NBER WORKING PAPER SERIES

OPTIMAL FINANCIAL TRANSACTION TAXES

Eduardo Dávila

Working Paper 27826
http://www.nber.org/papers/w27826

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2020

I am very grateful to Philippe Aghion, Adrien Auclert, Robert Barro, Roland Bénabou, Markus Brunnermeier, Eric Budish, 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 provided outstanding research assistance. Financial support from Rafael del Pino Foundation is gratefully acknowledged. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Optimal Financial Transaction Taxes
Eduardo Dávila
NBER Working Paper No. 27826
September 2020
JEL No. D61,G18,H21

ABSTRACT

This paper characterizes the optimal transaction tax in an equilibrium model of competitive financial markets. As long as investors hold heterogeneous beliefs that are not related to their fundamental trading motives and the planner calculates welfare using any single belief, a strictly positive tax is optimal, regardless of the magnitude of fundamental trading. Under some conditions, the optimal tax is independent of the belief used by the planner to calculate welfare. The optimal tax can be implemented by adjusting its value until observed 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. A calibration of the model consistent with empirically estimated volume semi-elasticities to tax changes and that features a 30% share of non-fundamental trading volume is associated with a 37bps optimal tax.

Eduardo Dávila
Yale University
Department of Economics
28 Hillhouse Avenue, Room 304
New Haven, CT 06520-8268
and NBER
eduardo.davila@yale.edu

An online appendix is available at: http://www.nber.org/data-appendix/w27838

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, 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

[1] Although independence between 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.

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 corresponds to 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 two assumptions: traded assets are in fixed supply and the planner does not seek to redistribute resources across investors. Intuitively, because welfare losses in this model arise from a distorted allocation of risk, only the dispersion in investors' beliefs — but not its 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 significant 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 observed volume equals fundamental volume. 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 a simple and easily implementable 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 static 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

[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.

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.

Finally, I establish the robustness of the results. I first characterize the optimal tax for more general specifications of beliefs and utility. In this more general case, investors' risk-adjusted beliefs — now computed using each investor's stochastic discount factor — and portfolio sensitivities are still the key determinants of the optimal tax. Importantly, the optimal tax formula of the baseline model remains valid as an approximation to the optimal tax in the general case when investors' stochastic discount factors are approximately constant. This result validates the analysis in the rest of the paper as an approximation to any specification of beliefs and preferences. 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.

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 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.

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.[4]

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 the beliefs of the investors 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.[5] 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 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. Xiong (2013) surveys 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 (2020) 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

[4] 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.

[5] 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).

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 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, $P_1 > 0$.[6] 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.[7] 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.

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

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

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

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

[7] 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.

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

(iv) Different beliefs: captured by $\mathbb{E}_i[D]$ (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.[8]

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, $\mathbb{E}_i[D] - A\mathbb{C}ov[M_{2i}, D] - A\mathbb{V}ar[D]X_{0i}$, is symmetric.*

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 terms of the numeraire, due at the time the transaction occurs, for both buyers and sellers, of

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

where $|\Delta X_{1i}| \equiv |X_{1i} - X_{0i}|$. The total tax revenue generated by the transaction is thus $2\tau P_1 |\Delta X_{1i}|$. 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.

[8] 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.

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 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 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.[9] 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.[10]

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

[9] 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.

[10] 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.

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

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

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 their 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.

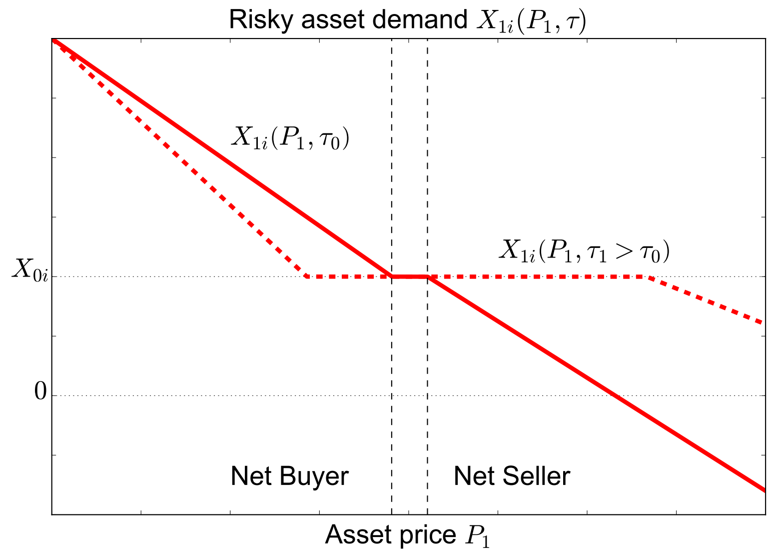


Figure 1: Risky asset demand

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.

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.[11]

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.[12] 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 set of buyers

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

[12] 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.

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 theories in which transaction taxes are desirable must rely on a mechanism through which reducing trading volume is welfare improving.[13]

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.

[13]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).

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-sump 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. 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.

4.1 Welfare criterion

In order to aggregate individual preferences, I assume that the planner maximizes the sum of investors' certainty equivalents. This is a standard approach in normative problems, and corresponds to maximizing a particular set of welfare weights, as explained in detail in the Appendix. 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.[14]

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

[14] 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.

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 is 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 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.[15]

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.

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

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

[15] 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.

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 about 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 $\mathbb{E}_p[D]$.[16]

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

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

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

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

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 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.

[16] 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 crucial 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

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

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,

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

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.[17]

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 the 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

[17] 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.

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.[18] However, because the planner employs a second-best policy instrument — a single linear tax — the portfolio 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.[19] Note that the weights assigned to buyers $\omega_i^{\mathcal{B}}$ and sellers $\omega_i^{\mathcal{S}}$ add up to one 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 belief 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:

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

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

[19]The presence of demand/portfolio sensitivities in optimal corrective tax formulas goes back to Diamond (1973), who analyzes corrective taxation with restricted instruments in a model of consumption externalities. See also Rothschild and Scheuer (2016) for a recent application of similar principles.

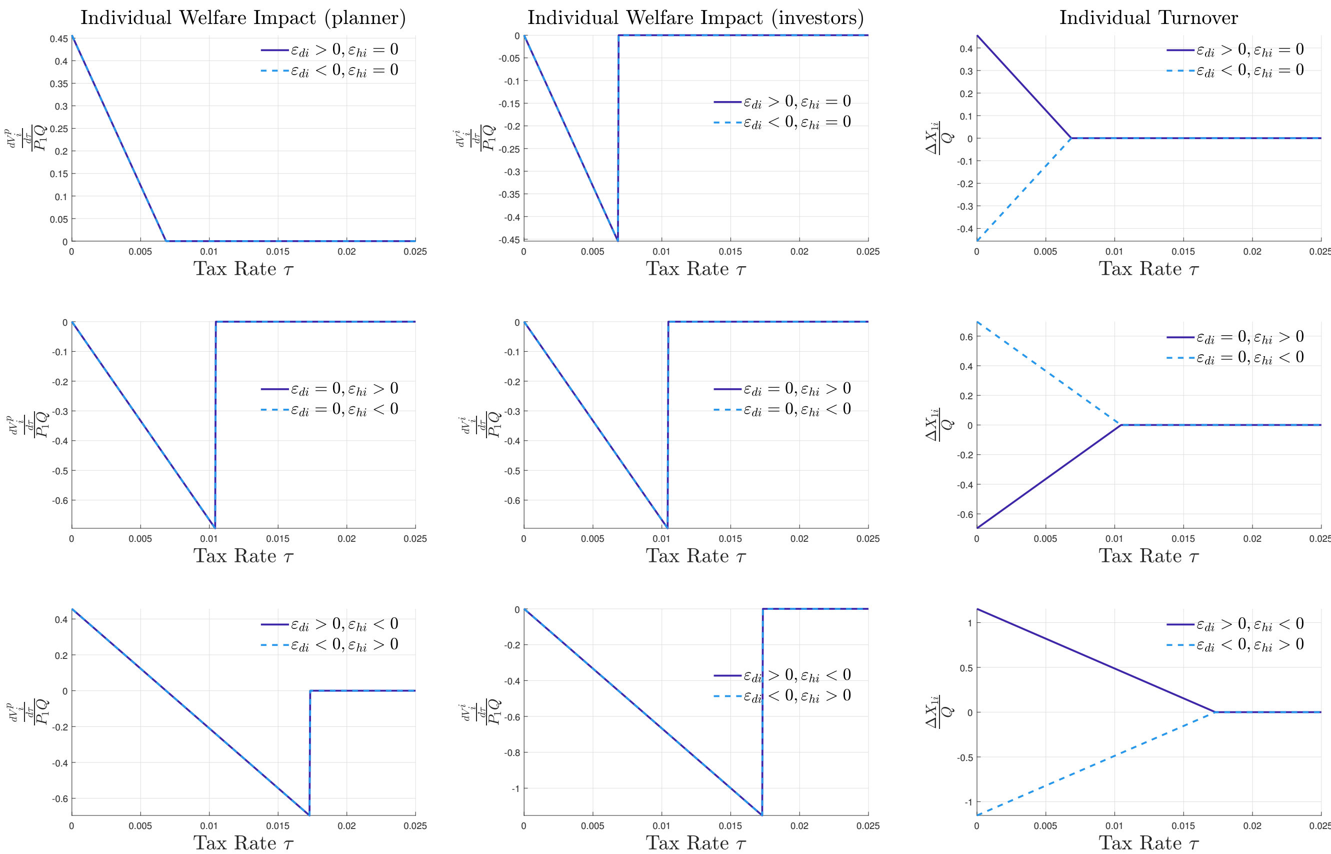


Figure 2: Individual marginal welfare impact (planner's and investors' perspective)

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 the each individual investor (middle plots), for different values of τ. These respectively correspond to Equations (10) and (30). All plots in Figure 2 satisfy Assumption [G], described on page 25, 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 ($\varepsilon_{hi} > 0$ 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_{hi} 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 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 i) that the planner's mean belief $\mathbb{E}_p[D]$ equals the average mean belief μ_d, and ii) an individually targeted rebate rule. See Figures A.8 and A.10b in the Appendix for an illustration of how the results change after relaxing both assumptions.

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 optimal tax is positive.

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 specifically 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 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.[20] 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 that are also optimistic (pessimist). 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.

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

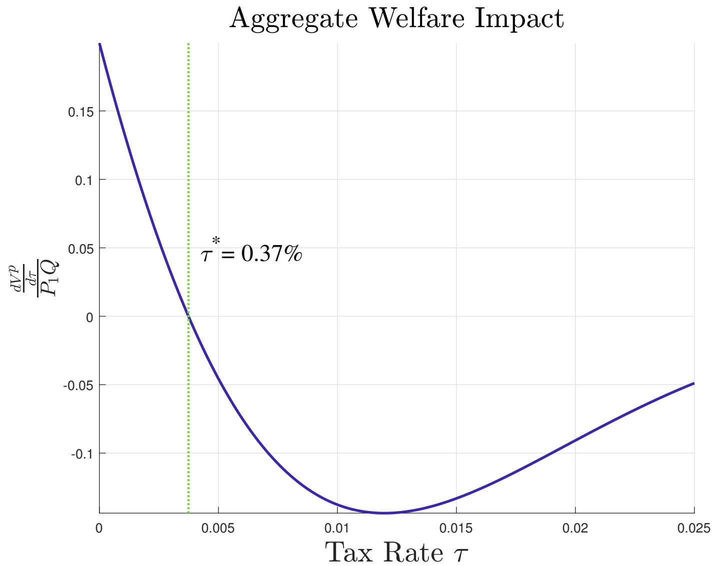


Figure 3: Aggregate marginal welfare impact

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 25, and uses 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 optimal tax, represented by a vertical dotted line, is $\tau^* = 0.37\%$.

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 belief. Leaving aside relative price changes and net transfers, these plots clearly illustrate that every investor perceives to be worse off using his own belief to compute his own welfare when facing a positive tax.

Under which conditions could a trading subsidy 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 is 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.

Two features of the economic environment are essential for this result to hold. First, it is key that 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, 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.

Second, the planner does not use the transaction tax with the purpose of redistributing resources across investors. Intuitively, the linearity of investors' certainty equivalents on the planner's expected payoff combined with the fact that the planner gives equal weight to the welfare gains/losses across investors in dollar terms guarantee that, after aggregating across investors, the optimal tax does not depend on the differences between investors' beliefs and the planner's belief, but only on the belief dispersion among investors.

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.[21] 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. 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 corresponds with a belief-neutral Pareto efficient allocation. The planner would need additional instruments to implement such 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 to 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

[21] I show in the Appendix that maximizing investors' certainty equivalents is identical to using generalized social welfare weights, using the terminology of Saez and Stantcheva (2016), which in turn can be mapped to traditional social welfare weights, as those used in the criterion of Brunnermeier, Simsek and Xiong (2014).

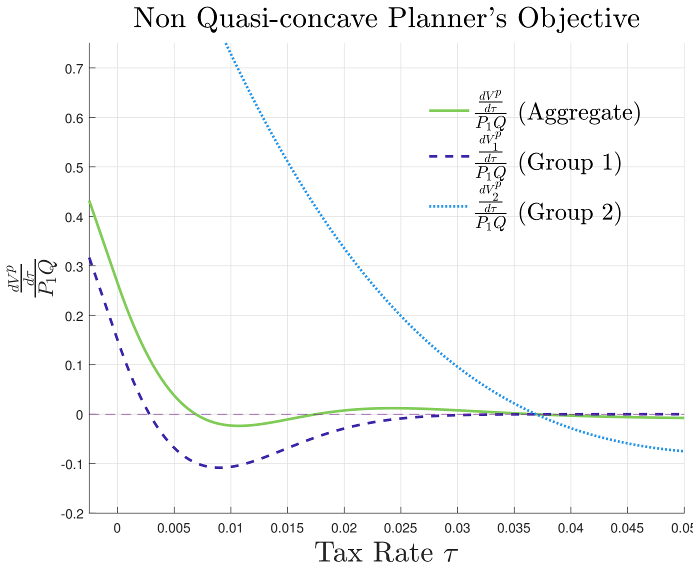


Figure 4: Failure of quasi-concavity of planner's objective

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 $\tau = 1.72\%$.

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' tail 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*.*

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 it 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 assessments. 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* $\mathbb{C}ov\left[M_{2i}, D\right]$ *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.[22] Under this alternative approach, the planner must adjust the tax rate until total trading volume equals fundamental volume. Proposition 2 provides a decomposition of trading volume into different components and describes a new implementation of the optimal policy that compares total trading volume with fundamental volume.

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

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

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

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

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

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

b) [Alternative optimal policy implementation] The planner can implement the optimal corrective policy by adjusting the tax rate until trading volume equals fundamental volume. Formally,

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

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 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

[22] 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.

to fundamental and non-fundamental components. 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 amount of fundamental trading volume, it can adjust the optimal tax until observed volume is commensurate with the appropriate amount of fundamental trading. 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. Alternatively, one can also reinterpret the optimal policy as setting a tax rate such that the tax-induced volume reduction equals non-fundamental volume, that is, setting τ^* so that $\Theta_\tau(\tau^*) = \Theta_{NF}(\tau^*)$. This alternative implementation, which is not a direct consequence of classic Pigouvian logic, relies on the ability to relate total trading volume to belief differences (the marginal distortion) and to the impact of the tax on trading (the direct effect of the policy instrument on investors' portfolios).

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 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.

Remark. (Practical advantage of trading volume implementation) The trading volume implementation of the optimal policy has two distinct practical advantages relative to directly using the optimal tax formula characterized in Equation (12). First, instead of having to explicitly recover or estimate investors' beliefs to set the optimal policy, it is sufficient to rely on a model that predicts the level of fundamental volume in the economy and then adjust the tax level under observed total volume reaches that level. Second, Equation (17) provides a tractable and intuitive approximation for the optimal tax with minimal informational requirements for a planner and under no restrictions regarding investors' trading motives.

5 Quantitative assessment

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, where $\rho \in [-1, 1]$ and $\sigma_d^2, \sigma_h^2 \geq 0$.[23]

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

Investors have identical preferences $A_i = A$ and hold identical initial asset positions $X_{0i} = X_0$.

The cross-sectional dispersion of investors' mean beliefs σ_d^2 and 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 (sellers) for fundamental motives. Making investors' preferences identical and assuming that they have identical asset holdings of the risky asset eliminates 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.

[23]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.

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 in 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 $\tau^* \in [0, \infty)$; while 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

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 same two 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 $\varepsilon_{\tau}^{\log \mathcal{V}} = \frac{d \log \mathcal{V}}{d\tau}$.[24]

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 $\varepsilon_{\tau}^{\log \mathcal{V}} = \frac{d \log \mathcal{V}}{d\tau}$, 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 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

[24] As implied by Equation (50) 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 $\varepsilon_{\tau}^{\log \mathcal{V}}\big|_{\tau=0}$.

statistics identified, e.g., the risk premium, the volume semi-elasticity 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.

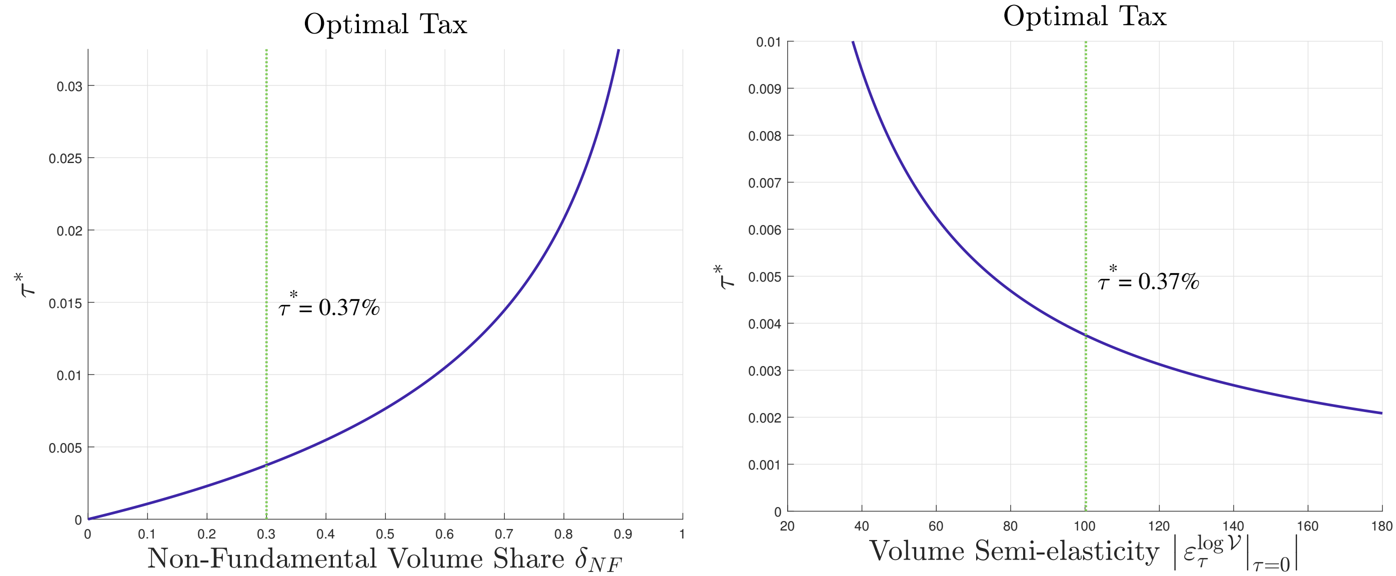
Optimal tax calibration Proposition 4 shows that finding an empirical counterpart of the optimal tax exclusively requires measures of $\varepsilon_{\tau}^{\log \mathcal{V}}$ and δ^{NF}, and that an estimate of the risk premium is necessary to compute welfare gains. Next, I describe how to find the plausible 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 $\varepsilon_{\tau}^{\log \mathcal{V}}\Big|_{\tau=0} = \frac{d \log \mathcal{V}}{d\tau}\Big|_{\tau=0} = \frac{-20\%}{0.2\%} = -100.$[25] As shown in the Appendix, in this model there is a tight relation between 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

[25] I adopt as reference the average estimate in Colliard and Hoffmann (2017) for stocks outside 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: The left panel in Figure 5 shows the optimal tax τ^* as a function of the share of non-fundamental trading volume δ^{NF}, defined in Equation (19). The right panel in Figure 5 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 $\tau^* = 0.37\%$.

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 6 illustrate the magnitudes implied by the model for the optimal tax and also provide a sensitivity analysis.

Figure 5a illustrates how the optimal tax varies as a function of the ratio 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

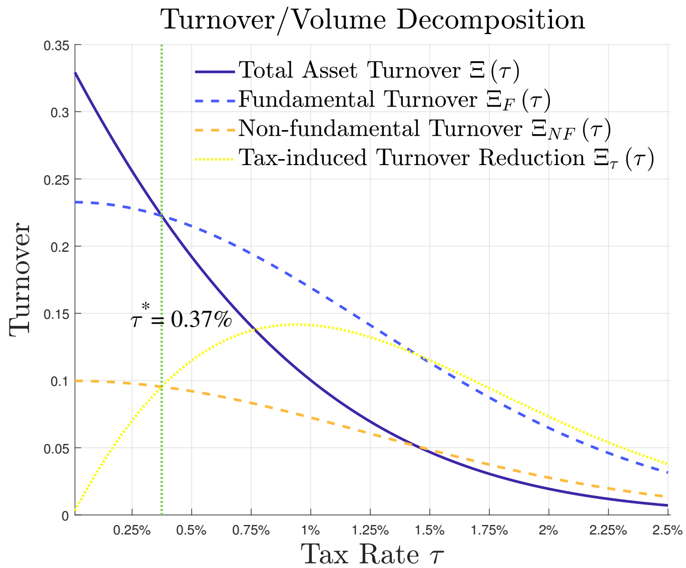


Figure 6: Trading volume implementation

Note: Figure 6 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 $\Xi(\tau) = \frac{P_1 \mathcal{V}(\tau)}{Q}$, fundamental turnover, given by $\Xi_F(\tau) = \frac{\Theta_F(\tau)}{Q}$, non-fundamental turnover, given by $\Xi_{NF}(\tau) = \frac{\Theta_{NF}(\tau)}{Q}$, and the tax induced turnover reduction $\Xi_\tau(\tau) = \frac{\Theta_\tau(\tau)}{Q}$. 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).

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 $\varepsilon_\tau^{\log \mathcal{V}}\Big|_{\tau=0} = -150$, $\tau^* = 0.23\%$. 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 6 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 component of volume grows rapidly at first before it starts decreasing monotonically, due to 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 component intersect as well at the same tax rate.

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.

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 $\left.\frac{\frac{dV^p}{d\tau}}{P_1 Q}\right|_{\tau=0} = \frac{1}{\Pi} \frac{1}{\left|\varepsilon_\tau^{\log \mathcal{V}}\right|_{\tau=0}\right|} \delta^{NF}$, 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.

It is worth concluding with two remarks. First, 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. Second, the quantitative results in this Section assume that the volume reduction associated with a tax increase represents a behavioral response and not tax avoidance. This is consistent with most of the derivations in the paper and is a reasonable assumption for small taxes. However, if tax enforcement is imperfect, one must interpret the optimal taxes reported in this section as upper bounds, and refine the analysis along the lines of Section F.3 in the Appendix.

6 Robustness of the results

Before concluding, I characterize the optimal financial transaction tax under more general assumptions on preferences and beliefs. Proposition 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. It also shows that the results derived in the rest of the paper remain valid as an approximation to more general models. Finally, based on results derived in the Appendix, I discuss several extensions to the baseline model.

6.1 General utility and arbitrary beliefs

This section extends the main results of the paper to 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

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

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

$$1 = R\mathbb{E}_i[m_i] \tag{21}$$

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

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 (22), which is the counterpart of Equation (5), does not hold.

As in the baseline model, the planner computes social welfare under a single belief. In this case, the planner uses uniform generalized welfare weights — using the terminology of Saez and Stantcheva (2016). As described in detail in the Appendix, this is equivalent to maximizing a weighted sum of investors' utilities. Proposition 5 characterizes the exact optimal tax in the general case, as well as its approximation when investors' stochastic discount factors are roughly constant. Importantly, the approximated optimal tax characterization in this general model turns out to be identical to the optimal tax characterization in the baseline model.

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

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

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

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

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

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

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

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

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

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

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.

b) [Approximation] In the limit in which investors' stochastic discount factors are approximately constant, formally, when $m_i \to \overline{m}$, where $\overline{m} \in \mathbb{R}_+$ 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 Equations (12) and (13). This expression is identical to the one in Proposition 1a).

The optimal tax characterized in Proposition 5a) 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 $\mathbb{E}_p[m_i D]$ (also $\mathbb{E}_p[m_i R]$) 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 $\theta(\tau^*)$ 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 $\frac{d\mathcal{V}^s(\tau^*)}{d\tau} = 0$, then $\theta(\tau^*) = 0$.

When investors stochastic discount factors are approximately constant, Proposition 5b) finds 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.[26] Proposition 5b) allows us to interpret the results of the baseline model 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 particular, it shows that investors' optimal portfolios in the general case map to those in the baseline model after linearly approximating investors' stochastic discount factor around their mean. This result implies that the volume decomposition characterized in Proposition 2 is also approximately valid more generally. 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.

6.2 Discussion of 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 1 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 1 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.

[26]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.

Finally, in a q-theory production economy, I show that a transaction tax generates additional first-order gains/losses as long as the planner's belief differs from the average belief of investors. In addition to the allocation of risk among investors, the level of aggregate risk and investment in a production economy also affects welfare. 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. In principle, the optimal tax formula in a production economy depends on the belief used by the planner. However, 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. Access to an additional policy instrument that targets aggregate investment would be optimal in this environment, allowing the planner to set the optimal transaction tax as in the baseline model, for any planner's belief.[27]

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.

Although the paper includes many extensions, there are related questions worth exploring. Understanding the normative implications of taxing financial transactions in models with endogenous learning dynamics or with rich wealth dynamics, when markets are decentralized, or when some investors have market power are fruitful avenues for further research.

[27] 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.

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:

$$J(X_{1i}) = [\mathbb{E}_i[D] - A_i \mathbb{C}ov[M_{2i}, D] - P_1] X_{1i} + P_1 X_{0i} - \tau P_1 |\Delta X_{1i}| + T_{1i} - \frac{A_i}{2} \mathbb{V}ar[D] X_{1i}^2 \tag{23}$$

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

$$J'(X_{1i}) = \mathbb{E}_i[D] - A_i \mathbb{C}ov[M_{2i}, D] - P_1 - \tau P_1 \operatorname{sgn}(\Delta X_{1i}) - A_i \mathbb{V}ar[D] X_{1i} = 0 \tag{24}$$

$$J''(X_{1i}) = -A_i \mathbb{V}ar[D] < 0. \tag{25}$$

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 $J''(\cdot) < 0$ and $\lim_{X_{1i} \to \pm\infty} J'(X_{1i}) = -\infty$, 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 $J''(\cdot) < 0$ and $\lim_{X_{1i} \to \pm\infty} J'(X_{1i}) = -\infty$, 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.[28] 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 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.

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

$$\frac{dX_{1i}}{d\tau} = \frac{\partial X_{1i}}{\partial \tau} \underbrace{\left[1 - (\operatorname{sgn}(\Delta X_{1i}) + \tau) \frac{\int_{i \in \mathcal{T}(P_1)} \frac{\operatorname{sgn}(\Delta X_{1i})}{A_i} dF(i)}{\int_{i \in \mathcal{T}(P_1)} \frac{1 + \operatorname{sgn}(\Delta X_{1i})\tau}{A_i} dF(i)}\right]}_{\equiv \varepsilon_i}, \tag{26}$$

where $\frac{\partial X_{1i}}{\partial \tau} = \frac{-P_1 \operatorname{sgn}(\Delta X_{1i})}{A_i \mathbb{V}ar[D]}$, $\frac{\partial X_{1i}}{\partial P} = \frac{-(1 + \operatorname{sgn}(\Delta X_{1i})\tau)}{A_i \mathbb{V}ar[D]}$, and it is straightforward to show that $\varepsilon_i > 0$ for both buyers and sellers. Equation (26) 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

$$\frac{dP_1}{d\tau} = \frac{\int_{i \in \mathcal{T}(P_1)} \frac{\partial X_{1i}}{\partial \tau} dF(i)}{-\int_{i \in \mathcal{T}(P_1)} \frac{\partial X_{1i}}{\partial P} dF(i)} = \frac{-\left(\int_{i \in \mathcal{B}(P_1)} \frac{P_1}{A_i \mathbb{V}ar[D]} dF(i) - \int_{i \in \mathcal{S}(P_1)} \frac{P_1}{A_i \mathbb{V}ar[D]} dF(i)\right)}{\int_{i \in \mathcal{T}(P_1)} \frac{1 + \operatorname{sgn}(\Delta X_{1i})\tau}{A_i \mathbb{V}ar[D]} dF(i)}. \tag{27}$$

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 (27) 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 for P_1 to take negative values, Equation (2) becomes $\tau|P_1||\Delta X_{1i}|$.

[28] 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 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

$$\frac{dV_i^p}{d\tau} = \left[\mathbb{E}_p[D] - \mathbb{E}_i[D] + \operatorname{sgn}(\Delta X_{1i}) P_1 \tau\right] \frac{dX_{1i}}{d\tau} - \Delta X_{1i} \frac{dP_1}{d\tau} + \frac{d\tilde{T}_{1i}}{d\tau}. \tag{28}$$

The derivation of Equation (28) 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

$$\frac{dV^p}{d\tau} = \int_{i \in \mathcal{T}(\tau)} \left[-\mathbb{E}_i[D] + \operatorname{sgn}(\Delta X_{1i}) P_1 \tau\right] \frac{dX_{1i}}{d\tau} dF(i), \tag{29}$$

where Equation (29) 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

$$\frac{dV_i^i}{d\tau} = \operatorname{sgn}(\Delta X_{1i}) P_1 \tau \frac{dX_{1i}}{d\tau} - \Delta X_{1i} \frac{dP_1}{d\tau} + \frac{d\tilde{T}_{1i}}{d\tau}. \tag{30}$$

Proposition 1. (Optimal financial transaction tax)

Proof. a) [Optimal tax formula] Starting from Equation (29), 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

$$\begin{aligned}\left.\frac{dV^p}{d\tau}\right|_{\tau=0} &= -\int_{i\in\mathcal{T}(0)} \mathbb{E}_i[D] \left.\frac{dX_{1i}}{d\tau}\right|_{\tau=0} dF(i) = -\mathbb{C}ov_F\left(\mathbb{E}_i[D], \left.\frac{dX_{1i}}{d\tau}\right|_{\tau=0}\right) \\ &= \frac{P_1}{\mathbb{V}ar[D]}\left[\mathbb{C}ov_F\left(\mathbb{E}_i[D], \frac{\mathbb{I}\left[\Delta X_{1i}|_{\tau=0} > 0\right]}{A_i}\right)\varepsilon_B - \mathbb{C}ov_F\left(\mathbb{E}_i[D], \frac{\mathbb{I}\left[\Delta X_{1i}|_{\tau=0} < 0\right]}{A_i}\right)\varepsilon_S\right],\end{aligned}$$

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

$$\Delta X_{1i} = \underbrace{\frac{\mathbb{E}_i[D] - \mathbb{E}_F[\mathbb{E}_i[D]] + A\mathbb{E}_F[\mathbb{C}ov[M_{2i}, D]] + A\mathbb{V}ar[D]Q}{A_i\mathbb{V}ar[D]}}_{\equiv Z_1} \underbrace{-\frac{\mathbb{C}ov[M_{2i}, D]}{\mathbb{V}ar[D]} - X_{0i}}_{\equiv Z_2}, \tag{31}$$

where Z_1 and Z_2 are defined in Equation (31) 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 $\left.\frac{dV}{d\tau}\right|_{\tau=0} > 0$ 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 (29). 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) are 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 (29).

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.[29] Formally, the second order condition of the planner's problem is given by

$$\frac{d^2V^p}{d\tau^2} = \frac{d\left(P_1\tau\right)}{d\tau}\int_{i\in\mathcal{T}(\tau)} \operatorname{sgn}\left(\Delta X_{1i}\right)\frac{dX_{1i}}{d\tau}dF\left(i\right) + \int_{i\in\mathcal{T}(\tau)}\left[-\mathbb{E}_i\left[D\right] + \operatorname{sgn}\left(\Delta X_{1i}\right)P_1\tau\right]\frac{d^2X_{1i}}{d\tau^2}dF\left(i\right) + \\ -\int_{\tilde{\mathcal{B}}(\tau)}\underbrace{\left[-\mathbb{E}_i\left[D\right] + P_1\tau\right]\frac{dX_{1i}}{d\tau}}_{\frac{dV_i^p}{d\tau}}dF\left(i\right) - \int_{\tilde{\mathcal{S}}(\tau)}\underbrace{\left[-\mathbb{E}_i\left[D\right] - P_1\tau\right]\frac{dX_{1i}}{d\tau}}_{\frac{dV_i^p}{d\tau}}dF\left(i\right), \tag{32}$$

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

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

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 (32). The first term in Equation (32) 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 $\int_{i\in\tilde{\mathcal{T}}(\tau)}\frac{dV_i^p}{d\tau}dF\left(i\right) > 0$, where $\tilde{\mathcal{T}}\left(\tau\right)$ 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 (32) that, at any interior optimum without extensive margin effects

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

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

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

[29]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.

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 symmetry, 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

$$\frac{\partial\mathbb{E}_{\mathcal{B}(\tau)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]}{\partial\tau}-\frac{\partial\mathbb{E}_{\mathcal{S}(\tau)}\left[\frac{\mathbb{E}_i[D]}{P_1}\right]}{\partial\tau}<2, \tag{33}$$

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 (33) 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

$$V^p_{notrade}=\left(\mathbb{E}_p[D]-A\mathbb{C}ov[M_2,D]\right)Q-\frac{A}{2}\mathbb{V}ar[D](Q)^2.$$

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

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

Since $\int X_{1i}dF(i)=Q$, a straight application of Jensen's inequality immediately implies that $V^p_{notrade}>V^p_{trade}$ whenever $X_{1i}\neq Q$ for at least a single investor. □

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

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

which allows us to write trading volume as

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

using the fact that

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

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

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

where each of the elements is given by

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

When Assumption [S] holds, $\frac{dX_{1i}}{d\tau}$ is constant across investors and $\kappa\left(P_1,\tau\right)=1$, 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

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

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

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

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

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

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

Using the fact that $\kappa\left(P_1,0\right)=1$ and $\Theta_\tau\left(0\right)=0$, this expression can be approximated around $\tau^*\approx0$ 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. □

References

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.

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.*

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.

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. 2020. "Trading Costs and Informational Efficiency." *Journal of Finance (forthcoming).*

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

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.

Rothschild, Casey, and Florian Scheuer. 2016. "Optimal Taxation with Rent-Seeking." *The Review of Economic Studies*, 83(3): 1225–1262.

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.

Santos, Tano, and Jose A. Scheinkman. 2001. "Competition Among Exchanges." *The Quarterly Journal of Economics*, 116(3): 1027–1061.

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.

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.*