Random Risk Aversion and Liquidity: a Model of Asset Pricing and Trade Volumes*

Fernando Alvarez† and Andy Atkeson‡

Abstract

We develop a new general equilibrium model of asset pricing and asset trading volume in which agents' motivations to trade arise due to uninsurable idiosyncratic shocks to agents' risk tolerance. In response to these shocks, agents trade to rebalance their portfolios between risky and riskless assets. We study a positive and a normative questions. The first question is the conditions under which trade volume becomes a pricing factor. We find that economies where its market wide risk aversion is positively correlated with trade volume have a higher risk premium for aggregate risk. Likewise, we find that asset whose cash flows are concentrated on times/states of high trading volume have higher prices, and lower risk premium. The mathematics of this pricing is the same as the one in Mankiw (1986) and Constantinides and Duffie (1996) regarding the role of idiosyncratic income shocks in pricing assets. The primary difference between our model and theirs, however, is that the shocks to risk tolerance in our model lead to a positive volume of trade while there is no trade in these other papers. The second question is the impact of taxes on trading on welfare. We show that such taxes have a first-order negative impact on ex-ante welfare, i.e. a small subsidy on trade has s first order improvement in ex-ante welfare. We compare this small linear tax/subsidy with the optimal non-linear tax/subsidy when we treat idiosyncratic risk tolerance to be private information. We find that the second best is closer to a subsidy to trade rather than a tax.

Preliminary and Incomplete

*We are grateful for comments from Martin Eichenbaum, Sergio Rebelo, Ivan Werning, Harald Uhlig and Lars Hansen. We thank for financial support the Goldman Sachs Global Markets Institute.

†Department of Economics, University of Chicago, NBER.

‡Department of Economics, University of California Los Angeles, NBER, and Federal Reserve Bank of Minneapolis.

1 Introduction

In this paper, we develop a new general equilibrium model of asset pricing and asset trading volume in which investors' motivations to trade arise due to uninsurable idiosyncratic shocks to investors' risk tolerance. In response to these shocks, investors trade to rebalance their portfolios between risky and riskless assets. We use this model to study a positive and a normative question: to what extent is trading volumes a factor that helps price assets? and what are the welfare implications of Tobin taxes and subsidies to asset trading? To answer the positive question we consider economies with "prudent" investors, as defined below. First, with positive trading volume, interest rates are lower than in otherwise identical representative agent economies with no rebalancing trade. Second, if the risk aversion of the representative investor is positively correlated with trade volume, the risk premium for aggregate risk is higher than in otherwise identical representative agent economies with no rebalancing trade. Third, asset whose cash flows are concentrated on times/states when trading volume is high sell at a higher price, i.e. they have lower expected returns. On the normative question posed above we show that taxes on rebalancing trade have a first-order negative impact on ex-ante welfare, i.e. a small subsidy on trade has s first order improvement in ex-ante welfare. We compare this small linear tax/subsidy with the optimal non-linear tax/subsidy when we treat idiosyncratic risk tolerance to be private information. We find that the solution of that mechanism design problem is closer to a subsidy to rebalancing trade rather than a tax to it.

The are two main mechanisms in play to obtain the answers the questions posed above. The first is that since in equilibrium an asset with non-diversifiable payoffs have a risk premium, an investor who suffer an idiosyncratic shock and becomes more risk averse effectively becomes poorer. That is, that investor obtain a lower welfare than an investor with average risk tolerance, since she wants to buy more of the risk-less asset by selling the risky asset, and thus obtaining a loss. We refer to this as a "rebalancing risk". The second is that we treat the idiosyncratic shocks to risk tolerance as uninsurable, i.e. we assume there are no market to hedge these idiosyncratic preference shocks. This means that in an equilibrium the risk tolerance shock (or rebalancing risk) act in the same way as uninsurable idiosyncratic shocks do in Mankiw (1986) and Constantinides and Duffie (1996). In their setup, as well as in ours, investors precautionary savings is the key property which induces risk premium due to uninsurable countercyclical risks –income in their case, rebalancing in ours. Moreover, the same mechanism explain why a Tobin tax exacerbates the assumed market incompleteness with respect to preference shocks, and thus such tax ends up having a first order ex-ante welfare loss. Finally, a mechanism design approach endogenizes this incompleteness under the assumption that such shocks are private information, and shows that the second best allocation is closer to have a subsidy to rebalancing rather than a tax.

Information and Market Structure. We consider a three period endowment economy, where, to simplify, consumption takes place only in the first and last periods. In period $t = 0$ all agents are identical, in period $t = 1$ all investors receive common signals about period $t = 2$ output, and each investors preferences for consumption at $t = 2$ are realized. These preference shocks can include both an idiosyncratic and an aggregate shift in risk tolerance. In period $t = 2$ aggregate output is realized. In our benchmark case we assume that time $t = 1$ idiosyncratic shocks to risk tolerance are uninsurable, but otherwise investors can trade in all assets at $t = 0$ and $t = 1$. In particular investors can trade on all asset that are indexed to the realization of aggregate output at $t = 2$. A Lucas's tree giving time $t = 2$ aggregate output is the only asset in positive net supply. Since investors are identical at $t = 0$ in equilibrium we can price asset but there will be no trade. All the trade takes place at period $t = 1$, and given our specification of preferences, will be due to the idiosyncratic shocks on risk preferences. It will turn out that in equilibrium it will suffice to trade bonds with uncontingent time $t = 2$ payments against the Lucas's tree. Mostly as an intermediate step, we are interested in the price of assets at time $t = 1$, as a function of the characteristic of their $t = 2$ cash flows and of the realization of the information at $t = 1$. Note that among the $t = 1$ information is the cross sectional distribution of risk tolerance, which will determine trade volume. Indeed trade volume equals a measure of dispersion on risk tolerance. We can price all asset at $t = 0$ even if there is no trade. In particular, we price assets whose cash-flows are more or less concentrated to states in $t = 1$ at which there is high or low trade volume, but that are otherwise identical. We also compare the $t = 0$ price of a Lucas' tree among economies which differ in the correlation between time $t = 1$ between dispersion of risk tolerance (i.e. on trade volume) and time $t = 2$ output. Before getting into the results we explain the nature of preferences.

Preferences. A key part of our model for its positive and normative implications is the specification of preferences. In particular the way that investors view at time $t = 0$ the prospect of a time $t = 1$ random shock to their risk tolerance. Yet to describe preferences it is easier to work backwards in time. We assume that after the realization of the $t = 1$ shocks to risk tolerance, investors use expected utility to value the time $t = 2$ random consumption. Specifically we assume that after the shock each investor has a utility function of the equicautious HARA family, which we index as $U_\tau(\cdot)$. At it is well known this family includes utility functions with constant relative risk aversion, constant absolute risk aversion, and quadratic utility, where the origin can be displaced from zero. Formally this is the class of utilities where risk tolerance is linear on consumption.[1] The intercept of the linear risk tolerance function, which we denote by τ is allows to be investor specific, and it is the preference shock that we consider. Note that in our specification risk tolerance can be increasing or decreasing in wealth, which is largely immaterial for our paper. For each realization of risk tolerance τ parameter and for each time $t = 2$ consumption bundle we define a time $t = 1$ certainty equivalent. Then the time $t = 0$ preferences are given by an additively separable utility V over time $t = 0$ consumption plus the discounted value of the expected utility over the time $t = 1$ *certainty equivalent* of continuation consumption, also computed with the utility function V. This gives a non-expected utility as of time $t = 0$ as in Kreps and Porteus (1978) or Selden (1978) type. For the particular case where the distribution of risk tolerance τ is degenerate, ex-ante preferences are exactly as in Selden (1978). In the general case, time $t = 0$ investors evaluate the prospects of preference shocks *only* by considering their effect on their implied certainly equivalent consumption. In particular we assume that investors are risk averse with respect to randomness on certainty equivalent consumption regardless of whether the variations comes from randomness on time $t = 2$ consumption or on time $t = 1$ preference shocks. This formulation, as opposed to simply add a shifter to preferences, isolates the effect of randomness of risk tolerance without having extra effects due to the particular cardinal representation of utility. Finally, this specification has been used in social choice theory when consider foundations for ex-ante Rawlsian's preferences "before the veil of ignorance" to take into account the effect of different realized risk tolerances –see Grant et al. (2010) or Mongin and Pivato (2015).

The assumption that the realized $U_\tau(\cdot)$ time $t = 1$ utility functions are of the HARA family has three important implications, which helps understand the logic of our asset pricing results. First, the two-mutual fund separation theorem hold form $t = 1$ to $t = 2$, and hence it suffices for investors to trade in a Lucas' tree and an uncontingent bond. Second, for the purpose of pricing at $t = 1$ securities that pay at $t = 2$ there is a representative investors whose preferences depend exclusively on the *average* of the risk tolerance parameter τ across investors. Additionally, since there is Gorman aggregation, trade volume depends exclusively on the *dispersion* of the risk

[1] Recall that risk tolerance is defined as the reciprocal of risk aversion.

tolerance parameter τ across investors.[2] Third, taking these two facts implies that time $t = 1$ asset pricing are independent of time $t = 1$ trade volume. Thus the only effect of trade volume on asset pricing comes from the marginal valuation that investor attach, from the point of view of time $t = 0$, to time $t = 1$ prices that occur at realizations with different trade volume. But since trade volume maps one-to-one with dispersion of the risk tolerance parameter τ's, then the impact in the time $t = 0$ marginal valuation depends on the correlation between time $t = 1$ trade volume and time $t = 1$ prices. In other words, the connection between trade volume and e-ante asset prices comes from investors valuation in the presence of the risk of rebalancing.

Rebalancing risk as income risk. The key property of our model is effect that the realization of different risk tolerance shock have in an investor both ex-post and ex-ante. The Arrow-Pratt theorem states that an investor with a risk tolerance lower than other, has also a lower certainty equivalent consumption. In our model all investor are ex-ante identical at $t = 0$, and thus have the same asset position right before the time $t = 1$ preference shock hits them, and hence face the same budget set at $t = 1$. Thus an investor with turns out to be risk averse ends up with lower certainty equivalent than a risk tolerant investor. Thus, in an environment with aggregate risk, the idiosyncratic risk tolerance shock is akin to a negative income shock in the sense that such a shock makes it more costly for that investor to attain any given level of certainty equivalent consumption. Hence certainty equivalence and rebalancing trade map one-to-one, and thus high trade volume at $t = 1$ is equivalent to high dispersion of time $t = 1$ certainty equivalent. Finally, since the assumed time $t = 0$ preferences are in terms of expected utility over *certainty equivalent* consumption, the equivalence of risk tolerance shock with an income shock is exact. Thus, as in Mankiw (1986) and Constantinides and Duffie (1996), the effect on time $t = 0$ marginal valuation of the idiosyncratic variation on *certainty* equivalent consumption depends on whether preferences feature precautionary savings. When $V''' > 0$, so investors are prudent, then states that correspond to high trade volume (i.e. high dispersion of certainty equivalence) are state with high marginal valuation. This effect explains the three positive asset pricing implications describe above.

Tobin Taxes. We also use our model to evaluate the impact on ex-ante welfare of a tax on asset transactions. Contrary to the standard public finance result that in an undistorted equilibrium a tax (or subsidy) has zero first order effect on welfare, in our case a transaction tax has a first order negative welfare effect. This is because it turns out that a transaction

[2] Recall that investors are identical at time $t = 0$ and that τ shocks are uninsurable, so investors start their time $t = 1$ with identical portfolios.

tax levied in the equilibrium with uninsurable risk tolerance shocks exacerbates the ex-ante distortion on certainty equivalent consumption. In other words, a small subsidy to trade has a first order ex-ante welfare improvement.

As background, note that standard welfare analyses taxes imply that, starting from the undistorted equilibrium, the introduction of such taxes have no first order impact on welfare because the envelope theorem ensures that the marginal impact on welfare from the distortion to trade is zero, and the standard welfare criterion is not impacted by the redistribution of resources that results from the differential incidence of the tax. In contrast to this standard result, we find that a transactions tax does have a first order negative impact on ex-ante welfare. In our setting, the envelope theorem still holds at time $t = 1$, so the first order welfare impact of the distortion to the allocation from $t = 1$ onwards of a transactions tax is still zero. However, in our environment, in contrast to the standard analysis, the redistribution of resources that results from the differential incidence of the tax has a first order negative impact on ex-ante welfare. In particular, it turns out that investors who experience negative shocks to their risk tolerance and wish to sell risky assets end up with lower certainty equivalence from the imposition of the tax, and those that experience positive shocks to their risk tolerance end up with higher certainty equivalence. Since in equilibrium these group of investors have discretely different marginal valuations, then ex-ante welfare is reduced by the imposition of a small tax. Put it differently, the incidence of the tax exacerbates the risk that investors already bear due to the uninsurable nature of risk tolerance shocks.

While this is the central idea of the effect of a tax, there are several nuanced effects that need to be discussed. Among them is why those that turn our less risk tolerant end up with lower certainty equivalence, i.e. how is the effect on equilibrium prices and why it does not depend on anything else. Also, int the general case of many values of the risk tolerance, there are effects of redistributions among those that become more risk tolerant and also among those that become less. These effects are discussed and analyzed in different results.

Finally we complement our analysis of a simple linear tax of trade rebated lump sum, with the analysis of the optimal non-linear tax-subsidy. We use a standard mechanism design problem, assuming that the realization of individual risk tolerance is private information for each investor. We think of this assumption as the natural explanation of why we assume that these risks are uninsurable. In this case we take that trade is observable and designed a non-linear menu of exposures to aggregate risk and uncontingent payoffs that maximize ex-ante utility. Investors chose where to locate on the menu of risky vs riskless assets, so that the menu implies a set of non-linear prices, or equivalently linear prices plus non-linear taxes. We use the optimal non-linear scheme judge the sense in which a subsidy to trade is a general feature of the optimal policy. While the optimal second best allocation is not exactly a subsidy to trade, it is closer to it than to a tax in trade. For instance, it does imply that investors end up with a more disperse allocation of risky exposures than in the equilibrium without taxes, i.e. it works as a *subsidy* to trade. Also the implied non-linear prices resemble more those of a Tobin subsidy than those of a Tobin tax.

1.1 Relation to the literature

There is a large theoretical and empirical literature on the relationship between trading volume and asset prices.

One branch of the literature on trading volume and asset pricing assumes that agents are different ex-ante in their trading behavior. Some agents are "noise traders" who buy and sell at intermediate dates with inelastic asset demands for exogenously specified reasons while other agents have elastic asset demands and are the marginal investors pricing assets in equilibrium. (See for example Shleifer and Summers (1990) and Shleifer and Vishny (1997)). As emphasized in the survey of this literature by Dow and Gorton (2006), in many models, noise traders systematically lose money because they tend to sell securities at low prices. One might interpret our model in which agents are identical ex-ante and then subject to idiosyncratic preference shocks as pricing the risk that one finds oneself wanting to sell risky securities at a time at which the price for these securities is low.

Rebalancing Trade. There is a large literature on rebalancing trade. For instance, Lo and Wang (2000) and Lo and Wang (2006) use a factor analysis on weekly trading volume of equities. They show that the de-trended cross sectional trade volume data has an important first component, which can be interpreted as rebalancing trade, accounting about 2/3 of the cross sectional variations. Yet, as they emphasize, this is far from being consistent with the two mutual fund separation theorem, and instead favor at least a second factor. There are also many recent studies of individual household portfolios, which take advantage of large administrative data sets coming from tax authorities, such as Calvet, Campbell, and Sodini (2009).

Shocks to hedging needs. Vayanos and Wang (2012) and Vayanos and Wang (2013) survey theoretical and empirical work on asset pricing and trading volume using a unifying three period model similar in structure to ours. In their model, agents are ex-ante identical in period $t = 0$ and they consume the payout from a risky asset in period $t = 2$. In period $t = 1$, agents receive non-traded endowments whose payoffs at $t = 2$ are heterogeneous in their correlation with the payoff from the risky asset. This heterogeneity motivates trade in the risky asset at $t = 1$ due to investors' heterogeneous desires to hedge the risk of their non-traded endowments. Vayanos and Wang focus their analysis on the impact of various frictions (participation costs, transactions costs, asymmetric information, imperfect competition, funding constraints, and search) on the model's implications for three empirical measures of the relationship between trading volume and asset pricing. The first of these measures is termed *lambda* and is the regression coefficient of the return on the risky asset between periods $t = 0$ and $t = 1$ on liquidity demanders signed volume. The second of these measures is termed *price reversal*, defined as minus the autocorrelation of the risk asset return between periods $t = 1$ and $t = 1$ and between $t = 1$ and $t = 2$. The third measure is the ex-ante expected returns on the risky asset between periods $t = 0$ and $t = 1$. Our focus differs from theirs in that we study the impact of the shocks that drive demand for trade at $t = 1$ on asset prices in a model without frictions and then consider the welfare implications of adding a trading friction in the form of a transactions tax. Yet we have shown that our setup is amenable to study frictions on trading, as we have done with our study of transaction taxes. We think our model can be easily combined with a framework in which trade is motivated by non-traded endowment shocks as well as shocks to risk tolerance. The advantage of such a model is that, by having both "demand" and "supply" shocks, it should yield better ways to identification of shocks.

Duffie, Gârleanu, and Pedersen (2005) study the relationship between trading volume and asset prices in a search model in which trade is motivated by heterogeneous shocks to agents' marginal utility of holding an asset. As they discuss, these preference shocks can be motivated in terms of random hedging needs -see also Uslu (2015).

Random Risk Tolerance. There is a small theoretical asset pricing literature that use random changes in risk tolerance. An early example, particularly related because it addresses properties of the volume of transactions, is Campbell, Grossman, and Wang (1993). The aims of that paper is to investigate the temporal patterns returns and volume, and use a stylized model with random risk aversion for investors, since they also consider different type of traders, those that trade for non-informational reasons, and those that trade on information. Additionally the paper has a substantive empirical analysis of the implications of the model. In Section 4 we define and distinguish between coincidental and direct connection between asset pricing and volume, to distinguish the mechanism at play in the model that paper vs the one in our paper.

On the pure portfolio side Steffensen (2011) analyzes the implications randomness on risk tolerance using expected utility. Gordon and St-Amour (2004) use a time separable utility with state dependent CRRA parameter to jointly fit consumption and asset pricing moments. Kozak (2015) uses time varying aversion in a representative agent model with non-separable preferences to model variations on the market price of risk. Kozak (2015) studies nature of time varying returns for bonds and stock returns. Kim (2014) uses Epstein-Zin preferences with a representative agent with time varying risk aversion. Kim (2014) uses a non-parametric estimates of risk aversion and finds strong evidence for its variability. Drechsler (2013), Bhandari, Borovička, and Ho (2016) use models where agents have time varying concerns for mode misspecification, which can be also interpreted as random risk aversion. Drechsler (2013) studies time varying returns, especially of volatility related derivatives. Bhandari, Borovička, and Ho (2016) combines time variable concern for misspecification with time series and survey data and analyze inflation expectations. Lenel (2016) also uses an Epstein-Zin model with random risk aversion, but he analyzes an equilibrium with incomplete markets. His interest is on the joint explanation of the holding of bonds and risky asset of different (ex-post) agent types and their returns.

The external habit formation model has, when one concentrates purely on the resulting stochastic discount factor, a form of random risk aversion that is nested by our equicautious HARA utility specification if agents have common CRRA preferences over consumption less the external habit parameter, which will correspond to τ (as in Campbell and Cochrane (1999). Bekaert, Engstrom, and Grenadier (2010) develop and estimate a version of Campbell and Cochrane (1999) where the ratio of consumption to habit also has independent random variation. They estimate a (linearized) version of the model and find a substantial role for independent shocks to the consumption/habit ratio, which have the interpretation of shocks to risk aversion. Guo, Wang, and Yang (2013) and Cho (2014) further investigates estimates of variations of this model. Our recursive definition of preferences isolates the shocks to risk tolerance, leaving only intertemporal preferences over the allocation of certainty equivalent consumption unchanged.[3]

The idea that idiosyncratic preference shocks impact investors' precautionary demand for

[3] The alert reader of Campbell and Cochrane (1999) will recognize the non-linear adjustment on that model to zero out the precautionary saving effect and obtain constant interest rates.

an asset (in this case money) is central to Lucas (1980). Also, the idea that shocks to the demand side for risky assets are important is emphasized by Albuquerque et al. (2016). The model in that paper, as well as several other related models, incorporate riskiness of preference shocks so that the model can account for the weak correlation with traditional supply side factors emphasized in the literature. We concentrate on the relationship between aggregate and idiosyncratic preference shocks so we can examine implied relationships between trade volume and asset pricing.

2 The Model

In this section, we describe our model environment and our specification of agents' preferences with random shocks to each agent's risk tolerance. We define optimal and equilibrium allocations and develop our asset pricing formulas. In the next section, we solve the model for a specific class of preferences and characterize the model's implications for asset prices and trading volume due to portfolio rebalancing.

Consider a three period economy with $t = 0, 1, 2$ and a continuum of measure one of agents. Agents are all identical at time $t = 0$. Agents consume in periods $t = 0$ and $t = 2$. Shocks to agents' risk tolerance are realized at $t = 1$.

There is an aggregate endowment of consumption available at $t = 0$ of $\bar{C}_0$. Agents face uncertainty over the aggregate endowment of consumption available at time $t = 2$, denoted by $y \in Y$. To simplify notation, we assume that Y is a finite set.

Agents face idiosyncratic and aggregate shocks to their preferences that are realized at $t = 1$. Heterogeneity in agents' preferences at time $t = 1$ motivates trade at $t = 1$ in claims to the aggregate endowment at $t = 2$. Preference types at $t = 1$ are indexed by τ with support $\tau \in \{\tau_1, \tau_2, \ldots, \tau_I\}$.

Uncertainty is described as follows. At time $t = 1$, an aggregate state $z \in Z$ is realized. Again, to simplify notation, we assume that Z is finite set and probabilities of z being realized at $t = 1$ are denoted by $\pi(z)$. The distribution of agents across types τ depends on the realized value of z, with $\mu(\tau|z)$ denoting the fraction of agents with realized type τ at $t = 1$ in state z. In describing agents' preferences below, we assume that the probability that an individual has realized type τ at $t = 1$ if state z is realized is also given by $\mu(\tau|z)$.

In addition, the conditional distribution of the aggregate endowment at $t = 2$ may also depends on z, with $\rho(y|z)$ denoting the probability of y being realized at $t = 2$ conditional on z being realized at $t = 1$. We denote the conditional mean and variance of the aggregate endowment at $t = 2$ by $\bar{y}(z)$ and $\sigma^2(z)$ respectively.

Allocations: An allocation in this environment is denoted by $\vec{c}(y; z) = \{C_0, c(\tau, y; z)\}$ where C_0 is the consumption of agents at $t = 0$ and $c(\tau, y; z)$ is the consumption at $t = 2$ of an agent whose realized type is τ if aggregate states z and y are realized.

Feasibility requires $C_0 = \bar{C}_0$ at $t = 0$ and, at $t = 2$

$$\sum_{\tau} \mu(\tau|z)\, c(\tau, y; z) = y \text{ for all } y \in Y \text{ and } z \in Z \tag{1}$$

2.1 Preferences

We describe agents' preferences at $t = 0$ (before z and their individual types are realized) over allocations $\vec{c}(y; z)$ by the utility function

$$V(C_0) + \beta \sum_{z} \left[\sum_{\tau} \mu(\tau|z) V \left(U_{\tau}^{-1} \left(\sum_{y} [U_{\tau}(c(\tau, y; z))\rho(y|z)] \right) \right) \right] \pi(z) \tag{2}$$

where V is some concave utility function. We refer to U_{τ} as agents' type-dependent sub-utility function.

Certainty Equivalent Consumption: It is useful to consider this specification of preferences in two stages as follows. In the first stage, consider the allocation of certainty equivalent consumption at $t = 1$ over states of nature z. For any allocation $\vec{c}(y; z)$, an agent whose realized type is τ at $t = 1$ has certainty equivalent consumption implied by the allocation to his or her type and the remaining risk over y in state z given by

$$C_1(\tau; z) \equiv U_{\tau}^{-1} \left(\sum_{y} U_{\tau}(c(\tau, y; z))\rho(y|z) \right) \tag{3}$$

Given this definition, in the second stage, we can write agents' preferences as of time $t = 0$ as expected utility over certainty equivalent consumption

$$V(C_0) + \beta \sum_{z} \left[\sum_{\tau} \mu(\tau|z) V\left(C_1(\tau; z)\right) \right] \pi(z) \tag{4}$$

Convexity of Upper Contour Sets: To ensure that agents' indifference curves define convex upper contour sets, we must restrict the class of subutility functions $U_\tau(c)$ that we consider to those for which, given z, certainty equivalence at time $t = 1$ as defined in equation (3) is a concave function of the underlying allocation $c(\tau, y; z)$ for each given τ and z at $t = 2$. Following Theorem 1 in Ben-Tal and Teboulle (1986), in the Appendix, we show that this is the case if and only if agents' risk tolerances, defined as $\mathcal{R}_\tau(c) \equiv -\frac{U'_\tau(c)}{U''_\tau(c)}$, are a concave function of consumption $c(\tau, y; z)$ for all types τ and realized z. One can verify by direct calculation that certainty equivalence is a concave function of the underlying allocations for subutility of the CRRA form in which agents differ in their coefficient of relative risk aversion. As we discuss below, this is also the case for the case of equicautious HARA utility functions that we consider as our leading example throughout the paper.

2.2 Equicautious HARA Utility

The specification of preferences we use in our leading example has subutility U_τ of the equicautious HARA utility class defined as

$$U_\tau(c) = \left(\frac{\gamma}{1-\gamma}\right)\left(\frac{c}{\gamma} + \tau\right)^{1-\gamma} \gamma \neq 1 \text{ for } \left\{c : \tau + \frac{c}{\gamma} > 0\right\} \tag{5}$$

$$U_\tau(c) = \log(c + \tau) \text{ for } \{c : \tau + c > 0\} \text{ for } \gamma = 1 \text{ for } \{c : \tau + c > 0\}\,, \text{ and} \tag{6}$$

$$U_\tau(c) = -\tau \, \exp\left(-c/\tau\right) \text{ as } \gamma \to \infty, \text{ for all } c\,. \tag{7}$$

This utility function is increasing and concave for any values of τ and γ as long as consumption belongs to the sets described above for each of the cases. To see this, we compute the first and second derivative as well as the risk tolerance function:

$$U'_\tau(c) = \left(\frac{c}{\gamma} + \tau\right)^{-\gamma} > 0\,,\; U''_\tau(c) = -\left(\frac{c}{\gamma} + \tau\right)^{-\gamma-1} < 0 \text{ and} \tag{8}$$

$$\mathcal{R}_\tau(c) \equiv -\frac{U'_\tau(c)}{U''_\tau(c)} = \frac{c}{\gamma} + \tau \tag{9}$$

Note that notation above assumes that γ is common across agents. Note also that $\gamma > 0$ gives decreasing absolute risk aversion and $\gamma < 0$ gives increasing absolute risk aversion. The sign of γ will turn out to be immaterial for the qualitative behavior of the model.

Type τ and the cost of certainty equivalent consumption: When agents have subutility U_τ of the equicautious HARA utility class, the interpretation of preference type τ is is that if

Figure 1: Event tree for 3-period model

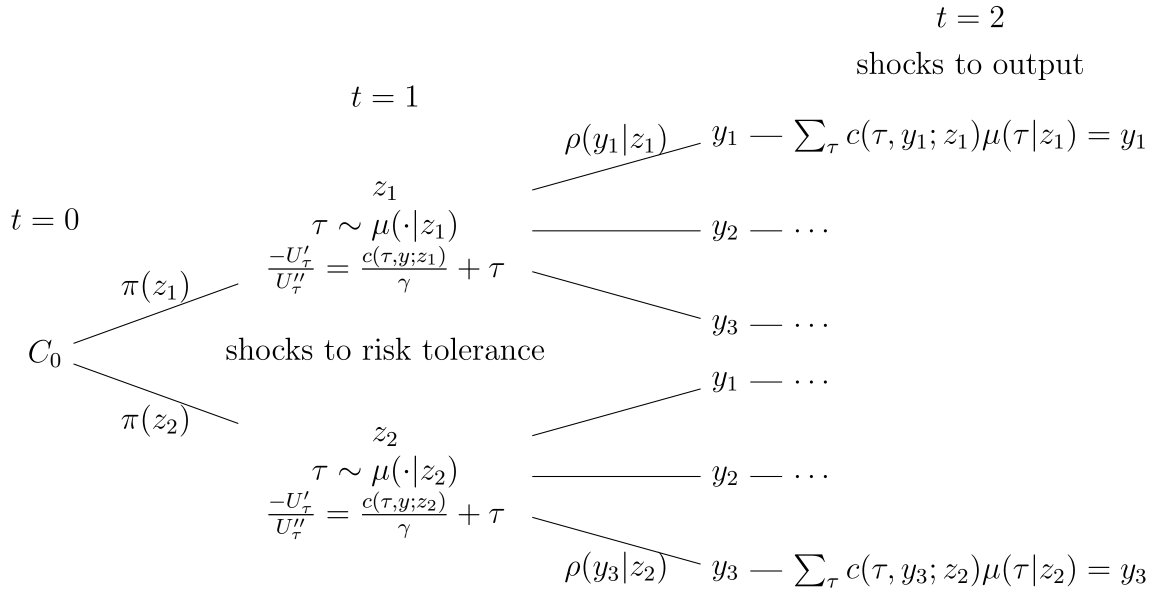


Figure for the case of two values for $z \in \{z_1, z_2\}$ and three values for $y \in \{y_1, y_2, y_3\}$.

$\tau > \tau'$, then at any level of consumption, an agent of type τ has higher risk tolerance than an agent of type τ'. Hence, the heterogeneity we consider with these preferences is purely in terms of the level of risk tolerance across agents. The Arrow-Pratt theorem then immediately implies that if, given z at $t = 1$, agents of type τ and τ' receive the same allocation at $t = 2$, i.e. if given z, $c(\tau, y; z) = c(\tau', y; z)$ for all y, then agents of type τ have higher certainty equivalent consumption at $t = 1$, i.e. $C_1(\tau; z) \geq C_1(\tau'; z)$. In this sense, for an individual agent, having type τ' realized at $t = 1$ is a negative shock relative to having type τ realized at $t = 1$ in that with preferences of type τ' it requires more resources for the agent to attain the same level of certainty equivalent consumption as an agent with preferences of type τ. In the context of the interpretation of the environment as a production economy, we say that it is more costly to produce certainty equivalent consumption for an agent of type τ' than it is to produce that certainty equivalent consumption for an agent of type τ.

We summarize the timing of the realization of uncertainty agents face in our model as in Figure 1.

2.3 Interpretation of Model as a Production Economy:

Our model economy is an endowment economy. To help in the interpretation of the asset pricing formulas below and in solving the model, it is useful to reinterpret our economy as a production economy in which agents face risk in the cost of producing certainty equivalent consumption specific to their type as indexed by τ. Specifically, consider an economy in which there is a continuum of agents who consume in periods $t = 0$ and $t = 1$ who face uncertainty at $t = 1$ over their "location", as indexed by τ. Let these agents have consumption denoted by $(C_0, C_1(\tau; z))$ and preferences over this consumption given by (4). Assume that, given a realization of z at $t = 1$, consumption $C_1(\tau; z)$ in location τ is produced using traded inputs $c(\tau, y; z)$ according to (3). We say that, given a realization of z, an allocation of certainty equivalent consumption across locations τ at $t = 1$, $\{C_1(\tau; z)\}$, is feasible if there exists an allocation of inputs $\vec{c}(y; z)$ that is feasible as in (1) and that delivers that vector of certainty equivalent consumption via (3).

Let $\mathbf{C}_1(z)$ denote the set of feasible allocations of certainty equivalent consumption at $t = 1$ given a realization of z. Note that this set is convex as long as certainty equivalent consumption at time $t = 1$ is a concave function of the underlying allocation at $t = 2$ as we have assumed. The set $\mathbf{C}_1(z)$ has the interpretation of a production possibility set whose shape is impacted by the aggregate shock z which determines the tolerance of agents for risk (through $\mu(\tau|z)$ and the quantity of risk to be borne (through $\rho(y|z)$). As we discuss below, the marginal cost of producing certainty equivalent consumption computed from this production possibility set plays an important role in asset pricing.

We next consider optimal allocations and the corresponding decentralization of those allocations as equilibria with complete asset markets.

2.4 Optimal Allocations

Consider a social planning problem of choosing an allocation $\vec{c}(y; z)$ to maximize welfare (2) subject to the feasibility constraints (1). We refer to the solution to this problem as the *ex-ante efficient or socially optimal allocation*. It will be useful to consider the solution of the social planning problem in two stages.

The first stage is to compute the set of feasible allocations of certainty equivalent consumption at $t = 1$ given z, denoted by $\mathbf{C}_1(z)$, and then solve the planning problem of choosing a feasible allocation of certainty equivalent consumption $\{C_0, C_1(\tau; z)\}$ to maximize (4) subject to those feasibility constraints. To characterize the sets $\mathbf{C}_1(z)$, we also consider efficient allocations as of $t = 1$ given z.

We say that a feasible allocation $\vec{c}(y; z)$ is *conditionally efficient* if, given a realization of z at $t = 1$, it solves the problem of maximizing the objective

$$\sum_{\tau} \lambda_\tau C_1(\tau; z)\mu(\tau|z) \tag{10}$$

given constraints (3) among feasible allocations given some vector of non-negative Pareto weights λ_τ. The frontier of the set of feasible allocations of certainty equivalent consumption $\mathbf{C}_1(z)$ is found by solving this maximization problem for all possible non-negative vectors of Pareto weights λ_τ.

We say that allocation $(C_0, C_1(\tau; z))$ of certainty equivalent consumption is *ex-ante efficient* if it solves the problem of maximizing (4) subject to feasibility constraints $C_0 \leq \bar{C}_0$ and the allocation of certainty equivalent consumption across types τ at $t = 1$ given z is feasible $(C_1(\tau; z) \in \mathbf{C}_1(z))$ for all z.

Clearly, the socially optimal allocation is also conditionally efficient.

2.5 Equilibrium Allocations with Complete Asset Markets

The Second Fundamental Welfare Theorem applies to this economy under our assumptions on preferences. Thus, corresponding to the socially optimal allocation is a decentralization of that allocation as an equilibrium allocation with complete markets. We consider the following specification of an equilibrium with complete asset markets.

We assume that all agents start at time $t = 0$ endowed with equal shares of the aggregate endowment of $\bar{C}_0$ at $t = 0$ and y at $t = 2$. In a first stage of trading at time $t = 0$, we assume that agents can trade type-contingent bonds whose payoffs are certain claims to consumption at time $t = 2$ conditional on aggregate state z and idiosyncratic type τ being realized at time $t = 1$. Let a single unit of such a contingent bond pay off one unit of consumption at $t = 2$ in all states y given that z and τ are realized at $t = 1$ and let $B(\tau; z)$ denote the quantity of such contingent bonds held by an agent in his or her portfolio. Let $Q(\tau; z)\mu(\tau|z)\pi(z)$ denote the price at $t = 0$ of such a contingent bond relative to consumption at $t = 0$. Each agents' budget constraint at this stage of trading is given by

$$C_0 + \sum_{\tau,z} Q(\tau; z)B(\tau; z)\mu(\tau|z)\pi(z) = \bar{C}_0 \tag{11}$$

The type-contingent bond market clearing conditions are $\sum_\tau \mu(\tau|z)B(\tau;z) = 0$ for all z.

In a second stage of trading at $t = 1$, agents can trade their shares of the aggregate endowment and the payoff from their portfolio of type-contingent bonds for consumption with a complete set of claims to consumption contingent of the realized value of y at $t = 2$. Let the price at $t = 1$ given z for a claim to one unit of consumption at $t = 2$ contingent on y being realized be denoted by $p(y;z)$. Agents' budget sets at $t = 1$ are contingent on the aggregate state z and their realized type τ and are given by

$$\sum_y p(y;z)c(\tau,y;z)\rho(y|z) \leq \sum_y p(y;z)\left[y + B(\tau;z)\right]\rho(y|z) \tag{12}$$

where the term y on the right hand side of the budget constraint refers to the agent's initial endowment of a share of the aggregate endowment at $t = 2$ and $B(\tau;z)$ refers to the agent's type-contingent bond that pays off in period $t = 2$ following the realization of τ and z at $t = 1$.

Complete Markets Equilibrium: An *equilibrium with complete asset markets* in this economy is a collection of asset prices $\{Q^*(\tau;z), p^*(y;z)\}$, a feasible allocation $\vec{c}^*(y;z)$, and type-contingent bondholdings at $t = 0$ $\{B^*(\tau;z)\}$ that satisfy the bond market clearing condition and that together solve the problem of maximizing agents' ex-ante utility (4) subject to the budget constraints (11) and (12).

In what follows, we refer to the ex-ante efficient allocation or, equivalently, the equilibrium allocation with complete markets, as the *optimal allocation*. We reserve the term *equilibrium allocation* for the allocation in the equilibrium with incomplete asset markets which we describe next.

2.6 Equilibrium with incomplete asset markets

We now consider equilibrium in an economy in which agents are not able to trade contingent claims on the realization of their type τ at $t = 1$. Instead, they can only trade claims contingent on aggregate states z and y. We are motivated to consider incomplete asset markets here by the possibility that the idiosyncratic realization of agents' preference types is private information and that opportunities for agents to retrade at $t = 1$ prevents the implementation of incentive compatible insurance contracts on agents' reports of their realized preference type τ.

We again consider equilibrium with two rounds of trading, one at $t = 0$ before agents' types are realized and one at $t = 1$ after the realization of agents' types. We assume that all agents start at time $t = 0$ endowed with equal shares of the aggregate endowment y. In a first stage of trading at time $t = 0$, we assume that agents can trade bonds whose payoffs are certain claims to consumption at time $t = 2$ conditional on aggregate state z being realized at time $t = 1$. Let a single unit of such a bond pay off one unit of consumption at $t = 2$ in all states y given that z is realized at $t = 1$ and let $B(z)$ denote the quantity of such bonds held by an agent in his or her portfolio. Let $Q(z)\pi(z)$ denote the price at $t = 0$ of such a bond. Each agents' budget constraint at this stage of trading is given by

$$C_0 + \sum_z Q(z)B(z)\pi(z) = \bar{C}_0 \tag{13}$$

with the bond market clearing conditions given by $B(z) = 0$ for all z.

In a second stage of trading at $t = 1$, as before, agents can trade their shares of the aggregate endowment and the payoff from their portfolio of bonds in exchange for a complete set of claims to consumption contingent of the realized value of y at $t = 2$. Agents' budget sets at $t = 1$ are contingent on the aggregate state z and are given by

$$\sum_y p(y;z)c(\tau, y;z)\rho(y|z) \leq \sum_y p(y;z)\left[y + B(z)\right]\rho(y|z) \tag{14}$$

We also use this decentralization to define a concept of equilibrium at time $t = 1$ conditional on a realization of z. Here we assume that at time $t = 1$ agents are each endowed with one share of the aggregate endowment y at $t = 2$ and a quantity of bonds $B(\tau; z)$ that are sure claims to consumption at $t = 2$. We require that, given z, the initial endowment of bonds satisfies the bond market clearing condition $\sum_\tau \mu(\tau|z)B(\tau;z) = 0$.

Conditional Equilibrium given z realized at $t = 1$: An *equilibrium conditional on* z and an allocation of bonds $\{B(\tau;z)\}$ is a collection of asset prices $\{p(y;z)\}$ and feasible allocation $\{c(\tau, y; z)\}$ that maximizes agents' certainty equivalent consumption (3) given the allocation of bonds and budget constraints (12) for all agents.

Clearly, from the two Welfare Theorems, every conditional equilibrium allocation is conditionally efficient and every conditionally efficient allocation is a conditional equilibrium allocation for some initial endowment of bonds.

Incomplete Markets Equilibrium: An *equilibrium with incomplete asset markets* in this economy is a collection of asset prices $\{Q^e(z), p^e(y;z)\}$ and a feasible allocation $\vec{c}^e(y;z)$ and bondholdings at $t = 0$ $\{B^e(z)\}$ that satisfy the bond market clearing condition and that together solve the problem of maximizing agents' ex-ante utility (4) subject to the budget constraints (13) and (14).

Note that since all agents are ex-ante identical, at date $t = 0$, they all hold identical bond portfolios $B^e(z) = 0$. This implies that we can solve for the equilibrium asset prices and quantities in two stages starting from $t = 1$ given a realization of z. Specifically, the equilibrium allocation of consumption at $t = 2$ conditional on z being realized at $t = 1$ is the conditional equilibrium allocation of consumption given z at $t = 1$ and initial bond holdings $B^e(z) = 0$ for all τ and z, and the allocation of certainty equivalent consumption at $t = 1$ given z, $\{C_1^e(\tau; z)\}$, is that implied by the conditional equilibrium allocation of consumption at $t = 2$. Likewise, equilibrium asset prices at $t = 1$, $p^e(y; z)$ are the conditional equilibrium asset prices at $t = 1$ given z. We refer to this conditional equilibrium as the *equal wealth conditional equilibrium* because in it all agents have identical portfolios comprised of one share of aggregate y and zero bonds.

2.7 Asset Pricing

We price assets at dates $t = 1$ and $t = 0$.

Risk Free Bond Prices at $t = 1$ In what follows, we choose to normalize asset prices at time $t = 1$ in each state z such that the price of a bond, i.e. a claim to a single unit of consumption at $t = 2$ for every realization of y, is equal to one. That is, in each equilibrium conditional on z, we choose the numeraire

$$\sum_y p(y; z)\rho(y|z)dy = 1, \tag{15}$$

Share Prices at $t = 1$ At $t = 1$, given state z, the price of a share of the aggregate endowment paid at $t = 2$ relative to that of a bond is given by

$$D_1(z) = \sum_y p(y; z)y\rho(y|z). \tag{16}$$

Since the price of a bond at this date and in this state is equal to one, $D_1(z)$ is also the level of this share price at $t = 1$ given state z.

Figure 2: Time line of 3-period model

	time $t = 0$	time $t = 1$	time $t = 2$
aggregate shocks:		$z \sim \pi(\cdot)$	$y \sim \rho(\cdot\|z)$
idiosyncratic shocks:		$U_\tau(\cdot)$ w/risk tolerance shock $\tau \sim \mu(\cdot\|z)$	
	C_0	certainty equivalent(s): $\bar{C}(z), C^e(\tau; z)$	$c(\tau, y; z)$
	price asset $P_0(d)$	rebalance portfolio, price assets $P_1(z; d)$	payoff $d(y; z)$

Asset prices at $t = 0$ We can price arbitrary claims to consumption at $t = 2$ with payoffs $d(y; z)$ contingent the realized aggregate states z and y as follows. Let

$$P_1(z; d) = \sum_y p(y; z) d(y; z) \rho(y|z) \tag{17}$$

denote the price at $t = 1$ of a security with payoffs $d(y; z)$ in period $t = 2$ given that state z is realized. Then the price of this security at $t = 0$ is

$$P_0(d) = \sum_z Q(z) P_1(z; d) \pi(z) \tag{18}$$

where, in the equilibrium with complete asset markets $Q^*(z) \equiv \sum_\tau Q^*(\tau; z)\mu(\tau|z)$, while in the equilibrium with incomplete asset markets $Q^e(z)$ are the equilibrium bond prices at date $t = 0$. Hence, the price at $t = 0$ of a riskless bond, i.e., a claim to a single unit of consumption at $t = 2$ for each possible realization of τ, z, and y, is given by $P_0(1) = \sum_z Q(z)\pi(z)$. We use the inverse of this price to define the risk free interest rate at $t = 0$ between periods $t = 0$ and $t = 1$ as $\bar{R}_0 = 1/P_0(1)$.

To summarize, the timing of trading and the notation for asset prices in our model is illustrated in Figure 2.

2.8 Preference Shocks and Asset Prices:

To gain intuition for how preference shocks impact asset pricing and to solve the model in the next section, it is useful to follow a two-stage procedure in solving for equilibrium exploiting the interpretation of our economy as a production economy.

In the first stage, we take as given the realized value of z at $t = 1$ and the payoffs from agents' bond portfolios (either $B^*(\tau; z)$ in the equilibrium with complete asset markets or $B^e(z)$ in the equilibrium with incomplete asset markets) and solve for the conditional equilibrium prices for contingent claims to consumption $p(y;z)$ and the corresponding conditional equilibrium allocation of consumption $c(\tau, y; z)$. These prices and this allocation satisfy the budget constraints (12) in the case with complete asset markets or (14) in the case with incomplete asset markets and the standard first order conditions

$$\frac{U'_\tau(c(\tau, y_1; z))}{U'_\tau(c(\tau, y_2; z))} = \frac{p(y_1; z)}{p(y_2; z)} \tag{19}$$

characterizing conditional efficiency for all types τ and all y_1, y_2.

Given a solution for contingent equilibrium prices $p(y;z)$, we can define for each type of agent a cost function for attaining (producing) a given level of certainty equivalent consumption at time $t = 1$ given z as

$$H_\tau(C_1; z) = \min_{c(y;z)} \sum_y p(y;z)c(y;z)\rho(y|z) \tag{20}$$

subject to the constraint that $c(y;z)$ delivers certainty equivalent consumption C_1 at $t = 1$ for an agent of type τ.

Using these cost functions, in the second stage, we can then compute the date $t = 0$ bond prices ($Q^*(\tau;z)$ in the equilibrium with complete asset markets and $Q^e(z)$ in the equilibrium with incomplete markets) that decentralize the equilibrium allocation of certainty equivalent consumption as follows.

In the case with complete asset markets, we analyze the problem for the consumer of choosing certainty equivalent consumption and bondholdings to maximize utility (4) subject to budget constraints (11) and (12) now restated as

$$H_\tau(C_1^*(\tau;z); z) = D_1^*(z) + B^*(\tau; z) \tag{21}$$

with $D_1^*(z)$ defined in (16) as the price of a share at $t = 1$ in state z. This problem has first order conditions

$$Q^*(\tau;z) = \beta \frac{V'(C_1^*(\tau;z))}{V'(C_0^*)} \Bigg/ \frac{\partial}{\partial C_1} H_\tau(C_1^*(\tau;z); z) \tag{22}$$

In the case with incomplete asset markets, we analyze the problem for the consumer of choosing certainty equivalent consumption and bondholdings to maximize utility (4) subject to budget constraints (13) and (14) restated as

$$H_\tau(C_1^e(\tau;z); z) = D_1^e(z) + B^e(z) \tag{23}$$

with $D_1^e(z)$ defined in (16) as the price of a share at $t = 1$ in state z. This problem has first order conditions

$$Q^e(z) = \beta \sum_{\tau} \left[\frac{V'(C_1^e(\tau;z))}{V'(C_0^e)} \Big/ \frac{\partial}{\partial C_1} H_\tau(C_1^e(\tau;z);z) \right] \mu(\tau|z) \tag{24}$$

The Marginal Cost of Certainty Equivalent Consumption: Our asset pricing formulas, (22) and (24) depend on the optimal and equilibrium allocations of certainty equivalent consumption and the marginal cost of providing that allocation of certainty equivalent consumption. Analysis of the cost minimization problem (20) yields that in in the socially optimal allocation, this marginal cost is given by

$$\frac{\partial}{\partial C_1} H_\tau(C_1^*(\tau;z);z) = \frac{U_\tau'(C_1^*(\tau;z))}{\sum_y U_\tau'(c^*(\tau,y;z))\rho(y|z)} \tag{25}$$

while in the equilibrium with incomplete markets it is given by

$$\frac{\partial}{\partial C_1} H_\tau(C_1^e(\tau;z);z) = \frac{U_\tau'(C_1^e(\tau;z))}{\sum_y U_\tau'(c^e(\tau,y;z))\rho(y|z)} \tag{26}$$

These expressions for the marginal cost of certainty equivalent consumption are hence a measure of the risk agents face in the conditional equilibrium at $t = 1$ given realized z in terms of the ratio of the marginal utility of certainty equivalent consumption at $t = 1$ relative to the expected marginal utility of consumption realized at $t = 2$.

3 Solving the Model with HARA Subutility

When agents have subutility functions of the equicautious HARA class (5), then our model is particularly tractable and it is possible to derive specific implications of the model for the relationship between asset prices and transactions volumes at $t = 1$.

The tractability of our model follows from four related properties of these preferences that follow from the assumption that all agents have linear risk tolerance with a common slope in consumption (determined by γ). We prove each of these properties in the appendix.

The four properties are as follows:

Gorman Aggregation: The first property is that, given a realization of z at $t = 1$, Gorman Aggregation holds in all conditional equilibria. That is, in all conditional equilibria at $t = 1$, asset prices $p(y;z)$ are independent of the allocation of bonds $B(\tau;z)$ at that date and also independent of moments of the distribution of types $\mu(\tau|z)$ other than the mean of this distribution defined by

$$\bar{\tau}(z) \equiv \sum_{\tau} \tau \mu(\tau|z). \tag{27}$$

This result allows us to solve for equilibrium asset prices at $t = 1$, $p^*(y; z)$ and $p^e(y; z)$ (both equal to $\bar{p}(y; z)$ defined in 28) in the complete and incomplete markets case directly from the parameters of the environment. Specifically, in all conditional equilibria,

$$\frac{U'_\tau(c(\tau, y_1; z))}{U'_\tau(c(\tau, y_2; z))} = \frac{U'_{\bar{\tau}(z)}(y_1)}{U'_{\bar{\tau}(z)}(y_2)} \equiv \frac{\bar{p}(y_1; z)}{\bar{p}(y_2; z)} \tag{28}$$

for all types τ and all y_1, y_2.

Linear Frontier of Feasible Allocations of Certainty Equivalent Consumption: The second property is that given subutility functions of the equicautious HARA class (5), the feasible sets of allocations of certainty equivalent consumption $\mathbf{C}_1(z)$ have a linear frontier. Specifically, all conditionally efficient allocations of consumption imply allocations of certainty equivalent consumption $C_1(\tau; z)$ that satisfy the pseudo-feasibility constraint

$$\sum_{\tau} \mu(\tau|z) C_1(\tau; z) = \bar{C}_1(z) \tag{29}$$

where

$$\bar{C}_1(z) \equiv U^{-1}_{\bar{\tau}(z)} \left(\sum_{y} U_{\bar{\tau}(z)}(y) \rho(y|z) \right) \tag{30}$$

is the certainty equivalent consumption of an agent with the average risk tolerance in the market who consumes the aggregate endowment at $t = 2$.

This characterization of the set of feasible allocations of certainty equivalent consumption $\mathbf{C}_1(z)$ implies that the socially optimal allocation of certainty equivalent consumption $C_1^*(\tau; z)$ solves the problem of maximizing welfare (4) subject to the pseudo-resource constraint (29). If the utility function over certainty equivalent consumption $V(C)$ is strictly concave, then the solution to this social planning problem is to have all agents receive the same certainty equivalent consumption at date $t = 1$, i.e. $C_1^*(\tau; z) = \bar{C}_1(z)$ for all τ.

This characterization of the set of feasible allocations of certainty equivalent consumption $\mathbf{C}_1(z)$ and equilibrium asset prices at $t = 1$ given in (28) also allows us to solve for the allocation of certainty equivalent consumption in the equilibrium with incomplete markets $C_1^e(\tau; z)$ as well, using the budget constraint (14) and imposing the bond market clearing condition $B^e(z) = 0$ for all z. Specifically, we have that this equilibrium allocation of certainty equivalent consumption is given by

$$C_1^e(\tau;z) = \bar{C}_1(z) + \left(\frac{\tau - \bar{\tau}(z)}{\frac{\bar{D}_1(z)}{\gamma} + \bar{\tau}(z)}\right)\left[\bar{C}_1(z) - \bar{D}_1(z)\right] \tag{31}$$

where $\bar{D}_1(z)$ is the price of a share of the aggregate endowment at $t = 1$ in state z defined in (16) using equilibrium asset prices $\bar{p}(y;z)$.

Type-independent marginal cost of certainty equivalent consumption. The third property is that, given subutility functions of the equicautious HARA class (5), for any conditionally efficient allocation of consumption together with the associated certainty equivalent consumptions, the marginal cost of delivering an additional unit of certainty equivalent consumption to any agent of type τ is independent of type and given by

$$\frac{\partial}{\partial C_1} H_\tau(C_1(\tau;z);z) = \frac{U'_{\bar{\tau}(z)}(\bar{C}_1(z))}{\sum_y U'_{\bar{\tau}(z)}(y)\rho(y|z)} \tag{32}$$

With these three results, we have a complete solution of the model for the optimal and equilibrium allocations and associated asset prices. We also wish to characterize the implications of the model for trading volumes. To do so, we use a fourth property of our preferences.

Two-Fund or Mutual Fund Separation Theorem Given subutility functions of the equicautious HARA class (5), a two fund theorem holds in all conditional equilibria. Specifically, all conditionally efficient allocations can be decentralized as conditional equilibria at $t = 1$ in which agents simply trade shares of the aggregate endowment at $t = 2$ and sure claims to consumption at $t = 2$. We use this result to derive our model's implications for trading volumes.

Consider first asset trade at $t = 1$ to decentralize the optimal allocation. Let $\phi^*(\tau;z)$ denote the post-trade quantity of shares of the aggregate endowment held by an agent of type τ at $t = 1$ given realized z. To implement the socially optimal allocation of certainty equivalent consumption, $C_1^*(\tau;z) = \bar{C}_1(z)$ for all τ, the quantity of shares purchased at $t = 1$ by this agent is

$$\phi^*(\tau;z) - 1 = \frac{\tau - \bar{\tau}(z)}{\frac{\bar{C}_1(z)}{\gamma} + \bar{\tau}(z)} \tag{33}$$

Hence, the volume of trade in shares at time $t = 1$ measured in terms of share turnover is given by

$$TV^*(z) = \frac{1}{2}\sum_\tau \frac{|\tau - \bar{\tau}(z)|}{\frac{\bar{C}_1(z)}{\gamma} + \bar{\tau}(z)}\mu(\tau|z) \tag{34}$$

This measure of trade volume is also a measure of the mean absolute deviation of agents' risk tolerances from the risk tolerance of the agent with average risk tolerance evaluated at the certainty equivalent level of consumption. Hence, in our model, there is a direct mapping between the dispersion in preference shocks agents face and observable trade in shares at time $t = 1$.

Consider now the equilibrium with incomplete asset markets. Let $\phi^e(\tau; z)$ denote the post-trade quantity of shares of the aggregate endowment held by an agent of type τ at $t = 1$ given realized z. To implement the equilibrium allocation of certainty equivalent consumption (31), the quantity of shares purchased at $t = 1$ by this agent is

$$\phi^e(\tau; z) - 1 = \frac{\tau - \bar{\tau}(z)}{\frac{\bar{D}_1(z)}{\gamma} + \bar{\tau}(z)} \tag{35}$$

Hence observed trade volumes are given by

$$TV^e(z) = \frac{1}{2} \sum_{\tau} \frac{|\tau - \bar{\tau}(z)|}{\frac{\bar{D}_1(z)}{\gamma} + \bar{\tau}(z)} \mu(\tau|z) \tag{36}$$

Hence, again this measure of trade volume is also a measure of the mean absolute deviation of agents' risk tolerances from the risk tolerance of the agent with average risk tolerance evaluated at consumption equal to the share price. In other words, observed share trade volumes again are a direct measure of the dispersion in agents' risk tolerances.

We summarize our solution of the model with both complete and incomplete asset markets in the following proposition.

Proposition 1. *Let $V(C)$ be strictly concave and let agents have type-dependent subutility functions of the equicautious HARA class (5) with $\frac{y}{\gamma} + \bar{\tau}(z) > 0$ for all y and z.*

(i) Asset prices at $t = 1$ in any conditional equilibrium are given by $\bar{p}(y; z)$ defined in (28) with $\sum_y \bar{p}(y; z)\rho(y|1) = 1$ as the numeraire. Hence, these are the asset prices at $t = 1$ that decentralize both the optimal and equilibrium allocations. The price of a share of the aggregate endowment at $t = 1$ given z is denoted $\bar{D}_1(z)$ and given by (16) at asset prices $\bar{p}(y, z)$.

(ii) The socially optimal allocation of certainty equivalent consumption is given by $C^*(0) = \bar{C}(0)$ and $C_1^*(\tau; z) = \bar{C}_1(z)$ defined in (30) while the allocation of certainty equivalent consumption in the equilibrium is incomplete asset markets is given by $C^e(0) = \bar{C}(0)$ and $C_1^e(\tau; z)$ given as in (31).

(iii) In the equilibrium with complete asset markets, date $t = 0$ bond prices $Q^*(\tau; z)$ are given by (22) evaluated at the optimal allocation of certainty equivalent consumption with common marginal cost of certainty equivalent consumption given as in (32). In the equilibrium with incomplete asset markets, date $t = 0$ bond prices $Q^e(z)$ are given by (24) evaluated at the equilibrium allocation of certainty equivalent consumption (31) with common marginal cost of certainty equivalent consumption given as in (32).

(iv) Agents can implement the socially optimal allocation of consumption at time $t = 2$, $c^*(\tau, y; z)$, by trading at $t = 1$ their one share of the aggregate endowment for $\phi^*(\tau; z)$ shares of the aggregate endowment y given as in (33) and holding the remainder of their portfolio in risk-free bonds. This leads to share turnover of $TV^*(z)$ as in (34). Agents can implement the incomplete markets equilibrium allocation of consumption at time $t = 2$, $c^e(\tau, y; z)$, by trading at $t = 1$ their one share of the aggregate endowment for $\phi^e(\tau; z)$ shares of the aggregate endowment y given as in (35) and holding $\bar{D}_1(z)(1 - \phi^e(\tau; z))$ risk-free bonds. This leads to share turnover of $TV^e(z)$ as in (36).

Proof. This proposition follows from the four properties of the equicautious HARA utility function described above. The restriction that $\frac{y}{\gamma} + \bar{\tau}(z) > 0$ for all possible values of y in the statement of this proposition is required to ensure that the HARA subutility is well defined for all agents in equilibrium for all values of y. Details are given in the appendix. □

3.1 Sensitivity of Equilibrium C.E. Consumption to τ

Note that in the optimal allocation, since $C_1^*(\tau; z) = \bar{C}_1(z)$ for all τ, agents do not face idiosyncratic risk with respect to their allocation of certainty equivalent consumption at $t = 1$. In contrast, in equilibrium, from equation (31), each agent's realized certainty equivalent consumption at $t = 1$ is a linear function of their realized τ. As we discuss in Section 4, this idiosyncratic risk agents face at $t = 0$ plays an important role in determining asset prices at $t = 1$. Here we discuss the sensitivity of equilibrium certainty equivalent consumption to realized risk tolerance τ.

Equilibrium certainty equivalent consumption (31) is an increasing function of agents' realized risk tolerance τ, with slope given by $(\bar{C}_1(z) - \bar{D}_1(z))/(\frac{\bar{D}_1(z)}{\gamma} + \bar{\tau}(z))$.

Consider first the term $\frac{\bar{D}_1(z)}{\gamma} + \bar{\tau}(z)$. That this term is positive follows from the restriction that $\frac{y}{\gamma} + \bar{\tau}(z) > 0$ for all possible values of y. Since the equilibrium risk-free interest rate between $t = 1$ and $t = 2$ is normalized to one, we must have $y_{min} \leq \bar{D}_1(z) \leq y_{max}$ and hence this term is positive as well.

Now consider the term $\bar{C}_1(z) - \bar{D}_1(z)$. Note that $\bar{C}_1(z)$ can be interpreted as the cost of purchasing the aggregate or average level of certainty equivalent consumption entirely through sure bonds. In contrast, since an agent with the average level of risk tolerance indexed by $\bar{\tau}(z)$ simply holds his or her one share of the aggregate endowment, $\bar{D}_1(z)$ is the cost that agent actually pays in the market to attain certainty equivalent consumption $\bar{C}_1(z)$. Clearly then this term is a measure of the aggregate consumption risk premium $\bar{C}_1(z) - \bar{D}_1(z) \geq 0$ and is larger the greater the amount of aggregate risk and the smaller is the average risk tolerance in the economy $\bar{\tau}(z)$.

The previous derivations highlight the importance of the difference between $\bar{C}(z)$ and $\bar{D}_1(z)$. The next proposition gives a characterization.

Proposition 2. *If $U_{\bar{\tau}(z)}(\cdot)$ is* any *strictly concave increasing function, and if $\sigma^2(z) \equiv Var(y|z) > 0$ then $\bar{C}_1(z) > \bar{D}_1(z)$. Moreover:*

$$\bar{C}_1(z) - \bar{D}_1(z) = \frac{\bar{D}_1(z)}{2} + o(\sigma^2(z))\,.$$

where $o(\sigma^2)$ denotes an expression of order smaller than $\sigma^2(z)$.

The proof can be found in the appendix.

Here we have only used concavity of $U_{\bar{\tau}(z)}$, hence this applies to all HARA utility functions for the representative agent. In this case we have:

$$-\frac{U''_{\bar{\tau}}(\bar{y})}{U'_{\bar{\tau}}(\bar{y})} = 1 \Big/ \left[\frac{\bar{y}}{\gamma} + \bar{\tau}\right]$$

Moreover, the difference $\bar{C}_1 - \bar{D}_1$, and the level of $\bar{C}_1$ can be approximated by $\bar{D}_1$, which in principle can be measured by the risk premium.

3.2 Solving the Model as an Endowment Shock Model:

When agents have subutility functions of the equicautious HARA class (5), then the equilibrium allocations of certainty equivalent consumption in our model and the associated date $t = 0$ asset prices are equivalent to those of the following economy with idiosyncratic endowment shocks but no preference shocks. This result, which we demonstrate here, follows from the properties of the equicautious HARA preferences used above. We spell out this mapping of the model to an endowment shock economy to highlight the mathematical connection between the role of idiosyncratic risk in certainty equivalent consumption at $t = 1$ due to risk tolerance shocks in shaping asset prices in our model to the role of idiosyncratic risk in consumption at $t = 1$ due to endowment shocks in shaping asset prices in Mankiw (1986) and Constantinides and Duffie (1996). Of course, in our model, the equilibrium allocation of certainty equivalent consumption at $t = 1$ is implemented with a positive volume of asset trade while there is no such trade in the endowment shock economy.

Consider an economy with two time periods, $t = 0$ and $t = 1$. Let agents face aggregate uncertainty indexed by z and idiosyncratic uncertainty indexed by τ. Let the probability that state z is realized at time $t = 1$ be given by $\tilde{\pi}(z)$ with change of measure

$$\tilde{\pi}(z) = \frac{J(z)\pi(z)}{\sum_{z'} J(z')\pi(z')}.$$

The term $J(z)$ is the inverse of the marginal cost of certainty equivalent consumption which, in equilibrium, is common to all agents and, from equation (32), is given by

$$J(z) \equiv \frac{\sum_y U'_{\bar{\tau}(z)}(y)\rho(y|z)}{U'_{\bar{\tau}(z)}(\bar{C}_1(z))} = \frac{\sum_y \left[\frac{y}{\gamma} + \bar{\tau}(z)\right]^{-\gamma} \rho(y|z)}{\left[\frac{\bar{C}_1(z)}{\gamma} + \bar{\tau}(z)\right]^{-\gamma}} \tag{37}$$

Note that in the case with CARA utility, we have $J(z) = 1$ for all z.[4]

Let the distributions of the idiosyncratic uncertainty faced by agents be given by $\mu(\tau|z)$. Assume that an agent who has realized type τ in state z has endowment at $t = 1$

$$Y_1(\tau; z) \equiv \bar{C}_1(z) + \left(\frac{\tau - \bar{\tau}(z)}{\frac{\bar{D}_1(z)}{\gamma} + \bar{\tau}(z)}\right) \left[\bar{C}_1(z) - \bar{D}_1(z)\right]$$

Let the allocation of consumption at $t = 1$ be denoted by $C_1(\tau; z)$. This allocation must satisfy the pseudo-resource constraint (29). As before, let all agents be endowed with $Y_0 = \bar{C}_0$ at time $t = 0$.

Let agents have preferences over allocations given by

$$V(C_0) + \tilde{\beta} \sum_z \sum_\tau V(C_1(\tau; z))\mu(\tau|z)\tilde{\pi}(z)$$

[4] For values of $\gamma < \infty$, we have the Taylor approximation around the conditional mean realization of the endowment, $\bar{y}(z)$

$$J(z) \approx 1 + \frac{\sigma_y^2(z)}{2} \frac{1/\gamma}{\left(\frac{\bar{y}(z)}{\gamma} + \bar{\tau}(z)\right)^2} \tag{38}$$

Hence, holding γ fixed, $J(z)$ is increasing in the conditional variance of the endowment, $\sigma_y^2(z)$, and decreasing in the average risk tolerance across agents $\bar{y}(z)/\gamma + \bar{\tau}(z)$ if and only if $\gamma > 0$.

with

$$\tilde{\beta} \equiv \beta \sum_z J(z)\pi(z)$$

In the equilibrium of this endowment shock economy with complete asset markets, let agents choose consumption $C(0), C_1(\tau; z)$ and bondholdings $B(\tau; z)$ to maximize utility subject to budget constraints (11) with Y_0 replacing $\bar{C}_0$ at $t = 0$ and, at $t = 1$

$$C_1(\tau; z) = Y_1(\tau; z) + B(\tau; z)$$

The bond market clearing conditions are given by $\sum_\tau \mu(\tau|z)B(\tau; z) = 0$ for all z.

In the equilibrium of this endowment shock economy with incomplete asset markets, let agents choose consumption $C(0), C_1(\tau; z)$ and bondholdings $B(z)$ to maximize utility subject to budget constraints (13) with Y_0 replacing $\bar{C}_0$ at $t = 0$ and, at $t = 1$

$$C_1(\tau; z) = Y_1(\tau; z) + B(z)$$

The bond market clearing conditions are given by $B(z) = 0$ for all z.

Proposition 3. *The equilibrium allocations $C_0, C_1(\tau; z)$ and date zero bond prices Q with complete and incomplete asset markets are equivalent to the equilibrium allocations of certainty equivalent consumption and date zero bond prices Q with complete and incomplete asset markets for the corresponding taste shock economies.*

Proof. Note that with the change of measure to $\tilde{\pi}(z)$ and the rescaling of the discount factor $\tilde{\beta}$, the bond pricing conditions (22) for complete asset markets and (24) for incomplete asset markets are the same in the two economies. Direct calculation then shows that the equilibrium allocations and date zero bond prices in our preference shock economy with complete and incomplete asset markets are also equilibrium allocations and bond prices in this endowment shock economy with complete and incomplete asset markets and vice versa. □

This proposition is also useful in establishing a bound on the extent of downside idiosyncratic risk to certainty equivalent consumption that agents can face in this economy. This bound on the downside risk that agents can face does put a bound on the extent to which this idiosyncratic risk can impact asset prices at $t = 0$. Specifically, note that the parameter restrictions we need to ensure that our HARA utility is well defined imply that the lowest possible endowment $Y_1(\tau; z)$ that can be realized is $\bar{D}_1(z)$.

This lower bound has a simple economic interpretation — an agent in our preference shock economy always has the option at $t = 1$ to trade his or her endowment of one share, at price $\bar{D}_1(z)$, for a portfolio comprised entirely of risk free bonds, hence ensuring certainty equivalent consumption of $\bar{D}_1(z)$. Thus, in the equilibrium with incomplete asset markets, the gap between the certainty equivalent consumption of the agent with the lowest realized risk tolerance and the average level of certainty equivalent consumption in the economy is always bounded above by the measure of the aggregate risk premium given by $\bar{C}_1(z) - \bar{D}_1(z)$. This bound restricts the downside risk that agents face ex-ante and hence the premia they are willing to pay at $t = 0$ to avoid the impact of this preference risk on their certainty equivalent consumption at $t = 1$.

4 Trade Volumes and Asset Prices

In proposition 1, we provided a complete characterization of equilibrium allocations and asset prices under the assumption that agents have subutility functions of the equicautious HARA class. We also characterized trade volumes in asset markets at $t = 1$ under the assumption that agents trade only shares of the aggregate endowment and risk free bonds. In this section, we study the implications of our model for the joint distribution of trade volumes and asset prices in greater detail.

From equations (34) and (36), we have that observed trade volumes in shares at $t = 1$ given z in the equilibrium with either complete or incomplete markets is determined by the dispersion in realized risk tolerances across agents measured as the mean absolute deviation of those risk tolerances relative to average risk tolerance. We say that our model implies a *direct* connection between trade volumes and asset prices if measures of the dispersion in realized risk tolerances across agents also appear in our formulas for asset pricing. We say that the predicted relationship between trade volumes and asset prices is *coincidental* if the model implies relationships between observed trade volumes and asset prices or returns only because of an assumed correlation between dispersion in realized risk tolerances and some other random variable important for risk pricing, such as average risk tolerance $\bar{\tau}(z)$, or the quantity of endowment risk remaining in the economy as encoded in the conditional distribution $\rho(y|z)$.

In this section, we first discuss our model's implications for observed trading volumes and the serial correlation of asset returns from $t = 0$ to $t = 1$ versus from $t = 1$ to $t = 2$. We are motivated to do so to compare our results to those of Campbell, Grossman, and Wang (1993). Here we find that this connection is *coincidental*.

We then discuss our model's implications for observed trading volumes and the expected returns as of $t = 0$ on assets with different payoffs. Here we find results similar to those in Mankiw (1986) and Constantinides and Duffie (1996) regarding idiosyncratic risk and asset pricing. There is a *direct* connection between dispersion in agents' realized risk tolerances, and hence trade volumes, and ex-ante asset prices. The relationship exists precisely because there is a direct connection between dispersion in agents' realized risk tolerances and dispersion in agents' realized certainty equivalent consumption at $t = 1$. To the extent that aggregate shocks to z index both shocks to dispersion in τ, through $\mu(\tau|z)$, and shocks to average risk tolerance $\bar{\tau}(z)$ and/or to the amount of aggregate risk, through $\rho(y|z)$, then this idiosyncratic risk is priced just as in Mankiw (1986) and Constantinides and Duffie (1996).

4.1 Trading Volumes and the Serial Correlation in Asset Returns

We now consider the potential of our model to generate relationships between observed trade volume and the serial correlation of asset returns from $t = 0$ to $t = 1$ versus from $t = 1$ to $t = 2$. Here we find that any such connection is coincidental.

Specifically, we have seen above that asset prices at time $t = 1$ for claims that pay off at $t = 2$ in any conditional equilibrium are given by $\bar{p}(y; z)$. Moreover, from equation (28), these asset prices are determined entirely by the realized average risk tolerance of agents in the economy $\bar{\tau}(z)$ and the remaining endowment risk in the economy. Hence, in any conditional equilibrium, the expected returns from $t = 1$ to $t = 2$ on any asset, given by

$$\frac{\mathbb{E}_1\left[d(y, z)|z\right]}{P_1(z; d)} \equiv \sum_y \frac{d(y; z)}{P_1(z; d)} \rho(y|z)$$

where $P_1(d; z)$ is defined as in (17) has no direct connection to trade volume realized at $t = 1$. Likewise, the gap between the realized return on this asset from $t = 0$ to $t = 1$ given by

$$\frac{P_1(z; d)}{P_0(d)} - \sum_z \frac{P_1(z; d)}{P_0(d)} \pi(z)$$

where $P_0(d)$ is defined as in (18) has no direct connection to trade volume realized at $t = 1$. Therefore, in the equilibrium with either complete or incomplete asset markets, any connection between realized trade volumes at $t = 1$ and the serial correlation of asset returns from $t = 0$ to $t = 1$ and $t = 1$ to $t = 2$ is *coincidental*.

We now discuss our model's implications for trading volume and expected excess returns as of date $t = 0$.

4.2 Trading Volume and Ex-ante Prices and Returns

From Proposition (1), we have that in the equilibrium with complete asset markets, all certainty equivalence consumption for all agents is insured against idiosyncratic shocks to risk tolerance and hence date $t = 0$ asset prices $Q^*(z) \equiv \sum_\tau Q^*(\tau; z)\mu(\tau|z)$ are not directly connected to the idiosyncratic shocks to risk tolerance that drive trading volume at $t = 1$.

In the equilibrium with incomplete asset markets, however, this is not the case. In that equilibrium, agents' certainty equivalent consumption is exposed to idiosyncratic shocks to their risk tolerance. In this section we consider the impact of these idiosyncratic shocks on asset pricing both in terms of additive expected excess returns and multiplicative expected excess returns.

The price at $t = 0$ of a riskless bond, i.e., a claim to a single unit of consumption at $t = 2$ for each possible realization of τ, z, and y, is given by $P_0(1) = \sum_z Q(z)\pi(z)$.

Consider a security with payoffs $d(y|z)$ at $t = 2$. The time $t = 0$ multiplicative excess return of a claim with consumption d at $t = 2$ is denoted by $\mathcal{E}_{0,2}(d)$ and it is defined as:

$$\mathcal{E}_{0,2}(d) \equiv \frac{\mathbb{E}_0\left[d(y,z)\right]}{P_0(d)} \Big/ \frac{1}{P_0(1)} \tag{39}$$

Analogously, the multiplicative excess return of a claim with consumption d at $t = 2$ bought at $t = 1$ in sate z is denoted by $\mathcal{E}_{1,2}(d)$ and it is defined as:

$$\mathcal{E}_{1,2}(z;d) \equiv \frac{\mathbb{E}_1\left[d(y,z)|z\right]}{P_1(z;d)} \Big/ \frac{1}{P_1(z;1)} = \frac{\mathbb{E}_1\left[d(y,z)|z\right]}{P_1(z;d)} \tag{40}$$

since we use the normalization $P_1(z; 1) = 1$ for all z. The expression for $P_1(z; d)$ are given by:

$$P_1(z;d) = \sum_y \bar{p}(y;z)d(y,z)\rho(y|z) \tag{41}$$

where $\bar{p}(y; z)$ are given by the representative agent marginal utility defined in (28). The expectations $\mathbb{E}[\cdot]$ are taken with respect to the statistical distribution, i.e., using the probability distributions π and ρ.

We have

$$Q^e(z) = \beta\frac{V'(\bar{C}_1(z))}{V'(\bar{C}_0)}J(z)L(z) \tag{42}$$

with

$$L(z) \equiv \sum_\tau \frac{V'(C_1^e(\tau;z))}{V'(\bar{C}_1(z))}\mu(\tau;z) \tag{43}$$

and $J(z)$ is defined as in equation (37).

Importantly, we can relate the bond prices at $t = 0$ of the economy with incomplete markets to the economy with complete markets as follows

$$Q^e(z) = Q^*(z)\, L(z)\,. \tag{44}$$

where $Q^*(z)$ is given by (22) and (25), with $\bar{C}_1$, C^e and D_1 are defined in (30), (31) and $D_1(z) = \sum_y \bar{p}(y; z) y \rho(y|z)$. This gives a complete characterization of asset prices for the incomplete market economy. Next we turn to its analysis based on these expressions.

Trade Volumes, $L(z)$, and Asset Prices: In the incomplete markets equilibrium, there is a direct connection between asset prices and the dispersion of the preference shocks τ realized at $t = 1$ in state z. This connection comes through the term $L(z)$ in Q^e. Under the assumption that $V''' > 0$, the term $L(z)$ is equal to one if there is no dispersion in τ and is strictly increasing in the dispersion in τ. Hence, the corresponding probabilities used in pricing assets, $\pi_{Q^e}(z)$, are strictly increasing in the dispersion in τ.

$$V'''(\cdot) \geq 0 \text{ implies } L(z) = \sum_\tau \frac{V'(C_1^e(\tau; z))}{V'(\bar{C}_1(z))} \mu(\tau; z) \geq 1$$

and using a Taylor expansion we have:

$$\begin{aligned} L(z) &\approx 1 + \frac{1}{2} \frac{V'''(\bar{C}_1(z))}{V'(\bar{C}_1(z))} \sum_\tau \left[C_1^e(\tau, z) - \bar{C}_1(z) \right]^2 \mu(\tau; z) \\ &= 1 + \frac{1}{2} \frac{V'''(\bar{C}_1(z))}{V'(\bar{C}_1(z))} \left[\bar{C}_1(z) - \bar{D}_1(z) \right]^2 \sum_\tau \left[\phi^e(\tau; z) - 1 \right]^2 \mu(\tau; z) \end{aligned} \tag{45}$$

Hence, if $V'''(\bar{C}_1(z)) > 0$, then our approximation to $L(z)$ is directly proportional to the variance of individual share trades times the square of the aggregate consumption risk premium as measured by $\left(\bar{C}_1(z) - \bar{D}_1(z)\right)^2$.

As an example consider the case of a uniform distribution, and then we consider a more general case. For a uniform distribution of τ, the mean absolute deviation of τ from $\bar{\tau}(z)$ is directly proportional to the standard deviation of τ and hence, in this case, to a second order approximation, data on the square of trading volume in state z is a valid proxy for the term $\sum_\tau \left(\phi^e(\tau; z) - 1\right)^2 \mu(\tau|z)$ in our approximation to $L(z)$.

Of course, the previous result that the square of trading volume is directly proportional to the dispersion of agents' marginal utilities of certainty equivalent consumption is special to the case of uniform shocks. More generally, if one had data on the distribution of trade sizes, one could potentially map data on trade volumes to empirical proxies for $L(z)$ using the relevant distributional assumptions. Moreover, since $|x-1|$ and $(x-1)^2$ are both convex functions, if we replace the distribution of risk-tolerance with one with a mean preserving spread, both trade volume and $L(z)$ increase. In this sense, both trade volumes and $L(z)$ are increasing in the dispersion of idiosyncratic risk-tolerance.

Overall, this analysis show that, using the definition stated above, trade volume and asset prices have a *directly* connection. We now use this connection to develop three results regarding the impact of trading volumes on asset pricing. The first one is a result about interest rates, the second one is a comparison of risk premium across economies with different patters of trading volume, and the third one is a comparison of the risk premium of different assets in the same economy.

For these results it is useful to collect two properties of asset prices. i) Using the option of re-trading we have that

$$P_0(d) = \sum_z Q^*(z)L(z)\pi(z)P_1(z;d) \tag{46}$$

for all assets with dividend d. This expression will be key for our first result on interest rates. Moreover, this expression together with the previous definitions of the excess returns give the following expression for the (inverse) time $t=0$ excess return as a weighted average of the time $t=1$ excess returns:

$$\frac{1}{\mathcal{E}_{0,2}(d)} = \sum_z \frac{Q^*(z)L(z)\pi(z)}{\sum_{z'} Q^*(z')L(z')\pi(z')} \frac{1}{\mathcal{E}_{1,2}(z;d)} \frac{\mathbb{E}_1\left[d(y,z)|z\right]}{\mathbb{E}_0\left[d(y,z)\right]} \tag{47}$$

This expression will be key to show the two results on excess returns. And ii) the only expression that involves the dispersion of $\mu(\cdot|z)$ is the term $L(z)$. The expressions for $Q^*(z)$, $\pi(z)$, $P_1(z;d)$, $\mathcal{E}_{1,2}(z;d)$ and $\mathbb{E}_1\left[d|z\right]/\mathbb{E}_0\left[d\right]$ are not functions of shape of $\mu(\cdot|z)$, and hence they are independent of trade volume. The excess return $\mathcal{E}_{1,2}(z;d)$ is identical to the one of a standard representative agent economy, so that the only feature of $\mu(\cdot|z)$ that matters is its first moment $\bar{\tau}(z)$.

Trade Volume and Interest Rates: We have the following comparative static result regarding the dispersion of shocks to risk tolerance and time $t=0$ bond prices

Proposition 4. *Consider two economies in which agents have the same preferences with* $V'''(\cdot) > 0$ *and face the same distribution of endowments,* $\bar{C}_0$, $\pi(z)$ *and* $\rho(y|z)$. *Assume that the distribution of shocks to risk aversion in the two economies* $\mu(\tau|z)$ *and* $\mu'(\tau;z)$ *are such that, for all* j, $\bar{\tau}(z_j) = \bar{\tau}'(z_j)$. *Then these two economies have the same equilibrium values of*

$\bar{C}_1(z)$ and $J(z)$, but, for each state z, the economy with the higher dispersion in shocks to risk aversion as measured by the a higher value of $L(z)$ has the higher equilibrium bond price at $t = 0$, $Q^e(z)$.

Proof. The proof is by direct calculation. □

Note that interest rates are the reciprocal of $\sum_z Q^e(z)\pi(z)$, so that the higher bond prices translate into lower interest rates.

Trading Volume and Expected Returns across economies: For the first result we compare an economy with the same dispersion of risk-tolerance across different states at $t = 1$ with one where the market-wide risk tolerance is negatively correlated with the dispersion of risk-tolerance. We find that if V displays prudence (i.e. if $V''' > 0$) then any cash-flow with systematic risk has a higher risk-premium in the economy in which dispersion is negatively correlated with risk tolerance.

Denote by $\tilde{\mu}(\cdot|z)$ the distribution of $(\tau - \bar{\tau}(z))/(\bar{D}(z)/\gamma + \bar{\tau}(z))$ conditional on z. We consider the following assumptions:

$$\text{If } z' > z \text{ then } \bar{\tau}(z') > \bar{\tau}(z) \text{ and} \tag{48}$$

$$\text{If } z' > z \text{ then } \tilde{\mu}(\cdot|z') \text{ is a less dispersed (in second order stochastic sense) than } \tilde{\mu}(\cdot|z) \tag{49}$$

In words, states with higher market wide risk tolerance have a lower dispersion of risk tolerance, and thus a lower volume of trade at $t = 1$. We say that an asset has systematic payoff exposure if d is increasing in y, and independent of z:

$$d(y', z') > d(y, z) \text{ for all } z, z' \text{ and } y' > y\,. \tag{50}$$

With this notation at hand we can state one of our main results:

Proposition 5. *Let the distribution of y conditional on z, $\rho(y|z)$ be constant across z. Consider two economies where shock z indexes market-wide risk tolerance as in (48). The first economy has constant dispersion on the idiosyncratic risk-tolerance across states at time $t = 1$, so $L(z) = L_1(z)$ is constant for all z. The second economy, has $\mu(\cdot|z)$ more dispersed for lower market-wide risk tolerance as defined in (49), so $L(z) = L_2(z)$ is decreasing in z. For both economies we assume that market wide risk tolerance $\bar{\tau}(z)$ is increasing in z. We fix the same asset $d(\cdot)$, with a systematic payoff exposure as defined in (50) in both economies. If investors are prudent* *(i.e they have precautionary savings motives, or $V''' > 0$), then the second economy (where the cross sectional dispersion in risk tolerance is negatively correlated with the market-wide risk tolerance) has a higher $t = 0$ expected excess return $\mathcal{E}_0(d)$.*

We can describe the second economy as one where the cross sectional dispersion in risk tolerance is negatively correlated with the market-wide risk tolerance, or equivalently as one where trade volume is positively correlated with the market-wide risk aversion. So that times of high market wide risk aversion, where risky asset sell at a low price, are times where lots of investors want to rebalance, making the second economy, ex-ante, risker than the first. In other words, in the second economy investors face lots of rebalancing risk.

This proposition parallels the results in Mankiw (1986) and Constantinides and Duffie (1996). In both papers the authors consider the level of excess expected returns when investors have uninsurable labor risk whose dispersion is correlated with the level of aggregate consumption. In both cases the authors show that this implies that the excess returns on an aggregate risky portfolio is smaller than an otherwise identical economy with complete markers or without dispersion on idiosyncratic income shocks.

Trading Volume and Expected Returns in the cross section of assets: In the second result we compare the risk premium across risky assets in the same incomplete market economy. In this case we find that if V displays prudence (i.e. if $V''' > 0$), asset whose cash-flows loads more into the time $t = 1$ states with higher dispersion, have higher prices or lower expected returns. Since trade volume is also given by a measure of dispersion, this second result means that assets whose cashflows load on states of high trade volume have low expected excess returns.

To make this result precise, we fix an economy with incomplete markets and compare the excess expected returns of assets with different exposures to the idiosyncratic dispersion of risk-tolerance. We assume that the average risk tolerance and the distribution of the endowment y conditional on z are both constant across z.

Proposition 6. *Consider an economy with incomplete markets with the same market-wide risk-tolerance and same conditional distribution of aggregate risk $\rho(y|z) = \bar{\rho}(y)$ for all states z. Assume that the states z are ordered in terms of dispersion of idiosyncratic risk-tolerance as in (49), so that $L(z)$ decreases with z. Consider two cash-flows, $\tilde{d}$ and d in the same economy, where $\tilde{d}$ loads more than d in states with higher dispersion of risk-tolerance in the following* *way:* $d(y,z) = \delta(y)\tilde{e}(z)$ *and* $\tilde{d}(y,z) = \delta(y)e(z)$ *with* $\tilde{e}(z)/e(z)$ *decreasing in* z*. Then, i) the time* $t = 1$ *conditional expected excess return are the same for both assets and all states* z*, i.e.,* $\mathcal{E}_{1,2}\big(z;\tilde{d}\big) = \mathcal{E}_{1,2}\big(z;d\big)$*, and ii) the time* $t = 0$ *excess expected return for the asset with higher exposure to trade is smaller, i.e.,* $\mathcal{E}_{0,2}\big(\tilde{d}\big) < \mathcal{E}_{0,2}\big(d\big)$*.*

This result gives the conditions under which trade volume acts as a pricing factor, i.e., the conditions under which the cross-sectional returns on assets (i.e., $\mathcal{E}_{0,2}(\tilde{d})$ vs $\mathcal{E}_{0,2}(d)$) depends on the correlation of returns with trade volume. In this case, the asset with dividend $\tilde{d}$, which has higher value when trade volume is high, and thus ex ante is a better hedge against the rebalance risk, has a higher price, i.e. it has a lower $t = 0$ excess expected return. The higher value of the asset with dividend $\tilde{d}$ is due to higher exposure of its dividends to trade volume, as captured by the term $\mathbb{E}_1[\tilde{d}|z]/\mathbb{E}_0[\tilde{d}] = \tilde{e}(z)/\mathbb{E}_0[\tilde{e}(z)]$.

5 Taxes on Trading and Ex-ante Welfare

In this section we consider the implications for welfare of a tax on trade in shares of the aggregate endowment at $t = 1$. We show that while a Tobin tax on trade has a zero first order effect on the socially optimal allocation, it has a first-order negative welfare effect on the equilbrium allocation. In other words, a Tobin subsidy to trade increases ex-ante welfare in equilibrium. The basic logic of this result is that a Tobin tax exacerbates the inefficient sharing of idiosyncratic preference risk in equilibrium. Agents who have negative risk tolerance shocks suffer a negative shock to certainty equivalent consumption in equilibrium. The Tobin tax also falls on them in terms of its tax incidence. Hence the tax exacerbates the inefficient sharing of risk in equilibrium.

To further understand the reasons behind this results, we also study a mechanism design problem where risk tolerance shocks τ are private information. We show that the fundamental tension in this environment is that efficient risk sharing calls for all agents to receive the same certainty equivalent consumption but that incentive compatibility requires that agents with higher risk tolerance receive higher certainty equivalent consumption. In this section we show that the optimal allocation is not incentive compatible and that the only incentive compatible mechanism based on transfers to agents contingent on their reported type with further trading allowed after transfers is the equal wealth equilibrium allocation. In the next section, we study the mechanism design problem in which the planner can control the consumption of agents and show that it shares certain features with a Tobin subsidy to trade.

5.1 Tobin Tax

In the standard equilibrium analysis of the welfare costs of a tax on the trading of any good in a setting in which agents have quasi-linear preferences for the good, the deadweight loss from the tax, measured by the Harberger triangle, is small in the sense that the derivative of welfare with respect to an increase in the tax is zero when evaluated at the undistorted equilibrium. We find the same result in our model. In addition, in the standard analysis, the incidence of the tax does not have a first order welfare effect either since there is no motive for redistribution in the initial equilibrium. This standard analysis carries through in our model at the optimal allocation.

In contrast, we find that when markets are incomplete in that agents cannot insure ex-ante against idiosyncratic shocks to risk aversion, at tax on trade in shares has a first-order negative impact on agents' ex-ante expected utility as of $t = 0$ due to the incidence of the tax.

In our analysis of a Tobin tax, we assume that there are two asset markets — one at $t = 0$ for bonds that pay off at $t = 1$ (contingent on realized τ and z in the complete markets economy and contingent only on z in the incomplete markets economy), and one at $t = 1$ in which agents trade shares of the aggregate endowment for sure claims to consumption at $t = 2$. Assume that trade in shares at $t = 1$ is taxed. Specifically, assume that there is a tax per share traded of ω such that if the seller receives price $D_1(z)$ for selling a share of the aggregate dividend at $t = 1$, the buyer pays $D_1(z) + \omega$, and the total revenue collected through this tax, equal to ω times the volume of shares traded, is rebated lump sum to all agents.

Here we use the result that since all agents are ex-ante identical, they do not trade bonds at $t = 0$ and hence they all hold a portfolio with no bonds and with a single share of the aggregate endowment at the start of period $t = 1$.

With this notation we define a conditional equilibrium with a transactions tax as follows.

Conditional Equilibrium with a share transactions tax. An *equilibrium conditional on z with a share transactions tax ω* is a share price $\{\bar{D}_1(z;\omega)\}$, transactions tax revenue rebate $T(z;\omega)$, post-trade holdings of share $s(\tau;z;\omega)$ that satisfy the market clearing condition

$$\sum_{\tau} s(\tau;z;\omega)\mu(\tau|z) = 0, \tag{51}$$

and corresponding allocation of consumption at $t = 2$, $c(\tau, y; z; \omega)$, that satisfy budget constraints,

$$c(\tau, y; z; \omega) = y + \left(\bar{D}_1(z; \omega) - \omega\right)(s(\tau; z; \omega) - 1) + \omega\, TV(z; \omega) + B(\tau; z)$$

if $s(\tau; z; \omega) \geq 1$ and

$$c(\tau, y; z; \omega) = y - \bar{D}_1(z; \omega)(s(\tau; z; \omega) - 1) + \omega\, TV(z; \omega) + B(\tau; z)$$

if $s(\tau; z; \omega) < 1$ where

$$TV(z; \omega) = \sum_{\tau : s(\tau; z; \omega) > 0} (s(\tau; z; \omega) - 1)\, \mu(\tau|z)$$

and that maximizes each agents' certainty equivalent consumption (3) among all share holdings and allocations of consumption that satisfy the budget constraints given the initial bondholdings, the share price, the tax, and the tax rebate.

Note that in the budget constraint we include the realized time $t = 1$ transfer $B(\tau, z)$ bought/sold at $t = 0$. In the case of an equal wealth equilibrium $B(\tau; z) = 0$. In the complete market case $B(\tau; z) = B^*(\tau, z)$:

$$B^*(\tau; z) = (\bar{\tau} - \tau) \frac{\bar{C}_1(z)}{\bar{\tau} + \frac{\bar{C}_1(z)}{\gamma}}$$

We denote by $C_1^i(\tau, \omega; z)$ the time $t = 1$ certainty equivalent consumption for agent with τ in state z for the conditional equilibrium with a transaction tax ω for $i \in \{*, e\}$ corresponding to the optimal allocation and equal wealth equilibrium.

Consider the following calculation of the change in ex-ante welfare from a marginal increase in the transactions tax ω starting from $\omega = 0$. Here we must compute

$$\frac{dW^i}{d\omega}|_{\omega=0} = \beta \sum_z \pi(z) \sum_\tau \mu(\tau|z) V'(C_1^i(\tau; z)) \frac{d}{d\omega} C_1(\tau; z) \tag{52}$$

where $C_1^i(\tau; z)$ is the undistorted allocation of certainty equivalent consumption (with $\omega = 0$) corresponding to either the equilibrium with complete or incomplete asset markets.

To compute this change in ex-ante welfare, we must compute the marginal change on the certainty equivalence consumption of a small tax, i.e. the derivative of $C_1^i(\tau; z; \omega)$ with respect to ω evaluated at $\omega = 0$. Using the envelope theorem, as well as the strong aggregation with equicautious HARA preferences in a conditional equilibrium we get that:

$$\frac{d}{d\omega}C_1^i(\tau;z;0) = \begin{cases} J(z)\left[(s^i(\tau;z;0)-1)\left(-\frac{\partial \bar{D}_1^i(z;0)}{\partial\omega}-1\right)+TV^i(0,z)\right] & \text{if } s^i(\tau;z;0)>1 \\ J(z)\,TV^i(0,z) & \text{if } s^i(\tau;z;0)=1 \\ J(z)\left[(s^i(\tau;z;0)-1)\left(-\frac{\partial \bar{D}_1^i(z;0)}{\partial\omega}\right)+TV^i(0,z)\right] & \text{if } s^i(\tau;z;0)<1 \end{cases} \tag{53}$$

for all τ, z and $i\in\{e,*\}$, and where $J(z)$ is given by expression (37).

Using (53) for either complete markets and incomplete markets, as well as market clearing for shares (51) we must have

$$\sum_\tau \mu(\tau|z)\frac{d}{d\omega}C_1^*(\tau;z) = \sum_\tau \mu(\tau|z)\frac{d}{d\omega}C_1^e(\tau;z) = 0$$

This result implies that a Tobin tax, at the margin, simply redistributed certainty equivalent consumption across agents. Since in the undistorted equilibrium with complete asset markets, $C_1^*(\tau;z)=\bar{C}_1(z)$ for all τ, the formula (52) then immediately implies the standard result that a share transactions tax has no first order impact on welfare starting from the undistorted equilibrium since all types of agents share the same initial marginal utilities of certainty equivalent consumption in each state z. We collect this result in a proposition.

Proposition 7. *Let $W^*(\omega;z)$ be the time $t=0$ ex-ante utility if at $t=1$ a Tobin tax ω is imposed in the complete markets equilibrium. This tax has a zero first order effect on welfare, i.e. $\frac{d}{d\omega}W^*(\omega;z)|_{\omega=0}=0$.*

In contrast, in the incomplete markets economy, the baseline equilibrium allocation of certainty equivalent consumption at $t=1$ is not socially efficient and, we show that under a wide set of assumptions a tax on trade in shares at $t=1$ has a first-order negative impact on ex-ante welfare because it exacerbates this baseline misallocation of certainty equivalent consumption.

Specifically, as shown in equation (31), certainty equivalent consumption for an agent with realized type τ in state z at $t=1$, is strictly increasing in the risk tolerance τ of that agent. Hence, if V is strictly concave, the marginal utility of certainty equivalent consumption for an agent with realized type τ in state z at $t=1$, $V'(C_1^e(\tau;z))$, is strictly decreasing in the risk tolerance of that agent. In this case, the restriction that the aggregate change in certainty equivalent consumption must be zero gives us that the total change in ex-ante welfare in equation (52) can be written

$$\frac{dW}{d\omega} = \beta\sum_z \pi(z)\mathbb{C}ov\left(V'(C_1(\tau;z)),\frac{d}{d\omega}C_1(\tau;z)\,|\,z\right) \tag{54}$$

where $\mathbb{C}ov(\cdot,\cdot\,|\,z)$ denotes the covariance of two variables dependent on τ conditional on z. As this result makes clear, the first-order impact on welfare of a tax on trading in shares is then determined by the question of whether it is agents with high or low marginal utilities of certainty equivalent consumption in the initial equilibrium allocation who bear the cost of the tax.

To study the incidence of a Tobin tax in an undistorted equilibrium with incomplete markets, we must solve for the terms for the changes in certainty equivalent consumption by type τ from equation (53) and then compute (54). We do so in steps as follows. We first compute the changes in demand for shares by each type τ of agent as a function of the change in price and lump sum transfer induced by the tax. We then compute the implied equilibrium change in price

$$\frac{\partial \bar{D}_1^e(z;0)}{\partial \omega}$$

and transfer from the share market clearing condition. We then compute (54) under various assumptions about the distribution of τ across agents. We first consider an economy with only two types of agents, $\tau \in \{\tau_1, \tau_2\}$ with $\tau_1 < \bar{\tau}(z) < \tau_2$ for all z, as long as $\mu(\tau_2)$ is not too small. Establishing this result relies on standard arguments about tax incidence inclusive of lump sum rebates of tax revenue. Then we extend the result to the case of a symmetric distribution μ as well as a regularity condition on V.

We first consider the derivatives of agents' demands for shares with respect to a change in the price of shares and a lump sum transfer. We evaluate these derivatives at the equilibrium allocation. Define $S(\tau; zD, T)$ as the optimal trade in shares for an investor with risk tolerance τ, facing a price D and receiving a transfer T. The first order condition for the risky asset trade is:

$$\mathbb{E}\left[U'_\tau\left(\,y + \left(S\left(\tau; z, D, T\right) - 1\right)\left(y - D\right) + T\,\right)\left(y - D\right)\,\mid z\right] = 0\,. \tag{55}$$

Differentiating this first order condition, and evaluating it at the equal wealth equilibrium we obtain:

Lemma 1. *Let $S(\tau; zD, T)$ be defined as the solution of (55) evaluated at the equilibrium price*

$D = \bar{D}_1(z)$ for equal wealth and at transfer $T = 0$. Then:

$$\frac{\partial S(\tau; z; D, T)}{\partial D} = \phi^e(\tau, z) \frac{\mathbb{E}\left[U'_{\bar{\tau}(z)}(y)\,|z\right]}{\mathbb{E}\left[U''_{\bar{\tau}(z)}(y)\,(y-D)^2\,|z\right]} + (\phi^e(\tau, z) - 1)\ \frac{\mathbb{E}\left[U''_{\bar{\tau}(z)}(y)\,(y-D)\,|z\right]}{\mathbb{E}\left[U''_{\bar{\tau}(z)}(y)\,(y-D)^2\,|z\right]}$$

$$\frac{\partial S(\tau; z; D, T)}{\partial T} = -\frac{\mathbb{E}\left[U''_{\bar{\tau}(z)}(y)\,(y-D)\,|z\right]}{\mathbb{E}\left[U''_{\bar{\tau}(z)}(y)\,(y-D)^2\,|z\right]}$$

Two comments about this lemma are in order as these results play an important role in our calculation of the welfare impact of a Tobin Tax. First, the result that the derivative of agents' demand for shares with respect to a transfer is common across all agents is simply an implication of the result that equicautious HARA preferences satisfy Gorman Aggregation. Second, the second term in the derivative of agents' demand for shares with respect to a change in price D, the term which reflects the income effects on demand from a change in price, cancel out when aggregated across all agents since the market for shares clears. This result is also an implication of the result that equicautious HARA preferences satisfy Gorman Aggregation. These two features of the demand for shares in our economy allow us to compute the change in price that arises from a change in the Tobin Tax simply as a function of trade volume and the fractions of agents who are buyers and sellers of shares. Other parameters of preferences do not enter into this calculation. We show this in the next proposition as follows.

Let $\bar{D}^e(z;\omega)$ be equilibrium price of a claim to the risky endowment with a tax on trade ω introduced in the equal wealth equilibrium and let $\bar{T}^e(z;\omega)$ be the associated lump sum transfer of tax revenues back to each agent. To compute the change in price and transfer due to a small Tobin Tax imposed on the equilibrium allocation, we differentiate the market clearing condition for shares

$$\sum_{\tau > \bar{\tau}(z)} \frac{\partial S(\tau; z; D, T)}{\partial D}\mu(\tau|z) + \frac{\partial \bar{D}_1^e(z;0)}{\partial \omega} \sum_{\tau} \frac{\partial S(\tau; z; D, T)}{\partial D}\mu(\tau|z) +$$
$$\frac{\partial \bar{T}_1^e(z;0)}{\partial \omega} \sum_{\tau} \frac{\partial S(\tau; z; D, T)}{\partial T}\mu(\tau|z) = 0$$

and the government budget constraint linking the tax to the lump sum transfer

$$\frac{\partial \bar{T}_1^e(z;0)}{\partial \omega} = \sum_{\tau > \bar{\tau}(z)} (\phi^e(\tau, z) - 1)\,\mu(\tau|z) = TV^e(z)$$

The first term in the derivative of the market clearing condition for shares is the direct impact of the Tobin Tax on the share demand of those agents who buy shares while the other two terms are the impact of the change in price and change in transfer on the demand for shares. The derivative of the government budget constraint follows from the standard result that if we start from a tax equal to zero, then the change in share demands has no first-order impact on revenue from a marginal increase in the tax.

Using lemma 1 and these conditions, we derive the following characterization for the impact of prices of a transaction tax:

Proposition 8. *Let $\bar{D}(z;\omega)$ be equilibrium price of a claim to the risky endowment with a tax on trade ω introduced in the equal wealth equilibrium. Assume, to simplify, that there are no marginal investors, i.e. μ has no mass point at $\tau = \bar{\tau}$. Then the price $\bar{D}(z;\omega)$ received by sellers decreases by the fraction of shares held post-trade by buyers times the Tobin Tax*

$$\frac{dD(z;0)}{d\omega} = -\sum_{\tau>\bar{\tau}} \phi^e(\tau;z)\mu(\tau|z) = -\left[TV^e(z) + \sum_{\tau>\bar{\tau}} \mu(\tau|z)\right] \in (-1\,,\,0)\;. \tag{56}$$

As mentioned above, this result gives us the simple result that the change in price from a Tobin Tax is simply a function of equilibrium post-trade portfolios and hence a function only of equilibrium trade volume and the fraction of agents who are buyers. No other preference parameters enter into the calculation.

We now use this result regarding the change in price due to a Tobin Tax together with equation (53) evaluated at equilibrium share holdings $\phi^e(\tau;z)$ to compute the incidence of the tax in terms of the changes in certainty equivalent consumption for agents with different realized values of τ.

We first show that if the distribution of preference shocks $\mu(\tau|z)$ is symmetric, then, on average, those experiencing negative shocks to risk tolerance (sellers of shares), lose certainty equivalent consumption and those experiencing positive shocks to risk tolerance (buyers of shares), gain certainty equivalent consumption. In fact, the next proposition shows that the average gains (or losses) for buyer (and sellers) of risk asset after the introduction of a small transaction tax ω are an extremely simple function of trade volume prior to the introduction of taxes.

Proposition 9. *Assume that $\mu(\cdot|z)$ is symmetric around $\bar{\tau}$. Then the average consumption equivalent gain among all buyers (respectively losses among sellers) of risky asset are propor-*

tional to the square of trade volume:

$$Avg.\ Gain\ Buyers \equiv \sum_{\tau > \bar{\tau}} \frac{d}{d\omega} C_1^e(\tau; z; 0) \frac{\mu(\tau|z)}{\sum_{\tau' > \bar{\tau}} \mu(\tau'|z)}\, \omega + o(\omega) = +2\, J(z)\, \left[TV^e(z)\right]^2\, \omega + o(\omega),$$
$$Avg.\ Loss\ Sellers \equiv \sum_{\tau < \bar{\tau}} \frac{d}{d\omega} C_1^e(\tau; z; 0) \frac{\mu(\tau|z)}{\sum_{\tau' < \bar{\tau}} \mu(\tau'|z)}\, \omega + o(\omega) = -2\, J(z)\, \left[TV^e(z)\right]^2\, \omega + o(\omega)\,.$$

Three comments about this proposition are in order. First a corollary of the previous proposition is that in the case of a symmetric distribution $\mu(\cdot|z)$ with only *two values of* τ, there is a first order welfare loss of introducing a transaction tax ω. This is because the marginal utility of buyers of risky assets is discretely below than the marginal utility of sellers. Second, since this results gives a strict inequality, it suggests that in the case of two values of τ one can relax the assumption of symmetry of $\mu(\cdot|z)$. Indeed Proposition 10 shows that. Third, and more subtly, the result in Proposition 9 does *not* imply that assuming symmetry there is a first order loss in welfare for a trade tax ω. The reason why this is not sufficient is that there can also be redistribution among sellers and or among buyers. Proposition 11 impose extra conditions on the utility function V so that these potential redistributional effects don't overcame the result.

We now prove our result that a Tobin-tax results in a first order welfare loss in an economy with only two possible realizations of τ. We then present this result in an economy with a symmetric distribution of shocks to risk tolerance $\mu(\tau|z)$.

Proposition 10. *Fix a z. Consider an economy with only two types of agents, $\tau \in \{\tau_1, \tau_2\}$ with $\tau_1 < \bar{\tau}(z) < \tau_2$ and thus $\phi^e(\tau_1; z) < 1 < \phi^e(\tau_2; z)$. Assume that V is strictly concave. Then, when agents have equicautious HARA preferences a tax on asset trade on the equal wealth equiibrium has a negative first order ex-ante welfare effect if and only if*

$$\frac{dW^e(0; z)}{d\omega} < 0 \iff \mu(\tau_2; z) > \frac{\phi^e(\tau_1|z)}{1 + \phi^e(\tau_1; z)} \tag{57}$$

Note that since $\phi^e(\tau_1; z) < 1$ then $\mu(\tau_2|z) > 1/2$ is a sufficient conditions for the Tobin tax to have a first order welfare loss. Also, symmetry of the distribution of τ implies $\mu(\tau_2|z) = 1/2$ and hence satisfies condition (57).

Now we extend the result to the case of a general symmetric distribution $\mu(\cdot, z)$ and where V is concave with derivatives that alternate signs.

Proposition 11. *Fix a state z. Assume that there are no marginal investors, i.e. $\mu(\cdot|z)$ has no mass point at $\tau = \bar{\tau}$, and that the distribution of τ is symmetric, i.e. $\mu(\bar{\tau} - a; z) = \mu(\bar{\tau} + a; z)$*

for all a. Furthermore assume that the ex-ante utility V is analytical, strictly increasing, and strictly concave, with all derivatives evaluated at $\bar{C}_1(z)$ alternating signs, i.e.:

$$sign\left(\frac{\partial^{n+1} V(C)}{\partial C^{n+1}}\right) = -sign\left(\frac{\partial^{n} V(C)}{\partial C^{n}}\right) \quad \text{evaluated at } C = \bar{C}_1(z), \text{ and all } n = 1, 2, 3, \ldots \tag{58}$$

Then, when agents have equicautious HARA preferences a tax on asset trade on the equal wealth equiibrium as a negative first order ex-ante welfare effect for each z, i.e.:

$$\frac{d}{d\omega} W^e(0; z) < 0 \tag{59}$$

Moreover, approximating the change on ex-ante utility in terms of moments of τ, and using the first leading term we obtain:

$$\frac{d}{d\omega} W^e(0; z) \approx J(z)\, V''(\bar{C}_1(z)) \left(\frac{\bar{C}_1(z) - \bar{D}_1(z)}{\left[\frac{\bar{D}_1(z)}{\gamma} + \bar{\tau}\right]^2} \right) TV^e(z)\, Var(\tau|z) \tag{60}$$

where TV^e is the trade volume in the equal wealth equilibrium.

A few comments are in order. First, the assumption that the derivatives of V change sign include the case of polynomials, such as quadratic utility. Second, since V is concave, this assumption is consistent with V displaying prudence, a key property that we use above in the asset pricing implications. Third, although we did not emphasize this in the statement of the proposition, in the proof we show that every extra term in the approximation corresponding to higher order derivatives is negative. Fourth, most commonly used utility functions satisfy the condition that derivatives of higher order change signs, such as all HARA utility functions. Finally, we can evaluate the expression (60) for the particular case where U_τ is CARA (so that $\gamma \to \infty$) and $y \sim N(\mu, \sigma^2(z))$. In this case $\bar{C}_1(z) - \bar{D}_1(z) = .5\sigma^2(z)/\bar{\tau}(z)$, and $J(z) = 1$. Additionally if $\mu(\cdot|z)$ is uniform then $Var(\tau|z) = [TV^e(z)]^2 [\bar{\tau}(z)]^2\, 64/12$. Hence, writing the expected utility $W^e(\omega; z)$ in ex-ante equivalent terms, i.e. defining $C^e_w(\omega; z)$ as $V(C^e_w(\omega; z)) = \sum_\tau V(C^e_a(\tau; z; \omega))\mu(\tau)$ we get:

$$\frac{d}{d\omega} C^e_w(0; z) \approx \frac{16}{3} \frac{V''(\bar{C}_1(z))}{V'(\bar{C}_1(z))} \frac{\sigma^2(z)}{2\bar{\tau}(z)} [TV^e(z)]^3$$

Thus the welfare loss of a Tobin tax is proportional to the curvature of the utility function V, the representative agent time $t = 1$ risk premium $.5\sigma^2(z)/\bar{\tau}(z)$, and the *cube of the trade volume.*

Note that in this economy, when a Tobin tax has a first order negative impact on welfare if applied to the equilibrium allocation with incomplete markets, then a Tobin subsidy to trade must have a positive first order impact on ex-ante welfare. This observation raises the question of what the optimal subsidy to trade looks like. We take up this question next.

6 A mechanism design approach to Tobin taxes

Now we consider a mechanism design approach in which we assume that agents' realized type τ is private information at $t = 1$. This is a natural assumption for the risk tolerance parameter τ, and a reasonable justification for the assumed of time $t = 0$ insurance at of the realization of time $t = 1$ value of τ in previous sections, i.e., a reasonable justification for the assumption of incomplete markets. In this section we discuss two cases, first briefly the one in which a mechanism designer is able to control the consumption of an agent, and then we turn in the next section to the main case where the designer must use an investor specific linear sharing rule.

We begin with two simple results regarding incentive compatible mechanisms in which the planner can fully control the consumption of each agent subject to an incentive compatibility constraint. Incentive compatibility is defined as follows. Consider a given allocation of consumption at $t = 2$ contingent on agents' announced type τ' at $t = 1$ and the realized value of y at $t = 2$ denoted by $c(\tau', y)$. In this section we suppress reference to the aggregate shock z realized at $t = 1$. The certainty equivalent consumption obtained by an agent of type τ who announces type τ' at $t = 1$ is given by

$$\mathbb{C}(\tau, \tau') = U_\tau^{-1} \left[\sum_y U_\tau(c(\tau', y))\rho(y) \right]$$

An allocation $\{c(\tau' y)\}$ for all τ, y is *incentive compatible* if

$$\mathbb{C}(\tau, \tau) \geq \mathbb{C}(\tau, \tau') \text{ for all } \tau, \tau'. \tag{61}$$

Lemma 2. *The first best allocation (i.e. the complete market equilibrium) is not incentive compatible.*

This lemma follows directly from the definition of risk tolerance. In the first best allocation, we have all agents receiving the same certainty equivalent consumption

$$\mathbb{C}(\tau, \tau) = \mathbb{C}(\tau', \tau').$$

But if $\tau' > \tau$ and there is any uncertainty in the allocation of consumption to type τ, then we have that the agent with higher risk tolerance obtains a higher certainty equivalent consumption from the allocation assigned to type τ than does that type, i.e.

$$\mathbb{C}(\tau, \tau) < \mathbb{C}(\tau, \tau')$$

But then incentive compatibility requires that

$$\mathbb{C}(\tau, \tau) < \mathbb{C}(\tau', \tau')$$

which is a contradiction.

This lemma highlights the fundamental tension in this economy. Risk sharing requires equating the allocation of certainty equivalent consumption across agents, but incentive compatibility implies that agents with higher risk tolerance must receive higher certainty equivalent consumption.

We now consider what incentive compatible allocations of consumption can be achieved if the planner makes transfers to agents at time $t = 1$ based on their announced risk tolerance and then lets agents trade based on these post-transfer endowments. The allocations implemented in this way are conditionally efficient. We now show that, when τ has a distribution with a density μ, then the equal wealth equilibrium allocation is the *only* incentive compatible allocation among conditionally efficient allocations. That is, the planner cannot improve on this allocation through a mechanism that makes transfers to agents based on their reported risk tolerance and allows agents to engage in trade after these transfers.

Lemma 3. *Assume that there are continuum of types of agents τ, and let $\mu(\tau)$ denote the strictly positive density of agents of type τ. Then the only conditionally efficient allocation that is incentive compatible is the equal wealth equilibrium allocation.*

We note that for this lemma we cannot dispense from τ having a density. For instance in the case of discrete distribution of τ the incentive compatibility constraint in the equal wealth equilibrium will be slack, since each agent will strictly prefer its equilibrium allocation in their budget set. This motivates us to study both the case with two values of τ and the case where τ has a continuous density.

6.1 Optimal Non-linear Tax-Subsidy

In the equilibrium with the Tobin tax, agents' consumption $c(\tau, y; z)$ at $t = 2$ for fixed τ and z is a linear function of realized y. We refer to the slope of consumption in y as the agent's exposure to aggregate risk or his or her shareholding. In the equilibrium with a Tobin tax investors face two prices, a high one for those than buy risky shares and a low price for those that sell risky shares. Recall that the ex-ante welfare losses comes from the redistributive properties of the tax incidence. In this section, we study the problem of optimal non-linear taxation. The optimality is with respect to the ex-ante utility and subject to the incentive compatibility constraint. Our interest is to compare the allocations under the second best mechanism with the ones under the Tobin tax or subsidy.

In the mechanism design problem that we study, the planner offers agents a menu of *linear sharing rules* for aggregate risk. The linear sharing rules have an agent specific exposure to the realization of aggregate risk y (a type of "risky equity") and an agent specific uncontingent transfer (a type of "uncontingent bond"). The incentive compatible mechanisms we study correspond to a menu of uncontingent bonds and risky equity from which the investor must chose only one point in the menu, so that they are not allowed to retrade. These mechanisms can be interpreted as a non-linear tax/subsidy on trade. Below we compare the features of this optimal tax/subsidy with the Tobin tax/subsidy analyzed above.

The case of two values of τ. Consider first the case of an economy with two possible realizations of $\tau \in \{\tau_1, \tau_2\}$ with $\tau_1 < \tau_2$. To simplify the notation, assume that there is only one possible value of z and suppress that in the notation. Then the mechanism design problem is one of choosing a menu of shares $\mathcal{S}(\tau)$ and bonds $\mathcal{B}(\tau)$ with corresponding allocation of consumption at $t = 2$, $c(\tau, y) = \mathcal{B}(\tau) + \mathcal{S}(\tau)y$ such that the implied certainty solves the problem of maximizing ex-ante welfare

$$W = \sum_{\tau} V\left(\mathbb{C}(\tau, \tau)\right)\mu(\tau)$$

subject to the constraints

$$\mathbb{C}(\tau, \tau') = U_{\tau}^{-1}\left[\sum_{y} U_{\tau}(\mathcal{S}(\tau')y + \mathcal{B}(\tau'))\rho(y)\right]$$

for all τ, τ',

$$\sum_{\tau} \mathcal{S}(\tau)\mu(\tau) = 1, \qquad \sum_{\tau} \mathcal{B}(\tau)\mu(\tau) = 0$$

and the incentive constraint that the agent with high risk tolerance does not want to report that he or she has low risk tolerance

$$\mathbb{C}(\tau_2, \tau_2) \geq \mathbb{C}(\tau_2, \tau_1)$$

The result that this is the only one of the two incentive constraints that is binding follows from the single crossing property of agents' indifference curves over shares s and bonds b that can be shown when agents have equicautious HARA preferences. Moreover, given that we know that the solution of this problem without the incentive constraint is not incentive compatible, we have that the incentive constraint must bind as an equality.

The solution to this mechanism design problem has several features in common with the allocation that arises if a small Tobin subsidy is imposed on the equal wealth equilibrium allocation.

First, the solution to this mechanism design problem must offer the risk averse agents higher certainty equivalent consumption than they achieve in the equal wealth equilibrium allocation (denoted with a superindex e), i.e. $\mathbb{C}(\tau_1, \tau_1) > \mathbb{C}^e(\tau_1)$, and vice-versa for the risk tolerant agents $\mathbb{C}(\tau_2, \tau_2) < \mathbb{C}^e(\tau_2)$. This result follows from the observation that the incentive constraint is slack at the equal wealth equilibrium allocation with discrete types since the risk tolerant agents are able to purchase the equilibrium allocation of the risk averse agents but choose note to. Because this constraint is slack in the equilibrium allocation, it is possible to strictly improve ex-ante welfare by transferring bonds from the risk tolerant agents to the risk averse agents and thus bring certainty equivalent consumption closer together. Hence, similar to the case of the Tobin subsidy, the optimal incentive compatible mechanism effects a transfer of certainty equivalent consumption from risk tolerant to risk averse agents via the incidence of the mechanism.

Second, the solution to the mechanism design problem allocates aggregate risk to the risk tolerant agent, i.e., $\mathcal{S}(\tau_2) > \mathcal{S}(\tau_1)$, and bonds $\mathcal{B}(\tau_2) < \mathcal{B}(\tau_1)$ such that the risk tolerant agents are indifferent between these two portfolios, $\mathbb{C}(\tau_2, \tau_2) = \mathbb{C}(\tau_2, \tau_1)$. We refer to this as saying that the incentive compatibility constraint binds.

Third, the solution to this mechanism design problem implies more trade than in the equal wealth equilibrium. The next result analyzes the special case of the case where $U_\tau(\cdot)$ has CARA utility function. In this case any conditionally efficient allocations has $\mathcal{S}(\tau)/\tau = 1/\bar{\tau}$ is independent of τ. Additionally, in the case where τ only takes two values the proposition below shows that the ex-ante welfare of the equal wealth equilibrium can be improved, and that this improvement has a more dispersed allocation of risk exposure, i.e. involves "more trade" than the equal wealth equilibrium.

Proposition 12. *Assume that τ takes two values $\{\tau_1, \tau_2\}$, and that $U_\tau(\cdot)$ has CARA, or equivalently that $\gamma = \infty$. In any allocation that has binding incentive constraint and that is condi-*

tionally efficient, ex-ante welfare can be improved. Then, denoting the allocation that improves ex-ante welfare with hats we have:

$$\frac{\hat{\mathcal{S}}(\tau_1)}{\tau_1} < \frac{\phi^e(\tau_1)}{\tau_1} = \frac{1}{\bar{\bar{\tau}}} = \frac{\phi^e(\tau_2)}{\tau_2} < \frac{\hat{\mathcal{S}}(\tau_2)}{\tau_2}$$

While the previous result just found an improvement in a feasible and incentive compatible allocation, under regularity conditions the same holds for the solution of the mechanism design problem.

Corollary 13. *If, under the same assumptions of Proposition 12, the set of feasible set and incentive compatible allocations is convex, then the second best allocation also has more dispersed risk exposure. Furthermore, a sufficient condition for the convexity of the feasible set is that $\varphi'''(\cdot) \leq 0$ and $\mu_2 \leq \mu_1$, where ϕ is defined in (64), and it is a function solely of the distribution of y.*

Forth, while in general we can just define a non-linear men, in this case we can devise a piecewise linear menu to decentralize the second best allocation. In particular, each agent has a piecewise linear budget set that is a budget set that arises from a fixed fee (in terms of bonds) to enter the market that is specific to buyers and sellers of shares, a Tobin subsidy of the shares purchased by agents of type τ_2 from agents of type τ_1 (so that the buying price is lower than the selling price), with a cap on the subsidy limiting it to the share sales $1 - \mathcal{S}(\tau_1)$ mandated by the mechanism (or equivalently, a tax on further sales of shares by agents of type τ_1 beyond the quantity $1 - \mathcal{S}(\tau_1)$).

The second best allocation can be decentralized using any non-linear budget set for agents that runs through these two portfolios and is everywhere below the indifference curves of both types of agents corresponding to the portfolios assigned to them.[5] Here we consider a decentralization that corresponds to a Tobin subsidy for trade financed by fixed fees for buyers and sellers to enter the market. This will imply separate budget sets for buyers and sellers of shares. As is the case for the equilibrium allocation, we assume that all agents enter the market at $t = 1$ with one share and zero bonds. Assume that agents announce that they wish to participate in the market either as a buyer or a seller. In doing so they pay a fee in terms of riskless bonds leaving them with their single share and bondholdings $\mathcal{B}^b$ or $\mathcal{B}^s$ for buyers and sellers of shares respectively. Buyers then purchase shares for bonds at price D^b and sellers sell shares for bonds at price D_s up to a cap on sales of shares at $1 - \mathcal{S}(\tau_1)$. We say that these budget

[5] In the next section we return to this general case and apply it to when τ has a distribution with a density.

sets decentralize the allocation that solves our mechanism design problem if agents are willing to enter the market either as buyers or sellers of shares corresponding to their realized risk tolerance τ_2 and τ_1 and then trade to the corresponding second-best portfolios $\mathcal{S}(\tau), \mathcal{B}(\tau)$.

We construct the budget set for buyers of shares as follows. Let the buying price for shares D^b correspond to the marginal rate of substitution between shares and bonds for the risk tolerant agent τ_2 at the portfolio $\mathcal{S}(\tau_2), \mathcal{B}(\tau_2)$. Let the entry fee for buyers of shares be set at

$$\mathcal{B}^b = D^b(\mathcal{S}(\tau_2) - 1) + \mathcal{B}(\tau_2)$$

so that the portfolios $\mathcal{S}(\tau_2), \mathcal{B}(\tau_2)$ and $1, \mathcal{B}^b$ both lie on this budget set with share buying price D^b.

We construct the budget set for sellers of shares as follows. Let the selling price for shares D^s correspond to the marginal rate of substitution between shares and bonds for the risk tolerant agent τ_2 at the portfolio assigned to the risk averse agent $\mathcal{S}(\tau_1), \mathcal{B}(\tau_1)$. Let the entry fee for sellers of shares be set at

$$\mathcal{B}^s = D^s(\mathcal{S}(\tau_1) - 1) + \mathcal{B}(\tau_1)$$

so that the portfolios $\mathcal{S}(\tau_1), \mathcal{B}(\tau_1)$ and $1, \mathcal{B}^s$ both lie on this budget set with share selling price D^s. Add to this budget set the constraint that agents cannot sell more that $1 - \mathcal{S}(\tau_1)$ shares.

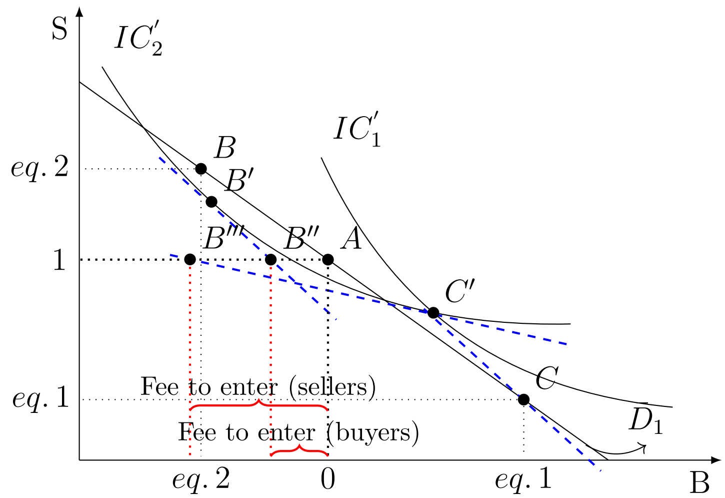
To see that this budget set decentralizes the allocation $\mathcal{S}(\tau), \mathcal{B}(\tau)$, note that the risk tolerant agents τ_2 are, by construction, indifferent between declaring themselves to be buyers or sellers of shares and trading to $\mathcal{S}(\tau_2), \mathcal{B}(\tau_2)$ and $\mathcal{S}(\tau_1), \mathcal{B}(\tau_1)$ respectively. Due to the single crossing property of agents' indifference curves, risk averse agents τ_1 strictly prefer to declare themselves sellers and are bound by the trading limit in trading to $\mathcal{S}(\tau_1), \mathcal{B}(\tau_1)$.

To see that this decentralization involves a Tobin subsidy to trade in the sense that the buying price D^b is lower than the selling price D^s, note that, by construction, both of these prices are determined by tangencies to different points on the same indifference curve for the risk tolerant agent τ_2. Because this agent's indifference curves define strictly convex upper contour sets, the share buying price D^b must be strictly less than the share selling price D^s because this price is tangent to a point with relatively more shares $\mathcal{S}(\tau_2) > \mathcal{S}(\tau_1)$ and fewer bonds $\mathcal{B}(\tau_2) < \mathcal{B}(\tau_1)$. An example of such decentralization can be seen in Figure 3.

This line of reasoning gives us the following proposition.

Proposition 14. *The allocation that solves this mechanism design problem has higher certainty equivalent consumption for the risk averse agents than these agents receive in the equal wealth*

Figure 3: Second Best with two values of τ



Eq.1 and Eq.2 denote the equal wealth equilibrium allocation. A is the initial endowment. B' and C' the second best allocation. The blue broken lines denote the piece-wise linear budget constraint for those declared as buyers and those declared as sellers of risky assets. B'' and B''' denote the point in the budget set of buyers (sellers) of risky asset after paying the entry free.

equilibrium $\mathbb{C}(\tau_1, \tau_1) > \mathbb{C}^e(\tau_1)$ and vice-versa for the risk tolerant agents $\mathbb{C}(\tau_2, \tau_2) < \mathbb{C}^e(\tau_2)$. This second-best allocation can be implemented as an equilibrium in which each agent has a piecewise linear budget set that is a budget set that arises from a fixed fee (in terms of bonds) to enter the market either as a buyer or a seller, a Tobin subsidy of the shares purchased by agents of type τ_2 from agents of type τ_1, with a cap on the subsidy limiting it to the share sales $1 - \mathcal{S}(\tau_1)$ mandated by the mechanism.

The case of a continuum of values of τ. We now consider the version of our economy in which the distribution of agents' types τ has a density $\mu(\tau)$. As we saw in Lemma 3, in this case, the equilibrium allocation is the only incentive compatible conditionally efficient allocation. Thus, a planner does not have room to simply transfer bonds from the risk tolerant to the risk averse agents without violating incentive compatibility as was the case in the economy with a discrete set of possible types. We now analyze this problem. We specialize our analysis to the case of CARA preferences.

The planner assign to each investor an uncontingent transfer $\mathcal{B}$ and a risk exposure to aggregate risk $\mathcal{S}$. There are two physical constraints for the planner. One is that the average risk exposure of investors should be one. The second is that the average uncontingent transfer across investors should be zero:

$$1 = \int \mathcal{S}(\tau)\mu(\tau)d\tau \tag{62}$$

$$0 = \int \mathcal{B}(\tau)\mu(\tau)d\tau \tag{63}$$

When agents have CARA preferences, the certainty equivalent consumption of the investor once the realization of her risk tolerance τ is known to her is given by: $\tau\varphi(s/\tau) + b$. In this case the certainty equivalent is $\tau\varphi(\mathcal{S}/\tau) + \mathcal{B} = -\tau \log \left(\int \exp\left(-[\mathcal{S}\, y/\tau]\right)\rho(y)dy\right) + \mathcal{B}$.

$$\varphi\left(x\right) = -\log E\left[e^{-xy}\right] = -\log \int e^{-xy}\rho(y)dy \tag{64}$$

We summarize the properties of φ as follows

$$\varphi(0) = 0\,,\ \varphi(1) > 0\,,\ \varphi'(x) > 0 \text{ for all } x \text{ if } y \geq 0 \ a.s.\,,\ \varphi'(0) = \mu_y\ , \tag{65}$$

$$\varphi''(x) < 0 \text{ for all } x\,,\ \varphi''(0) = -\sigma_y^2. \tag{66}$$

Normal y case. If $y \sim N(\mu_y, \sigma_y^2)$ then $\varphi\left(x\right) = x\mu_y - \frac{\sigma_y^2}{2}x^2$.

The realization of risk tolerance τ is private information of each investor. Incentive compatibility for an investor with risk tolerance τ is thus:

$$\tau\varphi\left(\frac{\mathcal{S}\left(\tau\right)}{\tau}\right) + \mathcal{B}(\tau) \geq \tau\varphi\left(\frac{\mathcal{S}\left(\tau'\right)}{\tau}\right) + \mathcal{B}(\tau') \ \text{ for all } \tau' \tag{67}$$

The planner wants to maximize ex-ante expected utility, where we assume that investor's take expected utility over their certainty equivalence using utility function V, which we assume to be strictly increasing and strictly concave.. Thus the planner seeks to maximize:

$$\int V\left(\tau\varphi\left(\frac{\mathcal{S}\left(\tau\right)}{\tau}\right) + \mathcal{B}(\tau)\right)\mu(\tau)d\tau \tag{68}$$

by choosing functions $s(\cdot)$ and $b(\cdot)$, subject to the physical constraints (62) and (63) and the incentive compatibility constraint (67) for each τ.

We think that this problem is close, but not identical to the one in Diamond (1998), which itself is a version of Mirrlees (1971) paper with quasilinear utility. For one, preferences are different, although they are quasilinear. Also the physical constraints are different. The constraints (62) is similar, yet not identical to the one in Diamond (1998). The constraint (63) has no analog in Diamond's.

In section (A.1) of the appendix we rewrite the planning problem in ways that it are standard in the literature, first replacing the global incentive compatibility constraint (67) by a local condition and the requirement of monotonicity, second rewriting the feasibility constraint, and third by rewriting the Lagrangian in a way that is easier to analyze similar to a Hamiltonean. In particular in section A.1 we show that solving the planning problem boils down to solving a system of two ordinary differential equations with two known boundary conditions, as well as two integral equations. The system of two ordinary differential equations are the first order conditions for the problem in terms of the path of $\mathcal{C}(\tau)$ and the path for the co-state variable corresponding to the local incentive compatibility constraint at each τ. This system of o.d.e.'s take as given the multipliers for the two feasibility constraints (62-63), which we denote by (θ_s, θ_b). The boundary condition at the lower and upper bound of τ for the costate are just two special cases of the first order conditions of the problem. In Lemma 5 we show that the o.d.e. system given the multipliers, has a unique solution which can be found by a convergent shooting algorithm. The two integral equations that we use to solve for the multipliers correspond to the two feasibility constraints. In lemma 6 we show that if $V(\cdot)$ is also CARA, so that $V(C) = -\tau_V \exp(-C/\tau_V)$ for some constant τ_V we can drop the feasibility constraint (63) and solve the first oder conditions entirely for the path of $x(\tau)$ and the ratio θ_s/θ_b, and analytically find the value of θ_b that impose feasibility of the uncontingent transfers.

The following proposition partially characterize the solution of the planning problem.

Proposition 15. *Assume that $\mu(\tau) > 0$ for all $\tau \in [\tau_L, \tau_H]$. Denote $x(\tau) \equiv \mathcal{S}(\tau)/\tau$. Let $\varphi'(x(\tau))$ be the shadow value of risk, and let θ_b and θ_s the Lagrange multipliers of the constraint (63) and (62) respectively. The shadow value of risk at the top and bottom are the same as the ratio of the Lagrange multipliers, but it is higher for intermediate values:*

$$\frac{\theta_s}{\theta_b} = \varphi'(x(\tau_H)) = \varphi'(x(\tau_L)) < \varphi'(x(\tau)) \text{ for all } \tau \in (\tau_L, \tau_H) \tag{69}$$

and at the extremes we have:

$$\frac{\tau_L}{x(\tau_L)} \frac{dx(\tau)}{d\tau}|_{\tau=\tau_L} = \frac{\theta_b - V'(\mathcal{C}(\tau_L))}{\theta_b} < 0 < \frac{\tau_H}{x(\tau_H)} \frac{dx(\tau)}{d\tau}|_{\tau=\tau_H} = \frac{\theta_b - V'(\mathcal{C}(\tau_H))}{\theta_b} \tag{70}$$

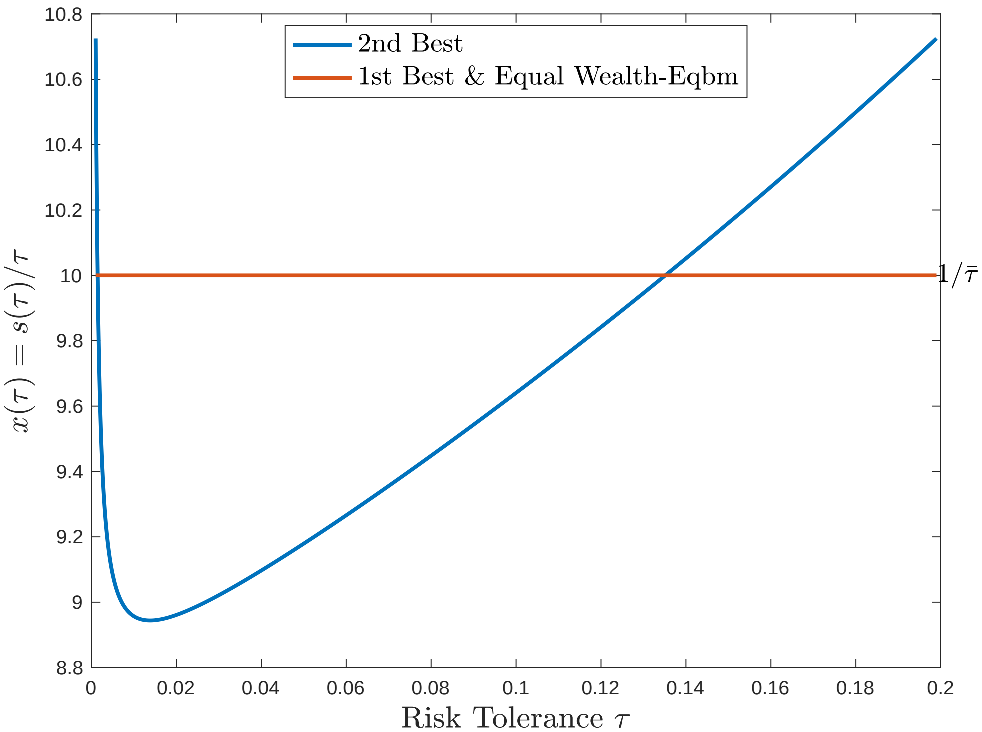
since

$$V'(\mathcal{C}(\tau_H)) < \theta_b = \int_{\tau_L}^{\tau_H} V'(\mathcal{C}(\tau_L))\, \mu(\tau) d\tau < V'(\mathcal{C}(\tau_L)) \,. \tag{71}$$

Moreover, a sufficient condition for $x(\tau)$ to be single peaked in τ is that the derivative of $\tau\mu'(\tau)/\mu(\tau)$ with respect to τ is positive evaluated at τ^* for which $x'(\tau^*) = 0$, i.e. that $\frac{\partial}{\partial \tau} \frac{\tau\mu'(\tau)}{\mu(\tau)}|_{\tau=\tau^*} \geq 0$.

This result is not surprising at all. It is the famous no distortions at the bottom and the top in Mirrlees model. This result, as shown by Seade (1977), requires bounded support for the types, continuous type density, and and interior allocations, which are all conditions satisfied in our setup. For the remaining of types, there is a distortion in the sense that they marginal rate of substitution $\varphi(x)$ is higher, so they must face a higher price.

Figure 4: Risk exposure, normalized relative to risk tolerance



CARA-CARA-Normal-Uniform case. Parameters $\mu_y = 1, \sigma = 0.07, \bar{\tau} = 1/10,\ \tau_L = 0.001, \tau_H = 0.199, \tau_V = \bar{\tau}$.

Figure 4 illustrates Proposition 15 by plotting $x(\tau) = \mathcal{S}(\tau)/\tau$ for the second best case and for the equilibrium with equal wealth (or any other conditionally efficient allocation). As stated in the proposition, x is U-shaped with the same values at both extremes. This is done for a particular numerical example where y is normally distributed and where τ is uniformly distributed –see parameter values are indicated in the note of the plot. With these values the price of risky asset relative to risk-less is $\bar{D}_1 = 0.951$, a 5% premiun.

We consider a decentralization of the optimal non-linear tax by the planner with a menu of contracts, letting the investor decide *only one* on that set. We let $\mathcal{M} = \{(S, B)\}$ be the menu of contracts offer to investors. Each point on the frontier of this set correspond to the values of $B = \mathcal{B}(\tau)$ and $S = \mathcal{S}(\tau)$ for some $\tau \in [\tau_L, \tau_H]$, where the functions $\mathcal{S}(\cdot), \mathcal{B}(\cdot)$ are the solution of the planning problem. Let $\mathbb{S} = \{S : S = \mathcal{S}(\tau) \text{ for some } \tau \in [\tau_L, \tau_H]\}$.

It is interesting to compare the slope of the frontier of $\mathcal{M}$ –given our knowledge of $\varphi'(x(\tau))$ with the slope of the budget line in the equal wealth equilibrium, as well as the slope of the locus of exposure to aggregate risk and transfer in the complete market equilibrium. In the last two cases the slopes are constant. The slope of the budget line in the equal wealth equilibrium is $dB^e/dS = -\bar{D}_1$. The slope on the first best or complete market is $dB^*/dS = -\bar{C}_1$. Now consider any allocation defined by functions $\tilde{x} : [\tau_L, \tau_H] \to \mathbb{R}$ and $\hat{\mathcal{B}} : [\tau_L, \tau_H] \to \mathbb{R}$ that are incentive compatible (IC) and feasible – in the sense that (63) and (62) hold. Define $\hat{\mathcal{M}}$ as the menu of contracts that decentralize the allocation $\hat{x}, \hat{\mathcal{B}}$. Note that $\hat{\mathcal{S}}(\tau) \equiv \hat{x}(\tau)\tau$. We have the following simple result:

Lemma 4. *The frontier of $\hat{\mathcal{M}}$ can be described as follows. Define $\bar{S}(B) \equiv \max B$ such that $(S, B) \in \hat{\mathcal{M}}$. The slope of the frontier for $\hat{\mathcal{M}}$ is given by:*

$$-\frac{d\bar{B}(\hat{\mathcal{S}}(\tau))}{dS} = \varphi'(\hat{x}(\tau)) \text{ for all } \tau \in [\tau_L, \tau_H] \tag{72}$$

$$-\frac{dB^e}{dS} = \varphi'(1/\bar{\tau}) < -\frac{dB^*}{dS} = \bar{\tau}\varphi(1/\bar{\tau}) \tag{73}$$

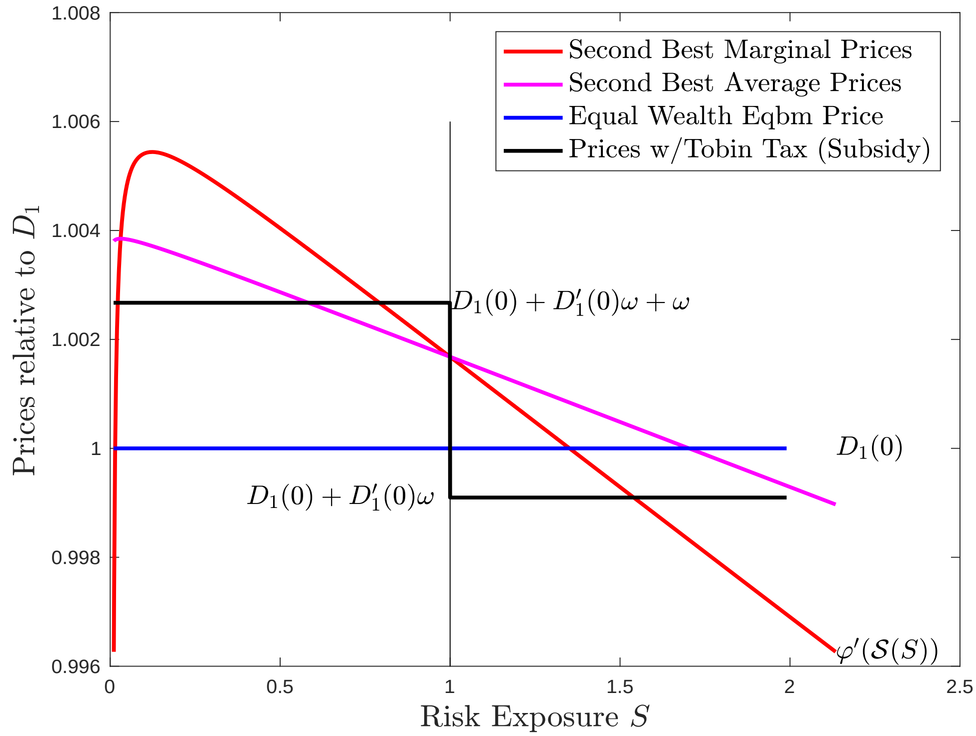
Normal y case. In the case in which $y \sim N(\mu_y, \sigma_y^2)$, then $\varphi'(x) = \mu_y - \sigma_y^2 x$ is linear. In this case we have that feasible and incentive compatible allocation $\hat{x}(\tau)$ must satisfy:

$$\int_{\tau_L}^{\tau_H} \frac{d\bar{B}(\tau\hat{x}(\tau))}{dS}\tilde{\mu}(\tau)d\tau = -\bar{D}_1 \text{ where } \tilde{\mu}(\tau) \equiv \frac{\tau\mu(\tau)}{\int_{\tau_L}^{\tau_H}\tau\mu(\tau)d\tau} \text{ for all } \tau \in [\tau_L, \tau_H] \tag{74}$$

So, in the CARA-Normal case, IC and feasibility implies the *weighted average of risk prices* should be the same as the one in the equal wealth incomplete market equilibrium. Since the optimal one is certainly feasible and incentive compatible, then, by combining (74) with the result from Proposition 15 for the optimal non-linear tax we obtain that the following pattern.

We illustrate the nature of the optimal non-linear tax as well as its differences with a Tobin tax/subsidy in Figure 5. We plot the value of $\varphi'\left(\hat{\mathcal{S}}\right)$ for different cases. In the case of the first best and of the equal wealth incomplete market equilibrium this price is equal to $\varphi'(1/\bar{\tau})$. In the Second Best we plot both the marginal price $\varphi'\left(\hat{\mathcal{S}}\right)$, as well as the average price, by integrating

Figure 5: Marginal & Avg. Prices in the 2^{nd} Best, and Prices in Equal Wealth Eqbm

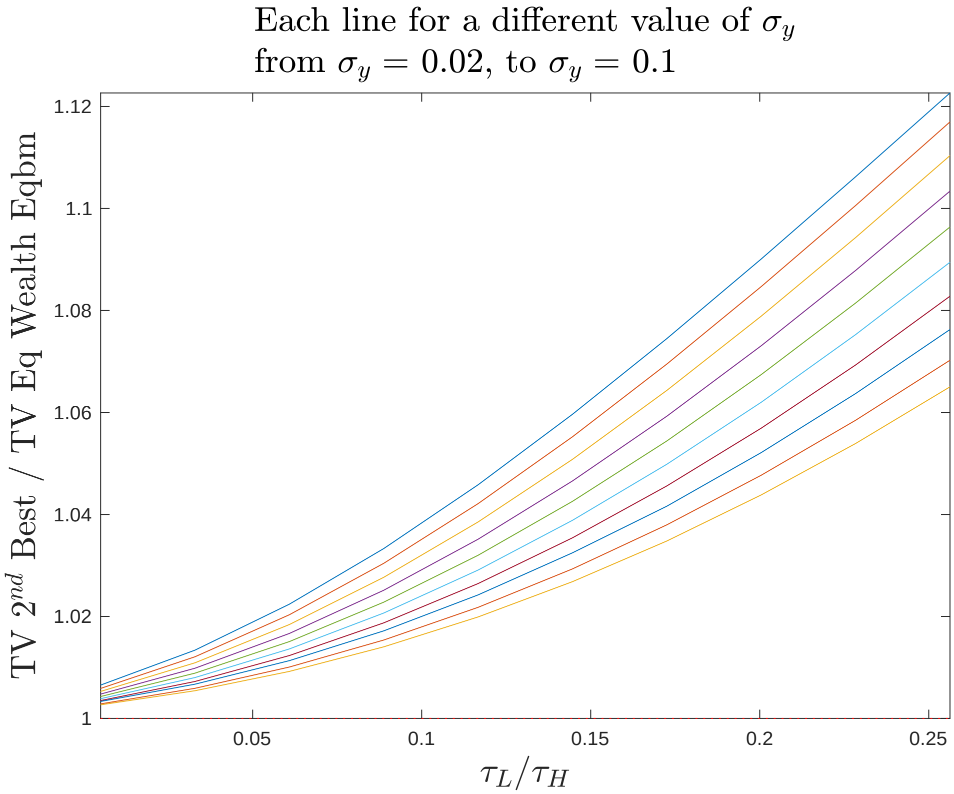


CARA-CARA-Normal-Uniform case. Parameters $\mu_y = 1, \sigma = 0.07, \bar{\tau} = 1/10,\ \tau_L = 0.001, \tau_H = 0.199, \tau_V = \bar{\tau},\ (\omega/D_1(0)) \times 100 = -0.358(36 basispoints)$, $TV^e = 0.2475$, $TV^{2^{nd}\,Best} = 0.2645$.

the marginal price from $\mathcal{S}$ to 1. We also plot the prices that correspond to the Tobin subsidy. These prices take two values, one for those buying and one for those selling risky assets. Since it is a Tobin subsidy, the price for those that sell –so they end up with $S < 1$– is higher than those that buy –so they end up with $S > 1$. The difference between the two prices is ω. We chose the Tobin tax so that it has the same transfer as the value of $\mathcal{B}(1)$ in the second best. This value has the interpretation of the tax or subsidy that goes to those that are assigned $S = 1$. The value of the prices for buyers and sellers are computed using Proposition 8. In this case it is tax, consistent with the idea of a Tobin subsidy to trade, finance with an entrance fee. Note that the pattern for the average prices in the second best "resemble" the prices for a Tobin subsidy.

Finally Figure 6 plots the ratio of the values of $E\left[|S-1|\right]$ computed for the Second Best

Figure 6: Trade Volume in the 2^{nd} Best relative to Trade Volume in Equal Wealth Equilibrium



CARA-CARA-Normal-Uniform case. Parameters $\mu_y = 1, \sigma_y \in [0.2, 0.1], \bar{\tau}_H = 1/10$, $\tau_L \in [0.001, 0.51]$, $\tau_v = \bar{\tau}$.

allocation relative to the Equal Wealth equilibrium. The ratios are computed for several values of σ_y and the ratio τ_L/τ_H for the same specification as in the previous plots. As can be seen, the trade volume in the second best is between 1% and 12% higher in the Second Best allocation as we compute both allocation for different parameter values.

References

Albuquerque, Rui, Eichenbaum Martin, Victor Xi Luo, and Sergio Rebelo. 2016. "Valuation risk and asset pricing." *The Journal of Finance* 71 (6):2861–2904. 10

Bekaert, Geert, Eric Engstrom, and Steven R. Grenadier. 2010. "Stock and bond returns with Moody Investors." *Journal of Empirical Finance* 17 (5):867 – 894. 9

Ben-Tal, Aharon and Marc Teboulle. 1986. "Expected utility, penalty functions, and duality in stochastic nonlinear programming." *Management Science* 32 (11):1445–1466. 12

Bhandari, Anmol, Jaroslav Borovička, and Paul Ho. 2016. "Identifying ambiguity shocks in business cycle models using survey data." Tech. rep., National Bureau of Economic Research. 9

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini. 2009. "Fight or Flight? Portfolio Rebalancing by Individual Investors*." *The Quarterly Journal of Economics* 124 (1):301. URL +http://dx.doi.org/10.1162/qjec.2009.124.1.301. 7

Campbell, John Y and John H Cochrane. 1999. "By force of habit: A consumption-based explanation of aggregate stock market behavior." *Journal of Political Economy* 107 (2):205–251. 9

Campbell, John Y, Sanford J Grossman, and Jiang Wang. 1993. "Trading volume and serial correlation in stock returns." *The Quarterly Journal of Economics* 108 (4):905–939. 8, 29

Cho, Sungjun. 2014. "What drives stochastic risk aversion?" *International Review of Financial Analysis* 34:44 – 63. 9

Constantinides, George M and Darrell Duffie. 1996. "Asset pricing with heterogeneous consumers." *Journal of Political Economy* 104 (2):219–240. 2, 5, 27, 30, 35

Diamond, Peter A. 1998. "Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates." *American Economic Review* :83–95. 52

Dow, James and Gary Gorton. 2006. "Noise traders." Tech. rep., National Bureau of Economic Research. 7

Drechsler, Itamar. 2013. "Uncertainty, time-varying fear, and asset prices." *The Journal of Finance* 68 (5):1843–1889. 9

Duffie, Darrell, Nicolae Gârleanu, and Lasse Heje Pedersen. 2005. "Over-the-counter markets." *Econometrica* 73 (6):1815–1847. 8

Gordon, Stephen and Pascal St-Amour. 2004. "Asset Returns and State-Dependent Risk Preferences." *Journal of Business & Economic Statistics* 22 (3):241–252. 9

Grant, Simon, Atsushi Kajii, Ben Polak, and Zvi Safra. 2010. "Generalized utilitarianism and Harsanyi's impartial observer theorem." *Econometrica* 78 (6):1939–1971. 4

Guo, Hui, Zijun Wang, and Jian Yang. 2013. "Time?Varying RiskReturn Trade?off in the Stock Market." *Journal of Money, Credit and Banking* 45 (4):623–650. 9

Kim, Kun Ho. 2014. "Counter-cyclical risk aversion." *Journal of Empirical Finance* 29:384 – 401. 9

Kozak, Serhiy. 2015. "Dynamics of bond and stock returns." Paper 1277, Ross School of Business. 9

Kreps, David and Evan L Porteus. 1978. "Temporal resolution of uncertainty and dynamic choice theory." *Econometrica* 46 (1):185–200. 4

Lenel, Moritz. 2016. "Safe Assets, Collateralized Lending and Monetary Policy." Tech. rep., Working paper. 9

Lo, Andrew W and Jiang Wang. 2000. "Trading volume: definitions, data analysis, and implications of portfolio theory." *Review of Financial Studies* 13 (2):257–300. 7

———. 2006. "Trading volume: Implications of an intertemporal capital asset pricing model." *The Journal of Finance* 61 (6):2805–2840. 7

Lucas, Robert E. 1980. "Equilibrium in a pure currency economy." *Economic Inquiry* 18 (2):203–220. 10

Mankiw, N Gregory. 1986. "The equity premium and the concentration of aggregate shocks." *Journal of Financial Economics* 17 (1):211–219. 2, 5, 27, 30, 35

Mirrlees, James A. 1971. "An exploration in the theory of optimum income taxation." *The Review of Economic Studies* 38 (2):175–208. 52

Mongin, Philippe and Marcus Pivato. 2015. "Social preference and social welfare under risk and uncertainty." *Oxford Handbook of Well-Being and Public Policy, Oxford* . 4

Seade, Jesus K. 1977. "On the shape of optimal tax schedules." *Journal of Public Economics* 7 (2):203–235. 54

Selden, Larry. 1978. "A new representation of preferences over" certain x uncertain" consumption pairs: The" ordinal certainty equivalent" hypothesis." *Econometrica: Journal of the Econometric Society* :1045–1060. 4

Shleifer, Andrei and Lawrence H Summers. 1990. "The noise trader approach to finance." *The Journal of Economic Perspectives* 4 (2):19–33. 7

Shleifer, Andrei and Robert W Vishny. 1997. "The limits of arbitrage." *The Journal of Finance* 52 (1):35–55. 7

Steffensen, Mogens. 2011. "Optimal consumption and investment under time-varying relative risk aversion." *Journal of Economic Dynamics and Control* 35 (5):659–667. 9

Uslu, Semih. 2015. "Pricing and liquidity in decentralized asset markets." *Working Paper* . 8

Vayanos, Dimitri and Jiang Wang. 2012. "Liquidity and asset returns under asymmetric information and imperfect competition." *Review of Financial Studies* 25 (5):1339–1365. 7

———. 2013. "Market liquidity: Theory and empirical evidence." In *Handbook of the Economics of Finance*, edited by George Constantinides, Milton Harris, and Rene Stulz, chap. 19. North Holland, Amsterdam. 7

A Appendix

Proof. (of Proposition 2) We omit the subindex z in several of the expressions to simplify the notation. For the same reason we omit the index $\bar{\tau}(z)$ for the function $U_{\bar{\tau}(z)}(\cdot)$. First we show that $\bar{C}_1 > \bar{D}_1$. To show this we construct the function $F(\theta) = \mathbb{E}\left[U\left((1-\theta)y + \theta\bar{C}_1\right)\right]$. Note F is strictly concave in θ since it is the expected value of the composition of a strictly concave function with a linear function. In particular $F''(\theta) = \mathbb{E}\left[U''\left((1-\theta)y+\theta\bar{C}_1\right)\left(\bar{C}_1 - y\right)^2\right] < 0$. Direct computation gives $F(0) = \mathbb{E}\left[U(y)\right]$ and $F(1) = U\left(\bar{C}_1\right)$, and thus by definition of $\bar{C}_1$ we have: $F(0) = F(1)$. Summarizing, F is a strictly concave function that attains the same value at $\theta = 0$ and $\theta = 1$, and hence the maximum of F is attained in $\theta \in (0,1)$. Thus, F must be strictly increasing at $\theta = 0$, or $F'(0) > 0$. Direct computation gives: $F'(0) = \mathbb{E}\left[U'(y)\left(\bar{C}_1 - y\right)\right]$ and using $F'(0) > 0$ we have: $\bar{C}_1\,\mathbb{E}\left[U'(y)\right] > \mathbb{E}\left[U'(y)\,y\right]$ which rearranging and using the definition of $\bar{D}_1$, gives the desired result.

Now we find the expression for the first order expansions. The one for $\bar{C}_1$ is well known. The for for $\bar{C}_1$ can be obtained by differentiating $\bar{C}_1$ with respect to σ^2. We begin by expanding $E[U'(y)] = U'(\bar{y}) + (1/2)U'''(\bar{y})\sigma^2 + o(\sigma^2)$ where $o(\sigma^2)$ denotes expression of order smaller than σ^2. And $E[U'(y)y] = U'(\bar{y})\bar{y} + (1/2)[U'''(\bar{y})\bar{y} + 2U''(\bar{y})]\sigma^2 + o(\sigma^2)$. Using this expansions,$\bar{D}_1$ is given by:

$$\bar{D}_1(\sigma^2) = \frac{U'(\bar{y})\bar{y} + (1/2)[U'''(\bar{y})\bar{y} + 2U''(\bar{y})]\sigma^2 + o(\sigma^2)}{U'(\bar{y}) + (1/2)U'''(\bar{y})\sigma^2 + o(\sigma^2)}\,.$$

Differentiating this expression with respect to σ^2, and evaluating at $\sigma^2 = 0$ we obtain:

$$\begin{aligned}\bar{D}_1'(0) &= \frac{(1/2)[U'''(\bar{y})\bar{y} + 2U''(\bar{y})]U'(\bar{y}) - U'(\bar{y})\bar{y}(1/2)U'''(\bar{y})}{U'(\bar{y})^2} \\ &= \frac{U''(\bar{y})U'(\bar{y})}{U'(\bar{y})^2} = \frac{U''(\bar{y})}{U'(\bar{y})}\,.\end{aligned}$$

Thus a Taylor expansion of $\bar{D}_1$ gives $\bar{D}_1 = \bar{y} + U''(\bar{y})/U'(\bar{y})\sigma^2 + o(\sigma^2)$, which is the desired expression.

Proof. (of Proposition 5) By the assumption that d we have that in both economies, $\mathbb{E}(d|z)/\mathbb{E}(d) = 1$ is independent of z we have that for first economy we have, using (47):

$$\frac{1}{\mathcal{E}_{0,2}(d)} = \sum_z \frac{Q^*(z)\pi(z')}{\sum_{z'} Q^*(z')\pi(z')}\left[\frac{1}{\mathcal{E}_{1,2}(z;d)}\right]$$

and for the second economy we have:

$$\frac{1}{\mathcal{E}_{0,2}(d)} = \sum_z \frac{Q^*(z)L_2(z)\pi(z')}{\sum_{z'} Q^*(z')L_2(z')\pi(z')}\left[\frac{1}{\mathcal{E}_{1,2}(z;d)}\right]$$

For both economies the terms $\mathcal{E}_{1,2}(z;d) \geq 1$ are the same, and equal to the one for a representative agent economy. Using that d has systematic exposure, these excess returns are increasing in market wide risk aversion, and hence decreasing in z, so that its reciprocal is $1/\mathcal{E}_{1,2}(z;d)$ is increasing in z. By assumption $L_2(z) \geq 1$ and decreasing in z. So that the induced distribution $Q^*(z)L_2(z)\pi(z')/[\sum_{z'} Q^*(z')L_2(z')\pi(z')]$ for the second economy is stochastically lower than the distribution $Q^*(z)\pi(z')/[\sum_{z'} Q^*(z')\pi(z')]$ corresponding to the first economy. Hence $1/\mathcal{E}_{0,2}(d)$, is smaller for the second economy than for the first economy, and thus its reciprocal, $\mathcal{E}_{0,2}(d)$, is higher for the second economy than for the first economy. □

Proof. (of Proposition 6) Direct computation $\mathbb{E}_1[\tilde{d}|z]/\mathbb{E}_0[\tilde{d}] = \tilde{e}(z)/\mathbb{E}_0[\tilde{e}(z)]$ and $\mathbb{E}_1[d|z]/\mathbb{E}_0[d] = \tilde{e}(z)/\mathbb{E}_0[\tilde{e}(z)] = e(z)/\mathbb{E}_0[e(z)]$. Using the specification for d and $\tilde{d}$ and that $\rho(y|z) = \bar{\rho}(y)$ we have

$$\mathcal{E}_{1,2}(z;\tilde{d}) = \frac{\tilde{e}(z)\sum_y \delta(y)\bar{\rho}(y)}{\tilde{e}(z)\sum_y \bar{p}(y)\delta(y)\bar{\rho}(y)} = \frac{e(z)\sum_y \delta(y)\bar{\rho}(y)}{e(z)\sum_y \bar{p}(y)\delta(y)\bar{\rho}(y)} = \mathcal{E}_{1,2}(z;d) \equiv \bar{\mathcal{E}}_{1,2}$$

Using that $\rho(y|z) = \bar{\rho}(y)$ we have that $Q^*z(z) = \bar{Q}^*$. Thus we have, sing (47)

$$\begin{aligned}
\frac{1}{\mathcal{E}_{0,2}(\tilde{d})} &= \frac{1}{\bar{\mathcal{E}}_{1,2}} \sum_z \frac{\bar{Q}^* L(z)\pi(z)}{\sum_{z'} \bar{Q}^* L(z')\pi(z')} \frac{\tilde{e}(z)}{\mathbb{E}_0\left[\tilde{e}(z)\right]} = \frac{1}{\bar{\mathcal{E}}_{1,2}} \sum_z \frac{L(z)\pi(z)}{\sum_{z'} L(z')\pi(z')} \frac{\tilde{e}(z)}{\mathbb{E}_0\left[\tilde{e}(z)\right]} \\
&= \frac{1}{\bar{\mathcal{E}}_{1,2}} \mathbb{E}_0\left[\frac{L(z)}{\mathbb{E}_0[L(z)]} \frac{\tilde{e}(z)}{\mathbb{E}_0\left[\tilde{e}(z)\right]}\right] \\
&= \frac{1}{\bar{\mathcal{E}}_{1,2}} \left\{1 + \mathbb{C}ov_0\left[\frac{L(z)}{\mathbb{E}_0[L(z)]}, \frac{\tilde{e}(z)}{\mathbb{E}_0[\tilde{e}(z)]}\right]\right\}
\end{aligned}$$

Likewise:

$$\frac{1}{\mathcal{E}_{0,2}(d)} = \frac{1}{\bar{\mathcal{E}}_{1,2}} \left\{1 + \mathbb{C}ov_0\left[\frac{L(z)}{\mathbb{E}_0[L(z)]}, \frac{e(z)}{\mathbb{E}_0[e(z)]}\right]\right\}$$

Since, by assumption, $\tilde{e}(z)/e(z)$ decreases with z then

$$\mathbb{C}ov_0\left[\frac{L(z)}{\mathbb{E}_0[L(z)]}, \frac{e(z)}{\mathbb{E}_0[e(z)]}\right] \leq \mathbb{C}ov_0\left[\frac{L(z)}{\mathbb{E}_0[L(z)]}, \frac{\tilde{e}(z)}{\mathbb{E}_0[\tilde{e}(z)]}\right]$$

Then $\frac{1}{\mathcal{E}_{0,2}(\tilde{d})} \geq \frac{1}{\mathcal{E}_{0,2}(d)}$ or $\mathcal{E}_{0,2}(\tilde{d}) \leq \mathcal{E}_{0,2}(d)$. □

Proof. (of Proposition 8) To render the notation manageable we suppress the z index all variables, and let $D(\omega) = \bar{D}_1(z;\omega)$. Under the assumption of no marginal investors, then for a small tax ω, then $S(\tau) > 1$ if $\tau > \bar{\tau}$ and $S(\tau) < 1$ if $\tau < \bar{\tau}$. We can differentiate market clearing to obtain:

$$
\begin{aligned}
0 = & \sum_{S(\tau)>1} \left\{ \frac{\partial S(\tau, D, 0)}{\partial D} \left[\frac{dD(0)}{d\omega} + 1 \right] + \frac{\partial S(\tau, D, 0)}{\partial T} TV \right\} \mu(\tau) \\
& + \sum_{\tau<\bar{\tau}} \left\{ \frac{\partial S(\tau, D, 0)}{\partial D} \frac{dD(0)}{d\omega} + \frac{\partial S(\tau, D, 0)}{\partial T} TV \right\} \mu(\tau)
\end{aligned}
$$

Rearranging we have:

$$
0 = \sum_{S(\tau)>1} \frac{\partial S(\tau, D, 0)}{\partial D} \mu(\tau) + \frac{dD(0)}{d\omega} \sum_{\tau} \frac{\partial S(\tau, D, 0)}{\partial D} \mu(\tau) + TV \sum_{\tau} \frac{\partial S(\tau, D, 0)}{\partial T} \mu(\tau)
$$

or

$$
\frac{dD(0)}{d\omega} = \frac{\sum_{S(\tau)>1} \frac{\partial S(\tau, D, 0)}{\partial D} \mu(\tau)}{-\sum_{\tau} \frac{\partial S(\tau, D, 0)}{\partial D} \mu(\tau)} + TV \frac{\sum_{\tau} \frac{\partial S(\tau, D, 0)}{\partial T} \mu(\tau)}{-\sum_{\tau} \frac{\partial S(\tau, D, 0)}{\partial D} \mu(\tau)}
$$

Using the characterization of the partial derivative of $S(\tau)$ with respect to D the previous lemma we have:

$$
\begin{aligned}
\sum_{\tau} \frac{\partial S(\tau, D, 0)}{\partial D} \mu(\tau) &= \sum_{\tau} \phi(\tau)\mu(\tau) \frac{\mathbb{E}\left[U'_{\bar{\tau}}(y)\right]}{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)^2\right]} + \sum_{\tau} \left(S(\tau) - 1\right) \mu(\tau) \frac{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)\right]}{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)^2\right]} \\
&= \frac{\mathbb{E}\left[U'_{\bar{\tau}}(y)\right]}{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)^2\right]}
\end{aligned}
$$

Likewise using the partial derivative of $S(\tau)$ with respect to T in the previous lemma we have:

$$
\sum_{\tau} \frac{\partial S(\tau, D, 0)}{\partial T} \mu(\tau) = -\frac{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)\right]}{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)^2\right]}
$$

and finally:

$$
\begin{aligned}
\sum_{S(\tau)>1} \frac{\partial S(\tau, D, 0)}{\partial D} \mu(\tau) = & \sum_{S(\tau)>1} \phi(\tau)\mu(\tau) \frac{\mathbb{E}\left[U'_{\bar{\tau}}(y)\right]}{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)^2\right]} \\
& + \sum_{S(\tau)>1} \left(S(\tau) - 1\right) \mu(\tau) \frac{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)\right]}{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)^2\right]}
\end{aligned}
$$

or using the expression for TV we have:

$$
\sum_{S(\tau)>1} \frac{\partial S(\tau, D, 0)}{\partial D} \mu(\tau) = \frac{\mathbb{E}\left[U'_{\bar{\tau}}(y)\right]}{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)^2\right]} \left(\sum_{S(\tau)>1} \phi(\tau)\mu(\tau) \right) + TV \frac{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)\right]}{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)^2\right]}
$$

Thus we have:

$$\frac{dD(0)}{d\omega} = \frac{\frac{\mathbb{E}[U'_{\bar{\tau}}(y)]}{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)^2\right]}\left(\sum_{S(\tau)>1}\phi(\tau)\mu(\tau)\right) + TV\frac{\mathbb{E}[U''_{\bar{\tau}}(y)(y-D)]}{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)^2\right]}}{-\frac{\mathbb{E}[U'_{\bar{\tau}}(y)]}{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)^2\right]}} - TV\frac{\frac{\mathbb{E}[U''_{\bar{\tau}}(y)(y-D)]}{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)^2\right]}}{-\frac{\mathbb{E}[U'_{\bar{\tau}}(y)]}{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)^2\right]}}$$

$$= \frac{\frac{\mathbb{E}[U'_{\bar{\tau}}(y)]}{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)^2\right]}\left(\sum_{S(\tau)>1}\phi(\tau)\mu(\tau)\right)}{-\frac{\mathbb{E}[U'_{\bar{\tau}}(y)]}{\mathbb{E}\left[U''_{\bar{\tau}}(y)(y-D)^2\right]}}$$

$$= -\sum_{S(\tau)>1}\phi(\tau)\mu(\tau) = -\sum_{\tau>\bar{\tau}}\phi(\tau)\mu(\tau) = -\sum_{\phi(\tau)>1}\phi(\tau)\mu(\tau) \in (-1\,,\,0)$$

since $\sum_\tau \phi(\tau)\mu(\tau) = 1$ and $\phi(\tau) \geq 0, \mu(\tau) \geq 0$ for all τ.

□

Proof. (of Proposition 10) Again we omit z to render the notation simpler. We want to compute:

$$\frac{d}{d\omega}C_1^e(\tau_1;z) = \frac{\mathbb{E}[U'_{\bar{\tau}}(y)]}{U_{\bar{\tau}}(\bar{C})}\left[-(\phi(\tau_1)-1)\frac{\partial}{\partial\omega}\bar{D}(0) + TV\right]$$

We have from the previous proposition:

$$\frac{\partial}{\partial\omega}\bar{D}(0) = -\mu(\tau_2)\phi(\tau_2) \text{ and } TV = (\phi(\tau_2)-1)\mu(\tau_2)$$

thus:

$$\frac{d}{d\omega}C_1^e(\tau_1;z) = \frac{\mathbb{E}[U'_{\bar{\tau}}(y)]}{U_{\bar{\tau}}(\bar{C})}\left[(\phi(\tau_1)-1)\mu(\tau_2)\phi(\tau_2) + (\phi(\tau_2)-1)\mu(\tau_2)\right]$$

$$= \frac{\mathbb{E}[U'_{\bar{\tau}}(y)]}{U_{\bar{\tau}}(\bar{C})}\mu(\tau_2)\left[\phi(\tau_1)\phi(\tau_2)-1\right]$$

$$= \frac{\mathbb{E}[U'_{\bar{\tau}}(y)]}{U_{\bar{\tau}}(\bar{C})}\phi(\tau_1)(1-\phi(\tau_1))\left[1-\mu(\tau_2)\frac{1+\phi(\tau_1)}{\phi(\tau_1)}\right]$$

so

$$\frac{d}{d\omega}C_1^e(\tau_1;z) = \frac{\mathbb{E}[U'_{\bar{\tau}}(y)]}{U_{\bar{\tau}}(\bar{C})}\phi(\tau_1)(1-\phi(\tau_1))\left[1-\mu(\tau_2)\frac{1+\phi(\tau_1)}{\phi(\tau_1)}\right]$$

and thus:

$$\frac{dW^e}{d\omega} = \beta\sum_z \pi(z)\left[V'(C_1^e(\tau_1;z)) - V'(C_1^e(\tau_2;z))\right]\frac{d}{d\omega}C_1^e(\tau_1;z)\mu(\tau_1;z) \text{ thus}$$

$$\frac{dW^e(0)}{d\omega} = \beta\sum_z \pi(z)\left[V'(C_1^e(\tau_1;z)) - V'(C_1^e(\tau_2;z))\right] \times$$

$$\frac{\mathbb{E}[U'_{\bar{\tau}}(y)|z]}{U_{\bar{\tau}}(\bar{C}_1(z))}\phi(\tau_1;z)(1-\phi(\tau_1;z))\left[1-\mu(\tau_2;z)\frac{1+\phi(\tau_1;z)}{\phi(\tau_1;z)}\right]\mu(\tau_1;z)$$

Since $V'' < 0$ and $C_1^e(\tau_1; z) < C_1^e(\tau_2; z)$ then

$$\frac{dW^e(0;z)}{d\omega} < 0 \iff \mu(\tau_2; z) > \frac{\phi(\tau_1; z)}{1 + \phi(\tau_1; z)} \tag{75}$$

□

Proof of Proposition (11). Again we omit z to render the notation easier to follow. Recall that in an equal wealth equilibrium:

$$\begin{aligned}
C_1^e(\tau) - \bar{C}_1(0) &= (\tau - \bar{\tau})\frac{\bar{C}_1 - \bar{D}_1}{\bar{\tau} + \bar{D}_1/\gamma} \equiv \chi(\tau - \bar{\tau}) \\
\phi^e(\tau) - 1 &= \frac{\tau - \bar{\tau}}{\bar{\tau} + \bar{D}_1/\gamma} \equiv \eta(\tau - \bar{\tau}) \\
TV^e &= \int_{\bar{\tau}}^{\tau_H} (\phi^e(\tau) - 1)\mu(\tau)d\tau = \eta \int_{\bar{\tau}}^{\tau_H} (\tau - \bar{\tau})\mu(\tau)d\tau \\
\bar{D}_1'(0) &= -\int_{\bar{\tau}}^{\tau_H} \phi^e(\tau)\mu(\tau)d\tau = -\left[\eta \int_{\bar{\tau}}^{\tau_H} (\tau - \bar{\tau})\mu(\tau)d\tau + \int_{\bar{\tau}}^{\tau_H} \mu(\tau)d\tau\right]
\end{aligned}$$

where $\eta \equiv 1/(\bar{\tau} + \bar{D}_1/\gamma)$ and $\chi \equiv (\bar{C}_1 - \bar{D}_1)/(\bar{\tau} + \bar{D}_1/\gamma)$. Also, since we assume that V is analytical we can write

$$V'(c) = \sum_{n=0}^{\infty} \frac{V^{n+1}(\bar{C}_1)}{n!} \left(c - \bar{C}_1\right)^n \text{ for any } c$$

where $V^{n+1}\left(\bar{C}_1\right) \equiv \partial^n V\left(\bar{C}_1\right)/\partial c^n$.

Using the expression (53), the ex-ante change on welfare of a small Tobin tax can be written as

$$\begin{aligned}
\frac{d}{d\omega}W^e(0;z) &= J(z)TV^e \int_{\tau_L}^{\tau_H} V'(C_1^e(\tau))\mu(\tau)d\tau \\
&- J(z)\bar{D}_1'(0) \int_{\tau_L}^{\tau_H} V'(C_1^e(\tau))\left(\phi^e(\tau) - 1\right)\mu(\tau)d\tau \\
&- J(z) \int_{\bar{\tau}}^{\tau_H} V'(C_1^e(\tau))\left(\phi^e(\tau) - 1\right)\mu(\tau)d\tau \\
&= J(z)\eta \left[\int_{\bar{\tau}}^{\tau_H} (\tau - \bar{\tau})\mu(\tau)d\tau\right] \left[\sum_{n=0}^{\infty} \chi^n \frac{V^{n+1}(\bar{C}_1)}{n!} \int_{\tau_L}^{\tau_H} (\tau - \bar{\tau})^n \mu(\tau)d\tau\right] \\
&- J(z)\bar{D}_1'(0)\,\eta \left[\sum_{n=0}^{\infty} \chi^n \frac{V^{n+1}(\bar{C}_1)}{n!} \int_{\tau_L}^{\tau_H} (\tau - \bar{\tau})^{n+1}\mu(\tau)d\tau\right] \\
&- J(z)\eta \left[\sum_{n=0}^{\infty} \chi^n \frac{V^{n+1}(\bar{C}_1)}{n!} \int_{\bar{\tau}}^{\tau_H} (\tau - \bar{\tau})^{n+1}\mu(\tau)d\tau\right]
\end{aligned}$$

We can rewrite this expression as:

$$\frac{d}{d\omega}W^e(0;z) = J(z)\eta\chi^n \frac{V^{n+1}(\bar{C}_1)}{n!} \times$$
$$\sum_{n=0}^{\infty}\left\{\left[\int_{\bar{\tau}}^{\tau_H}(\tau-\bar{\tau})\mu(\tau)d\tau\right]\left[\int_{\tau_L}^{\tau_H}(\tau-\bar{\tau})^n\mu(\tau)d\tau\right] - \bar{D}_1'(0)\left[\int_{\tau_L}^{\tau_H}(\tau-\bar{\tau})^{n+1}\mu(\tau)d\tau\right]\right.$$
$$\left. - \left[\int_{\bar{\tau}}^{\tau_H}(\tau-\bar{\tau})^{n+1}\mu(\tau)d\tau\right]\right\}$$

Now we analyze the term for each of the derivatives $V(\bar{C})$. For $n = 0, 2, 4$, we obtain:

$$0 = \eta V^1(\bar{C})\left\{\left[\int_{\bar{\tau}}^{\tau_H}(\tau-\bar{\tau})\mu(\tau)d\tau\right] - \bar{D}_1'(0)\times 0 - \int_{\bar{\tau}}^{\tau_H}(\tau-\bar{\tau})\mu(\tau)d\tau\right\}$$

We split the contribution of the remaining term into those with even and odd order of the derivatives of V. Using the symmetry of μ for these values of n:

$$\left[\int_{\bar{\tau}}^{\tau_H}(\tau-\bar{\tau})\mu(\tau)d\tau\right]\left[\int_{\tau_L}^{\tau_H}(\tau-\bar{\tau})^n\mu(\tau)d\tau\right] - \bar{D}_1'(0)\times 0 - \left[\int_{\bar{\tau}}^{\tau_H}(\tau-\bar{\tau})^{n+1}\mu(\tau)d\tau\right]$$
$$= \frac{1}{2}\left[\int_{\tau_L}^{\tau_H}|\tau-\bar{\tau}|\mu(\tau)d\tau\right]\left[\int_{\tau_L}^{\tau_H}|\tau-\bar{\tau}|^n\mu(\tau)d\tau\right] - \frac{1}{2}\left[\int_{\tau_L}^{\tau_H}|\tau-\bar{\tau}|^{n+1}\mu(\tau)d\tau\right] < 0$$

since $E[xy] = E[x]E[y] + Cov(x,y)$ can be applied to $x = |\tau-\bar{\tau}|$ and $y = |\tau-\bar{\tau}|^n$, which are clearly positively correlated. Finally since this term is multiplied by $V^{n+1}(\bar{C})$, which which for these n is positive by hypothesis, the terms with $n = 2, 4, 6, ...$ have a negative contribution to $\frac{d}{d\omega}W^e(0;z)$.

For $n = 1, 3, 5, ...$ we have, using the symmetry of μ:

$$\left[\int_{\bar{\tau}}^{\tau_H}(\tau-\bar{\tau})\mu(\tau)d\tau\right]\times 0 - D'(0)\left[\int_{\tau_L}^{\tau_H}(\tau-\bar{\tau})^{n+1}\mu(\tau)d\tau\right]\left[\int_{\bar{\tau}}^{\tau_H}(\tau-\bar{\tau})^{n+1}\mu(\tau)d\tau\right]$$
$$= -\bar{D}_1'(0)\int_{\tau_L}^{\tau_H}(\tau-\bar{\tau})^{n+1}\mu(\tau)d\tau - \frac{1}{2}\int_{\tau_L}^{\tau_H}(\tau-\bar{\tau})^{n+1}\mu(\tau)d\tau$$
$$\left[\frac{1}{2}+\eta\int_{\bar{\tau}}^{\tau_H}(\tau-\bar{\tau})\mu(\tau)d\tau\right]\int_{\tau_L}^{\tau_H}(\tau-\bar{\tau})^{n+1}\mu(\tau)d\tau - \frac{1}{2}\int_{\tau_L}^{\tau_H}(\tau-\bar{\tau})^{n+1}\mu(\tau)d\tau$$
$$= \eta\left[\int_{\bar{\tau}}^{\tau_H}(\tau-\bar{\tau})\mu(\tau)d\tau\right]\int_{\tau_L}^{\tau_H}(\tau-\bar{\tau})^{n+1}\mu(\tau)d\tau > 0$$

where we use that symmetry implies that $-\bar{D}_1'(0) = 1/2 + \eta\int_{\bar{\tau}}^{\tau_H}(\tau-\bar{\tau})\mu(\tau)d\tau$. Finally since this term is multiplied by $V^{n+1}(\bar{C})$, which for these n is negative by hypothesis, the terms with $n = 1, 3, 4, ...$ have a negative contribution to $\frac{d}{d\omega}W^e(0;z)$.

The expression for the approximation is obtained by using the term for $n = 1$. □

Proof. (of Lemma 2) Note that with our equicautious HARA preferences, we have that if $\tau > \tau'$ (so that agents of type τ have a higher risk tolerance than agents of type τ', then, for any allocation

$$\mathbb{C}(\tau, \tau') \geq \mathbb{C}(\tau', \tau')$$

with this being a strict inequality if $c(\tau', y)$ is risky in that

$$\mathbb{C}(\tau, \tau') < \sum_y c(\tau', y)\rho(y)$$

The first best allocation has risk so that, for τ, τ' such that $\tau > \tau' : \mathbb{C}(\tau, \tau') > \mathbb{C}(\tau', \tau')$. In the first best allocation: $\mathbb{C}(\tau, \tau) = \mathbb{C}(\tau', \tau')$. Hence $\mathbb{C}(\tau, \tau') > \mathbb{C}(\tau, \tau)$, which violates incentive compatibility. □

Proof. (of Lemma 3) Recall that conditionally efficient allocations $c(\tau, y)$ take the form given in equation (**??**) which we reproduce here as

$$c(\tau, y) = \phi(\tau)y + \gamma(\bar{\tau}\phi(\tau) - \tau)$$

One necessary condition for incentive compatibility is

$$\left.\frac{\partial}{\partial \tau'}\mathbb{C}(\tau, \tau')\right|_{\tau=\tau'} = 0$$

which can be written as

$$\sum_y U'_\tau(c(\tau, y))\left[y + \gamma\bar{\tau}\right]\rho(y)\phi'(\tau) = \sum_y U'_\tau(c(\tau, y))\gamma\rho(y)$$

or

$$\frac{\sum_y U'_\tau(c(\tau, y))\left[\frac{y}{\gamma} + \bar{\tau}\right]\rho(y)}{\sum_y U'_\tau(c(\tau, y))\rho(y)} = \frac{1}{\phi'(\tau)}$$

Using the form of conditionally efficient consumption given above together with the specification of $U_\tau(\cdot)$, we have

$$\frac{1}{\phi'(\tau)} = \frac{\sum_y \left(\frac{y}{\gamma} + \bar{\tau}\right)^{-\gamma}\left(\frac{y}{\gamma} + \bar{\tau}\right)\rho(y)}{\sum_y \left(\frac{y}{\gamma} + \bar{\tau}\right)^{-\gamma}\rho(y)} = \frac{D_1^*}{\gamma} + \bar{\tau}$$

This result together with the requirement that the shares $\int \phi(\tau)\mu(\tau)d\tau = 1$ integrate to one imply that

$$\phi(\tau) = \frac{\frac{D_1^*}{\gamma} + \tau}{\frac{D_1^*}{\gamma} + \bar{\tau}}$$

which is the form for $\phi(\tau)$ given in equation (35) for the equal wealth conditionally efficient allocation. □

Proof. (of Proposition 12) To simplify the notation in this proof we use the subindex 1 and 2 instead of the arguments τ_1 and τ_2 for all variables. We impose the binding IC constraint:

$$\mathcal{B}_2 + \tau_2\varphi(\mathcal{S}_2/\tau_2) = \mathcal{B}_1 + \tau_2\varphi(\mathcal{S}_1/\tau_2)$$

and the feasibility constraint:

$$0 = \mathcal{B}_2\mu_2 + \mathcal{B}_1\mu_1$$
$$1 = \mathcal{S}_2\mu_2 + \mathcal{S}_1\mu_1$$

Using these three constraints we can parametrize the set of allocations that are feasible and have binding IC, which involve four variables $(\mathcal{B}_1, \mathcal{S}_1, \mathcal{B}_2, \mathcal{S}_2)$, as a one dimensional manifold. In particular we write them as a function of $\mathcal{B}_1$.

The Certainty Equivalence for each type is thus:

$$\mathbb{C}_2 = \mathcal{B}_2 + \tau_2\varphi(\mathcal{S}_2/\tau_2)$$
$$\mathbb{C}_1 = \mathcal{B}_1 + \tau_1\varphi(\mathcal{S}_1/\tau_1)$$

We can thus write ex-ante welfare as a function of $\mathcal{B}_1$

$$E[V](\mathcal{B}_1) = V(\mathbb{C}_1)\mu_1 + V(\mathbb{C}_2)\mu_2$$

We show that evaluated at the feasible allocation with binding IC constraint that is conditionally efficient, then $\mathcal{S}_2/\tau_2 = \mathcal{S}_1/\tau_1 = x$ then $\frac{dE[V](\mathcal{B}_1)}{d\mathcal{B}_1} > 0$ and hence a decrease in $\mathcal{S}_1$ and an increase in $\mathcal{S}_2$ improve ex-ante welfare.

We consider the deviations in $(\mathcal{B}_1, \mathcal{S}_2, \mathcal{B}_2, \mathcal{S}_2)$ that are feasible and where the IC constraint hold with equality, i.e.:

$$d\mathcal{B}_2 = -\frac{\mu_1}{\mu_2}d\mathcal{B}_1$$
$$d\mathcal{S}_2 = -\frac{\mu_1}{\mu_2}d\mathcal{S}_1$$
$$d\mathcal{B}_2 + \varphi'(\mathcal{S}_2/\tau_2)d\mathcal{S}_2 = d\mathcal{B}_1 + \varphi'(\mathcal{S}_1/\tau_2)d\mathcal{S}_1$$

Substituting feasibility int the binding IC constraint:

$$
\begin{aligned}
-\frac{\mu_1}{\mu_2}d\mathcal{B}_1 - \varphi'(\mathcal{S}_2/\tau_2)\frac{\mu_1}{\mu_2}d\mathcal{S}_1 &= d\mathcal{B}_1 + \varphi'(\mathcal{S}_1/\tau_2)d\mathcal{S}_1 \\
-\left(\varphi'(\mathcal{S}_2/\tau_2)\frac{\mu_1}{\mu_2} + \varphi'(\mathcal{S}_1/\tau_2)\right)d\mathcal{S}_1 &= d\mathcal{B}_1\left(1+\frac{\mu_1}{\mu_2}\right) \\
-\frac{\left(1+\frac{\mu_1}{\mu_2}\right)}{\left(\varphi'(\mathcal{S}_2/\tau_2)\frac{\mu_1}{\mu_2} + \varphi'(\mathcal{S}_1/\tau_2)\right)}d\mathcal{B}_1 &= d\mathcal{S}_1
\end{aligned}
$$

The change on ex-ante utility is:

$$
\begin{aligned}
dE[V] &= V'(\mathbb{C}_1)d\mathbb{C}_1\mu_1 + V'(\mathbb{C}_2)d\mathbb{C}_2\mu_2 \\
&= V'(\mathbb{C}_1)\left[d\mathcal{B}_1 + \varphi'(\mathcal{S}_1/\tau_1)d\mathcal{S}_1\right]\mu_1 + V'(\mathbb{C}_2)\left[d\mathcal{B}_2 + \varphi'(\mathcal{S}_2/\tau_2)d\mathcal{S}_2\right]\mu_2
\end{aligned}
$$

Substituting feasibility:

$$
\begin{aligned}
dE[V] &= V'(\mathbb{C}_1)\left[d\mathcal{B}_1 + \varphi'(\mathcal{S}_1/\tau_1)d\mathcal{S}_1\right]\mu_1 - V'(\mathbb{C}_2)\left[d\mathcal{B}_1 + \varphi'(\mathcal{S}_2/\tau_1)d\mathcal{S}_1\right]\mu_1 \\
&= \left[V'(\mathbb{C}_1) - V'(\mathbb{C}_2)\right]d\mathcal{B}_1\mu_1 + \left[V'(\mathbb{C}_1)\varphi'(\mathcal{S}_1/\tau_1) - V'(\mathbb{C}_2)\varphi'(\mathcal{S}_2/\tau_2)\right]d\mathcal{S}_1\mu_1
\end{aligned}
$$

Substituting IC constraint:

$$
\frac{dE[V]}{\mu_1} = \left[V'(\mathbb{C}_1) - V'(\mathbb{C}_2)\right]d\mathcal{B}_1 - \frac{\left(1+\frac{\mu_1}{\mu_2}\right)\left[V'(\mathbb{C}_1)\varphi'(\mathcal{S}_1/\tau_1) - V'(\mathbb{C}_2)\varphi'(\mathcal{S}_2/\tau_2)\right]}{\left(\varphi'(\mathcal{S}_2/\tau_2)\frac{\mu_1}{\mu_2} + \varphi'(\mathcal{S}_1/\tau_2)\right)}d\mathcal{B}_1
$$

Denoting by $x_1 = \mathcal{S}_1/\tau_1$ and $x_2 = \mathcal{S}_2/\tau_2$ we have:

$$
\frac{1}{\mu_1}\frac{dE[V]}{d\mathcal{B}_1} = \left[V'(\mathbb{C}_1) - V'(\mathbb{C}_2)\right] - \frac{\left(1+\frac{\mu_1}{\mu_2}\right)\left[V'(\mathbb{C}_1)\varphi'(x_1) - V'(\mathbb{C}_2)\varphi'(x_2)\right]}{\left(\varphi'(x_2)\frac{\mu_1}{\mu_2} + \varphi'(x_1\tau_1/\tau_2)\right)}
$$

Note that if $x_1 = x_2 = x$ we have:

$$
\frac{1}{\mu_1}\frac{dE[V]}{d\mathcal{B}_1} = \left[V'(\mathbb{C}_1) - V'(\mathbb{C}_2)\right] - \frac{\left(1+\frac{\mu_1}{\mu_2}\right)}{\left(\frac{\mu_1}{\mu_2} + \frac{\varphi'(x\tau_1/\tau_2)}{\varphi'(x)}\right)}\left[V'(\mathbb{C}_1) - V'(\mathbb{C}_2)\right] > 0
$$

since $\mathbb{C}_1 < \mathbb{C}_2$ and hence $V'(\mathbb{C}_1) > V'(\mathbb{C}_2)$. Using the binding IC and feasibility we have, starting at $x_2 = x_1 = x$ –which characterize the constraint efficient allocations, then ex-ante welfare can be improved by increasing $\mathcal{B}_1$ and since

$$
d\mathcal{S}_1 = -\frac{\left(1+\frac{\mu_1}{\mu_2}\right)}{\left(\varphi'(\mathcal{S}_2/\tau_2)\frac{\mu_1}{\mu_2} + \varphi'(\mathcal{S}_1/\tau_2)\right)} < 0
$$

by decreasing $\mathcal{S}_1$. □

Proof. (of Corollary 13) Now we show that if $\mu_2/\mu_1 \leq 1$ and $\phi''' \leq 0$ then in the optimal allocation $x_2 > 1/\bar{\tau} > x_1$. To prove this we show that the set of feasible allocations is convex. If this set is convex, then $E[V](\mathcal{B}_1)$ must be concave, since V is a concave function. Hence if $\partial E[V](\mathcal{B}_1)/\partial \mathcal{B}_1 > 0$ evaluated at $x_1 = x_2 = 1/\bar{\tau}$, then the optimal must have $\mathcal{B}_1$ larger than that amount, and thus the xs must be more dispersed. To establish the convexity of the feasible set, note that the feasibility constraints are linear, so that the they define, as inequalities a convex set. The remaining constraint is incentive compatibility, which can be written as:

$$G(\mathcal{B}_2, \mathcal{B}_1, \mathcal{S}_2, \mathcal{S}_1) \geq 0 \text{ where}$$
$$G(\mathcal{B}_2, \mathcal{B}_1, \mathcal{S}_2, \mathcal{S}_1) \equiv \mathcal{B}_2 + \tau_2 \varphi(\mathcal{S}_2/\tau_2) - \mathcal{B}_1 - \tau_2 \varphi(\mathcal{S}_1/\tau_2)$$

If G is a concave function, then the set of values for which $G \geq 0$ is convex. Since G is linear in $\mathcal{B}_2$ and $\mathcal{B}_1$ it suffices to show that it is concave in $\mathcal{S}_1, \mathcal{S}_2$. We substitute the feasibility constraint for S in G obtaining:

$$G(\mathcal{B}_2, \mathcal{B}_1, \mathcal{S}_2, (1-\mathcal{S}_2\mu_2)/\mu_1) \equiv \mathcal{B}_2 + \tau_2 \varphi\left(\frac{\mathcal{S}_2}{\tau_2}\right) - \mathcal{B}_1 - \tau_2 \varphi\left(\frac{(1-\mathcal{S}_2\mu_2)/\mu_1}{\tau_2}\right)$$

Thus G is concave if and only $\frac{d^2}{d\mathcal{S}_2^2} G \leq 0$. Direct computation gives:

$$\frac{d}{d\mathcal{S}_2} G(\mathcal{B}_2, \mathcal{B}_1, \mathcal{S}_2, (1-\mathcal{S}_2\mu_2)/\mu_1) \equiv \varphi'\left(\frac{\mathcal{S}_2}{\tau_2}\right) + \varphi'\left(\frac{(1-\mathcal{S}_2\mu_2)/\mu_1}{\tau_2}\right)\frac{\mu_2}{\mu_1}$$
$$\frac{d^2}{d\mathcal{S}_2^2} G(\mathcal{B}_2, \mathcal{B}_1, \mathcal{S}_2, (1-\mathcal{S}_2\mu_2)/\mu_1) \equiv \frac{1}{\tau_2}\left[\varphi''\left(\frac{\mathcal{S}_2}{\tau_2}\right) - \varphi''\left(\frac{(1-\mathcal{S}_2\mu_2)/\mu_1}{\tau_2}\right)\left(\frac{\mu_2}{\mu_1}\right)^2\right]$$

Since $\varphi'' \leq 0$ and we have assumed that $\frac{\mu_2}{\mu_1} \leq 1$:

$$\begin{aligned}\frac{d^2}{d\mathcal{S}_2^2} G(\mathcal{B}_2, \mathcal{B}_1, \mathcal{S}_2, (1-\mathcal{S}_2\mu_2)/\mu_1) &= \frac{1}{\tau_2}\left[\varphi''\left(\frac{\mathcal{S}_2}{\tau_2}\right) - \varphi''\left(\frac{\mathcal{S}_1}{\tau_2}\right)\left(\frac{\mu_2}{\mu_1}\right)^2\right] \\ &\leq \frac{1}{\tau_2}\left[\varphi''\left(\frac{\mathcal{S}_2}{\tau_2}\right) - \varphi''\left(\frac{\mathcal{S}_1}{\tau_2}\right)\right] \leq 0\end{aligned}$$

where the last inequality follows by $\varphi''' \leq 0$ and if $\mathcal{S}_2 \geq \mathcal{S}_1$.

The last step is to show that $\mathcal{S}_2 \geq \mathcal{S}_1$. For this, as a contradiction, assume that $\mathcal{S}_2 < \mathcal{S}_1$. I this case note that:

$$\begin{aligned}\frac{1}{\mu_1}\frac{dE[V]}{d\mathcal{B}_1} &= [V'(\mathbb{C}_1) - V'(\mathbb{C}_2)] - \frac{\left(1+\frac{\mu_1}{\mu_2}\right)[V'(\mathbb{C}_1)\varphi'(\mathcal{S}_1/\tau_1) - V'(\mathbb{C}_2)\varphi'(\mathcal{S}_2/\tau_2)]}{\left(\varphi'(\mathcal{S}_2/\tau_2)\frac{\mu_1}{\mu_2} + \varphi'(\mathcal{S}_1/\tau_2)\right)} \\ &\geq [V'(\mathbb{C}_1) - V'(\mathbb{C}_2)] - \frac{\left(1+\frac{\mu_1}{\mu_2}\right)[V'(\mathbb{C}_1)\varphi'(\mathcal{S}_1/\tau_2) - V'(\mathbb{C}_2)\varphi'(\mathcal{S}_2/\tau_2)]}{\left(\varphi'(\mathcal{S}_2/\tau_2)\frac{\mu_1}{\mu_2} + \varphi'(\mathcal{S}_1/\tau_2)\right)}\end{aligned}$$

where the inequality follows from the concavity of φ and form $\tau_2 > \tau_1$. Rearranging the right hand side we can write:

$$\begin{aligned}\frac{1}{\mu_1}\frac{dE[V]}{d\mathcal{B}_1} &\geq V'(\mathbb{C}_1)\left[1-\frac{\left(1+\frac{\mu_1}{\mu_2}\right)\varphi'(\mathcal{S}_1/\tau_2)}{\varphi'(\mathcal{S}_2/\tau_2)\frac{\mu_1}{\mu_2}+\varphi'(\mathcal{S}_1/\tau_2)}\right]-V'(\mathbb{C}_2)\left[1-\frac{\left(1+\frac{\mu_1}{\mu_2}\right)\varphi'(\mathcal{S}_2/\tau_2)}{\varphi'(\mathcal{S}_2/\tau_2)\frac{\mu_1}{\mu_2}+\varphi'(\mathcal{S}_1/\tau_2)}\right]\\ &= V'(\mathbb{C}_1)\frac{\mu_1}{\mu_2}\left[\frac{\varphi'(\mathcal{S}_2/\tau_2)-\varphi'(\mathcal{S}_1/\tau_2)}{\varphi'(\mathcal{S}_2/\tau_2)\frac{\mu_1}{\mu_2}+\varphi'(\mathcal{S}_1/\tau_2)}\right]-V'(\mathbb{C}_2)\left[\frac{\varphi'(\mathcal{S}_1/\tau_2)-\varphi'(\mathcal{S}_2/\tau_2)}{\varphi'(\mathcal{S}_2/\tau_2)\frac{\mu_1}{\mu_2}+\varphi'(\mathcal{S}_1/\tau_2)}\right]\\ &= \left[V'(\mathbb{C}_1)\frac{\mu_1}{\mu_2}+V'(\mathbb{C}_2)\right]\left[\frac{\varphi'(\mathcal{S}_2/\tau_2)-\varphi'(\mathcal{S}_1/\tau_2)}{\varphi'(\mathcal{S}_2/\tau_2)\frac{\mu_1}{\mu_2}+\varphi'(\mathcal{S}_1/\tau_2)}\right]\end{aligned}$$

Thus if $\mathcal{S}_1 > \mathcal{S}_2$ then $\varphi'(\mathcal{S}_2/\tau_2)-\varphi'(\mathcal{S}_1/\tau_2) > 0$ and $\frac{dE[V]}{d\mathcal{B}_1} > 0$, which means that $\mathcal{S}_1 > \mathcal{S}_2$ can't be optimal. □

Examples of φ and Φ. *Normal case.* For the case where y is normal $N(\mu_y, \sigma_y^2)$ we have:

$$\Phi(x) = \varphi(x) - \varphi'(x)\,x = \frac{\sigma_y^2}{2}x^2\,,\ \Phi'(x) = \sigma_y^2 x > 0 \text{ and } \Phi''(x) = \sigma_y^2 > 0\,.$$

Poisson. If y is Poisson with mean μ_y then

$$\varphi(x) = -\mu_y\left(e^{-x}-1\right) \text{ and } \Phi''(x) = \mu_y e^{-x}\left[1-x\right] \tag{76}$$

so $\Phi''(x) > 0$ for $x < 1$ and $\Phi''(x) < 0$ for $x > 1$. Thus in this case, Φ' is not monotone.

Binominal. Suppose y is distributed as the outcome of n trials each with success with probability p. In this case:

$$\varphi(x) = -\log\left(1-p+pe^{-x}\right) \text{ and } \Phi''(x) = \frac{n(1-p)e^x}{\left[(1-p)e^x+p\right]^3}\left[x(1-p)e^x+p(1+x)\right] > 0 \tag{77}$$

In this case Φ' is monotone.

Exponential. Suppose that y is exponential with parameter λ. In this case

$$\varphi(x) = \log\left(\frac{\lambda+x}{\lambda}\right) \text{ and } \Phi''(x) = \frac{\lambda-x}{(\lambda+x)^3} \tag{78}$$

so $\Phi'' > 0$ if $x < \lambda$ and $\Phi'' < 0$ if $x > \lambda$.

A.1 Planning problem with private information, CARA case

To solve the planning problem we: i) rewrite the IC constraint, using its local version, ii) rewrite and characterize the local IC constraint, iii) provide two equivalent version of the problem, and iv) analyze the implications of the solution for the ration of $\mathcal{S}(\tau)/\tau$.

We will analyze the problem imposing only the local IC constraint. First, for given $\mathcal{S}(\cdot), \mathcal{B}(\cdot)$, define the value of reporting τ' for type τ:

$$\mathbb{C}(\tau, \tau') = \tau\varphi\left(\frac{\mathcal{S}(\tau')}{\tau}\right) + \mathcal{B}(\tau')$$

Thus, global IC for type τ can be written as: $\mathbb{C}(\tau, \tau) \geq (\tau, \tau')$ for all τ'. The condition that investors don't gain form a marginal deviation is:

$$\frac{\partial \mathbb{C}(\tau, \tau')}{\partial \tau'}|_{\tau'=\tau} = \varphi'\left(\frac{\mathcal{S}(\tau)}{\tau}\right)\mathcal{S}'(\tau) + \mathcal{B}'(\tau)$$

and defining

$$\mathcal{C}(\tau) = \tau\varphi\left(\frac{\mathcal{S}(\tau)}{\tau}\right) + \mathcal{B}(\tau)$$

we have:

$$\mathcal{C}'(\tau) = \varphi\left(\frac{\mathcal{S}(\tau)}{\tau}\right) - \varphi'\left(\frac{\mathcal{S}(\tau)}{\tau}\right)\frac{\mathcal{S}(\tau)}{\tau} + \varphi'\left(\frac{\mathcal{S}(\tau)}{\tau}\right)\mathcal{S}'(\tau) + \mathcal{B}'(\tau)$$

Thus we can replace the local IC by imposing:

$$\mathcal{C}'(\tau) = \varphi\left(\frac{\mathcal{S}(\tau)}{\tau}\right) - \varphi'\left(\frac{\mathcal{S}(\tau)}{\tau}\right)\frac{\mathcal{S}(\tau)}{\tau} \tag{79}$$

We also require that $\mathcal{C}'(\tau) > 0$. If these two conditions hold for all τ, then the global IC constraint for all τ.

Note that in the IC constraint we just use $\mathcal{S}(\tau)/\tau$ we write the problem in term of this variable. We also introduce a function Φ so we introduce the notation:

$$x(\tau) \equiv \frac{\mathcal{S}(\tau)}{\tau} \text{ and } \Phi(x) \equiv \phi(x) - \phi'(x)x\,. \tag{80}$$

We note the following properties of Φ:

$$\Phi(x) \equiv \varphi(x) - \varphi'(x)x \geq 0 \text{ with } \Phi'(x) = -\varphi''(x)x > 0 \text{ and} \tag{81}$$

$$\Phi''(x) = -\varphi''(x) - \varphi'''(x)x\,,\ \Phi''(0) = \sigma_y^2 > 0\,. \tag{82}$$

The first inequality follows from concavity of φ and $\varphi(0) = 0$. To see this, note that concavity implies that $\varphi(u) \leq \varphi(x) + \varphi'(x)(u - x)$ for all x, u, and using $u = 0$ we obtain the inequality. Note that $\Phi(x) > 0$, which implies that $\mathcal{C}'(\tau) > 0$. The second inequality also follows from concavity of φ.

We impose the constraint on the uncontingent transfers to obtain an expression for $\mathcal{C}_L$ and the path of $x's$. Recall that $\mathcal{B}(\tau) = \mathcal{C}(\tau) - \tau\varphi(x(\tau))$ so that

$$\int_{\tau_L}^{\tau_H} \mathcal{B}(\tau)\mu(\tau)d\tau \equiv \bar{B} = 0 \iff \bar{B} = \int_{\tau_L}^{\tau_H} \mathcal{C}(\tau)\mu(\tau)d\tau - \int_{\tau_L}^{\tau_H} \tau\varphi(x(\tau))\mu(\tau)d\tau = 0$$

Using the IC constraint we can write:

$$\mathcal{C}(\tau) = \mathcal{C}(\tau_L) + \int_{\tau_L}^{\tau} \Phi(x(t))dt \implies \int_{\tau_L}^{\tau_H} \mathcal{C}(\tau)\mu(\tau)d\tau = \mathcal{C}(\tau_L) + \int_{\tau_L}^{\tau_H} \left[\int_{\tau_L}^{\tau} \Phi(x(t))dt\right] \mu(\tau)d\tau \tag{83}$$

Thus we can write the constraint that $\bar{B} = 0$ in terms of the path $\{x(\tau)\}$ and $\mathcal{C}_L$

$$\bar{B} = \mathcal{C}(\tau_L) + \int_{\tau_L}^{\tau_H} \left[\int_{\tau_L}^{\tau} \Phi(x(t))dt\right] \mu(\tau)d\tau - \int_{\tau_L}^{\tau_H} \tau\varphi(x(\tau))\mu(\tau)d\tau = 0 \tag{84}$$

and integrating by parts and rearranging:

$$\bar{B} = \mathcal{C}(\tau_L) + \int_{\tau_L}^{\tau_H} \Phi(x(\tau)) \left[\int_{\tau_L}^{\tau} \mu(t)dt\right] d\tau - \int_{\tau_L}^{\tau_H} \tau\varphi(x(\tau))\mu(\tau)d\tau = 0 \tag{85}$$

The planning problem is thus:

$$\max_{\{x(\cdot),\mathcal{C}(\cdot)\}} \int_{\tau_L}^{\tau_H} V\left(\mathcal{C}(\tau)\right)\mu(\tau)d\tau \tag{86}$$

subject to:

$$\mathcal{C}'(\tau) = \Phi\left(x(\tau)\right) \text{ for all } \tau \in (\tau_L, \tau_H) \tag{87}$$

$$1 = \int_{\tau_L}^{\tau_H} x(\tau)\tau\mu(\tau)d\tau \text{ and} \tag{88}$$

$$0 = \mathcal{C}(\tau_L) + \int_{\tau_L}^{\tau_H} \Phi(x(\tau)) \left[\int_{\tau_L}^{\tau} \mu(t)dt\right] d\tau - \int_{\tau_L}^{\tau_H} \tau\varphi(x(\tau))\mu(\tau)d\tau \tag{89}$$

We can write the Lagrangian, with multiplier $\lambda(\tau)\mu(\tau)$ for the each IC constraint (87), multiplier θ_s for constraint (88), and multiplier θ_b for constraint (89). After using integration by parts the Lagrangian reads:

$$\begin{aligned} L = &\int_{\tau_L}^{\tau_H} V\left(\mathcal{C}(\tau)\right)\mu(\tau)d\tau + \int_{\tau_L}^{\tau_H} \lambda(\tau)\mu(\tau)\Phi(x(\tau))d\tau \\ &+ \int_{\tau_L}^{\tau_H} \mathcal{C}(\tau)\left[\lambda'(\mu)\mu(\tau) + \mu'(\tau)\lambda(\tau)\right]d\tau - \lambda(\tau)\mu(\tau)\mathcal{C}(\tau)|_{\tau_L}^{\tau_H} + \theta_s\left[1 - \int_{\tau_L}^{\tau_H} x(\tau)\tau\mu(\tau)d\tau\right] \\ &+ \theta_b\left[-\mathcal{C}(\tau_L) - \int_{\tau_L}^{\tau_H} \Phi(x(\tau))\left[\int_{\tau_L}^{\tau}\mu(t)dt\right]d\tau + \int_{\tau_L}^{\tau_H} \tau\varphi(x(\tau))\mu(\tau)d\tau\right] \end{aligned}$$

The first order conditions are:

$$\mathcal{C}(\tau): \; -V'\left(\mathcal{C}(\tau)\right)\mu(\tau) = \lambda'(\mu)\mu(\tau) + \mu'(\tau)\lambda(\tau) \text{ for } \tau \in (\tau_L, \tau_H) \tag{90}$$

$$\mathcal{C}(\tau_H): \; \mu(\tau_H)\lambda(\tau_H) = 0 \tag{91}$$

$$\mathcal{C}(\tau_L): \; \mu(\tau_L)\lambda(\tau_L) = \theta_b \tag{92}$$

$$\begin{aligned} x(\tau): \; &\theta_s\,\tau\,\mu(\tau) + \theta_b\,\Phi'(x(\tau))\left[\int_{\tau}^{\tau_H}\mu(t)dt\right] \\ &= \theta_b\,\tau\,\varphi'(x(\tau))\,\mu(\tau) + \Phi'(x(\tau))\mu(\tau)\lambda(\tau) \text{ for } \tau \in (\tau_L, \tau_H] \end{aligned} \tag{93}$$

Proof. (of Proposition 15) Rearranging the first order condition with respect to $x(\tau)$:

$$\frac{\theta_s}{\theta_b} - \varphi'(x(\tau)) = \frac{\Phi'(x(\tau))}{\tau\mu(\tau)}\left[\frac{\mu(\tau)\lambda(\tau)}{\theta_b} - \int_{\tau}^{\tau_H}\mu(t)dt\right] \tag{94}$$

The left hand side gives the different shadow prices, or implicit tax rates faced by agents. The right hand side determines the sign. It is proportional to the different of two functions, namely $\mu(\tau)\lambda(\tau)/\theta_b$ and $\int_{\tau}^{\tau_H}\mu(t)dt$. Both functions start at the value of one at τ_L and decrease to zero as τ increases to τ_H.

Evaluating the first order condition for x at τ_H and τ_L, and assuming that $\mu(\tau_L) > 0$ and $\mu(\tau_H) > 0$ we obtain

$$\frac{\theta_s}{\theta_b} - \varphi'(x(\tau_H)) = \frac{\theta_s}{\theta_b} - \varphi'(x(\tau_L)) \tag{95}$$

Then differentiating the first order condition of x with respect to τ:

$$\begin{aligned}
&\theta_s\left[\mu(\tau) + \tau\mu'(\tau)\right] + \theta_b\,\Phi''(x(\tau))x'(\tau)\left[\int_{\tau}^{\tau_H}\mu(t)dt\right] - \theta_b\,\Phi'(x(\tau))\mu(\tau)\\
&= \theta_b\,\varphi'(x(\tau))\left[\mu(\tau) + \tau\mu'(\tau)\right] + \theta_b\,\tau\,\varphi''(x(\tau))x'(\tau)\,\mu(\tau)\\
&+ \Phi''(x(\tau))x'(\tau)\mu(\tau)\lambda(\tau) - \Phi'(x(\tau))V'\left(\mathcal{C}(\tau)\right)\mu(\tau)
\end{aligned}$$

Rearranging:

$$x'(\tau) = \frac{\left[\varphi'(x(\tau)) - \frac{\theta_s}{\theta_b}\right]\left[\mu(\tau) + \tau\mu'(\tau)\right] + \Phi'(x(\tau))\mu(\tau)\left[1 - \frac{V'(\mathcal{C}(\tau))}{\theta_b}\right]}{\Phi''(x(\tau))\left[\int_{\tau}^{\tau_H}\mu(t)dt - \frac{\mu(\tau)\lambda(\tau)}{\theta_b}\right] - \tau\,\varphi''(x(\tau))\,\mu(\tau)}$$

Evaluating this at the extremes, using the values of $\lambda(\tau)\mu(\tau)$, and that $\Phi'(x) = -\varphi''(x)x$:

$$x'(\tau_L) = \frac{x(\tau_L)\left[\theta_b - V'\left(\mathcal{C}(\tau_L)\right)\right]}{\theta_b\tau_L} \text{ and } x'(\tau_H) = \frac{x(\tau_H)\left[\theta_b - V'\left(\mathcal{C}(\tau_H)\right)\right]}{\theta_b\tau_H}$$

The equality

$$\theta_b = \int_{\tau_L}^{\tau_H} V'\left(\mathcal{C}(\tau)\right)\mu(\tau)d\tau$$

follows by integrating with respect to τ both sides of the first order condition with respect to $\mathcal{C})(\tau)$ at $\tau \in (\tau_L, \tau_H)$ obtaining:

$$-\int_{\tau_L}^{\tau_H} V'\left(\mathcal{C}\left(\tau\right)\right)\mu\left(\tau\right)d\tau = \int_{\tau_L}^{\tau_H}\left[\lambda'(\tau)\mu(\tau) + \mu'(\tau)\right]d\tau = \lambda(\tau_H)\mu(\tau_H) - \lambda(\tau_L)\mu(\tau_L),$$

and evaluating the right hand side using the first oder conditions with respect to $\mathcal{C}(\tau)$ at the two extremes values, i.e. $\tau = \tau_L$ and $\tau = \tau_H$. Using that $\mathcal{C}(\tau)$ is increasing in τ:

$$V'\left(\mathcal{C}(\tau_H)\right) < \theta_b < V'\left(\mathcal{C}(\tau_L)\right)$$

Hence $x'(\tau_H) > 0 > x'(\tau_L)$. To show that $\theta_s/\theta_b - \varphi'(x(\tau)) < 0$ in the interior we analyze the function:

$$\Psi(\tau) \equiv \frac{\mu(\tau)\lambda(\tau)}{\theta_b} - \int_{\tau}^{\tau_H} \mu(t)dt$$

since we can write $\theta_s/\theta_b - \varphi'(x(\tau)) = \Phi'(x(\tau)/[\tau\mu(\tau)]\Psi(\tau)$. Thus we have:

$$\Psi(\tau) = \int_{\tau_L}^{\tau} \mu(t) \left[1 - \frac{V'(\mathcal{C}(t))}{\int_{\tau_L}^{\tau_H} V'(\mathcal{C}(s))\, \mu(\tau(s))ds}\right] dt \tag{96}$$

Note that

$$\Psi(\tau_L) = \Psi(\tau_H) = 0 \text{ and } \Psi'(\tau) = \mu(\tau)\left[1 - \frac{V'(\mathcal{C}(\tau))}{\theta_b}\right]$$

Using that $\mathcal{C}'(\tau) > 0$, and that $V'(\mathcal{C}(\tau_H)) < \theta_b < V'(\mathcal{C}(\tau_L))$, so that $\Psi'(\tau_L) < 0$, $\Psi'(\tau_H) > 0$ and $\Psi'(\tau^*) = 0$ at a unique value of τ for which $V'(\mathcal{C}(\tau^*)) = \theta_b$. Hence it has a unique minimum, and thus $\Psi(\tau) < 0$ for all $\tau \in (\tau_L, \tau_H)$. Since $\Phi'(x)/[\mu(\tau)\tau] > 0$ this gives the result that $\theta_s/\theta_b - \varphi'(x(\tau)) < 0$ in the interior.

Finally we show that $x(\tau)$ is single peaked in τ. For this we show that at τ^* for which $x'(\tau^*) = 0$ we have that $x''(\tau^*) > 0$. To do so we write:

$$\begin{aligned}
x'(\tau) &= \frac{f(\tau)}{g(\tau)} \text{ so that } x''(\tau^*) = \frac{f'(\tau^*)}{g(\tau^*)} \text{ where} \\
f(\tau) &= \left[\varphi'(x(\tau)) - \frac{\theta_s}{\theta_b}\right]\left[1 + \frac{\tau\mu'(\tau)}{\mu(\tau)}\right] + \Phi'(x(\tau))\left[1 - \frac{V'(\mathcal{C}(\tau))}{\theta_b}\right] \\
g(\tau) &= \frac{\Phi''(x(\tau))}{\mu(\tau)}\left[\int_{\tau}^{\tau_H} \mu(t)dt - \frac{\mu(\tau)\lambda(\tau)}{\theta_b}\right] - \tau\,\varphi''(x(\tau))
\end{aligned}$$

where we have used that $x'(\tau^*) = 0$ Since $\int_{\tau}^{\tau_H} \mu(t)dt - \frac{\mu(\tau)\lambda(\tau)}{\theta_b} > 0$, $\Phi''(x(\tau)) > 0$ and $\varphi''(x(\tau)) < 0$ we have that $sign\,(x''(\tau^*)) = sign\,(f'(\tau^*))$. Direct computation gives:

$$f'(\tau^*) = \left[\varphi'(x(\tau^*)) - \frac{\theta_s}{\theta_b}\right]\left[\frac{\partial}{\partial\tau}\frac{\tau\mu'(\tau)}{\mu(\tau)}\right]_{\tau=\tau^*} - \Phi'(x(\tau^*))\frac{V''(\mathcal{C}(\tau^*))}{\theta_b}\Phi(\tau^*)$$

where we have repeatedly used that $x'(\tau^*) = 0$. Since we have shown that $\varphi'(x(\tau^*)) - \frac{\theta_s}{\theta_b} > 0$, and we have that $\Phi(\tau^*) > 0$, $\Phi'(\tau^*) > 0$, and $V''(\mathcal{C}(\tau^*)) < 0$, then the assumption that $\frac{\partial}{\partial\tau}\frac{\tau\mu'(\tau)}{\mu(\tau)}|_{\tau=\tau^*} > 0$ implies that $f'(\tau^*) > 0$, hence x achieves a minimum at τ^*, and thus it is single peaked.

□

Structure of solution to planning problem. To solve the planning problem we first take as given (θ_s, θ_b) and convert its first order conditions in the solution of two o.d.e.'s subject to two known boundary conditions. The second step is to solve for the values of (θ_s, θ_b), using implied values for the o.d.e.'s for two integral equations, namely the feasibility conditions. We first turn to the description of the o.d.e. system given (θ_s, θ_b). For this we need to use the first order condition for $x(\tau)$ and solve it as a function of τ and λ. We denote such function as $x = X(\lambda, \tau; \theta_s, \theta_b)$. Then we use the following system for $\tau \in (\tau_L, \tau_H)$:

$$\mathcal{C}'(\tau) = \Phi\left(X\left(\lambda(\tau), \tau; \theta_s, \theta_b\right)\right) \tag{97}$$

$$\lambda'(\tau) = -V'\left(\mathcal{C}(\tau)\right) - \frac{\mu'(\tau)}{\mu(\tau)}\lambda(\tau) \tag{98}$$

with boundary conditions:

$$\lambda(\tau_H) = 0 \text{ and } \lambda(\tau_L) = \frac{\theta_b}{\mu(\tau_L)}\,.$$

To solve this two boundary problem we implement a shooting algorithm. We evaluate the system of (97)-(98) with initial condition $\lambda(\tau_L) = \theta_b/\mu(\tau_L)$ and some guess for the value for $\mathcal{C}(\tau_L)$. Then we check if the resulting value of $\lambda(\tau_H)$ satisfies the boundary condition, namely if $\lambda(\tau_H) = 0$. If not, we change the guess for $\mathcal{C}(\tau_L)$ and repeat the procedure. The next lemma ensures that there exists a unique solution that the this procedure converges by studying the properties of the implied mapping between $\mathcal{C}\tau_L$ and $\lambda(\tau_H)$.

Lemma 5. *Fix two arbitrary values of (θ_s, θ_b). Assume that $V'(\cdot)$ is strictly decreasing, and that there is a value for C_0 such that $V'(C_0) = \theta_b$. Let $\lambda_H(\mathcal{C}_L) = \lambda(\tau_H)$ be the value of the solution of the system of two ordinary difference equations (97)-(98) where $\lambda(\cdot)$ is evaluated at τ_H, taking as initial conditions $\mathcal{C}(\tau_L) = \mathcal{C}_L$ and $\lambda(\tau_L) = \theta_b/\mu(\tau_L)$. There exists a unique value $\mathcal{C}_L^*$ that solves: $\lambda_H(\mathcal{C}_L^*) = 0$. Furthermore, we can find an interval $\left[\underline{\mathcal{C}}_L\,, \bar{\mathcal{C}}_L\right]$ so that $\lambda_H(\underline{\mathcal{C}}_L) > 0 > \lambda_H(\bar{\mathcal{C}}_L)$ and $\frac{\partial}{\partial \mathcal{C}_L}\lambda_H(\mathcal{C}_L) > 0$ for all $\mathcal{C}_L \in [\underline{\mathcal{C}}_L\,, \mathcal{C}_L^*]$.*

Proof. (of lemma 5). First we note that $\lambda_H(\mathcal{C}_L) = 0$ if and only if $\theta_b = \int_{\tau_L}^{\tau_H} V'(\mathcal{C}(\tau, \mathcal{C}_L))\mu(\tau)d\tau$ where we let $\mathcal{C}(\tau, \mathcal{C}_L)$, and for future reference $\lambda(\tau, \mathcal{C}_L)$, the solution of the o.d.e. system with $\mathcal{C}_L = \mathcal{C}(\tau_L)$. Throughout this lemma we keep the initial condition $\lambda(\tau_L) = \theta_b/\mu(\tau_L)$. We proceed in three steps.

Step 1. We show that $\lambda_H(\underline{\mathcal{C}_L}) < 0$. To see this we use that the solution for the first order condition of $x(\tau)$ given λ and τ, which we denote $X(\lambda, \tau; \tau_b\tau_s)$ is bounded from above by $\bar{x}(\tau)$, the solution to:

$$\frac{\theta_s}{\theta_b} - \varphi'(\bar{x}(\tau)) = -\frac{\Phi'(\bar{x}(\tau))}{\tau\mu(\tau)} \int_{\tau}^{\tau_H} \mu(t)dt$$

Given this upper bound we can construct an upper bound for $\mathcal{C}(\tau_H, \mathcal{C}_L)$, namely:

$$\mathcal{C}(\tau_H, \mathcal{C}_L) \leq \mathcal{C}_L + (\tau_H - \tau_L)\bar{\bar{\Phi}} \equiv \mathcal{C}_L + (\tau_H - \tau_L) \max_{\tau} \Phi(\bar{x}(\tau))$$

Thus by setting $\mathcal{C}_L$ small enough $V'(\mathcal{C}(\tau, \mathcal{C}_L)) > \theta_b$, and hence $\theta_b < \int_{\tau_L}^{\tau_H} V'(\mathcal{C}(\tau, \mathcal{C}_L))\mu(\tau)d\tau$.

Step 2. We show that $\lambda_H(\underline{\mathcal{C}_L}) > 0$. Since $\mathcal{C}(\tau, \mathcal{C}_L)$ is increasing in τ, then for $\mathcal{C}(\tau_L)$ large enough $\theta_b > \int_{\tau_L}^{\tau_H} V'(\mathcal{C}(\tau, \mathcal{C}_L))\mu(\tau)d\tau$.

Step 3. We now show that $\lambda_H(\cdot)$ is strictly increasing whenever $\lambda_H < 0$. To see this we totally differentiate the system of o.d.e.'s with respect to $\mathcal{C}_L$:

$$\frac{\partial}{\partial \mathcal{C}_L}\lambda(\tau, \mathcal{C}_L) = -\frac{1}{\mu(\tau)} \left[\int_{\tau_L} V''(\mathcal{C}(t, \mathcal{C}_L)) \frac{\partial}{\partial \mathcal{C}_L}\mathcal{C}(t, \mathcal{C}_L)\mu(t)dt\right]$$

$$\frac{\partial}{\partial \mathcal{C}_L}\mathcal{C}(\tau, \mathcal{C}_L) = 1 + \int_{\tau_L}^{\tau} \frac{\partial \Phi(X(\lambda(t, \mathcal{C}_L), t))}{\partial x} \frac{\partial X(t, \mathcal{C}_L), t)}{\partial \lambda} \frac{\partial}{\partial \mathcal{C}_L}\lambda(t, \mathcal{C}_L)dt$$

for all $\tau \in [\tau_L, \tau_H]$. From the first order conditions we have:

$$\frac{\partial X(\lambda, \tau)}{\partial \lambda} = \frac{\Phi'(x)/(\tau\theta_b)}{\varphi''(x) - \frac{\Phi''(x)}{\tau\mu(\tau)}\Psi(\tau, \lambda)} \text{ where } x = X(\lambda, \tau) \text{ and } \Psi(\tau, \lambda) = \frac{\mu(\tau)\lambda}{\theta_b} - \int_{\tau}^{\tau_H} \mu(t)dt$$

We can also write:

$$\Psi(\tau, \lambda(\tau, \mathcal{C}_L)) = \frac{1}{\theta_b}\int_{\tau_L}^{\tau} [\theta_b - V'(\mathcal{C}(t, \mathcal{C}_L))]\mu(t)dt$$

Thus, if $\theta_b < \int_{\tau_L}^{\tau_H} V'(\mathcal{C}(\tau, \mathcal{C}_L))\mu(\tau)d\tau$ then $\Psi(\tau, \lambda(\tau, \mathcal{C}_L))$ is non-positive, single peaked and hence $\partial X/\partial\lambda > 0$ since $\varphi'' < 0$ and $\Phi'' > 0$. Evaluating the derivatives above for each τ, noticing that it is a recrusive system, we have that $\frac{\partial}{\partial \mathcal{C}_L}\lambda(\tau, \mathcal{C}_L) > 1$ and thus $\lambda_H(\mathcal{C}_L) > 0$. □

The next lemma shows that in the case of CARA utility function V, given a solution for the o.d.e.s –in particular given the path for $\{x(\tau)\}_{\tau_L}^{\tau_H}$– we can analytically solve for the value of θ_b and $C(\tau_L)$, keeping the ratio θ_s/θ_b constant which ensures that the feasibility constraint (88) for uncontingent claims holds.

Lemma 6. *Assume that V is a CARA utility function, i.e. $V(C) = -\tau_V \exp(-C/\tau_V)$ for some $\tau_V > 0$. Assume that the path $\{x(\tau)\}_{\tau_L}^{\tau_H}$ and ratio θ_s/θ_b solve the first order conditions with $\mathcal{C}(\tau_L) = 0$, i.e.*

$$\frac{\theta_s}{\theta_b} - \varphi'(x(\tau)) = \frac{\Phi'(x(\tau))}{\tau\mu(\tau)} \int_{\tau_L}^{\tau} \mu(t) \left[1 - \frac{V'\left(\int_{\tau_L}^{t} \Phi(x(t))dt\right)}{\int_{\tau_L}^{\tau_H} V'\left(\int_{\tau_L}^{t} \Phi(x(s))ds\right)\mu(\tau(s))ds}\right] dt \qquad (99)$$

for all $\tau \in [\tau_L, \tau_H]$, and if the path $\{x(\tau)\}_{\tau_L}^{\tau_H}$ also satisfies the feasibility constraint (88). Then the boundary $\mathcal{C}(\tau_L)$ and θ_b

$$\mathcal{C}(\tau_L) = -\int_{\tau_L}^{\tau_H} \Phi(x(\tau)) \left[\int_{\tau_L}^{\tau} \mu(t)dt\right] d\tau + \int_{\tau_L}^{\tau_H} \tau\varphi(x(\tau))\mu(\tau)d\tau$$
$$\theta_b = e^{-\mathcal{C}(\tau_L)/\tau_V} \int_{\tau_L}^{\tau_H} V'\left(\int_{\tau_L}^{\tau} \Phi(x(t))dt\right) \mu(\tau)d\tau$$

satisfy the feasibility condition (89), and thus give the complete solution to the problem.

Proof. (of lemma 6) The proof is immediate, since with CARA the first order condition (99) is independent of $\mathcal{C}(\tau_L)$. □

Proof. (of lemma 4). We will compute $d\bar{B}(\mathcal{S}(\tau))/dS = \mathcal{B}'(\tau)/\mathcal{S}'(\tau)$. We have $\mathcal{S}(\tau) = \tau x(\tau)$ so $\mathcal{S}'(\tau) = x(\tau) + \tau x'(\tau)$. Likewise we have: $\mathcal{B}(\tau) = \mathcal{C}'(\tau) - \varphi(\tau) - \tau\varphi(x'(\tau))x'(\tau)$. From the IC we have: $\mathcal{C}'(\tau) = \varphi(x(\tau)) - x(\tau)\varphi'(x(\tau))$. Thus combining them:

$$\frac{d\bar{B}(\mathcal{S}(\tau))}{dS} = \frac{\mathcal{B}'(\tau)}{\mathcal{S}'(\tau)} = \frac{-x(\tau)\varphi'(x(\tau)) - \tau\varphi(x'(\tau))x'(\tau)}{x(\tau) + \tau x'(\tau)} . = -\varphi'(x(\tau))$$

For the equal wealth incomplete market we have:

$$S^e = 1 + \frac{\tau - \bar{\tau}}{\bar{\tau} + \bar{D}_1/\gamma} \text{ and } B^e = -[\tau - \bar{\tau}]\frac{\bar{D}_1}{\bar{\tau} + \bar{D}_1/\gamma} \text{ so } \frac{dB^e}{dS} = -\bar{D}_1 = -\varphi(1/\bar{\tau})$$

For the complete market of first best allocation we have:

$$S^* = 1 + \frac{\tau - \bar{\tau}}{\bar{\tau} + \bar{C}_1/\gamma} \text{ and } B^* = -[\tau - \bar{\tau}]\frac{\bar{C}_1}{\bar{\tau} + \bar{C}_1/\gamma} \text{ so } \frac{dB^*}{dS} = -\bar{C}_1 = -\tau\varphi(1/\bar{\tau})$$

where we use that in both the equal wealth and complete market equilibrium x is constant, and hence it must be equal to $1/\bar{\tau}$. Also since the complete market allocation is conditionally efficient, then $x^*(\tau) = 1/\bar{\tau}$ for all τ. In the complete market allocation we have that $\bar{C}_1(\tau) = \tau\varphi(1/\bar{\tau}) + \bar{B}(\tau)$ is the same for all τ. Multiplying by $\mu(\tau)$, integrating it across τ, using that uncontingent transfers have zero expected value across τ's, we have $\bar{C}_1 = \bar{\tau}\varphi(1/\bar{\tau})$. Finally since φ is concave, and $\varphi(0) = 0$, then $\varphi(1/\bar{\tau}) > (1/\bar{\tau})\varphi'(1/\bar{\tau})$ or $\bar{C}_1 = \bar{\tau}\varphi(1/\bar{\tau}) > \varphi'(1/\bar{\tau}) = \bar{D}_1$. □