varepsilon_{hi}\middle|\,\Delta X_{1i}^- < 0\right] = \mathbb{E}\left[\varepsilon_{hi}\middle|\,\varepsilon_{di} - \varepsilon_{hi} + \tau P_1 < 0\right] \\
&= \rho_h^{\mathcal{BS}}\sigma_h\lambda_-\left(-\alpha_+\right) = -\frac{\sigma_h\left(\rho\sigma_d - \sigma_h\right)}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}\frac{\phi\left(-\alpha_+\right)}{\Phi\left(-\alpha_+\right)},
\end{aligned}$$

where α_+ is defined in Equation (40) and

$$\rho_h^{\mathcal{BS}} = \frac{\rho\sigma_d - \sigma_h}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}$$

. Therefore, it follows that

$$\mathbb{E}_{\mathcal{B}}\left[A\mathbb{C}ov\left[M_{2i}, D\right]\right] - \mathbb{E}_{\mathcal{S}}\left[A\mathbb{C}ov\left[M_{2i}, D\right]\right] = 2\frac{\sigma_h\left(\rho\sigma_d - \sigma_h\right)}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}\frac{\phi\left(\alpha_+\right)}{1-\Phi\left(\alpha_+\right)}.$$

Now we can express total trading volume in dollars as follows

$$\begin{aligned}
P_1\mathcal{V}\left(\tau\right) &= \frac{1}{2}\int_{i\in\mathcal{T}}\left(\left(-\frac{\partial X_{1i}}{\partial\tau}\right)\left(\mathbb{E}_i\left[D\right] - A_i\mathbb{C}ov\left[M_{2i}, D\right] - P_1\operatorname{sgn}\left(\Delta X_{1i}\right)\tau - A_i\mathbb{V}ar\left[D\right]X_{0i}\right)\right)dF\left(i\right) \\
&= \Theta_F\left(\tau\right) + \Theta_{NF}\left(\tau\right) - \Theta_\tau\left(\tau\right) \\
&= \frac{P_1}{A\mathbb{V}ar\left[D\right]}\left(1-\Phi\left(\alpha_+\right)\right)\left[\frac{\sigma_d\left(\sigma_d - \rho\sigma_h\right)}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}\lambda\left(\alpha_+\right) + \frac{-\sigma_h\left(\rho\sigma_d - \sigma_h\right)}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}\lambda\left(\alpha_+\right) - \tau P_1\right] \\
&= \frac{P_1}{A\mathbb{V}ar\left[D\right]}\left(1-\Phi\left(\alpha_+\right)\right)\left(-\tau P_1 + \sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}\lambda\left(\alpha_+\right)\right).
\end{aligned} \tag{47}$$

The fundamental component of trading volume can be expressed as follows

$$\begin{aligned}
\Theta_F\left(\tau\right) &= \frac{1}{2}\int_{i\in\mathcal{T}}\left(\frac{\operatorname{sgn}\left(\Delta X_{1i}\right)P_1}{A\mathbb{V}ar\left[D\right]}\right)\left(-A\mathbb{C}ov\left[M_{2i}, D\right]\right)dF\left(i\right) \\
&= -\frac{1}{2}\frac{P_1}{A\mathbb{V}ar\left[D\right]}\left(\int_{i\in\mathcal{B}}A\mathbb{C}ov\left[M_{2i}, D\right]dF\left(i\right) - \int_{i\in\mathcal{S}}A\mathbb{C}ov\left[M_{2i}, D\right]dF\left(i\right)\right) \\
&= -\frac{1}{2}\frac{P_1}{A\mathbb{V}ar\left[D\right]}\int_{i\in\mathcal{B}}dF\left(i\right)\left(\mathbb{E}_{\mathcal{B}}\left[A\mathbb{C}ov\left[M_{2i}, D\right]\right] - \mathbb{E}_{\mathcal{S}}\left[A\mathbb{C}ov\left[M_{2i}, D\right]\right]\right) \\
&= -\frac{1}{2}\frac{P_1}{A\mathbb{V}ar\left[D\right]}\left(1-\Phi\left(\alpha_+\right)\right)\left(\mathbb{E}_{\mathcal{B}}\left[A\mathbb{C}ov\left[M_{2i}, D\right]\right] - \mathbb{E}_{\mathcal{S}}\left[A\mathbb{C}ov\left[M_{2i}, D\right]\right]\right) \\
&= -\frac{P_1}{A\mathbb{V}ar\left[D\right]}\left(1-\Phi\left(\alpha_+\right)\right)\frac{\sigma_h\left(\rho\sigma_d - \sigma_h\right)}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}\frac{\phi\left(\alpha_+\right)}{1-\Phi\left(\alpha_+\right)} \\
&= \frac{P_1}{A\mathbb{V}ar\left[D\right]}\frac{\sigma_h\left(\sigma_h - \rho\sigma_d\right)}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}\phi\left(\alpha_+\right).
\end{aligned} \tag{48}$$

The non-fundamental component of trading volume can be expressed as follows

$$\begin{aligned}
\Theta_{NF}(\tau) &= \frac{1}{2}\int_{i\in\mathcal{T}}\left(\frac{\operatorname{sgn}(\Delta X_{1i})P_1}{A\mathbb{V}ar[D]}\right)\mathbb{E}_i[D]\,dF(i) \\
&= \frac{1}{2}\frac{P_1}{A\mathbb{V}ar[D]}\int_{i\in\mathcal{B}}dF(i)\left(\mathbb{E}_{\mathcal{B}}[\mathbb{E}_i[D]]-\mathbb{E}_{\mathcal{S}}[\mathbb{E}_i[D]]\right) \\
&= \frac{1}{2}\frac{P_1}{A\mathbb{V}ar[D]}(1-\Phi(\alpha))\left(\mathbb{E}_{\mathcal{B}}[\mathbb{E}_i[D]]-\mathbb{E}_{\mathcal{S}}[\mathbb{E}_i[D]]\right) \\
&= \frac{P_1}{A\mathbb{V}ar[D]}\frac{\sigma_d(\sigma_d-\rho\sigma_h)}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}\phi(\alpha_+)\,. \qquad (49)
\end{aligned}$$

The tax component of trading volume can be expressed as follows

$$\begin{aligned}
\Theta_{\tau}(\tau) &= \frac{1}{2}\tau P_1\int_{i\in\mathcal{T}}\left(\frac{\operatorname{sgn}(\Delta X_{1i})P_1}{A\mathbb{V}ar[D]}\right)\operatorname{sgn}(\Delta X_{1i})\,dF(i) \\
&= \frac{1}{2}\tau P_1\frac{P_1}{A\mathbb{V}ar[D]}\int_{i\in\mathcal{T}}dF(i) \\
&= \tau P_1\frac{P_1}{A\mathbb{V}ar[D]}\int_{i\in\mathcal{B}}dF(i) \\
&= \tau P_1\frac{P_1}{A\mathbb{V}ar[D]}(1-\Phi(\alpha_+))\,. \qquad (50)
\end{aligned}$$

Note that $\Theta_F(\tau)$ and $\Theta_{NF}(\tau)$ can take negative values for extreme values of ρ, that is, if $\sigma_h-\rho\sigma_d<0$ or $\sigma_d-\rho\sigma_h<0$. However, its sum $\Theta_F(\tau)+\Theta_{NF}(\tau)$ is always non-negative. To derive $\Theta_{NF}(\tau)$, we use the expression for $\mathbb{E}_{\mathcal{B}}[\mathbb{E}_i[D]]-\mathbb{E}_{\mathcal{S}}[\mathbb{E}_i[D]]$ given in Equation (45).

Note that the ratio of the non-fundamental to fundamental components of trading volume can be expressed as follows:

$$\frac{\Theta_{NF}(\tau)}{\Theta_F(\tau)}=\frac{\sigma_d(\sigma_d-\rho\sigma_h)}{\sigma_h(\sigma_h-\rho\sigma_d)}=\frac{\sigma_d^2-\rho\sigma_d\sigma_h}{\sigma_h^2-\rho\sigma_d\sigma_h}=\frac{\frac{\sigma_d}{\sigma_h}-\rho}{\frac{\sigma_h}{\sigma_d}-\rho}.$$

Three facts are worth highlighting. First, this ratio is independent of the tax rate τ. Second, when $\rho=0$, the ratio is exactly $\left(\frac{\sigma_d}{\sigma_h}\right)^2$. Third, when $\sigma_d=\sigma_h$, the ratio is equal to 1. More importantly, it is possible to compute the share of the non-fundamental component of trading to the sum of fundamental and non-fundamental components as follows

$$\delta^{NF}=\frac{\Theta_{NF}(\tau)}{\Theta_F(\tau)+\Theta_{NF}(\tau)}=\frac{\sigma_d(\sigma_d-\rho\sigma_h)}{\sigma_h(\sigma_h-\rho\sigma_d)+\sigma_d(\sigma_d-\rho\sigma_h)}=\frac{\sigma_d^2-\rho\sigma_d\sigma_h}{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}=\frac{\frac{\sigma_d}{\sigma_h}-\rho}{\frac{\sigma_h}{\sigma_d}-\rho+\frac{\sigma_d}{\sigma_h}-\rho}.$$

Figure A.4 illustrates how δ^{NF} varies with the level of $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}$ for different values of the correlation coefficient ρ — see also Figure A.3 above.[39] Whenever $\delta^{NF}\in[0,1]$, an increase in $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}$ increases the value of δ^{NF}, as expected. For a given value of $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}$, increasing ρ increases the share of non-fundamental trading volume δ^{NF} when $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}<1/2$, but decreases δ^{NF} otherwise. Note that, when $\rho>0$, δ^{NF} can take negative values or values larger than unity, which makes it harder to interpret as a share.

Finally, note that in terms of turnover, we can express $\Xi(\tau)$, $\Xi_F(\tau)$, $\Xi_{NF}(\tau)$, and $\Xi_\tau(\tau)$ as follows

$$\begin{aligned}
\Xi(\tau) &= \frac{\mathcal{V}(\tau)}{Q}=(1-\Phi(\alpha_+))\left(-\frac{\tau}{\Pi}\right)+\Xi(0)\sqrt{2\pi}\phi(\alpha_+) \qquad (51)\\
\Xi_F(\tau) &= \frac{\Theta_F(\tau)}{P_1Q}=\Xi(0)\sqrt{2\pi}\delta^{NF}\phi(\alpha_+) \\
\Xi_{NF}(\tau) &= \frac{\Theta_{NF}(\tau)}{P_1Q}=\Xi(0)\sqrt{2\pi}\left(1-\delta^{NF}\right)\phi(\alpha_+) \\
\Xi_\tau(\tau) &= \frac{\Theta_\tau(\tau)}{P_1Q}=\frac{\tau}{\Pi}(1-\Phi(\alpha_+))\,,
\end{aligned}$$

where α_+ is defined in Equation (40), the risk premium is denoted by $\Pi=\frac{A\mathbb{V}ar[D]Q}{P_1}$, and laissez-faire turnover $\Xi(0)$ is given by

$$\Xi(0)=\frac{1}{\sqrt{2\pi}}\frac{1}{\Pi}\frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}. \qquad (52)$$

[39] Note that $\frac{\sigma_d}{\sigma_h}$ and $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}$ are related through the fact that $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}=\frac{1}{1+\left(\frac{\sigma_d}{\sigma_h}\right)^{-2}}$.

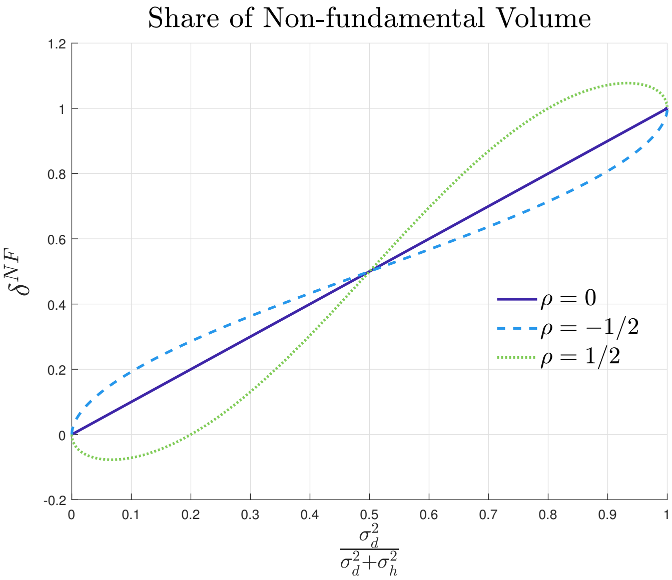


Figure A.4: Share of non-fundamental volume

Note: Figure A.4 shows how the share of non-fundamental trading volume δ^{NF}, defined in Equation (41), varies with the level of $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}$ for different values of the correlation coefficient ρ. Note that δ^{NF} is increasing in $\frac{\sigma_d^2}{\sigma_d^2+\sigma_h^2}$ (equivalently $\frac{\sigma_d}{\sigma_h}$) whenever $\delta^{NF} \in [0, 1]$.

Non-quasi-concave planner's objective Figure 4 shows the normalized aggregate welfare impact of a tax change from the planner's perspective in a scenario in which the planner's objective fails to be quasi-concave. The easiest scenario in which symmetry is preserved is one in which there are n groups of investors, with proportions π^n, some risk aversion and initial asset holdings, and whose fundamental and non-fundamental trading motives are centered around the same mean as follows

$$\mathbb{E}_i^n [D] \sim \mu_d + \varepsilon_{di}^n$$
$$A\mathbb{C}ov [M_{2i}, D] \sim \mu_h + \varepsilon_{hi}^n,$$

where $\mu_d \geq 0$ and $\mu_h = 0$ are the same across all n groups. The random variables ε_{hi}^n and ε_{di}^n are jointly normally distributed for i investors in group n as follows

$$\begin{pmatrix} \varepsilon_{di}^n \\ \varepsilon_{hi}^n \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} (\sigma_d^n)^2 & \rho_n \sigma_d^n \sigma_h^n \\ \rho_n \sigma_d^n \sigma_h^n & (\sigma_h^n)^2 \end{pmatrix} \right).$$

In this case, the equilibrium price still corresponds to Equation (6), and most of the equations derived in this section remain valid for each, including individual and aggregate turnover. In the environment considered in Figure 4, 90% of investors belong to group 1, while the remaining 10% belong to group 2. Group 1 investors have a turnover of 1/4 and a share of non-fundamental trading volume of 0.3. Group 2 investors have a turnover of 1 and a share of non-fundamental trading volume of 0.65. The risk premium is 1.5%.

The aggregate marginal impact of a tax change in this case is given by

$$\frac{\frac{dV^p}{d\tau}}{P_1 Q} = \sum_n \pi^n \frac{\frac{dV_n^p}{d\tau}}{P_1 Q}, \quad \text{where} \quad \frac{\frac{dV_n^p}{d\tau}}{P_1 Q} = \sqrt{2\pi} \Xi_n (0) \, 2 \left(1 - \Phi \left(\alpha_+^n\right)\right) \left(\delta_n^{NF} \lambda \left(\alpha_+^n\right) - \alpha^n\right),$$

where α_+^n, α^n, δ_n^{NF}, and $\Xi_n(0)$ are computed as in Equations (40), (41), and (52) for each group. Even though each of the group-specific $\frac{\frac{dV_n^p}{d\tau}}{P_1 Q}$ is quasi-concave, it is a fact that the sum of two quasi-concave functions is not quasi-concave, so the planner's objective $\frac{\frac{dV^p}{d\tau}}{P_1 Q}$ may or may not be quasi-concave.

Proposition 3. (Optimal tax and comparative statics)

Proof. a) It easily follows that, when $\rho \leq 0$ and $\sigma_d > 0$, Equation (46) has a strictly positive solution, since $\delta^{NF} > 0$.

b) From Equation (46), for any $\tau^* \geq 0$, the sign of $\frac{d\alpha^*}{d\left(\frac{\sigma_d}{\sigma_h}\right)}$ is determined by the sign of $-\frac{d\delta^{NF}}{d\left(\frac{\sigma_h}{\sigma_d}\right)}$. We can express $\frac{d\delta^{NF}}{d\left(\frac{\sigma_h}{\sigma_d}\right)}$ as follows

$$\frac{d\delta^{NF}}{d\left(\frac{\sigma_h}{\sigma_d}\right)} = \frac{-\rho\left(1+\left(\frac{\sigma_h}{\sigma_d}\right)^2 - 2\rho\frac{\sigma_h}{\sigma_d}\right) - \left(1-\rho\frac{\sigma_h}{\sigma_d}\right)2\left(\frac{\sigma_h}{\sigma_d}-\rho\right)}{\left(1+\left(\frac{\sigma_h}{\sigma_d}\right)^2 - 2\rho\frac{\sigma_h}{\sigma_d}\right)^2} = \frac{\rho\left(\frac{\sigma_h}{\sigma_d}\right)^2 - 2\frac{\sigma_h}{\sigma_d}+\rho}{\left(1+\left(\frac{\sigma_h}{\sigma_d}\right)^2 - 2\rho\frac{\sigma_h}{\sigma_d}\right)^2}.$$

When $\rho \leq 0$ this expression is everywhere negative implying that $\frac{d\alpha^*}{d\left(\frac{\sigma_d}{\sigma_h}\right)}$ is positive. Note that we can find the following cases, for $\rho \geq 0$:

$$\text{if} \begin{cases} \frac{\sigma_d}{\sigma_h} < \rho, & \alpha^* < 0 \\ \rho \leq \frac{\sigma_d}{\sigma_h} < \frac{1}{\rho}, & \alpha^* \geq 0 \\ \frac{\sigma_d}{\sigma_h} \geq \frac{1}{\rho}, & \alpha^* = \infty. \end{cases}$$

It can be established that $\rho\left(\frac{\sigma_h}{\sigma_d}\right)^2 - 2\frac{\sigma_h}{\sigma_d}+\rho$ is negative in between its two roots when $\rho \in (0,1]$. Its roots are given by $\frac{1\pm\sqrt{1-\rho^2}}{\rho}$. It is sufficient to show that $\frac{1-\sqrt{1-\rho^2}}{\rho} < \rho$ and that $\frac{1}{\rho} < \frac{1+\sqrt{1-\rho^2}}{\rho}$, which are trivially satisfied for any $\rho \in (0,1]$. This fact is sufficient to show the desired comparative statics for the $\rho > 0$ case.

c) When $\delta^{NF} < 0$, there is a unique negative solution to Equation (46) (a subsidy), given by $\alpha^* = \delta^{NF}\lambda(0)$. When $0 \leq \delta^{NF} < 1$, there is a uniquely optimal finite and positive tax, which solves $\delta^{NF}\lambda(\alpha^*) = \alpha^*$. It is easy to verify that Equation (46) has a single solution in that case. Third, when $\delta^{NF} \geq 1$, it is the case that $\frac{dV^p}{d\tau} > 0, \forall\tau$, so the optimal tax is $\tau^* = +\infty$. This follows from the fact that $\lambda(\alpha) > \alpha, \forall\alpha$, established in Section C.3. Formally,

$$\text{if} \begin{cases} \delta^{NF} < 0, & \tau^* < 0 \\ 0 \leq \delta^{NF} < 1, & 0 \leq \tau^* < \infty \\ \delta^{NF} \geq 1, & \tau^* = \infty. \end{cases}$$

□

C.2 Quantitative assessment

Proposition 4. (Optimal tax identification/Sufficient statistics)

Proof. a) Equation (46) establishes that α^* exclusively depends on the value of δ^{NF}, itself a function of $\frac{\sigma_d}{\sigma_h}$ and ρ. Therefore, for a given α^*, we can use the definition of α, given in Equation (40), to express τ^* as follows

$$\tau^* = \alpha^* \frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}.$$

Therefore, for a given α^*, it is sufficient to know $\frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}$ to find the optimal tax. A possible way of computing that object exploits the definition of volume semi-elasticity to tax changes. First, note that

$$\frac{d\log\mathcal{V}}{d\tau} = \frac{1}{\tau - \frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}\lambda(\alpha_+)}. \tag{53}$$

Therefore, if $\frac{d\log\mathcal{V}}{d\tau}$ is observed for a given value of τ, Equation (53) can be inverted to recover $\frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}$. In general, finding $\frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}$ involves solving a complicated non-linear equation. However, if one observes $\varepsilon_\tau^{\log\mathcal{V}}\big|_{\tau=0}$, then it is possible to find $\frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}$ explicitly, since

$$\varepsilon_\tau^{\log\mathcal{V}}\Big|_{\tau=0} = \frac{d\log\mathcal{V}}{d\tau}\Big|_{\tau=0} = -\frac{1}{\frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}\sqrt{\frac{2}{\pi}}}.$$

In that case,

$$\tau^* = -\frac{\alpha^*}{\varepsilon_\tau^{\log\mathcal{V}}\big|_{\tau=0}}\sqrt{\frac{\pi}{2}}.$$

This argument is sufficient to prove part a). Note that Equation (52) implies the relation

$$\frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}=\Pi\Xi\left(0\right)\sqrt{2\pi},$$

which allows us to write the optimal tax as $\tau^*=\Pi\Xi\left[0\right]\sqrt{2\pi}\alpha^*$. Hence, in this model, $\left.\varepsilon_\tau^{\log\mathcal{V}}\right|_{\tau=0}=-\frac{1}{2}\frac{1}{\Pi\Xi(0)}$, so the frequency of trading can be chosen to jointly match $\left.\varepsilon_\tau^{\log\mathcal{V}}\right|_{\tau=0}$, Π, and $\Xi\left(0\right)$.

b) The normalized aggregate marginal welfare impact of a tax change, already characterized in Equation (42), corresponds to

$$\frac{\frac{dV^p}{d\tau}}{P_1Q}=\sqrt{2\pi}\Xi\left(0\right)2\left(1-\Phi\left(\alpha_+\right)\right)\left(\delta^{NF}\lambda\left(\alpha_+\right)-\alpha\right).$$

In addition to knowing δ^{NF} and $\frac{d\log\mathcal{V}}{d\tau}$ for some value of τ, to compute $\frac{\frac{dV^p}{d\tau}}{P_1Q}$ it is now necessary to also know either laissez-faire turnover $\Xi\left[0\right]$ or the risk premium Π, which concludes the proof. □

Identification of individual marginal welfare gains It is worth making two additional remarks. First, combining Equation (39) with the fact that $\left.\varepsilon_\tau^{\log\mathcal{V}}\right|_{\tau=0}=-\frac{1}{2}\frac{1}{\Pi\Xi(0)}$, it follows easily that only information on two of the following three objects: $\left.\varepsilon_\tau^{\log\mathcal{V}}\right|_{\tau=0}$, Π, and $\Xi\left(0\right)$, is necessary to determine the distribution of investors individual turnover. Second, it is possible to recover the distribution of $\frac{\frac{dV_i^p}{d\tau}}{P_1Q}$ for individual investors for a given value of $\frac{\mathbb{E}_p[D]-\mu_d}{P_1}$. Note that $\frac{\mu_d}{P_1}=1+\Pi$. However, in this case, in addition to $\left.\varepsilon_\tau^{\log\mathcal{V}}\right|_{\tau=0}$, and Π or $\Xi\left(0\right)$, it is necessary to separately know $\frac{\sigma_d}{\sigma_h}$ and ρ. Formally, starting from Equation (43), we can express $\frac{\frac{dV_i^p}{d\tau}}{P_1Q}$ as follows

$$\frac{\frac{dV_i^p}{d\tau}}{P_1Q}=\left[\frac{\mathbb{E}_p\left[D\right]-\mu_d-\varepsilon_{di}}{P_1}+\mathrm{sgn}\left(\Delta X_{1i}\right)\tau\right]\frac{-\mathrm{sgn}\left(\Delta X_{1i}\right)}{\Pi}+\frac{\frac{d\tilde{T}_{1i}}{d\tau}}{P_1Q}.$$

Hence in this case, one needs to characterize the distributions of $\frac{\varepsilon_{di}}{P_1}$ and $\frac{\varepsilon_{hi}}{P_1}$, given by

$$\frac{\varepsilon_{di}}{P_1}\sim N\left(0,\frac{\sigma_d}{P_1}\right)\quad\text{and}\quad\frac{\varepsilon_{hi}}{P_1}\sim N\left(0,\frac{\sigma_h}{P_1}\right),$$

where

$$\frac{\sigma_d}{P_1}=\frac{\Pi\Xi\left(0\right)\sqrt{2\pi}}{\sqrt{1+\left(\frac{\sigma_h}{\sigma_d}\right)^2-2\rho\frac{\sigma_h}{\sigma_d}}}=\frac{\frac{1}{\left.\varepsilon_\tau^{\log\mathcal{V}}\right|_{\tau=0}}\sqrt{\frac{\pi}{2}}}{\sqrt{1+\left(\frac{\sigma_h}{\sigma_d}\right)^2-2\rho\frac{\sigma_h}{\sigma_d}}}$$

$$\frac{\sigma_h}{P_1}=\frac{\Pi\Xi\left(0\right)\sqrt{2\pi}}{\sqrt{1+\left(\frac{\sigma_d}{\sigma_h}\right)^2-2\rho\frac{\sigma_d}{\sigma_h}}}=\frac{\frac{1}{\left.\varepsilon_\tau^{\log\mathcal{V}}\right|_{\tau=0}}\sqrt{\frac{\pi}{2}}}{\sqrt{1+\left(\frac{\sigma_d}{\sigma_h}\right)^2-2\rho\frac{\sigma_d}{\sigma_h}}}.$$

The net trading positions for active investors as a function of $\frac{\varepsilon_{di}}{P_1}$ and $\frac{\varepsilon_{hi}}{P_1}$ are given by

$$\frac{\Delta X_{1i}}{Q}=\frac{\varepsilon_{di}-\varepsilon_{hi}-\mathrm{sgn}\left(\Delta X_{1i}\right)\tau P_1}{A\mathbb{V}ar\left[D\right]Q}=\frac{1}{\Pi}\left(\frac{\varepsilon_{di}-\varepsilon_{hi}}{P_1}-\mathrm{sgn}\left(\Delta X_{1i}\right)\tau\right).$$

Therefore, for given values of $\frac{\varepsilon_{di}}{P_1}$ and $\frac{\varepsilon_{hi}}{P_1}$, as well as Π, it is possible to compute $\frac{\Delta X_{1i}}{Q}$ as well as $\frac{\frac{dV_i^p}{d\tau}}{P_1Q}$ if $\frac{\frac{d\tilde{T}_{1i}}{d\tau}}{P_1Q}=0$. See Section E.4 for how to compute $\frac{\frac{d\tilde{T}_{1i}}{d\tau}}{P_1Q}$ with the same information under a uniform rebate rule.

Estimation of share of non-fundamental trading volume Here, I provide and study the properties of an estimation procedure for δ^{NF} based on information on individual investors' portfolio choices and hedging needs. I assume that the estimation is conducted in an economy with $\tau=0$, although the approach could be extended to scenarios with positive taxation.

First, note that Equation (36) implies that the following relation must hold for all investors:

$$\underbrace{\Delta X_{1i}}_{\text{observed}}=\underbrace{-\frac{\varepsilon_{hi}}{A\mathbb{V}ar\left[D\right]}}_{\text{observed}}+\underbrace{\frac{\varepsilon_{di}}{A\mathbb{V}ar\left[D\right]}}_{\text{error term}},$$

which implies that the following variance decomposition must hold

$$TSS \equiv \mathbb{V}ar_F\left[\Delta X_{1i}\right] = \frac{\mathbb{V}ar\left[\varepsilon_{di} - \varepsilon_{hi}\right]}{\left(A\mathbb{V}ar\left[D\right]\right)^2} = \frac{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}{\left(A\mathbb{V}ar\left[D\right]\right)^2},$$

where TSS denotes the total sum of squares.

Next, under the assumption that an econometrician observes ΔX_{1i} and ε_{hi} (or equivalently, $\mathbb{C}ov\left[M_{2i}, D\right]$, which would only change the analysis by including a constant), but cannot observe investors' beliefs (ε_{di}, or equivalently, $\mathbb{E}_D\left[D\right]$), this expression can be interpreted as a regression equation in which $\frac{\varepsilon_{di}}{A\mathbb{V}ar[D]}$ corresponds to an error term. Therefore, it is possible to recover from observables the explained sum of squares, which corresponds to

$$ESS \equiv \frac{\sigma_h^2}{\left(A\mathbb{V}ar\left[D\right]\right)^2},$$

where ESS denotes the explained sum of squares.

I then propose the following estimator for δ^{NF}:

$$\hat{\delta}^{NF} = 1 - \frac{ESS}{TSS}. \tag{54}$$

Note that this estimator recovers the ratio

$$1 - \frac{ESS}{TSS} = \frac{\sigma_d^2 - 2\rho\sigma_d\sigma_h}{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}, \tag{55}$$

which is always positive. By comparing Equation (55) with Equation (19), it becomes evident that $\hat{\delta}^{NF}$ may be biased when $\rho \neq 0$.[40] Formally, the relative bias of this estimator can be expressed as follows

$$\frac{\hat{\delta}^{NF} - \delta^{NF}}{\delta^{NF}} = -\frac{\rho}{\frac{\sigma_d}{\sigma_h} - \rho}. \tag{56}$$

Therefore, when $\rho = 0$, the estimator $\hat{\delta}^{NF}$, as defined in Equation (54) is an unbiased estimator of the share of non-fundamental trading. The magnitude of the bias is increasing in the magnitude of ρ. Note that when $\rho < 0$, the maximum relative bias is bounded by $\frac{1}{\frac{\sigma_d}{\sigma_h}+1}$, When $\rho > 0$, the bias could be arbitrarily large when $\frac{\sigma_d}{\sigma_h} \approx \rho$, but those are the cases in which $\delta^{NF} \approx 0$, so the overall impact on the optimal tax may still be small.

Conceptually, when beliefs and hedging motives are uncorrelated, it is possible to use observed investors' portfolio allocations and hedging needs to find the relevant ratio that determines the share of non-fundamental trading. However, when beliefs and hedging motives are correlated, there will be confounding effects unless one can observe individual beliefs too. Equation (56) is useful to understand the bias that arises in those cases.

C.3 Auxiliary results

The following auxiliary results are useful. These are well-known properties of the normal distribution — see, for instance, Greene (2003). Let's respectively denote by $\phi(\cdot)$ and $\Phi(\cdot)$ the p.d.f. and c.d.f. of the standard normal distribution. Note that $\phi(0) = \frac{1}{\sqrt{2\pi}}$.

Fact 1. *If $X \sim N\left(\mu, \sigma^2\right)$, then*

$$\mathbb{E}\left[X \left| X > a\right.\right] = \mu + \sigma\lambda\left(\alpha_x\right), \quad \text{where} \quad \lambda\left(\alpha_x\right) = \frac{\phi\left(\alpha_x\right)}{1 - \Phi\left(\alpha_x\right)} \quad \text{and} \quad \alpha_x = \frac{a - \mu}{\sigma}.$$

Fact 2. *If Y and Z have a bivariate normal distribution with means μ_y and μ_z, variances σ_y^2 and σ_z^2, and a correlation coefficient ρ_{yz}, then*

$$\begin{aligned} \mathbb{E}\left[Y \left| Z > a\right.\right] &= \mu_y + \rho_{yz}\sigma_y\lambda\left(\alpha_z\right), \quad \text{where} \quad \lambda\left(\alpha_z\right) = \frac{\phi\left(\alpha_z\right)}{1 - \Phi\left(\alpha_z\right)} \quad \text{and} \quad \alpha_z = \frac{a - \mu_z}{\sigma_z} \\ \mathbb{E}\left[Y \left| Z < b\right.\right] &= \mu_y + \rho_{yz}\sigma_y\lambda_-\left(\beta_z\right), \quad \text{where} \quad \lambda_-\left(\beta_z\right) = -\frac{\phi\left(\beta_z\right)}{\Phi\left(\beta_z\right)} \quad \text{and} \quad \beta_z = \frac{b - \mu_z}{\sigma_z}, \end{aligned}$$

and more generally

$$\mathbb{E}\left[Y \left| a < Z < b\right.\right] = \mu_y + \rho_{yz}\sigma_y \frac{\phi\left(\alpha_z\right) - \phi\left(\beta_z\right)}{\Phi\left(\beta_z\right) - \Phi\left(\alpha_z\right)}.$$

[40] The rest of the argument assumes that ESS and TSS can be estimated without bias.

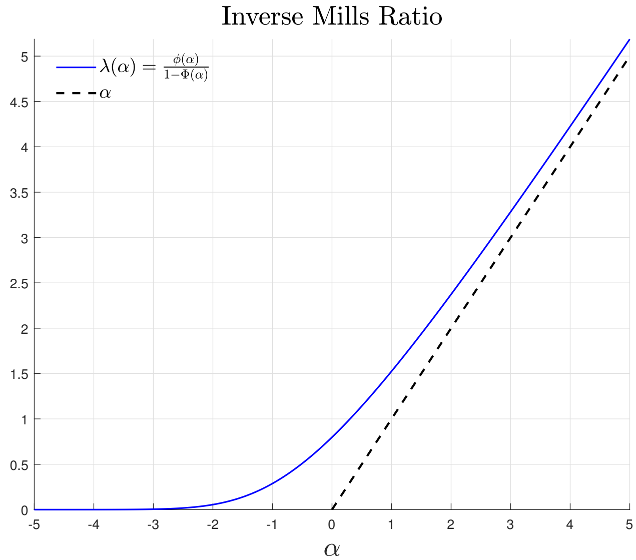


Figure A.5: Inverse Mills ratio

Note: Figure A.5 shows the inverse Mills ratio for the normal distribution: $\lambda(\alpha) = \frac{\phi(\alpha)}{1-\Phi(\alpha)}$.

Fact 3. *The function* $\lambda(\alpha) = \frac{\phi(\alpha)}{1-\Phi(\alpha)}$, *which corresponds to the hazard rate of the normal distribution, is also known as the inverse Mills ratio. It satisfies the following properties:*

1. $\lambda(0) = \sqrt{\frac{2}{\pi}}$, $\lambda(\alpha) \geq 0$, $\lambda(\alpha) > \alpha$, $\lambda'(\alpha) > 0$, and $\lambda''(\alpha) > 0$.
2. $\lim_{\alpha \to -\infty} \lambda(\alpha) = \lim_{\alpha \to -\infty} \lambda'(\alpha) = 0$, and $\lim_{\alpha \to \infty} \lambda'(\alpha) = 1$.
3. $\lambda(\alpha) < \frac{1}{\alpha} + \alpha$, $\lambda'(\alpha) = \frac{\phi'(\alpha_i)}{1-\Phi(\alpha_i)} + (\lambda(\alpha_i))^2 = \lambda(\alpha)(\lambda(\alpha) - \alpha) > 0$, $\lambda'(\alpha) < 1$, and $\lambda''(\alpha) \geq 0$.

C.4 Imperfect tax enforcement

Here, I relax the assumption of perfect tax enforcement. Initially, I derive the results starting from the baseline model presented in Section 2, but I often specialize the results to the environment used in the quantitative assessment in Section 5. These results underlie Figure 7 in the text.

Environment Investors can now trade in two different markets, A and B. Market A is meant to represent existing venues for trading and all trades in that market face a transaction tax τ. Market B seeks to represent trading venues outside of the regulatory scope of authorities. In market B, investors face a quadratic cost of trading, modulated by ν. When $\nu \to 0$, avoiding the tax is costless, and all trades migrate to market B when $\tau > 0$. When $\nu \to \infty$, no trade migrates to market B, nesting the baseline model. We will parametrize the model in terms of

$$\psi \equiv \frac{A\mathbb{V}ar[D]}{\nu} \geq 0.$$

In terms of ψ, when $\nu \to 0$, $\psi \to \infty$, and all trades migrate to market B when $\tau > 0$. When $\nu \to \infty$, $\psi \to 0$, and the model converges to the baseline model.

Formally, the consumption/wealth of a given investor i at date 2 is now given by

$$W_{2i} = M_{2i} + \left(X_{1i}^A + X_{1i}^B\right) D - \Delta X_{1i}^A P_1^A - \Delta X_{1i}^B P_1^B - \tau P_1^A \left|\Delta X_{1i}^A\right| - \frac{\nu}{2}\left(\Delta X_{1i}^B\right)^2 + T_{1i}.$$

The lump-sum transfer T_{1i} rebates tax revenues to investors, but not the trading costs associated with trading in market B. Note that the linear tax only affects trading in market A, while the quadratic cost only affects trading in market B. Below, I'll consider symmetric equilibria in which $P_1^A = P_1^B = P_1 > 0$, so the single market clearing condition is $\int \left(\Delta X_{1i}^A(P_1) + \Delta X_{1i}^B(P_1)\right) dF(i) = 0$. I further assume that $X_{0i}^A = Q$ and $X_{0i}^B = 0$, $\forall i$. This assumption allows us to interpret market B as a market that would not exist without taxation.

Equilibrium characterization Investors face the following well-behaved problem:

$$\max_{X_{1i}^A, X_{1i}^B} \left[\mathbb{E}_i\left[D\right] - A_i \mathbb{C}ov\left[M_{2i}, D\right]\right] X_{1i} - X_{1i}^A P_1^A - X_{1i}^B P_1^B - \tau P_1^A \left|\Delta X_{1i}^A\right| - \frac{\nu}{2}\left(\Delta X_{1i}^B\right)^2 - \frac{A_i}{2}\mathbb{V}ar\left[D\right]\left(X_{1i}\right)^2,$$

where $X_{1i} = X_{1i}^A + X_{1i}^B$. The optimality conditions of investors correspond to

$$X_{1i}^A + X_{1i}^B = \frac{\mathbb{E}_i\left[D\right] - A_i \mathbb{C}ov\left[M_{2i}, D\right] - P_1^A - \tau P_1^A \operatorname{sgn}\left(\Delta X_{1i}^A\right)}{A_i \mathbb{V}ar\left[D\right]} \tag{57}$$

$$X_{1i}^A + X_{1i}^B = \frac{\mathbb{E}_i\left[D\right] - A_i \mathbb{C}ov\left[M_{2i}, D\right] - P_1^B - \nu \Delta X_{1i}^B}{A_i \mathbb{V}ar\left[D\right]}. \tag{58}$$

Note that we expect $\Delta X_{1i}^B \neq 0$ to hold generically, so the second condition will always be satisfied. The first condition is only satisfied when $\Delta X_1^A \neq 0$. Note that we can express Equation (58) as

$$\Delta X_{1i}^B = \frac{A_i \mathbb{V}ar\left[D\right]}{A_i \mathbb{V}ar\left[D\right] + \nu}\left(\frac{\mathbb{E}_i\left[D\right] - A_i \mathbb{C}ov\left[M_{2i}, D\right] - \left(X_{0i}^A + X_{0i}^B\right) A_i \mathbb{V}ar\left[D\right] - P_1^B}{A_i \mathbb{V}ar\left[D\right]} - \Delta X_{1i}^A\right).$$

Note that, when both optimality conditions hold at the same time, it must also be the case that

$$\Delta X_{1i}^B = \frac{P_1^A - P_1^B + \tau P_1^A \operatorname{sgn}\left(\Delta X_{1i}^A\right)}{\nu}.$$

In the environment used in Section 5, which is symmetric, $P_1^A = P_1^B = P_1 = \mu_d - \mu_h - A\mathbb{V}ar\left[D\right] Q$. In that case, we can explicitly solve for the equilibrium portfolio allocations for ΔX_{1i}^B, which take the form

$$\Delta X_{1i}^B = \begin{cases} \frac{\varepsilon_{di} - \varepsilon_{hi}}{A\mathbb{V}ar[D] + \nu}, & \Delta X_{1i}^A = 0 \\ \frac{\operatorname{sgn}\left(\Delta X_{1i}^A\right)\tau P_1}{\nu}, & \Delta X_{1i}^A \neq 0, \end{cases}$$

and for the latent net trades ΔX_{1i}^{A+} and ΔX_{1i}^{A-}, given by

$$\Delta X_{1i}^{A+} = \frac{\varepsilon_{di} - \varepsilon_{hi} - (1+\psi)\tau P_1}{A\mathbb{V}ar\left[D\right]} \sim N\left(\frac{-\tau P_1}{A\mathbb{V}ar\left[D\right]}(1+\psi), \frac{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}{\left(A\mathbb{V}ar\left[D\right]\right)^2}\right) \tag{59}$$

$$\Delta X_{1i}^{A-} = \frac{\varepsilon_{di} - \varepsilon_{hi} + (1+\psi)\tau P_1}{A\mathbb{V}ar\left[D\right]} \sim N\left(\frac{\tau P_1}{A\mathbb{V}ar\left[D\right]}(1+\psi), \frac{\sigma_d^2 + \sigma_h^2 - 2\rho\sigma_d\sigma_h}{\left(A\mathbb{V}ar\left[D\right]\right)^2}\right). \tag{60}$$

Note that Equations (59) and (60) are the counterparts of Equations (37) and (38) in the model with perfect enforcement. As expected, when $\psi \to 0$, no trade migrates to market B, but when $\psi > 0$, the latent distribution of net purchases (sales) is shifted to left (right), implying that there will be fewer purchases (sales), and these will be made by investors with more extreme realizations of $\varepsilon_{di} - \varepsilon_{hi}$.

In terms of (latent) individual turnover, we can express $\frac{\Delta X_{1i}^{A+}}{Q}$ and $\frac{\Delta X_{1i}^{A-}}{Q}$ as follows

$$\frac{\Delta X_{1i}^{A+}}{Q} \sim N\left(\frac{\tau}{\Pi}(1+\psi), 2\pi\left(\Xi^A(0)\right)^2\right) \quad \text{and} \quad \frac{\Delta X_{1i}^{A-}}{Q} \sim N\left(\frac{-\tau}{\Pi}(1+\psi), 2\pi\left(\Xi^A(0)\right)^2\right), \tag{61}$$

where $\Pi = \frac{A\mathbb{V}ar[D]Q}{P_1} = \frac{\mu_d}{P_1} - 1$ denotes the risk premium and $\Xi^A(0)$ denotes laissez-faire aggregate turnover, defined in Equation (63) below. Note that Equation (61) is the counterpart of Equation (39).

Finally, note that $\frac{dX_{1i}^A}{d\tau}$ and $\frac{dX_{1i}^B}{d\tau}$ are given by

$$\frac{dX_{1i}^A}{d\tau} = \frac{-P_1 \operatorname{sgn}\left(\Delta X_{1i}^A\right)}{A\mathbb{V}ar\left[D\right]}(1+\psi) \quad \text{and} \quad \frac{dX_{1i}^B}{d\tau} = \frac{P_1 \operatorname{sgn}\left(\Delta X_{1i}^A\right)}{A\mathbb{V}ar\left[D\right]}\psi,$$

using the convention that $\operatorname{sgn}\left(\Delta X_{1i}^A\right) = 0$ if $\Delta X_{1i}^A = 0$. Therefore

$$\frac{dX_{1i}}{d\tau} = \frac{dX_{1i}^A}{d\tau} + \frac{dX_{1i}^B}{d\tau} = \frac{-P_1 \operatorname{sgn}\left(\Delta X_{1i}^A\right)}{A\mathbb{V}ar\left[D\right]},$$

which is identical to the model with perfect enforcement. This result implies that

$$\frac{dX_{1i}^A}{d\tau} = (1+\psi)\frac{dX_{1i}}{d\tau}.$$

Volume Exploiting the properties of truncated normal distributions, stated in Section C.3, we can express the two elements that will determine $\mathcal{V}^{\mathcal{A}}(\tau)$ as follows:

$$\mathbb{P}\left[\Delta X_{1i}^{A}>0\right]=1-\Phi\left(\alpha_{+}\right)$$
$$\mathbb{E}\left[\Delta X_{1i}^{A}\middle|\,\Delta X_{1i}^{A}>0\right]=\frac{1}{A\mathbb{V}ar\left[D\right]}\left(-\tau P_{1}\left(1+\psi\right)+\sqrt{\sigma_{d}^{2}+\sigma_{h}^{2}-2\rho\sigma_{d}\sigma_{h}}\lambda\left(\alpha_{+}\right)\right),$$

where $\lambda\left(\alpha_{+}\right)=\frac{\phi\left(\alpha_{+}\right)}{1-\Phi\left(\alpha_{+}\right)}$ corresponds to the inverse Mills ratio, whose properties are described in Section C.3, and α_{+} and α are defined as follows

$$\alpha_{+}=\max\left\{\alpha,0\right\}\quad\text{where}\quad\alpha=\frac{\tau P_{1}}{\sqrt{\sigma_{d}^{2}+\sigma_{h}^{2}-2\rho\sigma_{d}\sigma_{h}}}\left(1+\psi\right),\tag{62}$$

where $\phi\left(\cdot\right)$ and $\Phi\left(\cdot\right)$ respectively denote the p.d.f. and c.d.f. of the standard normal distribution. Note that the share of buyers (and sellers, by symmetry) in market A corresponds to $\mathbb{P}\left[\Delta X_{1i}^{A}>0\right]=\int_{i\in\mathcal{B}^{A}(\tau)}dF\left(i\right)$.

Trading volume in market A, expressed in shares, can be expressed as a function of the tax rate and primitives as follows

$$\begin{aligned}\mathcal{V}^{A}\left(\tau\right)&=\int_{i\in\mathcal{B}}\Delta X_{1i}^{A}dF\left(i\right)=\mathbb{P}\left[\Delta X_{1i}^{A}>0\right]\mathbb{E}\left[\Delta X_{1i}^{A}\middle|\,\Delta X_{1i}^{A}>0\right]\\&=\frac{1}{A\mathbb{V}ar\left[D\right]}\left(1-\Phi\left(\alpha_{+}\right)\right)\left(-\tau P_{1}\left(1+\psi\right)+\sqrt{\sigma_{d}^{2}+\sigma_{h}^{2}-2\rho\sigma_{d}\sigma_{h}}\lambda\left(\alpha_{+}\right)\right)\\&=\frac{1}{A\mathbb{V}ar\left[D\right]}\left(-\tau P_{1}\left(1+\psi\right)\left(1-\Phi\left(\alpha_{+}\right)\right)+\sqrt{\sigma_{d}^{2}+\sigma_{h}^{2}-2\rho\sigma_{d}\sigma_{h}}\phi\left(\alpha_{+}\right)\right),\end{aligned}$$

Therefore, we can express turnover in market A as follows:

$$\begin{aligned}\Xi^{A}\left(\tau\right)=\frac{\mathcal{V}^{A}\left(\tau\right)}{Q}&=\frac{1}{A\mathbb{V}ar\left[D\right]Q}\left(-\tau P_{1}\left(1+\psi\right)\left(1-\Phi\left(\alpha_{+}\right)\right)+\sqrt{\sigma_{d}^{2}+\sigma_{h}^{2}-2\rho\sigma_{d}\sigma_{h}}\phi\left(\alpha_{+}\right)\right)\\&=-\frac{\tau}{\Pi}\left(1+\psi\right)\left(1-\Phi\left(\alpha_{+}\right)\right)+\frac{\sqrt{\sigma_{d}^{2}+\sigma_{h}^{2}-2\rho\sigma_{d}\sigma_{h}}}{A\mathbb{V}ar\left[D\right]Q}\phi\left(\alpha_{+}\right)\\&=-\frac{\tau}{\Pi}\left(1+\psi\right)\left(1-\Phi\left(\alpha_{+}\right)\right)+\Xi^{A}\left(0\right)\sqrt{2\pi}\phi\left(\alpha_{+}\right),\end{aligned}\tag{63}$$

which implies that

$$\Xi^{A}\left(0\right)=\frac{\sqrt{\sigma_{d}^{2}+\sigma_{h}^{2}-2\rho\sigma_{d}\sigma_{h}}}{A\mathbb{V}ar\left[D\right]Q}\phi\left(0\right)=\frac{\sqrt{\sigma_{d}^{2}+\sigma_{h}^{2}-2\rho\sigma_{d}\sigma_{h}}}{P_{1}}\frac{1}{\frac{A\mathbb{V}ar\left[D\right]Q}{P_{1}}}\phi\left(0\right).$$

Note that

$$\frac{d\Xi^{A}\left(\tau\right)}{d\tau}=-\frac{P_{1}}{A\mathbb{V}ar\left[D\right]Q}\left(1-\Phi\left(\alpha_{+}\right)\right)\left(1+\psi\right)=-\frac{1}{\Pi}\left(1-\Phi\left(\alpha_{+}\right)\right)\left(1+\psi\right).$$

Trading volume in market B, expressed in shares, can be expressed as a function of the tax rate and primitives as follows

$$\begin{aligned}\mathcal{V}^{B}\left(\tau\right)&=\int_{i\in\mathcal{B}}\Delta X_{1i}^{B}dF\left(i\right)\\&=\int_{i\in\mathcal{B}}\Delta X_{1i}^{B}\mathbb{I}\left[\Delta X_{1i}^{A}=0\right]dF\left(i\right)+\int_{i\in\mathcal{B}}\Delta X_{1i}^{B}\mathbb{I}\left[\Delta X_{1i}^{A}>0\right]dF\left(i\right)\\&=\int_{i\in\mathcal{B}}\frac{\varepsilon_{di}-\varepsilon_{hi}}{A\mathbb{V}ar\left[D\right]+\nu}\mathbb{I}\left[\Delta X_{1i}^{A}=0\right]dF\left(i\right)+\int_{i\in\mathcal{B}}\frac{\operatorname{sgn}\left(\Delta X_{1i}^{A}\right)\tau P_{1}}{\nu}\mathbb{I}\left[\Delta X_{1i}^{A}>0\right]dF\left(i\right)\\&=\frac{1}{A\mathbb{V}ar\left[D\right]+\nu}\int\left(\varepsilon_{di}-\varepsilon_{hi}\right)\mathbb{I}\left[0<\varepsilon_{di}-\varepsilon_{hi}<\left(1+\psi\right)\tau P_{1}\right]dF\left(i\right)+\frac{\tau P_{1}}{\nu}\mathbb{P}\left[\Delta X_{1i}^{A}>0\right]\\&=\frac{1}{A\mathbb{V}ar\left[D\right]+\nu}\frac{\mathbb{E}\left[\varepsilon_{di}-\varepsilon_{hi}\middle|\,0<\varepsilon_{di}-\varepsilon_{hi}<\left(1+\psi\right)\tau P_{1}\right]}{\mathbb{P}\left[0<\varepsilon_{di}-\varepsilon_{hi}<\left(1+\psi\right)\tau P_{1}\right]}+\frac{\tau P_{1}}{\nu}\left(1-\Phi\left(\alpha_{+}\right)\right)\\&=\frac{\psi}{1+\psi}\frac{\sqrt{\sigma_{d}^{2}+\sigma_{h}^{2}-2\rho\sigma_{d}\sigma_{h}}}{A\mathbb{V}ar\left[D\right]}\left(\phi\left(0\right)-\phi\left(\alpha_{+}\right)\right)+\frac{\tau P_{1}}{\nu}\left(1-\Phi\left(\alpha_{+}\right)\right),\end{aligned}$$

where we use the fact that $\varepsilon_{di}-\varepsilon_{hi}\sim N\left(0,\sigma_{d}^{2}+\sigma_{h}^{2}-2\rho\sigma_{d}\sigma_{h}\right)$, as well as $\frac{A\mathbb{V}ar[D]}{A\mathbb{V}ar[D]+\nu}=\frac{\psi}{1+\psi}$, and where

$$\mathbb{P}\left[0<\varepsilon_{di}-\varepsilon_{hi}<\left(1+\psi\right)\tau P_{1}\right]=\Phi\left(\alpha_{+}\right)-\frac{1}{2}$$
$$\mathbb{E}\left[\varepsilon_{di}-\varepsilon_{hi}\middle|\,0<\varepsilon_{di}-\varepsilon_{hi}<\left(1+\psi\right)\tau P_{1}\right]=\sqrt{\sigma_{d}^{2}+\sigma_{h}^{2}-2\rho\sigma_{d}\sigma_{h}}\frac{\phi\left(0\right)-\phi\left(\alpha_{+}\right)}{\Phi\left(\alpha_{+}\right)-\frac{1}{2}},$$

and where α_+ is defined in Equation (62).

Therefore, we can express turnover in market B as follows:

$$\Xi^B(\tau) = \frac{\mathcal{V}^B(\tau)}{Q} = \frac{\psi}{1+\psi}\frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{A\mathbb{V}ar[D]Q}(\phi(0)-\phi(\alpha_+)) + \psi\tau\frac{P_1}{A\mathbb{V}ar[D]Q}(1-\Phi(\alpha_+))$$
$$= \frac{\psi}{1+\psi}\Xi^A(0)\left(1-\frac{\phi(\alpha_+)}{\phi(0)}\right) + \psi\frac{\tau}{\Pi}(1-\Phi(\alpha_+)),$$

where we use the fact that $\frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{A\mathbb{V}ar[D]Q} = \frac{\Xi^A(0)}{\phi(0)}$, the definition of the risk premium, and the fact that $\frac{A\mathbb{V}ar[D]}{A\mathbb{V}ar[D]+\nu} = \frac{\psi}{1+\psi}$ and $\nu = \frac{A\mathbb{V}ar[D]}{\psi}$.

Note that

$$\frac{d\Xi^B(\tau)}{d\tau} = -\frac{\psi}{1+\psi}\Xi^A(0)\frac{\phi'(\alpha_+)}{\phi(0)}\frac{d\alpha_+}{d\tau} - \psi\frac{\tau}{\Pi}\phi(\alpha_+)\frac{d\alpha_+}{d\tau} + \frac{\psi}{\Pi}(1-\Phi(\alpha_+)).$$

Therefore, we can write

$$\frac{\left.\frac{d\Xi^B(\tau)}{d\tau}\right|_{\tau=0}}{\left.\frac{d\Xi^A(\tau)}{d\tau}\right|_{\tau=0}} = \frac{\frac{\psi}{\Pi}(1-\Phi(\alpha_+))}{-\frac{1+\psi}{\Pi}(1-\Phi(\alpha_+))} = -\frac{\psi}{1+\psi}.$$

In this model, note also that

$$\left.\frac{d\log\mathcal{V}}{d\tau}\right|_{\tau=0} = \frac{-P_1}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}\lambda(0)},$$

where $\lambda(0) = \sqrt{\frac{2}{\pi}}$. Finally, note also that

$$\left.\frac{d\log\mathcal{V}^A}{d\tau}\right|_{\tau=0} = (1+\psi)\left.\frac{d\log\mathcal{V}}{d\tau}\right|_{\tau=0} \text{ and } \left.\frac{d\log\mathcal{V}^B}{d\tau}\right|_{\tau=0} = \psi\left.\frac{d\log\mathcal{V}}{d\tau}\right|_{\tau=0}.$$

Marginal welfare impact In general, we can characterize the optimal tax as follows.

Proposition 6. (Imperfect tax enforcement) *When investors can trade in an alternative market without facing the tax, the sign of the optimal tax is given by the sign of*

$$\left.\frac{dV^p}{d\tau}\right|_{\tau=0} = -\int\mathbb{E}_i[D]\left.\frac{dX_{1i}}{d\tau}\right|_{\tau=0}dF(i), \tag{64}$$

where $\frac{dX_{1i}}{d\tau} = \frac{dX_{1i}^A}{d\tau} + \frac{dX_{1i}^B}{d\tau}$. *The optimal financial transaction tax* τ^{**} *corresponds to*

$$\tau^{**} = \frac{\int\mathbb{E}_i[D]\frac{dX_{1i}}{d\tau}dF(i)}{\int\operatorname{sgn}(\Delta X_{1i})P_1\frac{dX_{1i}^A}{d\tau}dF(i)}. \tag{65}$$

Proof. After eliminating terms that do not affect the maximization problem, investors solve:

$$\max_{X_{1i}^A,X_{1i}^B}[\mathbb{E}_i[D]-A_i\mathbb{C}ov[M_{2i},D]-P_1]\left(X_{1i}^A+X_{1i}^B\right) - \tau P_1\left|\Delta X_{1i}^A\right| - \frac{A_i}{2}\mathbb{V}ar[D]\left(X_{1i}^A+X_{1i}^B\right)^2 - \frac{\nu}{2}\left(\Delta X_{1i}^B\right)^2 + T_{1i},$$

with interior optimality conditions shown in the text. The change in investors' certainty equivalents is given by

$$\frac{dV_i^p}{d\tau} = [\mathbb{E}_p[D]-\mathbb{E}_i[D]]\frac{dX_{1i}}{d\tau} + \operatorname{sgn}(\Delta X_{1i})P_1\tau\frac{dX_{1i}^A}{d\tau} - \Delta X_{1i}^A\frac{dP_1^A}{d\tau} - \Delta X_{1i}^B\frac{dP_1^B}{d\tau} + \frac{d\tilde{T}_{1i}}{d\tau},$$

which can be aggregated to find

$$\frac{dV^p}{d\tau} = \int\left([\mathbb{E}_p[D]-\mathbb{E}_i[D]]\frac{dX_{1i}}{d\tau} + \operatorname{sgn}(\Delta X_{1i})P_1\tau\frac{dX_{1i}^A}{d\tau}\right)dF(i).$$

Equations (64) and (65) follow immediately. □

It is possible to derive tighter results within the environment used in the quantitative assessment in Section 5. In this case, we can express the aggregate marginal welfare change $\frac{dV^p}{d\tau}$ as follows

$$
\begin{aligned}
\frac{dV^p}{d\tau} &= \int_{i\in\mathcal{T}^A(\tau)} \left[-\mathbb{E}_i\left[D\right] + \operatorname{sgn}\left(\Delta X_{1i}\right) P_1 \tau \left(1+\psi\right)\right] \frac{dX_{1i}}{d\tau} dF\left(i\right) \\
&= \frac{P_1}{A\mathbb{V}ar\left[D\right]} \left(\int_{i\in\mathcal{B}^A(\tau)} \varepsilon_{di} dF\left(i\right) - \int_{i\in\mathcal{S}^A(\tau)} \varepsilon_{di} dF\left(i\right) - \left(1+\psi\right)\tau P_1 2 \int_{i\in\mathcal{B}^A(\tau)} dF\left(i\right) \right) \\
&= \frac{P_1}{A\mathbb{V}ar\left[D\right]} \left(2\int_{i\in\mathcal{B}^A(\tau)} dF\left(i\right) \right) \left(\frac{\mathbb{E}_{\mathcal{B}^A(\tau)}\left[\mathbb{E}_i\left[D\right]\right] - \mathbb{E}_{\mathcal{S}^A(\tau)}\left[\mathbb{E}_i\left[D\right]\right]}{2} - \left(1+\psi\right)\tau P_1 \right) \\
&= \frac{P_1}{A\mathbb{V}ar\left[D\right]} \left(2\int_{i\in\mathcal{B}^A(\tau)} dF\left(i\right) \right) \left(\frac{\sigma_d\left(\sigma_d - \rho\sigma_h\right)}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}} \lambda\left(\alpha_+\right) - \left(1+\psi\right)\tau P_1 \right) \\
&= \frac{P_1}{A\mathbb{V}ar\left[D\right]} \sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}\, 2\left(1-\Phi\left(\alpha_+\right)\right) \left(\delta^{NF}\lambda\left(\alpha_+\right) - \alpha\right),
\end{aligned}
$$

where α_+ and α are defined in Equation (62) and where δ^{NF} is defined in Equation (41). Note that $\mathbb{E}_{\mathcal{B}(\tau)}\left[\mathbb{E}_i\left[D\right]\right] - \mathbb{E}_{\mathcal{S}(\tau)}\left[\mathbb{E}_i\left[D\right]\right]$ is explicitly computed in Equation (45).

Optimal tax Note that we can write the optimal tax, which in this case we denote by τ^{**}, as

$$
\begin{aligned}
\tau^{**} &= \frac{1}{1+\psi} \frac{\mathbb{E}_{\mathcal{B}(\tau^*)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right] - \mathbb{E}_{\mathcal{S}(\tau^*)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]}{2} \\
&= \frac{1}{1+\psi} \frac{\sigma_d\left(\sigma_d-\rho\sigma_h\right)}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}} \frac{1}{P_1} \lambda\left(\max\left\{ \frac{\tau^{**}P_1\left(1+\psi\right)}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}, 0 \right\} \right).
\end{aligned}
$$

We can rearrange this expression to find that

$$
\frac{\tau^{**}P_1\left(1+\psi\right)}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}} = \frac{\sigma_d\left(\sigma_d-\rho\sigma_h\right)}{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h} \lambda\left(\max\left\{ \frac{\tau^{**}P_1\left(1+\psi\right)}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}, 0 \right\} \right),
$$

which allows us to define $\alpha^* \equiv \frac{\tau^{**}P_1(1+\psi)}{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}$, implying that the fixed point that characterizes the optimal tax can be expressed as

$$
\alpha^* = \delta^{NF}\lambda\left(\max\left\{\alpha^*, 0\right\}\right),
$$

where δ^{NF} is defined in Equation (41) above. Note that $\alpha^\star$ is the same as in the model with perfect enforcement. Once a solution for α^* is found, the optimal tax with imperfect enforcement simply corresponds to

$$
\tau^{**} = \frac{1}{1+\psi} \alpha^* \frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}.
$$

Note that we can write

$$
\tau^{**} = \frac{1}{1+\psi}\tau^*, \quad \text{where} \quad \tau^* = \alpha^* \frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}
$$

is the optimal tax with perfect enforcement.

Finally, note that if one observes $\left.\varepsilon_\tau^{\log \mathcal{V}^A}\right|_{\tau=0}$, then it is possible to find $\frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}$ explicitly, since

$$
\left.\varepsilon_\tau^{\log \mathcal{V}^A}\right|_{\tau=0} = \left.\frac{d\log\mathcal{V}^A}{d\tau}\right|_{\tau=0} = -\frac{1+\psi}{\frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}\sqrt{\frac{2}{\pi}}} \Rightarrow \frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1} = -\frac{1+\psi}{\left.\varepsilon_\tau^{\log \mathcal{V}^A}\right|_{\tau=0}}\sqrt{\frac{\pi}{2}}.
$$

In that case,

$$
\tau^{**} = \frac{\alpha^*}{-\left.\varepsilon_\tau^{\log \mathcal{V}^A}\right|_{\tau=0}}\sqrt{\frac{\pi}{2}}.
$$

Therefore, the optimal tax is pinned down for a given value of $\left.\varepsilon_\tau^{\log \mathcal{V}^A}\right|_{\tau=0}$. Alternatively, if one observes $\left.\varepsilon_\tau^{\log \mathcal{V}}\right|_{\tau=0}$ and ψ, then

$$
\left.\varepsilon_\tau^{\log \mathcal{V}}\right|_{\tau=0} = \left.\frac{d\log\mathcal{V}}{d\tau}\right|_{\tau=0} = -\frac{1}{\frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}\sqrt{\frac{2}{\pi}}} \Rightarrow \frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1} = -\frac{1}{\left.\varepsilon_\tau^{\log \mathcal{V}}\right|_{\tau=0}}\sqrt{\frac{\pi}{2}}.
$$

In that case,

$$\tau^{**} = \frac{1}{1+\psi} \frac{\alpha^*}{-\left.\varepsilon_\tau^{\log \mathcal{V}}\right|_{\tau=0}} \sqrt{\frac{\pi}{2}}.$$

However, for given primitives that pin down $\frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1}$, which in this model is also given by

$$\frac{\sqrt{\sigma_d^2+\sigma_h^2-2\rho\sigma_d\sigma_h}}{P_1} = \Pi\Xi\left(0\right)\sqrt{2\pi},$$

the optimal tax is decreasing in ψ, as stated in the text.

Interpretation of ψ Finally, for the purpose of interpretation, note that it is useful to connect ψ and $\frac{1}{1+\psi}$ to different equilibrium objects. First, note that

$$\psi = \frac{\frac{dX_{1i}^A}{d\tau}}{\frac{dX_{1i}}{d\tau}} - 1 = \frac{\frac{dX_{1i}^B}{d\tau}}{-\frac{dX_{1i}}{d\tau}} = \frac{\left.\frac{d\Xi^B(\tau)}{d\tau}\right|_{\tau=0}}{-\left(\left.\frac{d\Xi^A(\tau)}{d\tau}\right|_{\tau=0} + \left.\frac{d\Xi^B(\tau)}{d\tau}\right|_{\tau=0}\right)},$$

which can be interpreted as the relative volume/turnover sensitives (for a small tax) in market B relative to both markets. Similarly, note that

$$\frac{1}{1+\psi} = \frac{\frac{dX_{1i}}{d\tau}}{\frac{dX_{1i}^A}{d\tau}} = \frac{\frac{dX_{1i}^A}{d\tau}+\frac{dX_{1i}^B}{d\tau}}{\frac{dX_{1i}^A}{d\tau}} = 1 + \frac{\frac{dX_{1i}^B}{d\tau}}{\frac{dX_{1i}^A}{d\tau}} = 1 + \frac{\left.\frac{d\Xi^B(\tau)}{d\tau}\right|_{\tau=0}}{\left.\frac{d\Xi^A(\tau)}{d\tau}\right|_{\tau=0}} = 1 + \frac{\left.\frac{d\Xi^B(\tau)}{d\tau}\right|_{\tau=0}}{\left.\frac{d\Xi^A(\tau)}{d\tau}\right|_{\tau=0}},$$

which can be interpreted as the total volume/turnover sensitives (for a small tax) in both markets relative to market A.

D Proofs and derivations: Section 6

As an intermediate step before proving Proposition 5, I characterize the optimal financial transaction tax under the general assumptions on preferences and beliefs described in Section 6. Lemma 5 shows that investors' expectations of asset payoffs — now computed using investors' own stochastic discount factor — and portfolio sensitivities are still the key variables that determine the optimal tax in this more general framework.

Lemma 5. (General utility and arbitrary beliefs) *a) [Optimal tax formula] The optimal financial transaction tax τ^* satisfies*

$$\tau^* = \tau^*_{\textit{risky}} + \theta\left(\tau^*\right)\tau^*_{\textit{risk-free}},$$

where $\tau^*_{\textit{risky}}$ and $\tau^*_{\textit{risk-free}}$ are given by

$$\tau^*_{\textit{risky}} = \frac{\Omega^r_{\mathcal{B}(\tau^*)} - \Omega^r_{\mathcal{S}(\tau^*)}}{2} \quad \textit{and} \quad \tau^*_{\textit{risk-free}} = \frac{\Omega^f_{\mathcal{B}(\tau^*)} - \Omega^f_{\mathcal{S}(\tau^*)}}{2},$$

where $\Omega^r_{\mathcal{B}(\tau^*)}$ is a weighted sum of the difference between risky-asset buyers' expected returns and the planner's expected returns, computed using the investors' stochastic discount factor m_i, given by

$$\Omega^r_{\mathcal{B}(\tau^*)} \equiv \int_{i\in\mathcal{B}(\tau^*)} \omega^{\mathcal{B}}_{i,r}\left(\tau^*\right)\left(\mathbb{E}_i\left[m_i\left(\tau^*\right)\frac{D}{P_1\left(\tau^*\right)}\right] - \mathbb{E}_p\left[m_i\left(\tau^*\right)\frac{D}{P_1\left(\tau^*\right)}\right]\right)dF\left(i\right), \quad \textit{with} \quad \omega^{\mathcal{B}}_{i,r}\left(\tau^*\right) \equiv \frac{\frac{dX_{1i}(\tau^*)}{d\tau}}{\int_{i\in\mathcal{B}(\tau^*)}\frac{dX_{1i}(\tau^*)}{d\tau}dF\left(i\right)},$$

where $\Omega^r_{\mathcal{S}(\tau^*)}$ is analogously defined for sellers, and where $\Omega^f_{\mathcal{B}(\tau^*)}$ is a weighted sum of the difference between risk-free asset buyers' and the planner's valuation of the risky-free asset, computed using the investors' stochastic discount factor m_i, given by

$$\Omega^f_{\mathcal{B}(\tau^*)} \equiv \int_{i\in\mathcal{B}^f(\tau^*)} \omega^{\mathcal{B}}_{i,f}\left(\tau^*\right)\left(\mathbb{E}_i\left[m_i\left(\tau^*\right)R\left(\tau^*\right)\right] - \mathbb{E}_p\left[m_i\left(\tau^*\right)R\left(\tau^*\right)\right]\right)dF\left(i\right), \quad \textit{with} \quad \omega^{\mathcal{B}}_{i,f}\left(\tau^*\right) \equiv \frac{\frac{dY_{1i}(\tau^*)}{d\tau}}{\int_{i\in\mathcal{B}^f(\tau^*)}\frac{dY_{1i}(\tau^*)}{d\tau}dF\left(i\right)},$$

where $\Omega^f_{\mathcal{S}(\tau^*)}$ is analogously defined for sellers, and where $\theta\left(\tau^*\right)$ denotes the relative marginal change in trading volume of the risk-free asset relative to the risky asset, given by

$$\theta\left(\tau^*\right) = \frac{\frac{d\mathcal{V}^s(\tau^*)}{d\tau}}{\frac{d\mathcal{V}^r(\tau^*)}{d\tau}},$$

where $\mathcal{V}^r(\tau^*) = \int_{i\in\mathcal{B}(\tau^*)} \Delta X_{1i}(\tau^*)\,dF(i)$ *and* $\mathcal{V}^s(\tau^*) = \int_{i\in\mathcal{B}^f(\tau^*)} \Delta Y_{1i}(\tau^*)\,dF(i)$, *and where* $\mathcal{B}$ $(\mathcal{S})$ *and* $\mathcal{B}^f$ $(\mathcal{S}^f)$ *respectively denote the sets of buyers (sellers) of the risky and risk-free assets.*

Proof. It is useful to express Equations (22) and (23) as follows

$$1 - R\mathbb{E}_p[m_i] = R(\mathbb{E}_i[m_i] - \mathbb{E}_p[m_i]) \tag{66}$$

$$P_1 - \mathbb{E}_p[m_i D] = \mathbb{E}_i[m_i D] - \mathbb{E}_p[m_i D] - \tau\,\mathrm{sgn}(\Delta X_{1i})\,P_1, \tag{67}$$

where the notation $\mathbb{E}_p[\cdot]$ represents expectations computed using the planner's belief.

The individual marginal welfare impact of a tax change from the planner's perspective is given by

$$\frac{dV_i^p}{d\tau} = U_i'(C_{1i})\left(-P_1\frac{dX_{1i}}{d\tau} - \frac{dY_{1i}}{d\tau} - \frac{dP_1}{d\tau}\Delta X_{1i} + \frac{d\hat{T}_{1i}}{d\tau}\right) + \beta_i\mathbb{E}_p\left[U_i'(C_{2i})\left(D\frac{dX_{1i}}{d\tau} + R\frac{dY_{1i}}{d\tau} + \frac{dR}{d\tau}Y_{1i}\right)\right].$$

After normalizing by the marginal value of wealth, we can express $\frac{\frac{dV_i^p}{d\tau}}{U_i'(C_{1i})}$ as follows

$$\begin{aligned}
\frac{\frac{dV_i^p}{d\tau}}{U_i'(C_{1i})} &= (-P_1 + \mathbb{E}_p[m_i D])\frac{dX_{1i}}{d\tau} + (-1 + \mathbb{E}_p[m_i]R)\frac{dY_{1i}}{d\tau} - \frac{dP_1}{d\tau}\Delta X_{1i} + \frac{dR}{d\tau}Y_{1i} + \frac{d\hat{T}_{1i}}{d\tau} \\
&= (\mathbb{E}_p[m_i D] - \mathbb{E}_i[m_i D] + \tau\,\mathrm{sgn}(\Delta X_{1i})\,P_1)\frac{dX_{1i}}{d\tau} \\
&\quad + R(\mathbb{E}_p[m_i] - \mathbb{E}_i[m_i])\frac{dY_{1i}}{d\tau} + \frac{dR}{d\tau}Y_{1i} - \frac{dP_1}{d\tau}\Delta X_{1i} + \frac{d\hat{T}_{1i}}{d\tau},
\end{aligned}$$

where the last line uses Equations (66) and (67). Finally, we can exploit market clearing and the fact the tax revenue is rebated to aggregate across investors and compute the aggregate welfare impact of a tax change from the planner's perspective, given by

$$\begin{aligned}
\int \frac{\frac{dV_i^p}{d\tau}}{U_i'(C_{1i})}dF(i) &= \int_{i\in\mathcal{T}(\tau)} (\mathbb{E}_p[m_i D] - \mathbb{E}_i[m_i D] + \tau\,\mathrm{sgn}(\Delta X_{1i})\,P_1)\frac{dX_{1i}}{d\tau}dF(i) \\
&\quad + R\int (\mathbb{E}_p[m_i] - \mathbb{E}_i[m_i])\frac{dY_{1i}}{d\tau}dF(i).
\end{aligned}$$

This expression is the counterpart to Equation (11) in the text, with the introduction of a new term capturing differences in the valuations of the risk-free asset between the investors' and the planner.

Consequently, at an interior optimum, the optimal tax must satisfy

$$\tau^* = \frac{\int_{i\in\mathcal{T}(\tau^*)} (\mathbb{E}_i[m_i(\tau^*)D] - \mathbb{E}_p[m_i(\tau^*)D])\frac{dX_{1i}}{d\tau}(\tau^*)\,dF(i) + R(\tau^*)\int (\mathbb{E}_i[m_i(\tau^*)] - \mathbb{E}_p[m_i(\tau^*)])\frac{dY_{1i}}{d\tau}(\tau^*)\,dF(i)}{\int_{i\in\mathcal{T}(\tau^*)} P_1(\tau^*)\,\mathrm{sgn}(\Delta X_{1i}(\tau^*))\frac{dX_{1i}}{d\tau}(\tau^*)\,dF(i)}, \tag{68}$$

which can be expressed as follows

$$\begin{aligned}
\tau^* = &\underbrace{\frac{\int_{i\in\mathcal{T}(\tau^*)}\left(\mathbb{E}_i\left[m_i(\tau^*)\frac{D}{P_1(\tau^*)}\right] - \mathbb{E}_p\left[m_i(\tau^*)\frac{D}{P_1(\tau^*)}\right]\right)\frac{dX_{1i}}{d\tau}(\tau^*)\,dF(i)}{2\int_{i\in\mathcal{B}(\tau^*)}\frac{dX_{1i}}{d\tau}(\tau^*)\,dF(i)}}_{=\frac{\Omega^r_{\mathcal{B}(\tau^*)} - \Omega^r_{\mathcal{S}(\tau^*)}}{2}} + \\
&+ \underbrace{\frac{\int_{i\in\mathcal{B}^f(\tau^*)}\frac{dY_{1i}}{d\tau}(\tau^*)\,dF(i)}{\int_{i\in\mathcal{B}(\tau^*)}\frac{dX_{1i}}{d\tau}(\tau^*)\,dF(i)}}_{=\theta(\tau^*)}\underbrace{\frac{\int(\mathbb{E}_i[m_i(\tau^*)R(\tau^*)] - \mathbb{E}_p[m_i(\tau^*)R(\tau^*)])\frac{dY_{1i}}{d\tau}(\tau^*)\,dF(i)}{2\int_{i\in\mathcal{B}^f(\tau^*)}\frac{dY_{1i}}{d\tau}(\tau^*)\,dF(i)}}_{\frac{\Omega^f_{\mathcal{B}(\tau^*)} - \Omega^f_{\mathcal{S}(\tau^*)}}{2}}.
\end{aligned}$$

Exploiting market clearing for both the risky and the safe asset, we can therefore define

$$\Omega^r_{\mathcal{B}(\tau^*)} \equiv \int_{i\in\mathcal{B}(\tau^*)}\left(\mathbb{E}_i\left[m_i(\tau^*)\frac{D}{P_1(\tau^*)}\right] - \mathbb{E}_p\left[m_i(\tau^*)\frac{D}{P_1(\tau^*)}\right]\right)\underbrace{\frac{\frac{dX_{1i}}{d\tau}(\tau^*)}{\int_{i\in\mathcal{B}(\tau^*)}\frac{dX_{1i}}{d\tau}(\tau^*)\,dF(i)}}_{=\omega_i^{\mathcal{B}}(\tau^*)}dF(i)$$

$$\Omega^r_{\mathcal{S}(\tau^*)} \equiv \int_{i\in\mathcal{S}(\tau)}\left(\mathbb{E}_i\left[m_i(\tau^*)\frac{D}{P_1(\tau^*)}\right] - \mathbb{E}_p\left[m_i(\tau^*)\frac{D}{P_1(\tau^*)}\right]\right)\underbrace{\frac{\frac{dX_{1i}}{d\tau}(\tau^*)}{\int_{i\in\mathcal{S}(\tau^*)}\frac{dX_{1i}}{d\tau}(\tau^*)\,dF(i)}}_{=\omega_i^{\mathcal{S}}(\tau^*)}dF(i).$$

Similarly, we can define

$$\Omega^{f}_{\mathcal{B}(\tau^*)} \equiv \int_{i\in\mathcal{B}^f(\tau^*)} \left(\left(\mathbb{E}_i\left[m_i\left(\tau^*\right)R\left(\tau^*\right)\right]-\mathbb{E}_p\left[m_i\left(\tau^*\right)R\left(\tau^*\right)\right]\right)\right)\underbrace{\frac{\frac{dY_{1i}}{d\tau}\left(\tau^*\right)}{\int_{i\in\mathcal{B}^f(\tau^*)}\frac{dY_{1i}}{d\tau}\left(\tau^*\right)dF\left(i\right)}}_{=\omega_i^{\mathcal{B}f}\left(\tau^*\right)}dF\left(i\right)$$

$$\Omega^{f}_{\mathcal{S}(\tau^*)} \equiv \int_{i\in\mathcal{S}^f(\tau^*)} \left(\mathbb{E}_i\left[m_i\left(\tau^*\right)R\left(\tau^*\right)\right]-\mathbb{E}_p\left[m_i\left(\tau^*\right)R\left(\tau^*\right)\right]\right)\underbrace{\frac{\frac{dY_{1i}}{d\tau}\left(\tau^*\right)}{\int_{i\in\mathcal{S}^f(\tau^*)}\frac{dY_{1i}}{d\tau}\left(\tau^*\right)dF\left(i\right)}}_{=\omega_i^{\mathcal{S}f}\left(\tau^*\right)}dF\left(i\right),$$

which shows the result. Note that $\theta\left(\tau^*\right)$ corresponds to the differential in volume sensitivities to tax changes between the risky asset and risk-free asset, that is

$$\theta\left(\tau^*\right)=\frac{\frac{d\mathcal{V}^s\left(\tau^*\right)}{d\tau}}{\frac{d\mathcal{V}^r\left(\tau^*\right)}{d\tau}},$$

where $\mathcal{V}^r\left(\tau^*\right)=\int_{i\in\mathcal{B}^r(\tau^*)}\Delta X_{1i}\left(\tau^*\right)dF\left(i\right)$ and $\mathcal{V}^s\left(\tau^*\right)=\int_{i\in\mathcal{B}^f(\tau^*)}\Delta Y_{1i}\left(\tau^*\right)dF\left(i\right)$. □

The optimal tax characterized in Lemma 5 has a similar structure to the optimal tax characterized in Proposition 1a) for the baseline model, since it also involves investors' beliefs and portfolio sensitivities. There are three major differences. First, the optimal tax formula now includes risk-adjusted expectations — through investors' stochastic discount factors m_i — of asset returns. By computing this risk-adjustment, the planner can flexibly account for how investors' beliefs affect their portfolio decisions.

Second, the terms $\Omega^r_{\mathcal{B}(\tau^*)}$ and $\Omega^r_{\mathcal{S}(\tau^*)}$ now include a weighted average of *differences* between investors' beliefs and the planner's belief about expected asset returns. In this case, the planner's belief does not drop out of the optimal tax formula, since now $\mathbb{E}_p\left[m_i D\right]$ (also $\mathbb{E}_p\left[m_i R\right]$) takes different values for each investor i. Intuitively, even though portfolio reallocations induced by a tax change still must add up to zero, the fact that different investors value cash-flows differently in different states fails to make the aggregate sum of the induced welfare changes zero-sum in this case.

Third, since now investors also have a consumption-savings decision, the planner finds it desirable to adjust the optimal tax of the risky asset to try to counteract the perceived distortions in investors' risk-free portfolio decisions. Consistent with the second-best Pigouvian logic of the tax, the value of $\theta\left(\tau^*\right)$ modulates how important the belief distortions in consumption-savings decisions are for τ^* depending on the relative sensitivities of trading volume to a tax change in the risky and risk-free asset markets. The expression for $\tau^*_{\text{risk-free}}$ mimics that of τ^*_{risky}, in that it reflects one-half of the difference between buyers and sellers differences between investors' and the planner's risk-adjusted risk-free returns. Intuitively, when risk-free asset volume does not change with the tax, that is, when $\frac{d\mathcal{V}^s\left(\tau^*\right)}{d\tau}=0$, then $\theta\left(\tau^*\right)=0$.

Proposition 5. (Optimal tax approximation)

Proof. Follows immediately from taking the limit in Equation (68). □

Investors' demand/volume approximation Starting from Equation (23), it is possible to recover investors' demand in the baseline model, after approximating investors' date 2 marginal utility around its mean. Formally, note that we can rewrite Equation (23) as follows

$$\begin{aligned}P_1 R\left(1+\operatorname{sgn}\left(\Delta X_{1i}\right)\tau\right)&=\mathbb{E}_i\left[\frac{U_i'\left(W_{2i}\right)}{\mathbb{E}_i\left[U_i'\left(W_{2i}\right)\right]}D\right]\\&=\mathbb{E}_i\left[D\right]+\mathbb{C}ov\left[\frac{U_i'\left(W_{2i}\right)}{\mathbb{E}_i\left[U_i'\left(W_{2i}\right)\right]},D\right].\end{aligned}\tag{69}$$

A first-order approximation of $U_i'\left(W_{2i}\right)$ around its mean implies that

$$U_i'\left(W_{2i}\right)\approx U_i'\left(\mathbb{E}\left[W_{2i}\right]\right)+U_i''\left(\mathbb{E}\left[W_{2i}\right]\right)\left(W_{2i}-\mathbb{E}\left[W_{2i}\right]\right),$$

which at the same time implies that

$$\frac{U_i'\left(W_{2i}\right)}{\mathbb{E}_i\left[U_i'\left(W_{2i}\right)\right]}\approx 1+\frac{U_i''\left(\mathbb{E}\left[W_{2i}\right]\right)}{U_i'\left(\mathbb{E}\left[W_{2i}\right]\right)}\left(W_{2i}-\mathbb{E}\left[W_{2i}\right]\right)\approx 1-A_i\left(W_{2i}-\mathbb{E}\left[W_{2i}\right]\right),$$

where $A_i = -\frac{U_i''(\mathbb{E}[W_{2i}])}{U_i'(\mathbb{E}[W_{2i}])}$. Consequently, we can express $\mathbb{C}ov\left[\frac{U_i'(W_{2i})}{\mathbb{E}_i\left[U_i'(W_{2i})\right]}, D\right]$ as follows

$$\mathbb{C}ov\left[\frac{U_i'(W_{2i})}{\mathbb{E}_i\left[U_i'(W_{2i})\right]}, D\right] \approx \mathbb{C}ov\left[1 - A_i\left(W_{2i} - \mathbb{E}\left[W_{2i}\right]\right), D\right]$$
$$\approx -A_i\left(\mathbb{C}ov\left[M_{2i}, D\right] + X_{1i}\mathbb{V}ar\left[D\right]\right).$$

Substituting into Equation (69), it immediately follows that

$$X_{1i} \approx \frac{\mathbb{E}_i\left[D\right] - A_i\mathbb{C}ov\left[M_{2i}, D\right] - P_1 R\left(1 + \operatorname{sgn}\left(\Delta X_{1i}\right)\tau\right)}{A_i\mathbb{V}ar\left[D\right]},$$

which corresponds to Equation (5) in the text. Consequently, thee volume decomposition in Proposition 2 also remains valid in approximate terms.

E Additional results

In this section, I provide extended remarks on several important issues that are only briefly mentioned in the body of the paper.

E.1 Extended comparison to alternative criteria

It is easier to illustrate the comparison between the welfare criterion in this paper and those of Gilboa, Samuelson and Schmeidler (2014) and Brunnermeier, Simsek and Xiong (2014) by exploring two different scenarios. Both scenarios can be seen as the following special case of the general model studied in the paper.

Environment Consider a environment in which there are two (types of) investors, indexed by $i = \{A, B\}$ — there should be no confusion between investor's A index and the risk aversion coefficient. All investors have identical risk aversion, so $A_i = A$, $\forall i$, and start with the same asset holdings, $X_{0i} = Q$, $\forall i$. Also, $M_{1i} = 0$ and $\mathbb{E}\left[M_{2i}\right] - \frac{A_i}{2}\mathbb{V}ar\left[M_{2i}\right] = 0$, $\forall i$. Investors' beliefs and hedging motives respectively correspond to

$$\mathbb{E}_A\left[D\right] = \mu_d + \varepsilon_d \quad \text{and} \quad A\mathbb{C}ov\left[M_{2A}, D\right] = \varepsilon_h$$
$$\mathbb{E}_B\left[D\right] = \mu_d - \varepsilon_d \quad \text{and} \quad A\mathbb{C}ov\left[M_{2B}, D\right] = -\varepsilon_h.$$

Hence, when $\varepsilon_d \geq 0$, as assumed in both scenarios, A investors will be relatively more optimistic than B investors, and would demand more of the risky asset. When $\varepsilon_h \leq 0$, A investors also have a higher demand for the risky asset relative to B investors, but this time because of hedging reasons. As in the body of the paper, I measure investors' welfare through their certainty equivalents, given by $\mathbb{E}\left[W_{2i}\right] - \frac{A}{2}\mathbb{V}ar\left[W_{2i}\right]$. In this environment, the certainty equivalents from the planner's perspective and the investors' perspective are respectively given by

$$V_i^p = \left(\mathbb{E}_p\left[D\right] - A\mathbb{C}ov\left[M_{2i}, D\right]\right)X_{1i} - \frac{A}{2}\mathbb{V}ar\left[D\right]\left(X_{1i}\right)^2$$
$$V_i^i = \left(\mathbb{E}_i\left[D\right] - A\mathbb{C}ov\left[M_{2i}, D\right]\right)X_{1i} - \frac{A}{2}\mathbb{V}ar\left[D\right]\left(X_{1i}\right)^2.$$

Figure A.6 shows the marginal welfare impact of an increase in X_{1A} in both cases, given by

$$\frac{dV_i^p}{dX_{1i}} = \mathbb{E}_p\left[D\right] - A\mathbb{C}ov\left[M_{2i}, D\right] - A\mathbb{V}ar\left[D\right]X_{1i} \tag{70}$$
$$\frac{dV_i^i}{dX_{1i}} = \mathbb{E}_i\left[D\right] - A\mathbb{C}ov\left[M_{2i}, D\right] - A\mathbb{V}ar\left[D\right]X_{1i}. \tag{71}$$

The three allocations worth discussing are summarized in Table 1. First, the no-trade allocation, characterized by $X_{1A} = X_{1B} = Q$. Second, the laissez-faire competitive equilibrium allocation, characterized by $X_{1A} = Q + \frac{\varepsilon_d - \varepsilon_h}{A\mathbb{V}ar[D]}$ and $X_{1B} = Q + \frac{-\varepsilon_d + \varepsilon_h}{A\mathbb{V}ar[D]}$. Finally, the competitive equilibrium under the optimal tax τ^*, characterized by $X_{1A} = Q + \frac{-\varepsilon_h}{A\mathbb{V}ar[D]}$ and $X_{1B} = Q + \frac{\varepsilon_h}{A\mathbb{V}ar[D]}$. In this last case, the optimal tax corresponds to $\tau^* = \frac{\varepsilon_d}{P_1}$. In any competitive equilibrium, the equilibrium price for any tax rate is given by $P_1 = \mu_d - A\mathbb{V}ar\left[D\right]Q$.

The first scenario considered here (pure betting) exclusively features non-fundamental trading. The second scenario has both fundamental and non-fundamental trading.

Allocations	Risky Asset Holdings X_{1A}	X_{1B}
No-trade	Q	Q
Competitive Equilibrium (laissez-faire)	$Q + \frac{\varepsilon_d - \varepsilon_h}{A\mathbb{V}ar[D]}$	$Q + \frac{-\varepsilon_d + \varepsilon_h}{A\mathbb{V}ar[D]}$
Competitive Equilibrium ($\tau = \tau^*$)	$Q + \frac{-\varepsilon_h}{A\mathbb{V}ar[D]}$	$Q + \frac{\varepsilon_h}{A\mathbb{V}ar[D]}$

Table 1: Relevant allocations

E.1.1 Pure betting scenario ($\varepsilon_d \geq 0$, $\varepsilon_h = 0$)

In the pure betting scenario, consistent with Lemma 4, the optimal tax implements the no-trade allocation. In this case, the no-betting Pareto criterion of Gilboa, Samuelson and Schmeidler (2014) does not rank the no-trade/optimal tax allocation relative to the laissez-faire allocation. It is evident that both investors prefer the laissez-faire allocation (given by the intersection point L_1^* in Figure A.6a) to the no-trade/optimal tax allocation (given by the intersection point L_0^* in Figure A.6a) when computing welfare using their own beliefs — this follows immediately by invoking the First Welfare Theorem in this context. However, for any single belief welfare assessment, investors prefer the no-trade allocation to the laissez-faire allocation. Therefore, the laissez-faire allocation fails to no-betting Pareto dominate the no-trade allocation, since it fails the second requirement of that criterion. At the same time, the no-trade/optimal tax allocation fails to no-betting Pareto dominate the laissez-faire allocation, since it fails the first requirement of that criterion.

In this case, using the criterion of Brunnermeier, Simsek and Xiong (2014), the optimal tax allocation is the best belief-neutral efficient allocation within the set of competitive equilibria with transfers for a specific set of welfare weights. In this particular scenario, the no-trade allocation is also a belief-neutral Pareto Efficient allocation, since it can be found as the solution to a planning problem. The latter is not a general result. If we included some investors who only trade for hedging reasons, the optimal-tax allocation would fail to be belief-neutral Pareto Efficient, although it would still select the best belief-neutral efficient allocation within the set of competitive equilibria with transfers for a specific set of welfare weights.

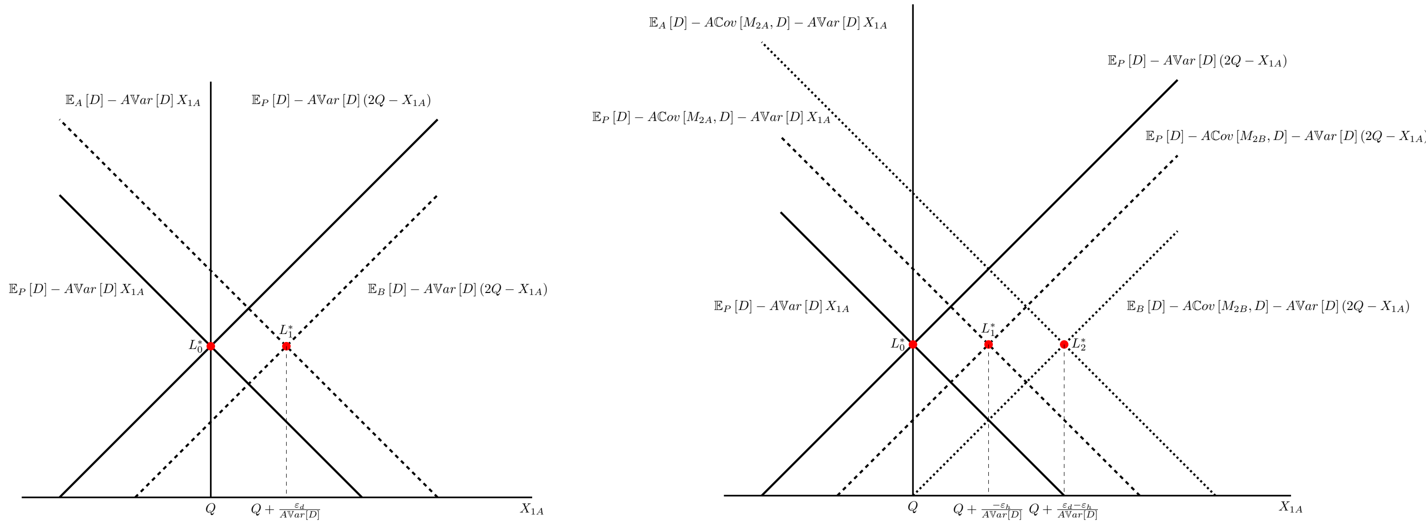
Finally, because single-belief welfare assessments are independent of the belief chosen, allocations that satisfy the no-betting Pareto dominance criterion also satisfy the Unanimity Pareto criterion proposed by Gayer et al. (2014), since these assessments are valid in particular for the beliefs of each of the investors in the economy. This remark also applies to the scenario discussed next.

E.1.2 Fundamental/Non-fundamental trading scenario ($\varepsilon_d \geq 0$, $\varepsilon_h \leq 0$)

In this scenario, trading volume is still positive under the optimal tax. The optimal tax allocation (given by the intersection point L_1^* in Figure A.6b) fails to no-betting Pareto dominate the laissez-faire equilibrium allocation (given by the intersection point L_2^* in Figure A.6b) and vice versa, following a similar logic to the one described in the pure-betting case. Investors prefer the laissez-faire allocation using their own beliefs, but there is no single belief assessment that prefers the laissez-faire allocation to the optimal tax allocation. However, in this scenario, the optimal tax allocation will no-betting Pareto dominate the no-trade allocation, since the two conditions of the criterion are satisfied. First, any single-belief assessment prefers the optimal tax allocation to no-trade. Second, all investors prefer the optimal tax allocation to no trading using their own beliefs.

As in the previous case, the optimal tax allocation is the best belief-neutral efficient allocation within the set of competitive equilibria with tax for some welfare weights and also characterizes a belief-neutral Pareto Efficient allocation. As discussed above, the optimal tax allocation would fail to be belief-neutral Pareto Efficient if we introduce any additional heterogeneity.

Even though I have chosen these two scenarios to illustrate the relation between the criterion used in the paper with the criteria of Gilboa, Samuelson and Schmeidler (2014) and Brunnermeier, Simsek and Xiong (2014), the Pareto notions in both papers will in general fail to be able to rank the optimal tax allocation characterized in this paper once there is rich heterogeneity among investors. In particular, different investors with a different fundamental/non-fundamental trading motive mix will disagree on the desirability of different tax levels. Similarly, including a set of optimistic sellers/pessimistic buyers will immediately generate winners and losers from any policy.



(a) Pure betting scenario

(b) Fundamental/Non-fundamental trading scenario

Figure A.6: Extended comparison with GSS and BSX

Note: Figure A.6 shows the marginal welfare impact of increasing X_{1A} for both A (downward sloping lines) and B investors (upward sloping lines) from the perspective of the planner (solid lines) and from the perspective of each individual set of investors (dashed and dotted lines). Figure A.6a illustrates the pure betting scenario ($\varepsilon_d > 0, \varepsilon_h = 0$). Figure A.6b illustrates the alternative scenario in which $\varepsilon_d \geq 0, \varepsilon_h \leq 0$. The intersection of the solid lines corresponds to the solution of the planner's problem. The intersection of the dashed lines corresponds to the laissez-faire equilibrium, which in this case can be computed as the solution to a planning problem in which investors' beliefs are respected. The intersection of the dotted lines in Figure A.6b corresponds to the allocation under the optimal tax. Both plots are drawn under the assumption that $\mathbb{E}_p[D] = \mu_d$, but note that varying $\mathbb{E}_p[D]$ only shifts the curves vertically, which has no impact on the ranking of allocations.

E.2 Transaction taxes vs. transaction subsidies

It is worth discussing two different issues that emerge when investors face transaction subsidies instead of taxes. First, I compare the properties of an investor's portfolio problem when they face a positive tax ($\tau > 0$) versus a subsidy ($\tau < 0$). Next, I show that trading subsidies can be implemented when paid on the net change of asset holdings over a given period, but cannot be implemented when paid on every purchase or sale.

Figure A.7 illustrates the objective function of an investor, defined in Equation (24), for different values of the tax/subsidy: $\tau = \{-0.4, 0, 0.4\}$. As formally shown above, exploiting Equations (25) and (26), the problem faced by investors features a concave kink when $\tau > 0$ and a convex kink when $\tau < 0$. When $\tau > 0$, investors may find optimal not to trade for some primitives, as in the case considered in the figure. When $\tau < 0$, investors always find optimal to trade for any set of primitives.

Next, I study the implementation problems that may arise under specific forms of trading subsidies. Throughout the paper, as implied by Equation (3), the tax/subsidy base is the change in an investor's net asset position between dates 0 and 1, given by $|\Delta X_{1i}| = |X_{1i} - X_{0i}|$. Now, I assume instead that the tax/subsidy base is given by the sum of net purchases and net sales of the risky asset. Formally, investor i's budget/wealth accumulation constraint is given by

$$W_{2i} = N_{2i} + X_{1i}D + (X_{0i}P_1 - X_{1i}P_1 - \tau P_1 (B_{1i} + S_{1i}) + T_{1i}),$$

where $\Delta X_{1i} = X_{1i} - X_{0i} = B_{1i} - S_{1i}$, and $B_{1i} \geq 0$ and $S_{1i} \geq 0$. That is, an investor who starts with x_0 shares, buys x_b shares and sells x_s shares, ends up with $x_1 = x_0 + x_b - x_s$ shares, so the tax/subsidy under the new assumption is $\tau P_1 (|x_b| + |x_s|)$ dollars, while the tax/subsidy under the assumption sustained in most of the paper is instead $\tau P_1 |x_b - x_s|$ dollars (equivalently, $\tau P_1 |x_1 - x_0|$).

Formally, an investor now chooses X_{1i}, B_{1i}, and S_{1i} in order to maximize

$$J(X_{1i}) = [\mathbb{E}_i[D] - A_i \mathbb{C}ov[M_{2i}, D] - P_1] X_{1i} - \tau P_1 (|B_{1i}| + |S_{1i}|) - \frac{A_i}{2} \mathbb{V}ar[D] (X_{1i})^2$$
$$- \eta (X_{1i} - X_{0i} - B_{1i} + S_{1i}) + \eta^B B_{1i} + \eta^S S_{1i},$$

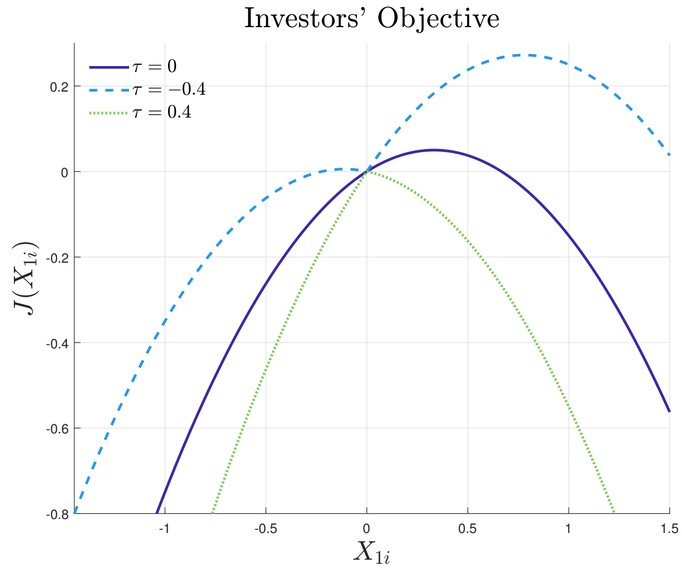


Figure A.7: Investors' objective for $\tau \gtreqless 0$

Note: Figure A.7 illustrates the objective function of an investor, $J(X_{1i})$, defined in Equation (24), for different values of the tax/subsidy: $\tau = \{-0.4, 0, 0.4\}$. The parameters used to draw these plots are $X_{0i} = 0$, $\mathbb{E}_i[D] = 1.3$, $\mathbb{V}ar[D] = 0.9$, $\mathbb{C}ov[M_{2i}, D] = 0$, $A = 1$, $P_1 = 1$, and $T_{1i} = 0$. These values are chosen to highlight the shape of the investors' objective, and are unrelated to those used for the quantitative exercise in Section 5 of the paper.

where $\eta \gtreqless 0$ denotes the Lagrange multiplier on the constraint $X_{1i} - X_{0i} = B_{1i} - S_{1i}$, and $\eta^B \geq 0$ and $\eta^S \geq 0$ respectively denote the Lagrange multipliers on the non-negativity constraints on buying and selling choices.

The first-order conditions to this problem are given by

$$\frac{\partial J}{\partial X_{1i}} = \mathbb{E}_i[D] - A_i \mathbb{C}ov[M_{2i}, D] - P_1 - A_i \mathbb{V}ar[D] X_{1i} - \eta = 0$$
$$\frac{\partial J}{\partial B_{1i}} = -\tau P_1 + \eta + \eta^B = 0$$
$$\frac{\partial J}{\partial S_{1i}} = -\tau P_1 - \eta + \eta^S = 0.$$

Note that $\frac{\partial^2 J}{\partial X_{1i}^2} = -A_i \mathbb{V}ar[D] < 0$, guaranteeing a well defined interior optimum for X_{1i}, but also that the problem is linear in B_{1i} and S_{1i}. Let's consider first the case in which $\tau > 0$. In that case, we can add the first-order conditions for B_{1i} and S_{1i}, which imply that $\eta^S + \eta^B = 2\tau P_1$. Therefore, at least one of the non-negativity constraints is binding at an optimum. Let's assume that it is optimal for an investor to be a buyer, in that case, $\eta^S > 0$ and $\eta^B = 0$, which implies that $X_{1i} - X_{0i} = B_{1i}$ and that X_{1i} is given by

$$\mathbb{E}_i[D] - A_i \mathbb{C}ov[M_{2i}, D] - P_1 - \tau P_1 - A_i \mathbb{V}ar[D] X_{1i} = 0,$$

exactly as implied by Equation (5). A parallel argument can be made when $\eta^S > 0$ and $\eta^B = 0$.

Let's consider next the more interesting subsidy case, in which $\tau < 0$. In that case, one can show that a perturbation in which an investor buys and sells a unit of the asset (often called a "wash trade") increases the investor objective by

$$\frac{\partial J}{\partial B_{1i}} + \frac{\partial J}{\partial S_{1i}} = -2\tau P_1 + \eta^B + \eta^S > 0,$$

which is a strictly positive value whenever $B_{1i} > 0$ and $S_{1i} > 0$. Consequently, we have found a feasible perturbation that increases the objective function everywhere, so an investor finds optimal to choose $B_{1i} = S_{1i} = \infty$, which achieves a value of ∞ and rules out the existence of an equilibrium. At the same time, there is a well defined interior solution for X_{1i}. Note that there are two interior optimum candidates for X_{1i} (as illustrated by the dashed line in Figure A.7), one in which $\frac{\partial J}{\partial X_{1i}} = 0$ is solved with $\eta = \tau P_1$, and another one in which $\eta = -\tau P_1$. Intuitively, under the new formulation for the tax/subsidy base, even though the size of net trades ΔX_{1i} is pinned down by the same forces as when the tax/subsidy base are net trades, investors can engage on "wash trades" to obtain an infinite profit by trading. These results show that the model predicts that all investors would execute wash trades, regardless of whether their trading motives are fundamental or non-fundamental.

E.3 Marginal welfare impact when $\mathbb{E}_p[D] \neq \mu_d$

Figure A.8 shows the normalized individual welfare impact of a tax change when the planner's mean belief $\mathbb{E}_p[D]$ is different from the average mean belief of investors μ_d. Using the same parameterization as in Figure 2, only the set of left plots, which show the normalized individual welfare impact from the planner's perspective $\frac{dV_i^p}{d\tau}$, change. Formally, note that we can express $\frac{\frac{dV_i^p}{d\tau}}{P_1 Q}$ as follows

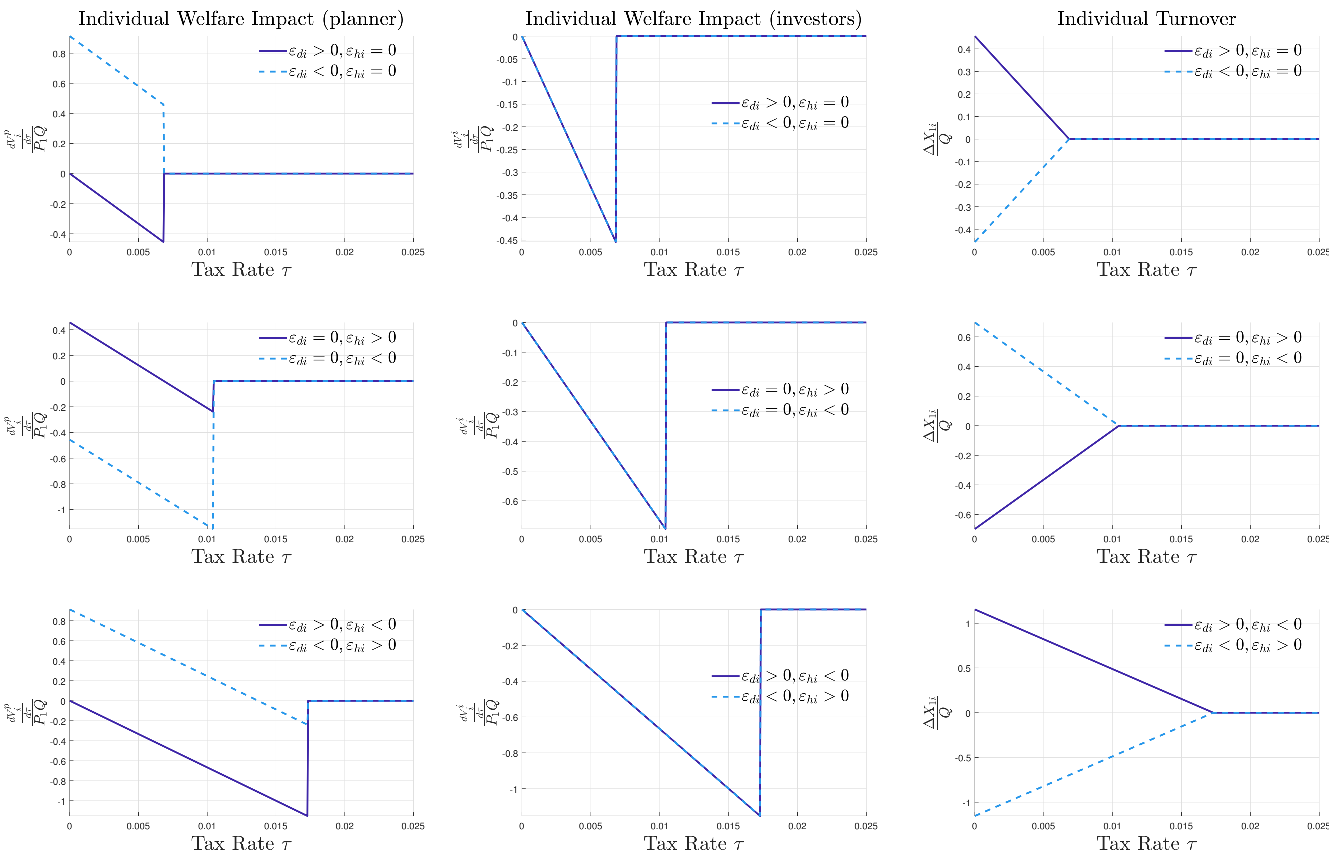


Figure A.8: Marginal welfare impact when $\mathbb{E}_p[D] \neq \mu_d$

Note: Figure A.8 is the counterpart of Figure 2 when assuming that the planner's belief is different from the average belief, that is, $\mathbb{E}_p[D] \neq \mu_d$. Specifically, Figure A.8 assumes that the price-normalized planner's belief $\frac{\mathbb{E}_p[D]}{P_1}$ is one standard deviation above the average on the distribution of investors. Figure 3 applies unchanged in this scenario. Note that the middle and right plots are identical to those in Figure 2.
The top row plots show a buyer ($\varepsilon_{hi} > 0$ and $\varepsilon_{di} = 0$) and a seller ($\varepsilon_{hi} > 0$ and $\varepsilon_{di} = 0$) who only trade for non-fundamental reasons. The middle row plots show a buyer ($\varepsilon_{hi} = 0$ and $\varepsilon_{di} < 0$) and a seller ($\varepsilon_{hi} = 0$ and $\varepsilon_{di} > 0$) who only trade for fundamental reasons. The bottom plots show a buyer ($\varepsilon_{hi} > 0$ and $\varepsilon_{di} < 0$) who buys for non-fundamental and fundamental reasons and a seller ($\varepsilon_{hi} < 0$ and $\varepsilon_{di} > 0$) who sells for non-fundamental and fundamental reasons. The values of ε_{di} and ε_{hi} correspond to one standard deviation of the distributions of $\frac{\varepsilon_{di}}{P_1}$ and $\frac{\varepsilon_{hi}}{P_1}$, respectively. All plots in Figure A.8 use the baseline calibration from Section 5.

$$\frac{\frac{dV_i^p}{d\tau}}{P_1 Q} = \left[\frac{\mathbb{E}_p[D] - \mu_d + \overbrace{\mu_d - \mathbb{E}_i[D]}^{-\varepsilon_{di}}}{P_1} + \operatorname{sgn}(\Delta X_{1i})\, \tau \right] \frac{\frac{dX_{1i}}{d\tau}}{Q} + \frac{\frac{d\tilde{T}_{1i}}{d\tau}}{P_1 Q}, \tag{72}$$

where $\frac{d\tilde{T}_{1i}}{d\tau} = \frac{dT_{1i}}{d\tau} - \frac{d(\tau P_1 |\Delta X_{1i}|)}{d\tau}$ is set to zero in both Figures 2 and A.8. Since Figure A.8 assumes that $\frac{\mathbb{E}_p[D]}{P_1} - \frac{\mu_d}{P_1} > 0$, Equation (72) clearly shows why the planner perceives a systematically higher positive welfare impact of a tax change for

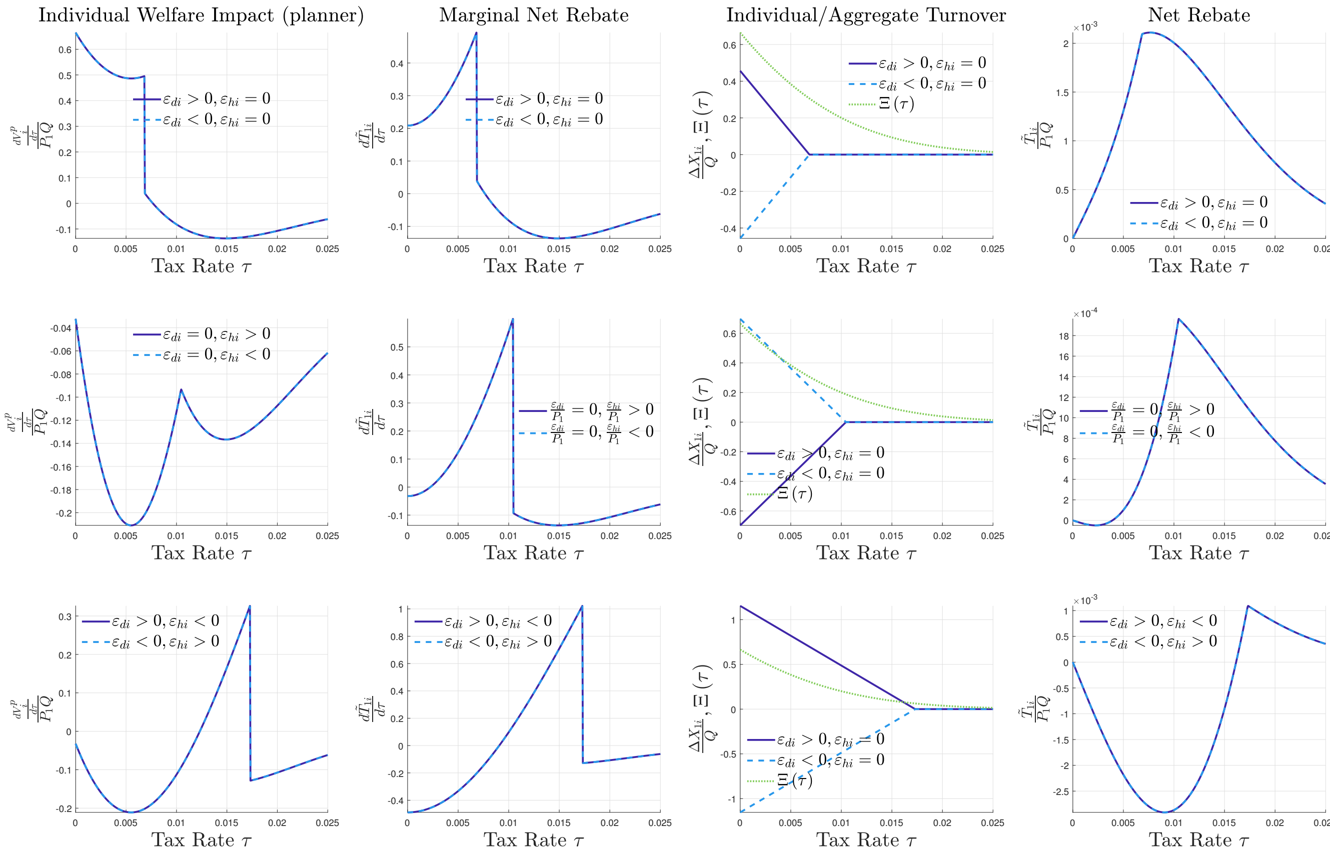


Figure A.9: Uniform rebate rule

Note: Figure A.9 is the counterpart of Figure 2 when instead of assuming an individually targeted rebate rule that fully offsets the tax liability ($\frac{d\tilde{T}_{1i}}{d\tau}=0$), it assumes a uniform rebate rule that distributes tax revenue equally among all investors, including those who do not trade, as described in Equation (73). The left plots show Equation (72), and the middle-left plots show Equation (75). The middle-right plots show individual and aggregate turnover, as defined in Equations (39) and (51), while the right plots show Equation (74).
The top row plots show a buyer ($\varepsilon_{di}>0$ and $\varepsilon_{hi}=0$) and a seller ($\varepsilon_{di}>0$ and $\varepsilon_{hi}=0$) who only trade for non-fundamental reasons. The middle row plots show a buyer ($\varepsilon_{di}=0$ and $\varepsilon_{hi}<0$) and a seller ($\varepsilon_{di}=0$ and $\varepsilon_{hi}>0$) who only trade for fundamental reasons. The bottom plots show a buyer ($\varepsilon_{di}>0$ and $\varepsilon_{hi}<0$) who buys for non-fundamental and fundamental reasons and a seller ($\varepsilon_{di}<0$ and $\varepsilon_{hi}>0$) who sells for non-fundamental and fundamental reasons. The values of ε_{di} and ε_{hi} correspond to one standard deviation of the distributions of $\frac{\varepsilon_{di}}{P_1}$ and $\frac{\varepsilon_{hi}}{P_1}$, respectively. All plots in Figure A.9 use the baseline calibration from Section 5.

sellers than for buyers. This occurs because the term $\left[\frac{\mathbb{E}_p[D]-\mu_d}{P_1}\right]\frac{\frac{dX_{1i}}{d\tau}}{Q}$ takes negative values for buyers, but positive values for sellers. Importantly, the aggregate welfare impact of a tax change, which still corresponds to the one shown in Figure 3, and consequently the optimal tax, are invariant to the level of the planner's mean belief, as shown in Lemma 2 and Proposition 1.

E.4 Uniform rebate rule

In order to clearly illustrate how taxes impact investors' welfare through changes in equilibrium allocations, the illustration of investors' individual marginal welfare impact in Figure 2 and others in this paper assume that the planner follows an individually targeted rebate rule. Under that rule, the net rebate received by each investor for every tax rate τ is 0, which implies that $\frac{d\tilde{T}_{1i}}{d\tau}=0$.

In this section, I assume that the planner follows a uniform rebate rule, in which every investor receives the same rebate, given by the average tax payment among all investors. Formally, the rebate received by an investor i is given by $T_{1i}=\frac{\int\tau P_1|\Delta X_{1i}|dF(i)}{\int dF(i)}$, which only depends on aggregate variables. It is evident that this revenue rule is revenue neutral in the aggregate, since $\int\tilde{T}_{1i}dF(i)=0$. Consequently, the net rebate received by an investor i under a uniform rebate rule can be expressed as follows

$$\tilde{T}_{1i} = \frac{\int \tau P_1 |\Delta X_{1i}| \, dF(i)}{\int dF(i)} - \tau P_1 |\Delta X_{1i}| . \tag{73}$$

Note that the individual net rebate and the marginal impact of a tax change of an investor's net rebate can respectively be expressed, once normalized by the value of risky asset, as follows

$$\frac{\tilde{T}_{1i}}{P_1 Q} = \tau \left(2 \frac{\mathcal{V}(\tau)}{Q} - \frac{|\Delta X_{1i}|}{Q} \right) = \tau \left(2\Xi(\tau) - \frac{|\Delta X_{1i}|}{Q} \right) \tag{74}$$

$$\frac{\frac{d\tilde{T}_{1i}}{d\tau}}{P_1 Q} = 2 \left(\Xi(\tau) + \tau \frac{\frac{d\mathcal{V}}{d\tau}}{Q} \right) - \frac{|\Delta X_{1i}| + \tau \frac{d|\Delta X_{1i}|}{d\tau}}{Q}, \tag{75}$$

where the elements of the marginal impact on the rebate can be expressed as

$$\Xi(\tau) = \frac{\mathcal{V}(\tau)}{Q} = (1 - \Phi(\alpha_+)) \left(-\frac{\tau}{\Pi} \right) + \Xi(0) \sqrt{2\pi} \phi(\alpha_+)$$

$$\tau \frac{\frac{d\mathcal{V}}{d\tau}}{Q} = -\frac{\tau}{\Pi} (1 - \Phi(\alpha_+)),$$

while the elements of the marginal impact on investor i's tax liability are given by

$$\frac{|\Delta X_{1i}|}{Q} = \left| \frac{1}{\Pi} \left(\frac{\varepsilon_{di} - \varepsilon_{hi}}{P_1} - \operatorname{sgn}(\Delta X_{1i}) \tau \right) \right| \text{if } i \text{ is a buyer or seller}$$

$$\frac{\frac{d|\Delta X_{1i}|}{d\tau}}{Q} = \begin{cases} 0, & \text{if } i \text{ is inactive} \\ -\frac{1}{\Pi}, & \text{if } i \text{ is a buyer or seller.} \end{cases}$$

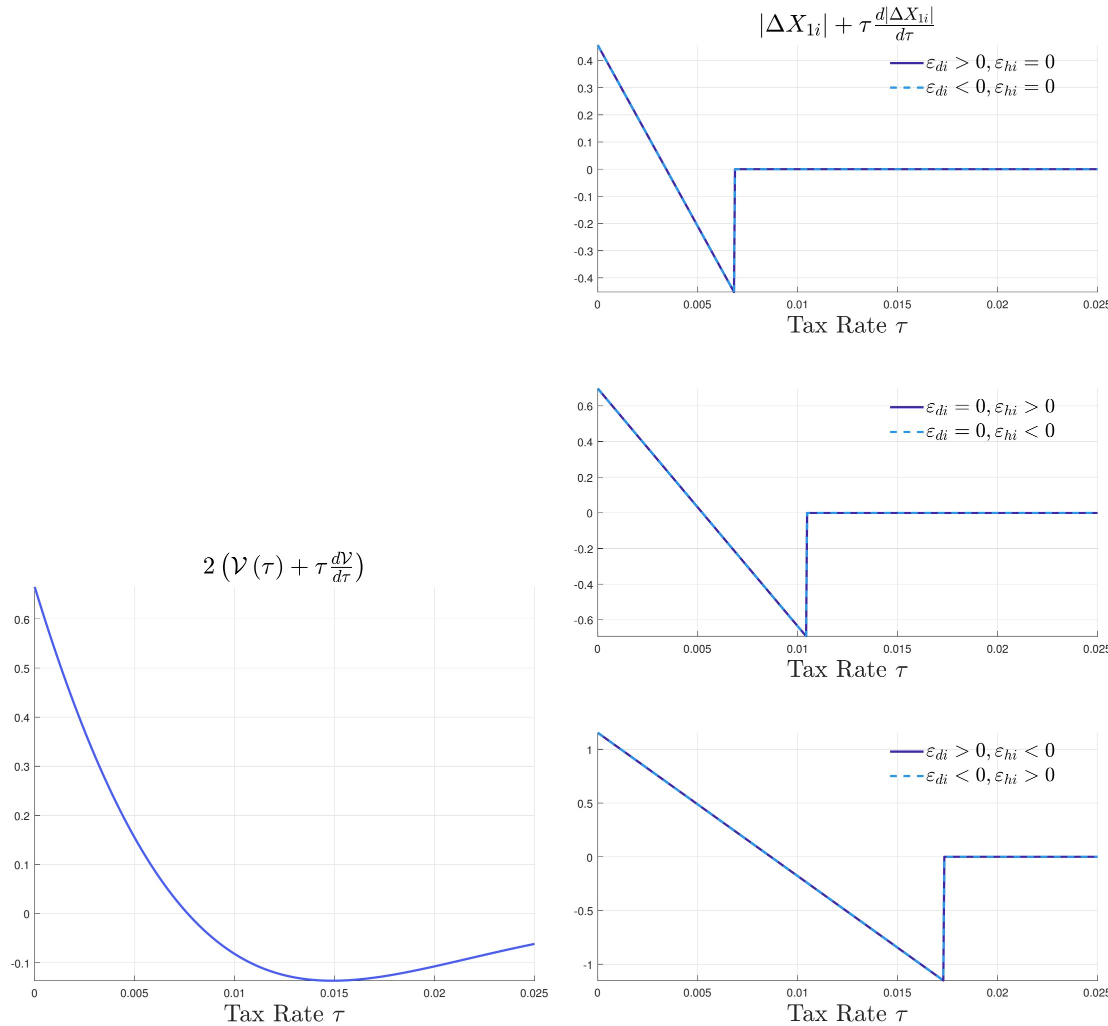
Figures A.9 and A.10 illustrate how using a uniform rebate rule changes the individual welfare impact of tax changes. The rows in Figures A.9 and A.10b parallel those in Figure 2. It's perhaps easiest to understand first the last column in Figure A.9. When $\tau = 0$, by construction the net rebate for all investors is also 0. When τ increases, the net transfer received by each agent depends on his position in the distribution of trades. For instance, the transfer for the bottom row investors decreases with the tax rate for a while, since their trading position gets further away from the average trade initially. When the tax is sufficiently large, their net trading position becomes closer to the average, making their net rebate even positive. After they stop trading, the individual net rebate received by each investor is simply decreasing on the total amount of revenue raised. The second column of these plots represents these effects in marginal terms, while the first column of these plots simply combines the first column of plots in Figure 2 with the second column of plots in this Figure A.9.

Figure A.10b separately illustrates the two components of Equation (75). Figure A.10b shows the marginal impact of an increase in the tax rate on an individual investor's rebate. Intuitively, the average tax payment initially increases with the tax rate up to a point, in which higher taxes reduce total revenue, and consequently the rebates for each investor. Figure A.10a shows the marginal impact of an increase in the tax rate on an individual investor's tax liability. An initial increase in the tax rate τ increases the investors' tax liability, but further increases in τ will reduce the total tax liability, once investors trade less and less until becoming inactive, defining an individual Laffer curve.

E.5 Equilibrium price changes

While the results in Section 5 have been derived in a symmetric environment in which Assumption [S] holds, a change in taxes may have an impact on asset prices, as shown in Lemma 1. These price movements may differentially affect buyers and sellers, as well as the optimal tax. Figure A.11 illustrates such effects in an asymmetric environment with three groups of investors. Figure A.11 is designed to show how $\frac{dP_1}{d\tau}$ can both be increasing and decreasing in the tax rate for different values of τ depending on how the composition of investors varies with the tax level, as shown in Lemma 1.

In this economy, as shown by the top-left plot, the price of the risky asset initially falls with the tax rate, since the share of buyers is initially higher than the share of sellers in the economy. When τ is sufficiently large, most of group 2 investors, which correspond to a set of optimistic investors with low belief dispersion and were buying the risky asset, stop trading, as shown by the bottom right plot, so the share of sellers becomes higher than the share of buyers, as shown by the top middle plot. This makes the asset price increasing in the tax rate, which is consistent with Lemma 1. At $\tau = 0$, this economy matches the same share of non-fundamental trading volume, risk premium, and aggregate turnover as the



(a) Marginal average transfer impact

(b) Marginal individual tax impact

Figure A.10: Marginal net rebate decomposition

Note: Figures A.10a and A.10b show the components of Equation (75). Figure A.10a shows the marginal impact on the investors' rebate of an increase in the tax rate. This plot effectively traces an aggregate revenue Laffer curve. Figure A.10b shows the marginal impact on investors' tax liability of an increase in the tax rate. These plots effectively trace individual revenue Laffer curves.

The top row plots of Figure A.10b show a buyer ($\varepsilon_{hi} > 0$ and $\varepsilon_{di} = 0$) and a seller ($\varepsilon_{hi} > 0$ and $\varepsilon_{di} = 0$) who only trade for non-fundamental reasons. The middle row plots show a buyer ($\varepsilon_{hi} = 0$ and $\varepsilon_{di} < 0$) and a seller ($\varepsilon_{hi} = 0$ and $\varepsilon_{di} > 0$) who only trade for fundamental reasons. The bottom plots show a buyer ($\varepsilon_{hi} > 0$ and $\varepsilon_{di} < 0$) who buys for non-fundamental and fundamental reasons and a seller ($\varepsilon_{hi} < 0$ and $\varepsilon_{di} > 0$) who sells for non-fundamental and fundamental reasons. The values of ε_{di} and ε_{hi} correspond to one standard deviation of the distributions of $\frac{\varepsilon_{di}}{P_1}$ and $\frac{\varepsilon_{hi}}{P_1}$, respectively. All plots in Figure A.10 use the baseline calibration from Section 5.

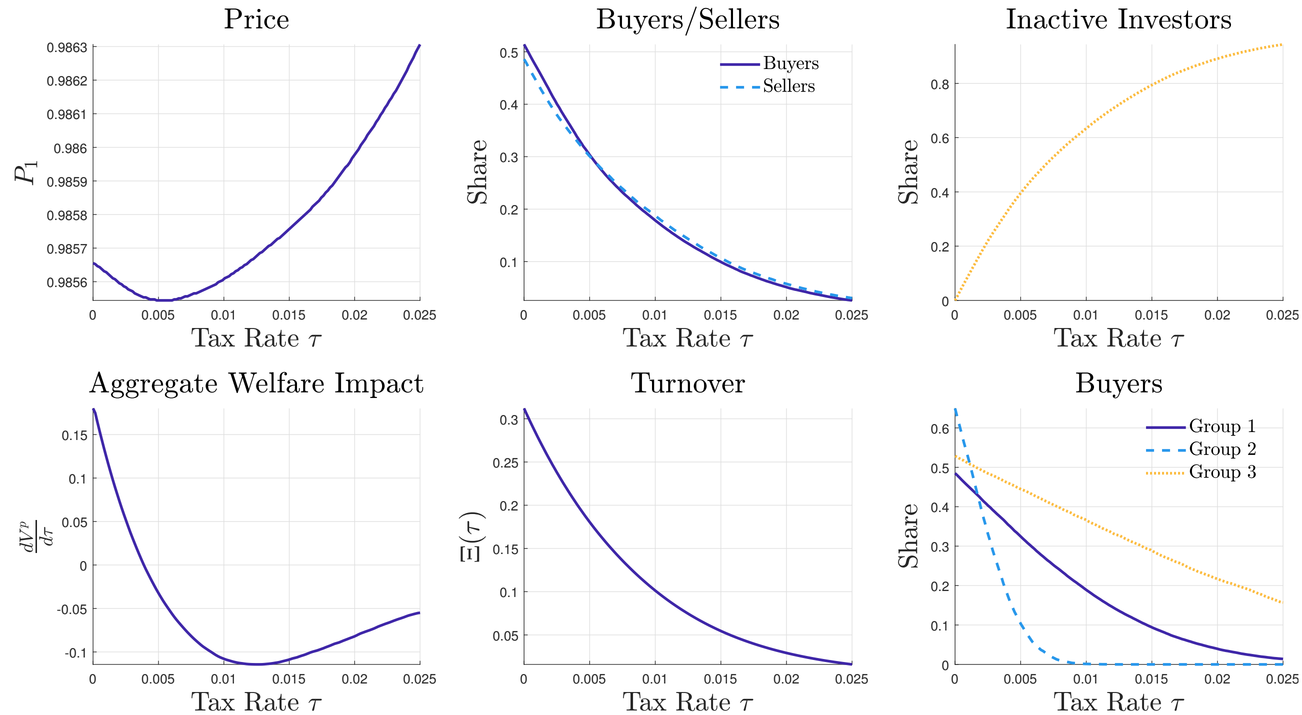


Figure A.11: Pecuniary effects

Note: The top left plot in Figure A.11 shows the equilibrium price P_1 as a function of the tax rate. The top middle plot shows the share of buyers and sellers in the economy as a function of the tax rate. The top right plot shows the share of inactive investors in the economy as a function of the tax rate. The bottom left plot shows the aggregate marginal welfare impact of a tax change. The bottom middle plot shows aggregate turnover. The bottom right plots shows the share of investors of each group who are net buyers of the risky asset as a function of the tax rate.

This simulation is calibrated to match the same three targets used in the quantitative assessment in Section 5. In particular, it features a 0.3 share of non-fundamental trading volume, a 1.5% risk premium, and an aggregate laissez-faire turnover of 31%. This calibration has three groups of investors. Group 1, with 75% of investors, has $\mu_d = 1$ and $\sigma_d = 0.0064$, group 2, with 15% of investors, has $\mu_d = 1.0016$ and $\sigma_d = 0.0016$, and group 3, with 10% of investors, has $\mu_d = 1.0024$ and $\sigma_d = 0.0128$. The optimal tax is $\tau = 0.39\%$. These plots show the outcome of a simulation of the model with $N = 100,000$ investors.

calibrated economy studied in Section 5. In terms of the optimal tax, this calibrated economy, which features equilibrium price changes features an optimal tax somewhat higher, but comparable, to the one in Section 5 ($\tau^* = 0.39\%$).

More broadly, it may seem that the planner in this paper only seeks to improve the welfare of investors with distorted beliefs. That need not be the case, because of general equilibrium effects like the price effect considered here. This is an important observation since there is often greater support for policies that protect bystanders from the mistakes of others than for policies that protect people from their own mistakes. It is true that tax policy can only increase social welfare if some investors hold heterogeneous beliefs. However, which particular investors benefit or lose from the tax policy may depend on the type of general equilibrium effects illustrated. Although the paper focuses on aggregate efficiency, this general equilibrium spillovers will in principle affect all investors in the economy.

E.6 Welfare aggregation

Up to Section 6, the paper makes use of a welfare criterion that involves adding up the sum of investors' certainty equivalents. More generally, in Section 6, in order to facilitate the aggregation of investors' utilities, this paper assumes that the planner uses *uniform generalized social welfare weights*, using the "generalized social welfare weight" terminology introduced in Saez and Stantcheva (2016). Formally, under this aggregation criterion, the marginal impact on aggregate welfare of a tax change can be computed as the sum of the marginal impact on investors' indirect utility, given by $\frac{dV_i}{d\tau}$ and measured in $\frac{\text{utils}}{\text{tax rate}}$, normalized by the investors' marginal value of unit of wealth, given by $U_i'(C_{1i})$ in the case with initial consumption (similarly defined in the case without out) and measured in $\frac{\text{utils}}{\text{dollars}}$, that is

$$\frac{dV^p}{d\tau} = \int \frac{\frac{dV_i^p}{d\tau}}{U_i'(C_{1i})} dF(i),$$

which is measured in $\frac{\text{dollars}}{\text{tax rate}}$. Normalizing investor i's indirect utility by the marginal utility of consumption allows us to add up the marginal impact of a tax change for each investor measured in dollars, which allows for a meaningful aggregation process. Under this welfare criterion, if $\frac{dV^p}{d\tau} > 0$ for a given level of τ, the winners of the policy can always locally compensate the losers in dollar terms, from the planner's perspective. Hence, local welfare comparisons under this approach become similar to those in problems with quasi-linearity or transferable utility. This is a natural assumption in models of corrective taxation with concave utility, and corresponds to a local Kaldor-Hicks interpretation. See Weyl (2019) for further references.

Note that this approach can be mapped to the conventional approach in which a planner maximizes a weighted utilitarian social welfare function built by adding up agents' indirect utility with the different weights. Formally, once we have found the optimal tax using the set of uniform generalized social welfare weights, we can find conventional utilitarian (linear) social welfare weights given by $\lambda_i = \frac{1}{U_i'(C_{1i}(\tau^*))}$ such that the solution to a social welfare maximization problem using those welfare weights also finds that τ^* is the optimal tax. Proposition 4 in Saez and Stantcheva (2016) characterizes this equivalence in the context of an income taxation problem. That result can be trivially extended to the environment studied in this paper.

Finally, note that optimal tax characterizations would look the same for a planner with access to individual specific lump-sum transfers. Formally, a planner with access to ex-ante lump-sum transfers would be able to equalize the marginal value of a unit of wealth across all investors, making the normalization irrelevant to derive optimal tax formulas. The downside of allowing for ex-ante transfers is that those would violate the anonymity assumption sustained throughout the paper.

E.7 Altruism

An alternative defense of the welfare criterion comes from considering altruism. As shown in Figure 2, given his own belief and leaving aside price changes and net rebates, each investor perceives that a transaction tax reduces his individual welfare. However, if investors are at all altruistic towards others, they will agree on implementing a positive tax if the planner finds a positive tax to be optimal. An altruistic investor perceives that a small tax creates a first-order gain for all other investors in the economy, at the cost of a second-order private loss. This approach is consistent with the political philosophy tradition of deliberative democracy, in which individuals think and decide together about what serves the common interest, provided this common interest does not harm much any given individual. The approach followed by the planner in this paper fits legal traditions that consider speculation as fraudulent, because each individual perceives a gain at the expense of others, as well as religious precepts questioning gambling. I formalize this notion as follows.

Let us assume that investors are altruistic of degree $\alpha \in [0,1]$. That is, they compute individual welfare as a linear combination between their on individualistic welfare and social welfare, as computed by the single-belief planner in this paper. More specifically, we'll assume that the planner's belief $\mathbb{E}_p[D]$ is exactly the same as investor i's, although as shown in the paper, the only key restriction is that social welfare is computed consistently using any single belief. Formally, denote by V_i^A the welfare of an altruistic investor with a degree of altruism α, in certainty equivalent terms. In that case, $V_i^A = (1-\alpha) V_i^i + \alpha V^p$, and consequently

$$\frac{dV_i^A}{d\tau} = (1-\alpha) \frac{dV_i^i}{d\tau} + \alpha \frac{dV^p}{d\tau}, \tag{76}$$

where $\frac{dV_i^i}{d\tau}$ denotes the marginal impact of a tax change from an individual's perspective, defined in Equation (44), and $\frac{dV_i^p}{d\tau}$ denotes the marginal impact of a tax change from the perspective of a hypothetical planner which uses the same belief as investor i, defined in Equation (43).

The actual optimal tax preferred by an altruistic investor varies with the degree of altruism α. However, it follows that every investor will prefer a positive tax whenever the planner in this paper prefers a positive tax. Formally,

$$\left.\frac{dV_i^A}{d\tau}\right|_{\tau=0} = \alpha \left.\frac{dV^p}{d\tau}\right|_{\tau=0},$$

since $\left.\frac{dV_i^i}{d\tau}\right|_{\tau=0} = 0$. Hence, whenever $\left.\frac{dV_i^p}{d\tau}\right|_{\tau=0} > 0$ is positive, it must be the case that $\left.\frac{dV_i^A}{d\tau}\right|_{\tau=0} > 0$. Intuitively, for a small tax, each individual investor perceives a second-order welfare loss individually but a first-order gain socially, so they all support a positive optimal tax. When τ is away from zero, every investor whose altruism degree α is less than one will prefer a smaller tax than the optimal tax chosen by the planner.

E.8 Linear combination between planner's belief and investors' beliefs

The results in the paper have been derived under the assumption that the planner maximizes welfare using a single distribution of payoffs to compute the welfare of all investors. It is straightforward to generalize the results to a planner that puts weight γ on social welfare computed with a single belief and weight $1-\gamma$ on social welfare computed respecting investors' individual beliefs. In that case, the new optimal tax τ_γ^* simply turns out to be a scaled version of the optimal tax characterized under the criterion proposed in this paper. Formally, the optimal tax for a planner who puts weight γ on social welfare as computed in this paper and weight $1-\gamma$ on social welfare as computed respecting investors' beliefs is given by

$$\tau_\alpha^* = \gamma \tau^*,$$

where τ^* is given by Equation (12). This result extends to the different extensions analyzed in this paper. This intermediate approach may be appealing to those who prefer a partially paternalistic welfare criterion.

E.9 Transaction taxes under incomplete markets

Even though the model in the paper features incomplete markets, the economy is constrained efficient since there is effectively a single market that clears. This section briefly describes how, beyond the role of differences in beliefs, the pecuniary impact of transaction taxes may by itself have a first-order effect on welfare.

Environment Consider an economy with investors indexed by i who solve the following dynamic problem

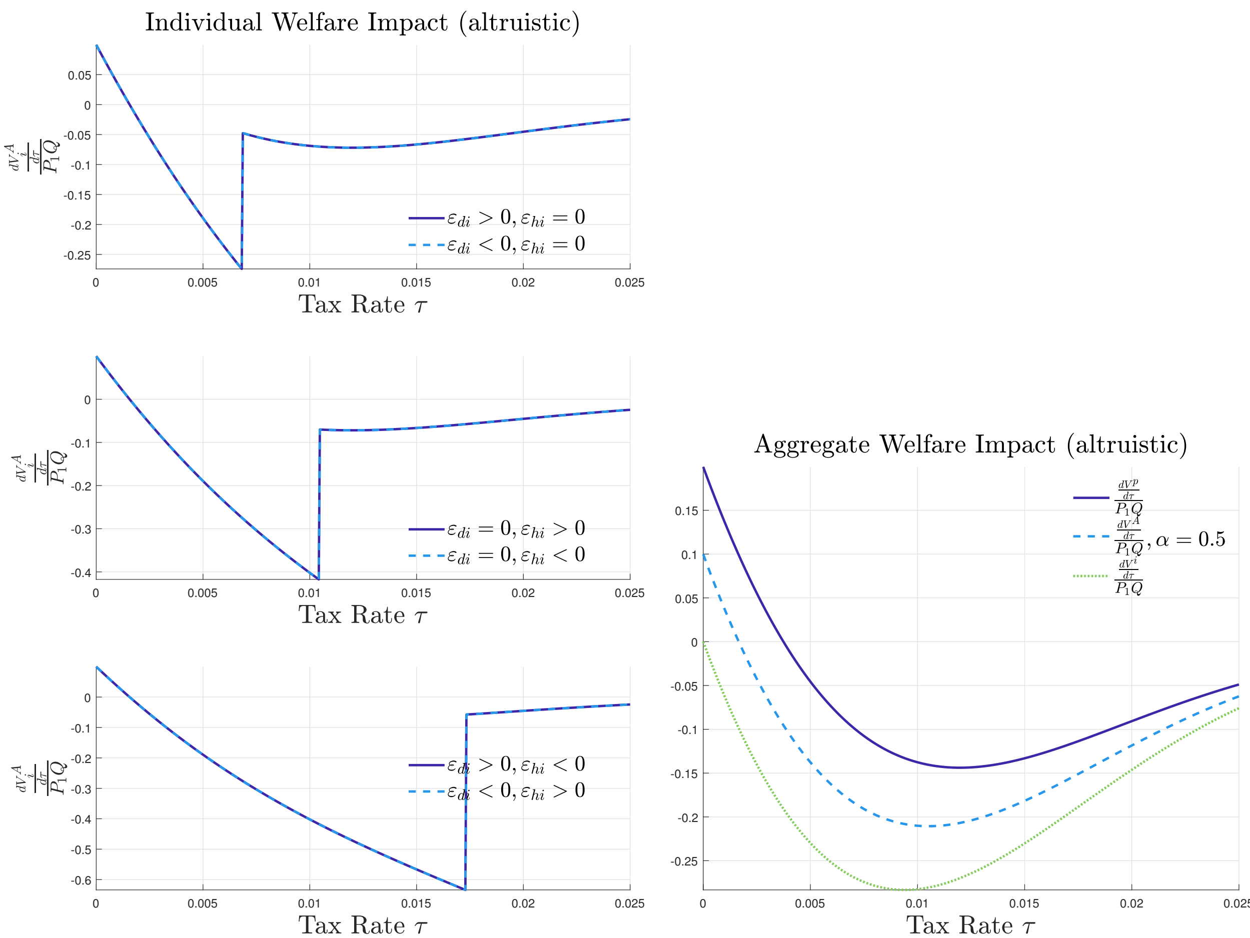
$$\max_{C_{ti}, X_{ti}, Y_{ti}} \mathbb{E}\left[\sum_{t=0}^{T} \beta^t U_i(C_{ti})\right],$$

where

$$C_{ti} = M_{ti} + X_{t-1i}(P_t + D_t) + T_{ti} - \tau P_t |\Delta X_{ti}| - X_{ti} P_t + R Y_{t-1i} - Y_{ti},$$

where C_{ti} denotes investors' consumption, M_{ti} denotes investors' stochastic endowment, D_t denotes the stochastic payoff of the risky asset and P_t its price, $T_{ti} - \tau P_t |\Delta X_{ti}|$ denotes the net rebate received by investors, Y_{ti} and X_{ti} respectively denote investors' holdings of the risky and the safe asset, and R denotes the elastic risk-free rate.

To highlight the role of market incompleteness, I assume that both the planner and the investors' in the economy share the same beliefs. It is straightforward to include differences in beliefs here, as shown in Section 6. For simplicity, I consider an individually targeted rebate rule. The equilibrium definition is standard.



(a) Individual welfare impact (altruistic)

(b) Aggregate welfare impact (altruistic)

Figure A.12: Altruism

Note: Figure A.12a shows the normalized individual marginal welfare impact of a tax change for specific investors from the perspective of an altruistic investor, as defined in Equation (76), for different values of τ. Figure A.12a is the counterpart of Figure 2, so the top row shows a buyer ($\varepsilon_{hi} > 0$ and $\varepsilon_{di} = 0$) and a seller ($\varepsilon_{hi} > 0$ and $\varepsilon_{di} = 0$) who only trade for non-fundamental reasons, the middle row plots shows a buyer ($\varepsilon_{hi} = 0$ and $\varepsilon_{di} < 0$) and a seller ($\varepsilon_{hi} = 0$ and $\varepsilon_{di} > 0$) who only trade for fundamental reasons, and the bottom plots show a buyer ($\varepsilon_{hi} > 0$ and $\varepsilon_{di} < 0$) who buys for non-fundamental and fundamental reasons and a seller ($\varepsilon_{hi} < 0$ and $\varepsilon_{di} > 0$) who sells for non-fundamental and fundamental reasons.

Figure A.12b shows the normalized the normalized aggregate marginal welfare impact of a tax change for a planner who maximizes investors' altruistic utility. The case with $\alpha = 1$ (solid blue line) corresponds to the case studied in the rest of the paper. The case $\alpha = 0$ (dotted green line) corresponds to the case in which individual welfare is compute respecting investors' beliefs. The intermediate case with $\alpha = 0.5$ is a linear combination between the two. When $\alpha = 1$, $\tau^* = 0.37\%$, when $\alpha = 0.5$, $\tau^* = 0.11\%$, while when $\alpha = 0$, $\tau^* = 0$.

All plots in Figure A.12 use the baseline calibration from Section 5.

Results I simply focus here on the first-order of a transaction tax on social welfare, computed as in Section 6, using uniform generalized social welfare weights.

Proposition 7. (Incomplete markets and common beliefs) *The marginal welfare impact of a tax change at $\tau = 0$, which can take any sign, is given by*

$$\left.\frac{dV^p}{d\tau}\right|_{\tau=0} = \int \sum_{t=0}^{T} \mathbb{E}\left[m_{ti} \Delta X_{ti} \left.\frac{dP_t}{d\tau}\right|_{\tau=0}\right] dF(i),$$

where m_{ti} denotes investor i's stochastic discount factor for date t, given by

$$m_{ti} = \frac{\beta^t U_i'(C_{ti})}{U_i'(C_{0i})}.$$

Proof. Note that

$$\frac{dV_i^p}{d\tau} = \mathbb{E}\left[\sum_{t=0}^{T} \frac{\beta^t U_i'(C_{ti})}{U_i'(C_{0i})} \frac{dC_{ti}}{d\tau}\right],$$

where $\frac{dC_{ti}}{d\tau} = \frac{dX_{t-1i}}{d\tau}(P_t + D_t) - X_{ti}P_t + R\frac{dY_{t-1i}}{d\tau} - \frac{dY_{ti}}{d\tau} - \Delta X_{ti}\frac{dP_t}{dt}$. Aggregating over all investors and using investors' individual optimality conditions immediately yields the result. ☐

When markets are incomplete, m_{ti} varies across the distribution of investors in every equilibrium. Since $\frac{dP_t}{d\tau}$ can be positive or negative, and $\int \Delta X_{ti} dF(i) = 0$, $\forall t$, it is straightforward to show that $\left.\frac{dV^p}{d\tau}\right|_{\tau=0}$ can take positive or negative values, which implies that a transaction tax or subsidy may or may not be welfare improving in an economy with incomplete markets. Whether a tax or subsidy is welfare improving will depend on the differences in investors' stochastic discount factors, changes in net trading positions, and the price impact of a tax. See Dávila and Korinek (2018) for a detailed discussion of these effects when characterizing constrained efficient allocations. If markets are complete when $\tau = 0$, then the sequence of m_{ti} is identical across investors, and $\left.\frac{dV^p}{d\tau}\right|_{\tau=0} = 0$, which is simple local version of the First Welfare Theorem. From the perspective of missing hedging markets, 7 also shows that price volatility is not the key variable to target. The correct target is whether transaction taxes modify prices to improve investors' insurance at the right times. More generally, note that pecuniary effects derived from wage changes or relative price changes in a multi-good economy would have similar effects.

E.10 General utility and arbitrary beliefs: quantitative assessment

Finally, it is worth exploring the quantitative implications of the model in the general non-linear case. In order to minimize the differences with the baseline calibration, I effectively fix investors' consumption at the initial date, so they only face a portfolio allocation problem and not a consumption-savings decision. I assume that investors' preferences are given by an isoelastic (CRRA) utility specification of the form $\frac{c^{1-\gamma}}{1-\gamma}$. All investors have the same initial asset holdings of the risky asset and have identical preferences. I also assume that M_{2i} and D are normally distributed and that investors have the same assessment about the unconditional first and second univariate moments of M_{2i}, as well as the variance of D. As in the baseline model, investors have different beliefs about the expected payoff of D, and they all have different hedging needs, modeled in this case in terms of their correlation coefficient between D and M_{2i}.[41]

In terms of solving the model, since there is no closed form to the portfolio choice problem in this case (even without transaction taxes), it is necessary to numerically solve investor's optimality condition, given by

$$\mathbb{E}_i\left[U_i'(W_{2i})(D - P_1(1 + \tau \operatorname{sgn}(\Delta X_{1i})))\right] = 0.$$

I use quadrature methods, with 15 points along each dimension of D and M_{2i}, to approximate the investors' expectations, as well as the planner's. I assume that investors receive individually targeted rebates, so $W_{2i} = M_{2i} + X_{1i}D + (M_{1i} - P_1\Delta X_{1i})$. Figure A.13 illustrates the results, which imply an optimal tax similar and of the same order of magnitude as the optimal tax in the baseline model.

[41] Since households' preferences feature an Inada condition, assuming that D and M_{2i} are normally distributed can be problematic in some calibrations. For the calibration used here, I ensure that consumption is always positive. Similarly, I ensure that the variance-covariance matrix of the random variables D and M_{2i} is positive semi-definite for each investor.

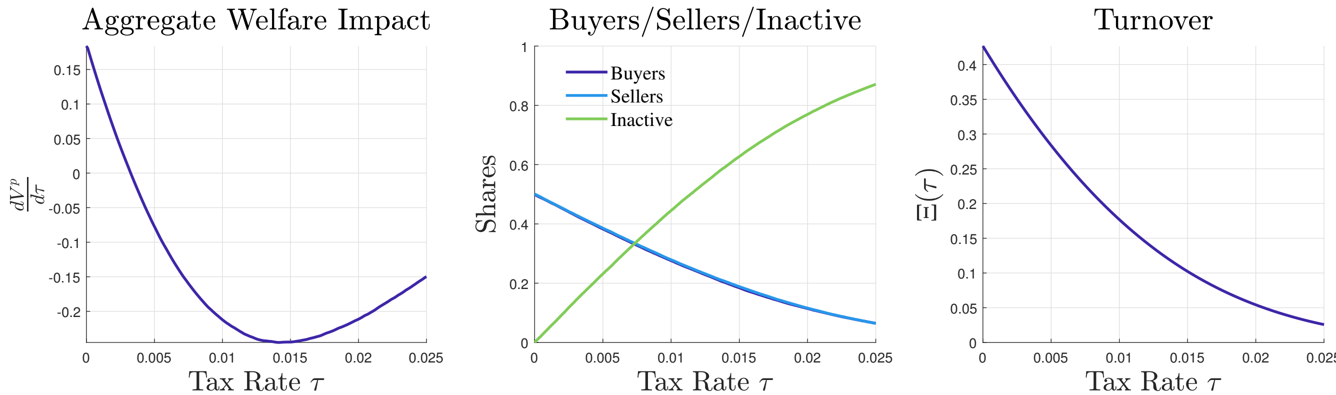


Figure A.13: Simulation with general preferences

Note: The left plot in Figure A.13 shows the aggregate welfare impact of a tax change for different values of τ. The middle plot shows the share of buyers, sellers, and investors who decide not trade for different values of τ. The right plots shows asset turnover for different values of τ. This simulation is calibrated to approximately match the same three targets used in the quantitative assessment in Section 5. In particular, it features a 34% share of non-fundamental trading volume (computed comparing a solution of the model with $\sigma_d = 0$ and $\sigma_h = 0$), a 1.53% risk premium, a volume semi-elasticity of -75, and a value of laissez-faire turnover of 42%. The optimal tax is $\tau = 0.33\%$. These plots show the outcome of a simulation of the model with $N = 100,000$ investors.

F Extensions

I now study multiple extensions of the benchmark model. Earlier versions of this paper included additional extensions.

F.1 Portfolio constraints: short-sale and borrowing constraints

Environment Although participants in financial markets face short-sale and borrowing constraints, investors in the baseline model face no restrictions when choosing portfolios. I now introduce trading constraints into the model as a pair of scalars for $\overline{g}_i$ and $\underline{g}_i$ for every investor i, such that

$$\underline{g}_i \leq X_{1i} \leq \overline{g}_i. \tag{77}$$

Both short-sale constraints and borrowing constraints are special cases of Equation (77). Short-sale constraints can be expressed as $X_{1i} \geq 0$, so $\underline{g}_i = 0$. Borrowing constraints can be mapped to a constraint of the form $X_{1i} \leq \overline{g}_i$.[42] Intuitively, an investor who wants to sufficiently increase his holdings of the risky asset must rely on borrowing. Hence, a borrowing limit is equivalent to an upper bound constraining the amount held of the risky asset.

Results and quantitative assessment The optimal portfolio is identical to the one in the baseline model, unless a constraint binds. In that case, X_{1i} equals the trading limit. Formally, the optimal portfolio decision for an investor is still given by Equation (5), unless the value of ΔX_{1i} implied by that equation violates the short-sale constraint, in which case $X_{1i} = \underline{g}_i$ or $X_{1i} = \overline{g}_i$The only substantial difference with the baseline is the expression that determines the equilibrium price, which is a slightly modified version of Equation (6) in the body of the paper. It is formally given by

$$P_1 = \frac{\int_{i \in \mathcal{U}} \frac{\mathbb{E}_i[D] - A_i \mathbb{C}ov[M_{2i}, D]}{A_i \mathbb{V}ar[D]} dF(i) - \int_{i \in \mathcal{T}} X_{0i} dF(i) - \int_{i \in \mathcal{C}} g_i(P_1) dF(i)}{\int_{i \in \mathcal{U}} \frac{1 + \text{sgn}(\Delta X_{1i})\tau}{A_i \mathbb{V}ar[D]} dF(i)},$$

where $\mathcal{T}$, $\mathcal{U}$, and $\mathcal{C}$ respectively denote the set of active investors, unconstrained investors, and constrained investors.

Proposition 8. (Trading constraints) *The optimal financial transaction tax τ^* satisfies exactly the same expression as in Equation (12).*

Proof. The marginal welfare impact of a tax change for investor i from the planner's perspective is given by

$$\frac{dV_i^p}{d\tau} = [\mathbb{E}_p[D] - \mathbb{E}_i[D] + \text{sgn}(\Delta X_{1i}) P_1 \tau] \frac{dX_{1i}}{d\tau} - \Delta X_{1i} \frac{dP_1}{d\tau} + \frac{d\tilde{T}_{1i}}{d\tau}$$

[42] An earlier version of this paper allowed for price dependent borrowing limits. In that case, the optimal tax must account for the pecuniary binding-constraint effects induced by tax changes.

for every investor, constrained or unconstrained. For this expression to be valid for constrained investors, it is essential that $\frac{dX_{1i}}{d\tau} = 0$ for those investors. Consequently, we can express aggregate welfare change as follows

$$\frac{dV^p}{d\tau} = \int_{i \in \mathcal{T}(\tau)} \left[-\mathbb{E}_i[D] + \text{sgn}(\Delta X_{1i}) P_1 \tau\right] \frac{dX_{1i}}{d\tau} dF(i),$$

as in Equation (30). The expression for the optimal tax follows immediately in that case. □

Intuitively, a marginal tax change at the optimum does not modify the portfolio allocation of constrained investors, leaving their welfare unchanged, i.e., for those investors $\frac{dX_{1i}}{d\tau} = 0$. Intuitively, investors who face trading constraints are infra-marginal for the determination of aggregate welfare.

Figure A.14 illustrates the role of short-sale constraints and how they affect the optimal tax within a model calibrated (without constraints and when $\tau = 0$) to match the share of non-fundamental trading volume, the level of turnover, and the risk premium, as in the quantitative assessment of the model in Section 5. Unsurprisingly, the top left plot in Figure A.14 shows that the equilibrium price when investors face short-sale constraints is higher than the equilibrium price when investors are unconstrained – this result is often attributed to Miller (1977). Since there are more buyers than sellers when investors are short-sale constrained, it is also natural to find that increases in the tax rate reduce the equilibrium price in the economy with short-sale constraints — the mechanism behind these effects was described in detail in Section E.5.

As expected, aggregate turnover is decreasing in the level of tax, while the share of inactive investors is increasing in the tax rate. In general, it is not possible to say in which direction introducing short-sale constraints shift the optimal tax, but it is possible to understand why in the model used in this section, the optimal tax is lower when there are short-sale constraints. In this case, starting from $\tau = 0$, the constraints will restrict the portfolio holdings of both fundamental and non-fundamental sellers. However, at the margin, only the non-fundamental/belief distortions part enters the welfare calculation of the planner. Since increasing the tax has now a smaller effect on restricting the trades of pessimists (since they already could not sell as much as they want and for them $\frac{dX_{1i}}{d\tau} = 0$), there is a lower marginal benefit from taxation and the optimal tax is lower.[43] Finally, the bottom right plot in Figure A.14 shows that higher tax rates reduce the number of short-sale constrained investors. In the limit, when τ is sufficiently large, all investors find optimal not to trade, so their short-sale constraints will necessarily be slack.

F.2 Pre-existing trading costs

Because actual investors face trading costs even when there are no taxes, one could wonder about the validity of the results derived around the point $\tau = 0$. Here, I show that the optimal tax formula is still valid as long as transaction costs are a mere compensation for the use of economic resources.[44]

Environment Investors now face transaction costs, regardless of the value of τ. These represent costs associated with trading, like brokerage commissions, exchange fees, or bookkeeping costs. Investors must pay a quadratic cost, parameterized by α, a linear cost η on the number of shares traded, and a linear cost ς on the dollar volume of the transaction. These trading costs are paid to a new group of investors (intermediaries), which facilitate the process of trading. Crucially, I assume that intermediaries make zero profits in equilibrium. Hence, wealth at date 2 for an investor i is now given by

$$W_{2i} = M_{2i} + X_{1i}D + X_{0i}P_1 - X_{1i}P_1 - |\Delta X_{1i}| P_1 (\tau + \varsigma) - \eta |\Delta X_{1i}| - \frac{\alpha}{2} (\Delta X_{1i})^2 + T_{1i}. \tag{78}$$

Tax revenues are rebated to investors through an arbitrary transfer, but not trading costs.

Results The demand for the risky asset takes a similar form as in the baseline model, featuring also an inaction region, now determined jointly by the trading costs and the transaction tax. The optimal portfolio given prices can be compactly written in the trade region as:

$$X_{1i} = \frac{\mathbb{E}_i[D] - A_i \mathbb{C}ov[M_{2i}, D] - P_1 (1 + \text{sgn}(\Delta X_{1i})(\tau + \varsigma)) - \text{sgn}(\Delta X_{1i}) \eta + \alpha X_{0i}}{A_i \mathbb{V}ar[D] + \alpha}.$$

All three types of trading costs — quadratic, linear in shares and linear in dollar value — shift investors' portfolios towards their initial positions. The equilibrium price is a slightly modified version of Equation (6).

[43] There is scope to explore the optimal determination of short-sale and borrowing constraints in a similar environment to the one considered here

[44] There is scope to study in more detail the interaction of trading costs in models that deliver endogenous bid-ask spreads in models with differentially informed investors, like that of Glosten and Milgrom (1985).

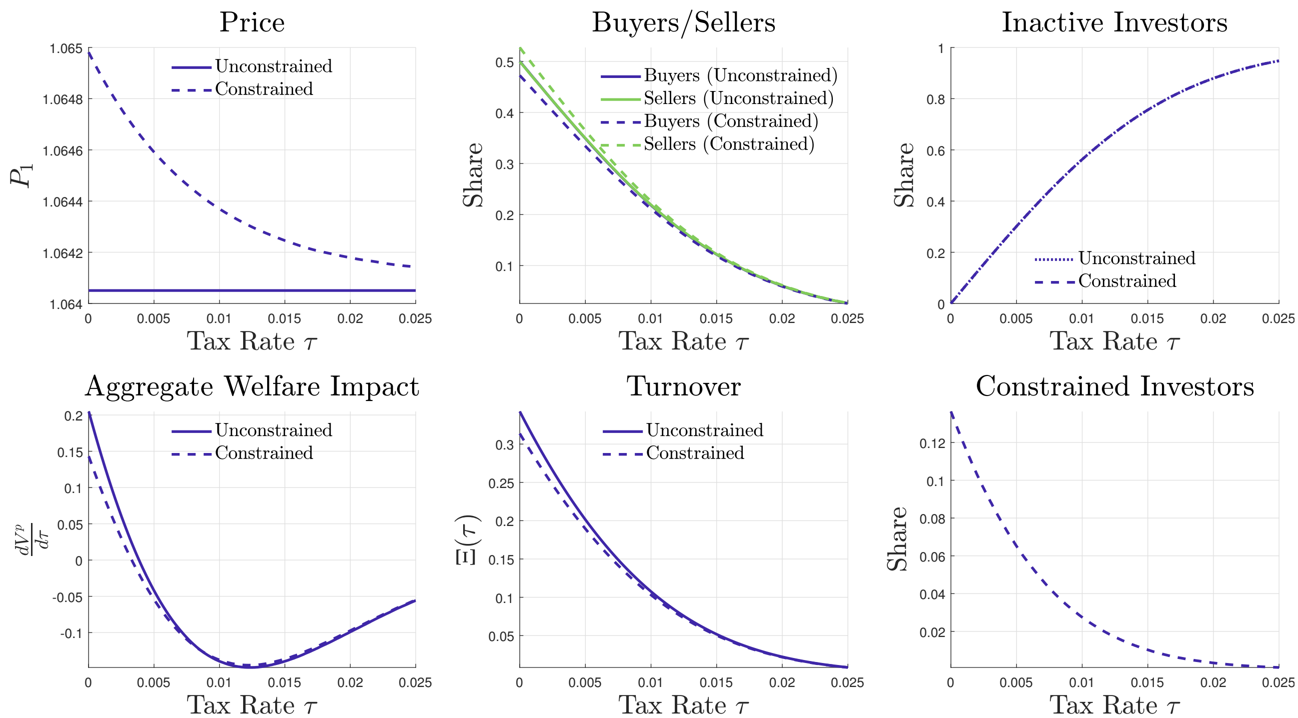


Figure A.14: Short-sale constraints

Note: The top left plot in Figure A.14 shows the equilibrium price P_1 as a function of the tax rate. The top middle plot shows the share of buyers and sellers in an economy that features short-sale constraints and in an economy in which all investors potentially face short-sale constraints. The top right plot shows the share of inactive investors in the economy as a function of the tax rate. The bottom left plot shows the aggregate marginal welfare impact of a tax change. The bottom middle plot shows aggregate turnover. The bottom right plots shows the share of investors of each group who are net buyers of the risky asset as a function of the tax rate. This unconstrained model of this simulation is calibrated to match the same three targets used in the quantitative assessment in Section 5. In particular, it features a 0.3 share of non-fundamental trading volume, a 1.5% risk premium, and an aggregate laissez-faire turnover of 34%. The constrained results introduce a short-sale constraint in that calibration that satisfies Assumption [G], assuming $\mu_d = 1.08$, $\sigma_d = 0.0075$, and $\rho = 0$. The optimal tax is $\tau = 0.39\%$ in the case without constraints and $\tau = 0.33\%$ in the case in which investors face short-sale constraints. These plots show the outcome of a simulation of the model with $N = 100,000$ investors.

When calculating welfare, the planner takes into account that investors must incur these costs when trading — this is the natural constrained efficient benchmark. The optimal tax formula remains unchanged when investors face transaction costs, as long as these trading costs represent exclusively a compensation for the use of economic resources.

Proposition 9. (Pre-existing trading costs) *When investors face trading costs as specified in Equation* (78)*, the optimal financial transaction tax τ^* satisfies exactly the same expression as in Equation* (12)*.*

Proof. Given investors' optimal portfolios, stated in the main text, it is straightforward to derive the equilibrium price, which is given by

$$P_1 = \frac{\int_{i\in\mathcal{T}} \frac{\mathbb{E}_i[D]-A_i\mathbb{C}ov[M_{2i},D]-\operatorname{sgn}(\Delta X_{1i})\eta+\alpha X_{0i}}{A_i\mathbb{V}ar[D]+\alpha} dF(i) - \int_{i\in\mathcal{T}} X_{0i} dF(i)}{\int_{i\in\mathcal{T}} \frac{(1+\operatorname{sgn}(\Delta X_{1i})(\tau+\varsigma))}{A_i\mathbb{V}ar[D]+\alpha} dF(i)}.$$

The certainty equivalent for an investor i from the planner's perspective is given by

$$V_i^p = \left(\mathbb{E}_p[D] - A_i\mathbb{C}ov[M_{2i}, D] - P_1\right) X_{1i} + P_1 X_{0i} - |\Delta X_{1i}| P_1\varsigma - \frac{\alpha}{2}(\Delta X_{1i})^2 - \frac{A_i}{2}\mathbb{V}ar[D](X_{1i})^2 + \tilde{T}_{1i}.$$

Note that only the resources corresponding to the transaction tax are rebated back to investors. All resources devoted to transaction costs are a compensation for the use of resources, so the planner does not have to account for them explicitly, since they form part of a zero profit condition. Hence, the marginal change in welfare for an investor i is given by

$$\frac{dV_i^p}{d\tau} = \begin{array}{l} \left(\mathbb{E}_p[D] - A_i\mathbb{C}ov[M_{2i}, D] - P_1 - \operatorname{sgn}(\Delta X_{1i}) P_1\varsigma - \eta\operatorname{sgn}(\Delta X_{1i})\right)\frac{dX_{1i}}{d\tau} \\ -\left(\alpha\Delta X_{1i} + A\mathbb{V}ar[D] X_{1i}\right)\frac{dX_{1i}}{d\tau} - \Delta X_{1i}\frac{dP_1}{d\tau} + \frac{d\tilde{T}_{1i}}{d\tau}. \end{array}$$

Substituting investors' first order conditions and exploiting market clearing, we can express $\frac{dV^p}{d\tau}$ as follows

$$\frac{dV^p}{d\tau} = \int \left[\mathbb{E}_p[D] - \mathbb{E}_i[D] + \operatorname{sgn}(\Delta X_{1i}) P_1\tau\right]\frac{dX_{1i}}{d\tau} dF(i).$$

Hence, the optimal tax has the same expression as in Proposition 1. ☐

The intuition behind Proposition 9 is similar to the baseline case. An envelope condition eliminates any term regarding transaction costs from $\frac{dV^p}{d\tau}$, because the planner must also face such costs, so the optimal tax looks identical to the one in the baseline model. This relies on the assumption that any economic profits made by the intermediaries who receive the transaction costs are zero — there cannot be economic rents. Proposition 9 does not imply that two overlapping authorities with taxation power would both impose the same τ^* twice. Assume for simplicity that they set taxes sequentially. The first authority would set the optimal tax according to Proposition 1, while the second authority, internalizing that the pre-existing tax is a mere transfer and does not correspond to a compensation for costs of trading, would set a zero tax. Alternatively, τ^* would characterize the sum of both taxes.

This result has further implications. First, although the optimal tax formula does not vary, an economy with transaction costs has less trade in equilibrium than one without transaction costs. Depending on whether this reduction in trading is of the fundamental type or not, the optimal tax may be larger or smaller. Transaction costs affect the optimal tax through changes in the identity of the marginal investors. Second, the mere existence of transaction costs does not provide a new rationale for further discouraging non-fundamental trading. Welfare losses must be traced back to wedges derived from portfolio distortions. Third, if transaction costs, that is, ς, η, and α, were endogenously functions of τ, as in richer models of the market microstructure, the planner would have to take into account those effects when solving for optimal taxes.[45] For instance, if a transaction tax endogenously increases trading costs, the optimal tax may be very small. However, if endogenously determined transaction costs are efficiently determined, the envelope theorem would still apply, leaving Proposition 9 unchanged.

Remark. (Trading costs instead risk-sharing distortions) In the model considered in this paper, the planner perceives that trading on belief differences is costly because it distorts investors risk-sharing decisions. In a model in which investors are risk-neutral but feature physical or technological costs of trading, a planner would have an identical rationale to discourage trading. For instance, if we assume in this Section that investors face a quadratic trading but $A_i = 0$, $\forall i$, it is possible to effectively re-derive the same expressions for welfare impact and optimal taxes as in the body of the paper.

[45] The Walrasian approach of this paper does not capture market microstructure effects. There is scope for understanding how transaction taxes affect market making and liquidity provision in greater detail, introducing, for instance, imperfectly competitive investors, search, or network frictions. The results of this paper would still be present regardless of the specific trading microstructure.

F.3 Multiple risky assets

Environment The results of the baseline model extend naturally to an environment with multiple assets. Now there are J risky assets in fixed supply, in addition to the risk-free asset. The $J \times 1$ vectors of total shares, equilibrium prices and dividend payments are respectively denoted by $\mathbf{q}$, $\mathbf{p}$, and $\mathbf{d}$.[46] Every purchase or sale of a risky asset faces an identical linear transaction tax τ. This is a further restriction on the planner's problem, since belief disagreements can vary across different assets, but the tax must be constant. Allowing for different taxes for different (groups of) assets is conceptually straightforward, following the logic of Section F.4.

The distribution of dividends $\mathbf{d}$ paid by the risky assets is a multivariate normal with a given mean and variance-covariance matrix $\mathbb{V}ar[\mathbf{d}]$. All investors agree about the variance, but an investor i believes that the mean of $\mathbf{d}$ is $\mathbb{E}_i[\mathbf{d}]$. We can thus write:

$$\mathbf{d} \sim_i N\left(\mathbb{E}_i[\mathbf{d}], \mathbb{V}ar[\mathbf{d}]\right),$$

where risk aversion A_i, and the vectors of initial asset holdings $\mathbf{x}_{0i}$, hedging needs $\mathbb{C}ov[M_{2i}, \mathbf{d}]$ and beliefs $\mathbb{E}_i[\mathbf{d}]$ are arbitrary across the distribution of investors. The wealth at $t = 2$ of an investor i is thus given by:

$$W_{2i} = M_{2i} + \mathbf{x}_{1i}'\mathbf{d} + \mathbf{x}_{0i}'\mathbf{p} - \mathbf{x}_{1i}'\mathbf{p} - \left|\mathbf{x}_{1i}' - \mathbf{x}_{0i}'\right| \mathbf{p}\tau + T_{1i}.$$

Results The first order condition (79) characterizes the solution of this problem for the set of assets traded:

$$\mathbf{x}_{1i} = \left(A_i \mathbb{V}ar[\mathbf{d}]\right)^{-1} \left(\mathbb{E}_i[\mathbf{d}] - A_i \mathbb{C}ov[M_{2i}, \mathbf{d}] - \mathbf{p} - \hat{\mathbf{p}}_i \tau\right), \tag{79}$$

where $\hat{\mathbf{p}}_i$ is a $J \times 1$ vector where row j is given by $\operatorname{sgn}(\Delta X_{1ij}) p_j$ and p_j denotes the price of asset j. If an asset j is not traded by an investors i, then $X_{1ij} = X_{0ij}$. If asset returns are independent, the portfolio allocation to every asset can be determined in isolation. Equilibrium prices are the natural generalization of the baseline model.

Proposition 10. (Multiple risky assets) *The optimal financial transaction tax τ^* when investors can trade J risky assets is given by*

$$\tau^* = \sum_{j=1}^{J} \omega_j \tau_j^*, \tag{80}$$

with weights ω_j and individual-asset taxes τ_j^* given by $\omega_j \equiv \frac{p_j \int \operatorname{sgn}(\Delta X_{1ij}) \frac{dX_{1ij}}{d\tau} dF(i)}{\sum_{j=1}^{J} p_j \int \operatorname{sgn}(\Delta X_{1ij}) \frac{dX_{1ij}}{d\tau} dF(i)}$ *and* $\tau_j^* \equiv \frac{\int \frac{\mathbb{E}_i[D_j]}{p_j} \frac{dX_{1ij}}{d\tau} dF(i)}{\int \operatorname{sgn}(\Delta X_{1ij}) \frac{dX_{1ij}}{d\tau} dF(i)}$.

Proof. After eliminating terms that do not affect the maximization problem, investors solve

$$\max_{\mathbf{x}_{1i}} \mathbf{x}_{1i}' \left(\mathbb{E}_i[\mathbf{d}] - A_i \mathbb{C}ov[M_{2i}, \mathbf{d}] - \mathbf{p}\right) - \left|\mathbf{x}_{1i}' - \mathbf{x}_{0i}'\right| \mathbf{p}\tau - \frac{A_i}{2} \mathbf{x}_{1i}' \mathbb{V}ar[\mathbf{d}] \mathbf{x}_{1i}.$$

Where I use $|\mathbf{x}_{1i}' - \mathbf{x}_{0i}'|$ to denote the vector of absolute values of the difference between both vectors. This problem is well-behaved, so the first order condition fully characterizes investors' optimal portfolios as long as they trade a given asset j

$$\mathbf{x}_{1i} = \left(A_i \mathbb{V}ar[\mathbf{d}]\right)^{-1} \left(\mathbb{E}_i[\mathbf{d}] - A_i \mathbb{C}ov[M_{2i}, \mathbf{d}] - \mathbf{p} - \hat{\mathbf{p}}_i \tau\right),$$

where $\hat{\mathbf{p}}_i$ is a $J \times 1$ vector where a given row j is given by $\operatorname{sgn}(\Delta X_{1ij}) p_j$. If an asset j is not traded by an investor i, then $X_{1ij} = X_{0ij}$. The inaction regions are defined analogously to the one asset case. Note that there exists a way to write optimal portfolio choices only with matrix operations; however, the notation turns out to be more cumbersome. The equilibrium price vector is given by

$$\mathbf{p} \odot \int \left(1 + \frac{\mathbf{s}_i}{A_i}\tau\right) dF(i) = \int \frac{\mathbb{E}_i[\mathbf{d}]}{A_i} dF(i) - \int \left(\mathbb{C}ov[M_{2i}, \mathbf{d}] + \mathbb{V}ar[\mathbf{d}] \mathbf{x}_{0i}\right) dF(i),$$

where I denote the element-by-element multiplication (Hadamard product) as $y \odot z$ and use $\mathbf{s}_i$ to denote a $J \times 1$ vector given by $\operatorname{sgn}(\Delta X_{1ij})$.

$$\frac{dV^p}{d\tau} = \int \left(\mathbb{E}[\mathbf{d}] - \mathbb{E}_i[\mathbf{d}] + \hat{\mathbf{p}}_i \tau\right)' \frac{d\mathbf{x}_{1i}}{d\tau} dF(i).$$

The marginal effect of varying taxes in social welfare is given by

$$\frac{dV^p}{d\tau} = \int \left[\left(\mathbb{E}[\mathbf{d}] - \mathbb{E}_i[\mathbf{d}] + \hat{\mathbf{p}}_i \tau\right)' \frac{d\mathbf{x}_{1i}}{d\tau} - \left(\mathbf{x}_{1i} - \mathbf{x}_{0i}\right)' \frac{d\mathbf{p}}{d\tau}\right] dF(i).$$

[46] I use bold lower-case letters to denote vectors but, for consistency, I keep the upper-case notation for holdings of a single asset.

This is a generalization of the one asset case. We can write in product notation

$$\int \sum_{j=1}^{J} \left(-\mathbb{E}_i\left[D_j\right] + \operatorname{sgn}\left(\Delta X_{1ij}\right) p_j \tau\right) \frac{dX_{1ij}}{d\tau} dF\left(i\right) = 0.$$

So the optimal tax becomes

$$\tau^* = \frac{\sum_{j=1}^{J} \int \mathbb{E}_i\left[D_j\right] \frac{dX_{1ij}}{d\tau} dF\left(i\right)}{\sum_{j=1}^{J} \int \operatorname{sgn}\left(\Delta X_{1ij}\right) p_j \frac{dX_{1ij}}{d\tau} dF\left(i\right)},$$

which can be rewritten as

$$\tau^* = \frac{\sum_{j=1}^{J} \int \operatorname{sgn}\left(\Delta X_{1ij}\right) p_j \frac{dX_{1ij}}{d\tau} dF\left(i\right) \tau_j^*}{\sum_{j=1}^{J} \int \operatorname{sgn}\left(\Delta X_{1ij}\right) p_j \frac{dX_{1ij}}{d\tau} dF\left(i\right)},$$

where $\tau_j^* = \frac{\int \frac{\mathbb{E}_i\left[D_j\right]}{p_j} \frac{dX_{1ij}}{d\tau} dF(i)}{\int \operatorname{sgn}\left(\Delta X_{1ij}\right) \frac{dX_{1ij}}{d\tau} dF(i)}$. And by defining weights $\omega_j = \frac{\int \operatorname{sgn}\left(\Delta X_{1ij}\right) p_j \frac{dX_{1ij}}{d\tau} dF(i)}{\sum_{j=1}^{J} \int \operatorname{sgn}\left(\Delta X_{1ij}\right) p_j \frac{dX_{1ij}}{d\tau} dF(i)}$, we recover Equation (80). □

The formula for τ_j^* is identical to the one in an economy with a single risky asset. The optimal tax in a model with J risky assets is simply a weighted average of all τ_j^*. The weights are determined by the relative marginal changes in (dollar) volume. Those assets whose volume responds more aggressively to tax changes carry higher weights when determining the optimal tax and vice versa.

F.4 Asymmetric taxes/Multiple tax instruments

In the baseline model, the only instrument available to the planner is a single linear financial transaction tax, which applies symmetrically to all investors. However, the planner could set different (linear) taxes for buyers and sellers.[47] Or, at least theoretically, even investor-specific taxes. In general, more sophisticated policy instruments bring the outcome of the planner's problem closer to the first-best, at the cost of increasing informational requirements.

Asymmetric taxes on buyers versus sellers Assume now that buyers pay a linear tax τ_B in the dollar volume of the transaction while sellers pay τ_S. Hence, the total tax liability is given by $\left(\tau_B + \tau_S\right) P_1 \left|\Delta X_{1i}\right|$. Outside of the inaction region, the optimal portfolio demand is given by

$$X_{1i} = \frac{\mathbb{E}_i\left[D\right] - A_i \mathbb{C}ov\left[M_{2i}, D\right] - P_1\left(1 + \mathbb{I}\left[\Delta X_{1i} > 0\right]\tau_B + \mathbb{I}\left[\Delta X_{1i} < 0\right]\tau_S\right)}{A_i \mathbb{V}ar\left[D\right]},$$

where $\mathbb{I}\left[\cdot\right]$ denotes the indicator function. This expression differs from (5) in that buyers now face a different tax than sellers. The equilibrium price is a natural extension of the one in the baseline model.

Proposition 11. (Asymmetric taxes on buyers versus sellers) *The pair of optimal financial transaction taxes for buyers and sellers, τ_B^* and τ_S^*, is characterized by the solution of the following system of non-linear equations:*

$$\tau_B^* + \tau_S^* = \frac{\int \frac{\mathbb{E}_i[D]}{P_1} \frac{dX_{1i}}{d\tau_B} dF\left(i\right)}{\int_{\mathcal{B}} \frac{dX_{1i}}{d\tau_B} dF\left(i\right)}, \qquad \tau_B^* + \tau_S^* = \frac{\int \frac{\mathbb{E}_i[D]}{P_1} \frac{dX_{1i}}{d\tau_S} dF\left(i\right)}{\int_{\mathcal{B}} \frac{dX_{1i}}{d\tau_S} dF\left(i\right)}. \tag{81}$$

Proof. The budget/wealth accumulation constraint for an investor in this case can be expressed as:

$$W_{2i} = M_{2i} + X_{1i} D + X_{0i} P_1 - X_{1i} P_1 - \tau_B P_1 \left|\Delta X_{1i}\right|_+ - \tau_S P_1 \left|\Delta X_{1i}\right|_- + T_{1i}.$$

The first order condition becomes

$$X_{1i} = \frac{\mathbb{E}_i\left[D\right] - A_i \mathbb{C}ov\left[M_{2i}, D\right] - P_1\left(1 + \mathbb{I}\left[\Delta X_{1i} > 0\right]\tau_B + \mathbb{I}\left[\Delta X_{1i} < 0\right]\tau_S\right)}{A_i \mathbb{V}ar\left[D\right]}.$$

With an equilibrium price given by

$$P_1 = \frac{\int_{i \in \mathcal{T}} \left(\frac{\mathbb{E}_i[D] - A_i \mathbb{C}ov[M_{2i}, D]}{A_i} - \mathbb{V}ar\left[D\right] X_{0i}\right) dF\left(i\right)}{\int_{i \in \mathcal{T}} \frac{1}{A_i} + \tau_B \int_{i \in \mathcal{B}} \frac{1}{A_i} - \tau_S \int_{i \in \mathcal{S}} \frac{1}{A_i} dF\left(i\right)}.$$

[47] Note that when investors have zero initial asset holdings of the risky asset, by setting an arbitrarily large tax on sellers, the planner is able to effectively impose a short-sale constraint.

In this case, we can write: $X_{1i}(\tau_i, P_1(\{\tau_j\}))$, where $\{\tau_j\}$ denotes a vector of taxes. This implies that $\frac{dX_{1i}}{d\tau_j} = \frac{\partial X_{1i}}{\partial \tau_j} + \frac{\partial X_{1i}}{\partial P_1}\frac{dP_1}{d\tau_j}$. The change in social welfare for an investor i when varying a tax τ_j, from a planner's perspective, is given by

$$\begin{aligned}\frac{dV_i^p}{d\tau_j} &= \left(\mathbb{E}_p[D] - A_i \mathbb{C}ov[M_{2i}, D] - P_1 - A_i X_{1i} \mathbb{V}ar[D]\right)\frac{dX_{1i}}{d\tau_j} - \Delta X_{1i}\frac{dP_1}{d\tau_j} + \frac{d\tilde{T}_{1i}}{d\tau_j} \\ &= \left(\mathbb{E}_p[D] - \mathbb{E}_i[D] + P_1\left(\mathbb{I}[\Delta X_{1i} > 0]\tau_B - \mathbb{I}[\Delta X_{1i} < 0]\tau_S\right)\right)\frac{dX_{1i}}{d\tau_j} - \Delta X_{1i}\frac{dP_1}{d\tau_j} + \frac{d\tilde{T}_{1i}}{d\tau_j}.\end{aligned}$$

The marginal aggregate welfare change of a tax change is given by

$$\frac{dV^p}{d\tau_j} = \int \left(\mathbb{E}_p[D] - \mathbb{E}_i[D] + P_1\left(\mathbb{I}[\Delta X_{1i} > 0]\tau_B - \mathbb{I}[\Delta X_{1i} < 0]\tau_S\right)\right)\frac{dX_{1i}}{d\tau_j} dF(i),$$

which is a generalization of Lemma 2. Under the usual differentiability and convexity assumptions, the optimal tax is characterized by $\frac{dV}{d\tau_j} = 0$, $\forall j$. This yields the following system of equation in the vector of taxes τ_B and τ_S

$$0 = \int \left(\mathbb{E}_p[D] - \mathbb{E}_i[D] + P_1 \operatorname{sgn}(\Delta X_{1i})\tau_i\right)\frac{dX_{1i}}{d\tau_j} dF(i), \ \forall j.$$

We can write

$$\frac{dV^p}{d\tau_j} = \int \left(\mathbb{E}_p[D] - \mathbb{E}_i[D]\right)\frac{dX_{1i}}{d\tau_j} dF(i) + P_1\left(\tau_B \int_{i\in\mathcal{B}} \frac{dX_{1i}}{d\tau_j} dF(i) - \tau_S \int_{i\in\mathcal{S}} \frac{dX_{1i}}{d\tau_j} dF(i)\right).$$

Using market clearing, we can find

$$\frac{dV^p}{d\tau_j} = -\int \mathbb{E}_i[D]\frac{dX_{1i}}{d\tau_j} dF(i) + P_1(\tau_B + \tau_S)\int_{i\in\mathcal{B}} \frac{dX_{1i}}{d\tau_j} dF(i),$$

which allows us to solve for $\tau_B + \tau_S$ as follows

$$\tau_B + \tau_S = \frac{\int \mathbb{E}_i[D]\frac{dX_{1i}}{d\tau_j} dF(i)}{P_1 \int_{i\in\mathcal{B}} \frac{dX_{1i}}{d\tau_j} dF(i)}, \ \forall_j.$$

In general this gives a system of non-linear equations in $\tau_B + \tau_S$. When there are two investors, the two equations become collinear, because of market clearing

$$\int \mathbb{E}_i[D]\frac{dX_{1i}}{d\tau_j} dF(i) = \left(\mathbb{E}_{\mathcal{B}}[D] - \mathbb{E}_{\mathcal{S}}[D]\right)\int_{i\in\mathcal{B}} \frac{dX_{1i}}{d\tau_j} dF(i).$$

Hence, in that case, only the sum of taxes is pinned down, and must satisfy

$$\tau_B + \tau_S = \frac{\mathbb{E}_{\mathcal{B}}[D] - \mathbb{E}_{\mathcal{S}}[D]}{P_1}.$$

□

The economic forces that shape the optimal values for τ_B^* and τ_S^* are the same as in the baseline model. Once again, the planner's belief is irrelevant for the optimal policy, which shows that that results are not sensitive to the use of more sophisticated policy instruments. Intuitively, the change in portfolio allocations induced by a marginal change in any instrument must cancel out in the aggregate. Equation (81) provides intuition for why all taxes in the baseline model are divided by 2; in that case, there exists a single optimality condition and $2\tau^* = \tau_B^* + \tau_S^*$. As long as there are more than two investors, this system has at least a solution. When there are two investors, the system is indeterminate and only the sum $\tau_B^* + \tau_S^*$ is pinned down. In that case, $\tau_B^* + \tau_S^* = \frac{\mathbb{E}_{\mathcal{B}}[D] - \mathbb{E}_{\mathcal{S}}[D]}{P_1}$.

Individual taxes/First-best Assume now that the planner can set investor specific taxes. This is an interesting theoretical benchmark, despite being unrealistic. For simplicity, I now assume that there is a finite number N of (types of) investors in the economy.

Proposition 12. (Individual taxes/First-best)

a) The first-best can be implemented with a set of investor specific taxes given by

$$\tau_i^* = \operatorname{sgn}(\Delta X_{1i})\frac{\mathbb{E}_i[D] - \Upsilon}{P_1}, \forall i = 1, \ldots, N, \tag{82}$$

where Υ is any real number; a natural choice for Υ is $\mathbb{E}_p[D]$.

b) The planner only needs $N-1$ taxes to implement the first-best in an economy with N investors.

Proof. a) In the case with I taxes and I investors, the first order conditions for the planner become

$$\frac{dV^p}{d\tau_j} = \sum_i \left(-\mathbb{E}_i\left[D\right] + P_1 \operatorname{sgn}\left(\Delta X_{1i}\right)\tau_i\right) \frac{dX_{1i}}{d\tau_j} F\left(i\right) = 0,\ \forall j.$$

This system of equations characterizes the set of optimal taxes. Note that one solution to this system is given by

$$-\mathbb{E}_i\left[D\right] + P_1 \operatorname{sgn}\left(\Delta X_{1i}\right)\tau_i = -F.$$

Where F is an arbitrary real number. Rearranging this expression we can find Equation (82).

b) Starting from the system of equations which characterizes the optimal set of taxes, we can write, using market clearing $F\left(j\right)\frac{dX_{1i}}{d\tau_j} + \sum_{i\neq j}\frac{dX_{1i}}{d\tau_j}F\left(i\right) = 0$, the following set of equations:

$$\sum_{i\neq j}\left(\mathbb{E}_j\left[D\right] - \mathbb{E}_i\left[D\right]\right)\frac{dX_{1i}}{d\tau_j}F\left(i\right) + P_1\left(\begin{array}{c} -\operatorname{sgn}\left(\Delta X_{1j}\right)\tau_j \sum_{i\neq j}\frac{dX_{1i}}{d\tau_j}F\left(i\right) \\ +\sum_i\left(\operatorname{sgn}\left(\Delta X_{1i}\right)\tau_i\frac{dX_{1i}}{d\tau_j}\right)F\left(i\right)\end{array}\right) = 0.$$

For all equations but for the one with respect to tax j. To show that this system only depends on $N-1$ taxes, we simply need to show that all $\frac{dX_{1i}}{d\tau_j}$ do not depend on the tax τ_j. Note that $\frac{dX_{1i}}{d\tau_j} = \frac{\partial X_{1i}}{\partial\tau_j} + \frac{\partial X_{1i}}{\partial P_1}\frac{dP_1}{d\tau_j}$. But when $i\neq j$ then $\frac{dX_{1i}}{d\tau_j}$ only depends on $\frac{dP_1}{d\tau_j}$ because $\frac{\partial X_{1i}}{\partial\tau_j}$ equals zero and $\frac{\partial X_{1i}}{\partial P_1}$ does not depend on τ_j. We just need to show that $\frac{dP_1}{d\tau_j}$ can be expressed as a function of all other taxes but τ_j. This can be easily shown combining the expressions used to show Lemma 1 with market clearing conditions. □

Proposition 12a) follows standard Pigouvian logic. The planner sets optimal individual taxes so that investors' portfolio choices replicate those of an economy with homogeneous beliefs. Note that the planner can use any belief Υ to implement the first-best allocation, as long as it is the same for all investors. In a production economy, the natural choice would be $\Upsilon = \mathbb{E}_p\left[D\right]$. Finally, because P_1 is a function of all taxes, Equation (82) also defines a system of non-linear equations.

Proposition 12b) shows that the first-best could be implemented with $N-1$ taxes. This occurs because the risky asset is in fixed supply. The logic behind this result is similar to Walras' law. For instance, when $N = 2$, a single tax which modifies directly the allocation of one of the investors necessarily changes the allocation of the other one through market clearing. Finally, note that Propositions 11 and 12 highlight that the optimal set of taxes in each case are independent of the planner's beliefs, as in Proposition 1.

F.5 Production

The results derived so far rely on the assumption that assets are in fixed supply. I now study how optimal policies vary when financial markets determine production by influencing the intertemporal investment decision in a standard price-taking environment — this is the role explored in classic q-theory models.

Environment There is a new group of agents in the economy who were not present in the baseline model: identical competitive producers in unit measure. Producers are indexed by k and maximize well-behaved time separable expected utility, with flow utility given by $U_k\left(\cdot\right)$. They have exclusive access to a technology $\Phi\left(S_{1k}\right)$, which allows them to issue or dispose of S_{1k} shares of the risky asset at date 1.[48] I refer to S_{1k}, which can be negative, as investment. The function $\Phi\left(\cdot\right)$ is increasing and strictly convex; that is, $\Phi'\left(\cdot\right) > 0$, $\Phi''\left(\cdot\right) > 0$. To ease the exposition, I assume throughout that $\Phi\left(S_{1k}\right) = \gamma_1\left|S_{1k}\right| + \frac{\gamma_2}{2}\left|S_{1k}\right|^2$, with $\gamma_1,\gamma_2 > 0$. Producers are initially endowed with E_{1k} units of consumption good (dollars) and can only borrow or save in the risk-free asset at a (gross) rate $R = 1$. Their endowment E_{2k} at date 2 is stochastic and follows an arbitrary distribution.

To avoid distortions in primary markets, the planner does not tax the issuance of new shares. Importantly, market clearing is now given by $\int X_{1i}dF\left(i\right) = Q + S_{1k}$. Total output at date 2 in this economy is endogenous and given by $D\left(Q + S_{1k}\right)$.

[48] A "tree" analogy can be helpful here. Assume that a share of the risky asset (i.e., a tree) entitles the owner to a dividend payment D (fruit). Producers can plant new trees or chop them at a cost $\Phi\left(S_{1k}\right)$, which they sell or buy at a price P_1. Producers would be willing to create trees until the marginal cost of producing a new tree/chopping and old tree $\Phi'\left(S_{1k}\right)$ equals the marginal benefit of selling/buying P_1. For consistency, any normalization concerning Q must also normalize $\Phi\left(\cdot\right)$.

Positive results Producers thus maximize

$$\max_{C_{1k},C_{2k},S_{1k},Y_k} U_k\left(C_{1k}\right)+\mathbb{E}\left[U_k\left(C_{2k}\right)\right],$$

with budget constraints $Y_k + C_{1k} = E_{1k} + P_1^s S_{1k} - \Phi\left(S_{1k}\right)$ and $C_{2k} = E_{2k} + Y_k$, where Y_k denotes the amount saved in the risk-free asset and P_1^s denotes the price faced by producers — the superscript s stands for supply. The optimality conditions for producers are given by

$$U_k'\left(C_{1k}\right)=\mathbb{E}\left[U_k'\left(C_{2k}\right)\right] \quad \text{and} \quad P_1^s=\Phi'\left(S_{1k}\right).$$

The first condition is a standard Euler condition for the risk-free asset. The second condition provides a supply curve for the number of shares. Combining this supply curve with the portfolio choices of investors, generates the following equilibrium price:

$$P_1=\left(1-\alpha\right)\gamma_1+\alpha P_1^e,$$

where the weight $\alpha \in [0,1]$ — defined in the Appendix — is higher when the adjustment cost is very concave (γ_2 is large) and P_1^e is essentially the same expression for the price that would prevail in an exchange economy, which is given in Equation (6). Intuitively, the equilibrium price is a weighted average of the exchange economy price and γ_1, which is the replacement cost of the risky asset with linear adjustments costs.

Allowing for production does not affect those positive properties of the model that matter for the determination of the optimal tax. An increase in the transaction tax can increase, reduce, or keep equilibrium prices (and investment) constant, but all buyers buy less and all sellers sell less.

Normative results The marginal welfare impact of a tax change in producers' welfare from the planner's perspective, when measured in dollars, is given by

$$\begin{aligned}\frac{\frac{dV_k^p}{d\tau}}{U_k'\left(C_{1k}\right)}&=\left[\frac{dP_1}{d\tau}S_{1k}+\left[P_1-\Phi'\left(S_{1k}\right)\right]\frac{dS_{1k}}{d\tau}-\frac{dY_k}{d\tau}\right]+\mathbb{E}\left[\frac{U_k'\left(C_{2k}\right)}{U_k'\left(C_{1k}\right)}\right]\frac{dY_k}{d\tau}\\&=\mathbb{E}\left[\frac{U_k'\left(C_{2k}\right)}{U_k'\left(C_{1k}\right)}\right]\frac{dP_1}{d\tau}S_{1k},\end{aligned}$$

where the second line follows by substituting producers' optimality conditions. Intuitively, because producers do not pay taxes and invest optimally given prices, a marginal tax change only modifies their welfare through the distributive price effects on the shares they issue/repurchase. When P_1 is high, producers enjoy a better deal selling shares than when P_1 is low. The envelope theorem eliminates from $\frac{dV_k^p}{d\tau}$ the direct effects caused by changes in producers portfolio or investment choices.

Proposition 13 assumes that the planner accounts for producers' certainty equivalents, and characterizes the optimal tax.

Proposition 13. (Optimal tax in production economies) *The optimal tax in a production economy is given by*

$$\tau^*=\frac{\int\left(\frac{\mathbb{E}_p[D]-\mathbb{E}_i[D]}{P_1}\right)\frac{dX_{1i}}{d\tau}dF\left(i\right)}{-\int \operatorname{sgn}\left(\Delta X_{1i}\right)\frac{dX_{1i}}{d\tau}dF\left(i\right)}=\left(1-\omega\right)\tau^*_{exchange}+\omega\tau^*_{production}, \tag{83}$$

where $\tau^*_{exchange}=\frac{\zeta(\tau)\mathbb{C}ov_{F,\mathcal{T}}\left[\frac{\mathbb{E}_i[D]}{P_1},\frac{dX_{1i}}{d\tau}\right]}{2\int_{i\in\mathcal{B}}\frac{dX_{1i}}{d\tau}dF(i)}$, $\tau^*_{production}=\frac{\mathbb{E}_p[D]-\mathbb{E}_{F,\mathcal{T}}[\mathbb{E}_i[D]]}{P_1}$ *and* $\omega<1$ *is given in the Appendix (*ω *is small in magnitude when* $\frac{dS_{1k}}{d\tau}\approx 0$ *and close to unity when* $\left|\frac{dS_{1k}}{d\tau}\right|$ *is large).* $\mathbb{E}_{F,\mathcal{T}}\left[\mathbb{E}_i\left[D\right]\right]$ *denotes the average belief in the population of active investors,* $\mathbb{C}ov_{F,\mathcal{T}}\left[\mathbb{E}_i\left[D\right],\frac{dX_{1i}}{d\tau}\right]$ *denotes a cross-sectional covariance among active investors and* $\zeta\left(\tau\right)\equiv\int_{i\in\mathcal{T}}dF\left(i\right)$ *is the share of active investors.*

Proof. The expression for the asset price in Equation (84) now yields the following demand curve for risky asset shares

$$P_1^e=\frac{\int_{i\in\mathcal{T}}\left(\frac{\mathbb{E}_i[D]}{\mathcal{A}_i}-A\left(\mathbb{C}ov\left[M_{2i},D\right]+\mathbb{V}ar\left[D\right]X_{0i}\right)\right)dF\left(i\right)-A\mathbb{V}ar\left[D\right]S_{1k}}{1+\tau\int_{i\in\mathcal{T}}\frac{\operatorname{sgn}(\Delta X_{1i})}{\mathcal{A}_i}dF\left(i\right)} \tag{84}$$

The demand by investors for the risky asset is identical to the baseline model. The equilibrium price is now determined by the intersection of Equation (84) and the supply curve, given by $P_1^s=\gamma_1+\gamma_2 S_{1k}$.

After writing the market clearing condition as $\int (X_{1i} - X_{0i})\, dF(i) = F(P_1)$, where $F(\cdot) = \Phi'^{-1}(\cdot)$ is an upward sloping function, we can derive $\frac{dP_1}{d\tau} = \frac{\int \frac{\partial X_{1i}}{\partial \tau} dF(i)}{F'(P_1) - \int \frac{\partial X_{1i}}{\partial P_1} dF(i)} = \frac{-P_1 \int \frac{\operatorname{sgn}(\Delta X_{1i})}{A_i \mathbb{V}ar[D]} dF(i)}{F'(P_1) + \int \frac{(1+\operatorname{sgn}(\Delta X_{1i})\tau)}{A_i \mathbb{V}ar[D]} dF(i)}$. $\frac{dP_1}{d\tau}$ can have any sign, depending on its numerator. We can write $\frac{dX_{1i}}{d\tau} = \frac{\partial X_{1i}}{\partial \tau}\varepsilon_i$, where ε_i, which is constant within buyers/sellers, can be expressed as $\varepsilon_i = 1 - (\operatorname{sgn}(\Delta X_{1i}) + \tau) \frac{1-H}{\frac{\mathbb{V}ar[D] F'(P_1)}{\int_{\mathcal{B}} \frac{1}{A_i} dF(i)} + 1 + \tau + H(1-\tau)}$ and $H \equiv \frac{\int_{i \in \mathcal{S}} \frac{1}{A_i} dF(i)}{\int_{i \in \mathcal{B}} \frac{1}{A_i} dF(i)} \in (0, \infty)$. It is easy to show that $\varepsilon_i > 0$, which proves the result.

The marginal change in social welfare is given by

$$\frac{dV^p}{d\tau} = \int \left[(\mathbb{E}_p[D] - \mathbb{E}_i[D] + \operatorname{sgn}(\Delta X_{1i}) P_1 \tau) \frac{dX_{1i}}{d\tau} - \Delta X_{1i} \frac{dP_1}{d\tau} \right] dF(i) + \frac{dP_1}{d\tau} S_{1k}$$

Using market clearing, the marginal change in social welfare can be expressed as follows

$$\frac{dV^p}{d\tau} = \int (\mathbb{E}_p[D] - \mathbb{E}_i[D] + \operatorname{sgn}(\Delta X_{1i}) P_1 \tau) \frac{dX_{1i}}{d\tau} dF(i)$$

Solving for τ^* in the previous expression yields Equation (83). We can re-write the numerator of the optimal tax as

$$\begin{aligned}
\int (\mathbb{E}_p[D] - \mathbb{E}_i[D]) \frac{dX_{1i}}{d\tau} dF(i) &= \zeta(\tau)\, \mathbb{E}_{F,\mathcal{T}} \left[(\mathbb{E}_p[D] - \mathbb{E}_i[D]) \frac{dX_{1i}}{d\tau} \right] \\
&= \zeta(\tau) \begin{pmatrix} \mathbb{C}ov_{F,\mathcal{T}} \left[\mathbb{E}_p[D] - \mathbb{E}_i[D], \frac{dX_{1i}}{d\tau} \right] \\ + \mathbb{E}_{F,\mathcal{T}} [\mathbb{E}_p[D] - \mathbb{E}_i[D]]\, \mathbb{E}_{F,\mathcal{T}} \left[\frac{dX_{1i}}{d\tau} \right] \end{pmatrix} \\
&= -\zeta(\tau)\, \mathbb{C}ov_{F,\mathcal{T}} \left[\mathbb{E}_i[D], \frac{dX_{1i}}{d\tau} \right] + (\mathbb{E}_p[D] - \mathbb{E}_{F,\mathcal{T}}[\mathbb{E}_i[D]]) \frac{dS_{1k}}{d\tau},
\end{aligned}$$

where we can define $\zeta(\tau) \equiv \int_{i \in \mathcal{T}} dF(i)$. This normalization by the number of active investors is necessary to use expectation and covariance operators. Using the fact that $\int_{i \in \mathcal{S}} \frac{dX_{1i}}{d\tau} dF(i) = \frac{dS_{1k}}{d\tau} - \int_{i \in \mathcal{B}} \frac{dX_{1i}}{d\tau} dF(i)$, the denominator in (83) can be expressed as $\int \operatorname{sgn}(\Delta X_{1i}) \frac{dX_{1i}}{d\tau} dF(i) = \int_{i \in \mathcal{B}} \frac{dX_{1i}}{d\tau} dF(i) - \int_{i \in \mathcal{S}} \frac{dX_{1i}}{d\tau} dF(i) = 2 \int_{i \in \mathcal{B}} \frac{dX_{1i}}{d\tau} dF(i) - \frac{dS_{1k}}{d\tau}$. By substituting and rearranging the previous two expressions in the optimal tax formula, we can write τ^*as follows

$$\tau^* = \underbrace{\frac{-2 \int_{i \in \mathcal{B}} \frac{dX_{1i}}{d\tau} dF(i)}{-2 \int_{i \in \mathcal{B}} \frac{dX_{1i}}{d\tau} dF(i) + \frac{dS_{1k}}{d\tau}}}_{\equiv 1-\omega} \underbrace{\frac{-\zeta(\tau)\, \mathbb{C}ov_{F,\mathcal{T}} \left[\frac{\mathbb{E}_i[D]}{P_1}, \frac{dX_{1i}}{d\tau} \right]}{-2 \int_{i \in \mathcal{B}} \frac{dX_{1i}}{d\tau} dF(i)}}_{\equiv \tau^*_{\text{exchange}}} + \underbrace{\frac{\frac{dS_{1k}}{d\tau}}{-2 \int_{i \in \mathcal{B}} \frac{dX_{1i}}{d\tau} dF(i) + \frac{dS_{1k}}{d\tau}}}_{\equiv \omega} \underbrace{\frac{\mathbb{E}_p[D] - \mathbb{E}_{F,\mathcal{T}}[\mathbb{E}_i[D]]}{P_1}}_{\equiv \tau^*_{\text{production}}}.$$

□

Corollary. (τ^* may depend on the planner's belief) *The optimal financial transaction tax in a production economy depends on the distribution of payoffs assumed by the planner. However, if the planner uses the average belief across investors, that is, $\mathbb{E}_p[D] = \mathbb{E}_{F,\mathcal{T}}[\mathbb{E}_i[D]]$ at the optimum, the optimal tax is identical to the one in the exchange economy and independent of the belief used by the planner.*[49]

The numerator in Equation (83), which evaluated at $\tau = 0$ determines the sign of the optimal tax can be decomposed in two terms

$$\int (\mathbb{E}_p[D] - \mathbb{E}_i[D]) \frac{dX_{1i}}{d\tau} dF(i) = -\zeta(\tau) \underbrace{\mathbb{C}ov_{F,\mathcal{T}} \left[\mathbb{E}_i[D], \frac{dX_{1i}}{d\tau} \right]}_{\text{Belief dispersion}} + \underbrace{(\mathbb{E}_p[D] - \mathbb{E}_{F,\mathcal{T}}[\mathbb{E}_i[D]]) \frac{dS_{1k}}{d\tau}}_{\substack{\text{Aggregate belief difference} \times \\ \text{Investment response}}}. \tag{85}$$

Because the second term in Equation (85) is in general non-zero when $\tau = 0$, we can say that belief distortions in production economies have an additional first-order effect on welfare. Again asset prices do not appear in optimal tax formulas, despite playing a role in determining allocations. All welfare losses must be traced back to distortions in "quantities", either in portfolio allocations, captured by $\frac{dX_{1i}}{d\tau}$, or in production decisions, captured by $\frac{dS_{1k}}{d\tau}$.

Intuitively, the optimal tax corrects two wedges created by heterogeneous beliefs. First, given an amount of aggregate risk, the optimal tax seeks to reduce the asset holding dispersion induced by disagreement — some investors are holding too much risk and some others too little risk. This is the same mechanism present in exchange economies. Second, as long

[49]The average belief may change if there are changes in the composition of marginal investors. For the irrelevance result to hold without further qualifications, the average belief for marginal investors must be invariant to the level of τ.

	Aggregate optimism	Aggregate pessimism
	$\mathbb{E}_F[\mathbb{E}_i[D]] > \mathbb{E}_p[D]$	$\mathbb{E}_F[\mathbb{E}_i[D]] < \mathbb{E}_p[D]$
$\int_{\mathcal{B}} \frac{1}{A_i} dF(i) > \int_{\mathcal{S}} \frac{1}{A_i} dF(i)$	$\omega\tau^*_{\text{production}} > 0$	$\omega\tau^*_{\text{production}} < 0$
$\int_{\mathcal{B}} \frac{1}{A_i} dF(i) < \int_{\mathcal{S}} \frac{1}{A_i} dF(i)$	$\omega\tau^*_{\text{production}} < 0$	$\omega\tau^*_{\text{production}} > 0$

Table 2: Sign of $\omega\tau^*_{\text{production}}$

as the average belief differs from the one used by the planner, the level of production in the economy is too high (low) when investors are on average too optimistic (pessimistic). This provides a second rationale for taxation. Intuitively, the investors in the economy hold too much aggregate risk when they are on average optimistic or too little when they are pessimistic.[50]

Let's describe the sign of $\omega\tau^*_{\text{production}}$ next. Belief dispersion is not sufficient anymore to pin down the sign of the optimal tax, which now also depends on whether $\mathbb{E}_p[D] - \mathbb{E}_F[\mathbb{E}_i[D]]$ and $\frac{dS_{1k}}{d\tau}$ have the same or opposite signs. Intuitively, if a marginal tax increase reduces (increases) investment at the margin when investors are too optimistic (pessimistic), a positive tax is welfare improving, and vice versa. Table 5 summarizes the conditions that determine the sign of the term associated with production.

Unlike in the exchange economy, in which the independence of beliefs justifies that the belief dispersion term is negative, it is not obvious whether we should expect $\omega\tau^*_{\text{production}}$ to be positive or negative. For investment to be reduced (increased) at the margin by a tax increase, it has to be the case that the (risk aversion adjusted) mass of buyers is larger (smaller) than the mass of sellers. In principle, the relation between the average belief distortion and the relative mass of buyers/sellers need not be linked, so the sign of $\omega\tau^*_{\text{production}}$ is theoretically ambiguous. Additional policy instruments, like short-sale of borrowing constraints, investment taxes, or active monetary policy can be used to target the production distortion induced by beliefs, allowing the transaction tax to be exclusively focused again on the dispersion of beliefs among investors.

Many informal discussions regarding the convenience of a transaction tax, following Tobin (1978), revolve around the notion that it would help reduce price volatility. Implicit in those discussions is the notion that high volatility is bad. The results in this section show that it is not price volatility, a variance, but whether investment (through prices) is lower when investors are optimistic and vice versa, a covariance, what captures the welfare consequences of a transaction tax in a production context.

The optimal tax can be expressed as a linear combination between the optimal tax in a (fictitious) exchange economy and the optimal tax in a (fictitious) production economy with a single investor with belief $\mathbb{E}_{F,\mathcal{T}}[\mathbb{E}_i[D]]$. The sensitivity of investment with respect to a tax change determines the relative importance of each term. Market clearing now implies that $\int \frac{dX_{1i}}{d\tau} dF(i) = \frac{dS_{1k}}{d\tau}$, which can take any positive or negative value. Hence, in production economies, the belief used by the planner to calculate welfare matters in general for the optimal policy. However, if the planner uses investors' average belief to calculate welfare, the belief used by the planner drops out of the optimal tax expression. Because of its importance, I state this result as a corollary of Proposition 13.

F.6 Tax on the number of shares

The baseline model assumes the tax is levied on the dollar value of a trade rather than on the number of shares traded to prevent investors from circumventing it by varying the effective number of shares traded through a reverse split. All results apply to taxes that depend on the number of shares with minor modifications.

When P_1 is exactly zero, a tax based on the dollar volume of the transaction is ineffective. However, a tax based on the number of shares traded $|\Delta X_{1i}|$ can be introduced to effectively tax the notional value of the contract. I extend here Proposition 1 to the case of taxes levied on the number of shares traded. In this case, the distinction between buyers and sellers is somewhat arbitrary, giving support to the idea that both sides of the market should face the same tax.

In the trade region, the optimal portfolio choice of an investor can be expressed as: $X_{1i} = \frac{\mathbb{E}_i[D] - A_i \mathbb{C}ov[M_{2i}, D] - P_1 - \text{sgn}(\Delta X_{1i})\tau}{A_i \mathbb{V}ar[D]}$. The equilibrium price becomes:

$$P_1 = \int_{i \in \mathcal{T}} \left(\frac{\mathbb{E}_i[D]}{\mathcal{A}_i} - A\left(\mathbb{C}ov[M_{2i}, D] - \mathbb{V}ar[D] X_{0i}\right) - \frac{\text{sgn}(\Delta X_{1i})}{\mathcal{A}_i}\tau \right) dF(i).$$

[50]If there were many produced risky assets, the welfare losses would capture the idea that belief distortions misallocate real investment across sectors in the economy. These results are available upon request.

The price correction is now additive rather than multiplicative. The value of $\frac{dV}{d\tau}$ corresponds to $\frac{dV}{d\tau} = \int \left[-\mathbb{E}_i\left[D\right] + \text{sgn}\left(\Delta X_{1i}\right)\tau\right] \frac{dX_{1i}}{d\tau} dF\left(i\right)$. The optimal tax now satisfies:

$$\tau^* = \frac{\int \mathbb{E}_i\left[D\right] \frac{dX_{1i}}{d\tau} dF\left(i\right)}{\int \text{sgn}\left(\Delta X_{1i}\right) \frac{dX_{1i}}{d\tau} dF\left(i\right)}.$$

This shows that optimal taxes in the paper are written in terms of returns because they are levied on the dollar value of the transaction. When they are levied on the number of shares, the dispersion in expected payoffs rather than the dispersion in expected returns becomes the welfare relevant variable.

F.7 Harberger calculation

The results derived so far rely on the assumption that the planner maximizes welfare using a single belief. However, it is straightforward to quantify the welfare loss induced by a tax increase assuming that all investors hold correct beliefs or that the planner assesses social welfare respecting individual beliefs. Under either of these assumptions, all trades are regarded as fundamental, so any tax induces a welfare loss. I derive a result analogous to Harberger (1964), whose triangle analysis can be traced back to Dupuit (1844).

Proposition 14. (Harberger (1964) revisited)

a) When investors hold identical beliefs or the planner respects individual beliefs when calculating social welfare, the marginal welfare loss generated by increasing the transaction tax at a level $\tilde{\tau}$, expressed as a money-metric (in dollars) at $t = 1$, is given by

$$\int \left.\frac{dV_i}{d\tau}\right|_{\tau=\tilde{\tau}} dF\left(i\right) = 2\tilde{\tau} P_1 \int_{i\in\mathcal{B}} \left.\frac{dX_{1i}}{d\tau}\right|_{\tau=\tilde{\tau}} dF\left(i\right) \leq 0, \tag{86}$$

where $i \in \mathcal{B}$ denotes that the integration is made only over the set of buyers and $\hat{V}_i$ denotes investors' certainty equivalents.

b) The marginal welfare loss of a small tax change around $\tau = 0$ can be approximated, using a second order Taylor expansion, by

$$\mathcal{L}\left(\tau\right) \equiv dV = \int \left. d\hat{V}_i\right|_{\tau=0} dF\left(i\right) \approx \tau^2 P_1 \int_{i\in\mathcal{B}} \left.\frac{dX_{1i}}{d\tau}\right|_{\tau=0} dF\left(i\right).$$

Proof. a) When there are no belief differences between investors and the planner or when the planner assesses social welfare respecting individual beliefs, we can write the marginal change in welfare as a money-metric (divided by investors' marginal utility) as

$$\left.\frac{dV_i}{d\tau}\right|_{\tau=\tilde{\tau}} = \text{sgn}\left(\Delta X_{1i}\right)\tilde{\tau} P_1 \frac{dX_{1i}}{d\tau} - \Delta X_{1i}\frac{dP_1}{d\tau}.$$

Adding up across all investors, and using the fact that $\int \text{sgn}\left(\Delta X_{1i}\right) \frac{dX_{1i}}{d\tau} dF\left(i\right) = 2\int_{i\in\mathcal{B}} \frac{dX_{1i}}{d\tau} dF\left(i\right)$, we then recover Equation (86).

b) The result in a) is an exact expression. However, we can write a second order approximation around $\tau = 0$ of the marginal change in social welfare. Note all terms corresponding to terms-of-trade cancel out after imposing market clearing, so I do not consider them. The first term of the Taylor expansion is given above. The derivative of the second term of the Taylor expansion is given by: $\text{sgn}\left(\Delta X_{1i}\right) P_1 \frac{dX_{1i}}{d\tau} + \text{sgn}\left(\Delta X_{1i}\right)\tau P_1 \frac{d^2 X_{1i}}{d\tau^2}$. Around $\tau = 0$, this becomes $\text{sgn}\left(\Delta X_{1i}\right) P_1 \frac{dX_{1i}}{d\tau}$. Hence, when $\tau = 0$ we can write:

$$\begin{aligned}\int \left. dV_i\right|_{\tau=0} dF\left(i\right) &\approx \int \text{sgn}\left(\Delta X_{1i}\right)\tau P_1 \left.\frac{dX_{1i}}{d\tau}\right|_{\tau=0} dF\left(i\right)\left(d\tau\right) \\ &\quad + \frac{1}{2}\int \left.\left(\text{sgn}\left(\Delta X_{1i}\right) P_1 \frac{dX_{1i}}{d\tau} + \text{sgn}\left(\Delta X_{1i}\right)\tau P_1 \frac{d^2 X_{1i}}{d\tau^2}\right)\right|_{\tau=0} dF\left(i\right)\left(d\tau\right)^2 \\ &= P_1 \tau^2 \int_{i\in\mathcal{B}} \left.\frac{dX_{1i}}{d\tau}\right|_{\tau=0} dF\left(i\right).\end{aligned}$$

□

This result provides a measure of welfare losses as a function of observables for any tax intervention. Given the money-metric correction, investors in this economy are willing to pay $\mathcal{L}\left(\tau\right)$ dollars to prevent a change in the tax rate. Note that this happens to correspond to the marginal change in revenue raised. Equation (86) derives an upper bound for the size of the welfare losses induced by taxation in the case in which all trades are deemed to be fundamental.

Equation (86) resembles the classic Harberger (1964) result about welfare losses in the context of commodity taxation.[51] However, the welfare loss in this case is given by twice the size of the tax, because the portfolio holdings of both buyers and sellers are distorted. Taxing a commodity distorts the amount consumed of a given good, reducing welfare. Taxing financial transactions distorts portfolio allocations, inducing investors to hold more or less risk than they should, also reducing welfare. The distortion created by a tax (approximately) grows with the square in this context of the model studied in this paper.

F.8 Disagreement about other moments

Environment Motivated by Proposition 5, in the baseline model, investors only disagree about the expected value of the payoff of the risky asset. I now assume that investors also hold distorted beliefs about their hedging needs $\mathbb{C}ov_i\left[M_{2i}, D\right]$ and about the variance of the payoff of the risky asset $\mathbb{V}ar_i\left[D\right]$.

Results The optimality condition presented in Equation (5) applies directly, after using the individual beliefs of each investor. Hedging needs enter additively, but perceived individual variances modify the sensitivity of portfolio demands with respect to the baseline case.

Market clearing determines the equilibrium price, given now by:

$$P_1 = \frac{\int_{i\in\mathcal{T}}\left(\frac{\mathbb{E}_i[D]}{\mathcal{AV}_i} - AV\left(\beta_{ii} + X_{0i}\right)\right)dF\left(i\right)}{1 + \tau\int_{i\in\mathcal{T}}\frac{\operatorname{sgn}(\Delta X_{1i})}{\mathcal{AV}_i}dF\left(i\right)},$$

where $AV \equiv \left(\int_{i\in\mathcal{T}}\frac{1}{A_i\mathbb{V}ar_i[D]}dF\left(i\right)\right)^{-1}$ is the harmonic mean of risk aversion coefficients and perceived variances for active investors and $\mathcal{AV}_i \equiv \frac{A_i\mathbb{V}ar_i[D]}{AV}$ is the quotient between investor i risk aversion times perceived variance and the harmonic mean. I define the regression coefficient (beta) of individual endowments M_{2i} on payoffs D perceived by investors by $\beta_{ii} = \frac{\mathbb{C}ov_i[M_{2i},D]}{\mathbb{V}ar_i[D]}$. Again, $\mathcal{T}$ denotes the set of active investors.

Proposition 15. (Disagreement about second moments)

a) The marginal change in social welfare from varying the financial transaction tax when investors disagree about second moments is given by:

$$\frac{dV^p}{d\tau} = \int\left[\left(-r_i\mathbb{E}_i\left[D\right] - A_i\mathbb{C}ov\left[M_{2i}, D\right]\left(1 - \frac{\beta_{ii}}{\beta_i}\right) + P_1 r_i\left(1 + \operatorname{sgn}\left(\Delta X_{1i}\right)\tau\right)\right)\frac{dX_{1i}}{d\tau}\right]dF\left(i\right), \tag{87}$$

where $r_i \equiv \frac{\mathbb{V}ar[D]}{\mathbb{V}ar_i[D]}$, $\beta_{ii} \equiv \frac{\mathbb{C}ov_i[M_{2i},D]}{\mathbb{V}ar_i[D]}$ *and* $\beta_i \equiv \frac{\mathbb{C}ov[M_{2i},D]}{\mathbb{V}ar[D]}$. *Note that* $r_i \in (0,\infty)$ *and* $\beta_i, \beta_{ii} \in (-\infty,\infty)$.

b) The optimal tax when investors disagree about second moments is given by:

$$\tau^* = \frac{\int\left(r_i\mathbb{E}_i\left[D\right] + A_i\mathbb{C}ov\left[M_{2i}, D\right]\left(1 - \frac{\beta_{ii}}{\beta_i}\right)\right)\frac{dX_{1i}}{d\tau}dF\left(i\right)}{P_1\int\left(r_i\left(1 + \operatorname{sgn}\left(\Delta X_{1i}\right)\right)\right)\frac{dX_{1i}}{d\tau}dF\left(i\right)}.$$

Proof. The optimal portfolio allocation for an investor i in his trade region is given by:

$$X_{1i} = \frac{\mathbb{E}_i\left[D\right] - A_i\mathbb{C}ov_i\left[M_{2i}, D\right] - P_1\left(1 + \operatorname{sgn}\left(\Delta X_{1i}\right)\tau\right)}{A_i\mathbb{V}ar_i\left[D\right]}.$$

The marginal change in welfare for an investor i is given by:

$$\frac{dV_i^p}{d\tau} = \left[\begin{array}{c}\left(\mathbb{E}_p\left[D\right] - A_i\mathbb{C}ov\left[M_{2i}, D\right] - P_1\right)\frac{dX_{1i}}{d\tau}\\ -r_i\left(\mathbb{E}_i\left[D\right] - A_i\mathbb{C}ov_i\left[M_{2i}, D\right] - P_1\left(1 + \operatorname{sgn}\left(\Delta X_{1i}\right)\tau\right)\right)\frac{dX_{1i}}{d\tau} - \Delta X_{1i}\frac{dP_1}{d\tau}\end{array}\right].$$

Where $r_i \equiv \frac{\mathbb{V}ar[D]}{\mathbb{V}ar_i[D]}$. The change in social welfare can then be written as:

$$\frac{dV^p}{d\tau} = \int\left(-r_i\mathbb{E}_i\left[D\right] - A_i\mathbb{C}ov\left[M_{2i}, D\right]\left(1 - \frac{\beta_{ii}}{\beta_i}\right) + P_1 r_i\left(1 + \operatorname{sgn}\left(\Delta X_{1i}\right)\tau\right)\right)\frac{dX_{1i}}{d\tau}dF\left(i\right).$$

Solving for τ in this equation, which corresponds to Equation (87) in the paper, delivers the expression for the optimal tax in Proposition 15b). □

[51] Although this result is intuitive, to my knowledge, it had not been derived before in the context of a portfolio choice problem. See Auerbach and Hines Jr. (2002) for a comprehensive analysis of tax efficiency results and Sandmo (1985) for a survey of results on how taxation affects portfolio allocations.

The formula for the optimal tax now incorporates hedging needs and modifies the weights given to investors' beliefs. An investor with correct beliefs about second moments has $r_i = 1$ and $\beta_{ii} = \beta_i$; in that case, we recover Equation (12). When investors perceive a high variance, that is, r_i is close to 0, they receive less weight in the optimal tax formula. The opposite occurs when they perceive a low variance. Intuitively, lower perceived variances amplify distortions in expected payoffs, and vice versa.

As in the baseline model, the planner does not need to know the value of $\mathbb{E}_p[D]$ to implement the optimal tax. However, if investors hold distorted beliefs about their hedging needs, the planner needs to know explicitly the magnitude of the mistake. Intuitively, there is no mechanism in the model which cancels out the mistakes in hedging made by investors. The sign of the optimal tax depends directly on the errors made by investors when hedging.

There are two interesting parameters restrictions. First, when investors with correct expected payoffs and hedging betas, that is $\frac{\beta_{ii}}{\beta_i} = 1$, disagree about variances, the optimal tax τ^* turns out to be:

$$\tau^* = \frac{\mathbb{E}[D] \int r_i \frac{dX_{1i}}{d\tau} dF(i)}{P_1 \int r_i \left(1 + \operatorname{sgn}(\Delta X_{1i})\right) \frac{dX_{1i}}{d\tau} dF(i)}.$$

The dispersion of variances, given by $\mathbb{C}ov_{F,\mathcal{T}}\left[r_i, \frac{dX_{1i}}{d\tau}\right]$, determines now the sign of the optimal tax. When r_i is constant (although not necessarily equal to one), the optimal tax becomes zero. This reinforces the intuition that belief dispersion is what matters for optimal taxes in an exchange economy. Intuitively, when buyers, with $\frac{dX_{1i}}{d\tau} < 0$, are relatively aggressive, that is, r_i is large, they are buying too much of the risky asset, so $\mathbb{C}ov_{F,\mathcal{T}}\left[r_i, \frac{dX_{1i}}{d\tau}\right]$ is negative and the optimal tax is positive, and vice versa.

Second, when investors have correct beliefs about the mean and the variance of expected returns, but hedge incorrectly, the optimal tax becomes:

$$\tau^* = \frac{\mathbb{V}ar[D] \int A_i (\beta_i - \beta_{ii}) \frac{dX_{1i}}{d\tau} dF(i)}{P_1 \int \operatorname{sgn}(\Delta X_{1i}) \frac{dX_{1i}}{d\tau} dF(i)}.$$

The optimal tax now has the opposite sign of $\mathbb{C}ov_{F,\mathcal{T}}\left[A_i (\beta_i - \beta_{ii}), \frac{dX_{1i}}{d\tau}\right]$. Intuitively, when buyers, with $\frac{dX_{1i}}{d\tau} < 0$, overestimate their need for hedging and end up buying too much of the risky asset — this occurs when $\beta_i - \beta_{ii} < 0$ — the optimal tax is positive, and vice versa.