The Private Memory of Aggregate Shocks*

Carlos E. da Costa† Vitor F. Luz‡
Fundação Getulio Vargas Yale University

February 10, 2010

Abstract

This paper addresses the interaction between privately observed idiosyncratic and publicly observed aggregate shocks in a dynamic Mirrlees's (1971) economy. Assuming both types of shocks to be independent from one another we show that constrained optimal allocations display memory with respect to *aggregate* shocks when both types are i.i.d.. Under a mild restriction on the nature of optimal allocations, the same is true for all less than perfectly persistent idiosyncratic shocks. Indeed, our numeric exercises indicate that the more persistent the private shock, the less memory allocations display, which indicates that the theoretical results relating presence of memory in the i.i.d. case and its absence in the perfect persistence case are not knife-edge. We use our findings to investigate some macroeconomic consequences of private information. Our results indicate that the model is capable of matching some qualitative features of the data: counter-cyclical labor wedges and risk premia.

1 Introduction

Following Wilson's (1968) landmark contribution, much has been learnt about optimal risk sharing. Absent asymmetric information the mutualization of private

*We thank Caio Almeida, Luis Braido, Vinicius Carrasco, Bev Dahlby, Humberto Moreira, Henry van Egteren, seminar participants at the Walton College and the University of Alberta for their invaluable comments. da Costa thanks the hospitality of MIT and gratefully acknowledges financial support from CNPq.

†carlos.eugenio@fgv.br
‡vluz@yale.edu

risks and its consequences for the dynamics of private consumption as a function of aggregate consumption are now well understood. In particular, conditional on aggregate consumption, any individuals current consumption ought to be independent of past information, i.e., first best allocations are memoryless.

If, however, there is private information with regards to idiosyncratic shocks, full insurance is no longer feasible. Contrary to the case with no private information, (constrained) optimal schemes exhibit memory, leading to extreme results in the long run.[1] Our knowledge of dynamic insurance in the presence of private information stems from a large literature, starting with the early contribution of Rogerson (1985) followed by the methodological advances found in Green (1987), Spear and Srivastava (1987) and Thomas and Worrall (1990) which built our current understanding of fundamental issues of dynamic insurance schemes. Most of this latter literature, however, focuses on the case in which only private risk exists, making assumptions on the nature of idiosyncratic shocks that ultimately lead to the elimination of aggregate risk – Phelan (1994) being a noteworthy exception.

In this paper we consider the interaction between aggregate shocks and private allocations which arises in a dynamic environment. Our prototype economy is a dynamic Mirrlees economy with a finite number of productivity levels and i.i.d. aggregate shocks. We assume that aggregate and private shocks are independent from one another and explore how constrained optimal insurance may still endogenously generate inter-temporal links in optimal allocations. That is, we eliminate all exogenous interactions between private and public shocks and investigate the interconnections displayed by the corresponding allocations.

We show a strong form of dependence between aggregate shocks and private allocations: allocations exhibit memory with respect to aggregate shocks. That is, today's allocation depends not only on current shock, as in Wilson (1968), and on yesterday's idiosyncratic shock, as in Rogerson (1985), but also on yesterday's aggregate shock, despite it's being public and independent of idiosyncratic shocks.

It is important to emphasize that our independence assumptions rules out nontrivial interactions between aggregate and private risk due to the reasons found in Holmström (1982). Namely, that the aggregate state may serve as a signal about

[1]The non-existence of non-degenerate steady-states and the associated immiseration results are now well understood consequences of back loading incentives that typically characterizes such environments.

the agents' private information. Instead, the rationale for our results is closer to Levin (2002). In our model, the resource constraint induces an interdependence in the provision of incentives across agents, while in his model, the same type of interdependence arises due to an endogenous restriction regarding what may be promised to the pool of agents in a credible way.[2] The rationale for backloading incentives is straightforward: someone who is hit by a bad shock need not suffer all its consequences today if, in the future, she may turn out to experience a good shock. The more persistent the shock, the less probable this event, thus the lower the incentives for backloading. In the limit, when shocks are perfectly persistent incentives are evenly spread across periods. With aggregate shocks, the distribution of incentives across states of nature becomes an issue as well, since the provision of incentives may alter the marginal value of aggregate resources.

Spreading incentives across state of nature is not identical to spreading shocks across periods, however. For the latter to be useful all that is needed is the existence of a (however small) probability that an agent reverses from one private state to the other. Spreading across states depends on the specifics of the curvature of utility functions.[3]

Phelan (1994) specifies a perpetual life OLG economy, assumes that individuals have constant absolute risk aversion—CARA—preferences, and restrict private shocks to be i.i.d.. He is, then, able to fully characterize the steady state of a dynamic moral hazard model with aggregate uncertainty. By eliminating income effects and killing the tendency for ever increasing inequality through the perpetual assumption he is able to find closed form solution for optimal allocations. Although this is not the focus of his paper, Phelan (1994) shows that optimal allocations depend on past aggregate variables in his prototype economy.

In contrast with Phelan (1994) we focus on a Mirrlees economy and allow for persistence in private shocks. Our numerical simulations indicate that the presence of memory which characterizes an environment with i.i.d. private shocks—

[2]We thank Vinicius Carrasco for pointing this out.

[3]In the class of utility function studied here, this reduces to the coefficient of risk aversion. As shown by Demange (2008) it is possible to define an incentive premium that makes the marginal utility of consumption for a representative consumer identical to the marginal value of resources for the society. Under our constant relative risk aversion — CRRA — specification, the sign of this premium only depends on whether risk aversion is greater or less than one. For log preferences this premium is zero.

reminiscent of Phelan (1994)—and the lack of memory which characterizes the pure heterogeneity case—explored by Werning (2007)—are not knife-edge results: there is a monotonic relationship between private persistence and aggregate memory.

Our focusing on a Mirrlees economy allows us to explore the 'macro' consequences, i.e., the behavior of aggregate variables of our prototype economy. [4] Specifically, we check if our model is capable of generating two stylized facts which are are hard to account for using a standard representative agent setting: the countercyclical movements in risk premia and labor wedges.

Our dynamic agency model is potentially useful for one to accommodate asset pricing anomalies. First, although idiosyncratic and aggregate shocks are independent, the same is not true for allocations. Second, to implement efficient allocations, agents must have restricted access to financial markets, a point that is crucial for an incomplete markets model to generate any type of interesting asset pricing behavior; a point emphasized by Cochrane (2001).

We first use the relevant pricing kernel as defined in Kocherlakota (2005) to generate the price for a risk free asset and a stock, understood as a claim to aggregate income. Next we use the aggregate data generated in our heterogeneous agent economy as if it were a representative agent economy to try and price these assets.[5] Although the model is capable of generating countercyclical risks premia, our oversimplified numerical examples do not allow for a quantitative assessment of the model, which is in contrast with Kocherlakota and Pistaferri (2007) or Kocherlakota and Pistaferri (2009) which use real consumption and asset pricing data.

As for the labor wedge, Chari et al. (2007) have shown that prototype economies can be mapped into a representative agent economies with time and state varying wedges. They go on to show that the labor wedge is key to explain business cycle stylized facts.[6] Because incentives in our model are provided in the leisure/consumption margin in a time and state varying fashion, our prototype

[4] Although most of the literature of the so-called Dynamic Public Finance Literature is normative, in recent years, a relatively large body of research has arisen on the Macroeconomics implications of private information—see Sleet (2006) for a brief account. The underlying idea is that many different institutional arrangements are capable of inducing any allocation. So one should focus on allocations, which is what we are able to, at least partially, characterize.

[5] Despite market incompleteness we shall see that we can use 'the' pricing kernel instead of 'a' pricing kernel in the current setting. See Kocherlakota (2005).

[6] See also Shimer (2009).

economy is potentially useful for understanding state-dependent labor wedges. For high levels of risk aversion, the wedge is countercyclical as seen in the data. For risk aversion less than unity, whilst our economy does generate state varying wedges, the covariance with aggregate product has the wrong sign. Optimal wedges are smaller in bad times while empirically it is the opposite.

Although memory and state varying wedges are related, the relationship is only indirect. In fact, it is non-separability that generates the result. In our environment we define a form of separability of allocations with respect to aggregate shocks. This form of separability plays an important role in our discussion of memory. Indeed, an allocation is memoryless only if it is separable in previous periods. It is separability that is crucial for generating state varying wedges. Due to the relationship between separability and memory, it is then the case that an allocation exhibits time varying wedges only if it displays memory.

The rest of the paper is organized as follows. Section 2 describes a general setting that encompasses both economies that we investigate. The Atkeson and Lucas (1992) framework is studied in section 2.1 while the Mirrlees (1971) setting is covered in section 3. The macro implications of our model are considered in 4. Numerical exercises are conducted in section 5. Section 6 concludes the paper.

2 Basic Setting

The economy is inhabited by a continuum of measure one of ex-ante identical individuals, each living for T periods. Agents have preferences defined on a consumption set $X \subseteq \mathbb{R}_+^T$. In every period, agents are subject to idiosyncratic shocks, $\theta_t \in \Theta = \{\theta(H), \theta(L)\}$, with $\theta(H) > \theta(L)$, which affects not only their welfare but also the way they rank different streams $\{x_t\}_{t=1}^T \in X^T$ of the relevant bundle. We use $\theta^t = (\theta_1, ..., \theta_t)$ to denote a history of idiosyncratic shocks up to period t. For any $\theta^{t+\tau}$ and θ^t, we say that $\theta^{t+\tau} \succ \theta^t$ if the first $t+1$ entries of $\theta^{t+\tau}$ are equal to θ^t.

Given a history of shocks, θ^T, individuals rank deterministic streams $\{x_t\}_{t=1}^T \in X^T$, according to

$$\sum_{t=1}^{T} \beta^{t-1} U\left(x_t, \theta^t\right),$$

where $U(., \theta) : \mathbb{R}_+^2 \to \mathbb{R}$ is a smooth, strictly increasing and strictly concave function, for all $\theta \in \Theta$.

These idiosyncratic shocks are not the only source of uncertainty in this economy. In each period, an aggregate shock represented by a random variable $z_t \in Z \equiv \{\overline{z}, \underline{z}\}$, with $\overline{z} > \underline{z} > 0$, affects the economy's technology. Let also $z^t = (z_1, ..., z_t)$ denote the history of aggregate shocks. We assume the aggregate shocks to be i.i.d. and distributed according to $\pi(z)$ ($\pi(z^t)$ is the product measure induced on Z^T).

Conditional on z^t, idiosyncratic shocks, θ^t, are drawn from independent distributions μ_t defined over Θ^t. We assume that a law of large numbers applies so that, at period t, state z^t, the cross-sectional distribution of agents coincides with the ex-ante distribution μ_t. We use $\mu_t\left(\theta^t|\theta^{t-1}\right)$ to denote the period t conditional distribution, following history θ^{t-1}.

Idiosyncratic shocks are private information while aggregate shocks are publicly observed.

An *allocation* is $x = \{x_t\}_{t=1}^T$, with $x_t : Z^t \times \Theta^t \to \mathbb{R}_+^L$ for each t, where $x_t(\theta^t, z^t)$ is a $(\theta^t, z^t)-$measurable function that denotes the bundle allocated to an agent with history θ^t at period t, when aggregate history is z^t. Agents' preferences are represented by a von Neumann-Morgenstern utility function,

$$\mathbb{E}\left[\sum_t \beta^{t-1} U\left(x_t(\theta^t, z^t), \theta_t\right)\right] = \sum_t \beta^{t-1} \sum_{z^t} \sum_{\theta^t} \mu_t\left(\theta^t\right) \pi_t(z^t) U\left(x_t(\theta^t, z^t), \theta_t\right). \quad (1)$$

Technology is represented by a transformation function $G : \mathbb{R}_+^L \times Z \to \mathbb{R}$. We say that an allocation x is *resource feasible* if

$$G\left(\sum\nolimits_{\theta^t} \mu_t\left(\theta^t\right) x_t(\theta^t, z^t), z_t\right) \leq 0, \text{ for all } t, z^t. \quad (2)$$

A benevolent planner maximizes individuals' expected utility. From a period 1 perspective this is equivalent to our assuming that the government maximizes a utilitarian social welfare function.

Due to the presence of private information, we write the planner's program as a mechanism design problem. Throughout the analysis we assume that the planner is endowed with a commitment technology. This will allow us to restrict implementation to direct revelation mechanisms, in which agents are asked to report their types at each period, and are assigned corresponding bundles.

Define a *reporting strategy,* $\sigma = \{\sigma_t\}_{t=1}^T$, as a sequence of mappings $\sigma_t : Z^t \times \Theta^t \to \Theta$, which associate to every history $\left(z^t, \theta^t\right)$ an announcement $\hat{\theta}$. Two things are worth mentioning here. First, since individuals cannot lie about the aggregate state of the economy, reports are restricted to idiosyncratic shocks. Announcement strategies may, nonetheless, depend on z^t. Second, we assume that the agent only announces the current shock, θ_t, in period t, and not the history θ^t. Alternatively the mechanism could be described by requiring the agent to announce θ^t instead. However, given that we assume perfect recall by the government, strategies that do not respect the condition $\sigma^t(\theta^t) = (\sigma^{t-1}(\theta^{t-1}), \theta')$ for some θ' are ruled out. As a consequence, asking θ_t each period is without loss of generality; our choice being due to notational simplicity. The set of all admissible strategies is represented by Σ. We also use $\sigma^{TT} = \left\{\sigma_t^{TT}\right\}_{t=1}^{T}$ to denote the truth-telling strategy $\sigma_t^{TT}\left(\theta^t\right) = \theta_t$ for all t.

Let $\mathcal{U}\left(x, \sigma\right)$ be the utility derived from an agent choosing reporting strategy σ given allocation x, i.e.,

$$\mathcal{U}\left(x, \sigma\right) \equiv \mathbb{E}\left[\sum\nolimits_t \beta^{t-1} U\left(x_t\left(\sigma^t(\theta^t, z^t), z^t\right), \theta_t\right)\right].$$

An allocation x is *incentive compatible* if

$$\mathcal{U}\left(x, \sigma^{TT}\right) \geq \mathcal{U}\left(x, \sigma\right), \text{ for all } \sigma \in \Sigma. \tag{3}$$

We shall (partially) characterize the solution to the problem of maximizing (1) subject to (2) and (3). We shall focus on the special case of a dynamic Mirrlees economy—Section 3. However, we shall first consider a taste shock example—Section 2.1— which will allow us to illustrate some of the concepts and issues to be discussed in our main setup. Before moving on, under our assumption that the sets of all possible shocks is finite, and assuming for the moment $T < \infty$, existence and uniqueness of this solution is trivially verified for both models.

2.1 A Preference Shock Example

Consider a two period version of Atkeson and Lucas (1992) with aggregate shocks. Individuals care only about consumption; $x = c \in \mathbb{R}_+$. An allocation, x, is therefore a $\hat{\theta}^t-$measurable mapping from history of announcements, $\hat{\theta}^t$, to a consumption stream, $c = \{c_t\}_{t=1}^{2} \in X \subseteq \mathbb{R}_+^2$. In each period, individuals are exposed to taste shocks that determine their marginal utility of consumption across states. That is, individuals preferences are of the form

$$U(c,\theta) = \begin{cases} \theta c^{1-\rho}/(1-\rho) & \text{, if } \rho > 0, \rho \neq 1, \\ \theta \log c & \text{, if } \rho = 1, \end{cases}$$

where θ is the taste parameter. For the moment let taste shocks be i.i.d.

The way in which the aggregate shock, z_t, determines the total amount of resources the economy is given by the feasibility condition: an allocation, c, is, feasible if and only if:

$$G\left(\sum\nolimits_{\theta^t} \mu\left(\theta^t\right) c(\theta^t, z^t), z_t\right) = \sum\nolimits_{\theta^t} \mu\left(\theta^t\right) c(\theta^t, z^t) - z_t \leq 0, \text{ for all } t, z^t.$$

As a benchmark case, consider the first best allocation,

$$c_t^{FB}(\theta^t, z^t) = z_t \theta_t^{\frac{1}{\rho}} \mathbb{E}\left[\theta^{\frac{1}{\rho}}\right]^{-1}.$$

Four important features of this allocation are noteworthy. The allocation is: *i)* independent of θ^{t-1}; *ii)* increasing in θ_t, *iii)* independent of z^{t-1} and; *iv)* linear in z_t.

With private information, things change substantially. Because temporary utility is defined in terms of a single good (consumption) in this setting, private information is revealed *only* through intertemporal trade-offs. For an allocation to be incentive compatible, more transfers today must generate lower expected transfers tomorrow, thus creating persistence of allocations, even though shocks are i.i.d. This is in opposition to *i*.

The last two features: independence with respect to z_{t-1}, and linearity in z_t are the ones we shall focus on. Note that, because aggregate shocks are public information and independent across periods, one might expect these properties to still characterize (constrained) efficient allocations. As we shall see, this is not true, in general. First, however, some definitions.

Definition 1 *If an allocation is such that in period t there are functions* $\eta : Z \rightarrow R_+$ *and* $\tilde{x} : \Theta^t \times Z^{t-1} \rightarrow R_+^L$ *such that* $x_t(\theta^t, z^t) = \tilde{x}_t(\theta^t, z^{t-1})\eta(z_t)$ *we say that it is* ***separable at*** *t.*

For our purposes it will also be important to define a function that displays this property in all periods.

Definition 2 *If an allocation is separable at t for all t we say it is* ***separable****.*

Separability is a property of allocations which will play a role in the discussions that follow. In the current setting, for example, first best allocation are characterized by a particular form of separability in which $\eta(z_t) = z_t$.

Finally, one last definition.

Definition 3 *If an allocation is such that, for all t, $x(\theta^t, z^t)$ is independent of z^{t-1}, i.e., it is of the form $\hat{x}(\theta^t, z_t)$, then we say it is* ***memoryless****.*

Note that the definition only applies to memory of past aggregate shocks. It is still the case that the allocation may, and will, depend on past idiosyncratic shocks.

Letting $w^*(\theta_1, z_1)$ define utility promises,

$$w^*(\theta_1, z_1) \equiv (1-\rho)^{-1} \sum_{\theta^2 \succ \theta_1} \sum_{z^2 \succ z_1} \mu_Z(z_2)\mu_\Theta(\theta)\theta c_2^*(\theta^2, z^2)^{1-\rho},$$

makes it clear that truthful revelation is obtained by trading-off current consumption for utility promises. As a consequence, no revelation of private information takes place in the last period, since there is no further consumption for screening the agents.[7]

Utility promises will take, in this case, the form

$$w^*(\theta_1, z_1) \equiv (1-\rho)^{-1}\mathbb{E}(\theta) \sum_{z_2 \in Z} \pi_1(z_2) c_2^*(\theta_1, z_1, z_2)^{1-\rho}.$$

In the last period, the only thing inherited from period one is a set of utility promises, which must be fulfilled. Using c^* to denote the (constrained) optimal allocation, we have that, conditional on period 1 promises, consumption in period 2 displays optimal risk sharing relative to the aggregate shock, i. e., for any $(\theta_1, z_1, \theta_1', z_1')$, $c_2^*(\theta_1, z_1, .)$ and $c_2^*(\theta_1', z_1', .)$ are perfectly correlated.

Also, utility promises fully determine the period 2 consumption distribution,

$$c_2^*(\theta_1, z_1, z_2) = z_2 \frac{w^*(\theta_1, z_1)^{\frac{1}{1-\rho}}}{\sum_{\theta_1 \in \Theta} \mu_1(\theta_1) w^*(\theta_1, z_1)^{\frac{1}{1-\rho}}}.$$

[7] Take any fixed $\theta', \theta'' \in \Theta$ and any $(\bar{\theta}, \bar{z}^2) \in \Theta \times Z^2$, then consider $\sigma = (\sigma_1^{TT}, \sigma_2)$ where $\sigma_2(\theta, \theta', \bar{z}^2) = \theta''$ and $\sigma = \sigma^{TT}$ otherwise, then $c_2(\bar{\theta}, \theta', \bar{z}^2) \geq c_2(\bar{\theta}, \theta'', \bar{z}^2)$. Analyzing the reverse deviation we get that $c(\bar{\theta}, \theta', \bar{z}^2) = c_2(\bar{\theta}, \theta'', \bar{z}^2)$. We, then, realize that c_2 can be represented as a function $c_2' : \Theta \times Z^2 \to \mathbb{R}_+$.

Note that last period allocation is separable, and there is perfect aggregate risk insurance, conditional on previous utility promises. Consumption in period 2 is perfectly correlated across agents and its distribution is determined exclusively by period 1 distribution of utility promises.

The direct relationship between utility promises in period 1 and consumption levels in period two imply that we can anticipate resource constraints in the last period by deriving the feasible subset of promises. This allows us to characterize and solve the planner's problem as a static one.

Solving this static planning problem one finds that, for $\rho \neq 1$.:

i) first period consumption is **not** separable on aggregate shock z_1, i. e., there exist **no** functions $\tilde{c}(\theta_1)$ and $\eta(z_1)$ such that $c_1^*(\theta_1, z_1) = \tilde{c}(\theta_1)\,\eta(z_1)$, and;

ii) second period consumption depends on period 1 aggregate shocks, i.e., $c_2 : Z^2 \times \Theta^2 \to \mathbb{R}_+$ cannot be independent of z_1.

Endowment shocks in period 1 could be accommodated by increasing proportionately each agent's consumption, but this is not optimal. Indeed, were we to use a proportional increase, the marginal rate of substitution between consumption in the two periods would vary in a non-linear fashion with z—except in the case $\rho = 1$—which would ultimately alter the incentives/insurance trade-off. Therefore, even though previous aggregate shocks are independent of present information structure of the economy, current consumption distribution may depend on the history of the economy. Finally note how separability in the first period and lack of memory are closely associated, a point we make more forcefully in the Mirrlees economy.

We have assumed idiosyncratic shocks to be i.i.d. Next, we provide a brief heuristic account of how the dependence on aggregate shocks may be influenced by persistence.

Let $\mathbb{E}[\theta'|\theta]$ denote the expected value of second period taste parameter θ' conditional on one's having realized θ in period 1. The marginal rate of substitution between future and current consumption is, then,

$$\left.\frac{dc_2}{dc_1}\right|_{\bar{U}} = \frac{\theta}{\mathbb{E}[\theta'|\theta]}\frac{c_1^{-\rho}}{c_2^{-\rho}}.$$

Monotonicity in θ_t follows immediately from incentive compatibility since single-crossing is obviously valid here. By single-crossing, we mean that, individuals can be ordered by the marginal rate of substitution between present and future consumption. It is also apparent from the expression above that, the higher the persistence, the smaller the difference between marginal rates of substitution between individuals. In particular, when persistence is complete, i.e., $\mathbb{E}[\theta'|\theta] = \theta$, the marginal rate of substitution is identical across agents and no insurance is possible.

Because, there are no transfers at the (constrained) optimum, there is no dependence on history; in particular, on aggregate history. There is a sense in which this result is of limited interest, however: memory with respect to aggregate shocks only disappears when individuals allocations are the ones they have if no transfers occur in any period. This will not be the case in the Mirrlees economy for which insurance is possible even in the last period, through the consumption/leisure tradeoff.

3 Mirrlees Economy

In this section, we consider a dynamic Mirrlees economy, where agents temporary preferences defined as a function of consumption c and effort l is of the form $u\left(c\right) - v\left(l\right)$, where $u\left(.\right)$ strictly increasing and concave and $v\left(.\right)$, strictly increasing and convex.

Now, θ is a productivity parameter which, combined with z, determines how many efficiency units, y, an agent produces per unit of effort, l; an agent with productivity θ produces $y = l\theta z$ with effort l if the aggregate state is z. Because our commodity space is now defined in terms of x, θ can be viewed as a taste shock, just as in the general presentation of the problem in Section 2. The technology is once again linear. One efficiency unit of labor produces one unit of consumption.[8] As in Mirrlees (1971), we redefine choice variables. Instead of (c, l) we consider $x = (c, y)$, where c is consumption and y denotes efficiency units of work.

The Social Planner's problem (program $\mathcal{P}_0$) is

$$\max \sum_t \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t)\mu_t\left(\theta^t\right)\left[u\left(c(\theta^t, z^t)\right) - v\left(\frac{y(\theta^t, z^t)}{\theta_t z_t}\right)\right] \tag{4}$$

[8] To map technology exactly in the terms of Section 2 we would instead define $\tilde{y} = l\theta$ and let

$$G\left(\sum \mu(\theta)c(\theta, z), \sum \mu(\theta)\tilde{y}(\theta, z), z\right) = \sum \mu(\theta)c(\theta, z) - \sum \mu(\theta)\tilde{y}(\theta, z)z \leq 0.$$

As it turns, the definition of variables we adopted is more convenient.

subject to

$$\sum_{\theta^t} \mu_t\left(\theta^t\right)\left[c(\theta^t, z^t) - y(\theta^t, z^t)\right] \leq 0 \tag{5}$$

and

$$\sum_t \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t)\mu_t\left(\theta^t\right)\left[u\left(c(\theta^t, z^t)\right) - v\left(\frac{y(\theta^t, z^t)}{\theta_t z_t}\right)\right] \geq$$
$$\sum_t \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t)\mu_t\left(\theta^t\right)\left[u\left(c\left(\sigma(\theta^t, z^t), z^t\right)\right) - v\left(\frac{y\left(\sigma(\theta^t, z^t), z^t\right)}{\theta_t z_t}\right)\right] \tag{6}$$

for all $\sigma \in \Sigma$.

It is a general feature of repeated screening and moral hazard problems that allocations keep track of individuals' private shocks histories. This allows the principal to go beyond repeating the static optimal allocation by linking future allocations to present type reports.[9]

This whole argument is false for the aggregate shock for one simple reason: it is observable. Since the aggregate state of the economy is assumed to be known by everyone, and is unrelated with idiosyncratic shocks, there is no obvious reason to believe that the optimal allocation will present history dependence with respect to this variable. Our first result makes this point clear by providing an example in which such dependence does not exist.

Proposition 1 *Let $u = \ln$ then there exist functions $\tilde{c}_t(\theta^t)$ and $\tilde{y}_t(\theta^t)$ such that $c_t(\theta^t, z^t) = \tilde{c}_t(\theta^t) z_t$ and $y_t(\theta^t, z^t) = \tilde{y}_t(\theta^t) z_t$.*

Proof. Let $u(c) = \ln(c)$ and consider the following transformation of variables $\tilde{c}(\theta^t, z^t) = c(\theta^t, z^t)/z_t$ and $\tilde{y}(\theta^t, z^t) = y(\theta^t, z^t)/z_t$. In this case we can write program $\mathcal{P}_0$ as

$$\sum_t \beta^{t-1} \sum_{z^t} \pi_t(z^t) \ln z_t + \max \sum_t \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t)\mu_t\left(\theta^t\right)\left[\ln \tilde{c}(\theta^t, z^t) - v\left(\frac{\tilde{y}(\theta^t, z^t)}{\theta_t}\right)\right]$$

[9]The extreme poverty result presented in Atkeson and Lucas (1992) and Phelan (1998) is a long run consequence of this feature of optimal allocations. At each period, the cross-sectional distribution of consumption and labor inherits the heterogeneity of previous period and brings a new source of variation, since you have to generate further distortions to separate agents with different types at the current period. The extra dispersion in allocations is independent of previous heterogeneity, assuming i.i.d. individual shocks, an assumption present in most of the literature. This pattern generates ever increasing inequality across agents.

subject to

$$\sum_{\theta^t} \mu_t\left(\theta^t\right)\left[\tilde{c}(\theta^t, z^t) - \tilde{y}(\theta^t, z^t)\right] \leq 0$$

and

$$\sum_t \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t) \mu_t\left(\theta^t\right)\left[\ln \tilde{c}(\theta^t, z^t) - v\left(\frac{\tilde{y}(\theta^t, z^t)}{\theta_t}\right)\right] \geq$$
$$\sum_t \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t) \mu_t\left(\theta^t\right)\left[\ln \tilde{c}\left(\sigma(\theta^t, z^t), z^t\right) - v\left(\frac{\tilde{y}\left(\sigma(\theta^t, z^t), z^t\right)}{\theta_t}\right)\right]$$

for all $\sigma \in \Sigma$. Note that z_t only appears additively in the objective function, which shows that, with some abuse in notation, $\tilde{c}(\theta^t, z^t) = \tilde{c}(\theta^t), \tilde{y}(\theta^t, z^t) = \tilde{y}(\theta^t)$. ■

The main normative goal of this paper is to find out whether this result applies more generally. As we shall see, this absence of memory is an very particular result. An heuristic account on why this property fails to be optimal in general, hence making memory relevant, goes as follows. Take the general dynamic Mirrlees economy, and define the temporary utility gap, associated with strategy σ at history (θ^t, z^t) as $\Delta^\sigma(\theta^t, z^t)$ as

$$\Delta^\sigma(\theta^t, z^t) = U\left(c(\theta^t, z^t), \frac{y(\theta^t, z^t)}{\theta_t z_t}\right) - U\left(c\left(\sigma(\theta^t, z^t), z^t\right), \frac{y\left(\sigma(\theta^t, z^t), z^t\right)}{\theta_t z_t}\right)$$

Consider the specific case $\Delta^\sigma(\theta^{t-1}, \bar{\theta}, z^{t-1}, z) =$

$$U\left(c(\theta^{t-1}, \bar{\theta}, z^{t-1}, z), \frac{y(\theta^{t-1}, \bar{\theta}, z^{t-1}, z)}{\bar{\theta} z}\right) - U\left(c(\theta^{t-1}, \underline{\theta}, z^{t-1}, z), \frac{y(\theta^{t-1}, \underline{\theta}, z^{t-1}, z)}{\bar{\theta} z}\right),$$

and assume that the constraint associated with the strategy associated with the strategy of telling the truth in all but history $(\theta^{t-1}, \bar{\theta}, z^{t-1}, z)$ binds. In this case, the difference between the continuation utilities must be equal to $\Delta^\sigma(\theta^{t-1}, \bar{\theta}, z^{t-1}, z)$. If the allocations do not display memory, continuation utilities cannot depend on z^t, in this case, either $\Delta(\theta^{t-1}, \bar{\theta}, z^{t-1}, z') = \Delta^\sigma(\theta^{t-1}, \bar{\theta}, z^{t-1}, z)$, or the same constraint cannot bind in the alternative state z'. This imposes a strong restriction on the possibilities of the planner, and will generate sub-optimal allocations, in general.

Let us now specialize to the class of iso-elastic utility functions. Hence, utility functions exhibit constant relative risk aversion in consumption and power disutility of labor effort,

$$u\left(c\right) = \frac{c^{1-\rho}}{1-\rho} \text{ and } v\left(l\right) = \frac{l^\gamma}{\gamma}, \tag{7}$$

for $\rho > 0$, $\rho \neq 1$ and $\gamma > 1$. For $\rho = 1$, $u(c) = \ln(c)$.

Under this iso-elastic preferences assumption, our first result relates memory to separability. To prove it, however, we first present the following lemma which will play an important role in deriving some of our main results.

Lemma 1 *At the optimum,*

$$\sum_{\theta^{t+1} \succ \theta^t} \mu(\theta^{t+1}|\theta^t) \frac{u(c(\theta^t, z^t))}{u'(c(\theta^{t+1}, z^{t+1}))} = \sum_{\bar{\theta}^{t+1} \succ \bar{\theta}^t} \mu(\bar{\theta}^{t+1}|\bar{\theta}^t) \frac{u(c(\bar{\theta}^t, z^t))}{u'(c(\bar{\theta}^{t+1}, z^{t+1}))},$$

for every, θ^t,$\bar{\theta}^t$, z^t and $z^{t+1} \succ z^t$.

Proof. Let $\{c, y\}$ be our candidate optimal allocation. We shall construct a new allocation $\{\hat{c}, \hat{y}\}$ as follows. We increase consumption in period t, public history z^t, for the agent with history θ^t in such a way that $u(\hat{c}(\theta^t, z^t)) = u(c(\theta^t, z^t)) + \Delta$. Because aggregate resources are fixed, this can only be done by reducing consumption for another individual, say, that with history $\bar{\theta}^t$. This agent's utility will be reduced by an amount ε. That is, $u(\hat{c}(\bar{\theta}^t, z^t)) = u(c(\bar{\theta}^t, z^t)) - \varepsilon$. For simplicity, consider $\mu(\theta^t) = \mu(\bar{\theta})$. For small enough Δ, the two are related by

$$\varepsilon = \frac{u'(c(\bar{\theta}^t, z^t))}{u'(c(\theta^t, z^t))} \Delta. \tag{8}$$

For this reform not to have impact on incentives or on any individual's utility, it must compensate the change in utility for every continuation history. Hence, for every $(\theta^{t+1}, z^{t+1}) \succ (\theta^t, z^t)$,

$$u(\hat{c}(\theta^{t+1}, z^{t+1})) = u(c(\theta^{t+1}, z^{t+1})) - \Delta.$$

Similarly, for every $(\bar{\theta}^{t+1}, z^{t+1}) \succ (\bar{\theta}^t, z^t)$,

$$u(\hat{c}(\bar{\theta}^{t+1}, z^{t+1})) = u(c(\bar{\theta}^{t+1}, z^{t+1})) + \varepsilon.$$

For every z^{t+1}, the cost of such reform is

$$-\sum_{\theta^{t+1} \succ \theta^t} \frac{\mu(\theta^{t+1}|\theta^t)}{u'(c(\theta^{t+1}, z^{t+1}))} \Delta + \sum_{\bar{\theta}^{t+1} \succ \bar{\theta}^t} \frac{\mu(\bar{\theta}^{t+1}|\bar{\theta}^t)}{u'(c(\bar{\theta}^{t+1}, z^{t+1}))} \varepsilon. \tag{9}$$

Substituting (8) in (9) it is then clear that, unless

$$\sum_{\theta^{t+1} \succ \theta^t} \mu(\theta^{t+1}|\theta^t) \frac{u'(c(\theta^t, z^t))}{u'(c(\theta^{t+1}, z^{t+1}))} = \sum_{\bar{\theta}^{t+1} \succ \bar{\theta}^t} \mu(\bar{\theta}^{t+1}|\bar{\theta}^t) \frac{u'(c(\bar{\theta}^t, z^t))}{u'(c(\bar{\theta}^{t+1}, z^{t+1}))},$$

there is a reform that preserves utility, preserves incentive compatibility and saves resources. ■

At any given period, individuals care about the current allocation and how the current report will affect their prospects in further periods. Lemma 1, above, states that, since it is always possible to distort future consumption to separate individuals today, differences in current consumption are preserved in the future in such a way as not to interfere with future incentive compatibilities.

Proposition 2 *With iso-elastic preferences, if period t consumption is not separable in z_t, then the allocation displays memory in $t+1$.*

Proof. If an allocation is not separable in t then, given any function $\eta(z)$ such that $c(\theta^t, z^t) = \tilde{c}(\theta^t, z^t)\eta(z_t)$ for some function $\tilde{c}(\theta^t, z^t)$, there must be a pair $\theta^t, \bar{\theta}^t$ such that $\tilde{c}(\theta^t, z^t)/\tilde{c}(\bar{\theta}^t, z^t)$ is a function of z_t. Combining Lemma 1 and Lemma 2, however, gives

$$\frac{\tilde{c}(\theta^t, z^t)^\rho}{\tilde{c}(\bar{\theta}^t, z^t)^\rho} = \frac{\mathbb{E}\left[\tilde{c}(\theta^t, \theta', z^{t+1})^\rho | \theta^t, z^{t+1}\right]}{\mathbb{E}\left[\tilde{c}(\bar{\theta}^t, \theta', z^{t+1})^\rho | \bar{\theta^t}, z^{t+1}\right]}. \tag{10}$$

Since the left hand side depends on z^t, so does the right hand side. Hence, either the numerator or the denominator varies with z^t, thus violating independence. ■

One interesting and straightforward consequence of (10) result is the case in which $u(\cdot)$ is a natural log. Consider $u = \ln$, then, letting $Y\left(z^t\right)$ denote total production, we have

$$\frac{c_t\left(\bar{\theta}^t, z^t\right)}{Y\left(z^t\right)} = \frac{\sum_{\theta'} \mu\left(\bar{\theta}^t, \theta_t | \bar{\theta}^t\right) c_{t+1}\left(\left(\bar{\theta}^t, \theta_t\right), \left(z^t, z\right)\right)}{\sum_{\theta^{t+1}} \mu\left(\theta^{t+1}\right) c_{t+1}\left(\theta^{t+1}, \left(z^t, z\right)\right)} = \frac{\mathbb{E}\left[c_{t+1} | \bar{\theta}^t, z^{t+1}\right]}{Y\left(z^{t+1}\right)}.$$

The fact that the expectation in the right hand side of the expression above is conditioned on z^{t+1} allows for a simple interpretation of the expression above. Namely, the expected share of aggregate output an agent gets to consume in any aggregate state tomorrow is equal to the share of current output that she consumes today. This means that, two agents are entitled to different shares of total output today, they will necessarily receive different shares of total output for at least one realization of private shocks and this will be true for all aggregate states.

More importantly for our discussion is the fact that, if the share of aggregate output that an agent with private history θ^t is entitled to in aggregate state z_t differs from the share she would be entitled to had state $\hat{z}_t$ realized, instead, then the expected share she would be entitled to in state (z^{t-1}, z_t, z) would differ from the expected share she is going to get in $(z^{t-1}, \hat{z}_t, z)$.

The consequence of Proposition 2 is that, with logarithmic preferences, it is either the case that optimal shares are independent of z or allocations must display memory with respect to aggregate shocks. Proposition 1 shows that it is the former, rather than the latter.

To check whether this result extends to other values of ρ, we make an additional assumption on the nature of the stochastic process associated with the idiosyncratic shock, namely we assume that it is a Markov process:

$$\mu(\theta^{t-1}, \theta, \theta' | \theta^{t-1}, \theta) = \mu(\theta, \theta' | \theta), \forall \theta^{t-1}$$

We shall also treat the case of perfect persistence: $\mu(\theta|\theta) = 1$ separately—Section 3.1. For now assume that $\mu(\theta|\theta) \in [.5, 1)$. Finally, in what follows we shall also restrict $T < \infty$.

Define $w(\theta^{t-1}, z^t)$ recursively as the expected continuation utility for an agent with history θ^t conditional on aggregate history z^{t+1},

$$\begin{aligned} w(\theta^{t-1}, z^t) = \sum_{\theta^t \succ \theta^{t-1}} \mu_t\left(\theta^t | \theta^{t-1}\right) \Bigg[& u\left(c(\theta^t, z^t)\right) - v\left(\frac{y(\theta^t, z^t)}{\theta_t z_t}\right) \\ & + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) w(\theta^t, z^{t+1}) \Bigg] \end{aligned} \tag{11}$$

Similarly, let $\bar{\theta}^{t-1} = (\theta^{t-2}, \bar{\theta}) \neq (\theta^{t-2}, \theta) = \theta^{t-1}$. We may, then, define

$$\begin{aligned} \hat{w}(\theta^{t-1}, z^t) = \sum_{\theta^t \succ \theta^{t-1}} \mu_t\left(\theta^t | \bar{\theta}^{t-1}\right) \Bigg[& u\left(c(\theta^t, z^t)\right) - v\left(\frac{y(\theta^t, z^t)}{\theta_t z_t}\right) \\ & + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) w(\theta^t, z^{t+1}) \Bigg] \end{aligned} \tag{12}$$

Following the steps of Fernandes and Phelan (2000) we can show—Lemma 4, in the appendix—that any incentive compatible allocation must satisfy the one period incentive constraint,

$$\begin{aligned} u\left(c(\theta^t, z^t)\right) - v\left(\frac{y(\theta^t, z^t)}{\theta_t z_t}\right) + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) w(\theta^t, z^{t+1}) \quad \geq \\ u\left(c(\theta^{t-1}, \bar{\theta}, z^t)\right) - v\left(\frac{y(\theta^{t-1}, \bar{\theta}, z^t)}{\theta_t z_t}\right) + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) \hat{w}(\theta^{t-1}, \bar{\theta}, z^{t+1}) \quad . \end{aligned} \tag{13}$$

To prove our main result we shall exploit these temporary incentive constraints. The idea is that, since the solution to the government's program is incentive compatible, it must satisfy the temporary constraints, (13). Now assume that the allocation does not exhibit memory, in which case, neither $w(\theta^t, z^{t+1})$ nor $\hat{w}(\hat{\theta}^t, z^{t+1})$ depend on z^t. For the $u = \ln$ case, under the transformation of variables used in the proof of Proposition 1, the incentive constraint becomes

$$\begin{aligned} \ln \tilde{c}(\theta^t, z^t) z_t - v\left(\frac{\tilde{y}(\theta^t, z^t)}{\theta_t}\right) + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) w(\theta^t, z_{t+1}) \quad &\geq \\ \ln \tilde{c}(\hat{\theta}^t, z^t) z_t - v\left(\frac{\tilde{y}(\hat{\theta}^t, z^t)}{\theta_t}\right) + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) \hat{w}(\hat{\theta}^t, z_{t+1}) & \end{aligned}$$

It is easy to see that z_t plays no role in providing incentives. Because z plays no role in the redefined resource constraint and is additive in the objective function it does not change the planner's optimization problem.

In the case of $\rho \neq 1$ another convenient change of variables makes the resource constraint and the objective function independent of z in the iso-elastic, $\rho \neq 1$ case. Namely, $c(\theta^t, z^t) = \tilde{c}(\theta^t, z^t)\eta(z_t)$, and $y(\theta^t, z^t) = \tilde{y}(\theta^t, z^t)\eta(z_t)$ where $\eta(z) = z^{\frac{\gamma}{\rho+\gamma-1}}$. The fact that it is multiplicative, however, makes $\eta(z_t)$ relevant for incentives.

In fact, under the assumption that the allocation is memoryless and using the suggested transformation of variables, the one period incentive constraint becomes

$$\begin{aligned} \eta(z_t)^{1-\rho} \left\{ \left[\frac{\tilde{c}(\theta^t, z^t)^{1-\rho}}{1-\rho} - \frac{\tilde{y}(\theta^t, z^t)^{\gamma}}{\theta^{\gamma}} \right] - \left[\frac{\tilde{c}(\hat{\theta}^t, z^t)^{1-\rho}}{1-\rho} - \frac{\tilde{y}(\hat{\theta}^t, z^t)^{\gamma}}{\theta^{\gamma}} \right] \right\} \quad &\geq \\ \beta \left\{ \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) \hat{w}(\hat{\theta}^t, z_{t+1}) - \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) w(\theta^t, z_{t+1}) \right\} \quad &. \end{aligned}$$

Assume that this expression to holds as an equality for both $z_t = \bar{z}$ and $z_t = \underline{z}$. In this case, because the right hand side is independent of z_t, the term in curly brackets in the left hand side must depend on z_t, which implies that period t allocation is not separable. This, however contradicts Proposition 2.

For the result to be valid, among other things we need to prove that there is at least one period t and one private history, θ^t such that the same one period deviation binds at the optimum at both states. This is a very weak condition, which does seem to be valid in general. Still, for now we shall just assume that this is the case.

Assumption A: At the optimum, for some $t < T$ there is a history (θ^{t-1}, z^{t-1}) such that $U(x(\theta^{t-1}, \theta, z^{t-1}, \bar{z})) = U(x(\theta^{t-1}, \tilde{\theta}, z^{t-1}, \bar{z})|\theta^{t-1}, \theta)$ and $U(x(\theta^{t-1}, \theta, z^{t-1}, \underline{z})) = U(x(\theta^{t-1}, \tilde{\theta}, z^{t-1}, \underline{z})|\theta^{t-1}, \theta)$.

Under this assumption it is possible to prove the following result.

Proposition 3 *Assume that $\rho \neq 1$ and that persistence is less than complete, then if Assumption A is valid the optimal allocation exhibits memory of past aggregate shocks.*

Proof. See Appendix. ■

Assumption A is valid if shocks are i.i.d. Hence, the following corollary is an immediate consequence of Proposition 3.

Corollary 1 *Assume that $\rho \neq 1$ and idiosyncratic shocks are i.i.d., then the optimal allocation exhibits memory of past aggregate shocks.*

Proposition 2 states that, for an allocation to be memoryless in t it must be separable in $\tau < t$, not that a separable allocation is necessarily memoryless. We cannot, therefore, infer from Proposition 3 whether the allocations associated with $\rho \neq 1$ are separable or not. For that we shall rely on our numeric exercises. The next Proposition, however, shows that separability always characterizes the period T allocation.

Proposition 4 *Period T allocation is separable. I.e., there are functions $\eta(z_T)$, $\tilde{c}(\theta^T, z^{T-1})$ and $\tilde{y}(\theta^T, z^{T-1})$ such that $c(\theta^T, z^T) = \tilde{c}(\theta^T, z^{T-1})\eta(z_T)$ and $y(\theta^T, z^T) = \tilde{y}(\theta^T, z^{T-1})\eta(z_T)$*

Proof. See Appendix. ■

Next we address the case of perfectly persistent shocks.

3.1 Perfect persistence

If shocks are permanent, i.e., if there is no uncertainty beyond the productivity one is born with, the problem becomes a particular case of Werning (2007). Indeed, let $\mu_{t+1}\left(\theta^{t-1}, \theta', \theta|\theta^{t-1}, \theta'\right) = 1$, $\theta = \theta'$, i.e., individual productivity levels are constant. Then, the Planner's problem is

$$\max \sum_{\theta} \mu\left(\theta\right) \sum_{t} \beta^{t-1} \sum_{z^t} \pi_t(z^t) \left[u\left(c(\theta, z^t)\right) - v\left(\frac{y(\theta, z^t)}{\theta z_t}\right)\right]$$

subject to

$$\sum_{\theta} \mu(\theta) \left[c(\theta, z^t) - y(\theta, z^t) \right] \leq 0$$

and

$$\sum_t \beta^{t-1} \sum_{z^t} \pi_t(z^t) \left[u\left(c(\theta, z^t)\right) - v\left(\frac{y(\theta, z^t)}{\theta z_t}\right) \right] \geq$$
$$\sum_t \beta^{t-1} \sum_{z^t} \pi_t(z^t) \left[u\left(c(\bar{\theta}, z^t)\right) - v\left(\frac{y(\bar{\theta}, z^t)}{\theta z_t}\right) \right]$$

for $\theta, \bar{\theta} = \theta(H), \theta(L)$.

It is important to note that there are only two incentive compatibility constraints, which are independent of time and aggregate state. We associate to the IC constraints the multipliers $\phi\left(\bar{\theta}|\theta\right)$ and $\phi\left(\theta|\bar{\theta}\right)$. The first order conditions associated with $c_t(\theta, z^t)$ and $y_t(\theta, z^t)$ are

$$\begin{aligned}
\left[1 + \sum_{\bar{\theta}} \left[\phi\left(\bar{\theta}|\theta\right) - \phi\left(\theta|\bar{\theta}\right)\right]\right] c_t\left(\theta, z^t\right)^{-\rho} &= \lambda\left(z^t\right), \\
\left[1 + \sum_{\bar{\theta}} \left[\phi\left(\bar{\theta}|\theta\right) - \phi\left(\theta|\bar{\theta}\right) \frac{\theta^\gamma}{\bar{\theta}^\gamma}\right]\right] \frac{y_t\left(\theta, z^t\right)^{\gamma-1}}{\left(\theta z_t\right)^\gamma} &= \lambda\left(z^t\right).
\end{aligned}$$

Hence, $c_t\left(\theta, z^t\right) = \lambda\left(z^t\right)^{\frac{-1}{\rho}} \kappa(\theta)$ and $y_t\left(\theta, z^t\right) = (\theta z_t)^{\frac{\gamma}{\gamma-1}} \lambda\left(z^t\right)^{\frac{1}{\gamma-1}} \omega(\theta)$, for some κ and ω. Substituting this in the resource constraints,

$$\sum \mu(\theta) \left[\lambda\left(z^t\right)^{\frac{-1}{\rho}} \kappa(\theta) - (\theta z_t)^{\frac{\gamma}{\gamma-1}} \lambda\left(z^t\right)^{\frac{1}{\gamma-1}} \omega(\theta)\right] = 0,$$

which means that $\lambda\left(z^t\right) = a z_t^{-\frac{\gamma\rho}{\rho+\gamma-1}}$. Therefore,

$$c_t\left(\theta, z^t\right) = \tilde{c}(\theta)\, z_t^{\frac{\gamma}{\rho+\gamma-1}}, y_t\left(\theta, z^t\right) = \tilde{y}(\theta)\, z_t^{\frac{\gamma}{\rho+\gamma-1}}.$$

Separability thus characterizes this pure heterogeneity case. As we shall se in the next section this has important consequences for the possibility of heterogeneity generating any type of interesting business cycle behavior.

4 Macro Consequences of Private Memory

Although most of the New Dynamic Public Finance Literature has been of a normative nature, a few attempts have been made to describe and test the macroeconomic 'predictions of those models' e.g., Kocherlakota and Pistaferri (2007), Kocherlakota and Pistaferri (2008), Kocherlakota and Pistaferri (2009), Ales and Maciero

(2008) and Sleet (2006). The underlying idea is that, because many different and sometimes informal institutions can implement the optimal allocations, one cannot 'a priori' rule out the possibility that current institutions are constrained efficient.

When focus is shifted from policies to allocations, one possible path is to use partial characterizations of optimal allocations to create testable restrictions in the observed data. This type of approach has been the focus of an important literature that has made intensive use of micro data: Townsend (1994), Ligon et al. (2002) etc.[10] Another path, the one we shall follow here, is to derive the macro consequences of these models. By macro consequences we simply mean the consequences for aggregate variables and prices. In taking this path we follow the lead of Kocherlakota and Pistaferri (2007) and Kocherlakota and Pistaferri (2009) who have used the pricing kernel associated with a dynamic Mirrlees economy to try and price the excess return of equity over risk free bonds and the conditional return on forward exchange rate.

We follow this agenda by addressing the business cycle behavior of the labor wedge. In a recent work, Chari et al. (2007) have shown how equilibrium allocations for disaggregated model economies are equilibrium allocations for a representative agent economy with suitably defined time-varying wedges. They use this observation to identify which particular wedges are more important to generate the observed behavior of macroeconomic aggregates in the representative agent models. Their results highlight the fact that for a disaggregated model to account for the business cycle behavior of macroeconomic aggregates must generate a countercyclical wedge in the associated representative consumer economy.

One of our goals in our numeric exercises is to check whether our model is capable of generating such countercyclical wedges. The current model is particularly promising in this regard for it endogenously generates wedges at an individual level. Whether these individual wedges imply a time-varying labor wedge for the representative agent,and whether this variation follows the pattern found in the data is one of the questions we try to answer in our numeric exercises. Before moving on to those results, some definitions and preliminary results are, however, needed.

We start by defining the aggregate variables we are going to use. Aggregate consumption is $C(z^t) \equiv \sum \mu_\Theta(\theta^t) c(\theta^t, z^t)$, while output is $Y(z^t) \equiv \sum \mu_\Theta(\theta^t) y(\theta^t, z^t)$.

[10] A very entertaining presentation of this approach is found in Townsend (1993)

Because there is no aggregate savings in our economy, $C(z^t) = Y(z^t)$ for all z^t.

For our purposes it will also be convenient to define a *backloading* measure. The idea to define backloading is similar to the one that allows us to define the seed values in Fernandes and Phelan (2000) auxiliary problem. Assume that the individual has learnt his type. What is the expected utility that she gets from the government utility maximization problem? This is the seed value. The measure of backloading we shall use here assumes that we calculate this starting one period ahead and assuming that nothing happened in the first period, which we call the 'one step ahead seed'. The seed and the one step ahead seed coincide up to a constant in the infinite horizon case, but will differ, in general, for a finite horizon. Our backloading measure is calculated by taking the difference between this 'one-step-ahead seed' and the utility promise made after the first period in the actual mechanism.

This measure is very easy to illustrate in the two period case that will be the focus of our numeric exercises. The 'one step ahead seed' is simply the utilitarian optimum with the relative measures of the two types given by the probabilities associated with each different type. One important characteristic with this measure is that it compares the optimal utility promise with another feasible one. As a consequence, this one step ahead seed can also be used to bound the expected utilities associated with the optimal allocations.

Backloading varies in opposite way with the business cycle for the two types. What our numeric results show is that whether backloading is pro-cyclical for the high or the low productivity agent depends on the risk aversion. It is pro-cyclical for the high productivity agent (hence, countercyclical for the low productivity agent) when $\rho = .5$ and countercyclical for the high productivity agent (hence, pro-cyclical for the low productivity agent) when $\rho = 5$.

4.1 Asset Pricing

Kocherlakota and Pistaferri (2007) and Kocherlakota and Pistaferri (2009) have shown how the pricing kernel associated with a dynamic Mirrlees economy is compatible with the behavior of excess returns in foreign and domestic markets. Using the fact that the pricing kernel is a function of the cross-sectional harmonic mean of marginal consumption—as derived by Kocherlakota (2005)—they use cross-sectional consumption data to estimate the preference parameters associated with a CRRA specification for three economies. Kocherlakota and Pistaferri (2009) show that a low coefficient of relative risk aversion is needed to account for the equity premium found in the data, while Kocherlakota and Pistaferri (2007) find similar low coefficients can account for the forward exchange rate premium.

A natural next step in this agenda is to calibrate such models to generate pricing data compatible with the underlying primitives of the economy.[11] It, thus, becomes important to understand how private information regarding idiosyncratic shocks interact with public aggregate ones to generate optimal allocations from which the pricing kernels are derived. This paper considers very simple two period economies that display these elements to explore how the dynamic nature of incentives induces persistence in allocations with respect to aggregate shocks. Hence, we are not doing a true calibration exercise, but pointing out the potential results to expect from such exercises.

First, to focus on asset pricing, we may define the 'representative agent's pricing kernel', $\Lambda(z^{t+1}) = C(z^{t+1})^{-\rho}/C(z^t)^{-\rho}$. Following the idea of using parameters estimated from micro-data to calibrate the model, we may compare Λ with the true pricing kernel,

$$Q(z^t|z^{t-1}) = \left[\sum_{\theta^t \succ \theta^{t-1}} \mu_\Theta(\theta^t|\theta^{t-1}) \frac{c(\theta^t, z^t)^\rho}{c(\theta^{t-1}, z^{t-1})^\rho}\right]^{-1}.$$

As we have shown, last period allocation exhibits a form of separability with respect to last period shock that has interesting consequences for the form the kernel varies with last period's shock. Another interesting feature has to do with the fact that heterogeneity per se has consequences for the magnitude of the pricing kernel, but not with the way in which it varies with the states of nature.

Using the pricing kernel $Q(z^t|z^{t-1})$ we can also derive the return of a risk free assets and the expected return of stocks.[12]

[11]Grossly speaking, while their exercises in akin to Hansen and Singleton (1982), which looked at some of the moment restrictions that needed to prevail in equilibrium a calibration exercise would be akin to Mehra and Prescott (1985), which callibrated the full economy to see if the model was capable of generating the behavior of asset prices anywhere near what is seen in the data.

[12]We take the economy's GDP to represent the stock's dividends. This idea has been recently criticized by, among others, advocates of long run risks as being the relevant cause of most problems with the Consumption Capital Asset Pricing Model, e,g, Bansal and Yaron (2004).

4.2 The Labor Wedge

The second issue we shall address is the labor wedge. Chari et al. (2007) show the important role played by the so-called labor wedge in generating the observed pattern of 'choices' for the representative agent through the business cycles. Many potential explanations for these wedges are considered in Shimer (2009) . We ask whether the underlying informational problems that characterize a repeated Mirrlees economy should be seriously considered as a potential explanation.

Chari et al. (2007) and Shimer (2009) point out to the fact that the labor wedge increases in economic downturns and decreases in good times. A dynamic Mirrlees economy may potentially allow one to take the real business cycle literature idea that fluctuations are optimal responses to real technological shock one step further and ask whether staty-varying labor wedges are optimal (constrained efficient, in this case) responses to real shocks in a world where asymmetric information constrains the possibility of full insurance.

Note that we are not siding with the view of the strand of the literature assumes that tax shocks *cause* business cycle, hence the observed correlation between recessions and labor wedge. Instead, in our model state varying wedges (are they to exist) are optimal responses to real shocks in a second best world.

To derive the economy's productivity, $W(z^t)$, we need first to define aggregate hours,

$$L(z^{t-1}, z) \equiv \sum_{\theta^{t-1}} \sum_{\theta} \mu_\Theta(\theta^{t-1}, \theta) \frac{y(\theta^{t-1}, \theta, z^{t-1}, z)}{\theta z}$$

and, finally, $W(z^t) = Y(z^t)/L(z^t)$.

We shall endow the representative agent with the same preferences as individuals. Hence,[13]

$$\mathcal{W}(C(z), Y(z), W(z)) = \frac{C(z)^{1-\rho}}{1-\rho} - \frac{Y(z)^\gamma}{\gamma W(z)^\gamma}.$$

Note also that without taxes, the representative agent's optimal labor/consumption choice would be given by $W(z)^\gamma C(z)^{-\rho} = Y(z)^{\gamma-1}$. Naturally, there are taxes, both in the model and in the real world, and this equality should not be observed in practice. We shall, then, define the labor wedge,

$$\tau(z) = 1 - W(z)^{-\gamma} Y(z)^{\rho+\gamma-1},$$

[13] In the spirit of a calibration exercise in which instead of allowing for free parameters, we take them from micro data. See Chari et al. (2009).

where we have used the fact that $C(z) = Y(z)$.[14]

4.3 Two Particular cases

Before moving on to the numeric exercises, we consider two specifications for our model that render very sharp predictions, which, however, eliminate most of the interesting aspects of our model: the case or perfectly persistent shocks, i.e., the case where heterogeneity but not idiosyncratic risks play a role, and the case of log preferences.

4.3.1 Perfectly Persistent Shocks

Our first result is that, in a repeated Mirrlees economy, which we have been calling the perfect persistence case, allocation is separable. The first consequence of separability is that the wedge is constant across states of nature.

We have seen that optimal allocations are of the form $c(\theta, z^t) = \tilde{c}(\theta)\eta(z_t)$, and $y(\theta, z^t) = \tilde{y}(\theta)\eta(z_t)$. In this case, $Y(z^t) = \eta(z_t) \sum \mu_\Theta(\theta^t)\tilde{y}(\theta^t)$. Hours, are, then

$$L(z^t) = \frac{\eta(z_t)}{z_t} \sum \mu(\theta) \frac{\tilde{y}(\theta)}{\theta},$$

while the estimated productivity is

$$W(z^t) = \frac{Y(z^t)}{L(z^t)} = z_t \frac{\sum \mu(\theta)\tilde{y}(\theta)}{\sum \mu(\theta)\tilde{y}(\theta)/\theta}.$$

Deriving the wedge is now a simple task

$$\tau(z^t) = 1 - \left\{ \frac{\sum \mu(\theta)\tilde{y}(\theta)/\theta}{\sum \mu(\theta)\tilde{y}(\theta)} \right\}^\gamma \left\{ \sum \mu(\theta)\tilde{y}(\theta) \right\}^{\rho+\gamma-1},$$

where we used the fact that $\eta(z_t)^{\rho+\gamma-1} z_t^\gamma = 1$.

As is clear from the right hand side of the expression above, the labor wedge is invariant with respect to z. We had already shown that labor wedges for each individual do not vary with z. This does not imply that the aggregate wedge will not vary. Changes in the distribution of income may generate such variation through composition effects.

The message we get from this case is that simply allowing for heterogeneity will not do. We need partially insured private risks to generate variation in wedges. In the next section we numerically explore how wedges vary with z.

[14] footnote discussing some aggregation issues

As for asset pricing note that the pricing kernel is, in the case of separability, and perfect persistence,

$$Q(z^t|z^{t-1}) = \frac{\eta(z_{t-1})^\rho}{\eta(z_t)^\rho}, \tag{14}$$

which is the pricing kernel associated with a representative consumer economy.

4.3.2 The Log case

As we have seen in Section 3, when preferences are of the form $u(c) = \ln(c)$, allocations take the form $(c(\theta^t, z^t), y(\theta^t, z^t)) = (\tilde{c}(\theta^t), \tilde{y}(\theta^t))z_t$. Using the same argument of Section 4.3.1, separability implies that the wedge does not depend on z_t. As for return on assets, note that,

$$Q(z^t|z^{t-1}) = \left[\sum_{\theta^t \succ \theta^{t-1}} \mu(\theta^t|\theta^{t-1}) \frac{\tilde{c}(\theta^t)}{\tilde{c}(\theta^{t-1})}\right]^{-1} \frac{z_{t-1}}{z_t}.$$

Now, stock returns are given by

$$\frac{C(z^t)}{Q(z^t|z^{t-1})} = z_{t-1} \left[\sum_{\theta^t \succ \theta^{t-1}} \mu(\theta^t|\theta^{t-1}) \frac{\tilde{c}(\theta^t)}{\tilde{c}(\theta^{t-1})}\right]^{-1} \sum_{\theta^t} \mu(\theta^t)\tilde{c}(\theta^t),$$

which is independent of z_t.

5 Numeric Exercises

In this section we conduct numerical exercises that help illustrate some of the properties that we have derived in Section 2. This exercise is similar to the one presented in Golosov et al. (2006), in which a two period Mirrlees economy with government spending shocks, that are public information. However, they have a different focus and only characterize how last period shocks affect optimal labor tax wedges. More precisely, Golosov et al. (2006) do not allow public shocks in the first period, which would allow them to address the persistence issues that is the focus of our work.

Kocherlakota (2005) provides a numerical example that is similar to ours in many ways, but analyzes the impact of individual shock persistence and the size of the public shocks on capital taxation, which is the main focus of his paper. He does not mention labor tax nor the possibility of persistent effects of aggregate shocks.

In our exercises we fix $\gamma = 2$, $\theta(L) = 1$, $\theta(H) = 2$, $\mu_1(L) = \mu_1(H) = .5$, $\pi_1(\underline{z}) = \pi_1(\overline{z}) = .5$, $\underline{z} = 1$ and $\overline{z} = 2$. Our measure of persistence, $\pi \in [0,1]$, is the value of the non-unit eigenvalue of the transition matrix for idiosyncratic shocks.[15] The probability of maintaining the same productivity in the next period, i. e., $\mu_\Theta(\theta, \theta|\theta) = \pi$ for all θ is directly associated with our measure in this two type context. In particular, because we vary π in $[0.5, 1]$, our persistence measure ranges between 0 and 1: where 0 persistence is i.i.d. case and 1 is the case of perfectly persistent shocks.

We analyze how relevant individual shock persistence is on the allocation and what is driving the non-existence of persistence of aggregate shocks effects in the model for the perfectly persistence case.

The first set of figures shows how consumption, output and marginal tax rates vary in each state for low and high type agents according to how persistent is the shock. For this set of figures we have chosen $\rho = 5$

Not unexpectedly, consumption is increasing in aggregate productivity for all agents. Also, in both states, agents who realize a high type have more consumption than agents who realize a low type. Low productivity individuals face a positive marginal tax rate, while high productivity individuals face a 0 marginal tax rate (we, thus, omit marginal tax rates for high types from the figures). This reproduces the classic result of a static Mirrlees economy, and indicates that only downward constraints bind at the optimum for both states.

Our emphasis in those figures is on the way allocations vary with persistence. It is apparent that while consumption for low productivity individuals decrease with persistence, it increases with persistence for a high productivity agent. The more persistent a shock is, the less one gains from backloading incentives. For the exact same reason, income displays the opposite behavior, although it varies with persistence much less significantly than income. A consequence of this behavior for consumption and income is that temporary utility declines with persistence for low productivity individuals and increases for high productivity individuals. It happens so dramatically that for very low levels of persistence the temporary

[15]We use the term transition matrix in the traditional sense for Markov chains. In our case the Markov nature of idiosyncratic shock is trivial.

utility is higher for the low productivity individuals.[16] Of course, for this type of utility reversion to take place in an incentive compatible way, promised utilities behave in the exact opposite way.

The second set of figures is intended to show the presence of memory. We compare, in percentage terms, consumption in each possible second period state with consumption in the first period. That is, for each type θ^2 and each second period aggregate state z_2, we calculate $c(\theta^2, \bar{z}, z_2)/c(\theta^2, \underline{z}, z_2)$ Consumption growth declines with persistence for those who realized high productivity in the first period and increases with persistence for those who realized a low type.

Allocations display memory of previous aggregate shocks, in our simple model. Second period consumption depends on first period aggregate shock. Better aggregate states in the past are associated with smaller differences in consumption of low and high type, which points out to the fact that the value of backloading incentives in the first period is higher in bad economic states. Given the extreme variations in aggregate productivity that we have assumed, memory is apparent. For more realistic variations in aggregate productivity, memory does not seem to be quantitatively relevant.

In the third set of figures we show how backloading varies with persistence and across states of nature. Also important for the discussions that will follow, we compare backloading for the cases $\rho = 5$ (figures in the left) and $\rho = 1/2$ (figures in the right). As expected, our backloading measure is positive for the high productivity individual, which means that he defers some of his utility to the second period, and negative for the low productivity individual, meaning that he borrows from future utility. Another interesting although anticipated aspect is that backloading decreases in absolute value with persistence.

What is more of a novelty is the way in which backloading varies across states of nature for the two different levels of risk aversion. While backloading is more intense in the high state for $\rho = 5$ it is less intense for the $\rho = 1/2$. As a consequence, more redistribution is made using temporary utility in the high state for $\rho = 1/2$. The practical consequence is shown in the next set of figures which compare marginal tax rates for the two cases in the first period.

It is apparent that marginal tax rates are larger in the low state for $\rho = 5$ and

[16] To make it sound a little less paradoxical, recall that the utility for a low type is higher in the first best.

smaller for $\rho = 1/2$.

The last set of figures deals with the macro consequences of our micro facts. The single most important aspect is that labor wedges vary with output in the first period, provided that persistence is not full.

As we have shown they do not vary with current (although they do vary with previous) output in the second period, but it does vary with past aggregate shock. When $\rho = 1/2$ the labor wedge varies in the opposite direction from what has been documented; see Shimer (2009). That is, wedges are pro-cylcical. When, however, $\rho = 5$, labor wedges are countercyclical, in agreement with the stylized facts about business cycles.

Finally, with regards to asset pricing the representative agent mis-prices both, the risk free bond and the stock if persistence is not full. The amount of mis-pricing does not seem to be very important when compared to the absolute size of the returns we are considering. Once again, differences in the behavior of a representative consumer when compared to our model depends on the existence of idiosyncratic risk. Finally note that the model does generate state varying risk premia, but very similar to what we would observe with a representative agent.

6 Conclusion

Our numeric exercises show that a countercyclical labor wedge may characterize constrained efficient allocations in a dynamic Mirrlees economy. One important aspect of our findings is that we have imposed complete independence between private and aggregate shocks. All movements in labor wedges are, thus, endogenously generated.

Asset pricing implications of our economy do not seem to diverge too much from the aggregate consumer's one. This result, which is in contrast with Kocherlakota and Pistaferri (2007), Kocherlakota and Pistaferri (2008) and Kocherlakota and Pistaferri (2009), should be taken with some caution. First, as we have emphasized in the previous paragraph, the distribution of productivity is independent of aggregate shock, in our model. Second, we take dividends to be a claim to GDP, a procedure that has been under increasing criticism.

The model presented in the numeric example is very simple with only two periods and states of the world. There is still much to be done to advance our comprehension of the intricate relationship between constrained optimal insurance under aggregate uncertainty and its macroeconomic implications.

This should not be an easy task. As we have shown here, the business cycle movements of labor wedge—at least in the separable iso-elastic preference case—depend on a form of non-separability that is associated with the presence of memory with respect to aggregate shocks. In other words, for the model to endogenously produce any interesting business cycle pattern, constrained efficient allocations must depend on aggregate shocks in a non-trivial fashion. This type of memory makes it harder to generate a stationary environment where a solution to a fully dynamic model can be handled with well known computational techniques.[17]

Ours is just another step toward evaluating the potential gains from assessing the importance of endogenous market incompleteness in understanding macroeconomic patterns. Although our preliminary findings are not quantitatively meaningful, one interesting lesson we get from them is that counter-cyclical business cycle wedges and risk premia need not suggest inefficiencies in a second best world.

References

Ales, Lawrence and Pricila Maciero, "Accounting for private information," Working papers, University of Minnesota 2008.

Atkeson, Andrew and Robert E. Lucas, "On Efficient Distribution with Private Information," *Review of Economic Studies*, 1992, *59*, 427–453.

Bansal, Ravi and Amir Yaron, "Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, August 2004, *59* (4), 1481–1509.

[17] Using a perpetual life framework, Phelan (1994) focused on showing how some simplifying assumptions (CARA preferences are crucial here) allows for the use of recursive methods to a setting with both idiosyncratic and aggregate shocks. His analysis does provide results regarding the interaction between the two types of shocks. More recently Scheuer (2009) and Demange (2008) study the interaction between risk sharing and moral hazard. Both Scheuer (2009) and Demange (2008) models are static, which means that they cannot address memory. We must also mention the numerical exercises conducted by Golosov et al. (2006) in a two period Mirrlees economy with aggregate risk. Golosov et al. (2006) do not exploit models that generate memory of aggregate shocks. Also related are the numerical exercises found in Kocherlakota (2005).

Chari, Varadarajan V., Patrick J. Kehoe, and Ellen R McGrattan, "Business Cycle Accounting," *Econometrica*, May 2007, *75* (3), 781–836.

_ , _ , and _ , "New Kenesian Models: Not Yet Useful for Policy Analysis," *American Economic Journal: Macroeconomics*, 2009, *1* (1), 242–266.

Cochrane, John H., *Asset pricing*, Princeton University Press, 2001.

Demange, Gabrielle, "Sharing Aggregate Risk Under Moral Hazard," Technical Report 2008-27, Paris School of Economics 2008.

Fernandes, Ana and Christopher Phelan, "A Recursive Formulation for Repeated Agency with History Dependence," *Journal of Economic Theory*, 2000, *91*, 223–247.

Golosov, Mikhail, Aleh Tsyvinski, and Iván Werning, "New Dynamic Public Finance: A User's Guide," *NBER Macroeconomics Annual 2006*, 2006.

Green, Edward, "Contractual Arrangements for Intertemporal Trade," in Edward Prescott and Neil Wallace, eds., *Lending and the Smoothing of Uninsurable Income*, Minneapolis: University of Minnesota Press, 1987.

Hansen, Lars P. and Kenneth Singleton, "Generalized instrumental variables estimation of nonlinear expectations models," *Econometrica*, 1982, *50* (5), 1269–1286.

Holmström, Bengt, "Moral Hazard in Teams," *Bell Journal of Economics*, 1982, *10* (0), 74–91.

Kocherlakota, Narayana, "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica*, 2005, *73*, 1587–1621.

_ and Luigi Pistaferri, "Household Heterogeneity and Real Exchange Rates," *Economic Journal*, 2007, *117*, C1–C25.

_ and _ , "Inequality and Real Exchange Rates," *Journal of The European Economic Association*, 2008, *6*.

_ and _ , "Asset pricing implications of Pareto optimality with private information," *Journal of Political Economy*, 2009, *0*, forthcoming.

Levin, Jonathan, "Multilateral Contracting and the Employment Relationship," *Quarterly Journal of Economics*, 2002, *117* (3), 1075–1103.

Ligon, Ethan, Jonathan Thomas, and Tim Worrall, "Informal Insurance Arrangements with Limited Commitment: Theory and Evidence from Village Economies," *Review of Economic Studies*, 2002, *69*, 209–244.

Mehra, Rajnish and Edward C. Prescott, "The equity risk premium: A puzzle," *Journal of Monetary Economics*, 1985, *15* (2), 145–61.

Mirrlees, James A., "An Exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies*, 1971, *38*, 175–208.

Phelan, Christopher, "Incentives and Aggregate Shocks," *The Review of Economic Studies*, 1994, *61*, 681–700.

Phelan, Cristopher, "On the Long-Run Implications of Repeated Moral Hazard," *Journal of Economic Theory*, 1998, *79*, 174–191.

Rogerson, William P., "Repeated Moral Hazard," *Econometrica*, 1985, *53* (1), 69–76.

Scheuer, Florian, "Pareto-Optimal Taxation with Aggregate Uncertainty and Financial Markets," 2009. MIT working paper.

Shimer, Robert, "Convergence in Macroeconomics," *American Economic Journal: Macroeconomics*, 2009, *1* (1), 280–297.

Sleet, Christopher, "Endogenously Incomplete Markets: Macroeconomic Implications," 2006. Mimeo. Tepper Business School.

Spear, Stephen E. and Sanjay Srivastava, "On Repeated Moral Hazard with Discounting," *Review of Economic Studies*, 1987, *54*, 599–617.

Thomas, Jonathan and Tim Worrall, "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem.," *Journal of Economic Theory*, 1990, *51*, 367–390.

Townsend, Robert, "Risk and Insurance in Village India," *Econometrica*, 1994, *62* (3), 539–591.

Townsend, Robert M., *The Medieval Village Economy*, Princeton University Press, 1993.

Werning, Iván, "Optimal Fiscal Policy with Redistribution," *Quarterly Journal of Economics*, 2007, *122*, 925–967.

Wilson, Robert, "The Theory of Syndicates," *Econometrica*, 1968, *36*, 119–132.

A Proof of Proposition 3

The first step of our proof is to define a cost minimization program that the efficient allocation must solve. We then show that, associated with this program is an auxiliary problem—a finite version of Fernandes and Phelan (2000)—which has a nice recursive structure which will be used to partially characterize the efficient allocations.

Toward this goal, using Lemma 1, define $Q(z^{t+1}|z^t)$ through

$$Q(z^{t+1}|z^t) \equiv \left[\sum_{\theta^{t+1}\succ\theta^t} \mu(\theta^{t+1}|\theta^t) \frac{u'(c(\theta^t, z^t))}{u'(c(\theta^{t+1}, z^{t+1}))} \right]^{-1}.$$

Now consider the cost minimization problem (program $\mathcal{P}_1$)

$$\min \sum_{z^t} Q(z^t) \sum_{\theta^t} \mu_t\left(\theta^t\right) \left[c(\theta^t, z^t) - y(\theta^t, z^t)\right]$$

subject to

$$\sum \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t)\mu_t\left(\theta^t\right) \left[u\left(c(\theta^t, z^t)\right) - v\left(\frac{y(\theta^t, z^t)}{\theta_t z_t}\right)\right] = w(\theta_0, z_1)$$

and

$$\sum_t \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t)\mu_t\left(\theta^t\right) \left[u\left(c(\theta^t, z^t)\right) - v\left(\frac{y(\theta^t, z^t)}{\theta_t z_t}\right)\right] \geq$$
$$\sum_t \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t)\mu_t\left(\theta^t\right) \left[u\left(c\left(\sigma(\theta^t, z^t), z^t\right)\right) - v\left(\frac{y\left(\sigma(\theta^t, z^t), z^t\right)}{\theta_t z_t}\right)\right]$$

for all $\sigma \in \Sigma$, where $Q(z^t) = \prod_{\tau=1}^t Q(z^{\tau+1}|z^\tau)$ and $w(\theta_0, z_1)$ is the expected utility associated with an individual with 'seed' value θ_0.

Lemma 2 *At the optimum, there are no idle resources in any period t, and any state z_t.*

Proof. Assume that there are idle resources at a given state z_t. Split them across agents in such a way that utility is increased by the same amount for all agents. This reform is incentive compatible since preferences are additively separable between consumption and effort. This contradicts welfare maximization. ■

Lemma 3 *If allocation $\{c, y\}$ solves $\mathcal{P}_0$ it also solves $\mathcal{P}_1$.*

Proof. From Lemma 2, period by period, flow utility is attained in the most efficient, i.e., cost saving, way with $\{c, y\}$. Thus any possible gain should come from redistribution of resources across periods. But, $Q(z^t)$ was constructed in the exact way as to make any possible gains from intertemporal transfers to disappear. ■

Given the 'price' process $Q(z^t)$, program $\mathcal{P}_1$ has a nice separable structure that we shall exploit. Hence, in what follows, we shall focus on program $\mathcal{P}_1$.

Lemma 4 *An allocation is incentive compatible if and only if it satisfies inequalities (13).*

Proof. Because our proof is a simple adaptation of Theorem 2.1 in Fernandes and Phelan (2000) we shall just sketch its steps. First we show that if an allocation satisfies (13) then, following any history of past announcements $\sigma\left(\theta^{t-1}, z^{t-1}\right)$ and true history (θ^{t-1}, z^{t-1}), the allocation is incentive compatible after history and announcements $\left(\left(\theta^{t-1}, z^{t-1}\right), \sigma\left(\theta^{t-1}, z^{t-1}\right)\right)$. Suppose that this is not the case, then, following some $\left(\left(\theta^{t-1}, z^{t-1}\right), \sigma\left(\theta^{t-1}, z^{t-1}\right)\right)$ it is possible to find a continuation $\varsigma_t = (\tilde{\theta}^t, ..., \tilde{\theta}^{t+s}, ...)$ that yields higher expected utility than telling the truth. But, in this case, following the strategy of telling the truth until period $t-1$ and following the continuation strategy ς_t yields more utility than truthtelling, which is a contradiction with the fact that the allocation is incentive compatible. To show that (6) implies (13) is immediate. Just note that if there were any one period deviation that yielded more expected utility following any history that occurs with positive probability (since we are not considering the perfectly persistent case, this means all finite histories), then the strategy of only lying at the specific node in which one period deviation is welfare increasing; i.e., yields a larger expected utility than truthtelling. This contradicts the assumption that (4) is satisfied. To prove the converse we assume that (13) is satisfied for all (θ^t, z^t) For the last period, this simply amounts to static incentive compatibility. Now assume that the continuation strategy of telling the truth from period t on delivers higher expected utility than any other continuation strategy. In this case, (13) guarantees that the continuation strategy of telling the truth from period $t-1$ on delivers more expected utility than any other continuation strategy starting in period $t-1$. Because telling the truth is optimal in period T, then it is optimal in all periods. ■

For the preferences defined in Section 3, consider the following transformation of variables,

$$\mathfrak{u}(\theta^t, z^t) = \frac{c(\theta^t, z^t)^{1-\rho}}{1-\rho}$$

and

$$\mathfrak{v}(\theta^t, z^t) = \frac{y(\theta^t, z^t)^{\gamma}}{\gamma \theta_t^{\gamma} z_t^{\gamma}}.$$

It will also be convenient to define

$$\mathfrak{u}^{\sigma}(\theta^t, z^t) = \frac{c(\sigma(\theta^t, z^t), z^t)^{1-\rho}}{1-\rho}$$

and

$$\mathfrak{v}^{\sigma}(\theta^t, z^t) = \frac{y(\sigma(\theta^t, z^t), z^t)^{\gamma}}{\gamma \theta_t^{\gamma} z_t^{\gamma}}.$$

Note that

$$\mathfrak{u}^{\sigma}(\theta^t, z^t) = \mathfrak{u}(\sigma(\theta^t, z^t), z^t)$$

while

$$\mathfrak{v}^{\sigma}(\theta^t, z^t) = \frac{y(\sigma(\theta^t, z^t), z^t)^{\gamma}}{\gamma \theta_t^{\gamma} z_t^{\gamma}} = \mathfrak{v}(\sigma(\theta^t, z^t), z^t)\frac{\sigma_t(\theta^t, z^t)^{\gamma}}{\theta_t^{\gamma}}$$

Finally let $C(\mathfrak{u}(\theta, z)) = \mathfrak{u}(\theta, z)^{\frac{1}{1-\rho}}[1-\rho]^{\frac{1}{1-\rho}}$ and $V(\mathfrak{v}(\theta, z)) = \gamma^{\frac{1}{\gamma}}\mathfrak{v}(\theta, z)^{\frac{1}{\gamma}}\theta z$.

The first thing we shall do is to recursively define the expected continuation utility for an agent with history θ^t conditional on aggregate history z^{t+1} as

$$w(\theta^{t-1}, z^t) = \sum_{\theta^t \succ \theta^{t-1}} \mu_t\left(\theta^t | \theta^{t-1}\right)\left[\mathfrak{u}(\theta^t, z^t) - \mathfrak{v}(\theta^t, z^t) + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) w(\theta^t, z^{t+1})\right]$$

for $t < T$ and

$$w(\theta^{T-1}, z^T) = \sum_{\theta^T \succ \theta^{T-1}} \mu_T\left(\theta^T | \theta^{T-1}\right)\left[\mathfrak{u}(\theta^T, z^T) - \mathfrak{v}(\theta^T, z^T)\right],$$

for $t = T$.

Similarly, for $\bar{\theta}^{t-1} = (\theta^{t-2}, \bar{\theta}) \neq (\theta^{t-2}, \theta) = \theta^{t-1}$. We then define

$$\hat{w}(\theta^{t-1}, z^t) = \sum_{\theta^t \succ \theta^{t-1}} \mu_t\left(\theta^t | \bar{\theta}^{t-1}\right)\left[\mathfrak{u}(\theta^t, z^t) - \mathfrak{v}(\theta^t, z^t) + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) w(\theta^t, z^{t+1})\right]$$

for $t < T$ and

$$\hat{w}(\theta^{T-1}, z^T) = \sum_{\theta^T \succ \theta^{T-1}} \mu_T \left(\theta^T | \bar{\theta}^{T-1}\right) \left[\mathfrak{u}(\theta^T, z^T) - \mathfrak{v}(\theta^T, z^T)\right],$$

for $t = T$.

Assuming that the 'seed values' θ_0 are publicly known, and starting from z_1 (i.e., using $\pi(z_1) = 1$) we may define as in Fernandes and Phelan (2000) an auxiliary problem of the form:

$$\min \sum_t \sum_{\theta^t} \sum_{z^t} Q(z^t) \left[C(\mathfrak{u}(\theta^t, z^t)) - V(\mathfrak{v}(\theta^t, z^t))\right]$$

subject to

$$\sum_t \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t) \mu_t \left(\theta^t | \theta_0\right) \left[\mathfrak{u}(\theta^t, z^t) - \mathfrak{v}(\theta^t, z^t)\right] = w(\theta_0, z_1),$$

$$\sum_t \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t) \mu_t \left(\theta^t | \bar{\theta}_0\right) \left[\mathfrak{u}(\theta^t, z^t) - \mathfrak{v}(\theta^t, z^t)\right] = \hat{w}(\theta_0, z_1),$$

and

$$\sum_t \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t) \mu_t \left(\theta^t\right) \left[\mathfrak{u}(\theta^t, z^t) - \mathfrak{v}(\theta^t, z^t)\right] \geq$$
$$\sum_t \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t) \mu_t \left(\theta^t\right) \left[\mathfrak{u}(\sigma(\theta^t, z^t), z^t) - \mathfrak{v}(\sigma(\theta^t, z^t), z^t) \frac{\sigma_t(\theta^t, z^t)^\gamma}{\theta_t^\gamma}\right]$$

for all $\sigma \in \Sigma$, which defines a function $\xi(w(\theta_0, z_1), \hat{w}(\theta_0, z_1))$.

Our goal is to characterize as much as possible any solution to this problem. Since any solution to $\mathcal{P}_1$ is also a solution to this problem, this will allow us to partially characterize the solution to $\mathcal{P}_1$.

We start with the period T problem,

$$\min \sum_{\theta^T \succ \theta^{T-1}} \mu_T \left(\theta^T | \theta^{T-1}\right) \left[C(\mathfrak{u}(\theta^T, z^T)) - V(\mathfrak{v}(\theta^T, z^T))\right]$$

subject to

$$\sum_{\theta^T \succ \theta^{T-1}} \mu_T \left(\theta^T | \theta^{T-1}\right) \left[\mathfrak{u}(\theta^T, z^T) - \mathfrak{v}(\theta^T, z^T)\right] = w(\theta^{T-1}, z^T),$$

$$\sum_{\theta^T \succ \theta^{T-1}} \mu_T \left(\theta^T | \theta^{T-1}\right) \left[\mathfrak{u}(\theta^T, z^T) - \mathfrak{v}(\theta^T, z^T)\right] = \hat{w}(\theta^{T-1}, z^T),$$

and

$$\mathfrak{u}(\theta^T, z^T) - \mathfrak{v}(\theta^T, z^T) \geq \mathfrak{u}(\theta^{T-1}, \bar{\theta}, z^T) - \mathfrak{v}(\theta^{T-1}, \bar{\theta}, z^T)\frac{\bar{\theta}^\gamma}{\theta^\gamma}.$$

The problem is strictly convex and defines a strictly convex value function

$$\xi_T(w(\theta^{T-1}, z^T), \hat{w}(\theta^{T-1}, z^T)).$$

Assume now that period t problem defines a strictly convex function

$$\xi_t(w(\theta^{t-1}, z^t), \hat{w}(\theta^{t-1}, z^t))$$

and consider period $t-1$ problem,

$$\min \Bigg[\sum_{\theta^{t-1} \succ \theta^{t-2}} \mu_{t-1}\left(\theta^{t-1}|\theta^{t-2}\right) C(\mathfrak{u}(\theta^{t-1}, z^{t-1})) - V(\mathfrak{v}(\theta^{t-1}, z^{t-1})) \\ + \sum_{z^{t-1} \succ z^{t-2}} Q(z^t)\xi_t(w(\theta^{t-1}, z^t), \hat{w}(\theta^{t-1}, z^t)) \Bigg]$$

subject to

$$\sum_{\theta^{t-1} \succ \theta^{t-2}} \mu_t\left(\theta^{t-1}|\theta^{t-2}\right) \Bigg[\mathfrak{u}(\theta^{t-1}, z^{t-1}) - \mathfrak{v}(\theta^{t-1}, z^{t-1}) \\ + \beta \sum_{z^{t-1} \succ z^{t-2}} \pi(z^t) w(\theta^{t-1}, z^t) \Bigg] = w(\theta^{t-2}, z^{t-1}),$$

$$\sum_{\theta^{t-1} \succ \theta^{t-2}} \mu_t\left(\theta^{t-1}|\bar{\theta}^{t-2}\right) \Bigg[\mathfrak{u}(\theta^{t-1}, z^{t-1}) - \mathfrak{v}(\theta^{t-1}, z^{t-1}) \\ + \beta \sum_{z^{t-1} \succ z^{t-2}} \pi(z^t) w(\theta^{t-1}, z^t) \Bigg] = \hat{w}(\theta^{t-2}, z^{t-1}),$$

and

$$\begin{aligned} \mathfrak{u}(\theta^{t-1}, z^{t-1}) - \mathfrak{v}(\theta^{t-1}, z^{t-1}) + \beta \sum \pi(z^t) w(\theta^{t-1}, z^t) &\geq \\ \mathfrak{u}(\theta^{t-2}, \bar{\theta}, z^{t-1}) - \mathfrak{v}(\theta^{t-2}, \bar{\theta}, z^{t-1})\frac{\bar{\theta}^\gamma}{\theta^\gamma} + \beta \sum \pi(z^t) \hat{w}(\theta^{t-1}, z^t) &. \end{aligned}$$

Once again, this defines a strictly convex problem, which solution we represent with the strictly convex cost function $\xi_{t-1}(w(\theta^{t-2}, z^{t-1}), \hat{w}(\theta^{t-2}, z^{t-1}))$. Hence, the whole program is convex and we can characterize the solution to period t problem by solving the associated Lagrangian.

Proof. of Proposition 3: Assume that the optimal allocation does not display memory for $t = T$. In particular, this implies that utility promises are independent of z_{T-1}. Define $\hat{c}(\theta^t, z^t) = c(\theta^t, z^t)/\eta(z_t)$ and $\hat{y}(\theta^t, z^t) = y(\theta^t, z^t)/\eta(z_t)$

$$\xi(w(\theta^{t-1}, z^t), \hat{w}(\theta^{t-1}, z^t)) = \min \sum_{\theta^t \succ \theta^{t-1}} \mu_t\left(\theta^t | \theta^{t-1}\right) \Bigg[\eta(z_t)\left[\hat{c}(\theta^t, z^t) - \hat{y}(\theta^t, z^t)\right] \\ + \sum_{z^{t+1} \succ z^t} Q(z^{t+1}) \xi(w(\theta^t, z_{t+1}), \hat{w}(\theta^t, z_{t+1}))\Bigg]$$

subject to

$$\sum_{\theta^t \succ \theta^{t-1}} \mu_t\left(\theta^t | \theta^{t-1}\right) \Bigg[\eta(z_t)^{1-\rho}\left[\frac{\hat{c}(\theta^t, z^t)^{1-\rho}}{1-\rho} - \frac{\hat{y}(\theta^t, z^t)^{\gamma}}{\gamma \theta_t^{\gamma}}\right] \\ + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) w(\theta^t, z_{t+1})\Bigg] = w(\theta^{t-1}, z_t),$$

$$\sum_{\theta^t \succ \theta^{t-1}} \mu_t\left(\theta^t | \hat{\theta}^{t-1}\right) \Bigg[\eta(z_t)^{1-\rho}\left[\frac{\hat{c}(\theta^t, z^t)^{1-\rho}}{1-\rho} - \frac{\hat{y}(\theta^t, z^t)^{\gamma}}{\gamma \theta_t^{\gamma}}\right] \\ + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) w(\theta^t, z_{t+1})\Bigg] = \hat{w}(\theta^{t-1}, z_t),$$

and

$$\left[\frac{\hat{c}(\theta^t, z^t)^{1-\rho}}{1-\rho} - \frac{\hat{y}(\theta^t, z^t)^{\gamma}}{\gamma \theta_t^{\gamma}}\right] - \left[\frac{\hat{c}\left(\theta^{t-1}, \theta, z^t\right)^{1-\rho}}{1-\rho} - \frac{\hat{y}\left(\theta^{t-1}, \theta, z^t\right)^{\gamma}}{\gamma \theta_t^{\gamma}}\right] \geq \\ \beta \left\{\sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) \left[\hat{w}(\theta^{t-1}, \theta, z_{t+1}) - w(\theta^t, z_{t+1})\right]\right\} \eta(z_t)^{\rho-1} \tag{15}$$

The term in curly brackets in the right hand side of expression (15) is independent of z^t since the allocation is memoryless. But, in this case, the left hand side must depend on z^t, meaning that it is not separable. However, from Proposition 2, this cannot be the case. ■

B Appendix: Mirrlees' period T allocation

Proof. Consider the problem being analyzed,

$$\max_{c,y} \sum_{t=1,2} \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_1(z^t) \mu_t\left(\theta^t\right) \left[u\left(c(\theta^t, z^t)\right) - v\left(\frac{y(\theta^t, z^t)}{\theta_t z_t}\right)\right]$$

subject to

$$\sum_{\theta^t} \mu_t\left(\theta^t\right)\left[c(\theta^t, z^t) - y(\theta^t, z^t)\right] \leq 0$$

and

$$\sum_{t=1,2} \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t)\mu_t\left(\theta^t\right)\left[u\left(c(\theta^t, z^t)\right) - v\left(\frac{y(\theta^t, z^t)}{\theta_t z_t}\right)\right] \geq$$
$$\sum_{t=1,2} \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t)\mu_t\left(\theta^t\right)\left[u\left(c\left(\sigma(\theta^t, z^t), z^t\right)\right) - v\left(\frac{y\left(\sigma(\theta^t, z^t), z^t\right)}{\theta_t z_t}\right)\right]$$

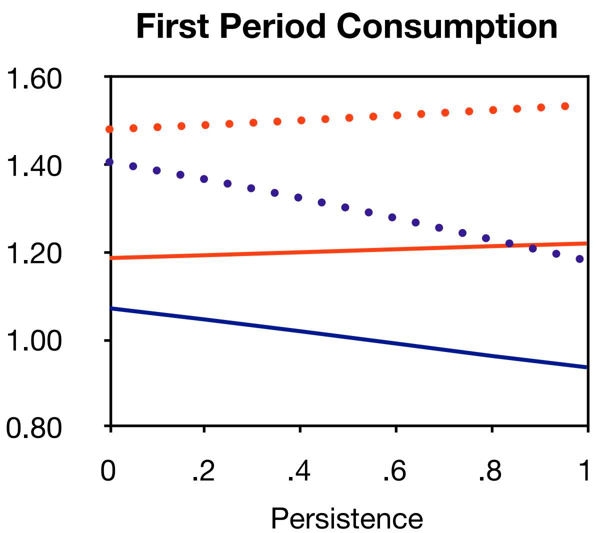
Suppose that the variable of choice is $\tilde{c}$ and $\tilde{y}$, but that the actual consumption and labor implemented in period T are given by $c_T\left(.\right) = \tilde{c}_T\left(.\right) z_T^{\overline{\frac{\gamma}{\gamma+\rho-1}}}$ and $y_T\left(.\right) = \tilde{y}_T\left(.\right) z_T^{\overline{\frac{\gamma}{\gamma+\rho-1}}}$, since $\tilde{c}_T$ and $\tilde{y}_T$ also depend on z_T potentially, this change of variable is without loss of generality. Assume that $c_t = \tilde{c}_t$ and $y_t = \tilde{y}_t$ for $t < T$. Given that we are in a finite period model, the incentive compatibility constraint can be written as a restriction that agents do not lie at period t given any history and that they do not plan to lie from t onwards. Then the last period incentive compatibility becomes

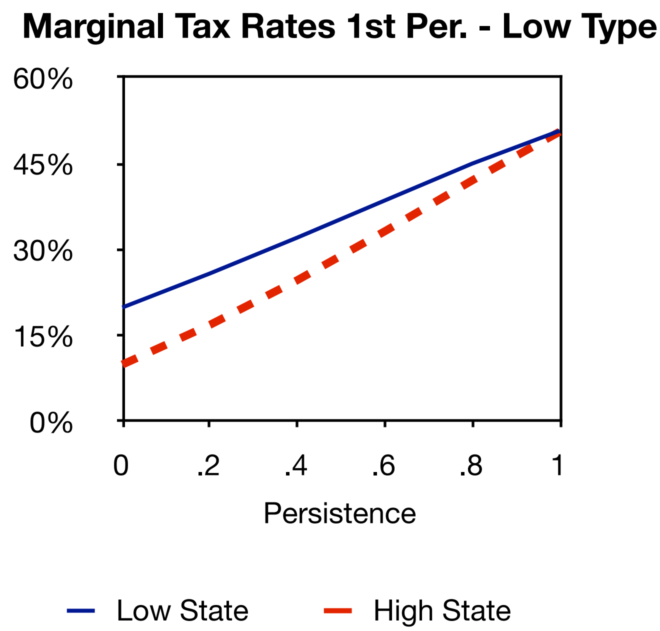
$$z_T^{\frac{(1-\rho)\gamma}{\gamma+\rho-1}}\left\{u\left(\tilde{c}(\theta^{T-1}, \theta, z^t)\right) - v\left(\frac{\tilde{y}(\theta^{T-1}, \theta, z^T) z_T^{\overline{\frac{\gamma}{\gamma+\rho-1}}}}{\theta_T}\right)\right\} \geq$$
$$z_T^{\frac{(1-\rho)\gamma}{\gamma+\rho-1}}\left\{u\left(\tilde{c}\left(\theta^{T-1}, \theta', z^t\right)\right) - v\left(\frac{\tilde{y}\left(\theta^{T-1}, \theta', z^T\right)}{\theta_T}\right)\right\}.$$

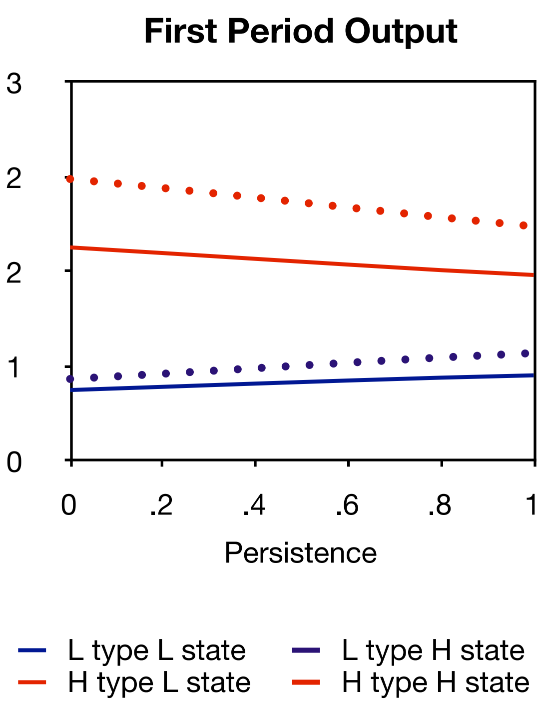
And the resource constraints turns into

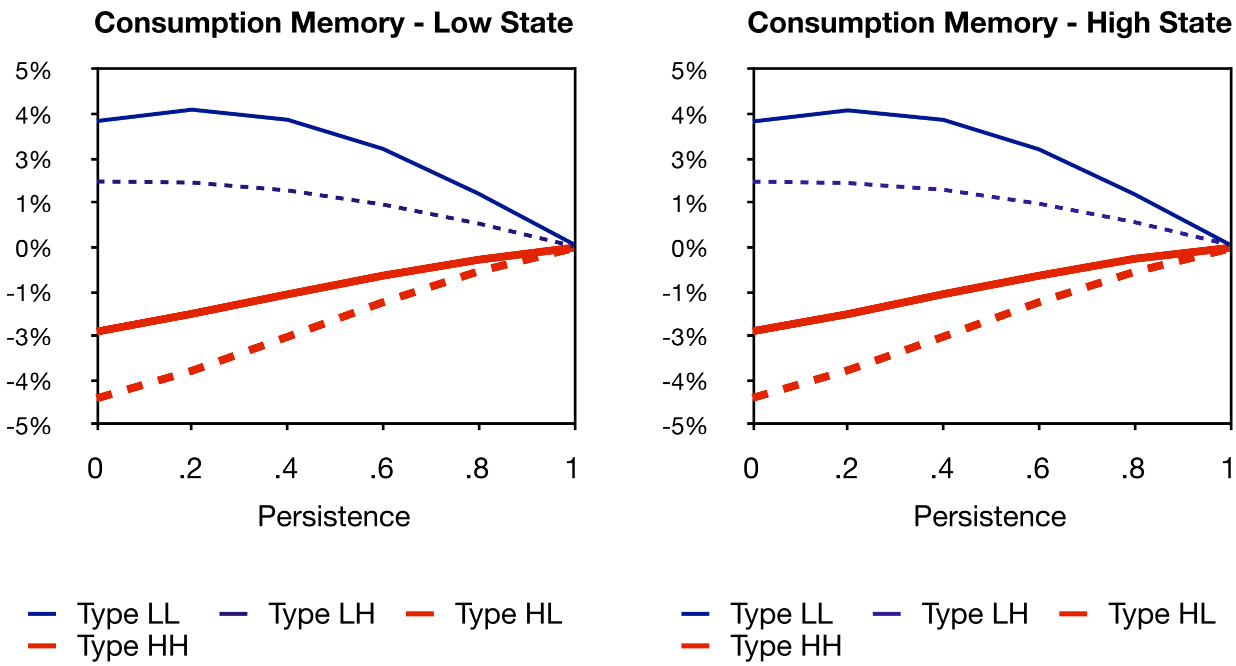
$$z_T^{\overline{\frac{\gamma}{\gamma+\rho-1}}} \sum_{\theta^T} \mu_t\left(\theta^T\right)\left[\tilde{c}_T(\theta^T, z^T) - \tilde{y}_T(\theta^T, z^T)\right] \leq 0.$$

Then the constraints on $\tilde{c}$ and $\tilde{y}$ do not depend explicitly on the aggregate shock in the last period. From this and the concavity of the objective function, we get that $\tilde{c}$ and $\tilde{y}$ do not depend on z_T. ■

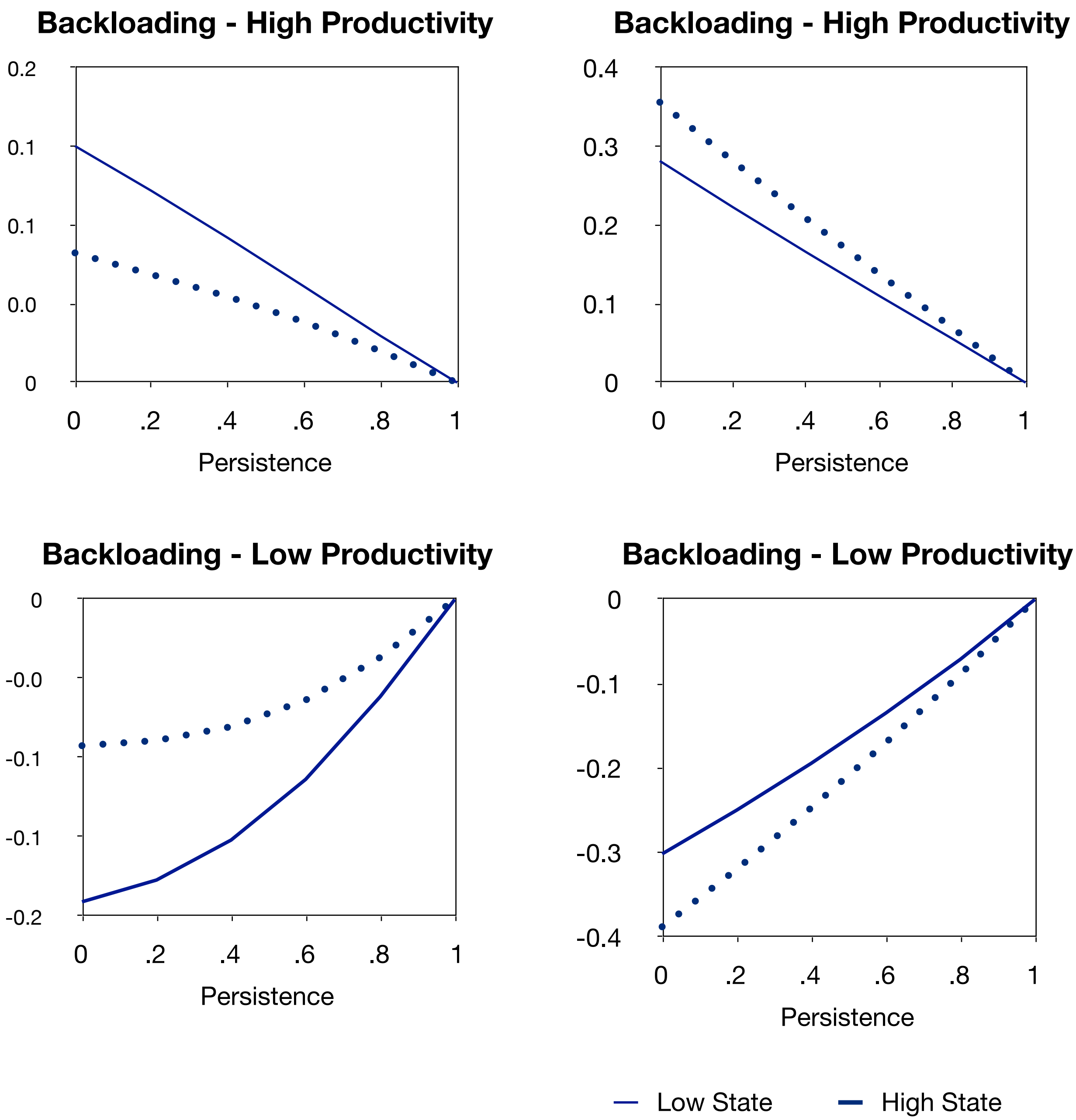

First Period Consumption
1.60
1.40
1.20
1.00
0.80
0
.2
.4
.6
.8
1
Persistence

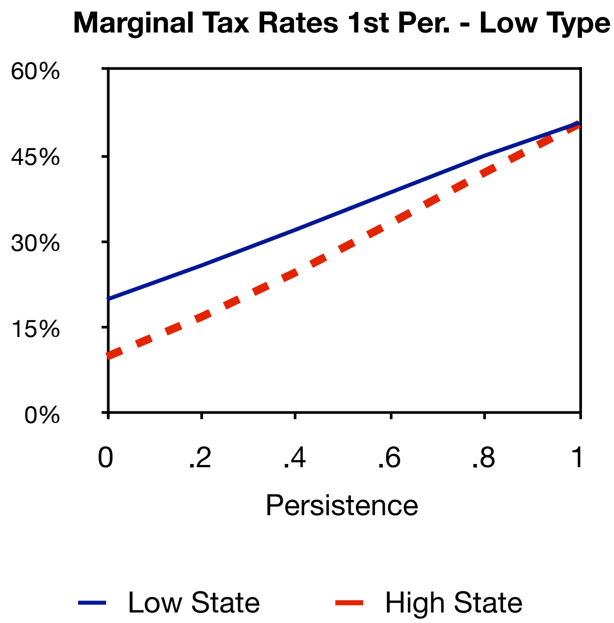

Marginal Tax Rates 1st Per. - Low Type
60%
45%
30%
15%
0%
0
.2
.4
.6
.8
1
Persistence
Low State
High State


First Period Output
3
2
2
1
0
0
.2
.4
.6
.8
1
Persistence
L type L state
L type H state
H type L state
H type H state

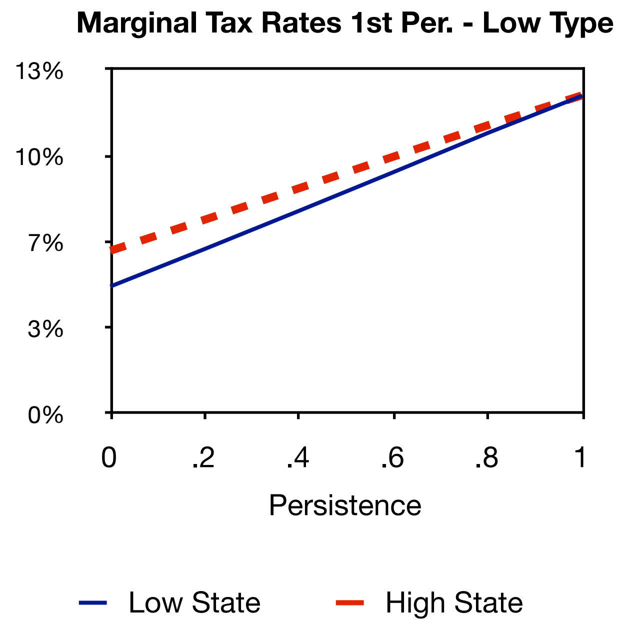


Consumption Memory - Low State
5%
4%
3%
1%
0%
-1%
-3%
-4%
-5%
0
.2
.4
.6
.8
1
Persistence
Type LL
Type LH
Type HL
Type HH
Consumption Memory - High State
5%
4%
3%
1%
0%
-1%
-3%
-4%
-5%
0
.2
.4
.6
.8
1
Persistence
Type LL
Type LH
Type HL
Type HH

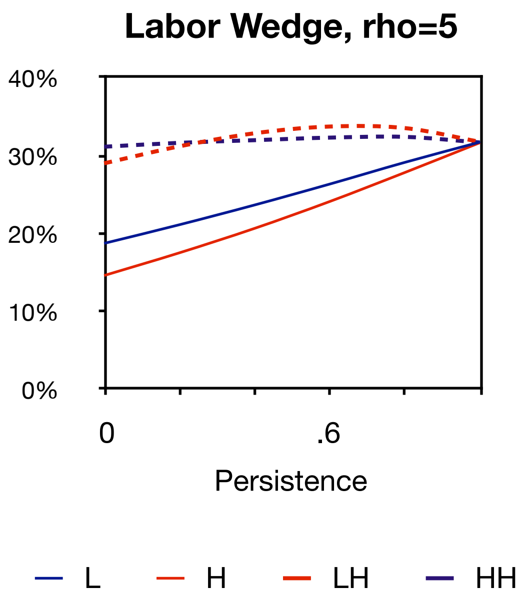

Backloading - High Productivity
0.2
0.1
0.1
0.0
0
0
.2
.4
.6
.8
1
Persistence
Backloading - High Productivity
0.4
0.3
0.2
0.1
0
0
.2
.4
.6
.8
1
Persistence
Backloading - Low Productivity
0
-0.0
-0.1
-0.1
-0.2
0
.2
.4
.6
.8
1
Persistence
Backloading - Low Productivity
0
-0.1
-0.2
-0.3
-0.4
0
.2
.4
.6
.8
1
Persistence
Low State
High State

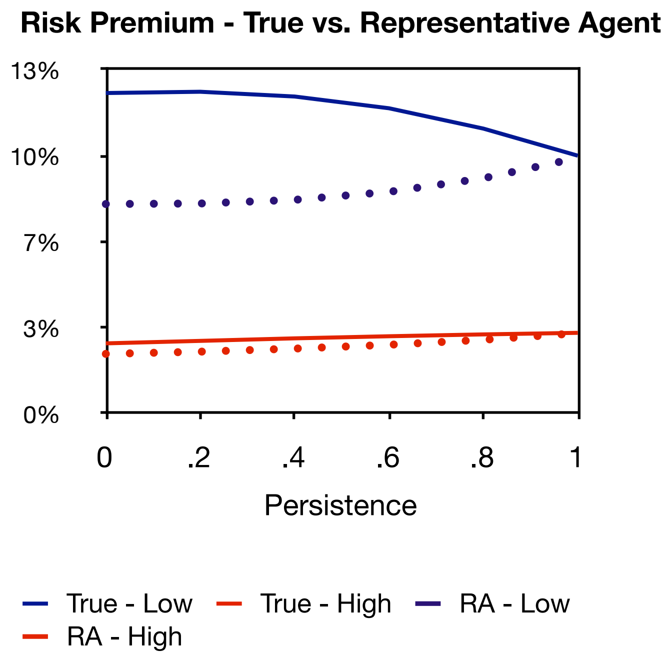


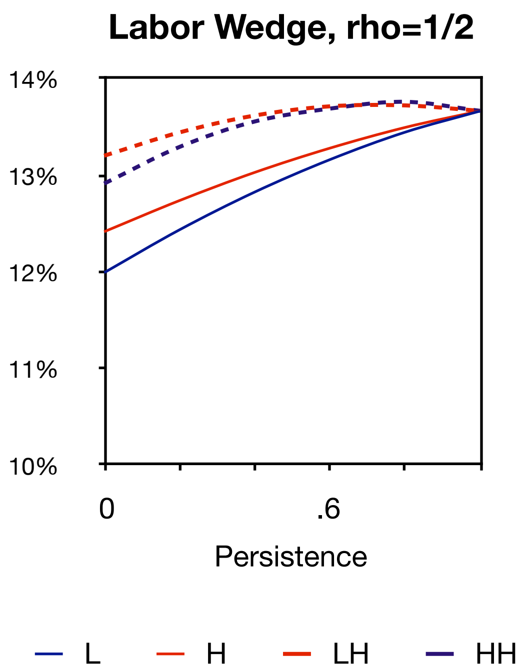
Marginal Tax Rates 1st Per. - Low Type
60%
45%
30%
15%
0%
0
.2
.4
.6
.8
1
Persistence
Low State
High State

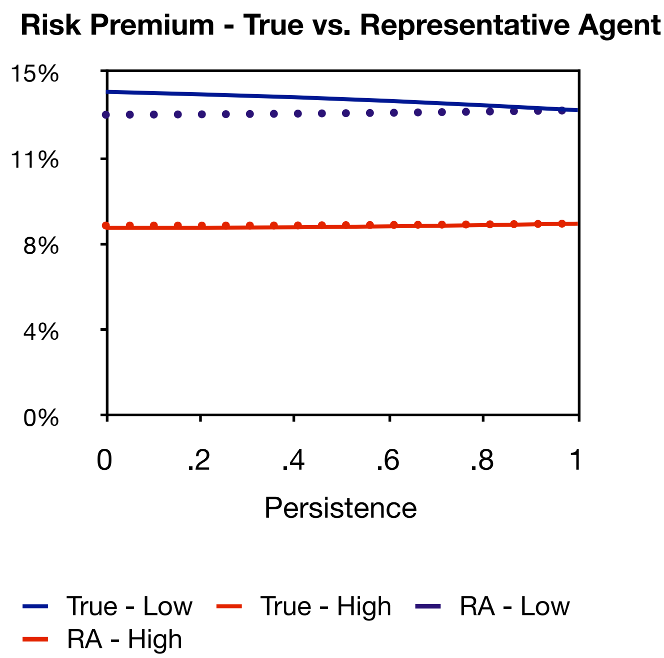


Marginal Tax Rates 1st Per. - Low Type
13%
10%
7%
3%
0%
0
.2
.4
.6
.8
1
Persistence
Low State
High State


Labor Wedge, rho=5
40%
30%
20%
10%
0%
0
.6
Persistence
L
H
LH
HH


Risk Premium - True vs. Representative Agent
13%
10%
7%
3%
0%
0
.2
.4
.6
.8
1
Persistence
True - Low
True - High
RA - Low
RA - High


Labor Wedge, rho=1/2
14%
13%
12%
11%
10%
0
.6
Persistence
L
H
LH
HH


Risk Premium - True vs. Representative Agent
15%
11%
8%
4%
0%
0
.2
.4
.6
.8
1
Persistence
True - Low
True - High
RA - Low
RA - High