Optimal Financial Transaction Taxes

Eduardo Dávila∗
NYU Stern
May 2015

Abstract

This paper characterizes the optimal linear financial transaction tax in an equilibrium model of competitive financial markets. As long as investors hold heterogeneous beliefs about the returns of assets in fixed supply and the planner calculates welfare using any single belief, we expect the optimal tax to be positive, even when a fraction of trading is fundamental. Strikingly, the optimal tax is independent of the belief used by the planner to calculate welfare. If assets are in elastic supply, as in q-theory, the sign of the optimal tax is indeterminate, although the optimal tax remains unchanged if the planner uses the average belief of investors to calculate welfare. In dynamic environments, the optimal tax is lower when investors trade more frequently.

JEL Classification: D61, H21, G18

Keywords: financial transaction tax, Tobin tax, belief disagreement, non-fundamental trading, optimal taxation

∗Contact: edavila@stern.nyu.edu. I would like to especially thank John Campbell, Emmanuel Farhi, Jeremy Stein and Alp Simsek for many helpful comments and advice. I am very grateful to Philippe Aghion, Adrien Auclert, Robert Barro, Markus Brunnermeier, Raj Chetty, Chris Edmond, Jordi Gali, Xavier Gabaix, Gita Gopinath, Robin Greenwood, Sam Hanson, Oliver Hart, Nathan Hendren, David Laibson, Greg Mankiw, Sendhil Mullainathan, Jonathan Parker, Adriano Rampini, David Scharfstein, Andrei Shleifer, Adi Sunderam, Luis Viceira and seminar participants at many institutions and conferences for comments. Financial support from Rafael del Pino Foundation is gratefully acknowledged.
1 Introduction

Should we set a financial transaction tax to curb speculation generated by distorted beliefs? Or would a “Tobin” tax prevent welfare enhancing trades from being executed? These questions routinely arise after periods of turmoil in financial markets. For instance, Tobin’s well-known 1972 speech follows the collapse of the Bretton Woods system, Stiglitz (1989) and Summers and Summers (1989) write on the subject after the 1987 crash and, spurred by the 2008 financial crisis, the European Commission seems eager at present to impose a transaction tax. However, as recently posed by Cochrane (2013), a financial transaction tax may seem “the perennial favorite answer in search of a question”. This paper studies the welfare implications of taxing financial transactions in an equilibrium model of competitive financial markets.

Financial markets play three distinct roles in this paper. First, they allow investors to conduct fundamental trading. Fundamental trading allows for risk sharing among similar investors or risk transfer to those investors who can better bear risk. It also allows for liquidity or life-cycle trading needs, as well as trading for market-making or (limited) arbitrage purposes. Second, financial markets allow investors to engage in betting or gambling — I refer to this as non-fundamental trading. In this paper, I capture this notion by assuming that some trades made by investors are due to having different beliefs. Third, as in q-theory models, financial markets determine production and investment by affecting the relative scarcity of a given asset/technology, enticing investors to supply more or less of it.

I initially lay out a static CARA-Normal model of an exchange economy as a benchmark to understand the tradeoffs between fundamental and non-fundamental trading. First, I show that, as long as investors hold heterogeneous beliefs and the planner calculates welfare using a single belief, it is always optimal to set a corrective policy, which can be a tax or a subsidy, depending on the primitives of the economy. I also show that the optimal corrective policy for most economies is a strictly positive tax. Even though introducing a (small) transaction tax distorts investors’ portfolio allocations towards their no-trade positions, the reduction in fundamental trading creates a second-order welfare loss while the reduction in non-fundamental trading creates a first-order gain. Intuitively, because fundamental trades are done optimally, the envelope theorem guarantees that the (local) welfare loss derived from reducing them is second-order. However, preventing non-fundamental trades creates a first-order welfare gain, because these trades are done suboptimally when a single belief is used to calculate welfare.

Second, I characterize the optimal transaction tax. The optimal tax globally balances the welfare gains generated by reducing non-fundamental trades with the welfare losses generated by distorting fundamental trades. The cross-sectional covariance between investors’ beliefs and equilibrium portfolio tax sensitivities becomes the single sufficient statistic for the optimal tax. Once this covariance is known, no other information about the environment is needed to determine the optimal tax. This result is due to the corrective nature of the tax. Strikingly, the belief used by the planner to calculate welfare turns out to be irrelevant to determine the optimal transaction tax. This result relies on two assumptions: a) traded assets are in fixed supply and b) the planner does not seek to redistribute income across investors.

Moreover, by analogy to the Harberger (1964) triangle calculation, I provide a formula for the upper bound of welfare losses induced by a marginal tax change when all trades are deemed fundamental. I also show that the planner’s problem is in general non-convex, so different levels of transaction
taxes may yield similar levels of welfare. This result, which has important practical implications, seems surprising in an economy in which investors face a convex problem and there are no income effects. This non-convexity arises because the distribution of marginal investors changes in the extensive margin for different levels of the transaction tax. Furthermore, I show how the optimal policy can be implemented using trading volume as an intermediate target. This alternative approach allows the planner to implement the optimal transaction tax without knowing investors’ beliefs, simply by varying the tax level until observed volume equals fundamental volume.

Within the static framework, I explore several extensions. First, I show that the optimal tax formula derived in the baseline model remains unchanged when there are pre-existing trading costs, as long as these are compensation for the use of economic resources, not economic rents. I then show that, when pre-existing trading costs reduce fundamental trading relatively more than non-fundamental trading, the optimal transaction tax can be increasing in the level of trading costs and vice versa, a seemingly counterintuitive result. Second, in an environment with multiple risky assets, the optimal tax becomes a weighted average of the optimal tax for each individual asset. A planner optimally gives more weight to those assets whose volume is more sensitive to tax changes. Third, the optimal tax formula in a model with general utility and an arbitrary form of disagreement collapses to the one in the baseline model in the limit when marginal utility is constant, that is, when the risks faced by investors are small relative to their risk bearing capacity. This shows that the results of the baseline model hold in general as a first-order approximation to any economy. Finally, I show that the planner would need investor-specific taxes in order to implement the first-best outcome.

Subsequently, I introduce production. In a q-theory production economy, a (small) transaction tax creates an additional first-order gain or loss as long as the planner’s belief differs from the average belief of investors. In addition to the appropriate allocation of risk among investors, the level of aggregate risk in the economy and the allocation of investment across sectors also matter for welfare. If a marginal tax increase reduces (increases) investment at the margin when investors are too optimistic (pessimistic) with respect to the planner, a positive tax is welfare improving, and vice versa. Unlike in exchange economies, there is no clear theoretical prediction for whether the optimal policy is a tax or a subsidy. 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, which implies that the planner does not actually need to know the average belief of investors to determine the optimal tax.

Finally, I study dynamics and show that the dynamic structure of the economy affects the magnitude of the optimal tax. Controlling for the level of static disagreement, the optimal tax in a dynamic environment is smaller when investors alternate between buying and selling over time. Intuitively, a transaction tax is more effective with forward-looking investors who buy and sell at high frequencies, since the anticipation of future taxes further reduces the incentives to trade today. Buy-and-hold investors are much less affected by the transaction tax so, if they predominate, a larger optimal tax is needed. This result is consistent with Tobin’s insight that high frequency trading is more affected by a transaction tax, but it drastically shifts the emphasis towards identifying first the source of the trading distortion and then adjusting the magnitude of the optimal corrective tax depending on the patterns of trading across investors.
The main contribution of this paper is not to point out that some form of intervention can improve welfare when investors hold heterogeneous beliefs and the planner calculates welfare using an alternative belief: that should not be too surprising. The key contribution of this paper is to sharply characterize the form of the optimal intervention very generally and to show that the thorny issue of which believe to pick to calculate welfare does not arise in very natural environments, in particular, in exchange economies or in production economies in which the planner adopts the average belief to calculate welfare. I should also emphasize that the goal of this paper is not to evaluate the vague proposals currently made by policymakers regarding transaction taxes, which are only loosely grounded in economic theory. Instead, this paper systematically derives, from first principles, the normative implications of taxing financial transactions in a general environment that encompasses both fundamental and non-fundamental trading. This is the first paper to do so.

Related Literature

This paper belongs to the literature that analyzes the Tobin (1978) proposal of introducing financial transaction taxes as a way to improve the societal performance of financial markets. Although Tobin originally 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 lay out several arguments that support the implementation of a financial transaction tax; Ross (1989) takes the opposite view, highlighting many of the apparent contradictions of those verbal arguments. Roll (1989) and Schwert and Seguin (1993) also verbally discuss related issues. Campbell and Froot (1994), several chapters in ul Haq, Kaul and Grunberg (1996) and Jones (1997) are representative samples of empirical work in the area. See McCulloch and Pacillo (2011) for a recent survey of the empirical literature. It is puzzling that almost forty years after Tobin’s original piece, there has been no systematic theoretical study on the normative implications of taxing financial transactions. This paper provides a first step in that direction.

This paper is directly related to the growing literature that evaluates welfare under belief disagreements in financial markets. Weyl (2007) studies the efficiency of arbitrage in an economy in which some investors have mistaken beliefs. Brunnermeier, Simsek and Xiong (2014) propose a criterion to evaluate welfare in models with belief heterogeneity: they assess efficiency by evaluating welfare under a convex combination of the beliefs of the investors in the economy. Gilboa, Samuelson and Schmeidler (2014) and Gayer et al. (2014) present refined Pareto criteria that identify negative-sum betting situations. No-Betting Pareto requires that there exists a single belief that, if shared, implies that all agents are better off by trading. Unanimity Pareto requires that every agent perceives, using his own belief, that all agents are better off by trading. 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 exactly the goal of this paper. Blume et al. (2013) propose a different criterion: their planner evaluates welfare under the worst case scenario among all possible belief assignments. Their results relate to the Friedman (1953) selection hypothesis, in the sense that long-run wealth accumulation

---

1I make an effort to point out throughout the paper logical mistakes made in such proposals.
is crucial for them to determine welfare. They quantitatively analyze several restrictions on trading, but do not characterize optimal policies. Heyerdahl-Larsen and Walden (2014) have recently proposed an alternative 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. In addition to characterizing optimal policy measures for the first time within this line of work, the approach of this paper has the upshot that, by exploiting the structure of the economy, is able to find conditions under which the preferred allocation for the planner turns out to be independent of which belief is used to calculate welfare.

This paper also relates to the sizable literature on belief disagreements, speculation and trading, represented by Harrison and Kreps (1978), Hong and Stein (1999), Scheinkman and Xiong (2003), Geanakoplos (2010) or Simsek (2013a). In particular, Scheinkman and Xiong (2003) analyze the positive implications of a transaction tax in a model with belief disagreement; however, they do not make any statements about welfare. Panageas (2005) and Simsek (2013b) study implications for production and risk sharing of speculative trading motives. See Xiong (2012) for a recent survey of this line of work. More broadly, in the sense that 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, used in Subrahmanyam (1998) and Dow and Rahi (2000) to study transaction taxes, is not appropriate to study the welfare implications of policies. In particular, it is 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.

My results are also related to the growing literature on behavioral welfare economics. The emphasis on the envelope theorem is reminiscent of Akerlof and Yellen (1985), which analyze how individual near-rational behavior can have first-order effects in the aggregate. My paper is more directly related to O’Donoghue and Rabin (2006), in which commitment problems due to hyperbolic discounting — a form of behavioral distortion — for a fraction of investors create room for public intervention through taxation. Sandroni and Squintani (2007) and Spinnewijn (2012) analyze behavioral biases in insurance markets and unemployment search respectively. See Bernheim and Rangel (2009) and Mullainathan, Schwartzstein and Congdon (2012) for recent surveys of behavioral welfare economics and the recent work by Farhi and Gabaix (2015) for a systematic analysis of optimal taxation with behavioral agents.

Finally, the literature on transaction costs is formally related to this paper, since a financial 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 survey. While those papers focus on the positive implications of exogenously given transaction costs/taxes, this paper focuses on the effect of transaction taxes on welfare and on its optimal determination. I explicitly relate in the text the positive results of the paper to this work when appropriate.

Outline Section 2 presents and solves the baseline model and section 3 executes the normative analysis, delivering the main results. Section 4 analyzes several extensions within the static exchange economy model while sections 5 and 6 respectively allow for production and dynamics. Section 7 puts the results on perspective and section 8 concludes. All proofs, derivations and additional results are in the online appendix.
2 Baseline model

This section presents a static model with CARA utility and investors who disagree only about expected returns. When transaction taxes are zero, this environment is identical to Lintner (1969), who relaxes the standard CAPM by allowing for heterogeneous beliefs among investors.

2.1 Environment

**Investors** Time is discrete, there are two dates \( t = \{1, 2\} \) and there is a unit measure of investors. Investors are indexed by \( i \) and they are distributed according to a probability distribution \( F; \) therefore \( \int dF (i) = 1. \) The distribution \( F \) can be discrete, continuous or a mixture of both.

Investors choose their portfolio allocation at \( t = 1 \) and consume at \( t = 2. \) They maximize expected utility with preferences that feature constant absolute risk aversion. Therefore, each investor maximizes:

\[
    E_i [U_i (W_{2i})] \quad \text{with} \quad U_i (W_{2i}) = -e^{-A_i W_{2i}},
\]

where (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 \( A_i \equiv -\frac{U''_i}{U'_i}, \) can vary freely in the distribution of investors. The index \( i \) in the expectation is necessary because every investor holds heterogeneous beliefs about expected returns. For now, investors only disagree about the first moment of the distribution of returns.

**Market structure and beliefs** There exists a riskless asset in elastic supply with a (exogenously determined) gross interest rate \( R = 1. \) There is a single risky asset with exogenously fixed supply \( Q \geq 0. \) The price of the risky asset at \( t = 1 \) is denoted by \( P_1 \) and is quoted in terms of an underlying good (dollar), which acts as numeraire. The initial holdings of the risky asset at \( t = 1, \) given by \( X_{0i}, \) are arbitrary across the distribution of investors. All together, the investors must hold the total supply \( Q, \) therefore \( \int X_{0i} dF (i) = Q. \) For now, investors face no constraints when choosing portfolios: they can borrow and short sell freely.\(^2\)

The risky asset yields a dividend \( D \) in period 2, which is normally distributed with some mean and variance \( Var [D]. \) An investor \( i \) believes that \( D \) is normally distributed with mean \( E_i [D] \) and variance \( Var [D], \) that is:

\[
    D \sim N (E_i [D], Var [D])
\]

Investors do not learn from each other, or from the price, and agree to disagree in the Aumann (1976) sense.\(^3\) For now, the pattern of belief disagreement, which is a primitive of the model, can vary freely in the distribution of investors. Nothing prevents investors from having correct beliefs; those investors can represent market makers or (limited) arbitrageurs.

\(^2\)The extension to the case with \( R \neq 1 \) is straightforward. The online appendix deals with short-sale and borrowing constraints.

\(^3\)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 models of Eyster and Rabin (2005); Eyster, Rabin and Vayanos (2013). As long as the transaction tax does not affect the belief formation process, the results of this paper would go through unchanged, independently of the origin of investors’ beliefs.
Hedging needs  Every investor has a stochastic endowment at $t = 2$, denoted by $E_{2i}$, which is normally distributed and potentially correlated with $D$. This endowment captures the fundamental risks associated with the normal economic activity of each investor. The exposure of an investor $i$ to this underlying risk is captured by the covariance $\text{Cov} [E_{2i}, D]$. The size and magnitude of the hedging needs can vary freely in the distribution of investors. For now, investors hold correct beliefs about $\text{Cov} [E_{2i}, D]$. Without loss of generality, I assume that $\mathbb{E} [E_{2i}] - \frac{A_i}{2} \text{Var} [E_{2i}] = 0$ and normalize the initial endowment $E_{1i}$ to zero for all investors.

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

(a) Different hedging needs: captured by $\text{Cov} [E_{2i}, D]$

(b) Different risk aversion: captured by $A_i$

(c) Different initial asset holdings: captured by $X_{0i}$

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

The first three correspond to fundamental reasons 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. For positive purposes, all four reasons are equally valid: the assumed welfare criterion will make the last reason non-fundamental.

2.1.1 Policy instrument: a linear financial transaction tax

Tax definition  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 $\tau$ paid per dollar traded in the risky asset. In particular, a change in the asset holdings of the risky asset $|X_{1i} - X_{0i}|$ at a price $P_1$ faces a total tax in terms of the numeraire, due at the same time the transaction occurs, for both buyers and sellers, of:

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

where $|\Delta X_{1i}| \equiv |X_{1i} - X_{0i}|$. Total tax revenue generated by that transaction is thus $2\tau |P_1| |\Delta X_{1i}|$. I restrict $\tau$ to be in a closed interval $[\tau, \bar{\tau}]$ such that $-1 < \tau$ and $\bar{\tau} < 1$. In general, the use of the absolute value for the price in (2) is needed because asset prices can be negative in this model, as they are in many markets for derivative contracts. I restrict soon the analysis to situations with strictly positive prices.

Analytically, nothing prevents the tax to be negative (a subsidy). However, an anonymous linear trading subsidy can never be implemented: investors would continuously exchange assets, making infinite profits. Therefore, when I refer to subsidies in this paper, I implicitly assume that the planner is able to rule out this type of “wash trades”.

\footnote{Having multiple sources of fundamental trading — while not necessary — is important to show that they all enter symmetrically in optimal tax formulas.}
**Linearity and anonymity**  I restrict the analysis to linear taxes with the intention of being realistic. There is a powerful argument for the use of linear (as opposed to non-linear) taxes in this environment: linear taxes are the most robust to sophisticated trading schemes. For example, a lump-sum 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. This 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 by taxing harmful trades on an individual basis: this is an implausible assumption. More generally, any policy which allows the planner to directly target bad trades would improve welfare — see section 4.4 for a characterization of the first-best in this model and the online appendix for a more detailed discussion of policy instruments.

**No tax avoidance**  Furthermore, I assume that investors cannot avoid paying transaction taxes, either by trading secretly, within an intermediary that cancels opposites trades, or by moving to a different exchange. This behavior is optimal when the penalties from evasion are sufficiently large, provided the taxable event is appropriately defined. Those issues are critical for implementing financial transaction taxes in practice, but they are not the focus of this paper.

**Revenue rebate and lump-sum redistribution**  Lastly, since this paper focuses on the corrective (Pigovian) 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.\(^5\) Under CARA utility, the rebate that each particular investor receives is irrelevant to determine trading behavior, although variations in the individual level of the transfers change wealth and marginal utility. For clarity, I assume that every (group of) investor(s) \(i\) receives a rebate \(T_{1i}\) equal to his (their) own tax liability, that is \(T_{1i} = \tau |P_1| \Delta X_{1i}|.\) Investors do not internalize the rebate, since they are assumed to be small. It is important that tax revenue is rebated and not wasted.

The planner is able redistribute wealth across investors in a lump-sum way ex-ante. This assumption allows the planner to exclusively focus on efficiency considerations. This paper is silent on how to spend any tax revenue generated.

**Investors’ budget constraints**  Hence, consumption/wealth of a given investor \(i\) at \(t = 2\) is composed of the stochastic endowment \(E_{2i},\) the stochastic payoff of the risky asset \(X_{1i}D\) and the return on the investment in the riskless asset. This includes the net purchase or sale of the risky asset \((X_{0i} - X_{1i}) P_1,\)

\(^5\)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 only analyzes corrective taxes. Sandmo (1975) shows that corrective taxes are additive with respect to the optimal revenue raising taxes; 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 the case of transaction taxes.
the total tax liability $-\tau |P_i| |\Delta X_{1i}|$ and the lump-sum transfer $T_{1i}$. It can be written as:

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

### 2.1.2 Equilibrium definition

A competitive equilibrium with taxes is defined as a portfolio allocation $X_{1i}$ for every investor, a price $P_1$ and set of lump-sum transfers $T_{1i}$ such that: a) investors maximize expected utility in $X_{1i}$, subject to their budget constraint (3); b) the price $P_1$ is such that the market for the risky asset clears, that is $\int X_{1i} dF(i) = Q$ and c) tax revenues are rebated in a lump-sum fashion to investors, that is $\int T_{1i} dF(i) = \tau |P_1| \int |\Delta X_{1i}| dF(i)$.

For the rest of the paper, I assume that the fundamentals of the economy are such that the price of the risky asset is always positive, that is $P_1 > 0$. The online appendix provides a sufficient condition. Hence, I use $P_1$ instead of $|P_1|$. This assumption simplifies the number of cases to consider and is without loss of generality. The analysis when $P_1$ can be potentially negative is tedious but straightforward and it is available under request.

### 2.2 Equilibrium characterization

I initially solve for investors portfolio demands given prices. Subsequently, I characterize the equilibrium price and allocations.

**Investors’ problem** Because of CARA utility and normality of returns, the demand for the risky asset of every investor is derived by solving a mean-variance problem in $X_{1i}$. In particular, after getting rid of terms that do not affect the maximization problem, investors solve:

$$\max_{X_{1i}} |E_i[D] - A_i \text{Cov}[E_{2i}, D] - P_1 X_{1i} - \tau P_1 |\Delta X_{1i}| - \frac{A_i}{2} \text{Var}[D]| X_{1i}^2$$

The first order condition of this problem yields the following demand for the risky asset:\(^6\)

$$X_{1i} = \begin{cases} \frac{E_i[D] - A_i \text{Cov}[E_{2i}, D] - P_1 (1+\tau)}{A_i \text{Var}[D]}, & \Delta X_{1i} > 0 \quad \text{Buying} \\ X_{0i}, & \Delta X_{1i} = 0 \quad \text{No Trade} \\ \frac{E_i[D] - A_i \text{Cov}[E_{2i}, D] - P_1 (1-\tau)}{A_i \text{Var}[D]}, & \Delta X_{1i} < 0 \quad \text{Selling} \end{cases} \quad (4)$$

The optimal portfolio choice is represented in figure 1, which shows the net change in asset holdings $\Delta X_{1i}$ for an investor $i$ as a function of his initial asset holdings $X_{0i}$.

The presence of linear transaction taxes modifies the optimal portfolio allocation in 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 $P_1$, the tax reduces the amount of trading $|\Delta X_{1i}|$ at the intensive margin for both buyers and sellers. Second, in addition to a reduction in trading at the margin, a linear tax implies that some investors decide not to trade altogether, creating an inaction region. In

---

\(^6\)The problem solved by investors is convex, so first order conditions are necessary and sufficient to characterize optimal asset holdings. Furthermore, this fact, combined with the absence of income effects, are sufficient to guarantee existence and uniqueness of the competitive equilibrium given a tax $\tau$. 

---

8
particular, when taxes are positive, for a given price $P_1$, the no-trade region for an individual investor $i$ is given by:

$$\frac{\mathbb{E}_i[D] - A_i\text{Cov}[E_{2i}, D] - P_1 (1 + \tau)}{A_i\text{Var}[D]} < X_{0i} < \frac{\mathbb{E}_i[D] - A_i\text{Cov}[E_{2i}, D] - P_1 (1 - \tau)}{A_i\text{Var}[D]}$$  \hspace{1cm} (5)

If the initial holdings of the risky asset $X_{0i}$ are not too far from the optimal allocation without taxes $\frac{\mathbb{E}_i[D] - A_i\text{Cov}[E_{2i}, D] - P_1}{A_i\text{Var}[D]}$, an investor decides not to trade. When $\tau = 0$, the no-trade region ceases to exist.

Linear transaction costs naturally create inaction regions. The envelope theorem, which plays an important role when deriving the optimal tax results, is also key to generate 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. It is thus optimal for him not to trade. If, for instance, taxes were quadratic instead of linear, the marginal welfare loss induced by the tax around the optimum would also be second-order, eliminating the inaction region.

**Equilibrium price** Given the optimal portfolio allocation derived in (4) and the market clearing condition $\int X_1 dF(i) = Q$, the equilibrium price of the risky asset can be written as:

$$P_1 = \frac{\mathbb{E}_i[D]}{A_i}\left(\frac{\mathbb{E}_i[D]}{A_i} - A_i \left(\text{Cov}[E_{2i}, D] + \text{Var}[D] X_{0i}\right)\right) \frac{dF(i)}{1 + \tau \frac{\text{sgn}(\Delta X_{1i})}{A_i}}$$  \hspace{1cm} (6)

where $A \equiv \left(\int_{i \in \mathcal{T}} A_i dF(i)\right)^{-1}$ is the harmonic mean of risk aversion coefficients for active investors and $A_i \equiv \frac{A}{\mathcal{T}}$ is the quotient between the risk aversion coefficient of investor $i$ and the harmonic mean. I denote by $i \in \mathcal{T}$ that all integrations are made only over investors who are actively trading in equilibrium.\footnote{Equation (5) defines explicitly the set of active investors. Because the area of integration implied by $i \in \mathcal{T}$ depends on the equilibrium price, equation (6) provides an implicit characterization of $P_1$.} Intuitively, only marginal investors determine directly the equilibrium price. I use $\text{sgn} (\cdot)$ to denote the sign function, assuming that $\text{sgn} (0) = 0$. 

Figure 1: Optimal portfolio allocation/Inaction region

![Figure 1: Optimal portfolio allocation/Inaction region](image-url)
The numerator of the equilibrium price has two components. The first one is a risk aversion weighted average of the expected payoff of the risky asset. The second one 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 $\int_{i \in T} \text{Cov} [E_{2i}, D] dF(i)$. The second one is the product of the the variance of the risky asset $\text{Var}[D]$ with the number of shares initially held by investors $\int_{i \in T} X_0(i) dF(i)$.

The tax rate $\tau$ only appears directly in the denominator of the equilibrium price; it also appears indirectly in the limits of integration. When $\tau = 0$, the numerator of (6) completely determines the equilibrium price. The following lemma characterizes how the equilibrium price varies with $\tau$.

**Lemma 1. (Effect of $\tau$ on prices)** An increase in the transaction tax can increase, reduce or keep constant the equilibrium price. When $\int_{i \in T} \frac{\text{sgn}(\Delta X_i)}{A_i} dF(i)$ is positive/negative/zero, $\frac{dP_1}{d\tau}$ is negative/positive/zero.

The key determinant of $\frac{dP_1}{d\tau}$ is $\int_{i \in T} \frac{\text{sgn}(\Delta X_i)}{A_i} dF(i)$, which is the weighted difference between buyers’ and sellers’ price elasticities — as shown in the appendix. When this term is positive, increasing $\tau$ reduces the buying pressure by more than the selling pressure, reducing the equilibrium price, and vice versa. When the difference between buyers’ and sellers’ elasticities is zero — this occurs, for instance, when all investors have identical risk aversion coefficients and the mass of buyers equals the mass of sellers — the equilibrium price is independent of the tax. This is an interesting benchmark.\(^8\)

The result that equilibrium prices can increase in the level of transaction costs exists in the general equilibrium transaction costs literature. It is originally derived by Vayanos (1998) and it contrasts with the results by Amihud and Mendelson (1986), who use an arbitrage argument to show that assets subject to transaction costs/taxes must trade at a discount. Both results can be easily reconciled. If two identical assets are held in equilibrium but only one of them faces trading costs, by arbitrage, this one must trade at a lower price. However, the overall fundamental value for both assets, determined by equilibrium forces, can go up or down, as shown in lemma 1.\(^9\)

**Equilibrium portfolio allocations** The allocation of the risky asset in equilibrium can be found by substituting the equilibrium price, given by (6), into the portfolio demand by every investor, given by (4). The following property about equilibrium allocations is useful for the subsequent welfare analysis. I denote the set of active buyers by $B$ and the set of active sellers by $S$ — the identity of the investors in each of the sets varies with $\tau$.

**Lemma 2. (Effect of $\tau$ on allocations)**

- a) An increase in the transaction tax $\tau$ reduces the equilibrium asset holdings for every buyer, that is, $\frac{dX_{1i}}{d\tau} < 0$ for $i \in B$, and increases the equilibrium asset holdings for every seller, that is, $\frac{dX_{2i}}{d\tau} > 0$, for $i \in S$.

- b) An increase in the transaction tax $\tau$ reduces volume, that is, $\int_{i \in B} \frac{dX_{1i}}{d\tau} dF(i) < 0$.

\(^8\) Standard supply-demand logic — as in Mankiw (2011) — concludes that any tax reduces the equilibrium pre-tax price (which corresponds to $P_1$ in this model). Why can prices go up in a financial market? In this model, all investors determine the “demand side” of the model, so prices go up as long as the excess demand for the risky asset goes up, since the supply of assets $Q$ is fixed. Varying $\tau$ is effectively varying the excess demand curve. See the online appendix for a graphical representation.

\(^9\) Two final remarks about equilibrium prices. First, the equilibrium price $P_1$ is differentiable in $\tau$ almost everywhere. When the distribution $F$ is everywhere continuous, $P_1$ is continuously differentiable. Second, the equilibrium price is only well defined when there is trade in equilibrium. For a sufficiently large $\tau$, it is optimal for every investor to keep his initial shares; trade stops and there exists no equilibrium price.
Intuitively, when $\tau$ increases, buyers buy fewer shares of the risky asset and sellers sell fewer shares of the risky asset. This result is intuitive but not immediately obvious, because equilibrium allocations $X_{1i}(\tau, P_1)$ depend on $\tau$ both directly and through the equilibrium price. The equilibrium response of $X_{1i}$ to a change in the tax $\tau$ is thus given by:

$$
\frac{dX_{1i}}{d\tau} = \frac{\partial X_{1i}}{\partial \tau} + \frac{\partial X_{1i}}{\partial P_1} \frac{dP_1}{d\tau}
$$

(7)

As shown in (4), the first term $\frac{\partial X_{1i}}{\partial \tau}$ is negative for buyers and positive for sellers. The second term $\frac{\partial X_{1i}}{\partial P_1}$ is negative for both buyers and sellers and, as stated in lemma 1, $\frac{dP_1}{d\tau}$ can have any sign. Lemma 2 guarantees that the effect of price changes in equilibrium allocations is not strong enough, concluding that $\frac{dX_{1i}}{d\tau}$ and $\frac{\partial X_{1i}}{\partial \tau}$ have the same sign. The absence of income effects, implied by the CARA specification, allows to derive this result. The result about equilibrium volume follows directly.\(^{10}\)

**Summary of positive results**

Summing up, the positive analysis of the model concludes:

(a) Increasing transaction taxes always reduces volume in equilibrium. At the margin, every buyer buys fewer shares and every seller sells fewer shares.

(b) Increasing transaction taxes can increase, reduce or keep unchanged asset prices and returns. It follows logically that the same conclusion must apply to price volatility in a dynamic context.

The existing empirical evidence supports the positive predictions of the model. A robust empirical fact is that volume goes down (up) after an increase (reduction) in transaction taxes/costs — although tax evasion might a be confounding factor. However, the evidence of the effect on prices is mixed. Some studies find an increase in price volatility, but others find no significant changes or even a reduction. Asset prices usually fall at impact following a tax increase, but seem to recover over time. See Roll (1989), Umlauf (1993), Hau (2006), the review articles by Campbell and Froot (1994), Habermeier and Kirilenko (2003), McCulloch and Pacillo (2011) and the recent work on the European Transaction Tax by Colliard and Hoffmann (2013) and Coelho (2014) for evidence supporting lemmas 1 and 2. Buss et al. (2013) and Adam et al. (2014) also find consistent results in quantitative models.

3 **Normative analysis: optimal financial transaction tax**

After characterizing price and equilibrium allocations for a given financial transaction tax, I solve for the welfare maximizing value of $\tau$.

3.1 **Welfare criterion**

The major contribution of this paper resides in the normative analysis of the policy of taxing financial transactions. To carry out this task, I must take a stand on how to evaluate social welfare when investors

\(^{10}\)Because $X_{1i}$ is continuous in $\tau$ and $P_1$, the equilibrium allocations inherit the differentiability properties of $\frac{dP_1}{d\tau}$.  

hold heterogeneous beliefs. I approach this problem by making two assumptions about the planner’s social welfare function.

First, the planner initially characterizes the Pareto frontier in the economy by maximizing an arbitrarily weighted sum of utilities of the investors. This approach is standard and should raise no particular concerns.

Second, the planner calculates individual welfare using a single probability distribution about payoffs. This distribution will be necessarily different than the one held by most investors. In practice, I initially assume that the planner maximizes welfare using an arbitrary distribution. Subsequently, I point out under which circumstances the optimal policy does not depend on the distribution used by the planner — in those cases, the planner only assumes that there exists a single distribution of payoffs.

Two different questions emerge here. First, does the planner respect subjective beliefs when calculating social welfare? The answer to this question is negative. The planner does not respect subjective beliefs. In that sense, the approach followed by the planner is paternalistic. Second, does it matter which distribution is used by the planner to calculate welfare? Under certain (plausible) circumstances specified below, the planner does not need to impose any probability distribution to implement the optimal policy, only the logical consistency requirement of choosing a single distribution. In those cases, any criticism about paternalistic policies on the grounds that the planner must have better information than economic agents does not apply.

Belief disagreement among investors can be interpreted as a device to model departures from full rationality in information processing. Instead of modeling a specific rationality failure, this paper takes the distribution of beliefs as a primitive. Consistent with that interpretation, there are two arguments that justify the welfare criterion adopted in this paper. The first one relies on the idea that rational investors cannot agree to disagree when their posteriors are common knowledge — see the discussion in Morris (1995). How can the planner respect investors’ beliefs when they are inconsistent with one another? If we assume that there is a single correct belief but different investors hold different beliefs, all of them (but one) must be wrong. Alternatively, a veil of ignorance interpretation is also consistent with this welfare criterion. If investors acknowledge that they may wrongly hold different beliefs, they would happily implement, at an ex-ante stage, a tax policy that curtails trading.

Note that, given his own ex-post belief, leaving aside price changes, a given investor believes that a transaction tax reduces his individual welfare. However, an altruistic investor will always vote for a positive tax. Intuitively, an altruistic investor perceives that a small tax has a second-order loss for him but it creates a first-order gain for every other investor in the economy. There is scope to explore in detail political economy considerations in this environment, although I abstract from those in this paper.

---

11This is a controversial issue. In addition to the work discussed on the literature section, see Kreps (2012), Cochrane (2014) and Duffie (2014) for some reflections on this topic. Duffie (2014), in particular, challenges policy treatments of speculative trading motivated by differences in beliefs. He raises philosophical/axiomatic challenges and a practical challenge. The results of this paper directly address the practical challenge, which questions the ability of enforcement agencies to distinguish between belief-motivated trade and trade motivated by welfare enhancing activities.
3.2 Optimal tax policy

After presenting the welfare criterion used by the planner, I characterize the optimal tax. Social welfare, denoted by $V$, can thus be written as the sum of the indirect utility for every investor $i$, under the distribution of returns used by the planner, that is:

$$V(\tau) = \int \lambda_i V_i dF(i),$$

where $\lambda_i > 0$ is the welfare weight assigned by the planner to an investor $i$.

$V_i$ denotes the indirect utility for an investor $i$ from the planner’s perspective, which can be expressed as:

$$V_i \equiv \mathbb{E} [U_i(W_{2i})] = -e^{-A_i} \hat{V}_i \quad \text{with} \quad \hat{V}_i \equiv \mathbb{E} [W_{2i}(X_{1i}, P_1)] - \frac{A_i}{2} \text{Var} [W_{2i}(X_{1i}, P_1)],$$

where $W_{2i}$ is a function of the portfolio allocation optimally chosen by investors $X_{1i}$ and the equilibrium price $P_1$. $\hat{V}_i$ is the certainty equivalent from the planner’s perspective. Note that the expectation used to calculate the indirect utility $V_i$ does not have an individual subscript $i$, because it is taken under the planner’s distribution.

Substituting $W_{2i}$ explicitly and using the fact that tax proceeds are rebated lump-sum to every investor, i.e., $T_{1i} = \tau P_1 |\Delta X_{1i}|$, the indirect utility of an investor $i$ from the planner’s perspective can be written as:

$$V_i = -e^{-A_i} \left( (\mathbb{E}[D] - A_i \text{Cov}[E_{2i}, D] - P_1) X_{1i} + P_1 X_{1i} - \frac{A_i}{2} \text{Var}[D](D)(X_{1i})^2 \right),$$

where $X_{1i}$ is chosen by every investor according to the first order condition described in (4). Note that both $X_{1i}$ and $P_1$ are in principle functions of $\tau$.

The optimal tax $\tau^*$ is defined by $\tau^* = \arg \max_{\tau} V(\tau)$. The first step to characterize $\tau^*$ is to find $\frac{dV_i}{d\tau}$, the marginal effect of varying taxes on the indirect utility of investor $i$ from the planner’s perspective. This is given by:

$$\frac{dV_i}{d\tau} = A_i e^{-A_i} \hat{V}_i \frac{d\hat{V}_i}{d\tau} = \mathbb{E} [U_i'(W_{2i})] \frac{d\hat{V}_i}{d\tau}$$

The value of $\frac{dV_i}{d\tau}$ is given by the (dollar) change in the certainty equivalent $\frac{d\hat{V}_i}{d\tau}$, valued at the expected marginal utility of wealth from the planner’s perspective $\mathbb{E} [U_i'(W_{2i})]$. The marginal change in the certainty equivalent $\frac{d\hat{V}_i}{d\tau}$ can be expressed as:

$$\frac{d\hat{V}_i}{d\tau} = \left[ (\mathbb{E}[D] - \mathbb{E}_i[D]) + \text{sgn}(\Delta X_{1i}) P_1 \tau \right] \frac{dX_{1i}}{d\tau} - \Delta X_{1i} \frac{dP_1}{d\tau}$$

(8)

The expression for $\frac{d\hat{V}_i}{d\tau}$ consists of two terms.\footnote{Note that the derivation of (8) uses the envelope theorem for the choice of $X_{1i}$ and the extensive margin choice between trading and not trading, which are both made optimally. Note also that when an investor decides not to trade, the marginal change $\frac{dX_{1i}}{d\tau}$ is zero and there are no terms-of-trade-effects, because $X_{1i} = X_{0i}$. Hence, a marginal tax change has no effect at all on the welfare of those investors who decide not to trade.} The first term captures how welfare varies through changes in the portfolio allocation. Intuitively, a change in the allocation $\frac{dX_{1i}}{d\tau}$ changes the certainty equivalent of investor $i$ only when there exists a wedge in his portfolio demand. In this case, a first wedge arises when $\mathbb{E}[D] \neq \mathbb{E}_i[D]$, due to the belief distortion. A second wedge arises when the tax is strictly positive, that is, when $\text{sgn}(\Delta X_{1i}) P_1 \tau \neq 0$. In a model with homogenous beliefs and no taxes, a
marginal increase in $\tau$ has no effect on welfare through changes in the portfolio allocation. The second term in $\frac{dV_i}{d\tau}$ captures terms-of-trade effects. If $P_1$ increases with $\tau$, the buyers of the risky asset are worse off, since they now buy a more expensive asset: the terms-of-trade of their purchase have worsened. However, sellers are better off, since they receive a higher price for their sale. The opposite occurs when prices decrease with $\tau$.

**Assumptions** Before characterizing the main results of this section, I introduce assumptions [NR] and [OBPS].

**Assumption.** [NR] *(No Redistribution)* The planner has access to ex-ante lump-sum transfers across investors. This condition guarantees that $\lambda_i E\left[U_i'(W_{2i})\right] = \delta, \forall i$, where $\delta > 0$, for any distribution of welfare weights $\lambda_i$ and for any value of $\tau$.

Assumption [NR] allows the planner to separate redistributional considerations from efficiency considerations. It is equivalent to assuming quasilinear preferences in a commodity taxation problem and, in this environment, it is equivalent to assuming that the planner maximizes investors’ certainty equivalents. This is a standard assumption in models of corrective taxation with concave utility — see, for instance, Atkinson and Stiglitz (1980).

**Assumption.** [OBPS] *(Optimists are mostly Buyers/Pessimists are mostly Sellers)* The cross-sectional covariance of expected payoffs $E_i[D]$ with respect to marginal changes in equilibrium allocations $\frac{dX_{1i}}{d\tau}$ at $\tau = 0$ is negative, that is:

$$\text{Cov}_F \left( E_i[D], \left. \frac{dX_{1i}}{d\tau} \right|_{\tau=0} \right) = \int E_i[D] \left. \frac{dX_{1i}}{d\tau} \right|_{\tau=0} dF(i) < 0,$$

where $\text{Cov}_F[\cdot]$ denotes the covariance in the cross-section of the distribution of investors and the equality follows by using market clearing.

Assumption [OBPS] holds when optimistic investors happen to be (on average) the buyers of the risky asset and pessimistic investors happen to be the sellers. If all trading is driven by belief disagreement, assumption [OBPS] always holds — in that case, optimists buy and pessimists sell. However, because investors also trade due to fundamental reasons, it is possible that optimistic investors happen to be net sellers in equilibrium. For example, this situation occurs when workers are overoptimistic about the returns of their own companies and do not sufficiently hedge their implicit labor market risk.\(^{13}\) Assumption [OBPS] implies that this phenomenon is not too prevalent in the economy without taxes.

As long as there is some belief disagreement in this economy, we expect assumption [OBPS] to hold purely on theoretical grounds — the appendix provides a precise theoretical argument. 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, we expect this covariance to be negative. Alternatively, if fundamental trading is orthogonal to non-fundamental

---

\(^{13}\) Benartzi (2001), Massa and Simonov (2006) and Døskeland and Hvide (2011) provide evidence of this phenomenon. Other behavioral biases — see Campbell (2006) — may give rise to the same pattern.
trading, we expect [OBPS] to hold.\footnote{Note that [OBPS] is an assumption on the primitives of the economy. I do not substitute for \( \frac{dX_i}{d\tau} \) because it is not particularly illustrative. A simple example can give further intuition for how assumption [OBPS] may arise. Assume that all buyers hold identical beliefs \( \mathbb{E}_B[D] \), all sellers hold identical beliefs \( \mathbb{E}_S[D] \) and also assume that \( \mathbb{E}_B[D] > \mathbb{E}_S[D] \). In that case: \( \text{Cov}_F \left( \mathbb{E}_i[D], \frac{dX_i}{d\tau} \right) = (\mathbb{E}_B[D] - \mathbb{E}_S[D]) \int_{i \epsilon B} \frac{dX_i}{d\tau} dF(i) < 0 \), since the first term is positive by assumption and the second one negative by lemma 2.}

The body of evidence accumulated in the behavioral finance literature — as surveyed by Hong and Stein (2007) or Barber and Odean (2013) — suggesting that belief distortions are an important driver of the observed trading volume in financial markets would also yield empirical support to assumption [OBPS].

**Main results** Propositions 1 and 2 present the main results of this paper.

**Proposition 1. (Effect of a marginal tax change on welfare)**

a) The marginal change in social welfare induced by varying the financial transaction tax \( \tau \) is given by:

\[
\frac{dV}{d\tau} = \int \lambda_i \mathbb{E}[U_i^f(W_{2i})] \begin{bmatrix} \mathbb{E}[D] - \mathbb{E}_i[D] + \text{sgn}(\Delta X_{1i}) P_1 \tau \frac{dX_i}{d\tau} - \Delta X_{1i} \frac{dP_1}{d\tau} \end{bmatrix} dF(i) \tag{9}
\]

b) Under assumption [NR], the marginal change in social welfare induced by varying the financial transaction tax becomes:

\[
\frac{dV}{d\tau} = \int [-\mathbb{E}_i[D] + \text{sgn}(\Delta X_{1i}) P_1 \tau] \frac{dX_i}{d\tau} dF(i) \tag{10}
\]

c) Under assumptions [NR] and [OBPS], the marginal change in social welfare induced by introducing a very small tax is strictly positive. Locally around \( \tau = 0 \), there is a first-order gain from reducing trades driven by belief disagreement but there is a second-order loss from reducing fundamental trades, that is:

\[
\left. \frac{dV}{d\tau} \right|_{\tau = 0} = -\int \mathbb{E}_i[D] \left. \frac{dX_i}{d\tau} \right|_{\tau = 0} dF(i) > 0
\]

Equation (9) is the weighted sum of the effect of the tax in the indirect utilities of all market participants. The weight given by the planner to the changes in the certainty equivalent of every investor is a combination of the social welfare weight and the (expected) marginal utility. Note that all welfare losses due to belief disagreement ultimately arise because investors hold too much or too little of the risky asset: their risk-return tradeoff is distorted.

After imposing assumption [NR], two terms drop out from (9). First, terms-of-trade effects cancel out in the aggregate after imposing market clearing. Intuitively, changes in equilibrium prices are simply a transfer of resources, which are irrelevant for a planner without a redistributive goal in a static setup. Second, the term containing the planner’s belief \( \mathbb{E}[D] \) also cancels out after imposing market clearing. Because of its conceptual importance, I state this result as a corollary of proposition 1.

**Corollary. (Irrelevance of planner’s belief)** The optimal financial transaction tax does not depend on the distribution of beliefs used by the planner to calculate welfare.
The fact that the optimal transaction tax is identical for any belief chosen by the planner to calculate welfare is an important takeaway of this paper. In that case, there is no need to rely on criteria that use a convex combination of beliefs, like Brunnermeier, Simsek and Xiong (2014), or worst case scenarios, like Blume et al. (2013). Any belief used by the planner implements the same optimal policy.

Two characteristics of the economic environment are essential to derive this corollary. First, 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 less, that is, \( \int \frac{dX_{1i}}{d\tau} dF(i) = 0 \). In that case, only relative asset holdings matter for welfare. Second, the planner does not want to redistribute wealth across investors, that is, \( \lambda_i \mathbb{E} \left[ U_i' \left( W_{2i} \right) \right] \) is constant. A planner who redistributes resources overweights the distortions of investors with higher welfare weights, making the belief he uses important because tax-induced changes in allocations do not cancel out.

I emphasize the local result from proposition 1c because of its generality. A small transaction tax reduces equally fundamental and non-fundamental trades. However, as long as a fraction of investors, no matter how small, have heterogeneous beliefs, we expect [OBPS] to hold, which implies that the marginal reduction in volume created by a tax unequivocally increases social welfare.

I now characterize the optimal financial transaction tax.\(^{15}\)

\[ \text{Proposition 2. (Optimal financial transaction tax)} \]

a) Under assumption [NR], the optimal financial transaction tax is characterized by:

\[
\tau^* = \frac{\int \frac{E_i[D]}{P_1} \frac{dX_{1i}}{d\tau} dF(i)}{\int \text{sgn} \left( \Delta X_{1i} \right) \frac{dX_{1i}}{d\tau} dF(i)} = \frac{\Omega_B - \Omega_S}{2},
\]

where \( \Omega_B \) is a weighted average of buyer’s expected returns, defined by:

\[
\Omega_B \equiv \int_{i \in B} \omega^B_i E_i[D] \frac{dX_{1i}}{d\tau} dF(i), \quad \text{with} \quad \omega^B_i \equiv \frac{\int_{i \in B} \frac{dX_{1i}}{d\tau} dF(i)}{\int_{i \in B} \frac{dX_{1i}}{d\tau} dF(i)},
\]

and \( \Omega_S \) is analogously defined for sellers.

b) Under assumptions [NR] and [OBPS], the optimal financial transaction tax is strictly positive. If assumption [OBPS] does not hold, the optimal policy could be a positive, negative or zero tax.

The expression for \( \tau^* \) optimally trades off the distortion induced by belief disagreement (numerator) against the loss in fundamental trading (denominator). The numerator of (11) can be written as the cross sectional covariance for active investors of expected returns with equilibrium portfolio tax sensitivities, denoted by \( \text{Cov}_{F,T} \cdot \cdot \cdot \cdot \cdot \) times the fraction of active investors, denoted by \( \zeta(\tau) \equiv \int_{i \in T} dF(i) > 0 \), that is:

\[
\int \frac{E_i[D]}{P_1} \frac{dX_{1i}}{d\tau} dF(i) = \zeta(\tau) \text{Cov}_{F,T} \left[ \frac{E_i[D]}{P_1}, \frac{dX_{1i}}{d\tau} \right]
\]

When this covariance is negative, \( \tau^* \) is positive and vice versa. This occurs because the denominator, which can be written as twice the marginal change in trading volume, that is:

\[
\int \text{sgn} \left( \Delta X_{1i} \right) \frac{dX_{1i}}{d\tau} dF(i) = 2 \int_{i \in B} \frac{dX_{1i}}{d\tau} dF(i),
\]

\(^{15}\)Because the planner’s problem is not convex in \( \tau \), first order conditions — like equation (11) — are necessary but not sufficient to characterize the optimal \( \tau^* \) when \( V \) is differentiable. See section 3.5.2 for a more detailed discussion.
is strictly negative. Because prices, demand responses and limits of integration are functions of taxes, equation (11) provides an implicit representation for $\tau^*$. This is a general feature of optimal taxation exercises.

Intuitively, the magnitude of the optimal tax depends on how much buyers’ beliefs differ from sellers’ beliefs, which is a measure of excess trading. The planner overweighs the beliefs of investors with higher sensitivities $\frac{dX_{1i}}{d\tau}$, because they are the most responsive to tax policy. Note that the weights assigned to active traders $\omega_i^B$ and $\omega_i^S$ add up to one, and the weights of investors who do not trade are exactly zero.

**Sufficient statistics**  It is remarkable that only the joint distribution of investors’ beliefs and investors’ equilibrium portfolio tax sensitivities directly determine the optimal tax, independently of the rest of the structure of the model. In particular, the cross-sectional covariance between investors’ beliefs and equilibrium portfolio tax sensitivities becomes the relevant sufficient statistic to determine the optimal tax.\(^{16}\) This result is due to the Pigovian/corrective nature of the optimal tax: Pigovian/corrective taxes are set to target marginal distortions, which in this case are investors’ beliefs.

When all portfolio tax sensitivities are identical in magnitude, that is, $|\frac{dX_{1i}}{d\tau}|$ is constant, $\omega_i^B$ and $\omega_i^S$ become constants. In that case, $\tau^*$ only depends on the difference between the average beliefs of buyers and the average beliefs of sellers. Hence, those two moments of the distribution of beliefs turn out to be the single sufficient statistics for the determination of the optimal tax, since portfolio tax sensitivities drop out of the optimal tax formula. This is an important benchmark for a hypothetical implementation of a tax, since it has minimal informational requirements for the planner. The online appendix contains a parametric example for the distribution of beliefs (normally distributed). In that case, the optimal tax is an increasing monotonic function of the variance of the distribution of beliefs.

### 3.3 Remarks

I summarize several implications of propositions 1 and 2 in a series of remarks.

**Remark 1.** *The total amount of fundamental trading only determines the optimal tax through the identity of the marginal investors.*

It may be natural to think that the optimal tax formula should feature a tradeoff between measures of fundamental trading — like the absolute value of the covariance terms $\text{Cov}[E_{2i}, D]$ or the dispersion among investors risk aversion coefficients — versus measures of non-fundamental trading. This logic is flawed. Intuitively, the optimal tax only depends on the marginal distortions on trading, i.e. $\frac{dX_{1i}}{d\tau}$, but not on the total amount of trading.

However, fundamental trading does affect the optimal tax through the limits of integration — which determine, on the extensive margin, who are the marginal investors in equilibrium for every value $\tau$ — and potentially through portfolio tax sensitivities. In other words, a marginal change in beliefs $E_i [D]$ for active investors always changes the value of the optimal tax. However, a marginal change in measures

\(^{16}\)The sufficient statistic terminology is often used to refer to measurable variables. Whether beliefs can be inferred from data remains an open question. The sufficient statistic logic is similar to the one behind the CAPM, in which the beta of an asset becomes sufficient to determine expected returns. It is also similar to the logic behind consumption based asset pricing models, in which the consumption process, independently of how it is generated, is sufficient to determine asset prices and expected returns.
of fundamental trading only modifies the optimal tax by varying the composition or the tax sensitivities of investors who actively trade for a given $\tau$.

Remark 2. Investors who decide not to trade are inframarginal for the determination of the optimal tax.

Only the beliefs and portfolio tax sensitivities of the marginal investors matter for the determination of the optimal tax in equilibrium. Intuitively, the envelope theorem also acts at the margin of trading versus not trading, so those investors who optimally decide not to trade become inframarginal from an optimal taxation perspective. Hence, the optimal tax accounts for the fact that some trades cease to occur when transaction taxes are imposed.

Quadratic taxes — which are often used as a tractable approximation for linear taxes — do not allow for extensive margin adjustments, because all investors are marginal for all values of $\tau$. Hence, quadratic taxes fail to capture potential non-convexities in the planner’s problem, as discussed in section 3.5.2.

Remark 3. The dispersion of beliefs across investors, regardless of their average, is what determines the optimal tax.

For instance, if all investors agree about the expected payoff the risky asset, that is, $E_i[D]$ is constant, the optimal tax is $\tau^* = 0$, independently of whether their average belief is different from the belief used by the planner. This situation occurs naturally when all investors hold correct beliefs about the distribution of $D$, the standard common prior/full rationality result. However, if all investors were equally mistaken, the optimal tax would still be zero. Only belief dispersion matters because the equilibrium portfolio allocation, which is what ultimately matters for welfare, only depends on the relative dispersion of beliefs.

Remark 4. The optimal policy can be a subsidy, that is, $\tau^* < 0$.

Although the exposition focuses on the $\tau^* > 0$ case — the relevant case under [OBPS] —, 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. Intuitively, if investors trade too little, it is optimal for the planner to encourage them to trade more with a subsidy.

Remark 5. Changes in trading volume, and not price changes, are the welfare relevant sufficient statistics.

Although many informal arguments surrounding transaction taxes focus on their effects on prices, changes in social welfare must be eventually traced back to changes in allocations/volume. From a welfare standpoint, the correct question is not whether prices are right, but whether every investor is holding the right amount of risk.

Remark 6. The restriction to a single linear tax implies that portfolio tax sensitivities $\frac{dX_i}{d\tau}$ determine the weights given to individual beliefs in the optimal tax formula.

Classic Pigovian logic suggests that an optimal corrective tax only depends on the magnitude of the distortion but not on elasticities. However, the optimal tax in (11) depends on portfolio tax sensitivities. This occurs because the planner uses a second-best policy instrument. Intuitively, the planer would like to eliminate every belief distortion individually. However, because he has a single policy instrument, it turns out to be optimal to weigh distortions according to demand sensitivities. The planner weighs more the distortions of the most responsive investors.

The presence of demand sensitivities in optimal corrective tax formulas goes back to Diamond (1973), who analyzes corrective taxation with restricted policy instruments in a model of consumption
externalities. My results are directly related to the separable case analyzed in that paper.\textsuperscript{17}

### 3.4 Numerical simulations

To provide further intuition, I illustrate the previous results with two numerical examples. In both examples, I make some further assumptions, summarized in table 1. First, all investors have the same absolute risk aversion coefficient $A_i = 1$ and the same initial holdings of the risky asset $X_{0i} = Q = 1$. Therefore, hedging needs become the only source of fundamental trading. Second, the planner uses welfare weights $\lambda_i = 1$ and [NR] holds, so investors’ welfare is given by their certainty equivalents. Third, the standard deviation of the risky asset is 4 and the planner calculates welfare using $E[D] = 100$.

$$
\begin{align*}
E[D] &= 100 \\
Var[D] &= 16 \\
\lambda_i &= 1, \forall i \\
E[U'_i(W_2)] &\text{ constant, } \forall i \\
X_{0i} &= Q = 1, \forall i
\end{align*}
$$

Table 1: Parameter choices for both examples

Although I use reasonable values, a serious calibration is beyond the scope of this paper.

**Example 1. (Only optimists and pessimists/No fundamental trading)** There are two groups of investors. A fraction $\pi_H$ of them is of type $i = H$ and the remainder fraction $\pi_L = 1 - \pi_H$ is of type $i = L$, where $H$ and $L$ correspond to high (optimists) and low (pessimists) valuations for the risky asset. The rest of parameters are in Table 2.

<table>
<thead>
<tr>
<th>Optimistic Buyers</th>
<th>$\pi_H = 0.5$</th>
<th>$E_H[D] = 106$</th>
<th>$Cov[E_{2H}, D] = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pessimistic Sellers</td>
<td>$\pi_L = 0.5$</td>
<td>$E_L[D] = 96$</td>
<td>$Cov[E_{2L}, D] = 0$</td>
</tr>
</tbody>
</table>

Table 2: Parameters figure 2 (example 1)

Under this parametrization there is no fundamental trading. The difference in expected payoffs with respect to the planner by optimists corresponds to 1.5 standard deviations and the one by pessimists to a single standard deviation. Figure 2 shows welfare relative to the $\tau = 0$ case, equilibrium allocations, price and volume for different values of $\tau$.\textsuperscript{18} The optimal tax is $\tau^* = 5.88\%$ and the welfare gain at the optimum is 0.86\%. As emphasized in remark 1 above, starting from this parametrization, an increase in the amount of fundamental trading would not change the optimal tax. Investors would start to trade, but the optimal tax would not change.

Because buyers and sellers are symmetric in all dimensions but beliefs, the equilibrium price is constant for any level of the tax. Volume decreases in the level of the tax, as expected. When $\tau > 5.88\%$, trading stops and all investors simply keep their initial holdings of the risky asset. It is not a coincidence.

\textsuperscript{17}Most of the results in Diamond (1973) focus on how the optimal corrective tax should account for demand inter-linkages. Those issues never arise in this paper, because the portfolio allocation of an investor $i$ does not depend directly (only through prices) on the beliefs of other investors.

\textsuperscript{18}The top left plot in figures 2 and 3 shows the indirect utility for each investor from the planner’s perspective $V_i$ and social welfare $V$ in percent deviations from the $\tau = 0$ value. The bottom left one shows the equilibrium price. The top right plot shows the equilibrium levels of trade for each investor and the bottom right one shows the total amount of trade in equilibrium, given by $\int_{c \in B} X_{1i} dF(i)$. The vertical purple dashed lines represent tax levels at which at least one of the investors stops trading. The vertical black dashed line represents the optimal tax level.
that the optimal tax is the smallest one which eliminates trade. More generally, it is easy to show that if all trades are due to disagreement, the optimal tax completely shuts down trading.

The welfare of the optimists is monotonically increasing in the tax. However, the welfare of pessimists peaks at a tax rate $\tau = 4.71\%$ and then it decreases. Note that the equilibrium price is “too high” from the planner’s viewpoint, since $\frac{E_H[D] + E_L[D]}{2} = 101 > E[D] = 100$, which favors sellers. Intuitively, pessimists, who are perceived as less mistaken than optimists, are making a good deal — from the planner’s perspective — when selling their shares to optimists at a high price. However, the optimal tax disregards these considerations: proposition 2 shows that social welfare is independent of the planner’s belief.

**Example 2. (All types of investors)** This new example introduces a group of well calibrated buyers and a group of pessimistic sellers. Now, there is a fraction $\pi_{HA}$ of optimists who buy the asset with no hedging motive and a fraction $\pi_{HB}$ who happen to sell the risky asset for risk sharing reasons. A fraction $\pi_{C}$ of calibrated investors must buy the risky asset for risk sharing reasons and the remainder fraction $\pi_{L}$ of investors are pessimists with no hedging motive. Table 3 describes the parametrization.
These parameters imply that roughly 65% of the trading volume at $\tau = 0$ is fundamental and the remaining 35% is due to disagreement. The optimal tax is $\tau^* = 2.01\%$ and the welfare gain at the optimum is 0.11%. This example shows how asset prices can fall or rise with larger taxes. Volume and portfolio allocations behave as expected.

Around $\tau = 0$, the slope of the indirect utility for a given investor can be positive, negative or zero, even when prices are locally constant. In particular, investors $HA$ and $L$ have positive slopes, which implies that they experience a first-order gain from having a positive tax — this feature is identical to the previous example. The slope of the calibrated investor $C$, who only trades for fundamental reasons, is flat, since he only experiences a second-order loss. And the slope of investor $HB$ is strictly negative; any tax induces a first-order loss for him, since it moves him further away from his optimum, starting from a point at which he was not already optimizing from the planner’s perspective. This is the effect that, when sufficiently prevalent — when [OBPS] does not hold — would call for a subsidy as the optimal
policy. That said, because [OBPS] holds in this example, the optimal tax is still positive.

3.5 Further results

3.5.1 Harberger revisited

The results derived so far rely on the assumption that the planner maximizes welfare using a single belief. However, it is straightforward to quantify the welfare loss induced by a tax increase assuming that all investors hold correct beliefs or that the planner assess social welfare respecting individual beliefs. Under either of these assumptions, all trades are regarded as fundamental, so any tax induces a welfare loss. I derive a result analogous to Harberger (1964).

**Proposition 3. (Harberger (1964) revisited)**

a) When investors hold correct beliefs or the planner respects individual beliefs when calculating social welfare, the marginal welfare loss generated by increasing the transaction tax at a level $\tau$, expressed as a money-metric (in dollars) at $t = 1$, is given by:

$$\int \frac{d\hat{V}_i}{d\tau} \bigg|_{\tau=\tilde{\tau}} dF (i) = 2\tilde{\tau} P_1 \int_{i \in B} \frac{dX_{1i}}{d\tau} \bigg|_{\tau=\tilde{\tau}} dF (i) \leq 0,$$

(12)

where $i \in B$ denotes that the integration is made only over the set of buyers and $\hat{V}_i$ denotes investors’ certainty equivalents.

b) The marginal welfare loss of a small tax change around $\tau = 0$ can be approximated, using a second order Taylor expansion, by:

$$dV = \int d\hat{V}_i|_{\tau=0} dF (i) \approx \tilde{\tau}^2 P_1 \int_{i \in B} \frac{dX_{1i}}{d\tau} \bigg|_{\tau=0} dF (i)$$

(13)

This result provides a measure of welfare losses as a function of observables for any tax intervention. Given the money-metric correction, investors in this economy are willing to pay $\mathcal{L} (\tau)$ dollars to prevent a change in the tax rate. Note that this happens to correspond to the marginal change in revenue raised. Equation (12) derives an upper bound for the size of the welfare losses induced by taxation in the case in which all trades are deemed to be fundamental.

Equation (12) resembles the classic Harberger (1964) result about welfare losses in the context of commodity taxation. However, the welfare loss in this case is given by twice the size of the tax, because the portfolio holdings of both buyers and sellers are distorted. Taxing a commodity distorts the amount consumed of a given good, reducing welfare. Taxing financial transactions distorts portfolio allocations, inducing investors to hold more or less risk than they should, also reducing welfare. The idea that the distortion created by a tax (approximately) grows with its square is also associated to Harberger (1964) — equation (13) presents the equivalent result in the context of this model.

19 Although this result is intuitive, to my knowledge, it had not been derived before in the context of a portfolio choice problem. See Auerbach and Hines Jr (2002) for a comprehensive analysis of tax efficiency results and Sandmo (1985) for a survey of results on how taxation affects portfolio allocations.
3.5.2 Non-convexity of the planner’s problem

I now describe and interpret the convexity properties of the planner’s problem. In general, there is no guarantee that the planner’s problem is convex on \( \tau \) for a general distribution \( F \). This may seem surprising, given that the problem solved by investors is well behaved and that there are no complementarities or income effects at play. The non-convexity of the planner’s problem is due to changes in the composition of marginal investors for different values of \( \tau \). It may well be the case that both a very small tax, with minor effects, and a very large one, which would reduce volume greatly, yield similar welfare levels.

**Proposition 4. (Non-convex planner’s problem)**

a) The planner’s problem is in general non-convex.

b) The non-convexity only arises when the composition of active investors in the economy varies with \( \tau \).

The best way to understand this result is graphically. Figure 4 shows an example of a very non-convex problem. In this extreme example, there are six investors who trade for both fundamental and non-fundamental reasons – their characteristics are in table 4. In this particular example, [OBPS] does not hold, but the optimal tax happens to be positive — this can only happen if the planner’s problem is non-convex.

<table>
<thead>
<tr>
<th>( \pi_i )</th>
<th>0.2</th>
<th>0.2</th>
<th>0.2</th>
<th>0.2</th>
<th>0.1</th>
<th>0.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_i )</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( \text{Cov} [E_{2i}, D] )</td>
<td>-20</td>
<td>20</td>
<td>50</td>
<td>-50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \mathbb{E} [D] )</td>
<td>100</td>
<td>100</td>
<td>140</td>
<td>60</td>
<td>161</td>
<td>39</td>
</tr>
</tbody>
</table>

Table 4: Parameters figure 4

It is easy to see that, as long as the identity of the marginal investors does not change (between any two vertical purple dashed lines), social welfare is convex. The fact that the planner’s problem is non-convex has an important economic interpretation. Intuitively, a planner should always think whether a tax increase improves social welfare given the set of marginal investors. For instance, imagine that a group of fundamental investors cease to trade after a tax increase. If the remaining investors trade mostly for non-fundamental reasons, the planner would now have a further incentive to increase taxes, since now there are relatively fewer fundamental investors being distorted at the margin by a tax increase. In practical terms, this logic can be important for high frequency investors, which are very affected by transaction taxes — as shown in section 6 — and can become inframarginal even for very small levels of \( \tau \). This non-convexity can only arise with linear taxes; quadratic taxes, which are often used as an approximation, necessarily make the problem solved by the planner convex.

---

20 I relegate the study of the differentiability properties to the appendix, since it has little economic content. Essentially, if the distribution of investors is continuous, the planner’s problem is differentiable. Only for simplicity, all examples in this paper contain a finite number of investors, which creates kinks on social welfare, but those kinks can be easily smoothed out.

21 The top left plot in figure 4 shows the equilibrium price and the top right one the equilibrium allocations for each investor. The bottom left one shows social welfare (measured in units of the numeraire at \( t = 0 \)) and the bottom right one shows the value of the marginal change in social welfare \( \frac{dV}{d\tau} \). The vertical purple dashed lines represent tax levels at which at least one of the investors stops trading. The vertical black dashed line represents the optimal tax level.

23
3.5.3 An intermediate target implementation

As described above, the distribution of beliefs is the key object that determines the optimal tax. Given that it is not obvious how to recover investors’ beliefs, an alternative route to implement the optimal policy is setting the optimal tax to target a certain level of trading volume. Under some conditions, adjusting the tax rate until total volume equals fundamental volume allows the planner to implement the optimal policy.

**Proposition 5. (Intermediate target implementation)**

a) Total volume can be decomposed in the following way:

\[
\int_{i \in B} \Delta X_{1i} dF(i) = \Theta_F + \Theta_{NF} - \Theta_{\tau},
\]

where \( \Theta_F, \Theta_{NF} \) and \( \Theta_{\tau} \) are defined in the appendix.

b) The optimal tax policy can be implemented by guaranteeing that total volume equals fundamental volume, i.e., \( \Theta_F = \Theta_{NF} \), when \( \int_{i \in B} \frac{1}{A_i} dF(i) = \int_{i \in S} \frac{1}{A_i} dF(i) \), which is equivalent to assume that \( \frac{dP_1}{d\tau} = 0 \).

---

I use the intermediate target nomenclature by analogy with the literature on optimal monetary policy. In this model, equilibrium volume becomes an intermediate target to implement optimal portfolio allocations.
Proposition 5a decomposes total volume into three sources. The first one is volume driven by differences in endowments, risk aversion, and hedging needs. I call this source fundamental trading. The second one is volume driven by belief disagreement. I call this sources non-fundamental trading. The third one is the reduction in volume induced by the transaction tax. I refer to this term as the tax induced trading distortion. Proposition 5b shows that, if the planner can use a model that credibly predicts the amount of fundamental trading and takes the view that residual volume is non-fundamental, it can adjust the optimal tax until observed volume corresponds with the theoretical amount of fundamental trading.\footnote{This approach has an important downside. If the model used by the planner is mispecified and fails to capture fundamental reasons for trading, the attributed amount of non-fundamental trading and the level of the induced optimal tax will be too high.}

Although enlightening, the intermediate target implementation is less general than the expression for the optimal tax in proposition 2. The optimal tax result is more general — at least as an approximation, as shown in section 4.3 — because it targets beliefs directly, which are the source of marginal distortions: that is the Pigovian principle. The total amount of volume is an infra-marginal object which relies more heavily on the structure of the model; for instance, the decomposition result is only valid in the CARA-Normal environment and the implementation of the optimal policy through volume further requires that the transaction tax does not modify equilibrium prices.

## 4 Static extensions

Within the CARA-Normal static framework, I now analyze, one at a time, several extensions of the baseline model. To ease the exposition, I only describe the essential differences of every extension with respect to the baseline model and use assumption [NR] throughout. I also omit definitions of equilibrium — they are all standard — and most regularity conditions. Finally, I only focus on $\tau^*$, omitting the discussion of $\frac{dV}{d\tau}$.

### 4.1 Pre-existing trading costs

Because actual investors face trading costs even when there are no taxes, it is natural to think that the results derived around the point $\tau = 0$ are never valid. This logic is flawed. Propositions 1 and 2 still apply as long as transaction costs are a mere compensation for the use of economic resources.

**Assumptions**  Investors now face transaction costs, regardless of the value of $\tau$. These represent costs associated with trading, like brokerage commissions, exchange fees or bookkeeping costs. Investors must pay a quadratic cost, parametrized by $\alpha$, a linear cost $\eta$ on the number of shares traded and a linear cost $\psi$ on the dollar volume of the transaction.\footnote{We do not analyze fixed costs for technical reasons. Fixed costs introduce non-convexities, which would substantially change the equilibrium characterization, as in Garleanu, Panageas and Yu (2013).} These trading costs are paid to a new group of agents (intermediaries) which facilitate the process of trading. Crucially, I assume that intermediaries make zero profits in equilibrium. Hence, wealth at $t = 2$ for an investor $i$ is now given by:

$$W_{2i} = E_{2i} + X_{1i}D + \left( X_{1i}P_1 - X_{1i}P_1 - |\Delta X_{1i}| \right) \left( \tau + \psi - \eta |\Delta X_{1i}| - \frac{\alpha}{2} (\Delta X_{1i})^2 + T_{1i} \right)$$

The transfer exclusively rebates to the investor the amount paid as a tax, that is, $T_{1i} = \tau |P_1| |\Delta X_{1i}|$.\footnote{This approach has an important downside. If the model used by the planner is mispecified and fails to capture fundamental reasons for trading, the attributed amount of non-fundamental trading and the level of the induced optimal tax will be too high.}
Results  The demand for the risky asset takes a similar form as in the baseline model, featuring also an inaction region, now determined jointly by the trading costs and the transaction tax. The optimal portfolio given prices can be compactly written in the trade region as:

\[
X_{1i} = E_{\tilde{t}}[D] - A_i \text{Cov}[E_{2i}, D] - P_1 (1 + \text{sgn}(\Delta X_{1i}) (\tau + \psi)) - \text{sgn}(\Delta X_{1i}) \eta + \alpha X_{0i}
\]

All three types of trading costs — quadratic, linear in shares and linear in dollar value — shift investors’ portfolios towards their initial positions. The equilibrium price is a slightly modified version of (6).

When calculating welfare, the planner takes into account that investors must incur in these costs when trading — this is the natural constrained efficient benchmark. The optimal tax formula remains unchanged when investors face transaction costs, as long as these trading costs represent exclusively a compensation for the use of economic resources.

Proposition 6. (Pre-existing trading costs) a) Under assumption [NR], when investors face trading costs as specified in (14), the expression for \( \tau^* \) is identical to the one in proposition 2.

The intuition behind proposition 6 is similar to the baseline case. An envelope condition eliminates any term regarding transaction costs from \( \frac{dV}{d\tau} \), because the planner must also face such costs, so the optimal tax looks identical to the one in the baseline model. This relies on the assumption that the economic profit made by the intermediaries who receive the transaction costs is zero — there cannot be economic rents.\(^25\)

This result has further implications. First, although the optimal tax formula does not vary, an economy with transaction costs has less trade in equilibrium than one without transaction costs. Depending on whether this reduction in trading is of the fundamental type or not, the optimal tax may be larger or smaller — the online appendix presents a numerical example in which the optimal tax is larger when there are transaction costs. Transaction costs affect the optimal tax through changes in the identity of the marginal investors. Second, the mere existence of transaction costs does not provide a new rationale for further discouraging non-fundamental trading.\(^26\) Welfare losses must be traced back to wedges derived from portfolio distortions. Third, if transactions costs, that is, \( \psi, \eta \) and \( \alpha \), were endogenously functions of \( \tau \), as in richer models of the market microstructure, the planner would have to take into account those effects when solving for optimal taxes.\(^27\) For instance, if a transaction tax endogenously increases trading costs, the optimal tax may be very small. However, if endogenously determined transaction costs are efficiently determined, the envelope theorem would still apply, leaving proposition 6 unchanged — the values of \( \frac{dX_{1i}}{d\tau} \) would subsume any microstructure effect.

\(^{25}\)Does proposition 6 imply that if there were two authorities with taxation power, they would both impose the same \( \tau^* \) twice? Of course not. Assume for simplicity that they set taxes sequentially. The first authority would set the optimal tax according to proposition 2. The second authority, internalizing that the pre-existing tax is a mere transfer and does not correspond to a compensation for costs of trading, would set a zero tax. Alternatively, \( \tau^* \) would characterize the sum of both taxes.

\(^{26}\)It is sometimes argued — as in the leading example of Brunnermeier, Simsek and Xiong (2014) — that excess trading is costly because there are transaction costs.

\(^{27}\)The Walrasian approach of this paper does not capture market microstructure effects. There is scope for understanding how transaction taxes affect market making and liquidity provision in greater detail, introducing, for instance, imperfectly competitive investors, search or network frictions. The results of this paper would still be present regardless of the specific trading microstructure.
4.2 Multiple risky assets

Assumptions The results of the baseline model extend naturally to an environment with multiple assets. Now there are \( J \) risky assets in fixed supply, in addition to the riskless asset. The \( J \times 1 \) vectors of total shares, equilibrium prices and dividend payments are respectively denoted by \( q, \ p \) and \( d \). Every purchase or sale of a risky asset faces an identical linear transaction tax \( \tau \). This is a further restriction on the planner’s problem, since belief disagreements can vary across different assets, but the tax must be constant. Allowing for different taxes for different (groups of) assets is conceptually straightforward, following the logic of section 4.4.

The distribution of dividends \( d \) paid by the risky assets is a multivariate normal with a given mean and variance-covariance matrix \( \text{Var}[d] \). All investors agree about the variance, but an investor \( i \) believes that the mean of \( d \) is \( \mathbb{E}_i[d] \). We can thus write:

\[
d \sim_i N \left( \mathbb{E}_i[d], \text{Var}[d] \right)
\]

Risk aversion \( A_i \), and the vectors of initial asset holdings \( x_{0i} \), hedging needs \( \text{Cov}[E_{2i},d] \) and beliefs \( \mathbb{E}_i[d] \) are arbitrary across the distribution of investors. The wealth at \( t = 2 \) of an investor \( i \) is thus given by:

\[
W_{2i} = E_{2i} + x_{i1}d + (x_{0i}p - x'_{i1}p - |x_{1i}' - x_{0i}'| p\tau + T_{1i})
\]

Results The first order condition (15) characterizes the solution of this problem for the set of assets traded:

\[
x_{1i} = (A_i\text{Var}[d])^{-1} \left( \mathbb{E}_i[d] - A_i\text{Cov}[E_{2i},d] - p - \hat{p}_i\tau \right),
\]

where \( \hat{p}_i \) is a \( J \times 1 \) vector where row \( j \) is given by \( \text{sgn}(\Delta X_{1ij}) \) \( p_j \) and \( p_j \) denotes the price of asset \( j \). If an asset \( j \) is not traded by an investors \( i \), then \( X_{1ij} = X_{0ij} \). If asset returns are independent, the portfolio allocation to every asset can be determined in isolation. Equilibrium prices are the natural generalization of the baseline model.

Proposition 7. (Multiple risky assets)

a) Under assumption [NR], the optimal tax when investors can trade \( J \) risky assets is given by:

\[
\tau^* = \sum_{j=1}^{J} \omega_j \tau^*_j
\]

With weights \( \omega_j \) and individual-asset taxes \( \tau^*_j \) given by \( \omega_j = \frac{p_j \int \text{sgn}(\Delta X_{1ij}) \frac{dX_{1ij}}{\tau} dF(i)}{\sum_{j=1}^{J} p_j \int \text{sgn}(\Delta X_{1ij}) \frac{dX_{1ij}}{\tau} dF(i)} \) and \( \tau^*_j = \frac{\int \mathbb{E}_i[d_j] \frac{dX_{1ij}}{\tau} dF(i)}{\int \text{sgn}(\Delta X_{1ij}) \frac{dX_{1ij}}{\tau} dF(i)} \).

The formula for \( \tau^*_j \) is identical to the one in an economy with a single risky asset. The optimal tax in a model with \( J \) risky assets is simply a weighted average of all \( \tau^*_j \). The weights are determined by the relative marginal changes in (dollar) volume. Those assets whose volume responds more aggressively to tax changes carry higher weights when determining the optimal tax and vice versa.

---

28I use bold lower-case letters to denote vectors but, for consistency, I keep the upper-case notation for holdings of a single asset.

29The model with many assets contains interesting predictions for how the set of assets traded in equilibrium endogenously adjust to tax changes. There is scope for much further analysis on that question.
4.3 General utility and arbitrary beliefs

Assumptions This section extends the main results to an environment with investors who face a non-trivial consumption choice between periods, have general utility specifications, and disagree about probability distributions in an arbitrary way. Investors’ beliefs are now modeled as a change of measure with respect to the probability measure used by the planner, which (jointly) determines the realization of all random variables — dividends and endowments — in the model. The conditional beliefs of an investor $i$ about $t = 2$ uncertainty are determined by a Radon-Nikodym derivative $Z_i$, which is absolutely continuous with respect to the actual probability distribution. This random variable $Z_i$ captures any discrepancy between the probability assessments of the planner and the ones used by the investors.\footnote{31}

Investors thus maximize:

$$\max_{C_{1i},X_{1i},Y_{1i}} U_i(C_{1i}) + \beta E_i \left[ U_i(C_{2i}) \right]$$

With budget constraints:

$$C_{1i} + X_{1i}P_1 + Y_{1i} = E_1 + X_{0i}P_i - \tau P_i \mid \Delta X_{1i} \mid + T_{1i} \quad \text{and} \quad C_{2i} = E_{2i} + X_{1i}D + Y_{1i},$$

where $Y_{1i}$ denotes the amount invested in the safe asset.

Results Investors’ consumption and portfolio allocations are determined by a pair of Euler equations for the risky and the safe asset:

$$U_i'(C_{1i}) = \beta E_i \left[ U_i'(C_{2i}) \right] \quad \text{and} \quad U_i'(C_{1i}) P_1 (1 + \text{sgn} \left( \Delta X_{1i} \right) \tau) = \beta E_i \left[ U_i'(C_{2i}) D \right],$$

which combined with market clearing fully characterize the equilibrium. In this general case it is impossible to solve for allocations and prices in closed form.

Proposition 8. (General utility and arbitrary beliefs)

a) The optimal tax in this general framework is given by:

$$\tau^* = \frac{\int \lambda_i U_i'(C_{1i}) \left( \text{Cov} \left[ Z_i \beta \frac{U_i'(C_{2i})}{U_i'(C_{1i})} D \right] \frac{dX_{1i}}{d\tau} + \text{Cov} \left[ Z_i \beta \frac{U_i'(C_{2i})}{U_i'(C_{1i})} \right] \frac{dY_{1i}}{d\tau} - \Delta X_{1i} \frac{dP_i}{d\tau} \right) dF(i)}{\int \lambda_i U_i'(C_{1i}) P_1 \text{sgn} \left( \Delta X_{1i} \right) \frac{dX_{1i}}{d\tau} dF(i)} \quad (17)$$

b) When marginal utilities are approximately constant, that is $U_i'(\cdot) \approx 1$, and $\lambda_i$ is constant across investors, the expression for $\tau^*$ is identical to the one in proposition 2.

The numerator of the optimal tax in the general case captures the distortions on investors’ Euler equations from the planner’s perspective. The first term is the non-linear counterpart of the numerator in the baseline model. Note that the belief used by the planner matters to determine the optimal tax, since it is needed to calculate the covariance with the Radon-Nikodym derivative.\footnote{32} The second term

\footnote{30}Because now income effects matter, it is important in theory who receives the tax rebate. However, if the total tax liability is small, income effects should be negligible.

\footnote{31}This should not be surprising since this general formulation encompasses different forms of belief distortions. For instance, it allows investors to be mistaken about their own hedging needs. The online appendix shows that, even in the CARA-normal model, the planner would need to know the appropriate hedging needs of investors who hedge incorrectly to determine the optimal tax.
corresponds to the mistake perceived by the planner regarding the holdings of the safe asset \( (R = 1) \) — the planner tilts the transaction tax to correct savings distortions. The third term simply captures changes in the terms-of-trade. Assumption [NR] is not needed in this case when the planner is a strict utilitarian.

Proposition 8b shows that the optimal tax found in the baseline model works as an approximation to the optimal tax in the general case when the risks faced by investors are small in comparison to their risk bearing capacity.\(^{32}\) Hence, the tractable CARA-Normal results of the baseline model can be seen as a first order approximation to more general models in which income effects are not particularly strong.

4.4 Asymmetric taxes/Multiple tax instruments

In the baseline model, the only instrument available to the planner is a single linear financial transaction tax which applies symmetrically to all investors. However, the planner could set different (linear) taxes for buyers and sellers. Or, at least theoretically, even investor-specific taxes. In general, more sophisticated policy instruments bring the outcome of the planner’s problem closer to the first-best, at the cost of increasing informational requirements.

Asymmetric taxes on buyers versus sellers

Assume now that buyers pay a linear tax \( \tau_B \) in the dollar volume of the transaction while sellers pay \( \tau_S \). Hence, total tax revenue is given by \( (\tau_B + \tau_S) P_1 |\Delta X_{1i}| \). Outside of the inaction region, the optimal portfolio demand is given by:

\[
X_{1i} = \frac{E_i[D] - A_i \text{Cov}[E_{2i}, D] - P_1 (1 + I_2 \left[ \Delta X_{1i} > 0 \right] \tau_B + I_2 \left[ \Delta X_{1i} < 0 \right] \tau_S)}{A_i \text{Var}[D]}. 
\]

where \( I_2 [\cdot] \) denotes the indicator function. This expression differs from (4) in that buyers now face a different tax than sellers. The equilibrium price is a natural extension of the one in the baseline model.

Proposition 9. (Asymmetric taxes on buyers versus sellers)

a) Under assumption [NR], the pair of optimal taxes \( \{\tau_B^*, \tau_S^*\} \) is characterized by the solution of the following system of non-linear equations:

\[
\begin{align*}
\tau_B^* + \tau_S^* &= \frac{\int E_i[D] \frac{dX_{1i}}{d\tau_B} dF(i)}{\int B \frac{dX_{1i}}{d\tau_B} dF(i)} , \\
\tau_B^* + \tau_S^* &= \frac{\int E_i[D] \frac{dX_{1i}}{d\tau_S} dF(i)}{\int B \frac{dX_{1i}}{d\tau_S} dF(i)}. 
\end{align*}
\]

The economic forces that shape the optimal values for \( \tau_B^* \) and \( \tau_S^* \) are the same as in the baseline model. Once again, the planner’s belief is irrelevant for the optimal policy. This shows that, as long as [NR] holds, the use of more sophisticated policy instruments does not require knowledge of the actual distribution of payoffs. Intuitively, the change in portfolio allocation induced by a marginal change in any instrument must cancel out in the aggregate. Equation (18) provides intuition for why all taxes in the baseline model are divided by 2; in that case, there exists a single optimality condition and \( 2\tau^* = \tau_B^* + \tau_S^* \).

\(^{32}\)This result is related to the classic Arrow-Pratt approximation (Arrow (1971), Pratt (1964)), which shows that the solution to the portfolio problem with CARA utility and normal distribution of returns is an approximation to any portfolio problem for small gambles. This is also an approximation to continuous time diffusion models.
As long as there are more than two investors, this system has at least a solution. When there are two investors, the system is indeterminate and only the sum $\tau_B + \tau_S$ is pinned down. In that case, $\tau_B + \tau_S = \frac{E_B[D] - E_S[D]}{P_1}$.

**Individual taxes/First-best**

Assume now that the planner can set investor specific taxes. This is an interesting theoretical benchmark, despite being unrealistic. For simplicity, I now assume that there is a finite number $N$ of (types of) investors in the economy.

**Proposition 10. (Individual taxes/First-best)**

a) The first-best can be implemented with a set of investor specific taxes given by:

$$\tau_i^* = \text{sgn}(\Delta X_i) \frac{E_i[D] - Y}{P_1}, \forall i = 1, \ldots, N, \quad (19)$$

where $Y$ is any real number; a natural choice for $Y$ is $E[D]$.

b) Under assumption [NR], the planner only needs $N - 1$ taxes to implement the first-best in an economy with $N$ investors. If [NR] does not hold, $N$ taxes are needed.

Proposition 10a follows standard Pigovian logic. The planner sets optimal individual taxes so that investors portfolio choices replicate those of a economy with homogenous beliefs. Note that the planner can use any belief $Y$ to implement the first-best allocation, as long as it is the same for all investors. In a production economy, the natural choice would be $Y = E[D]$. Finally, because $P_1$ is a function of all taxes, equation (19) also defines a system of non-linear equations.

Proposition 10b shows that the first-best could be implemented with $N - 1$ taxes when [NR] holds. This occurs because the risky asset is in fixed supply. The logic behind this result is similar to Walras’ law. For instance, when $N = 2$, a single tax which modifies directly the allocation of one of the investors necessarily changes the allocation of the other one through market clearing.

**5 Production**

The results derived so far rely on the assumption that assets are in fixed supply. I now study how optimal policies vary when financial markets determine production by influencing the intertemporal investment decision in a standard price-taking environment — this is the role explored in q-theory models, as Tobin (1969) and Hayashi (1982).

**Assumptions** There is a new group of agents in the economy who were not present in the baseline model: identical competitive producers in unit measure. Producers are indexed by $k$ and maximize well-behaved time separable expected utility, with flow utility given by $U_k(\cdot)$. They have exclusive access to a technology $\Phi(S_{1k})$, which allows them to issue or dispose of $S_{1k}$ shares of the risky asset at $t = 1$.\(^{33}\)

\(^{33}\) A “tree” analogy can be helpful here. Assume that a share of the risky asset (i.e., a tree) entitles the owner to a dividend payment $D$ (fruit). Producers can plant new trees or chop them at a cost $\Phi(S_{1k})$, which they sell or buy at a price $P_1$. Producers would be willing to create trees until the marginal cost of producing a new tree/chopping and old tree $\Phi'(S_{1k})$ equals the marginal benefit of selling/buying $P_1$. For consistency, any normalization concerning $Q$ must also normalize $\Phi(\cdot)$. 

30
to $S_{1k}$, which can be negative, as investment. The function $\Phi(\cdot)$ is increasing and strictly convex; that is, $\Phi'(\cdot) > 0$, $\Phi''(\cdot) > 0$. To ease the exposition, I assume throughout that $\Phi(S_{1k}) = \gamma_1 |S_{1k}| + \frac{\gamma_2}{2} |S_{1k}|^2$, with $\gamma_1, \gamma_2 > 0$. Producers are initially endowed with $E_{1k}$ units of consumption good (dollars) and can only borrow or save in the riskless asset at a (gross) rate $R = 1$. Their endowment $E_{2k}$ at $t = 2$ is stochastic and follows an arbitrary distribution.

To avoid distortions in primary markets, the planner does not tax the issuance of new shares. Importantly, market clearing is now given by $\int X_t dF(i) = Q + S_{1k}$. Total output at $t = 2$ in this economy is endogenous and given by $D(Q + S_{1k})$.

**Positive results** Producers thus maximize:

$$\max_{C_{1k}, C_{2k}, S_{1k}, Y_k} U_k(C_{1k}) + \mathbb{E}[U_k(C_{2k})]$$

With budget constraints $Y_k + C_{1k} = E_{1k} + P^s_1 S_{1k} - \Phi(S_{1k})$ and $C_{2k} = E_{2k} + Y_k$, where $Y_k$ denotes the amount saved in the riskless asset and $P^s_1$ denotes the price faced by producers — the superscript $s$ stands for supply. The optimality conditions for producers are given by:

$$U'_k(C_{1k}) = \mathbb{E}[U'_k(C_{2k})] \quad \text{and} \quad P^s_1 = \Phi'(S_{1k})$$

The first condition is a standard Euler condition for the riskless asset. The second condition provides a supply curve for the number of shares. Combining this supply curve with the portfolio choices of investors, generates the following equilibrium price:

$$P_1 = (1 - \alpha) \gamma_1 + \alpha P^s_1,$$

where the weight $\alpha \in [0, 1]$ — defined in the appendix — is higher when the adjustment cost is very concave ($\gamma_2$ is large) and $P^s_1$ is essentially the same expression for the price that would prevail in an exchange economy, which is given in equation (6). Intuitively, the equilibrium price is a weighted average of the exchange economy price and $\gamma_1$, which is the replacement cost of the risky asset with linear adjustments costs.

The following lemma shows that allowing for production does not affect those positive properties of the model that matter for the determination of the optimal tax.

**Lemma 3. (Effect of $\tau$ on prices and allocations)** Lemmas 1 and 2 remain valid. Hence, an increase in the transaction tax can increase, reduce or keep equilibrium prices (and investment) constant, but all buyers buy less and all sellers sell less.

**Normative results** Accounting for producers, social welfare is now defined as:

$$V(\tau) = \int \lambda_i V_i dF(i) + \lambda_k V_k,$$

where $V_k$ and $\lambda_k$ respectively denote the indirect utility and the welfare weight of producers. The change in producers’ welfare induced by a marginal change in the tax is given by:

$$\frac{dV_k}{d\tau} = U'_k(C_{1k}) \left[ \frac{dP_1}{d\tau} S_{1k} + [P_1 - \Phi'(S_{1k})] \frac{dS_{1k}}{d\tau} - \frac{dY_k}{d\tau} \right] + \mathbb{E} \left[ U'_k(C_{2k}) \right] \frac{dY_k}{d\tau} = \mathbb{E} \left[ U'_k(C_{2k}) \right] \frac{dP_1}{d\tau} S_{1k},$$
where the second line follows by substituting producers’ optimality conditions. Intuitively, because producers do not pay taxes and invest optimally given prices, a marginal tax change only modifies their welfare through the terms-of-trade on the shares they issue/repurchase. When \( P_i \) is high, producers enjoy a better deal selling shares than when \( P_i \) is low. The envelope theorem eliminates from \( \frac{dV_k}{d\tau} \) the direct effects caused by changes in producers portfolio or investment choices.

Proposition 11 imposes assumption [NR] directly, assuming that it also applies to producers’ welfare, and characterizes the optimal tax.

**Proposition 11. (Optimal tax in production economies)**

a) Under assumption [NR], the optimal tax in a production economy is given by:

\[
\tau^* = \frac{\int \left( \frac{E[D]-E_i[D]}{P_i} \right) \frac{dX_{ii}}{d\tau} dF(i)}{-\int \text{sgn}(\Delta X_{1i}) \frac{dX_{ii}}{d\tau} dF(i)} = (1 - \omega) \tau^\text{exchange} + \omega \tau^\text{production},
\]

(20)

where \( \tau^\text{exchange} = \frac{\zeta(\tau) \text{Cov}_{F,T} \left[ \frac{E[D]-E_i[D]}{P_i} \frac{dX_{ii}}{d\tau}, \frac{dX_{1i}}{d\tau} \right]}{2 \int_{I\in B} \frac{dX_{ii}}{d\tau} dF(i)} \), \( \tau^\text{production} = \frac{E[D]-E_i[D]}{P_i} \) and \( \omega < 1 \) is given in the appendix (\( \omega \) is small in magnitude when \( \frac{dS_{1i}}{d\tau} \approx 0 \) and close to unity when \( \frac{dS_{1i}}{d\tau} \) is large). \( E_{F,T} [E_i [D]] \) denotes the average belief in the population of active investors, \( \text{Cov}_{F,T} \left[ E_i [D], \frac{dX_{1i}}{d\tau} \right] \) denotes a cross-sectional covariance among active investors and \( \zeta (\tau) \equiv \int_{I\in T} dF(i) \) is the fraction of active investors.

The optimal tax can be written as a linear combination between the optimal tax in a (fictitious) exchange economy and the optimal tax in a (fictitious) production economy with a single investor with belief \( E_{F,T} [E_i [D]] \). The sensitivity of investment with respect to a tax change determines the relative importance of each term.

Market clearing now implies that \( \int \frac{dX_{ii}}{d\tau} dF(i) = \frac{dS_{1i}}{d\tau} \), which can take any positive or negative value. Hence, in production economies, the belief used by the planner to calculate welfare matters in general for the optimal policy. However, if the planner uses investors’ average belief to calculate welfare, the belief used by the planner drops out of the optimal tax expression. Because of its importance, I state this result as a corollary of proposition 11.

**Corollary. (\( \tau^* \) may depend on planner’s belief)** Even when assumption [NR] holds, the optimal financial transaction tax in a production economy depends on the distribution of payoffs assumed by the planner. However, if the planner uses the average belief across investors, that is, \( E [D] = E_{F,T} [E_i [D]] \) at the optimum, the optimal tax is identical to the one in the exchange economy and independent of the belief used by the planner.\(^{34}\)

The numerator in equation (20), which evaluated at \( \tau = 0 \) determines the sign of the optimal tax can be decomposed in two terms:

\[
\int (E [D] - E_i [D]) \frac{dX_{ii}}{d\tau} dF(i) = -\zeta (\tau) \text{Cov}_{F,T} \left[ E_i [D], \frac{dX_{1i}}{d\tau} \right] + (E [D] - E_{F,T} [E_i [D]]) \frac{dS_{1i}}{d\tau},
\]

(21)

\(^{34}\)The average belief may change if there are changes in the composition of marginal investors. For the irrelevance result to hold without further qualifications, I implicitly assume that the average belief for marginal investors remains constant for all \( \tau \).
Because the second term in (21) is in general non-zero when $\tau = 0$, we can say that belief distortions in production economies have an additional first-order effect on welfare. Again asset prices do not appear in optimal tax formulas, despite playing a role in determining allocations. All welfare losses must be traced back to distortions in “quantities”, either in portfolio allocations, captured by $\frac{dX_i}{dt}$, or in production decisions, captured by $\frac{dS_k}{dt}$.

Intuitively, the optimal tax corrects two wedges created by heterogeneous beliefs. First, given an amount of aggregate risk, the optimal tax seeks to reduce the dispersion in asset holdings induced by disagreement — some investors are holding too much risk and some others too little risk. This is the same mechanism present in exchange economies. Second, as long as the average belief differs from the one used by the planner, the level of production in the economy is too high (low) when investors are on average too optimistic (pessimistic). This provides a second rationale for taxation. Intuitively, the investors in the economy hold too much aggregate risk when they are on average optimistic or too little when they are pessimistic.\(^{35}\)

**Sign of $\omega^{\tau^*_\text{production}}$** Assumption [OBPS] is not sufficient anymore to pin down the sign of the optimal tax, which now also depends on whether $\mathbb{E}_F[\mathbb{E}_i[D]] - \mathbb{E}_F[\mathbb{E}_i[D]]$ and $\frac{dS_k}{dt}$ have the same or opposite signs. Intuitively, if a marginal tax increase reduces (increases) investment at the margin when investors are too optimistic (pessimistic), a positive tax is welfare improving, and vice versa. Table 5 summarizes the conditions that determine the sign of the term associated to production.

<table>
<thead>
<tr>
<th>Aggregate optimism</th>
<th>Aggregate pessimism</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{E}_F[\mathbb{E}_i[D]] &gt; \mathbb{E}_F[D]$</td>
<td>$\mathbb{E}_F[\mathbb{E}_i[D]] &lt; \mathbb{E}_F[D]$</td>
</tr>
<tr>
<td>$\int_B \frac{1}{\lambda_i} dF(i) &gt; \int_S \frac{1}{\lambda_i} dF(i)$</td>
<td>$\omega^{\tau^*_\text{production}} &gt; 0$</td>
</tr>
<tr>
<td>$\int_B \frac{1}{\lambda_i} dF(i) &lt; \int_S \frac{1}{\lambda_i} dF(i)$</td>
<td>$\omega^{\tau^*_\text{production}} &lt; 0$</td>
</tr>
</tbody>
</table>

Table 5: Sign of $\omega^{\tau^*_\text{production}}$

Unlike in the exchange economy, in which there are theoretical and empirical arguments that support assumption [OBPS], it is not obvious whether we should expect $\omega^{\tau^*_\text{production}}$ to be positive or negative. As shown in lemma 3, for investment to be reduced (increased) at the margin by a tax increase, it has be the case that the (risk aversion adjusted) mass of buyers is larger (smaller) than the mass of sellers. In principle, the relation between the average belief distortion and the relative mass of buyers/sellers need not be linked, so the sign of $\omega^{\tau^*_\text{production}}$ is theoretically ambiguous.\(^{36}\)

I conclude this section with the following remark.

**Remark 7.** The covariance between investors being too optimistic/pessimistic on average and the effect of a transaction tax on prices is what determines the convenience of a transaction tax from a welfare standpoint in a production economy, not the effect of the tax in price volatility.

\(^{35}\)If there were many produced risky assets, the welfare losses would capture the idea that belief distortions misallocate real investment across sectors in the economy. These results are available under request.

\(^{36}\)It is possible to show that other policy instruments, like short-sale constraints, borrowing constraints, or investment taxes, can be welfare improving when dealing with belief distortions in production economies. Those results are available under request.
Many informal discussions regarding the convenience of a transaction tax, following Tobin (1978), revolve around the notion that it would help reduce price volatility — implicit in those discussions is the notion that high volatility is bad. The results in this section show that it is not price volatility — a variance — but whether investment (through prices) is lower when investors are optimistic and vice versa — a covariance — what captures the welfare consequences of a transaction tax in a production context.

6 Dynamics

I now extend the results to dynamic environments. This analysis shows that, controlling for the magnitude of disagreement, a lower optimal tax is needed when investors buy and sell frequently. For clarity, I study a three-period extension of the baseline model, which is sufficient to convey the main theoretical results.

Assumptions

Now there is an additional period in this economy, so \( t = \{1, 2, 3\} \). Figure 5 represents the timeline.

\[
\begin{align*}
X_{1t} & \quad \text{chosen} \\
D_2 & \quad \text{paid} \\
X_{2t} & \quad \text{chosen} \\
D_3 & \quad \text{paid} \\
\text{Investors consume } W_{3t}
\end{align*}
\]

Figure 5: Three period dynamics

Investors choose portfolios at \( t = 1 \) and \( t = 2 \) to maximize expected utility of consumption at the final period, that is, \( E_i \left[-e^{-A_i W_{3t}}\right] \). The risky asset yields a dividend \( D_2 \) in period 2 and a dividend \( D_3 \) in period 3. Every investor holds different beliefs about the expected payoff of the risky asset. An investor \( i \) believes that \( D_t \) is normally distributed with mean \( E_i \left[D_t\right] \) and variance \( Var \left[D_t\right] \), that is:

\[
D_t \sim_i N \left( E_i \left[D_t\right], Var \left[D_t\right] \right), \quad t = 2, 3
\]

Investors’ beliefs are predetermined and dogmatic, and do not change when time and uncertainty unfold.\(^{37}\) The planner calculates welfare using the mean belief \( E \left[D_t\right] \). Investors’ budget constraints can thus be written as:

\[
\begin{align*}
W_{1t} &= E_{1t} + X_{0t}P_1 \\
W_{2t} &= E_{2t} + (D_2 + P_2) X_{1t} - X_{1t}P_1 - \tau |P_1| \Delta X_{1t} + T_{1t} + W_{1t} \\
W_{3t} &= E_{3t} + D_3 X_{2t} - X_{2t}P_2 - \tau |P_2| \Delta X_{2t} + T_{2t} + W_{2t}
\end{align*}
\]

I continue to restrict the set of policy measures to a constant linear transaction tax under commitment. The use of time varying taxes raises a different set of concerns — related to credibility and commitment — which I do not address in this paper.

\(^{37}\)Note that these assumptions imply that investors have perfect foresight and fully agree about the price \( P_2 \), independently of the realization of \( D_2 \), which allows to solve the model analytically. A previous version of this paper, which analyzed a more general multiperiod model, relaxed that feature and several others. Those results are available under request.
Positive results The problem faced by investors and the equilibrium at \( t = 2 \) is identical to the baseline model. I focus instead on \( t = 1 \). The optimal portfolio choice of a given investor at \( t = 1 \), is given by:

\[
X_{1i} = \begin{cases} 
X_{0i}; & \Delta X_{1i} = 0 \quad \text{No Trade at } t = 1 \\

v \hat{X}_{1i} + (1-v) \hat{X}_{2i}; & \Delta X_{1i} \neq 0, \Delta X_{1i} = 0 \quad \text{Trade at } t = 1; \text{No Trade at } t = 2 , \\

\hat{X}_{1i}; & \Delta X_{1i} \neq 0, \Delta X_{2i} \neq 0 \quad \text{Trade at } t = 1; \text{Trade at } t = 2 
\end{cases}
\]

where

\[
\hat{X}_{ti} = \frac{\mathbb{E}_i[D_{t+1}] + P_{t+1} (1 + \text{sgn}(\Delta X_{t+1})) \tau - A_i \text{Cov}[E_{t+1}, D_{t+1}] - P_t (1 + \text{sgn}(\Delta X_{ti}) \tau)}{A_i \mathbb{V}ar[D_{t+1}]}, \quad (22)
\]

and \( \nu = \frac{\mathbb{V}ar[D_3]}{\mathbb{V}ar[D_2] + \mathbb{V}ar[D_3]} \).

As in the baseline model, investors may decide not to trade. If they decide to trade at period \( t \), their portfolio decision will depend on whether they are planning to trade in the next period or not. If they are not going to trade in the next period, their optimal portfolio allocation is a weighted average of the optimal portfolio allocations at \( t \) and \( t + 1 \). If they are going to trade in the next period, the optimal portfolio now contains a forward-looking correction, which corresponds to the term multiplying the optimal tax in the numerator of (22), which can be written as:

\[
-\tau \frac{P_{t+1}}{P_t} \text{sgn}(\Delta X_{ti}) \text{sgn}(\Delta X_{t+1}) \equiv \kappa_{ti}
\]

Dynamic correction

Assuming, to simplify the interpretation, that \( \frac{P_{t+1}}{P_t} \approx 1 \), we can write the dynamic correction at \( t = 1 \) as:

\[
1 - \kappa_{ti} = \begin{cases} 
2 & \text{if } \text{sgn}(\Delta X_{1i}) \text{sgn}(\Delta X_{2i}) < 0 \quad (\kappa_{ti} = -1) \\
1 & \text{if } \text{sgn}(\Delta X_{2i}) = 0 \quad (\kappa_{ti} = 0) \\
0 & \text{if } \text{sgn}(\Delta X_{1i}) \text{sgn}(\Delta X_{2i}) > 0 \quad (\kappa_{ti} = 1)
\end{cases}
\]

Intuitively, when investors trade in one direction at period \( t = 1 \) and in the opposite one at \( t = 2 \), they anticipate that a transaction tax will hit him twice, once when buying and another time when selling. When investors buy-and-hold, only the time \( t \) distortion modifies the portfolio decision — this is exactly what happens in the static model. Finally, when investors trade in the same direction at periods \( t \) and \( t + 1 \), a transaction tax does not affect at all period \( t \) portfolio decision. A persistent buyer (seller) is aware that every (marginal) share purchased at \( t = 1 \) has the extra benefit of saving the \( t = 2 \) tax, rendering the tax irrelevant. This logic extends to the case with \( P_{t+1}/P_t \neq 1 \), with minor modifications.\(^{38}\)

The determination of equilibrium prices is straightforward.

\(^{38}\)Because the tax is paid on the dollar value of the transaction, investors tilt their purchasing decision towards periods with lower prices. For instance, when prices are steeply rising, a persistent buyer has a further incentive to buy early. Therefore, expected price growth can amplify or dampen the forward-looking effect. For the interpretation of the results, I implicitly assume that this effect is never powerful enough to turn \((1 - \kappa_{ti})\) negative.
Normative results

The optimal tax can be written as a (corrected) weighted average of the taxes that would prevail in static economies.

Proposition 12. (Optimal tax in dynamic environments)

a) Under assumption [NR], the optimal tax in an economy with two trading periods is given by:

\[ \tau^* = \omega_1 f_1 \tau_1^* + \omega_2 \tau_2^*, \]

where \( f_1, \omega_1 \) and \( \omega_2 \) given in the appendix, are such that \( \omega_1 + \omega_2 = 1 \) and \( f_1 \in \left[ \frac{1}{2}, \infty \right) \), and \( \tau_i^* \) denotes the optimal static tax as defined in proposition 2.

The term \( f_1 \) is linked to the forward-looking nature of investors’ decisions. If \( f_1 = 1 \), the optimal tax becomes a weighted average of static taxes, with weights determined by relative portfolio responses. This result is similar to the multiple risky asset case. Intuitively, the planner seeks to correct a weighted average of the belief distortions in the economy. Unsurprisingly, the weights are determined according to portfolio tax sensitivities. The belief distortions in periods in which volume is more responsive to tax changes have a higher weight \( \omega_t \) when setting the optimal tax.

When \( f_1 < 1 \), forward-looking investors trade on opposite directions, which makes them more responsive to the tax. Therefore, controlling for the level of disagreement induced trading \( \tau_i^* \), less weight is given to the optimal tax in the first period. The opposite occurs when \( f_1 > 1 \). In that case, most of the investors are persistent buyers/sellers, which makes them insensitive to tax changes. Hence, for a given level of disagreement induced trading, measured in the form of static taxes, a lower optimal tax is required when forward-looking investors alternate between being buyers and sellers. On the contrary, a higher tax is optimal if they are persistent buyers/sellers. Intuitively, very large taxes are needed to dissuade investors who are planning to build long-term positions over time.

Note that the planner exploits the fact that trades at high frequencies are particularly discouraged when solving for the optimal tax.\(^{39}\) In other words, the planner endogenously uses the well known fact that small trading costs can have large effects in dynamic models to design optimal policies. Trading dynamics directly influence the magnitude of the tax — through the set of weights \( \omega_t \) and \( f_t \) — but not necessarily its sign (i.e., if all \( \tau_i^* \) are positive, \( \tau^* \) must be positive too). The rationale for using corrective policies at all has to be traced back to the existence of belief distortions.

I conclude with several implications of the results in this section. First, because it has little effect on persistent buyers/sellers, the results of this section suggest that a transaction tax does not seem to be a good instrument to correct bubbles, understanding these as persistent fads that move asset prices. Second, in a more general dynamic model, imposing a transaction tax would make the markets incomplete. In that case, even for a planner who respects investors’ beliefs a transaction tax could be welfare improving if it improves risk sharing or welfare reducing if it worsens risk sharing.\(^{40}\) That rationale in favor or against taxing financial transactions is perfectly valid but orthogonal to the one

\(^{39}\)Tobin (1978) already acknowledges that the kind of transaction tax analyzed in this paper would have disproportionate effects over short horizon investments. The planner in this paper endogenously exploits this property of the tax when setting the optimal policy.

\(^{40}\)A previous version of this paper studied that case. Those results are available under request.
analyzed in this paper. Third, despite capturing dynamics, this paper does not model explicitly the role of modern high frequency trading, as described, for instance, by Budish, Cramton and Shim (2013). However, the dynamic results in this paper imply that investors who trade at very high frequencies would instantly become inframarginal for any tax level, no matter how small, necessarily making the planner’s problem non-convex. If high frequency trading happens to be of large fundamental value, the optimal tax might be very close to zero.

7 Final remarks

Before concluding, I would like to make three final remarks regarding the applicability and generality of the approach used in this paper.

General approach to normative problems with belief heterogeneity Any problem in which a planner does not respect investors’ beliefs to calculate welfare, including the one tackled in this paper, can be approached in two stages. First, we can characterize the solution to the planner’s problem taking as given the planner’s belief. This exercise allows us to identify wedges in investors’ decisions, as well as the optimal policy. Although overruling investors’ beliefs creates a rationale for intervention, the form of such intervention and the welfare losses induced by belief distortions are not obvious. For instance, adding a new set of agents (producers) to the baseline model introduces a new rationale for taxation, because investors’ beliefs affect investment in equilibrium. This paper focuses on this first stage, characterizing optimal policies given a planner’s belief.

If the optimal policy turns out to be independent of the belief used by the planner — as in proposition 2 — no further analysis is needed. If not, a second stage involves choosing the belief used by the planner. In those cases, the welfare criteria proposed in Brunnermeier, Simsek and Xiong (2014), Gilboa, Samuelson and Schmeidler (2014), Gayer et al. (2014) or Blume et al. (2013) can be used. Both approaches are complementary.

Relation to Tobin’s argument The theory developed in this paper differs substantially from the one in Tobin (1978). In a nutshell, Tobin postulates a) that prices are excessively volatile and b) that a transaction tax is a good instrument to reduce price volatility. His argument supporting a positive tax relies on two claims. On the positive side, he argues that transaction taxes reduce price volatility. On the normative side, he argues that price volatility is harmful.

On the contrary, on the positive side, this paper shows that the effect of transaction taxes on the level and volatility of asset prices is indeterminate, which refutes Tobin’s mechanism. This paper shows instead that transaction taxes are an effective instrument to discourage trading and reduce volume. Hence, any robust theory in which optimal transaction taxes are positive has to incorporate a mechanism by which reducing trading volume is welfare improving. On the normative side, this paper studies a mechanism by which the allocation of risk in the economy improves when investors are prevented from trading too much. Moreover, when investment reacts to asset prices, this paper shows that the convenience of a transaction tax depends on the covariance between the marginal effect of a tax change in investment (and prices) and the aggregate belief distortion. It is theoretically ambiguous whether a
transaction tax is welfare improving — by inducing investors to produce efficiently — in production economies.

**Correcting internalities vs. externalities** It may seem that the planner in this paper only seeks to improve the welfare of investors with distorted beliefs. That need not be the case, because of general equilibrium effects. It is true that tax policy can only increase social welfare if some investors hold heterogeneous beliefs. However, which particular investors benefit or lose from the tax policy depends on general equilibrium effects. In particular, when the beliefs of a given group of investors become more distorted, they may end up being better off. This can occur when the changes in the terms-of-trade of their portfolio happen to favor them. Although the paper focuses on aggregate efficiency, those spillovers are present throughout the paper, specially in the extensions with production and multiple assets.

Because of the Walrasian nature of this paper, all equilibrium spillovers induced by belief distortions work through prices. There is scope to explore whether, in richer models, investors with distorted beliefs can have a disproportionally large effect on other market participants — for instance, following the logic in Haltiwanger and Waldman (1985, 1989), by introducing some form of strategic complementarity, as in models with positive feedback traders, herding behavior or credit constraints.

### 8 Conclusion

This paper studies the welfare implications of taxing financial transactions in an equilibrium model of competitive financial markets. 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, no matter how small, hold heterogeneous beliefs, we expect the optimal tax to be positive. When financial markets determine production, the optimal transaction tax may be higher or lower, although the optimal tax for a planner who adopts the average belief among investors remains unchanged. Moreover, the optimal tax accounts for the fact that trading at higher frequencies is particularly discouraged. The frequency of trading in a given economy affects the magnitude of the optimal tax but the pattern of beliefs determines its sign.

There are many extensions of this paper that deserve to be explored. Understanding the normative implications of taxing financial transactions in models with endogenous learning dynamics or rich wealth dynamics, in search or networks environments and when some investors have market power are significant ones. The normative implications of taxing transaction taxes for information diffusion and acquisition are also relevant. These are all important questions that I leave for future research.

---

41 This is an important remark since there is often greater support for policies that protect bystanders from the mistakes of others than for policies that protect people from their own mistakes.

42 A previous version of this paper provided an example in which a transaction tax created a new first-order loss by preventing informed traders to disseminate information through trading. In ongoing work, I show in a similar environment that a (quadratic, for technical reasons) transaction tax can either increase or decrease the amount of information in the economy (i.e., the signal-to-noise ratio), depending on the portfolio sensitivities of the affected investors. However, because in models with disperse information there are pre-existing informational externalities, it is not clear whether increasing or decreasing the amount of information in the economy is welfare improving or vice versa. Any result on information diffusion and acquisition is complementary to the results of this paper.
Finally, this paper has implicitly assumed perfect tax enforcement. Whether legal and practical concerns regarding tax avoidance are sufficiently important to advise against implementing any financial transaction tax remains an open question.

References


Online appendix (not for publication)

A Proofs and derivations

Section 2

Sufficient condition for $P_1 > 0$

The exact sufficient 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:

$$E_i[D] > A_i \left( \text{Cov} [E_{Zi}, D] + \text{Var} [D] \frac{Q}{Q} \right), \quad \forall i$$

This condition is sufficient but not necessary. Any combination of expected payoffs and risk bearing capacity which makes the equilibrium price, given in equation (6), positive for the relevant range of $\tau$ would work.

Lemma 1. (Effect of $\tau$ on prices)

a) Substituting (7) into the derivative of the budget constraint $\int \frac{dX_{1i}}{d\tau} dF(i) = 0$, we can find that $\frac{dP}{d\tau} = -\int_{\tau} \frac{dX_{1i}}{dF(i)}$. Note that this derivation takes into account the fact that a marginal tax change modifies the extensive margin choice of some investors, who decide whether to trade or not; these effects do not affect price changes because the net demand of an investor indifferent at the extensive margin is zero — this follows directly from Leibniz rule. Taking derivatives directly on (4), we can also find that $\frac{\partial X_{1i}}{\partial \tau} = -\frac{P_i \text{sgn}(\Delta X_{1i})}{A_i \text{Var}[D]}$ and $\frac{\partial X_{1i}}{\partial p} = -\frac{(1+\text{sgn}(\Delta X_{1i})\tau)}{A_i \text{Var}[D]}$. Therefore, we have that $\frac{dP}{d\tau} = -\int_{\tau} \frac{P_i \text{sgn}(\Delta X_{1i})}{A_i \text{Var}[D]} dF(i) - \int_{\tau} \frac{\text{sgn}(\Delta X_{1i})}{A_i} dF(i) = \text{Var}[D] \left( \frac{\partial (\int_{\tau} X_{1i} |_{\tau=0} dF(i))}{\partial p} - \frac{\partial (\int_{\tau} X_{1i} |_{\tau=0} dF(i))}{\partial \tau} \right)$.

Lemma 2. (Effect of $\tau$ on allocations)

a) It is sufficient to show that the sign of $\frac{dX_{1i}}{d\tau}$ is identical to the one of $\frac{\partial X_{1i}}{\partial \tau}$. The total derivative of portfolio allocations in equilibrium with respect to taxes is given by equation (7). Hence, as derived in the proof of lemma 1, $\frac{dP}{d\tau} = -\int_{\tau} \frac{P_i \text{sgn}(\Delta X_{1i})}{A_i \text{Var}[D]} dF(i)$ and $\frac{\partial X_{1i}}{\partial p} = -\frac{(1+\text{sgn}(\Delta X_{1i})\tau)}{A_i \text{Var}[D]}$, we find:

$$\frac{dX_{1i}}{d\tau} = \frac{\partial X_{1i}}{\partial \tau} \left[ 1 - (\text{sgn}(\Delta X_{1i}) + \tau) \frac{\int \text{sgn}(\Delta X_{1i}) dF(i)}{\int \frac{1+\text{sgn}(\Delta X_{1i})\tau}{A_i} dF(i)} \right],$$

where $\epsilon_i$, which is constant within buyers/sellers, can be rewritten as $\epsilon_i = 1 - (\text{sgn}(\Delta X_{1i}) + \tau) \frac{1-H}{(1+\tau)+(1-\tau)H}$, where $H \equiv \frac{\int_{\tau} dF(i)}{\int_{\tau} dF(i)} \in (0, \infty)$ is the ratio of sellers risk aversion to buyers risk aversion. It is straightforward to show that $\epsilon_i > 0$ for both buyers and sellers. This proves lemma 2.
Section 3

Assumption [NR]. (No Redistribution)

When the planner has access to ex-ante lump sum transfers, he solves:

$$\max_{\{T_{ii}\}} \int \lambda_i V_i (T_{ii}) dF (i) - \delta \int T_{ii} dF (i),$$

where the last term represents the total amount of transfers. Hence the optimality condition in $T_{ii}$ is

$$\lambda_i \frac{dV_i}{dT_{ii}} = \delta, \quad \forall i,$n where it is easy to show that $\frac{dV_i}{dT_{ii}} = \mathbb{E} [U'_i (W_{2i})].$

Assumption [OBPS]. (Optimists Buyers/Pessimists Sellers)

The marginal gain/loss in welfare at $\tau = 0$ is given by $- \int \mathbb{E}_i [D] \frac{dX_{ii}}{d\tau} |_{\tau=0} dF (i)$. We can define

$$\int \mathbb{E}_i [D] \frac{dX_{ii}}{d\tau} |_{\tau=0} dF (i) \equiv \mathbb{E}_F \left[ \mathbb{E}_i [D] \frac{dX_{ii}}{d\tau} |_{\tau=0} \right],$$

and then using properties of the covariance, find that

$$\mathbb{E}_F \left[ \mathbb{E}_i [D] \frac{dX_{ii}}{d\tau} |_{\tau=0} \right] = \text{Cov}_F \left( \mathbb{E}_i [D], \frac{dX_{ii}}{d\tau} |_{\tau=0} \right).$$

Market clearing implies that

$$\mathbb{E}_F \left[ \frac{dX_{ii}}{d\tau} |_{\tau=0} \right] = 0,$n hence $\int \mathbb{E}_i [D] \frac{dX_{ii}}{d\tau} |_{\tau=0} dF (i) = \text{Cov}_F \left( \mathbb{E}_i [D], \frac{dX_{ii}}{d\tau} |_{\tau=0} \right).$ This condition can be rewritten as a function of exogenous variables after substituting for $\frac{dX_{ii}}{d\tau} |_{\tau=0}$.

Theoretical justification for assumption [OBPS] Assumption [OBPS] determines the sign of the optimal tax. I now provide a purely theoretical argument that show that assumption [OBPS] must hold in expectation. Using the result from lemma 1 that shows that $\frac{dX_{ii}}{d\tau} = \frac{dX_{ii}}{d\tau} \epsilon_i$, where $\epsilon_i > 0$, we can write:

$$\text{Cov}_F \left[ \mathbb{E}_i [D], \frac{dX_{ii}}{d\tau} \right] = \frac{-P_1}{A \text{Var} [D]} \text{Cov}_F \left[ \mathbb{E}_i [D], \text{sgn} \left( \mathbb{E}_i [D] - A \text{Cov} [E_{2i}, D] - P_1 (1 + \text{sgn} (\Delta X_{1i}) \tau) - X_{0i} \right) \epsilon_i \right]$$

The argument can be made from equation (23): as long as fundamental trading needs are orthogonal to investors’ beliefs, we expect [OBPS] to hold. The term $\mathbb{E}_i [D]$ appears on both random variables that form part of the covariance. If the rest of the terms of $\text{sgn} (\Delta X_{1i})$ were constant for all $i$ — that would correspond to the case when all trading is driven by disagreement, the covariance would simply become (approximately, disregarding the $\epsilon_i$ correction, which simply changes weights) the cross sectional variance of $\mathbb{E}_i [D]$, which is strictly positive as long as investors disagree. Fundamental trading can make some optimists sellers and some pessimists buyers, but in order to overturn [OBPS], a particular pattern of fundamental trading is required.

Proposition 1. (Effect of a marginal tax change on welfare)

a) By taking the derivative of social welfare, it follows that $\frac{dV_i}{d\tau} = \int \lambda_i \frac{dV_i}{d\tau} dF (i)$. Taking the derivative of $V_i$ with respect to $\tau$, we can write:

$$\frac{dV_i}{d\tau} = \mathbb{E} \left[ U'_i (W_{2i}) \right] \left[ \mathbb{E} [D] - A_i \text{Cov} [E_{2i}, D] - P_1 - A_i X_{1i} \text{Var} [D] \right] \frac{dX_{ii}}{d\tau} - \Delta X_{1i} \frac{dP_1}{d\tau}$$

The derivative $\frac{dX_{ii}}{d\tau}$ is the equilibrium portfolio response. For those investors who decide not trade $\frac{dV_i}{d\tau} = 0$, because $\frac{dX_{ii}}{d\tau} = 0$ and $X_{1i} = X_{0i}$. This equation uses the envelope theorem at the extensive margin between trading and no trading. Substituting the first order condition (4) into the previous expression, we find:

$$\frac{dV_i}{d\tau} = \mathbb{E} \left[ U'_i (W_{2i}) \right] \left[ \mathbb{E} [D] - \mathbb{E}_i [D] + \text{sgn} (\Delta X_{1i}) P_1 \frac{dX_{ii}}{d\tau} - \Delta X_{1i} \frac{dP_1}{d\tau} \right]$$
Substituting this last expression into \( \frac{dV}{d\tau} \), we recover the result in proposition 1a.

b) Imposing \( \lambda_i E_i \left[ U_i' \left( W_{2i} \right) \right] = 1 \) allows us to write:

\[
\frac{dV_i}{d\tau} = \int \left[ E \left[ D \right] - E_i \left[ D \right] + \text{sgn} \left( \Delta X_{1i} \right) P_1 \frac{dX_{1i}}{d\tau} \right] dF \left( i \right)
\]

Market clearing implies that \( \int \Delta X_{1i} dF \left( i \right) = 0 \), which allows to cancel the term multiplying \( \frac{dP_1}{d\tau} \). Market clearing also implies that \( \int \frac{dX_{1i}}{d\tau} dF \left( i \right) = 0 \), which allows to cancel the term that corresponds to the planner’s measure \( E \left[ D \right] \). This yields equation (10).

c) Follows directly by imposing \( \tau = 0 \) in (10).

**Proposition 2. (Optimal financial transaction tax)**

a) By solving for \( \tau^* \) in (10), we recover equation (11). By rearranging terms, we can write:

\[
\tau^* = \frac{1}{2} \frac{\int_{i \in B} E_i \left[ D \right] \frac{dX_{1i}}{d\tau} dF \left( i \right)}{\int_{i \in B} \frac{dX_{1i}}{d\tau} dF \left( i \right)} = \frac{1}{2} \left[ \int_{i \in B} \frac{E_i \left[ D \right]}{P_1} \frac{dX_{1i}}{d\tau} dF \left( i \right) \right] - \frac{1}{2} \int_{i \in \bar{S}} \frac{E_i \left[ D \right]}{P_1} \frac{dX_{1i}}{d\tau} dF \left( i \right),
\]

where I have used the fact that \( \int_{i \in B} E_i \left[ D \right] \frac{dX_{1i}}{d\tau} dF \left( i \right) = \int_{i \in \bar{S}} E_i \left[ D \right] \frac{dX_{1i}}{d\tau} dF \left( i \right) \).  

b) Because \(-1 < \tau < 1 \) and \( V \left( \cdot \right) \) is continuous, by the extreme value theorem we conclude that there exists a maximum. I show when proving proposition 4 that the planner’s problem is strictly convex when \( \tau \leq 0 \). This fact, combined with assumption [OBPS], guarantees that the maximum occurs in the \( \tau > 0 \) region. Because of the non-convexity of the problem, equation (11) may have multiple solutions, which correspond to local optima, but only one of them will be the global optimum.

**Proposition 3. (Harberger (1964) revisited)**

a) When there is no dispersion in the beliefs of investors and the planner or the planner assesses social welfare respecting individual beliefs, we can write the marginal change in welfare as a money-metric (divided by investors’ marginal utility) as:

\[
\left. \frac{d\hat{V}_i}{d\tau} \right|_{\tau=\hat{\tau}} = \frac{\frac{dV_i}{d\tau}}{E \left[ U_i' \left( W_{2i} \right) \right]} = \text{sgn} \left( \Delta X_{1i} \right) \frac{\hat{\tau} P_1 \frac{dX_{1i}}{d\tau} - \Delta X_{1i} \frac{dP_1}{d\tau}}{E \left[ U_i' \left( W_{2i} \right) \right]},
\]

Adding up across all investors, and using the fact that \( \int \text{sgn} \left( \Delta X_{1i} \right) \frac{dX_{1i}}{d\tau} dF \left( i \right) = 2 \int_{i \in B} \frac{dX_{1i}}{d\tau} dF \left( i \right) \), we then recover equation (12).

b) The result in a) is an exact expression. However, we can write a second order approximation around \( \tau = 0 \) of the marginal change in social welfare. Note all terms corresponding to terms-of-trade cancel out after imposing market clearing, so I do not consider them. The first term of the Taylor expansion is given above. The derivative of the second term of the Taylor expansion is given by: \( \text{sgn} \left( \Delta X_{1i} \right) P_1 \frac{d^{2}X_{1i}}{d\tau^2} + \text{sgn} \left( \Delta X_{1i} \right) \tau P_1 \frac{dX_{1i}}{d\tau} \). Around \( \tau = 0 \), this becomes \( \text{sgn} \left( \Delta X_{1i} \right) P_1 \frac{d^{2}X_{1i}}{d\tau^2} \). Hence, when \( \tau = 0 \) we can write:

\[
\int d\hat{V}_i \big|_{\tau=0} dF \left( i \right) \approx \int \text{sgn} \left( \Delta X_{1i} \right) \tau P_1 \frac{dX_{1i}}{d\tau} \big|_{\tau=0} dF \left( i \right) \left( d\tau \right) + \frac{1}{2} \left( \text{sgn} \left( \Delta X_{1i} \right) P_1 \frac{dX_{1i}}{d\tau} + \text{sgn} \left( \Delta X_{1i} \right) \tau P_1 \frac{d^{2}X_{1i}}{d\tau^2} \right) \big|_{\tau=0} dF \left( i \right) \left( d\tau \right)^2
\]

\[
= P_1 \hat{\tau}^2 \int_{i \in B} \frac{dX_{1i}}{d\tau} \big|_{\tau=0} dF \left( i \right)
\]

43
Proposition 4. (Non-convex planner’s problem)

a) The example in the text is enough to support this claim.

b) We can write $\frac{dV}{d\tau}$ as:

$$
\frac{dV}{d\tau} = \int [-E_i[D] + \text{sgn}(\Delta X_{1i}) P_1 \tau^i dF(i) = - \int \mathbb{E}_i[D] \frac{dX_{1i}}{d\tau} dF(i) + 2P_1 \tau \int_{i \in \mathcal{B}} \frac{dX_{1i}}{d\tau} dF(i) 
$$

$$
= \frac{-P_1}{\mathbb{V}ar[D]} \left[- \int \mathbb{E}_i[D] \frac{\text{sgn}(\Delta X_{1i})}{A_i} \epsilon_i dF(i) + 2P_1 \tau \epsilon \int_{i \in \mathcal{B}} \frac{1}{A_i} dF(i) \right],
$$

where I have used the result from lemma 1 that $\frac{dX_{1i}}{d\tau} = \frac{\partial X_{1i}}{\partial \epsilon_i}$ and market clearing. Hence, as long as there are no extensive margin changes, the sign of $\frac{d^2V}{d\tau^2}$ is only determined by $\frac{d(P_1 \tau)}{d\tau}$. But it is easy to show that $\frac{d(P_1 \tau)}{d\tau} = P_1 \left(1 - \frac{\tau \int \text{sgn}(\Delta X_{1i}) dF(i)}{\int \frac{1 + \text{sgn}(\Delta X_{1i})}{A_i} dF(i)} \right) > 0$. This shows that non-convexities can only arise when the composition of active investors in the economy varies with $\tau$.

This result creates an asymmetry between the $\tau > 0$ and the $\tau \leq 0$ region. In the $\tau \leq 0$ region, social welfare is always convex because there are no extensive margin changes. When $\tau > 0$, we would have to impose restrictions on the distribution of investors to find a unique maximum. Even if [OBPS] does not hold, we can only guarantee that there will be a local maximum in the $\tau \leq 0$ region, but the global maximum could entail a positive tax, as in the the example in the text.

Differentiability

From equation (6) it is straightforward to show that the equilibrium price $P_1$ is continuous in $\tau$. As shown in the proof of lemma 1, for the values of $\tau$ such that $P_1$ is differentiable we have that $\frac{dP_1}{d\tau} = -\frac{P_1 \int_{i \in \mathcal{T}} \frac{\text{sgn}(\Delta X_{1i})}{A_i} dF(i)}{\int_{i \in \mathcal{T}} \frac{1 + \text{sgn}(\Delta X_{1i})}{A_i} dF(i)}$. Hence, when the distribution $F$ of investors is continuous across all dimensions of heterogeneity, $\frac{dP_1}{d\tau}$ is continuous, since both numerator and denominator change smoothly with $\tau$, which implies that $P_1$ is differentiable. When $F$ has mass points, $\frac{dP_1}{d\tau}$ is discontinuous for those values of $\tau$ in which a positive measure of investors becomes inactive, since the left and right limits to $\int_{i \in \mathcal{T}} \frac{\text{sgn}(\Delta X_{1i})}{A_i} dF(i)$ differ at that value of $\tau$. Under the assumed regularity conditions on $F$, $P_1$ is differentiable almost everywhere.

Because $X_{1i}$ is continuous in $\tau$ and $P_1$, the equilibrium allocations inherit the differentiability properties of $\frac{dP_1}{d\tau}$. Hence, social welfare $V(\tau)$ also inherits the differentiability properties of $\frac{dP_1}{d\tau}$. Throughout the paper I implicitly assume that it is never the case that the planner’s problem reaches its optimum at a point of non-differentiability. A sufficient condition for this to hold is that the distribution of investors is continuous – I use a discrete number of investors in the simulations only for numerical tractability.

Summing up: when the distribution of investors $F$ is continuous, $P_1$, $X_{1i}$ and $V$ are differentiable. When $F$ has mass points, there are a finite number of values for $\tau$ at which $P_1$, $X_{1i}$ and $V$ are non-differentiable.

Proposition 5. (Intermediate target implementation)

a) We can decompose total volume as:

$$
\int_{i \in \mathcal{B}} \Delta X_{1i} dF(i) = \frac{1}{2} \left( \int_{i \in \mathcal{B}} \Delta X_{1i} dF(i) - \int_{i \in \mathcal{S}} \Delta X_{1i} dF(i) \right) = \Theta_F + \Theta_{NF} + \Theta_{\tau}
$$
where

\[ \Theta_{NF} \equiv \frac{1}{2} \left( \int_{j \in B} \frac{E_i[D]}{P_i} A_i \text{Var}[D] dF(i) - \int_{i \in S} \frac{E_i[D]}{P_i} A_i \text{Var}[D] dF(i) \right) \]

\[ \Theta_F \equiv \frac{1}{2} \left( \int_{i \in S} \frac{\text{Cov}[E_{2i}, D]}{\text{Var}[D]} dF(i) - \int_{i \in B} \frac{\text{Cov}[E_{2i}, D]}{\text{Var}[D]} dF(i) \right) \]

\[ + \left( \int_{i \in S} \frac{P_i}{A_i \text{Var}[D]} dF(i) - \int_{i \in B} \frac{P_i}{A_i \text{Var}[D]} dF(i) \right) \]

\[ \Theta_{\tau} \equiv -\frac{1}{2} \tau \left( \int_B \frac{P_i}{A_i \text{Var}[D]} dF(i) + \int_S \frac{P_i}{A_i \text{Var}[D]} dF(i) \right) \]

Note that \( \Theta_F \) can be decomposed into a risk sharing motive, due to different hedging needs, a risk transfer motive, due to different levels of risk aversion among investors, and a motive linked to starting with different asset holdings. Note also that \( \Theta_{\tau} \) can be written as:

\[ \Theta_{\tau} = \frac{1}{2} \tau \left( \int_B \frac{\partial X_{1i}}{\partial \tau} dF(i) - \int_S \frac{\partial X_{1i}}{\partial \tau} dF(i) \right) = \frac{1}{2} \tau \left( \int \text{sgn}(\Delta X_{1i}) \frac{\partial X_{1i}}{\partial \tau} dF(i) \right) \]

b) The value of \( \tau \) that makes \( \Theta_F = \Theta_{\tau} \) is

\[ \tau = \frac{\int_{i \in B} \frac{E_i[D]}{P_i} A_i \text{Var}[D] dF(i) - \int_{i \in S} \frac{E_i[D]}{P_i} A_i \text{Var}[D] dF(i)}{\int \text{sgn}(\Delta X_{1i}) \frac{\partial X_{1i}}{\partial \tau} dF(i)} = \frac{\int \frac{E_i[D]}{P_i} \frac{\partial X_{1i}}{\partial \tau} dF(i)}{\int \text{sgn}(\Delta X_{1i}) \frac{\partial X_{1i}}{\partial \tau} dF(i)} = \frac{\int \frac{E_i[D]}{P_i} \frac{\partial X_{1i}}{\partial \tau} dF(i)}{\int \text{sgn}(\Delta X_{1i}) \frac{\partial X_{1i}}{\partial \tau} dF(i)} \]

which equals the optimal tax as long as \( \frac{\partial X_{1i}}{\partial \tau} = \frac{dX_{1i}}{d\tau} \). But this only occurs when \( \frac{dP_i}{d\tau} = 0 \) or, equivalently, when \( \int_B \frac{1}{A_i} dF(i) = \int_S \frac{1}{A_i} dF(i) \).

Section 4

Proposition 6. (Pre-existing trading costs)

Given investors’ optimal portfolios, stated in the main text, it is straightforward to derive the equilibrium price, which is given by:

\[ P_1 = \frac{\int_{i \in T} \frac{E_i[D]}{A_i \text{Var}[D] + \alpha} \frac{\partial X_{1i}}{\partial \tau} dF(i) - \int_{i \in T} X_{0i} \frac{1 + \text{sgn}(\Delta X_{1i})(\tau + \psi)}{A_i \text{Var}[D] + \alpha}}{\int_{i \in T} X_{0i}} \]

Indirect utility for an investor \( i \) from the planner’s perspective is given by:

\[ V_i = -e^{-A_i \left( \frac{(E[D]) - A_i \text{Cov}[E_{2i}, D] - P_i X_{1i} + P_i X_{0i} - |\Delta X_{1i}| P_i \psi - \frac{\partial}{\partial \tau} (\Delta X_{1i})^2}{A_i \text{Var}[D] + \alpha} \right) \]

Note that only the resources corresponding to the transaction tax are rebated back to investors. All resources devoted to transaction costs are a compensation for the use of resources, so the planner does not have to account for them explicitly, since they form part of a zero profit condition.

Hence, the marginal change in welfare for an investor \( i \) is given by:

\[ \frac{dV_i}{d\tau} = \mathbb{E} [U_i'(W_2)] \left( (E[D] - A_i \text{Cov}[E_{2i}, D] - P_i - \text{sgn}(|\Delta X_{1i}|) P_i \psi - \text{sgn}(|\Delta X_{1i}|) \frac{dX_{1i}}{d\tau} - \alpha |\Delta X_{1i}| + A_i \text{Var}[D] X_{1i}) \frac{dX_{1i}}{d\tau} - \Delta X_{1i} \frac{dP_i}{d\tau} \right) \]

45
By substituting investors’ first order conditions, we find:

$$\frac{dV_i}{d\tau} = \mathbb{E}[U_i'(W_{2i})] \left[ \mathbb{E}[D] - \mathbb{E}_i[D] + \text{sgn} (\Delta X_{ij}) p_j \tau \right] \frac{dX_{ij}}{d\tau} - \Delta X_{ij} \frac{dP_i}{d\tau}$$

Which is identical to the condition derived when showing proposition 1. It follows that the optimal tax has the same expression as in (11).

**Proposition 7. (Multiple risky assets)**

After getting rid of terms that do not affect the maximization problem, investors solve:

$$\max_{\mathbf{x}_{ij}} (\mathbb{E}_i[\mathbf{d}] - A_i \text{Cov}[E_{2i}, \mathbf{d}] - \mathbf{p}) - |\mathbf{x}'_{ij} - \mathbf{x}_{0i}'| \mathbf{p} \tau - \frac{A_i}{2} \mathbf{x}'_{ij} \text{Var}[\mathbf{d}] \mathbf{x}_{ij}$$

Where I use $|\mathbf{x}'_{ij} - \mathbf{x}_{0i}'|$ to denote the vector of absolute values of the difference between both vectors. This problem is convex, so the first order condition fully characterizes investors’ optimal portfolios as long as they trade a given asset $j$:

$$\mathbf{x}_{ij} = (A_i \text{Var}[\mathbf{d}])^{-1} (\mathbb{E}_i[\mathbf{d}] - A_i \text{Cov}[E_{2i}, \mathbf{d}] - \mathbf{p} - \hat{\mathbf{p}}_j \tau),$$

where $\hat{\mathbf{p}}_j$ is a $J \times 1$ vector where a given row $j$ is given by $\text{sgn} (\Delta X_{1ij}) p_j$. If an asset $j$ is not traded by an investor $i$, then $X_{1ij} = X_{0ij}$. The inaction regions are defined analogously to the one asset case. Note that there exists a way to write optimal portfolio choices only with matrix operations; however, the notation turns out to be more cumbersome. The equilibrium price vector is given by:

$$\mathbf{p} = \int \left( 1 + \frac{s_i}{A_i} \right) dF (i) = \int \frac{\mathbb{E}_i[\mathbf{d}]}{A_i} dF (i) - \int (\text{Cov}[E_{2i}, \mathbf{d}] + \text{Var}[\mathbf{d}] \mathbf{x}_{0i}) dF (i)$$

Where I denote element-by-element multiplication as $y \times z$ and use $s_i$ to denote a $J \times 1$ vector given by $\text{sgn} (\Delta X_{1ij})$.

$$\frac{dV}{d\tau} = \int \lambda_i \mathbb{E}[U_i'(W_{2i})] \left[ (\mathbb{E}_i[\mathbf{d}] - \mathbb{E}_i[\mathbf{d}] + \hat{\mathbf{p}}_i \tau)' \frac{d\mathbf{x}_{ij}}{d\tau} \frac{dF (i)}{d\tau} - (\mathbf{x}_{ij} - \mathbf{x}_{0ij})' \frac{dp}{d\tau} \right] dF (i)$$

This is a generalization of the one asset case. Under assumption [NR], we can write in product notation:

$$\int \sum_{j=1}^J \left( -\mathbb{E}_i[D_j] + \text{sgn} (\Delta X_{1ij}) p_j \tau \right) \frac{dX_{ij}}{d\tau} dF (i) = 0$$

So the optimal tax becomes:

$$\tau^* = \frac{\sum_{j=1}^J \int \mathbb{E}_i[D_j] \frac{dX_{ij}}{d\tau} dF (i)}{\sum_{j=1}^J \int \text{sgn} (\Delta X_{1ij}) p_j \frac{dX_{ij}}{d\tau} dF (i)}$$

Which can be rewritten as:

$$\tau^* = \frac{\sum_{j=1}^J \int \text{sgn} (\Delta X_{1ij}) p_j \frac{dX_{ij}}{d\tau} dF (i) \tau_j^*}{\sum_{j=1}^J \int \text{sgn} (\Delta X_{1ij}) p_j \frac{dX_{ij}}{d\tau} dF (i)}$$

Where $\tau_j^* = \frac{\int \mathbb{E}_i[D_j] \frac{dX_{ij}}{d\tau} dF (i)}{\int \text{sgn} (\Delta X_{1ij}) \frac{dX_{ij}}{d\tau} dF (i)}$. And by defining weights $\omega_i = \frac{\int \text{sgn} (\Delta X_{1ij}) p_j \frac{dX_{ij}}{d\tau} dF (i)}{\sum_{j=1}^J \int \text{sgn} (\Delta X_{1ij}) p_j \frac{dX_{ij}}{d\tau} dF (i)}$, we recover equation (16).
Proposition 8. (General utility and arbitrary beliefs)

a) Social welfare is given by \( V(\tau) = \int \lambda_i V_i dF(i) \), where \( V_i \) denotes indirect utility from the planner’s perspective, that is:

\[
V_i = U_i (E_1 + X_0 P_1 - X_{ii} P_1 - Y_{ii}) + \beta E [U_i (E_{2i} + X_{ii} D + Y_{ii})],
\]

where \( X_{ii} \) and \( Y_{ii} \) are chosen optimally. We could thus write the marginal change in welfare as:

\[
\frac{dV}{d\tau} = \int \lambda_i U'_i (C_{ii}) \left( \frac{P_i \text{sgn}(\Delta X_{ii}) \tau - \beta \text{Cov} \left[ Z_{ii}, \frac{U_i(C_{ii})}{U'_i(C_{ii})} D \right]}{-\beta \text{Cov} \left[ Z_{ii}, \frac{U_i(C_{ii})}{U'_i(C_{ii})} D \right] - \Delta X_{ii} \frac{dP_i}{d\tau}} \right) dF(i)
\]

Solving for the optimal tax yields equation (17) in the paper.

b) Assuming that \( \lambda_i = 1 \), taking the limit when \( U'_i (C_{ii}) \) and using the optimality condition for \( Y_{ii} \), we find:

\[
\tau^* = \frac{\int \text{Cov} \left[ Z_{ii}, \frac{D}{P_i} \right] \frac{dX_{ii}}{d\tau} dF(i)}{\int \text{sgn}(\Delta X_{ii}) \frac{dX_{ii}}{d\tau} dF(i)} = \frac{\int E_i [D] \frac{dX_{ii}}{d\tau} dF(i)}{\int \text{sgn}(\Delta X_{ii}) \frac{dX_{ii}}{d\tau} dF(i)},
\]

using the fact that \( E_i [D] = \text{Cov}(Z_{ii}, D) + E [D] \).

Proposition 9. (Asymmetric taxes on buyers versus sellers)

a) The budget constraint for an investor in this case can be written as:

\[
W_{2i} = E_{2i} + X_{ii} D + (X_0 P_1 - X_{ii} P_1 - \tau_B |P_1| \Delta X_{ii} + \tau_S |P_1| \Delta X_{ii} + T_{ii})
\]

The first order condition becomes:

\[
X_{ii} = \frac{E_i [D] - A_i \text{Cov} [E_{2i}, D] - P_1 (1 + \mathbb{I} [\Delta X_{ii} > 0] \tau_B + \mathbb{I} [\Delta X_{ii} < 0] \tau_S)}{A_i \text{Var} [D]}
\]

With an equilibrium price:

\[
P_1 = \frac{\int_{i \in T} \left( \frac{E_i [D] - A_i \text{Cov} [E_{2i}, D] - \text{Var} [D] X_{ii}}{A_i} \right) dF(i)}{\int_{i \in T} \frac{1}{A_i} + \tau_B \int_{i \in B} \frac{1}{A_i} - \tau_S \int_{i \in S} \frac{1}{A_i} dF(i)}
\]

In this case we can write: \( X_{ii} (\tau_j, P_1 (\{\tau_j\}), \) where \( \{\tau_j\} \) denotes a vector of taxes. This implies that \( \frac{dX_{ii}}{d\tau_j} = \frac{dX_{ii}}{d\tau_j} + \frac{dX_{ii}}{dP_j} \frac{dP_j}{d\tau_j} \). The change in social welfare for an investor \( i \) when varying a tax \( \tau_j \), from a planner’s perspective, is given by:

\[
\frac{dV}{d\tau_j} = \mathbb{E} \left[ U'_i (W_{2i}) \right] \left( (E_i [D] - A_i \text{Cov} [E_{2i}, D] - P_1 - A_i X_{ii} \text{Var} [D]) \frac{dX_{ii}}{d\tau_j} - \Delta X_{ii} \frac{dP_1}{d\tau_j} \right)
\]

Social welfare is then:

\[
\frac{dV}{d\tau_j} = \int \lambda_i \mathbb{E} \left[ U'_i (W_{2i}) \right] \left( (E_i [D] - E_i [D] + P_1 (\mathbb{I} [\Delta X_{ii} > 0] \tau_B - \mathbb{I} [\Delta X_{ii} < 0] \tau_S)) \frac{dX_{ii}}{d\tau_j} - \Delta X_{ii} \frac{dP_1}{d\tau_j} \right) dF(i)
\]

Any tax change has two direct effects. First, it marginally affects those investors who pay that tax at the margin. Second, it moves prices. This price change creates two effects. There is a first effect working through terms-of-trade. A second effect works through demand changes. Under the usual
differentiability and convexity assumptions, the optimal tax is characterized by \( \frac{dV}{d\tau_j} = 0, \forall j \). This yields a system of equation in the vector of taxes:

\[
0 = \int \lambda_i \mathbb{E} \left[ U'_i (W_{2i}) \right] \left( (\mathbb{E} [D] - \mathbb{E}_i [D]) + P_1 \text{sgn} (\Delta X_{1i}) \tau_i \right) \frac{dX_{1i}}{d\tau_j} - \Delta X_{1i} \frac{dP_i}{d\tau_j} \right) dF (i) , \forall j
\]

This equation characterizes a system of equations in \( \tau_B \) and \( \tau_S \).

When [NR] holds, that is \( \lambda_i \mathbb{E} \left[ U'_i (W_{2i}) \right] = 1, \forall i \), we can write:

\[
\frac{dV}{d\tau_j} = \int (\mathbb{E} [D] - \mathbb{E}_i [D]) \frac{dX_{1i}}{d\tau_j} dF (i) + P_1 \int_{i \in B} \frac{dX_{1i}}{d\tau_j} dF (i) - \tau_B \int_{i \in B} \frac{dX_{1i}}{d\tau_j} dF (i) - \tau_S \int_{i \in S} \frac{dX_{1i}}{d\tau_j} dF (i)
\]

Using market clearing, we can find:

\[
\frac{dV}{d\tau_j} = - \int \mathbb{E}_i [D] \frac{dX_{1i}}{d\tau_j} dF (i) + P_1 (\tau_B + \tau_S) \int_{i \in B} \frac{dX_{1i}}{d\tau_j} dF (i)
\]

Solving for \( \tau_B + \tau_S \):

\[
\tau_B + \tau_S = \frac{\int \mathbb{E}_i [D] \frac{dX_{1i}}{d\tau_j} dF (i)}{P_1 \int_{i \in B} \frac{dX_{1i}}{d\tau_j} dF (i)} , \forall j
\]

In general this gives a system of non-linear equations in \( \tau_B + \tau_S \). When there are two investors, the two equations become collinear, because of market clearing:

\[
\int \mathbb{E}_i [D] \frac{dX_{1i}}{d\tau_j} dF (i) = (\mathbb{E}_B [D] - \mathbb{E}_S [D]) \int_{i \in B} \frac{dX_{1i}}{d\tau_j} dF (i)
\]

In that case only the sum of taxes is pinned down:

\[
\tau_B + \tau_S = \frac{\mathbb{E}_B [D] - \mathbb{E}_S [D]}{P_1}
\]

**Proposition 10. (Individual taxes/First-best)**

a) In the case with \( I \) taxes and \( I \) investors, the first order conditions under assumption [NR] for the planner become:

\[
\frac{dV}{d\tau_j} = \sum_i \left( -\mathbb{E}_i [D] + P_1 \text{sgn} (\Delta X_{1i}) \tau_i \right) \frac{dX_{1i}}{d\tau_j} F (i) = 0, \forall j
\]

This system of equations characterizes the set of optimal taxes. Note that one solution to this system is given by:

\[
-\mathbb{E}_i [D] + P_1 \text{sgn} (\Delta X_{1i}) \tau_i = -F
\]

Where \( F \) is an arbitrary real number. Rearranging this expression we can find equation (19).

b) Starting from the system of equations which characterizes the optimal set of taxes, we can write, using market clearing \( F (j) \frac{dX_{1j}}{d\tau_j} + \sum_{i \neq j} \frac{dX_{1j}}{d\tau_j} F (i) = 0 \), the following set of equations:

\[
\sum_{i \neq j} \left( \mathbb{E}_i [D] - \mathbb{E}_i [D] \right) \frac{dX_{1i}}{d\tau_j} F (i) + P_1 \left( -\text{sgn} (\Delta X_{1i}) \tau_i \sum_{i \neq j} \frac{dX_{1i}}{d\tau_j} F (i) + \sum_i \left( \text{sgn} (\Delta X_{1i}) \tau_i \frac{dX_{1i}}{d\tau_j} F (i) \right) \right) = 0
\]

For all equations but for the one with respect to tax \( j \). To show that this system only depends on \( N - 1 \) taxes, we simply need to show that all \( \frac{dX_{1i}}{d\tau_j} \) do not depend on the tax \( \tau_j \). Note that \( \frac{dX_{1i}}{d\tau_j} = \frac{\partial X_{1i}}{\partial \tau_j} + \frac{\partial X_{1i}}{\partial \tau_j} \frac{dP_i}{d\tau_j} \). But when \( i \neq j \) then \( \frac{dX_{1i}}{d\tau_j} \) only depends on \( \frac{dP_i}{d\tau_j} \) because \( \frac{\partial X_{1i}}{\partial \tau_j} \) equals zero and \( \frac{\partial X_{1i}}{\partial \tau_j} \) does not depend on \( \tau_j \). We just need to show that \( \frac{dP_i}{d\tau_j} \) can be written as a function of all other taxes but \( \tau_j \). This can be easily shown combining the expressions used to show lemma 1 with market clearing conditions.
Section 5

The expression for the asset price in (24) now yields a demand curve for shares.

\[ P_t^* = \int_{\mathcal{E}} \left( \frac{E[D]}{A_t} - \Lambda \left( \text{Cov} \left[ E_2, D \right] + \text{Var} \left[ D \right] X_0 \right) \right) dF(i) - \text{AVar} \left[ D \right] S_{1k} \]

The demand by investors for the risky asset is identical to the baseline model. The equilibrium price is now determined by the intersection of equation (24) and the supply curve, given by \( P_t^* = \gamma_1 + \gamma_2 S_{1k} \).

Lemma 3. (Effect of \( \tau \) on prices and allocations)

a) After writing the market clearing condition as \( \int (X_{1i} - X_0) dF(i) = F(P_t) \), where \( F(\cdot) = \Phi^{-1}(\cdot) \) is an upward sloping function, we can derive \( \frac{dP_t}{d\tau} = \frac{\int \Phi^{-1}(\frac{\partial \text{Cov}(\tau)}{\partial \tau}) dF(i)}{F'(P_t) - \int \frac{\partial \text{Var}(\tau)}{\partial \tau} dF(i)} \). Using the same argument as in lemma 1, \( \frac{dP_t}{d\tau} \) can have any sign, depending on its numerator. Using a similar approach to lemma 2, we can write \( \frac{dX_{1i}}{d\tau} = \frac{\partial X_{1i}}{\partial \tau} \epsilon_i \), where \( \epsilon_i \), which is constant within buyers/sellers, can be written as \( \epsilon_i = 1 - (\text{sgn} (\Delta X_{1i} + \tau)) \frac{1-H}{\int_{B} \frac{1}{F(i)} + 1 + \tau + H(1-\tau)} \) and \( H = \int_{B} \frac{1}{F(i)} \) \( \in (0, \infty) \). It is easy to show that \( \epsilon_i > 0 \), which proves the result.

Proposition 11. (Optimal tax with production)

a) The marginal change in social welfare is given by:

\[ \frac{dV}{d\tau} = \int \lambda_i \text{E} \left[ U'_i(W_2) \right] \left( \text{E}[D] - \text{E}_i[D] + \text{sgn} (\Delta X_{1i}) P_t \right) \frac{dX_{1i}}{d\tau} dF(i) \]

Imposing assumption [NR], and using market clearing, the marginal change in social welfare as:

\[ \frac{dV}{d\tau} = \int \left( \text{E}[D] - \text{E}_i[D] + \text{sgn} (\Delta X_{1i}) P_t \right) \frac{dX_{1i}}{d\tau} dF(i) \]

b) Solving for \( \tau^* \) in the previous expression yields equation (20). We can re-write the numerator of the optimal tax as:

\[ \int (\text{E}[D] - \text{E}_i[D]) \frac{dX_{1i}}{d\tau} dF(i) = \zeta(\tau) \text{E}_{F,T} \left[ \text{E}[D] - \text{E}_i[D] \right] \frac{dX_{1i}}{d\tau} \]

\[ = \zeta(\tau) \left( \text{Cov}_{F,T} \left[ \text{E}[D] - \text{E}_i[D], \frac{dX_{1i}}{d\tau} \right] + \text{E}_{F,T} \left[ \text{E}[D] - \text{E}_i[D] \right] \frac{dX_{1i}}{d\tau} \right) \]

\[ = -\zeta(\tau) \text{Cov}_{F,T} \left[ \text{E}_i[D], \frac{dX_{1i}}{d\tau} \right] + \text{E}[D] - \text{E}_{F,T} \left[ \text{E}_i[D] \right] \frac{dS_{1k}}{d\tau} \]

Where we define \( \zeta(\tau) = \int_{F,T} dF(i) \); this normalization by the number of active investors is necessary to use expectation and covariance operators. Using the fact that \( \int_{F,T} dX_{1i} dF(i) = \frac{dS_{1k}}{d\tau} - \int_{E} dX_{1i} dF(i) \), the denominator in (20) can be written as: \( \int \text{sgn} (\Delta X_{1i}) dX_{1i} dF(i) = \int_{E} dX_{1i} dF(i) - \int_{E} dX_{1i} dF(i) = 2 \int_{E} dX_{1i} dF(i) - \frac{dS_{1k}}{d\tau} \). By substituting and rearranging the previous two expressions in the optimal tax formula, we can write \( \tau^* \) as:

\[ \tau^* = \frac{-2 \int_{E} \frac{dX_{1i}}{d\tau} dF(i)}{-2 \int_{E} \frac{dX_{1i}}{d\tau} dF(i) + \frac{dS_{1k}}{d\tau}} \]

\[ \equiv \tau^*_{\text{exchange}} \]

\[ \equiv \tau^*_{\text{production}} \]
Section 6

The value function at $t = 1$ for an investor reads $V_{ii} (X_{0i}, W_{ii}) = \max_{X_i} \mathbb{E} [V_{ii} (X_{1i}, W_{ii})] = e^{-A_i \hat{\tau}}$, where the certainty equivalent $\hat{\tau}$, after normalizing $\mathbb{E} [E_i] - \frac{A_i}{2} \text{Var} [E_i] = 0$, is given by:

$$\hat{\tau}_i = (\mathbb{E}_i [D_3] - P_2 - A_i \text{Cov} [E_{3i}, D_3]) X_{2i} - \frac{A_i}{2} \text{Var} [D_3] (X_{2i})^2 - \tau |P_2| |\Delta X_{2i}| + T_{2i}$$

$$+ (\mathbb{E}_i [D_2] + P_2 - P_1 - A_i \text{Cov} [E_{2i}, D_2]) X_{1i} - \frac{A_i}{2} \text{Var} [D_2] (X_{1i})^2 - \tau |P_1| |\Delta X_{1i}| + T_{1i} + E_{ii} + X_{0i} P_i,$$

and $X_{2i}$ is given by an equivalent expression to (4). The first order condition for $X_{1i}$ is given by:

$$\mathbb{E}_i [D_2] + P_2 (1 + \text{sgn} (\Delta X_{2i}) \tau) - P_1 (1 + \text{sgn} (\Delta X_{1i}) \tau) - A_i \text{Cov} [E_{2i}, D_2] - A_i \text{Var} [D_2] X_{1i}$$

$$+ \frac{dX_{2i}}{dX_{1i}} [\mathbb{E}_i [D_3] - P_2 (1 + \text{sgn} (\Delta X_{2i}) \tau) - A_i \text{Cov} [E_{3i}, D_3] - A_i \text{Var} [D_3] X_{2i}] = 0$$

When $\frac{dX_{2i}}{dX_{1i}} = 0$, the optimal $X_{1i}$ corresponds to (22). When $\frac{dX_{2i}}{dX_{1i}} = 1$, the case in which an investor does not buy or sell at period 2, we can rewrite the optimal $X_{1i}$ as:

$$X_{1i} = \nu \hat{X}_{1i} + (1 - \nu) \hat{X}_{2i}, \quad \text{where}$$

$$\hat{X}_{1i} = \frac{\mathbb{E}_i [D_2] + P_2 (1 + \text{sgn} (\Delta X_{2i}) \tau) - P_1 - A_i \text{Cov} [E_{2i}, D_2] - P_2 (1 + \text{sgn} (\Delta X_{2i}) \tau)}{A_i \text{Var} [D_2]} \quad \text{and} \quad \nu = \frac{\mathbb{E}_i [D_3] - A_i \text{Cov} [E_{3i}, D_3] - P_2 (1 + \text{sgn} (\Delta X_{2i}))}{A_i \text{Var} [D_3]}.$$

Proposition 12. (Optimal tax in dynamic environments)

a) After imposing assumption [NR], we can write $\frac{dV}{dT} = \int \frac{dV}{dT} dF (i)$ as:

$$\frac{dV}{dT} = \int \left( \left[ - \mathbb{E}_i [D_3] + \text{sgn} (\Delta X_{2i}) P_2 \tau \right] \frac{dX_{2i}}{dT} + \left[ - \mathbb{E}_i [D_2] - \text{sgn} (\Delta X_{2i}) P_2 \tau + \text{sgn} (\Delta X_{1i}) P_1 \tau \right] \frac{dX_{1i}}{dT} \right) dF (i)$$

And we can solve for $\tau$ and find:

$$\tau^* = \frac{\int \mathbb{E}_i [D_2] \frac{dX_{2i}}{dT} dF (i) + \int \mathbb{E}_i [D_3] \frac{dX_{1i}}{dT} dF (i)}{\int \text{sgn} (\Delta X_{1i}) \left[ \frac{1 - P_2}{P_1} \text{sgn} (\Delta X_{1i}) \text{sgn} (\Delta X_{2i}) \right] \frac{dX_{1i}}{dT} dF (i) + \int \text{sgn} (\Delta X_{2i}) \frac{dX_{2i}}{dT} dF (i)}$$

The optimal tax can be rewritten in the following way:

$$\tau^* = \omega_1 f_1 \tau_1^* + \omega_2 f_2^*,$$

where $f_1 \equiv \int \frac{\mathbb{E}_i [D_2]}{\text{sgn} (\Delta X_{1i}) \frac{dX_{2i}}{dT}} dF (i)$ and $f_2 \equiv \int \frac{\mathbb{E}_i [D_3]}{\text{sgn} (\Delta X_{2i}) \frac{dX_{1i}}{dT}} dF (i)$ — which are the counterparts of equation (11) — and weights $\omega_1$, $\omega_2$ and $f_1$ given by:

$$\omega_1 = \frac{P_1 \int \text{sgn} (\Delta X_{1i}) (1 - \kappa_{1i}) \frac{dX_{1i}}{dT} dF (i)}{P_1 \int \text{sgn} (\Delta X_{1i}) (1 - \kappa_{1i}) \frac{dX_{1i}}{dT} dF (i) + P_2 \int \text{sgn} (\Delta X_{2i}) \frac{dX_{2i}}{dT} dF (i)}$$

$$\omega_2 = \frac{P_2 \int \text{sgn} (\Delta X_{2i}) \frac{dX_{2i}}{dT} dF (i)}{P_1 \int \text{sgn} (\Delta X_{1i}) (1 - \kappa_{1i}) \frac{dX_{1i}}{dT} dF (i) + P_2 \int \text{sgn} (\Delta X_{2i}) \frac{dX_{2i}}{dT} dF (i)}$$

Note that $f_1$ has a minimum at $\frac{1}{2}$, when $1 - \kappa_{1i} = 2$ for all investors, but it is unbounded above when $1 - \kappa_{1i} \to 0$ for most investors.
B Remarks about modeling choices

Baseline model

- **Sources of belief disagreement:** because it is not crucial for the results, I do not model explicitly the sources of disagreement. Disagreement naturally arises when investors have different learning mechanisms. For instance, they may be overconfident about some pieces of information, as in Scheinkman and Xiong (2003), but they could be subject to alternative behavioral biases. See, for instance, the discussion in Shleifer (2000). Under the “common prior” assumption — see Morris (1995) — there exists a one-to-one mapping between behavioral departures from Bayesian updating in processing information and heterogeneous priors/belief disagreement.

- **Instruments:** as shown in subsection 4.4, because of the second-best nature of the problem addressed, additional instruments are in general welfare improving. For instance, short-sale constraints and borrowing constraints are potentially useful policy instruments, especially in the model with production. Transaction taxes which phase out at a given horizon are also natural policies to use. Allowing some groups of investors to be exempt from paying the tax could also be in theory welfare improving, if they can be properly targeted.

- **Capital taxation:** there exists a well established literature studying the effects of capital taxation on risk-taking – see, for instance, Sandmo (1985) and Atkinson and Stiglitz (1980). Unfortunately, that literature does not provide sharp predictions of the effects of capital taxation on portfolio allocations. Note also that capital taxation would not respect the anonymity assumption.

- **Untaxed riskless asset:** the fact that trading in the riskless asset is not subject to a tax does not modify the main conclusions of the paper. Taxing equally all tradable assets in a portfolio choice problem acts as a lump-sum tax, since every investor is forced to purchase at least an asset to “store” his wealth and all assets face an equivalent tax. With an additional margin of adjustment (e.g., consumption), a trading tax would have an additional effect similar to classic capital taxation rather than trading distortions.

- **Tax definition:** taxes could alternatively be imposed on the total amount traded over a period of time and not on the per trade amount. The more agent-dependent a tax is defined, as opposed to transaction-dependent, the larger the incentive to get around it. An individual specific tax can be overcome at least in two ways. First, investors could create multiple entities and trade with each of them until facing the minimum marginal tax. Second, they could distribute trades among periods to minimize tax liabilities. See Campbell and Froot (1994) for related arguments.

- **Risk neutrality:** welfare losses in this paper arise from the curvature of the utility function. When all investors are risk neutral, any portfolio allocation or production decision yields the same social welfare for the planner. For instance, the planner of this paper would suggest that all allocations in Scheinkman and Xiong (2003) are equally efficient, independently of the level of overconfidence.

- **Exogenous markets:** I assume throughout that the set of assets traded/active markets is exogenously given. Nonetheless, modifying transaction taxes may dramatically affect the set of assets traded. In subsection 4.2, we can think that the planner is endogenously taking into account the possibility that in some market trade ceases to occur. There is scope for better understanding how transaction taxes may shut down specific markets or create new ones.

- **Labor supply/additional choice variables:** The results of the paper do not change when investors’ incomes are endogenously generated. For instance, assume that investors decide to work $N_{it}$ hours at a wage $W_{it}$, maximizing flow utility given by $U_i(C_{it}, N_{it})$. In that case, the optimal tax formula remains unchanged after substituting $U_i(C_{it}, N_{it})$ with $\frac{\partial U_i(C_{it}, N_{it})}{\partial C_{it}}$. If consumption and labor are separable, all the previous results apply directly. This logic, which is identical to the one used to characterize standard Euler equations, applies to any other static choice variable.
C Further intuition on $\frac{dP_1}{d\tau}$

The following three figures provide further intuition for why the price level may go up when transaction taxes are higher. Demand $H$ and $L$ represent the optimal portfolio condition for an optimist and a pessimist respectively – given by equation (4). Total demand is the sum of both curves.

The left figure shows the equilibrium in the baseline model, in which assets are in fixed supply, when $\tau = 0$. In equilibrium, the selling and buying distances — shown in brackets — are equal to each other, once normalized by the number of buyers and sellers. The solid red line (the “demand” curve) shifts up or down (as described in the discussion of lemma 1) when the transaction tax varies.

The middle figure shows a new equilibrium with a positive tax. The dotted blue lines (which are discontinuous capturing the possibility of inaction) are the new levels of demand. Whenever the demand by the $H$ investor is more sensitive to the tax than the demand of the $L$ investor, the price goes down and vice versa. The right figure allows for a flexible supply, as in the production case in section 5. This dampens the effect of a tax on prices, since investment adjusts endogenously.

Figure A.1: Intuition on $\frac{dP_1}{d\tau}$

D Tax on number of shares

I assume in the baseline model that the tax is levied on the dollar value of a trade rather than on the number of shares traded to prevent investors from circumvent it by varying the effective number of shares traded — through a reverse split. All results apply to taxes that depend on the number of shares with minor modifications.

When $P_1$ is exactly zero, a tax based on the dollar volume of the transaction is ineffective. However, a tax based on the number of shares traded $|\Delta X_{1i}|$ can be introduced to effectively tax the notional value of the contract. I extend here propositions 1 and 2 to the case of taxes levied on the number of shares traded. In this case, the distinction between buyers and sellers is somewhat arbitrary, giving support to the idea that both sides of the market should face the same tax.

In the trade region, the optimal portfolio choice of an investor can be written as: $X_{1i} = \frac{E_i[D]-A_i Cov[E_{2i},D] - P_i - sgn(\Delta X_{1i})\tau}{A_i Var[D]}$. The equilibrium price becomes:

$$P_1 = \int_{i \in T} \left( \frac{E_i[D]}{A_i} - A_i \left( Cov[E_{2i},D] - Var[D] X_{0i} \right) - \frac{sgn(\Delta X_{1i}) \tau}{A_i} \right) dF(i)$$

The price corrections is now additive rather than multiplicative. The equivalent to equation (10) is

$$\frac{dV}{d\tau} = \int \left[ -E_i[D] + sgn(\Delta X_{1i}) \tau \right] \frac{dX_{1i}}{d\tau} dF(i).$$

The equivalent to equation (11) is now:

$$\tau^* = \frac{\int E_i[D] \frac{dX_{1i}}{d\tau} dF(i)}{\int sgn(\Delta X_{1i}) \frac{dX_{1i}}{d\tau} dF(i)}$$
This shows that optimal taxes in the paper are written in terms of returns because they are levied on the dollar value of the transaction. When they are levied on the number of shares, the dispersion in expected payoffs rather than the dispersion in expected returns becomes the welfare relevant variable.

E  Parametric belief distribution

Starting from equation (11), it is possible to write the optimal tax as a function of moments of the distribution of beliefs, once a parametric distribution is chosen.

For simplicity, I assume that investors are fully symmetric and that they only trade due to disagreement. Assume that equation (11) holds and that the distribution of beliefs for a fraction $\theta$ of the investors is a normally distributed with mean $\mu$ and variance $\sigma^2$. All remaining $1 - \theta$ investors believe that the expectation of the dividend process equals $\mu$. Note that an investor with belief $x$ only trades in this case when the tax is lower than the absolute value of $x - \mu$. The optimal tax is thus given by the following equation:

$$\tau^* = \theta \frac{\sigma}{P_1} \Lambda \left( \frac{\tau^*}{\sigma P_1} \right),$$

where $\Lambda (\cdot)$ denotes the hazard rate of the normal distribution. It can be easily shown that the value of the tax is an increasing function of $\sigma$. Intuitively, the larger the belief dispersion, the larger the magnitude of non-fundamental trading and the larger the need for a corrective tax. Similar derivations can be done with distributions that truncate easily, like Pareto or Exponential distributions.

F  Linear combination between planner’s belief and investors’ beliefs

Throughout the paper, I assume that the planner maximizes welfare using a single distribution of payoffs for all investors. It is straightforward to generalize the results to a planner that puts weight $\alpha$ on his own belief and weight $1 - \alpha$ on the belief of each investor. In that case, the new optimal tax $\tau^*_\alpha$ looks turns out to be a linear combination of both taxes. In the baseline model, because the optimal tax for a planner that respects investors’ beliefs is $\tau^* = 0$, the optimal tax becomes:

$$\tau^*_\alpha = \alpha \tau^*$$

Where $\tau^*$ is given by equation (11). The same logic applies to other extensions of the baseline model. The case with $\alpha = 1$ is the leading case analyzed in the paper. The case with $\alpha = 0$ is the one used in section 3.5.1.

G  Example with transaction costs

Example 3. (Pre-existing linear costs)

I set $\eta = 0$ and $\alpha = 0$, and compare how the equilibrium and the optimal tax changes when the linear dollar transaction cost $\psi$ varies. Figure A.2 compares trading volume, allocation, prices and welfare (relative to the zero tax case) for $\psi = 0$ and $\psi = 2\%$ in an economy defined in table 5.

<table>
<thead>
<tr>
<th></th>
<th>$\pi_H = 0.1$</th>
<th>$\mathbb{E}_H [D] = 110$</th>
<th>$\text{Cov} [E_{2H}, D] = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pessimistic Sellers</td>
<td>$\pi_L = 0.1$</td>
<td>$\mathbb{E}_L [D] = 92$</td>
<td>$\text{Cov} [E_{2L}, D] = 0$</td>
</tr>
<tr>
<td>Calibrated Buyers</td>
<td>$\pi_{CA} = 0.4$</td>
<td>$\mathbb{E}_{CA} [D] = 100$</td>
<td>$\text{Cov} [E_{2CA}, D] = -5$</td>
</tr>
<tr>
<td>Calibrated Sellers</td>
<td>$\pi_{CB} = 0.4$</td>
<td>$\mathbb{E}_{CB} [D] = 100$</td>
<td>$\text{Cov} [E_{2CB}, D] = 5$</td>
</tr>
</tbody>
</table>

Table 6: Parameters example with transaction costs
In general, when $\psi$ is very small, the optimal tax does not change. However, for larger values of $\psi$, like 2%, changes on the extensive margin directly affect the optimal tax. When $\psi$ is larger, volume is always lower in equilibrium and prices can move in any direction — the logic from lemmas 1 and 2 applies directly here. Welfare is lower in absolute terms — this is intuitive, since now resources are spent in any transaction — but in general the optimal tax with transaction costs can be higher or lower than without transaction costs. This particular parametrization illustrates the possibility that the optimal tax $\tau^*$ can be higher when $\psi = 2\%$ than when there are no transaction costs. This occurs because transactions costs disproportionately reduce trades made due to fundamental reasons. Hence, given that the composition of trading after facing transaction costs, the planner sets a higher transaction since it deters a larger number of non-fundamental trades at the margin.\footnote{The top left plot shows social welfare $V$ for both values of $\psi$. It is expressed in relative terms with respect to the $\tau = 0$ case. The bottom left one shows the equilibrium price. The top right plots show the equilibrium levels of trade for each investor and the bottom right one shows the total amount of trade in equilibrium, given by $\int_{B} X_{1i} dF(i)$. With the exception of the equilibrium allocation plots, the dashed lines represent the $\psi = 2\%$ case, while the solid lines represent the $\psi = 0\%$ case. The vertical dashed lines denote the optimal tax for the $\psi = 2\%$ case, while the vertical solid lines denote the optimal tax for the $\psi = 0\%$ case.}

### H Disagreement about other moments

**Assumptions**

In the baseline model, investors only disagree about the expected value of the payoff of the risky asset. I now assume that investors also hold distorted beliefs about their hedging needs $\text{Cov}_i [E_{2i}, D]$ and about the variance of the payoff of the risky asset $\text{Var}_i [D]$. 

---

54
Results

The optimality condition presented in (4) applies directly, after using the individual beliefs of each investor. Hedging needs enter additively, but perceived individual variances modify the sensitivity of portfolio demands with respect to the baseline case.

Market clearing determines the equilibrium price, given now by:

$$P_1 = \frac{\int_{i \in T} \left( \frac{E_i[D]}{AV_i} - AV \left( \beta_{ii} + X_0i \right) \right) dF(i)}{1 + \tau \int_{i \in T} \frac{\text{sgn}(\Delta X_i)}{AV_i} dF(i)},$$

where $AV \equiv \left( \int_{i \in T} \frac{1}{AV_i} dF(i) \right)^{-1}$ is the harmonic mean of risk aversion coefficients and perceived variances for active investors and $AV_i \equiv \frac{AV_i}{\text{Var}[D]^i}$ is the quotient between investor $i$ risk aversion times perceived variance and the harmonic mean. I define the regression coefficient (beta) of individual endowments $E_i$ on payoffs $D$ perceived by investors by $\beta_{ii} = \frac{\text{Cov}[E_i,D]}{\text{Var}[D]^i}$. Again, $T$ denotes the set of active investors.

**Proposition 13. (Disagreement about second moments)**

a) Under assumption [NR], the marginal change in social welfare from varying the financial transaction tax when investors disagree about second moments is given by:

$$\frac{dV}{d\tau} = \int \left[ (-r_i E_i[D] - A_i \text{Cov}[E_{2i},D] \left( 1 - \frac{\beta_{ii}}{\beta_i} \right) + P_1 r_i (1 + \text{sgn}(\Delta X_i) \tau) \right] \frac{dX_{ii}}{d\tau} dF(i),$$

where $r_i \equiv \frac{\text{Var}[D]^i}{\text{Var}[D]_i}$, $\beta_{ii} \equiv \frac{\text{Cov}[E_{2i},D]}{\text{Var}[D]^i}$, and $\beta_i \equiv \frac{\text{Cov}[E_{2i},D]}{\text{Var}[D]}$. Note that $r_i \in (0,\infty)$ and $\beta_{ii}, \beta_i \in (-\infty,\infty)$.

b) Under assumption [NR], the optimal tax when investors disagree about second moments is given by:

$$\tau^* = \frac{\int \left( r_i E_i[D] + A_i \text{Cov}[E_{2i},D] \left( 1 - \frac{\beta_{ii}}{\beta_i} \right) \right) \frac{dX_{ii}}{d\tau} dF(i)}{P_1 \int (r_i (1 + \text{sgn}(\Delta X_i))) \frac{dX_{ii}}{d\tau} dF(i)}

Proof. See below. 

The formula for the optimal tax now incorporates hedging needs and modifies the weights given to investors’ beliefs. An investor with correct beliefs about second moments has $r_i = 1$ and $\beta_{ii} = \beta_i$; in that case, we recover (11). When investors perceive a high variance, that is, $r_i$ is close to 0, they receive less weight in the optimal tax formula. The opposite occurs when they perceive a low variance. Intuitively, lower perceived variances amplify distortions in expected payoffs, and vice versa.

As in the baseline model, the planner does not need to know the value of $E[D]$ to implement the optimal tax. However, if investors hold distorted beliefs about their hedging needs, the planner needs to know explicitly the magnitude of the mistake. Intuitively, there is no mechanism in the model which cancels out the mistakes in hedging made by investors. The sign of the optimal tax depends directly on the errors made by investors when hedging.

There are two interesting parameters restrictions. First, when investors with correct expected payoffs and hedging betas, that is $\frac{\beta_{ii}}{\beta_i} = 1$, disagree about variances, the optimal tax $\tau^*$ turns out to be:

$$\tau^* = \frac{\mathbb{E}[D] \int r_i \frac{dX_{ii}}{d\tau} dF(i)}{P_1 \int r_i (1 + \text{sgn}(\Delta X_i)) \frac{dX_{ii}}{d\tau} dF(i)}$$

The dispersion of variances, given by $\text{Cov}_{T_i} \left[ r_i \frac{dX_i}{d\tau} \right]$, determines now the sign of the optimal tax. When $r_i$ is constant (although not necessarily equal to one), the optimal tax becomes zero. This reinforces the intuition that belief dispersion is what matters for optimal taxes in an exchange economy. Intuitively,
when buyers, with \( \frac{dX_1}{dt} < 0 \), are relatively aggressive, that is, \( r_i \) is large, they are buying too much of the risky asset, so \( \text{Cov}_{F,T} \left[ r_i, \frac{dX_1}{dt} \right] \) is negative and the optimal tax is positive, and vice versa.

Second, when investors have correct beliefs about the mean and the variance of expected returns, but hedge incorrectly, the optimal tax becomes:

\[
\tau^* = \frac{\mathbb{E} [D] \int A_i (\beta_i - \beta_{ii}) \frac{dX_1}{dt} dF (i)}{P_1 \int \text{sgn} (\Delta X_{1i}) \frac{dX_1}{dt} dF (i)}
\]

The optimal tax now has the opposite sign of \( \text{Cov}_{F,T} \left[ A_i (\beta_i - \beta_{ii}), \frac{dX_1}{dt} \right] \). Intuitively, when buyers, with \( \frac{dX_1}{dt} < 0 \), overestimate their need for hedging and end up buying too much of the risky asset — this occurs when \( \beta_i - \beta_{ii} < 0 \) — the optimal tax is positive, and vice versa.

### I Portfolio constraints: short-sale and borrowing constraints

#### Assumptions

Although participants in financial markets face short-sale and borrowing constraints, investors in the baseline model face no restrictions when choosing portfolios. I now introduce trading constraints into the model as a pair of functions \( \bar{g}_i (\cdot) \) and \( \underline{g}_i (\cdot) \) for every investor \( i \), which can depend on equilibrium prices,\(^{44}\) such that:

\[
\underline{g}_i (P_1) \leq X_{1i} \leq \bar{g}_i (P_1) \tag{26}
\]

Both short-sale constraints and borrowing constraints are special cases of (26). Short-sale constraints are in general price independent and can be written as \( X_{1i} \geq 0 \). Borrowing constraints can be modeled by choosing \( \bar{g}_i (P_1) \) appropriately, such that \( X_{1i} \leq \bar{g}_i (P_1) \). Intuitively, an investor who wants to sufficiently increase his holdings of the risky asset must rely on borrowing. Hence, a borrowing limit is equivalent to an upper bound constraining the amount held of the risky asset.

#### Results

The optimal portfolio is identical to the one in the baseline model, unless a constraint binds. In that case, \( X_{1i} \) equals the trading limit. The equilibrium price is a slightly modified version of (6).

**Proposition 14. (Trading constraints)**

a) Under assumption \([\text{NR}]\), the optimal tax when investors face trading constraints is given by:

\[
\tau^* = \frac{\mathbb{E} [D] \int A_i \left( \bar{g}_i (P_1) - g_i (P_1) \right) g_i' (P_1) \frac{dP_1}{dt} dF (i)}{P_1 \int \text{sgn} (\Delta X_{1i}) \frac{dX_1}{dt} dF (i)},
\]

where \( i = C \) denotes the set of investors with binding trading constraints and \( \hat{X}_{1i} \) denotes the optimal unconstrained portfolio holding for a constrained investor, given in the appendix. Note that I have used the fact that \( \frac{dX_1}{dt} = g_i' (P_1) \frac{dP_1}{dt} \) for constrained investors.

**Proof.** See below. \( \square \)

When trading constraints do not depend on prices that is, \( g_i' (P_1) = 0 \), the optimal tax formula is identical to the one of the baseline model. In those cases, changes in taxes do not modify the portfolio allocation of constrained investors, leaving their welfare unchanged, i.e., for those investors \( \frac{dX_1}{dt} = 0 \). Intuitively, investors with price independent trading constraints are inframarginal for price determination.

\(^{44}\)I assume that investors’ choice sets remain convex after imposing (26). By \( g_i (\cdot) \), I denote either \( \bar{g}_i (\cdot) \) or \( \underline{g}_i (\cdot) \).
When trading constraints depend on prices, the optimal policy takes these effects into account. A marginal tax change modifies asset prices and consequently portfolio allocations for constrained investors; this portfolio change has a first-order effect on welfare. The size of the correction has three determinants. First, it depends on how far the actual portfolio allocation is from the unconstrained portfolio allocation, given by how much the constrained allocation \( g_i(P_1) \) differs from the optimal unconstrained allocation \( \hat{X}_{1i} \). Second, it depends on how sensitive the equilibrium restriction is with respect to asset prices — this is captured by \( g_i'(P_1) \). Third, it depends on how equilibrium prices react to tax changes \( \frac{dP_1}{d\tau} \). If prices remain constant after varying \( \tau \), that is \( \frac{dP_1}{d\tau} = 0 \), the optimal tax formula does not change.

Proofs

**Proposition 13. (Disagreement about second moments)** The optimal portfolio allocation for an investor \( i \) in his trade region is given by:

\[
X_{1i} = \frac{E_i[D] - A_i\text{Cov}_i[E_{2i}, D] - P_1(1 + \text{sgn}(\Delta X_{1i}) \tau)}{A_i\text{Var}_i[D]}
\]

The marginal change in welfare for an investor \( i \) is given by:

\[
\frac{dV_i}{d\tau} = \left[ -r_i(E_i[D] - A_i\text{Cov}_i[E_{2i}, D] - P_1(1 + \text{sgn}(\Delta X_{1i}) \tau)) \frac{dX_{1i}}{d\tau} - \Delta X_{1i} \frac{dP_1}{d\tau} \right]
\]

Where \( r_i = \frac{\text{Var}_i[D]}{\text{Var}_i[D]} \). The change in social welfare can then be written, under assumption [NR], as:

\[
\frac{dV}{d\tau} = \int \left( -r_iE_i[D] - A_i\text{Cov}_i[E_{2i}, D] \left( 1 - \frac{\beta_{1i}}{\beta_i} \right) + P_1 r_i (1 + \text{sgn}(\Delta X_{1i}) \tau) \right) \frac{dX_{1i}}{d\tau} dF(i)
\]

Solving for \( \tau \) in this equation, which corresponds to equation (25) in the paper, delivers the expression for the optimal tax in proposition 13b.

**Proposition 14. (Trading constraints)** a) The equilibrium price is given by:

\[
P_1 = \frac{\int_{i \in \mathcal{T}, U} \frac{E_i[D] - A_i\text{Cov}_i[E_{2i}, D]}{A_i\text{Var}_i[D]} - \int_{i \in \mathcal{T}} X_{0i} - \int_{i \in C} g_i(P_1)}{\int_{i \in \mathcal{T}, U} \frac{1 + \text{sgn}(\Delta X_{1i}) \tau}{A_i\text{Var}_i[D]}}
\]

where \( i \in \mathcal{T}, U \) denotes the set of active unconstrained investors.

The change in social welfare for an investor \( i \), from a planner’s perspective, is given by:

\[
\frac{dV_i}{d\tau} = \mathbb{E} \left[ U_i'(W_{2i}) \right] \left[ \left( E_i[D] - A_i\text{Cov}_i[E_{2i}, D] - P_1 - A_iX_{1i}\text{Var}_i[D] \right) \frac{dX_{1i}}{d\tau} - \Delta X_{1i} \frac{dP_1}{d\tau} \right]
\]

I use \( i = U \) to denote unconstrained investors and \( i = C \) for constrained investors. Substituting the optimality condition:

\[
\frac{dV_i}{d\tau} \bigg|_{i = U} = \mathbb{E} \left[ U_i'(W_{2i}) \right] \left[ \left( E_i[D] - E_i[D] + \text{sgn}(\Delta X_{1i}) P_1 \tau \right) \frac{dX_{1i}}{d\tau} - \Delta X_{1i} \frac{dP_1}{d\tau} \right]
\]

\[
\frac{dV_i}{d\tau} \bigg|_{i = C} = \mathbb{E} \left[ U_i'(W_{2i}) \right] \left[ \left( E_i[D] - A_i\text{Cov}_i[E_{2i}, D] - P_1 - A_i g_i(P_1) \text{Var}_i[D] \right) \frac{dX_{1i}}{d\tau} - \Delta X_{1i} \frac{dP_1}{d\tau} \right]
\]

We can show that the term multiplying \( \frac{dX_{1i}}{d\tau} \) for constrained investors is positive when:

\[
g(P_1) < \frac{E_i[D] - A_i\text{Cov}_i[E_{2i}, D] - P_1}{A_i\text{Var}_i[D]}
\]

57
and negative otherwise. The welfare change for constrained investors can be rewritten, by substituting the (shadow) first order condition as:

\[
\frac{dV_i}{d\tau} = \mathbb{E} [U_i' (W_{2i})] \left[ \mathbb{E} [D] - \mathbb{E}_i [D] + A_i \text{Var} [D] (\hat{X}_{1i} - g (P_1)) + P_1 \text{sgn} (\Delta X_{1i}) \tau \right] \frac{dX_{1i}}{d\tau} - \Delta X_{1i} \frac{dP_1}{d\tau}
\]

Under assumption [NR], we can write social welfare as:

\[
\frac{dV}{d\tau} = \int [\mathbb{E}_i [D] + \text{sgn} (\Delta X_{1i}) P_1 \tau] \frac{dX_{1i}}{d\tau} dF (i) + \int_{i \in C} A_i \text{Var} [D] (\hat{X}_{1i} - g (P_1)) \frac{dX_{1i}}{d\tau} dF (i)
\]

Where I use \( C \) to denote the set of constrained investors and \( \hat{X}_{1i} \) is given by the individual first order condition in equation (4). After substituting for constrained investors \( \frac{dX_{1i}}{d\tau} = g' (P_1) \frac{dP_1}{d\tau} \), we can write the optimal tax as:

\[
\int [\mathbb{E}_i [D] + \text{sgn} (\Delta X_{1i}) P_1 \tau] \frac{dX_{1i}}{d\tau} dF (i) + \int_{i \in C} A_i \text{Var} [D] (\hat{X}_{1i} - g (P_1)) g' (P_1) \frac{dP_1}{d\tau} dF (i) = 0
\]

\[
\tau^* = \frac{\int \mathbb{E}_i [D] \frac{dX_{1i}}{d\tau} dF (i) - \int_{i \in C} A_i \text{Var} [D] (\hat{X}_{1i} - g (P_1)) g' (P_1) \frac{dP_1}{d\tau} dF (i)}{P_1 \int \text{sgn} (\Delta X_{1i}) \frac{dX_{1i}}{d\tau} dF (i)}
\]