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

Fernando Alvarez[†] and Andy Atkeson[‡]

Abstract

Grossman, Campbell, and Wang (1993), and Pastor and Stambaugh (2003), among others present evidence that liquidity based on a measure of trading volume behaves as a factor in accounting for expected returns on risky assets. We present a tractable theoretical model where trade volume is a pricing factor, beyond the standard ones. In the model agents experience idiosyncratic shocks to risk aversion and these shocks drive both trading volume and asset returns. Dispersion of the idiosyncratic shocks to risk aversion result in trade, and investors regard these shocks as a risk. Just as is the case in the models of asset pricing with idiosyncratic shocks to income studied by Mankiw (1986) and Constantinedes and Duffie (1996), covariance between shocks to the risk aversion of the average investor and to the dispersion of idiosyncratic shocks to risk aversion result in these risks being priced in the cross section of asset returns. Intuitively, each investor is concerned about the risk that he or she will want to sell risky assets at a time in which the price for such assets is low if he or she experiences a higher than average shock to risk aversion at the same time that the risk aversion of the average investor is high. In this way, our model delivers a simple theoretical foundation for the motivating facts regarding trading volume and asset pricing. We also study the impact of taxes on trading on welfare in the incomplete market case and show that such taxes have a first-order negative impact on ex-ante welfare, i.e. a subsidy on trade improves ex-ante welfare. We compare this tax/subsidy with the optimal non-linear tax/subsidy when we treat individual risk tolerance as private information.

Preliminary and Incomplete

[†]Department of Economics, University of Chicago, NBER.

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

1

1 Introduction

We develop a theoretical model of liquidity risk where we obtain that, as some of the empirical literature suggests, trading volume acts as a pricing factor. Here we are thinking in particular of the work of Pastor and Stambaugh JPE (2003) which builds on the model and findings of Campbell, Grossman, and Wang QJE (1993). Given the importance of trading in our setup, the model is particularly suitable to study the welfare effect of trading costs such as taxes on asset trade.

Specifically, we consider a model in which agents experience both aggregate and idiosyncratic shocks to their risk tolerance. In this model, there is a positive volume of trade in intermediate periods because agents with different shocks to their risk tolerance wish to rebalance their portfolios to reflect their changing attitudes towards risk. Aggregate shocks to risk tolerance result in changes in the market price of risk at intermediate dates as well. We have developed a tractable framework in which we can solve for equilibrium and analyze asset prices both at intermediate dates and ex-ante if agents have equicautious HARA preferences when they trade at intermediate dates. In this framework, we are able to draw direct (mathematical) analogies to the results of Mankiw (1986) and Constantinides and Duffie (1996) on the impact of idiosyncratic and aggregate income shocks on asset prices.

The logic of why idiosyncratic and aggregate shocks to risk tolerance that lead to idiosyncratic and aggregate shocks to agents' desired trades might impact the pricing of assets ex-ante is as follows. The logic of the Arrow-Pratt theorem gives us that a preference shock which reduces an agents' risk tolerance in an environment in which aggregate risk is priced is akin to a negative income shock in the sense that such a shock makes it more costly for that agent to attain any given level of certainty equivalent consumption. A risk tolerant agent is content to bear a large amount of aggregate risk and hence can purchase a portfolio yielding a high level of certainty equivalent consumption at a low price because much of that portfolio is purchased at a discount determined by the aggregate risk premium. In contrast, a risk averse agent is highly averse to bearing aggregate risk and hence must pay full price for a portfolio of safe securities to obtain the same level of certainty equivalent consumption.

To the extent that these shocks are common to all investors, these shocks constitute an aggregate risk that is priced ex-ante, but these common shocks to risk tolerance do not lead to trade in intermediate periods. To the extent that these shocks to risk tolerance are idiosyncratic to individual investors, they constitute an idiosyncratic risk to the marginal utility of certainty

equivalent consumption, and data on trading volumes in intermediate periods constitute a valid empirical proxy for the dispersion in these idiosyncratic shocks. As is the case with idiosyncratic endowment shocks, the question of whether these idiosyncratic shocks to risk tolerance are priced in assets ex-ante depends on whether agents can insure themselves against these idiosyncratic shocks through asset markets, whether agents have precautionary savings motives ex-ante, and whether the dispersion in these shocks is correlated with aggregate shocks to risk tolerance and quantities of aggregate endowment risk.

We also use our model to evaluate the impact on ex-ante welfare of a tax on asset transactions in intermediate periods. Standard welfare analyses of sales 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 when the tax revenue is rebated lump sum. In contrast to this standard result, we find that a transactions tax does have a first order negative impact on ex-ante welfare when agents are not able to insure themselves against their idiosyncratic risk tolerance shocks in asset markets and their realized preferences are of the equicautious HARA class. In this setting, the envelope theorem still holds, so the first order welfare impact of the distortion to trade volumes from 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 when tax revenue is rebated lump sum does have a first order impact on ex-ante welfare. Those agents who experience negative shocks to their risk tolerance and wish to sell risky assets end up worse off from the imposition of the tax because they have relatively inelastic demand for risky assets. Hence standard tax incidence arguments imply that they pay more of the tax net of revenue rebates than do agents who experience positive shocks to their risk tolerance and thus wish to buy risky assets in equilibrium. Since ex-ante, agents have an unmet demand to insure themselves against negative risk tolerance shocks in the initial undistorted equilibrium, a transactions tax has a negative first order impact on welfare because it exacerbates the impact of idiosyncratic preference shocks on equilibrium risk sharing. At the margin, ex-ante welfare would be improved by a subsidy to trade.

Description of the model. We develop a simple three period model with $t = 0, 1, 2$. The three periods are, starting from the end: $t = 2$ where the aggregate endowment is realized and

consumption takes place, $t = 1$ where aggregate and idiosyncratic shocks to risk tolerance are realized and where investors can rebalance their portfolio, and $t = 0$ where assets are priced and initial consumption takes place. There is no consumption at $t = 1$, so all trade in assets corresponds to portfolio rebalancing.

An allocation in this environment is an assignment of consumption to agents in period $t = 0$ and consumption in period $t = 2$ contingent on aggregate and idiosyncratic shocks to agents' risk tolerance at $t = 1$ and the aggregate endowment realized at $t = 2$. We define agents' preferences over allocations recursively. As of period $t = 1$, once the aggregate and idiosyncratic shocks to agents' risk tolerances have been realized, each agent has realized subutility U_τ indexed by their realized type τ that they use to evaluate their expected utility and corresponding certainty equivalent consumption at $t = 1$ from the allocation of consumption at $t = 2$ assigned to their realized type. Agents' ex-ante preferences are then defined as expected utility of certainty equivalent consumption at $t = 0$ and $t = 1$ defined with respect to a common strictly concave utility function V. With this recursive specification of preferences, we can separate the impact of shocks to risk tolerance at $t = 1$ on attitudes towards the intertemporal allocation of consumption between period $t = 0$ and later periods. This recursive definition of preferences also has some grounding in the social choice literature (see, for example, Grant, Kajii, Polak, and Safra 2010) formalizing utilitarian preferences when the prospects of different risk tolerance is interpreted as identity risk. Furthermore in decision theory literature, combining the certainty equivalent with different utility functions gives the representation of cautious expected utility, see Cerreia-Vioglio, Dillenberger, and Ortoleva 2015, which gives a related representation.

Our model is particularly tractable when the subutility function U_τ is of the equicautious HARA class. For this specification, the type τ is a parallel shift in the agents' risk tolerance at $t = 1$ as a function of the level of their consumption. (Recall that risk tolerance is the inverse of the coefficient of absolute risk aversion). Hence, the preference shocks that we consider pure shocks to the level of an agents' risk tolerance.

Given our recursive definition of preferences, a negative shock to risk tolerance at $t = 1$ is analogous to a negative shock to one's endowment in units of certainty equivalent consumption at that date by the logic of the Arrow-Pratt Theorem —- for any stochastic assignment of consumption at $t = 2$, a more risk tolerant agent has higher certainty equivalent consumption at $t = 1$ than does a less risk tolerant agent. If preferences U_τ are of the equicautious HARA class, we can make this analogy between preference shocks and endowment shocks more precise as

these preferences display four properties that make solving for the equilibrium highly tractable.

These properties are as follows. With preferences of the equicautious HARA class, agents' asset demands at $t = 1$ display Gorman Aggregation. That is, we can solve for the prices at $t = 1$ of assets that pay off at $t = 2$ as if the economy had a representative agent with the average realized risk tolerance, and hence these asset prices are impacted only by aggregate shocks to risk tolerance. With this result we can solve directly for the set of feasible allocations of certainty equivalent consumption at $t = 1$ given the realized aggregate shock to risk tolerances and we find that this set has a linear frontier. This finding gives us the result that the socially optimal allocation of certainty equivalent consumption assigns the same certainty equivalent consumption to all agents at both $t = 0$ and $t = 1$ regardless of idiosyncratic shocks to risk tolerance. Hence, in the socially optimal allocation, agents are fully insured against idiosyncratic risk and hence this risk is not priced in assets at $t = 0$. We then consider the equilibrium allocation of certainty equivalent consumption which arises in an economy with incomplete asset markets in which agents can trade assets at $t = 0$ with payoffs contingent on aggregate shocks to risk preferences but not contingent on idiosyncratic shocks to risk preferences. Because agents are ex-ante identical, they do not trade these contingent securities at $t = 0$, and hence the equilibrium allocation of certainty equivalent consumption at $t = 1$ is the feasible allocation of certainty equivalent consumption at $t = 1$ that costs the same for each agent, where the agents' risk tolerance τ and the equilibrium asset prices at $t = 1$ determine the cost to that agent of attaining any given level of certainty equivalent consumption. With preferences of the equicautious HARA class we are able to characterize these cost functions and hence fully characterize the equilibrium allocation of certainty equivalent consumption and the equilibrium asset prices at $t = 0$. Finally, in order to derive the model's implications for trading volumes, we make use of the property that for preferences of this class, a two-fund theorem holds. Thus, we can implement the equilibrium allocation with trade only in shares of the aggregate endowment and risk free bonds. With these results, we are able to make the mathematical mapping between our model and a model with idiosyncratic endowment shocks precise.

We then use our model to explore the relationship between model implied trading volumes and asset prices. One of our central results is the certainty equivalent consumption assigned to a given agent at $t = 1$ in the equilibrium with incomplete markets is equal to the average level of certainty equivalent consumption plus a term that reflects the impact of the idiosyncratic shocks to risk tolerance τ. In equilibrium, this term reflecting idiosyncratic risk is the product

of that agents' equilibrium net trade in shares times a measure of the aggregate risk premium on shares. In this way, the model implies that if one had data on the full distribution of net trades in shares of the aggregate endowment undertaken by each agent and a measure of the aggregate risk premium on those shares, one would have a full description of the distribution of idiosyncratic consumption risk agents' experienced. Data on aggregate trade volumes is moment of this distribution (one-half the mean absolute deviation of net trades), and hence serves as a proxy for the data needed to measure the idiosyncratic consumption risk agents face in different states of nature. The model implies that data on trading volume must be interacted with data on the aggregate risk premium on shares to fully understand the idiosyncratic consumption risk agents face in equilibrium in different states of nature. The basic intuition is that an agent who experiences a large negative shock to his or her risk tolerance finds it very costly in terms of lost certainty equivalent consumption to rebalance his or her portfolio from risky shares to safe bonds if risky shares are trading at a large discount relative to safe bonds. In contrast, the loss in certainty equivalent consumption for this agent is not so large if risky shares are trading at only a small discount relative to safe bonds. We derive several formulas regarding the joint distribution of observed trading volume and aggregate risk premia at $t = 1$ and our model-implied asset prices at $t = 1$. These include formulas that compare aggregate risk premia at $t = 0$ across economies with higher or lower trade volumes and that compare risk premia observed in the cross section of assets at $t = 0$ in a single economy.

We then turn to our analysis of the impact of taxes on share trade at $t = 1$ on ex-ante welfare at $t = 0$. Here, because our model is tractable, we are able to solve for the incidence of the tax net of revenue rebates and establish our result that such a tax has a negative first order impact on ex-ante welfare. We see the approach we take to analyzing the welfare impact of transactions tax in terms of its incidence and hence its impact on the sharing of "liquidity risk" as the main contribution of this part of the paper.

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.

The idea that idiosyncratic preference shocks impact investors' precautionary demand for an asset (in this case money) is central to Lucas (1980). The observation that if agents have CARA preferences in the model of that paper, then the preference shocks in that model are isomorphic to endowment shocks is a clear antecedent to our result that, with our recursive formulation of preferences with equicautious HARA subutility, aggregate and idiosyncratic shocks to risk tolerance are isomorphic to aggregate and idiosyncratic shocks to endowments of certainty equivalent consumption. This equivalence result then allows us to map mathematically the asset pricing implications of shocks to risk tolerance in our framework into the asset pricing implications of endowment shocks studied in Mankiw (1986) and Constantinides and Duffie (1996). We see the difference here as primarily one of mapping models to data. In models in which agents' trade due to heterogeneous endowment shocks, empirical proxies for the risk that agents face correspond to observed income risk and/or trading volumes driven by fluctuations in individuals' savings rates. In our framework, empirical proxies for the risk that agents face corresponds to trading volumes driven by individuals' portfolio rebalancing rather than fluctuations in individuals' savings rates. Perhaps such a framework is more empirically relevant given the extremely high transactions volumes observed in asset markets.

Shocks to hedging needs Vayanos and Wang (2012) and (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. In future work we will explore more closely the extent to which our results hold in a framework in which trade is motivated by non-traded endowment shocks rather than shocks to risk tolerance.

Duffie, Garleanu, 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).

Risk tolerance and external habits. 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 (as in Campbell and Cochrane). In that model, shifts in the external habit parameter shift agents' risk tolerance and, to the extent that this external habit is stochastic, correspond to random shocks to investors' risk tolerance. In that model, shifts in the external habit parameter also impact agents' intertemporal elasticity of substitution. Our recursive definition of preferences isolates the shocks to risk tolerance, leaving intertemporal preferences over the allocation of certainty equivalent consumption unchanged.

The idea that shocks to the demand side for risky assets are important is emphasized by Albuquerque, Eichenbaum, and Rebelo (2015). 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 Three Period Model

Consider a three period economy with $t = 0, 1, 2$ and a continuum of measure one of agents. Agents are all identical at time $t = 0$. There is an aggregate endowment of consumption available at $t = 0$ of $\bar{C}_0$. Agents face uncertainty over the aggregate endowment that is realized at time $t = 2$, denoted by $y \in Y$. To simplify notation, we assume that Y is a finite set.

Agents also face idiosyncratic and aggregate shocks to their preferences. Specifically, at time $t = 0$, agents do not know which type of preferences they will have at time $t = 1$. 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, with Z being a finite set and probabilities 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 $t = 2$ also depends on z, with $\rho(y|z)$ denoting the density of y conditional on z. We denote the conditional mean and variance of the aggregate endowment at $t = 2$ by $\bar{y}(z)$ and $\sigma^2(z)$ respectively.

Allocations: Consumption occurs at $t = 0$ and $t = 2$. 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 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

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

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

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

$$C_1(\tau; z) \equiv U_\tau^{-1} \left(\mathbb{E} \left[U_\tau(c(\tau, y; z))|z \right] \right) = U_\tau^{-1} \left(\sum_y U_\tau(c(\tau, y; z))\rho(y|z) \right) \tag{3}$$

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 are convex, 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)[1], 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.

With this assumption regarding preferences, it is then immediate that the First and Second Welfare Theorems will apply in this environment if we assume asset markets that are complete with respect to both aggregate and idiosyncratic uncertainty.

---

[1]Theorem 1 in "Expected Utility, Penalty Functions, and Duality in Stochastic Nonlinear Programming" by Aharon Ben-Tal and Marc Teboulle, Management Science, Vol. 32, No. 11 (Nov., 1986)

Feasibility of Certainty Equivalent Consumption: In analyzing equilibria in two stages, it will be useful for us to consider the allocation of certainty equivalent consumption at time $t = 1$, $\{C_1(\tau; z)\}$ corresponding to any allocation $\vec{c}(y; z) = \{C_0, c(\tau, y; z)\}$. We say that an allocation of certainty equivalent consumption at $t = 1$, $\{C_1(\tau; z)\}$, is feasible if there exists a feasible allocation $\vec{c}(y; z)$ that delivers that vector of certainty equivalent consumption. 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 equivalence at time $t = 1$ is a concave function of the underlying allocation at $t = 2$ as we have assumed.

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: The interpretation of preference type τ is is that if $\tau > \tau'$, 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
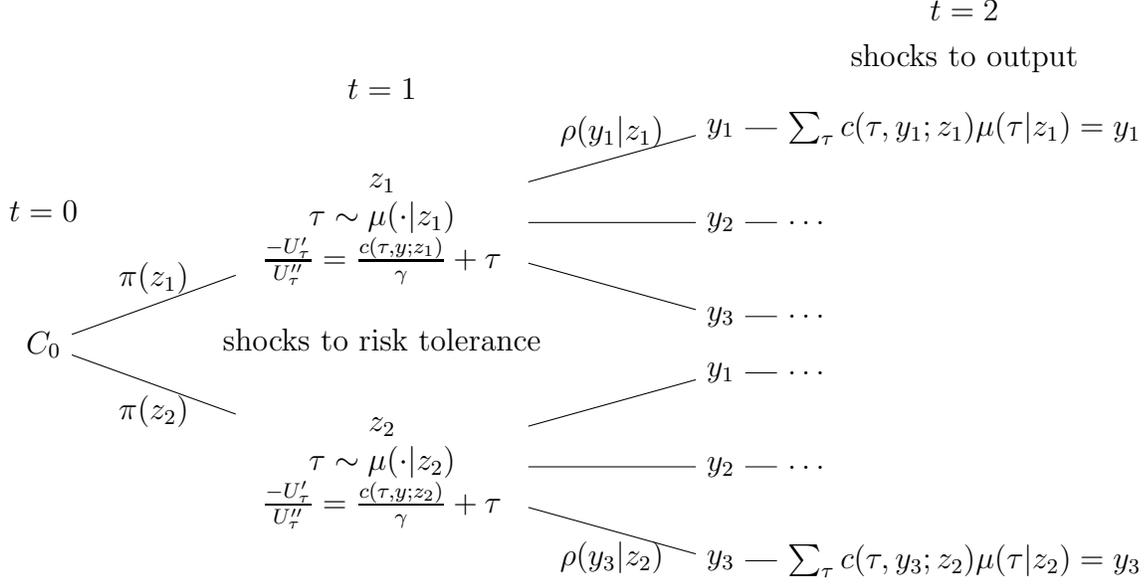
11

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 $y \in \{y_1, y_2, y_3\}$.

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.

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

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

2.2   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 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 is *ex-post efficient* if, given a realization of z at $t = 1$, it solves the problem of maximizing the objective

$$\sum_\tau \lambda_\tau \sum_y U_\tau(c(\tau, y; z)) \rho(y|z) \mu(\tau|z) \tag{10}$$

among feasible allocations given some vector of non-negative Pareto weights λ_τ. Clearly, the socially optimal allocation is also ex-post efficient.

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

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

13

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).

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.

2.3 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

14

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

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

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 endowments.

2.4 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.

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.

We are interested in the dynamics of asset returns from period $t = 0$ to $t = 1$ and from $t = 1$ to $t = 2$ and their relationship with transactions volumes at $t = 1$. The realized return on a security d from $t = 1$ to $t = 2$ given realized y and z is $R_2(y, z; d) = d(y; z)/P_1(z; d)$ and hence the expected return on this security at $t = 1$ given z is

$$\mathbb{E}\left[R_2(y, z; d)|z\right] = \frac{\sum_y d(y; z)\rho(y|z)}{P_1(z; d)} \tag{19}$$

The realized return on this security from $t = 0$ to $t = 1$ given realized z is

$$R_1(z; d) = \frac{P_1(z; d)}{P_0(d)} \text{ and } \mathbb{E}\left[R_1(z; d)\right] = \frac{\mathbb{E}\left[P_1(z; d)\right]}{P_0(d)} \tag{20}$$

To measure risk premium of a security with payoff d, depending on the circumstances, we will work with the multiplicative expected excess return on that security from $t = 0$ to $t = 1$, denoted by $\mathcal{E}_{0,1}(d)$, with the additive expected excess return as follows:

$$\mathcal{E}_{0,1}(d) \equiv \frac{\mathbb{E}[R_1(z, d)]}{\bar{R}_0} \text{ so that } \mathbb{E}\left[R_1(z; d)\right] - \bar{R}_0 \equiv \left[\mathcal{E}_{0,1}(d) - 1\right]\bar{R}_0 \tag{21}$$

As is standard, equation (18) can be used to price asset returns from $t = 0$ to $t = 1$ as

$$\mathcal{E}_{0,1}(d) - 1 = -Cov\left(Q(z), R_1(z; d)\right) \tag{22}$$

We also measure the risk premium of a security with payoff d with the multiplicative expected excess return on that security from $t = 0$ to $t = 2$ measured as the ratio of the cost of purchasing at $t = 0$ a sure claim to the expected dividend of that security at $t = 2$ relative to the price of that security at $t = 0$. We write this measure of the multiplicative expected excess return as

$$\mathcal{E}_{0,2}(d) = \frac{P_0(1)\mathbb{E}(d)}{P_0(d)} = \frac{P_0(1)\left[\sum_z \sum_y d(y; z)\rho(y|z)\pi(z)\right]}{P_0(d)} \tag{23}$$

If we define the multiplicative expected excess return on a security d from $t = 1$ to $t = 2$ conditional on z being realized at $t = 1$ to the be the ratio of the cost of purchasing at $t = 1$ a sure claim to the conditional expectation of the dividend d relative to the price of purchasing the security at $t = 1$, i.e.

$$\mathcal{E}_{1,2}(z, d) = \frac{\mathbb{E}(d|z)}{P_1(z, d)} = \frac{\sum_y d(y; z)\rho(y|z)}{P_1(z, d)}$$

we then have that the inverse of the multiplicative expected excess returns can be written

$$\frac{1}{\mathcal{E}_{0,2}(d)} = \sum_z \pi_Q(z)\frac{\mathbb{E}(d|z)}{\mathbb{E}(d)}\frac{1}{\mathcal{E}_{1,2}(z, d)} \tag{24}$$

17

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 $d(y; z)$

where $\pi_Q(z)$ is the change of measure

$$\pi_Q(z) = \frac{Q(z)\pi(z)}{\sum_z Q(z)\pi(z)} \tag{25}$$

2.5   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.

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

$$\frac{U'_\tau(c(\tau, y_1; z))}{U'_\tau(c(\tau, y_2; z))} = \frac{p(y_1; z)}{p(y_2; z)} \tag{26}$$

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 a given level of certainty equivalent consumption at time $t = 1$ given z as

$$H_\tau(C_1; z) = \min_{c(y; z)} \sum_y p(y; z)c(y; z)\rho(y|z) \tag{27}$$

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

$$H_\tau(C_1^*(\tau; z); z) = D_1^*(z) + B^*(\tau; z) \tag{28}$$

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

$$Q^*(\tau; z) = \beta \frac{V'(C_1^*(\tau; z))}{V'(C_0^*)} \Big/ \frac{\partial}{\partial C_1} H_\tau(C_1^*(\tau; z); z) \tag{29}$$

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

$$H_\tau(C_1^e(\tau; z); z) = D_1^e(z) + B^e(z) \tag{30}$$

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

$$Q^e(z) = \beta \sum_\tau \left[\frac{V'(C_1^e(\tau; z))}{V'(C_0^e)} \Big/ \frac{\partial}{\partial C_1} H_\tau(C_1^e(\tau; z); z) \right] \mu(\tau|z) \tag{31}$$

The Marginal Cost of Certainty Equivalent Consumption: Our asset pricing formulas, (29) and (31) 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 (27) yields that in in the socially optimal allocation, this marginal cost is given by

$$\frac{\partial}{\partial C_1} H_\tau(C_1^*(\tau; z); z) = \frac{U_\tau'(C_1^*(\tau; z))}{\sum_y U_\tau'(c^*(\tau, y; z))\rho(y|z)} \tag{32}$$

while in the equilibrium with incomplete markets it is given by

$$\frac{\partial}{\partial C_1} H_\tau(C_1^e(\tau; z); z) = \frac{U_\tau'(C_1^e(\tau; z))}{\sum_y U_\tau'(c^e(\tau, y; z))\rho(y|z)} \tag{33}$$

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 utility

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 $t = 1$. This tractability arises from four related properties of these preferences. We prove each of these properties in the appendix.

Gorman Aggregation: Given subutility functions of the equicautious HARA class (5), Gorman aggregation holds in all conditional equilibria. That is, in all conditional equilibria, asset prices $p(y; z)$ are independent of the initial endowment of bonds $B(\tau; z)$ and also independent of moments of the distribution of types $\mu(\tau|z)$ other than the mean of this distribution. Specifically, define

$$\bar{\tau}(z) \equiv \sum_{\tau} \tau \mu(\tau|z). \tag{34}$$

Then, in all conditional equilibria,

$$\frac{U'_\tau(c(\tau, y_1; z))}{U'_\tau(c(\tau, y_2; z))} = \frac{U'_{\bar{\tau}(z)}(y_1)}{U'_{\bar{\tau}(z)}(y_2)} \equiv \frac{\bar{p}(y_1; z)}{\bar{p}(y_2; z)} \tag{35}$$

for all types τ and all y_1, y_2.

The intuition for this result is that feasibility implies that the average risk tolerance in the market is given by

$$R_{\bar{\tau}(z)}(y) = \frac{y}{\gamma} + \bar{\tau}(z)$$

in all conditional equilibria because all agents have linear risk tolerance with a common slope in consumption (determined by γ).

This result allows us to solve for equilibrium prices $p^*(y; z)$ and $p^e(y; z)$ (both equal to $\bar{p}(y; z)$) in the complete and incomplete markets case directly from the parameters of the environment. Moreover, share prices $D_1^*(z) = D_1^e(z) = \bar{D}_1(z)$ where $\bar{D}_1(z)$ is defined from prices $\bar{p}(z; z)$ and equation (16). Accordingly, we are also able to solve for the cost functions $H_\tau(C_1; z)$ directly from the parameters of the environment.

Linear Frontier of Feasible Allocations of Certainty Equivalent Consumption: 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

$$\sum_\tau \mu(\tau|z)C_1(\tau; z) = \bar{C}_1(z) \tag{36}$$

where

$$\bar{C}_1(z) \equiv U^{-1}_{\bar{\tau}(z)}\left(\sum_y U_{\bar{\tau}(z)}(y)\rho(y|z)\right) \tag{37}$$

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 result 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 (36). 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. The corresponding bondholdings in the equilibrium with complete asset markets are then given from the budget constraint (28) evaluated at this optimal allocation of consumption. Clearly, since the cost of delivering a given amount of certainty equivalent consumption is higher for agents who are less risk tolerant, agents who experience a low realized risk tolerance τ relative to the average $\bar{\tau}(z)$ receive a transfer insuring them against the welfare consequences of this negative shock in terms of the payoff from their portfolio of bonds funded by a transfer from those agents who experience a high realized risk tolerance τ relative to the average $\bar{\tau}(z)$.

One can also use the Gorman aggregation result to solve for the allocation of certainty equivalent consumption in the equilibrium with incomplete markets, $C_1^e(\tau; z)$ using the budget constraint (30) and imposing the bond market clearing condition $B^e(z) = 0$ for all z. The result (36) implies that this equilibrium allocation of certainty equivalent consumption is given by

$$C_1^e(\tau; z) = \bar{C}_1(z) + \left(\frac{\tau - \bar{\tau}(z)}{\frac{\bar{D}_1(z)}{\gamma} + \bar{\tau}(z)}\right)\left[\bar{C}_1(z) - \bar{D}_1(z)\right] \tag{38}$$

where $\bar{D}_1(z)$ is the price of a share of the aggregate endowment at $t = 1$ in state z.

It is straightforward to show that with incomplete asset markets, equilibrium certainty equivalent consumption (38) 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 $\bar{C}_1(z) - \bar{D}_1(z)$. 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 $\bar{\tau}(z)$.

That the term $\frac{\bar{D}_1(z)}{\gamma} + \bar{\tau}(z) > 0$ follows from the restrictions on parameters we must make to ensure that the HARA subutility is well defined. Specifically, for the utility of the agent with average risk tolerance to be well defined, we must have $\frac{y}{\gamma} + \bar{\tau}(z) > 0$ 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.

Type-independent marginal cost of certainty equivalent consumption 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

$$\frac{\partial}{\partial C_1} H_\tau(C_1(\tau; z); z) = \frac{U'_{\bar{\tau}(z)}(\bar{C}_1(z))}{\sum_y U'_{\bar{\tau}(z)}(y)\rho(y|z)} \tag{39}$$

With these three results, we have a complete solution of the model for the equilibria with complete and with incomplete asset markets. This allows us to develop a complete characterization of allocations and asset prices as functions of the risks over preference shocks τ and endowments y. We also wish to characterize the implications of the model for trading volumes and asset prices. To do so, we use a fourth result.

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 as follows.

Consider first the equilibrium with complete asset markets. 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. The quantity of shares purchased at $t = 1$ by this agent is

$$\phi^*(\tau; z) - 1 = \frac{\tau - \bar{\tau}(z)}{\frac{\bar{C}_1(z)}{\gamma} + \bar{\tau}(z)} \tag{40}$$

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

$$TV^*(z) = \frac{1}{2} \sum_\tau \frac{|\tau - \bar{\tau}(z)|}{\frac{\bar{C}_1(z)}{\gamma} + \bar{\tau}(z)} \mu(\tau|z) \tag{41}$$

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. Then the quantity of shares purchased at $t = 1$ by this agent is

$$\phi^e(\tau; z) - 1 = \frac{\tau - \bar{\tau}(z)}{\frac{\bar{D}_1(z)}{\gamma} + \bar{\tau}(z)} \tag{42}$$

Hence observed trade volumes are given by

$$TV^e(z) = \frac{1}{2} \sum_\tau \frac{|\tau - \bar{\tau}(z)|}{\frac{\bar{D}_1(z)}{\gamma} + \bar{\tau}(z)} \mu(\tau|z) \tag{43}$$

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.*

Then 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 (37) 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 (38).*

Asset prices at $t = 1$ *in equilibrium both with complete and incomplete asset markets are given by* $p^*(y; z) = p^e(y; z) = \bar{p}(y; z)$ *defined in (35) with* $\sum_y \bar{p}(y; z)\rho(y|1) = 1$ *as the numeraire. 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* $\bar{p}(y, z)$.

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 (40) and holding the remainder of their portfolio in risk-free bonds. This leads to share turnover of* $TV^*(z)$ *as in (41). 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 (42) 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 (43).*

In the equilibrium with complete asset markets, date $t = 0$ *bond prices* $Q^*(\tau; z)$ *are given by (29) evaluated at the optimal allocation of certainty equivalent consumption with common marginal cost of certainty equivalent consumption given as in (39). In the equilibrium with incomplete asset markets, date* $t = 0$ *bond prices* $Q^e(z)$ *are given by (31) evaluated at the equilibrium allocation of certainty equivalent consumption (38) with common marginal cost of certainty equivalent consumption given as in (39).*

Proof. This proposition follows from the four properties of the equicautious HARA utility function described above. Details are given in the appendix. □

3.1 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 endowment shocks but no preference shocks. This result, which we demonstrate here, follows from the properties of the equicautious HARA preferences used above. In particular, through Gorman aggregation, we can solve for equilibrium asset prices at $t = 1$ ($\bar{p}(y; z)$ and hence $\bar{D}_1(z)$) independently of

agents choice of bondholdings at $t = 0$. The linear frontier of feasible allocations of certainty equivalent consumption allows us to treat certainty equivalent consumption as a good that can be transferred readily across agents, with $\bar{C}_1(z)$ serving as the aggregate endowment of this good. The common marginal cost of certainty equivalent consumption for all agents allows us to account for the role of this cost in asset pricing at $t = 0$ simply through a change in measure. We now describe this result in greater detail.

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 (39), is given by

$$J(z) \equiv \frac{\sum_y U'_{\bar{\tau}(z)}(y)\rho(y|z)}{U'_{\bar{\tau}(z)}(\bar{C}_1(z))} = \frac{\sum_y \left[\frac{y}{\gamma} + \bar{\tau}(z)\right]^{-\gamma}\rho(y|z)}{\left[\frac{\bar{C}_1(z)}{\gamma} + \bar{\tau}(z)\right]^{-\gamma}} \tag{44}$$

Note that in the case with CARA utility, we have $J(z) = 1$ for all z. For values of $\gamma < \infty$, we have the Taylor approximation around the conditional mean realization of the endowment, $\bar{y}(z)$

$$J(z) \approx 1 + \frac{\sigma_y^2(z)}{2} \frac{1/\gamma}{\left(\frac{\bar{y}(z)}{\gamma} + \bar{\tau}(z)\right)^2} \tag{45}$$

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 $\gamma > 0$.

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 resource constraint (36). 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

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

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 2. *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 (29) for complete asset markets and (31) 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. □

Note that in this economy, the agents' endowments $Y_1(\tau; z)$ have a sensitivity to the idiosyncratic realization τ that depends on aggregates $\bar{D}_1(z)/\gamma + \bar{\tau}(z)$ and $\bar{C}_1(z) - \bar{D}_1(z)$ reflecting the average risk tolerance in the economy. This interaction of aggregate and idiosyncratic risk plays an important role in shaping date zero asset prices as described below.

Note as well 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 this preference risk.

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 under both complete and incomplete asset markets in greater detail.

From equations (41) and (43), 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 similar 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, in the economy with incomplete asset markets, results similar to those in Mankiw (1986) and Constantinedes and Duffie (1996). There is a direct connection between dispersion in agents' realized risk tolerances (and hence trade volumes) and ex-ante asset prices. Here, the idiosyncratic risk agents face in terms of their equilibrium certainty equivalent consumption as a result of their preference shocks can have an impact on asset prices quite similar to that discussed in those earlier papers. In fact, in the case with CARA preferences, which is the limiting case as $\gamma \to \infty$, our model is isomorphic to one in which agents experience idiosyncratic income shocks as in Mankiw (1986). At the end of the paper we show that this result can be extended to a dynamic economy as in Constantinedes 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 $t = 2$.

We have seen above that asset prices at time $t = 1$ for claims that pay off at $t = 2$ are given by $\bar{p}(y; z)$ in the equilibrium with either complete or incomplete asset markets. Moreover, from equation (35), 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 both the equilibria with complete and incomplete asset markets, asset prices at $t = 1$ and returns from $t = 1$ to $t = 2$ bear no direct connection to dispersion in risk tolerances across agents and hence no direct connection to observed trade volumes. Therefore, in the equilibrium with either complete or incomplete asset markets, any connection between realized trade volumes at $t = 1$ and expected asset returns from $t = 1$ to $t = 2$ is coincidental.

We illustrate these properties in three examples. For simplicity, specialize our preferences to the CARA case with $U_\tau = -\tau \exp(-c/\tau)$ which is the limit of our HARA case as $\gamma \to \infty$. The equilibrium prices and quantities in this case are given as in Proposition 1 with $\gamma \to \infty$ and the marginal cost of a unit of certainty equivalent consumption is equal to one for all z.

In the first example we consider, there is negative serial correlation in asset returns because variation in these returns is driven by shocks to average risk tolerance. However, because these shocks are common to all agents, they do not generate trade at $t = 1$.

Example 1: Common shocks to risk aversion Consider an economy in which the conditional distribution of the endowment y does not vary with z. Specifically, assume that

$y \sim N(\bar{y}, \sigma^2)$. Let the number of types $I = 2$, with $\tau_1 < \tau_2$. Let there be two aggregate states $z \in \{z_1, z_2\}$, with $\mu(\tau_1; z_1) = 1$ and $\mu(\tau_2; z_2) = 1$. Hence $\bar{\tau}(z_1) = \tau_1$ and $\bar{\tau}(z_2) = \tau_2$. With CARA preferences and a normal endowment, we have that share prices in period $t = 1$ in state $z = z_j$ are given by $P_1(z_j) = \bar{y} - \sigma^2/\bar{\tau}(z_j)$, and hence the expected excess return on a share of the aggregate endowment between $t = 1$ in state z and $t = 2$ is

$$\mathbb{E}\left[R_2(y; z)|z = z_j\right] = \frac{\bar{y}}{(\bar{y} - \sigma^2/\bar{\tau}(z_j))}$$

The realized return on a share of the aggregate endowment between periods $t = 0$ and $t = 1$ contingent on z_j being realized is directly proportional to

$$\frac{P_1(z_j)}{P_0} \propto \frac{\bar{y} - \sigma^2/\bar{\tau}(z_j)}{\bar{y}}$$

The correlation of the realized return between $t = 0$ and $t = 1$ with the expected return between $t = 1$ and $t = 2$ is then

$$Corr\left(\frac{1}{P_1(z)}, P_1(z)\right) = Corr\left(\frac{1}{\bar{y} - \sigma^2/\bar{\tau}(z)}, \bar{y} - \sigma^2/\bar{\tau}(z)\right) < 0$$

Hence, it is the case that returns on a share of the aggregate endowment have negative serial correlation, i.e. a low realized return between periods $t = 0$ and $t = 1$ is followed by a high expected return between periods $t = 1$ and $t = 2$. Note, however, in this case, there is no trade in shares at $t = 1$ because all shocks to risk tolerance are common across agents.

In the second example we consider, there is negative serial correlation in asset returns because variation in these returns is driven by shocks to average risk tolerance, as in Example 1. However, now there is a positive volume of trade at $t = 1$ because these shocks to risk tolerance have both an idiosyncratic and an aggregate component.

Example 2: Idiosyncratic shocks to risk aversion Let the number of types $I = 4$, with $\tau_1 < \tau_2$ and $\tau_3 < \tau_4$. Let there be two aggregate states $z \in \{z_1, z_2\}$, with $\mu(\tau_1; z_1) = \mu(\tau_2; z_1) = 1/2$ and $\mu(\tau_3; z_2) = \mu(\tau_4; z_2) = 1/2$. Hence $\bar{\tau}(z_1) = (\tau_1 + \tau_2)/2$ and $\bar{\tau}(z_2) = (\tau_3 + \tau_4)/2$.

With these assumptions, asset prices are given as in example 1 above, but, from equations (41) and (43), the volume of trade in shares is now given by

$$TV^*(z_1) = TV^e(z_1) = \frac{1}{2}\frac{\tau_2 - \tau_1}{\bar{\tau}(z_1)}$$

and

$$TV^*(z_2) = TV^e(z_2) = \frac{1}{2}\frac{\tau_4 - \tau_3}{\bar{\tau}(z_2)}$$

Note that here the state contingent volume of trade in shares is determined by the dispersion in the idiosyncratic shocks to risk aversion around average risk aversion as measured by $\tau_2 - \tau_1$ and $\tau_4 - \tau_3$. But again, here asset prices are determined entirely by the mean shock to risk aversion. Hence, any relationship between trading volume and the serial correlation of returns implied by the model with complete markets arises entirely from direct assumptions about how the dispersion of shocks to risk aversion is correlated with mean shocks to risk aversion.

Campbell, Grossman, and Wang (1993) demonstrate that the serial correlation of excess returns on a stock index is closer to negative one when trading volume is high[2] and closer to zero when it is low. They then construct a model economy in which agents face random shocks to their risk aversion and construct an example equilibrium in which the serial correlation of returns on the stock market as a whole is negative when trading volume is high and closer to zero when trading volume is low. The key idea in their example is that there are two different types of shocks that drive fluctuations in realized excess returns. One shock is a signal of future cash flows. This shock drives realized excess returns but it has no impact on expected excess returns going forward and no impact on trading volume. The second shock is a shock to preferences that simultaneously alters average risk aversion and the dispersion of risk aversion across agents. We illustrate the central ideas underlying their example in our next example.

Example 3: Trading Volume and the Serial Correlation of Returns   To accommodate two types of aggregate shocks, let there be four states in Z, with $z \in \{z_1, z_2, z_3, z_4\}$, and five possible types of preferences for agents τ, with $\tau \in \{\tau_1, \tau_2, \ldots, \tau_5\}$. Assume that the four states z have equal probability, so $\pi(z_j) = 1/4$. We let states z_1 and z_2 correspond to shocks to average risk aversion and the dispersion of risk aversion as in Example 2, and we let states z_3 and z_4 correspond to shocks to the expected dividend from the aggregate endowment.

As in Example 2, let $\tau_1 < \tau_2$ and $\tau_3 < \tau_4$. Let $\mu(\tau_1; z_1) = \mu(\tau_2; z_1) = 1/2$ and $\mu(\tau_3; z_2) = \mu(\tau_4; z_2) = 1/2$. Hence $\bar{\tau}(z_1) = (\tau_1 + \tau_2)/2$ and $\bar{\tau}(z_2) = (\tau_3 + \tau_4)/2$. Assume that $\bar{\tau}(z_1) < \bar{\tau}(z_2)$. If we let $\bar{\tau}$ denote the unconditional mean of $\bar{\tau}(z)$, choose $\tau_5 = \bar{\tau}$ and $\mu(\tau_5; z_3) = \mu(\tau_5; z_4) = 1$. Hence, here, in states z_1 and z_2 there are shocks to average risk aversion and the dispersion of risk aversion. In states z_3 and z_4, all agents have risk aversion equal to the unconditional expectation of average risk aversion.

Assume that the aggregate endowment $y \sim N(\bar{y}(z), \sigma^2(z))$, with $\bar{y} = \sum_{j=1}^{4} \bar{y}(z)/4$ equal to the unconditional expectation of $\bar{y}(z)$ and $\bar{\sigma}^2$ the unconditional expectation of $\sigma^2(z)$. Assume

---

[2]Actually, the deviation of trading volume from a trend.

that $\bar{y}(z_1) = \bar{y}(z_2) = \bar{y}$ and $\bar{y}(z_3) < \bar{y} < \bar{y}(z_4)$. Assume that $\sigma^2(z_3)$ and $\sigma^2(z_4)$ are chosen such that

$$\frac{\bar{y}(z_3)}{\bar{y}(z_3) - \sigma^2(z_3)/\bar{\tau}} = \frac{\bar{y}(z_4)}{\bar{y}(z_4) - \sigma^2(z_4)/\bar{\tau}}$$

Hence, in states z_1 and z_2, the distribution of the aggregate endowment is equal to its unconditional distribution while this conditional distribution varies across states z_3 and z_4. Note, however, that the expected excess return on a share of the aggregate endowment from period $t = 1$ to $t = 2$ is equal in states z_3 and z_4.

The results in examples 1 and 2 directly imply that returns on shares of the aggregate endowment will be negatively serially correlated in states z_1 and z_2 and that trading volume will be positive as a result of dispersion in agents' realized risk aversion. In contrast, trading volume is zero in states z_3 and z_4 and, by assumption, the expected excess returns on a share of the aggregate endowment from period $t = 1$ to $t = 2$ is equal in states z_3 and z_4. Hence, as in Grossman, Campbell, and Wang (1996), the serial correlation of returns on the aggregate endowment is negative when trading volume is positive and zero when trading volume is zero.

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

4.2   Trading Volume and Date $t = 0$ Asset Prices:

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 multiplicative expected excess returns and additive expected excess returns as defined in (21).

The following two restatements of incomplete markets equilibrium bond prices at $t = 0$ are useful to us below.

First we consider a restatement of incomplete markets equilibrium bond prices that is useful for computing additive risk premia

$$Q^e(z) = \left[\beta \frac{V'(\bar{C}_1(z))}{V'(\bar{C}_0)} + \beta \frac{\Delta(z)}{V'(\bar{C}_0)} \right] J(z) \tag{46}$$

31

with

$$\Delta(z) \equiv \sum_{\tau} \left[V'(C_1^e(\tau; z)) - V'(\bar{C}_1(z)) \right] \mu(\tau; z) \tag{47}$$

and $J(z)$ is defined as in equation (44). We then have that bond prices at $t = 0$ with incomplete markets are given by

$$Q^e(z) = Q^*(z) + \frac{\beta}{V'(\bar{C}_0)} J(z) \Delta(z)$$

We now wish to use these bond prices together with equation (22) to price the expected excess returns on assets in this economy from period $t = 0$ to $t = 1$, as defined in equation (20), on any asset with cash flows $d(y; z)$ in period $t = 2$. Recall that, due to Gorman aggregation, the price of this asset at $t = 1$ defined in equation (17), $P_1(z; d)$, is independent of the dispersion of realized risk tolerances across agents and hence is not directly connected to observed trade volumes at $t = 1$. In contrast, the price of this asset at $t = 0$ is potentially directly connected to observed trade volumes at $t = 0$ as follows.

This expression (46) for bond prices with incomplete markets implies that expected excess returns on assets in this economy from period $t = 0$ to $t = 1$ are given as in (22) by

$$\mathcal{E}_1^e(d) - 1 = - Cov \left(Q^*(z) + \frac{\beta}{V'(\bar{C}_0)} J(z) \Delta(z), R_1(z; d) \right)$$

which can be written as

$$\mathcal{E}_1^e(d) - 1 = (\mathcal{E}^*(d) - 1) - \frac{\beta}{V'(\bar{C}_0)} Cov\left(J(z) \Delta(z), R_1(z; d) \right) \tag{48}$$

That is, equation (48) implies that with incomplete markets, assets are prices at $t = 0$ according to a two factor model. The first factor is the pricing factor that pertains with complete asset markets. The second factor contains the term $J(z)\Delta(z)$ where $J(z)$ is defined in equation (44). Note that $J(z)$ is not directly affected by the dispersion of realized risk tolerances across agents and hence it is not directly connected to observed trade volumes at $t = 1$. In contrast, the term $\Delta(z)$ is directly connected to the dispersion in realized risk tolerances across agents and hence is directly connected to observed trade volumes. We discuss that connection in greater detail below.

Next we consider a restatement of incomplete markets equilibrium bond prices that is useful for computing multiplicative risk premia from $t = 0$ to $t = 2$. We have

$$Q^e(z) = \beta \frac{V'(\bar{C}_1(z))}{V'(\bar{C}_0)} J(z) L(z) \tag{49}$$

with

$$L(z) \equiv \sum_{\tau} \frac{V'(C_1^e(\tau; z))}{V'(\bar{C}_1(z))} \mu(\tau; z) \qquad (50)$$

and $J(z)$ is defined as in equation (44). We then have that bond prices at $t = 0$ with incomplete markets are given by

$$Q^e(z) = Q^*(z)L(z)$$

This gives us that the change of measures in the complete and incomplete asset markets equilibria use in computing inverse multiplicative expected excess returns in equation (24) are given by

$$\pi_{Q^*}(z) = \frac{Q^*(z)\pi(z)}{\sum_z Q^*(z)\pi(z)}$$

and

$$\pi_{Q^e}(z) = \frac{Q^*(z)L(z)\pi(z)}{\sum_z Q^*(z)L(z)\pi(z)} = \frac{\pi_{Q^*}(z)L(z)}{\sum_z \pi_{Q^*}(z)L(z)} \qquad (51)$$

Trade Volumes and $\Delta(z)$ Since the allocation of certainty equivalent consumption must satisfy the pseudo-resource constraint (36), the magnitude of $\Delta(z)$ for any given realization of z is a function of the dispersion in equilibrium certainty equivalent consumption across agents from dispersion in shocks to risk aversion across agents in that state z as well as the magnitude of the convexity or concavity of $V'(\cdot)$ as determined by the third derivative of the utility function $V(\cdot)$ for certainty equivalent consumption.[3] If $V'''(\cdot) > 0$, then, with dispersion in certainty equivalent consumption across agents of different types τ, we have $\Delta(z) > 0$, and, with $V'''(\cdot) < 0$, we have $\Delta(z) < 0$. In the borderline case with quadratic utility (so $V'''(\cdot) = 0$), we have $\Delta(z) = 0$.

We these results, we have the following comparative static result regarding the dispersion of shocks to risk aversion and bond prices used in asset pricing at $t = 0$.

Lemma 1. *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 $\Delta(z)$ has the higher equilibrium bond price at $t = 0$, $Q^e(z)$.*

---

[3]The effect of this third derivative of the utility function for certainty equivalent consumption on asset pricing is also related to its effect on the strength of the precautionary savings motive given fixed asset prices. See, for example Kimball (1990) for a discussion of the role of this third derivative of the utility function in determining the strength of the precautionary savings motive.

Proof. The proof is by direct calculation. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We can say a bit more, informally, by taking a second order Taylor approximation to agents' marginal utilities of certainty equivalent consumption

$$\Delta(z) \approx \frac{\beta}{2} \frac{V'''(\bar{C}_1(z))}{V'(\bar{C}_0)} \sum_{\tau} \left(C_1^e(\tau; z) - \bar{C}_1(z) \right)^2 \mu(\tau|z) =$$

$$\frac{\beta}{2} \frac{V'''(\bar{C}_1(z))}{V'(\bar{C}_0)} \left(\bar{C}_1(z) - \bar{D}_1(z) \right)^2 \sum_{\tau} \left(\phi^e(\tau; z) - 1 \right)^2 \mu(\tau|z)$$

where the second equality follows from (38) and (42). Hence, we have immediately that, holding fixed $\bar{\tau}(z)$, which determines $\bar{C}(z)$ and $\bar{D}_1(z)$, if $V''' > 0$, then to a second order approximation, $\Delta(z)$ is increasing in the variance of agents trades of shares at $t = 1$ and hence increasing in the idiosyncratic shocks to risk aversion as well as in the aggregate consumption risk premium at $t = 1$ given by $\bar{C}_1(z) - \bar{D}_1(z)$.

Now we relate equilibrium trade volume with $\Delta(z)$. From our first order Taylor approximation to $\Delta(z)$, we see that this term is directly proportional to $V'''(\bar{C}_1(z)) \sum_{\tau} \left(\phi^e(\tau; z) - 1 \right)^2 \mu(\tau|z)$, with $V'''(\bar{C}_1(z)) > 0$ if agents display precautionary savings in terms of certainty equivalent consumption. Hence, $\Delta(z)$ is increasing in the variance of individual agents' share transactions. Hence $\Delta(z)$ is directly connected to observed share trade volume at $t = 1$ in state z, $TV^e(z)$, to the extent that the variance of individuals' share trades is connected to aggregate trading volume.

First we 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 $\Delta(z)$.

Of course, this 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 $\Delta(z)$ 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 $\Delta(z)$ increase. In this sense, both trade volumes and $\Delta(z)$ are increasing in the dispersion of idiosyncratic risk-tolerance.

Trade Volumes, $L(z)$, and Asset Prices: Our change of measure formulas (51) are somewhat more convenient for deriving analytical results about asset risk premia. Note that in the asset pricing formula (24), the cash flows on the asset d the expected excess returns from period $t = 1$ and $t = 2$, $\mathcal{E}_{1,2}(d; z)$, and the probabilities $\pi_{Q^*}(z)$ are all independent of the dispersion of the preference shocks τ realized at $t = 1$ in state z. Thus, we confirm again that with complete asset markets, there is no direct connection between observed trade volumes and expected asset returns.

In contrast, 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 (51). 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.

In terms of magnitudes, to a first order Taylor approximation, we have

$$L(z) \approx 1 + \frac{V'''(\bar{C}_1(z))}{V'(\bar{C}_1(z))} \left(\bar{C}_1(z) - \bar{D}_1(z) \right)^2 \sum_\tau \left(\phi^e(\tau; z) - 1 \right)^2 \mu(\tau|z)$$

Hence, as is the case with $\Delta(z)$, 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 $\left(\bar{C}_1(z) - \bar{D}_1(z) \right)^2$.

We now use our results to develop two results regarding the impact of trading volumes on asset pricing. The first one is a comparison of risk premium across economies with different patters of trading volume, and the second a comparison of the risk premium of different assets in the same economy.

Correlation of Trading Volume and Average Risk Tolerance: For the first result we compare a 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 dispersion of risk-tolerance. We find that if V displays prudence (i.e. if $V''' > 0$) then any cash-flow with systematic risk has higher risk-premium in the economy in which dispersion is negatively correlated with risk tolerance.

Let 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:

If $z' > z$ then $\bar{\tau}(z') < \bar{\tau}(z)$ and $\qquad (52)$

If $z' > z$ then $\tilde{\mu}(\cdot; z')$ is a less dispersed (in second order stochastic sense) than $\mu(\cdot|; z)$. $\quad (53)$

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

$$d(y', z') > d(y, z) \text{ for all } z, z' \text{ and } y' > y. \qquad (54)$$

Proposition 3. *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 (52). The first economy has constant dispersion on the idiosyncratic risk-tolerance, 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 in (53), so $L(z) = L_2(z)$ is decreasing in z. For both economies fix the same asset $d(\cdot)$, with a systematic payoff exposure as in (54). 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 positively associated with the market-wide risk tolerance) has higher expected excess returns.*

Proof. Given the assumption that d has systematic exposure, we have that in both economies, $\mathbb{E}(d|z)/\mathbb{E}(d)$ is an increasing function of z. Given the assumption that the conditional distribution of y is constant across z and that average risk tolerance $\bar{\tau}(z)$ is increasing in z, we have that the inverse of the conditional expected excess return on d, $1/\mathcal{E}_{1,2}(z, d)$, is also increasing in z. Given our assumption that the liquidity weights in the first economy $L_1(z)$ are constant across z and that the liquidity weight in the second economy $L_2(z)$ are decreasing in z, we have that the expectation on the right hand side of equation (24) is higher in the first economy than the second. Hence this implies that expected excess returns from period $t = 0$ to $t = 2$ is lower in the first economy than the second. $\qquad \square$

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 the dispersion on idiosyncratic income shocks.

Correlation of Asset Payoffs with Trading Volumes 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, the second corollary means that assets whose cashflows load on states of high trade volume have low expected excess returns. We believe that this corollary illustrate that our model exactly reproduces the logic of the quote of Pastor and Stambaugh (2003) reproduced above.

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 4. *Consider an economy with incomplete markets with the same market-wide risk-tolerance and same conditional distribution of aggregate risk across states as in (**??**). Assume that the states are order in terms of dispersion of idiosyncratic risk-tolerance as in (53). Consider two cash-flows, $\tilde{d}$ and d where $\tilde{d}$ loads more in states with hither dispersion of idiosyncratic risk-tolerance. In particular, assume that $(\mathbb{E}[\tilde{d}|z]/\mathbb{E}[\tilde{d}])(1/\mathcal{E}_{1,2}(z,\tilde{d}))$ is decreasing in z, while $(\mathbb{E}[d|z]/\mathbb{E}[d])(1/\mathcal{E}_{1,2}(z,d))$ is independent of z. Then the asset with cash-flow $\tilde{d}$ has higher price, i.e $\mathcal{E}_{0,2}(\tilde{d}) < \mathcal{E}_{0,2}(d)$.*

Proof. It follows directly from the observation that $L(z)$ is decreasing in z. $\square$

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 $t = 1$. In this section we show that a Tobin tax on trade has a zero first order effect in the complete market equilibrium, but instead it has a first order negative welfare effect in the incomplete market case. In other words, a Tobin subsidy to trade increases ex-ante welfare. To further understand the reasons behind this results, we study a mechanism design problem where risk tolerance shocks τ are private information. We show in that setup that the second best allocation share some features of a Tobin subsidy to trade.

5.1 Tobin Tax

In the standard general 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 if we start from the baseline equilibrium in the complete markets economy. 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 $t = 0$.

The line of argument for these results is as follows. Recall that in a setting with quasi-linear utility, the standard argument establishing the zero derivative of welfare with respect to a marginal increase in a commodity tax has two components. The first component is an envelope condition that states that the marginal impact on welfare that arises from the induced change in quantities consumed by buyers and sellers is zero for each agent in the economy because these agents had chosen quantities optimally in the original equilibrium and because there are no wealth effects on consumption due to quasi-linear utility, which make aggregation straightforward. Under our assumption that agents' type dependent sub-utility function is equicautious HARA, we have that at the time of trade at $t = 1$ the *average* welfare effect in across agents is zero in equilibrium, this is indeed the reason why there is aggregation. This is the case in both the complete markets and incomplete markets economy. Hence, this first component of the standard analysis of deadweight loss is also equal to zero in our model, regardless of whether there are complete or incomplete markets at $t = 0$ for claims contingent on idiosyncratic shocks to risk aversion.

The second component of the standard argument establishing the zero derivative of welfare with respect to a marginal change in a commodity tax concerns the welfare analysis of the incidence of the tax. Specifically, if a commodity tax is applied in a market with positive trade in the undistorted equilibrium, with the tax revenue rebated back lump sum to all agents in the economy, a redistribution of wealth, as measured by the distribution across agents of consumption of the good that enters utility linearly, occurs. Some agents will gain and some will lose from the tax. In our model, this redistribution corresponds to a redistribution of risk free bond holdings, and hence a redistribution of certainty equivalent consumption $C_1(\tau; z)$ across agents of different types τ at $t = 1$ in aggregate state z. In the standard analysis of the welfare

effects of a commodity tax this has no aggregate effect, due to quasilinearity of preferences. In our economy with complete asset markets, this redistribution has no first order effects on ex-ante welfare (expected utility as of $t = 0$), because the baseline socially efficient equilibrium allocation has the marginal utility of certainty equivalent consumption equated across agents ($V'(C_1(\tau_i; z) = V'(C_1(\tau_j, z))$) in each aggregate state z.

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 $t = 2$.[4] 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 $t = 2$, 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 are simply endowed 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

$$\sum_\tau s(\tau; z; \omega)\mu(\tau|z) = 0, \tag{55}$$

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

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

---

[4]The restrictions to two assets is immaterial since with equicautios HARA utilities the two-mural fund separation theorem hold.

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 equivalence for agent with τ in state z for the conditional equilibrium with a transaction tax ω for $i \in \{*, e\}$ corresponding to the complete market and equal wealth equilibrium. We are interested in 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 equicautios HARA preferences in a conditional equilibrium we get that:

$$\frac{d}{d\omega}C_1^i(\tau; z; 0) = \begin{cases} \frac{\mathbb{E}[U'_{\bar{\tau}(z)}(y)|z]}{U'_{\bar{\tau}(z)}(\bar{C})}\left[(s^i(\tau; z; 0) - 1)\left(-\frac{\partial \bar{D}_1^i(z;0)}{\partial\omega} - 1\right) + TV^i(0, z)\right] & \text{if } s^i(\tau; z; 0) > 1 \\ \frac{\mathbb{E}[U'_{\bar{\tau}(z)}(y)|z]}{U'_{\bar{\tau}(z)}(\bar{C})}TV^i(0, z) & \text{if } s^i(\tau; z; 0) = 1 \\ \frac{\mathbb{E}[U'_{\bar{\tau}(z)}(y)|z]}{U'_{\bar{\tau}(z)}(\bar{C})}\left[(s^i(\tau; z; 0) - 1)\left(-\frac{\partial \bar{D}_1^i(z;0)}{\partial\omega}\right) + TV^i(0, z)\right] & \text{if } s^i(\tau; z; 0) < 1 \end{cases}$$

(56)

for all τ, z and $i \in \{e, *\}$.

To make this standard argument regarding the impact of a transactions tax on welfare in the equilibrium with complete markets more formal, 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

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

(57)

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.

Note that we can analyze the effect of a trade tax z by z, and so we define $W^i(\omega; z)$ as the effect of the tax conditional on z so:

$$W^i(\omega) = \sum_z \pi(z)W^i(\omega; z) \text{ where } W^i(\omega; z) = \beta \sum_\tau V\left(C^i(\tau; z; \omega)\right)\mu(\tau'z) \text{ thus}$$

$$\frac{dW^i(\omega)}{d\omega}|_{\omega=0} = \sum_z \pi(z)\frac{dW^i(\omega; z)}{d\omega}|_{\omega=0} \text{ for } i \in \{e, *\}.$$

Using (56) for either complete markets and incomplete markets, as well as the market clear for shares (55) we must have

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

Since in the undistorted equilibrium with complete asset markets, $C_1^*(\tau;z) = \bar{C}_1(z)$ for all τ, the formula (57) 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 5. *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 market 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 (38), 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 (57) can be written

$$\frac{dW}{d\omega} = \beta\sum_z \pi(z)\mathbb{C}ov\left(V'(C_1(\tau;z)), \frac{d}{d\omega}C_1(\tau;z)\,|\,z\right) \tag{58}$$

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.

What we now show is that the reallocation of certainty equivalent consumption that results from the incidence of a tax on transactions in shares at $t = 1$ exacerbates this baseline misallocation of certainty equivalent consumption, reducing certainty equivalent consumption for risk

averse agents and raising it for risk tolerant 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 also explain why the case general case requires extra assumptions.

For both cases it is convenient first to establish an intermediate result on the effect on the equilibrium price D of a small transaction tax. Define $S(\tau; zD, T)$ as the optimal trade in shares for an investor with τ, facing a price D and receiving a transfer T. The first order condition for the risky asset trade is:

$$\mathbb{E}\left[U'_\tau\left(y + (S(\tau; z, D, T) - 1)(y - D) + T\right)(y - D) \mid z\right] = 0. \tag{59}$$

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

Lemma 2. *Let $S(\tau; zD, T)$ be defined as the solution of* (59) *evaluated at the equilibrium price $D = \bar{D}_1(z)$ for equal wealth and at transfer $T = 0$. Then:*

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

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

where $S(\tau; z; D, T) = \phi^e(\tau, z)$ denotes agent's τ holding of risky asset in the equal wealth equilibrium conditional on z.

Using this lemma we can derive the following characterization for the impact of prices of a transaction tax:

Proposition 6. *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 less than the Tobin tax:*

$$\frac{D(z; 0)}{d\omega} = -\sum_{\tau > \bar{\tau}} \phi^e(\tau; z)\mu(\tau|z) = -\left[TV^e(z) + \sum_{\tau > \bar{\tau}}\mu(\tau|z)\right] \in (-1, 0). \tag{60}$$

With only two types for τ, to prove the result that ex-ante welfare goes down, we simply need to show that

$$\frac{d}{d\omega} C_1^e(\tau_1; z; \omega) = \left(\frac{d}{d\omega} \bar{D}_1(z) + \mu(\tau_1; z) \right) (1 - \phi^e(\tau_1; z)) < 0,$$

i.e. that, as the share transactions tax drives down the selling price of shares, the decline in revenue that the risk averse agent gets from selling some of his shares for bonds is larger than the increase in the rebate revenue that he receives as a transfer from the implementation of the policy. This result follows from a standard analysis of the incidence of the transactions tax on buyers and sellers of shares. Specifically, we have the following result with two types of agents.

Proposition 7. *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 equilibrium as a negative first order ex-ante welfare effect for each z, i.e.:*

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

$$\frac{\mathbb{E}[U_{\bar{\tau}}'(y)|z]}{U_{\bar{\tau}}(\bar{C}_1(z))} \phi(\tau_1; z)(1 - \phi(\tau_1; z)) \left[1 - \mu(\tau_2; z) \frac{1 + \phi(\tau_1; z)}{\phi(\tau_1; z)} \right] \mu(\tau_1; z) \tag{61}$$

and thus

$$\frac{dW^e(0; z)}{d\omega} < 0 \iff \mu(\tau_2; z) > \frac{\phi(\tau_1; z)}{1 + \phi(\tau_1; z)} \in \left(0, \frac{1}{2} \right) \tag{62}$$

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 (62).

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 8. *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 alternating signs, i.e.:*

$$sign \left(\frac{\partial^{n+1} V(c)}{\partial C^{n+1}} \right) = -sign \left(\frac{\partial^n V(c)}{\partial C^n} \right) \quad \text{for all } c, \text{ and all } n = 1, 2, 3, \dots \tag{63}$$

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.:

$$\frac{d}{d\omega} W^e(0; z) < 0 \tag{64}$$

43

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

$$\frac{d}{d\omega}W^e(0;z) \approx V''(\bar{C}_1(z)) \left(\frac{\bar{C}_1(z) - \bar{D}_1(z)}{\left[\frac{\bar{D}_1(z)}{\gamma} + \bar{\tau} \right]^2} \right) TV^e(z) \, Var(\tau|z) \tag{65}$$

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

5.2   Optimal non-linear tax-subsidy

Here we consider several results taking a mechanism design approach in which we presume 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 of why markets for insurance at time $t = 0$ of the realization of time $t = 1$ value of τ are absent.

We discuss briefly the case where a mechanism designer is able to control the consumption of an agent, and then turn to the case where the designer must use an investor specific linear rule. The linear sharing rules have an agent specific exposure to the realization of aggregate risk y and a agent specific transfer. This linear sharing rule restrict the planner to use the same type of transfer than a market with an aggregate stock and a bond, and hence make the comparison with a Tobin tax/subsidy transparent. .

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

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

Lemma 3. *The first best allocation (i.e. the complete market equilibrium) is not incentive compatible.*

Now we move to the analysis of conditionally efficient allocation. Recall that the set of conditionally efficient allocations coincide with the set of equilibrium for arbitrary distribution

of time $t = 1$ wealth, so they include the complete market equilibrium as well as the equal wealth incomplete market equilibrium. The equal wealth equilibrium is clearly incentive compatible, since the budget constraint is the same for all agents. We now argue that, when τ has a distribution with a density μ then it is the the *only* incentive compatible allocation among the conditionally efficient allocation.

Lemma 4. *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.

The equilibrium with the Tobin tax is a linear tax, and hence it makes agents to face two prices, a high one for those than buy risky shares and a low price for those that sale risky shares. Recall that the ex-ante welfare losses comes from the redistributive properties of the tax incidence. Hence in the remaining of 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 patter on taxes/subsidy with the ones for the Tobin tax. For now we concentrate on the case of CARA preferences.

In the case of CARA with a Tobin tax the equilibrium values of $s(\tau)/\tau$ equal to one constant for low τ and $s(\tau)/\tau$ equals a different, lower constant for high levels of τ. They reflect the two prices faced by investors, as well as as the relatively simpler form of the optimal portfolio with CARA utility function.

We restrict our attention to linear sharing rules, i.e. rules so that consumption c of an investor at $t = 2$ is given by $c(\tau, y) = \mathcal{B}(\tau) + \mathcal{S}(\tau)y$. While this is not restrictive for HARA preferences when there is no private information That is, the more general problem is to assign $c(y, \tau')$ when investor reports τ' at $t = 1$ and the aggregate consumption is y at $t = 2$. Nevertheless we found it interesting since it allows to use the same two "securities", an aggregate stock and a bond that are sufficient in an equilibrium but it allows to have an investor-τ specific price, and thus it is comparable with the case of a Tobin tax.

Motivated by Lemma 4 we will assume that τ is distributed with a density $\mu(\cdot)$. The planner assign to each investor an uncontingent transfer $\lfloor$ 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:

$$1 = \int \mathcal{S}(\tau)\mu(\tau)d\tau \tag{66}$$

$$0 = \int \mathcal{B}(\tau)\mu(\tau)d\tau \tag{67}$$

The certainty equivalent consumption of the investor once the realization of her risk tolerance τ is known to her is given by: $\tau\varphi(s/\tau) + b$. The function $\tau\varphi(s/\tau)$ is strictly increasing and strictly concave in s.

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

$$\tau\varphi\left(\frac{\mathcal{S}(\tau)}{\tau}\right) + \mathcal{B}(\tau) \geq \tau\varphi\left(\frac{\mathcal{S}(\tau')}{\tau}\right) + \mathcal{B}(\tau') \quad \text{for all } \tau' \tag{68}$$

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:

$$\int V\left(\tau\varphi\left(\frac{\mathcal{S}(\tau)}{\tau}\right) + \mathcal{B}(\tau)\right)\mu(\tau)d\tau \tag{69}$$

by choosing functions $s(\cdot)$ and $b(\cdot)$, subject to the physical constraints (66) and (67) and the incentive compatibility constraint (68) for each τ.

In this case the certainty equivalent is $\tau\varphi(\mathcal{S}/\tau) + \mathcal{B} = -\tau\log\left(\int\exp\left(-[\mathcal{S}\,y/\tau]\right)\rho(y)dy\right) + \mathcal{B}$.

$$\varphi(x) = -\log E\left[e^{-xy}\right] = -\log\int e^{-xy}\rho(y)dy \tag{70}$$

We summarize the properties of φ

$$\varphi(0) = 0\,, \ \varphi(1) > 0\,, \ \varphi'(x) > 0 \text{ for all } x \text{ if } y \geq 0 \ a.s.\,, \ \varphi'(0) = \mu_y\,, \tag{71}$$

$$\varphi''(x) < 0 \text{ for all } x\,, \ \varphi''(0) = -\sigma_y^2. \tag{72}$$

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

We think that this problem is close, but not identical to the one in Diamond's 1998 paper, which itself is a version of Mirrlees classic 1971 paper with quasilinear utility. For one, preferences are different, although they are quasilinear. Also the physical constraints are different. The constraints (66) is similar, yet not identical to the one in Diamond 98. The constraint (67) has no analog in Diamond's.

46

Proposition 9. *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* (67) *and* (66) *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:*

$$\frac{\theta_s}{\theta_b} = \varphi'(x(\tau_H)) = \varphi'(x(\tau_L)) < \varphi'(x(\tau)) \text{ for all } \tau \in (\tau_L, \tau_H) \tag{73}$$

and at the extremes we have:

$$\frac{1}{x(\tau_L)} \frac{dx(\tau)}{d\tau}\Big|_{\tau=\tau_L} = \frac{[\theta_b - V'(\mathcal{C}(\tau_L))]}{\tau_L} < 0 < \frac{1}{x(\tau_H)} \frac{dx(\tau)}{d\tau}\Big|_{\tau=\tau_H} = \frac{[\theta_b - V'(\mathcal{C}(\tau_H))]}{\tau_H} \tag{74}$$

since

$$V'(\mathcal{C}(\tau_H)) < \theta_b = \int_{\tau_L}^{\tau_H} V'(\mathcal{C}(\tau_L)) \mu(\tau) d\tau < V'(\mathcal{C}(\tau_L)) . \tag{75}$$

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.

We imagine the decentralization of the optimal non-linear tax as the planner designing a menu of contracts, and letting the agent deciding which one to pick. 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 (67) and (66) 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 5. *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:*

$$-\frac{d\bar{B}(\hat{\mathcal{S}}(\tau))}{dS} = \varphi'(\hat{x}(\tau)) \ \textit{for all} \ \tau \in [\tau_L, \tau_H] \tag{76}$$

$$-\frac{dB^e}{dS} = \varphi'(1/\bar{\tau}) < -\frac{dB^*}{dS} = \bar{\tau}\varphi(1/\bar{\tau}) \tag{77}$$

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:

$$\int_{\tau_L}^{\tau_H} \frac{d\bar{B}(\tau \hat{x}(\tau))}{dS} \tilde{\mu}(\tau) d\tau = -\bar{D}_1 \ \text{where} \ \tilde{\mu}(\tau) \equiv \frac{\tau\mu(\tau)}{\int_{\tau_L}^{\tau_H} \tau\mu(\tau)d\tau} \ \text{for all} \ \tau \in [\tau_L, \tau_H] \tag{78}$$

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 (78) with the result from Proposition 9 for the optimal non-linear tax we obtain that the following pattern.

A Appendix

Proof. (of Proposition 6) 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:

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

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:

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

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:

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

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:

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

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

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

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

$\square$

Proof. (of Proposition 7) Again we omit z to render the notation simpler. We want to compute:

$$\frac{d}{d\omega}C^e_1(\tau_1;z) = \frac{\mathbb{E}[U'_{\bar{\tau}}(y)]}{U_{\bar{\tau}}(\bar{C})}\left[-(\phi(\tau_1)-1)\frac{\partial}{\partial\omega}\bar{D}(0) + TV\right]$$

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:

$$\frac{d}{d\omega}C^e_1(\tau_1;z) = \frac{\mathbb{E}[U'_{\bar{\tau}}(y)]}{U_{\bar{\tau}}(\bar{C})}\left[(\phi(\tau_1)-1)\mu(\tau_2)\phi(\tau_2) + (\phi(\tau_2)-1)\mu(\tau_2)\right]$$

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

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

so

$$\frac{d}{d\omega}C^e_1(\tau_1;z) = \frac{\mathbb{E}[U'_{\bar{\tau}}(y)]}{U_{\bar{\tau}}(\bar{C})}\phi(\tau_1)(1-\phi(\tau_1))\left[1 - \mu(\tau_2)\frac{1+\phi(\tau_1)}{\phi(\tau_1)}\right]$$

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

$$\frac{dW^e(0; z)}{d\omega} < 0 \iff \mu(\tau_2; z) > \frac{\phi(\tau_1; z)}{1 + \phi(\tau_1; z)} \tag{79}$$

$\square$

Proof of Proposition (8). Again we omit z to render the notation easier to follow. Recall that in an equal wealth equilibrium:

$$C_1^e(\tau) - \bar{C}_1(0) = (\tau - \bar{\tau}) \frac{\bar{C}_1 - \bar{D}_1}{\bar{\tau} + \bar{D}_1/\gamma} \equiv \chi(\tau - \bar{\tau})$$

$$\phi^e(\tau) - 1 = \frac{\tau - \bar{\tau}}{\bar{\tau} + \bar{D}_1/\gamma} \equiv \eta(\tau - \bar{\tau})$$

$$TV^e = \int_{\bar{\tau}}^{\tau_H} (\phi^e(\tau) - 1)\mu(\tau)d\tau = \eta \int_{\bar{\tau}}^{\tau_H} (\tau - \bar{\tau})\mu(\tau)d\tau$$

$$\bar{D}'_1(0) = -\int_{\bar{\tau}}^{\tau_H} \phi^e(\tau)\mu(\tau)d\tau = -\left[\eta \int_{\bar{\tau}}^{\tau_H} (\tau - \bar{\tau})\mu(\tau)d\tau + \int_{\bar{\tau}}^{\tau_H} \mu(\tau)d\tau \right]$$

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

In this case the change on welfare of a small Tobin tax can be written as

$$
\frac{d}{d\omega}W^e(0;z) = TV^e \int_{\tau_L}^{\tau_H} V'(C_1^e(\tau))\mu(\tau)d\tau
$$

$$
- \bar{D}_1'(0) \int_{\tau_L}^{\tau_H} V'(C_1^e(\tau))\left(\phi^e(\tau) - 1\right)\mu(\tau)d\tau
$$

$$
- \int_{\bar{\tau}}^{\tau_H} V'(C_1^e(\tau))\left(\phi^e(\tau) - 1\right)\mu(\tau)d\tau
$$

$$
= \eta \left[\int_{\bar{\tau}}^{\tau_H}(\tau - \bar{\tau})\mu(\tau)d\tau\right]\left[\sum_{n=0}^{\infty}\chi^n \frac{V^{n+1}(\bar{C}_1)}{n!}\int_{\tau_L}^{\tau_H}(\tau - \bar{\tau})^n\mu(\tau)d\tau\right]
$$

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

$$
- \eta\left[\sum_{n=0}^{\infty}\chi^n \frac{V^{n+1}(\bar{C}_1)}{n!}\int_{\bar{\tau}}^{\tau_H}(\tau - \bar{\tau})^{n+1}\mu(\tau)d\tau\right]
$$

We can rewrite this expression as:

$$
\frac{d}{d\omega}W^e(0;z) = \eta\chi^n \frac{V^{n+1}(\bar{C}_1)}{n!} \times
$$

$$
\sum_{n=0}^{\infty}\left\{\left[\int_{\bar{\tau}}^{\tau_H}(\tau - \bar{\tau})\mu(\tau)d\tau\right]\left[\int_{\tau_L}^{\tau_H}(\tau - \bar{\tau})^n\mu(\tau)d\tau\right] - \bar{D}_1'(0)\left[\int_{\tau_L}^{\tau_H}(\tau - \bar{\tau})^{n+1}\mu(\tau)d\tau\right]\right.
$$

$$
\left. - \left[\int_{\bar{\tau}}^{\tau_H}(\tau - \bar{\tau})^{n+1}\mu(\tau)d\tau\right]\right\}
$$

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 $n = 1, 3, 5, \ldots$ 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 $n = 1, 3, 4, \ldots$ 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 3) 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. $\square$

Proof. (of Lemma 4) 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

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

53

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 (42) for the equal wealth conditionally efficient allocation. □

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

$$\varphi(x) = -\mu_y \left(e^{-x} - 1\right) \text{ and } \Phi''(x) = \mu_y e^{-x} [1 - x] \tag{80}$$

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:

$$\varphi(x) = -\log\left(1 - p + p e^{-x}\right) \text{ and } \Phi''(x) = \frac{n(1-p)e^x}{[(1-p)e^x + p]^3} [x(1-p)e^x + p(1+x)] > 0 \tag{81}$$

In this case Φ' is monotone.

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

$$\varphi(x) = \log\left(\frac{\lambda + x}{\lambda}\right) \text{ and } \Phi''(x) = \frac{\lambda - x}{(\lambda + x)^3} \tag{82}$$

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

Solving Planning Problem. 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:

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

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:

$$\mathcal{C}'(\tau) = \varphi \left(\frac{\mathcal{S}(\tau)}{\tau} \right) - \varphi' \left(\frac{\mathcal{S}(\tau)}{\tau} \right) \frac{\mathcal{S}(\tau)}{\tau} \tag{83}$$

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:

$$x(\tau) \equiv \frac{\mathcal{S}(\tau)}{\tau} \text{ and } \Phi(x) \equiv \phi(x) - \phi'(x)x. \tag{84}$$

We note the following properties of Φ:

$$\Phi(x) \equiv \varphi(x) - \varphi'(x)x \geq 0 \text{ with } \Phi'(x) = -\varphi''(x)x > 0 \text{ and} \tag{85}$$

$$\Phi''(x) = -\varphi''(x) - \varphi'''(x)x, \quad \Phi''(0) = \sigma_y^2 > 0. \tag{86}$$

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:

$$\mathcal{C}(\tau) = \mathcal{C}(\tau_L) + \int_{\tau_L}^{\tau} \Phi(x(t))dt \implies \int_{\tau_L}^{\tau_H} \mathcal{C}(\tau)\mu(\tau)d\tau = \mathcal{C}(\tau_L) + \int_{\tau_L}^{\tau_H} \left[\int_{\tau_L}^{\tau} \Phi(x(t))dt\right]\mu(\tau)d\tau \tag{87}$$

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

$$\bar{B} = \mathcal{C}(\tau_L) + \int_{\tau_L}^{\tau_H} \left[\int_{\tau_L}^{\tau} \Phi(x(t))dt\right]\mu(\tau)d\tau - \int_{\tau_L}^{\tau_H} \tau\varphi(x(\tau))\mu(\tau)d\tau = 0 \tag{88}$$

and integrating by parts and rearranging:

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

The planning problem is thus:

$$\max_{\{x(\cdot),\mathcal{C}(\cdot)\}} \int_{\tau_L}^{\tau_H} V\left(\mathcal{C}(\tau)\right)\mu(\tau)d\tau \tag{90}$$

subject to:

$$\mathcal{C}'(\tau) = \Phi\left(x(\tau)\right) \text{ for all } \tau \in (\tau_L, \tau_H) \tag{91}$$

$$1 = \int_{\tau_L}^{\tau_H} x(\tau)\tau\mu(\tau)d\tau \text{ and} \tag{92}$$

$$0 = \mathcal{C}(\tau_L) + \int_{\tau_L}^{\tau_H} \Phi(x(\tau))\left[\int_{\tau_L}^{\tau} \mu(t)dt\right]d\tau - \int_{\tau_L}^{\tau_H} \tau\varphi(x(\tau))\mu(\tau)d\tau \tag{93}$$

We can write the Lagrangian, with multiplier $\lambda(\tau)\mu(\tau)$ for the each IC constraint (91), multiplier θ_s for constraint (92), and multiplier θ_b for constraint (93). After using integration by parts the Lagrangian reads:

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

The first order conditions are:

$$\mathcal{C}(\tau): \ -V'\left(\mathcal{C}(\tau)\right)\mu(\tau) = \lambda'(\mu)\mu(\tau) + \mu'(\tau)\lambda(\tau) \text{ for } \tau \in (\tau_L, \tau_H) \tag{94}$$

$$\mathcal{C}(\tau_H): \ \mu(\tau_H)\lambda(\tau_H) = 0 \tag{95}$$

$$\mathcal{C}(\tau_L): \ \mu(\tau_L)\lambda(\tau_L) = \theta_b \tag{96}$$

$$x(\tau): \ \theta_s\,\tau\,\mu(\tau) + \theta_b\,\Phi'(x(\tau))\left[\int_\tau^{\tau_H}\mu(t)dt\right]$$
$$= \theta_b\,\tau\,\varphi'(x(\tau))\,\mu(\tau) + \Phi'(x(\tau))\mu(\tau)\lambda(\tau) \text{ for } \tau \in (\tau_L, \tau_H] \tag{97}$$

Proof. (of Proposition 9) Rearranging the first order condition with respect to $x(\tau)$:

$$\frac{\theta_s}{\theta_b} - \varphi'(x(\tau)) = \frac{\Phi'(x(\tau))}{\tau\mu(\tau)}\left[\frac{\mu(\tau)\lambda(\tau)}{\theta_b} - \int_\tau^{\tau_H}\mu(t)dt\right] \tag{98}$$

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

$$\frac{\theta_s}{\theta_b} - \varphi'(x(\tau_H)) = \frac{\theta_s}{\theta_b} - \varphi'(x(\tau_L)) \tag{99}$$

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

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

Rearranging:

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

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

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

We have that

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

where the first one follows from differentiating the Lagrangian, and the second by increasing $\mathcal{C}(\tau)$ by $\bar{B}$ at all τ, and keeping $x(\tau)$ constant. 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 inteior 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 $\theta_s/\theta_b - \varphi'(x(\tau)) = \Phi'(x(\tau)/[\tau\mu(\tau)]\Psi(\tau)$. Note that

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

Using that $\mathcal{C}'(\tau) > 0$, and that $V'\left(\mathcal{C}(\tau_H)\right) < \theta_b < V'\left(\mathcal{C}(\tau_L)\right)$, so that $\Psi'(\tau_L) < 0$, $\Psi'(\tau_H) > 0$ and $\Psi'(\tau^*) = 0$ at a unique value of τ for which $V'\left(\mathcal{C}(\tau^*)\right) = \theta_b$. 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 our result. $\qquad\square$

Proof. (of lemma 5). 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 $\bar{C}_1 = \bar{\tau}\varphi(1/\bar{\tau}) > \varphi'(1/\bar{\tau}) = \bar{D}_1$. $\qquad\square$