Separability and Memory: Micro Causes, Macro Consequences.*

Carlos E. da Costa Vitor F. Luz

February 15, 2011

Abstract

In a multi-period Mirrlees [1971] setting individuals face privately observed shocks to their ability. Each individual's productivity is the product of this idiosyncratic ability and an aggregate component which is also stochastic but publicly observed. Using a separable iso-elastic specification for preferences, we investigate optimal risk sharing and its macro consequences in this environment. We show that efficient allocations display memory with respect to past aggregate shocks for all levels of risk aversion different from one. This dependence on past aggregate shocks decreases with the persistence of idiosyncratic shocks. At the limit, when the economy becomes a pure heterogeneity one, the constrained efficient allocation is independent of past aggregate shocks and displays a strong form of separability with respect to aggregate shocks. Indeed, these two properties, separability and lack of memory, are shown to be connected: an allocation is memoryless only if it separable. Non-separability is interesting in itself for it is associated with state-varying labor wedges at the micro level. Using numerical examples we show that this non-separability induces *efficient counter-cyclical labor wedges* at the macro level as well, when risk aversion is greater than one. **JEL classification:** H21; E60**.** **Keywords:** Dynamic Mirrlees Economy; Counter-cyclical labor wedges.

*Previously, *The Private Memory of Aggregate Shocks*. In preparing this manuscript we greatly benefited from our discussions and from the comments from Caio Almeida, Luis Braido, Vinicius Carrasco, Bev Dahlby, Mike Golosov, Humberto Moreira, Aleh Tsyvinski, Henry van Egteren and seminar participants at the Walton College and the University of Alberta. da Costa thanks the hospitality of MIT and gratefully acknowledges financial support from CNPq. We retain sole responsibility for all remaining errors.

1 Introduction

Following Wilson [1968]'s landmark contribution, much has been learnt about optimal risk sharing. The mutualization of private risks and its consequences for the dynamics of private consumption as a function of aggregate consumption are now well understood: conditional on aggregate consumption, an 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 are history dependent.[1] Most of this latter literature focuses on the case in which only private risk exists by 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 agency 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 the endogenously generated interconnections displayed by the efficient allocations.

We show that private allocations exhibit memory with respect to aggregate shocks. That is, today's allocation depends not only on the current aggregate shock—as in Wilson [1968]— and on yesterday's idiosyncratic shock—as in Golosov et al. [2003]—but also on yesterday's aggregate shock, despite the fact that the latter is public and independent of idiosyncratic shocks. Our assumptions on the nature of shocks rules out non-trivial 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 the agents' private in-

[1] The literature on dynamic insurance in the presence of private information is now an extensive one. Starting with the early contribution of Rogerson [1985] followed by the methodological advances found in Green [1987], Spear and Srivastava [1987], Thomas and Worrall [1990] and Atkeson and Lucas [1992] it has established robust predictions about the time series behavior of allocations with somewhat surprising consequences: the non-existence of non-degenerate steady-states and the associated immiseration results being one of the foremost consequences of the back loading incentives which typically characterizes such environments.

formation. Instead, the rationale for our results is closer to Levin [2002],[2] in the sense that aggregate shocks affect the marginal cost of incentives in the current and future periods.[3]

We finally run some numeric examples to: i) illustrate some of the results obtained analytically; ii) produce 'aggregate' data to check how the labor wedge varies along the economic cycle.

Chari et al. [2007], , among others, have shown that the labor wedge associated with a representative agent economy varies counter-cyclically along the business cycle. Because incentives in our model are provided in the leisure/consumption margin in a time and state varying fashion, our prototype economy is potentially useful for understanding state-dependent labor wedges. To summarize our findings, for risk aversion greater than one, the data generated by our prototype economy has the same property one finds in the data: a representative agent would have to be facing counter-cyclical labor wedges to make these choices. Although the results are reversed when risk aversion is less than one, our results raise the possibility that counter-cyclical wedges may characterize efficient responses to aggregate shocks.

For wedges to vary with the aggregate state of the economy, allocations must display a strong form of non-separability with respect to aggregate shocks. Non-separability also plays an important role in our discussion of memory: an allocation is memoryless only if it is separable in previous periods. Due to the relationship between separability and memory, an allocation exhibits state-varying wedges only if it displays memory.

From a purely conceptual perspective this result ties labor wedge variation to persistence of aggregate allocations with respect to temporary shocks. From a computational perspective this poses an important challenge for future calibration exercises, since interesting business cycle dynamics comes at the cost of an increased dimensionality problem. Or taken from a different angle, this separability result should be seen as an important

[2]Here, the resource constraint induces an interdependence in the provision of incentives across agents while in Levin [2002] the same type of interdependence arises due to an endogenous restriction on what may be credibly promised to the pool of agents. We thank Vinicius Carrasco for pointing this out.

[3]For the class of utility function studied here, the marginal cost of providing incentives is completely pinned down by the coefficient of risk aversion. Indeed, Demange [2008] defines the incentive premium as the transfer that makes the marginal utility of consumption for a representative consumer identical to the marginal value of resources for the society. With constant relative risk aversion — CRRA — preferences the sign of this premium only depends on whether risk aversion is greater than one or not.

cautionary note. Assumptions commonly used to eliminate history dependence—e.g., ln utility—are bound to eliminate interesting business cycle variations.

The rest of the paper is organized as follows. After a brief account of the related literature that is still part of this Introduction, the core of the paper starts in Section 2, where the environment is described. Our main theoretical results are presented in Section 3. Section 4 defines macro-aggregates for our prototype economy and discusses some of the implications of our underlying assumptions for their behavior. Numerical exercises are conducted in Section 5. Section 6 concludes the paper. All proofs are collected in the appendix.

Related Literature The consequences for efficient allocations of the interaction between privately observed idiosyncratic risk and publicly observed aggregate risk was, to the best of our knowledge, first analyzed by Phelan [1994]. Using a perpetual life overlapping generations model, Phelan [1994] provides a full characterization of the steady state of a dynamic moral hazard model with i.i.d. shocks and aggregate uncertainty, when preferences are of the constant absolute risk aversion—CARA—type. By eliminating income effects through the CARA assumption and by killing the tendency for ever increasing inequality through the use of an OLG formulation Phelan [1994] is able to find a closed form solution for optimal allocations. As it turns out, optimal allocations *do* exhibit memory of aggregate shock.

The main difference between our work and Phelan [1994] is that our focus is on a dynamic Mirrlees' economy, whilst his is on a repeated moral hazard economy. This turns out to be a crucial difference for our purposes for one of our main goals is to understand how labor wedges vary along the business cycles. This question can neither be addressed in Green [1987] nor in Phelan [1994] settings. Of course, this comes at a cost of our not being able to provide a full solution to the model.

Another difference is that we consider different levels of persistence of private shocks. Our numerical simulations indicate that the presence of memory which characterizes an environment with i.i.d. private information found in Phelan [1994] and the lack of memory which characterizes the pure heterogeneity case found in Werning [2007] are not knife-edge results. In fact, we find a monotonic relationship between private persistence and aggregate memory: the more persistent the private shocks the less dependent is the allocation on past aggregate shocks.

Finally, agents in our economy have preferences of the constant relative risk aversion—CRRA—type. We do so not so much for the sake of realism. Given all the simplifying assumptions we use this would be too naive. Instead, we aim at focusing on the effect of incentives on the marginal value of resources. Recent work by Demange [2008] has addressed the interaction between risk sharing and moral hazard–see also Scheuer [2009]. She shows how the presence of moral hazard may increase or reduce the marginal value of resources depending on the coefficient of relative risk aversion. In particular, when risk aversion is equal to one, the ln case, moral hazard does not affect the marginal value of resources. In our model, this is the case in which the allocations do not display memory. As we have suggested, because the dependence on aggregate shock is through the resource constraint, the fact that under ln preferences agency does not affect the value of income leads to our result. Although bothScheuer [2009] and Demange [2008] study the interaction between risk sharing and moral hazard, both work in a static setting, 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, nor do they emphasize the behavior of aggregate variables. Also related are the numerical exercises found in Kocherlakota [2005]. The emphasis in this case is on optimal taxes on capital returns. Finally Kocherlakota and Pistaferri [2007, 2009] explore a specific macro consequence of the combination of private idiosyncratic shock and public aggregate shock in a Mirrlees setting, namely its asset pricing implications.

We claim not that we are 'accounting' for the stylized facts. The simplifying assumptions we adopt in our normative framework weakens any claim of a qualitative account of the data. That is, although we would want to interpret our results in the spirit of Townsend [1993], where the underlying idea is that many different institutional arrangements are capable of inducing a given allocation,[4] our model is too stylized to allow one

[4]Indeed, one important lesson we take from the enormous mechanism design literature is that we should focus on allocations first, if not exclusively, in the case of positive questions–e.g., Ligon et al. [2002], Attanasio and Pavoni [2008]. In our case, the fact that we do not see labor income taxes vary at the frequency required

to take it to the data and check its explanation power. Rather we view our result as opening the possibility that the countercyclical behavior of the labor wedge found in the data need not indicate a form of inefficiency of responses to business cycle fluctuations.

2 Basic Setting

The economy is inhabited by a continuum of measure one of ex-ante identical individuals, each living for $T < \infty$ periods.Agents have preferences defined over deterministic sequences of consumption-effort bundles, $(c, y) \equiv \{c_t, l_t\}_{t=1}^{T}$ represented by an additively separable temporary utility

$$U(c, y) = \sum_{t=1}^{T} \beta^t \{u(c_t) - v(l_t)\}$$

where c_t is period t consumption and l_t is period t effort. We assume u and v are smooth strictly increasing functions. Both u and $-v$ are strictly concave.

Every period individuals realize a temporary productivity shock, $\theta_t \in \Theta = \{\bar{\theta}, \underline{\theta}\}$, with $\bar{\theta} > \underline{\theta}$. 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.

These idiosyncratic shocks are not the only source of uncertainty in this economy. In each period, an aggregate shock represented by a random variable $z_t \in Z \equiv \{\bar{z}, \underline{z}\}$, with $\bar{z} > \underline{z} > 0$, affects the economy's technology. Given an aggregate state z_t and individual who realized productivity θ_t in period t produces output $y_t = l_t z_t \theta_t$ with effort l_t. In what follows we shall focus on sequences of consumption/output, $(c, y) = \{c_t, y_t\}_{t=1}^{T}$ instead.

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 $\pi(z)$. $\pi_t(z^t)$ 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,

by the model does not mean that other formal or informal arrangements may generate these state-varying wedges. Along these lines, a relatively large body of research in the realm of the New Dynamic Public Finance—NDPF—literature has considered macroeconomic implications of private information. Examples are , Ales and Maciero [2008], etc. See Sleet [2006] for a brief account.

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}. We further assume that the stochastic process $\{\theta_t\}_{t\geq 0}$ is a first order Markov process.

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

Given a history of shocks, (θ^t, z^t), individuals rank deterministic streams $\{(c_t, y_t)\}_{t=1}^T \in R_+^{2T}$, according to

$$\sum_{t=1}^{T} \beta^{t-1}\left\{u(c_t) - v\left(\frac{y_t}{z_t\theta_t}\right)\right\}.$$

An *allocation* is $(c, y) \equiv \{(c_t, y_t)\}_{t=1}^T$, with $(c_t, y_t) : Z^t \times \Theta^t \rightarrow \mathbb{R}_+^2$ 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,

$$\sum_t \beta^{t-1} \sum_{z^t}\sum_{\theta^t} \mu_t\left(\theta^t\right)\pi_t(z^t)\left\{u(c_t) - v\left(\frac{y_t}{z_t\theta_t}\right)\right\}. \tag{1}$$

Technology is very simple. Each unit of output produces one unit of consumption in each period. There is no storage or any other form of investment in this economy.

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 $\sigma_t : Z^t \times \Theta^t \rightarrow \Theta$, which associates 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.

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

$$\mathcal{U}(c, y; \sigma) \equiv \mathbb{E}\left[\sum_t \beta^{t-1}\left\{u(c_t\left(\sigma^t(\theta^t, z^t), z^t\right)) - v\left(\frac{y_t\left(\sigma^t(\theta^t, z^t), z^t\right)}{z_t\theta_t}\right)\right\}\right].$$

An allocation (c, y) is *incentive compatible* if

$$\mathcal{U}(c, y) \geq \mathcal{U}(c, y; \sigma), \text{ for all } \sigma \in \Sigma, \tag{2}$$

where $\mathcal{U}(c, y)$ is the utility associated with the truth-telling strategy, $\sigma_t\left(\theta^t, z^t\right) = \theta^t$ for all t, θ^t, z^t.

3 Efficient Allocations

In this section we derive some of the properties of the solution to the problem of maximizing (1) subject to feasibility and (2), which we call the social planner's problem (program $\mathcal{P}_0$).

$$\max \sum_t \beta^{t-1} \sum_{\theta^t}\sum_{z^t} \pi_t(z^t)\mu_t\left(\theta^t\right)\left[u\left(c_t(\theta^t, z^t)\right) - v\left(\frac{y_t(\theta^t, z^t)}{\theta_t z_t}\right)\right] \tag{3}$$

subject to

$$\sum_{\theta^t} \mu_t\left(\theta^t\right)\left[c_t(\theta^t, z^t) - y_t(\theta^t, z^t)\right] \leq 0, \tag{4}$$

$\forall t, z^t$, and

$$\begin{aligned}&\sum_t \beta^{t-1} \sum_{\theta^t}\sum_{z^t} \pi_t(z^t)\mu_t\left(\theta^t\right)\left[u\left(c_t(\theta^t, z^t)\right) - v\left(\frac{y_t(\theta^t, z^t)}{\theta_t z_t}\right)\right] \geq \\ &\sum_t \beta^{t-1} \sum_{\theta^t}\sum_{z^t} \pi_t(z^t)\mu_t\left(\theta^t\right)\left[u\left(c_t\left(\sigma_t(\theta^t, z^t), z^t\right)\right) - v\left(\frac{y_t\left(\sigma_t(\theta^t, z^t), z^t\right)}{\theta_t z_t}\right)\right]\end{aligned} \tag{5}$$

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 current type reports.[5]

We start by deriving an inter-temporal restriction on allocations that will play an important role in what follows. The result is akin to the inverse Euler equation derived by Kocherlakota [2005]. However, because there is no technology to transfer resources across periods, however, the following lemma simply relates allocations along different paths.

Lemma 1 *At the optimum,*

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

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

At any given period, individuals are concerned 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.

As we have stated, in this section we want to explore how private allocations are affected by aggregate shocks. Toward this end we define some properties that allocations may (or not) display.

Definition 1 *If an allocation is such that in period t there are functions $\eta : Z \to R_+$ and $(\tilde{y}_t, \tilde{c}_t) : \Theta^t \times Z^{t-1} \to R^2_+$ such that $(c_t(\theta^t, z^t), y_t(\theta^t, z^t)) = (\tilde{c}_t(\theta^t, z^{t-1}), \tilde{y}_t(\theta^t, z^{t-1}))\eta(z_t)$ we say that it is **separable at** t.*

[5] 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. With i.i.d. private shocks, an assumption frequently adopted in the literature, the extra dispersion in allocations is independent of previous heterogeneity. This pattern generates ever increasing inequality across agents.

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

Definition 3 *If an allocation is such that, it is independent of z^{t-s} for all t,s, z^t, i.e., it is of the form*

$$(c_t(\theta^t, z^t), y_t(\theta^t, z^t)) = (\hat{c}_t(\theta^t, z_t), \hat{y}_t(\theta^t, z_t))$$

then we say it is ***memoryless****.*

For the rest of the paper we also restrict our investigation to the class of iso-elastic utility functions: constant relative risk aversion in consumption and power disutility of labor effort.

One of the reasons that we specialize to this class of preferences is that it is probably the most widely used specification of preferences, in macroeconomics. Moreover, this type of preferences imposes a lot of structure in optimal allocations for some special situations. In particular, we shall see that separability will characterize first best allocations when these preferences are used. Separability will, in this sense, become an interesting benchmark.

Formally, we assume that

$$u(c) = \frac{c^{1-\rho}}{1-\rho} \text{ and } v(l) = \frac{l^\gamma}{\gamma}, \tag{6}$$

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

Let us start our characterization of optimal allocations by deriving the benchmark first-best allocation. That is, assume that there is no asymmetric information, then the optimal allocations are given by,

$$c_t^{FB}(\theta^t, z^t) := \eta(z_t) \left\{ \sum \mu(\theta_t) \theta^{\frac{\gamma}{\gamma-1}} \right\}^{\frac{\gamma-1}{\gamma+\rho-1}}$$

and

$$y_t^{FB}(\theta^t, z^t) := \eta(z_t) \left\{ \sum \mu(\theta_t) \theta^{\frac{\gamma}{\gamma-1}} \right\}^{-\frac{\rho}{\gamma+\rho-1}} \theta_t^{\frac{\gamma}{\gamma-1}},$$

where the superscript FB stands for first best and $\eta(z_t) = z_t^{\frac{\gamma}{\gamma+\rho-1}}$. The first best allocation is both separable and memoryless.

Next, we shall show that this type of separability and lack of memory that we observe in the first best will also characterize a repeated Mirrlees economy, for which there is heterogeneity but no idiosyncratic risk. In this setting, which is a particular case of Werning [2007], there is no uncertainty beyond the productivity one is born with.

We shall call this the pure heterogeneity case, which we view as the perfect persistence limit of our economy. Indeed, if idiosyncratic shocks are Markov, this results from our choice of: $\mu(\theta', \theta|\theta') = 1$ for $\theta' = \theta$, $\mu(\theta', \theta|\theta') = 0$ for $\theta' \neq \theta$. Naturally, because there is no idiosyncratic risk there is no proper insurance to speak of. Instead we solve the Utilitarian planner's problem [6].

Proposition 1 *The efficient allocation for the pure heterogeneity economy is of the form*

$$(c_t(\theta, z^t), y_t(\theta, z^t)) = (\tilde{c}(\theta), \tilde{y}(\theta))\eta(z_t)$$

with $\eta(z) = z^{\frac{\gamma}{\gamma+\rho-1}}$*, i.e., it is separable and memoryless.*

As we shall see, a consequence of Proposition 1 is that simply allowing for heterogeneity will not suffice to generate any interesting business cycle behavior beyond what obtains with a suitably defined representative agent framework. This result is still interesting for a future understanding of what is *not* behind our results. In particular, it highlights the fact that it is not composition effects that drives our findings in Section 5.

So far we have considered environments for which the optimal allocations did not possess any kind of memory. We already knew from standard results in optimal risk sharing that there was no reason to tie current allocation to past shocks in the first best. As for the pure heterogeneity case, once again, because there is no idiosyncratic risk, our result implies that current allocation only depends on current aggregate shock.

Whether introducing idiosyncratic risk should change the way allocations vary with past aggregate risk is not obvious. The first best solution illustrates the fact that introducing this type of dependence is costly. Besides, since the aggregate state of the economy is assumed to be known by everyone, and since it is unrelated with idiosyncratic shocks, the public signal is not informative of an individual's current or past behavior.

[6] In the same way as Werning [2007], this characterization is true for any Constrained Pareto Optimal allocation.

Note, however, that these examples are special in that in one case we had idiosyncratic risk but no asymmetric information, while in the other we had asymmetric information but no idiosyncratic risk. As a consequence, in that optimal allocations do not display memory from either type of shock. At least since Rogerson [1985], however, we know that optimal allocations display memory of private shocks in repeated agency models. Would the introduction of idiosyncratic risk also affect the memory of aggregate shocks?

Before we tackle this question another interesting aspect of the previous results would have probably caught the attention of the careful reader. In all of the examples above, both separability and lack of memory were present. It makes one wonder whether these two properties are linked in a model with both asymmetric information and idiosyncratic risk. They are, according to the next proposition.

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

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.

To prove Proposition 2 we use the restriction on allocations that result from Lemma 1, which, when $u(\cdot) = \ln(.)$ reduces to

$$\frac{c(\theta^t, z^t)}{c(\hat{\theta}^t, z^t)} = \frac{\mathbb{E}\left[c(\theta^t, \theta', z^{t+1})|\theta^t, z^{t+1}\right]}{\mathbb{E}\left[c(\hat{\theta}^t, \theta', z^{t+1})|\hat{\theta}^t, z^{t+1}\right]}. \tag{7}$$

By adding (7) across θ^t, we obtain

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

where $Y\left(z^t\right)$ denotes total production.

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 important 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 3, below, shows that it is the former, rather than the latter.

Proposition 3 *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 $y_t(\theta^t, z^t) = \tilde{y}_t(\theta^t) z_t$.*

We have just shown that, when $\rho = 1$,[7] the same type of separability and lack of memory with respect to aggregate shocks that characterizes efficient allocations in the first best and in the pure heterogeneity world is present in an environment with idiosyncratic risk and asymmetric information. The main normative goal of this paper is to find out whether this result applies more generally.

As we shall see, absence of memory is a very particular result. The important thing about ln preferences is that they make the marginal value of income independent of the agency problem, as shown by Demange [2008]. This is not true for $\rho \neq 1$. An heuristic account of why this property fails to be optimal in general, i.e., of why memory is relevant, obtains when we consider the following two period example.

Example 1 *Let $U_t(\theta^t, z^t)$ denote the temporary utility associated with type θ^t when aggregate history is z^t. Let, $U_t(\theta^t, z^t | \tilde{\theta}^t)$ denote the temporary utility associated with type $\tilde{\theta}^t$ when aggregate history is z^t and the agent announced θ^t.*

Hence,

$$U_t(\theta^t, z^t) = u(c_t(\theta^t, z^t)) - v\left(\frac{c_t(\theta^t, z^t)}{\theta_t z_t}\right), \tag{8}$$

and

$$U_t(\theta^t, z^t | \tilde{\theta}^t) = u(c_t(\theta^t, z^t)) - v\left(\frac{c_t(\theta^t, z^t)}{\tilde{\theta}_t z_t}\right). \tag{9}$$

[7] Note that we never used $v(l) = l^\gamma / \gamma$.

We can write the two period problem as

$$\max \sum \mu(\theta_1) \left\{ U_1(\theta_1, z_1) + \beta \sum \sum \pi(z_2) \mu(\theta^2 | \theta_1) U_2(\theta^2, z_1, z_2) \right\}$$

subject to

$$U_2(\theta^2, z_1, z_2) \geq U_2(\tilde{\theta}^2, z_1, z_2 | \theta^2), \tag{10}$$

$$\begin{aligned} U_1(\theta_1, z_1) + \beta \sum \sum \pi(z_2) \mu(\theta^2 | \theta_1) U_2(\theta^2, z_1, z_2) &\geq \\ U_1(\tilde{\theta}_1, z_1 | \theta_1) + \beta \sum \sum \pi(z_2) \mu(\theta_1, \theta_2 | \theta_1) U_2(\tilde{\theta}_1, \theta_2, z_1, z_2), & \end{aligned} \tag{11}$$

and

$$\sum \mu_t(\theta^t) \left\{ c_t(\theta^t, z^t) - y_t(\theta^t, z^t) \right\} \leq 0. \tag{12}$$

Incentive constraint (10) requires truth-telling in period 2 *regardless of whether the agent announced truthfully or not in the first period. Using (9), however, it is easy to see that*

$$U_t(\theta^t, z^t | \tilde{\theta}^t) = U_t(\theta^t, z^t | \bar{\theta}^{t-1}, \tilde{\theta}_t), \tag{13}$$

For all $\theta^t, \tilde{\theta}^t, \bar{\theta}^{t-1}, z^t$*. The consequence is that imposing*

$$U_2(\theta_1, \tilde{\theta}, z^2 | \theta_1, \theta) \leq U(\theta_1, \theta, z^2)$$

in period 2 *suffices.*

Once we have imposed (10), first period incentive constraint only requires that telling the truth in both periods yields a higher expected utility than announcing falsely in the first period and telling the truth in period 2.

Assume the allocation is memoryless, that is $(c_2(\theta^2, z^2), y_2(\theta^2, z^2)) = (\hat{c}_2(\theta^2, z_2), \hat{y}_2(\theta^2, z_2))$*, then* $U_2(\theta^2, z^2) = \hat{U}_2(\theta^2, z_2)$*. Now, using (13) constraint (11) becomes*

$$U_1(\theta_1, z_1) - U_1(\tilde{\theta}_1, z_1 | \theta_1) \geq \beta \sum \sum \pi(z_2) \mu(\theta_1, \theta_2 | \theta_1) \left[\hat{U}_2(\tilde{\theta}_1, \theta_2, z_2) - \hat{U}_2(\theta^2, z_2) \right]. \tag{14}$$

Because the right hand side of (14) does not depend on z_1*, if the constraint binds in both agregate states, e.g., the most productive agent in period 1 not pretending to be the least productive one,*[8] *then, it must be the case that*

$$U_1(\bar{\theta}, \bar{z}) - U_1(\underline{\theta}, \bar{z} | \bar{\theta}) = U_1(\bar{\theta}, \underline{z}) - U_1(\underline{\theta}, \underline{z} | \bar{\theta}).$$

[8] This is necessarily the case in our separable specification when idiosyncratic shocks are i.i.d.

Finally note that using Proposition 2, this requires separability in period 1 *and* $\eta(\bar{z}) = \eta(\underline{z})$.[9] *A contradiction.*

To generalize the result for multiple periods, we first show that incentive compatibility is equivalent to being immune to one node deviation strategies—Lemma 2, in the appendix. Then, we define $w_t(\theta^{t-1}, z^t)$ as the expected continuation utility for an agent with history θ^{t-1} conditional on aggregate history z^t, so that it satisfies

$$w_t(\theta^{t-1}, z^t) = \sum_{\theta^t \succ \theta^{t-1}} \mu_t\left(\theta^t|\theta^{t-1}\right)\left[u\left(c_t(\theta^t, z^t)\right) - v\left(\frac{y_t(\theta^t, z^t)}{\theta_t z_t}\right)\right.$$
$$\left. + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) w_{t+1}(\theta^t, z^{t+1})\right]. \tag{15}$$

We also define as $\hat{w}_t(\theta^{t-1}, z^t|\bar{\theta})$ as the continuation utility of an individual with history $(\theta^{t-2}, \bar{\theta})$ gets from telling the truth up to period $t-2$ and announcing θ_{t-1} in period $t-1$ aggregate history z^{t-1} when he has actually realized $\bar{\theta}$. Given that idiosyncratic shocks are Markovian, everything that happened up to period $t-1$ only affects the distribution of types in period t, so that from period $t+1$ on the distribution of future shocks are identical for agent type-θ^{t-1} and $(\theta^{t-2}, \bar{\theta})$. Formally,

$$\hat{w}_t(\theta^{t-1}, z^t|\bar{\theta}) = \sum_{\theta} \mu\left(\theta|\bar{\theta}\right)\left[u\left(c_t(\theta^{t-1}, \theta, z^t)\right) - v\left(\frac{y_t(\theta^{t-1}, \theta, z^t)}{\theta z_t}\right)\right.$$
$$\left. + \beta \sum_{z_{t+1}} \pi(z_{t+1}) w_{t+1}(\theta^{t-1}, \theta, z^{t+1})\right]. \tag{16}$$

In the Appendix (Lemma 2) we show that an allocation is incentive compatible if and only if one period deviation proofness is satisfied, i.e.,

$$u\left(c(\theta^t, z^t)\right) - v\left(\frac{y(\theta^t, z^t)}{\theta_t z_t}\right) + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) w(\theta^t, z^{t+1}) \quad \geq$$
$$u\left(c(\theta^{t-1}, \bar{\theta}, z^t)\right) - v\left(\frac{y(\theta^{t-1}, \bar{\theta}, z^t)}{\theta_t z_t}\right) + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) \hat{w}(\theta^{t-1}, \bar{\theta}, z^{t+1}) \quad . \tag{17}$$

[9] We have also used the fact that that if $c_1(\theta_1, z_1)$ is separable, then $y_1(\theta_1, z_1)$ is also separable, and $\eta(z) = z^{\frac{\gamma}{\gamma+\rho-1}}$. This is simple to see in this two type two period example using the first order conditions of the associated Lagrangian.

Equation (17) is the multi-period version of (11) in example 1. A similar argument allows us to convey the reason for the optimality of memory in this general case. Indeed, consider a transformation of variables similar to the one used in the proof of Proposition 3 for the case $\rho \neq 1$. Let, $c(\theta^t, z^t) = \tilde{c}(\theta^t, z^t)\eta(z_t)$, and $y(\theta^t, z^t) = \tilde{y}(\theta^t, z^t)\eta(z_t)$ where $\eta(z) = z^{\frac{\gamma}{\rho+\gamma-1}}$.

If the allocation is memoryless from period $t+1$ on, then it is separable in consumption in t according to Proposition 2. It is also possible to show that separability in consumption implies separability in output—Lemma 3—and that $\eta(z) = z^{\frac{\gamma}{\rho+\gamma-1}}$ —Lemma 4. Using lack of memory and separability, the one period incentive constraint becomes

$$\left[\frac{\tilde{c}_t(\theta^t, z^{t-1})^{1-\rho}}{1-\rho} - \frac{\tilde{y}_t(\theta^t, z^{t-1})^{\gamma}}{\theta^{\gamma}}\right] - \left[\frac{\tilde{c}_t(\hat{\theta}^t, z^{t-1})^{1-\rho}}{1-\rho} - \frac{\tilde{y}_t(\hat{\theta}^t, z^{t-1})^{\gamma}}{\theta^{\gamma}}\right] \geq$$
$$\beta\left\{\sum_{z^{t+1}\succ z^t} \pi(z^{t+1})\hat{w}_{t+1}(\hat{\theta}^t, z_{t+1}) - \sum_{z^{t+1}\succ z^t} \pi(z^{t+1}) w_{t+1}(\theta^t, z_{t+1})\right\} \quad \eta(z_t)^{\rho-1}. \tag{18}$$

Assume that this expression holds as an equality, i.e., the constraint binds, for both $z_t = \bar{z}$ and $z_t = \underline{z}$. Because the left hand side is independent of z_t, the term in right hand side cannot depend on z_t. But the right hand side of (18) does depend on z^t. It is interesting now to compare with the $u = \ln$ case. Under the transformation of variables used in the proof of Proposition 3, the incentive constraint becomes

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

It is easy to see that z_t cancels out and 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. Contrary to the ln case in which the transformation, $\eta(z_t)$, is additive in utility the transformation in (18) is multiplicative, which is not neutral in the additively separable structure that characterizes the von-Neumann Morgenstern preferences.

Equation (18) also is suggestive of how allocations should depart from separability. As we know, it is efficient to have separable memoryless allocations when there are no incentive problems to be handled. Now, start from the separable allocation used in (18), and assume the constraint binds in state $\underline{z}$. As z increases, if the term in curly brackets is positive, the constraint will relaxed for $\rho < 1$ and violated for $\rho > 1$. It will be optimal to increase the left hand side, in the first case and it will be necessary to reduce it in the second. This is done by spreading or approaching temporary utilities, which is done efficiently by distorting less or more the labor wedge. As a consequence this adjustment process induces a procyclical tax in the first case and a countercyclical in the second.

An important assumption used in our heuristic demonstration of our memory result was that the analogous constraint binds in two different aggregate states of the world; i.e., there is at least one period t and one private history, θ^t such that the same one period deviation binds at the optimum in both aggregate states of the world. 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.[10]

Assumption A: At the optimum, for some $t < T$ there is a history (θ^{t-1}, z^{t-1}) such that

$$u\left(c(\theta^{t-1}, \theta, z^{t-1}, \bar{z})\right) - v\left(\frac{y(\theta^{t-1}, \theta, z^{t-1}, \bar{z})}{\theta_t z_t}\right) + \beta \sum_{z_{t+1}} \pi(z_{t+1}) w(\theta^{t-1}, \theta, z^{t-1}, \bar{z}, z_{t+1}) =$$

$$u\left(c(\theta^{t-1}, \hat{\theta}, z^{t-1}, \bar{z})\right) - v\left(\frac{y(\theta^{t-1}, \hat{\theta}, z^{t-1}, \bar{z})}{\theta_t z_t}\right) + \beta \sum_{z_{t+1}} \pi(z_{t+1}) \hat{w}(\theta^{t-1}, \hat{\theta}, z^{t-1}, \bar{z}, z_{t+1} | \theta),$$

and

$$u\left(c(\theta^{t-1}, \theta, z^{t-1}, \underline{z})\right) - v\left(\frac{y(\theta^{t-1}, \theta, z^{t-1}, \underline{z})}{\theta_t z_t}\right) + \beta \sum_{z_{t+1}} \pi(z_{t+1}) w(\theta^{t-1}, \theta, z^{t-1}, \underline{z}, z_{t+1}) =$$

$$u\left(c(\theta^{t-1}, \hat{\theta}, z^{t-1}, \underline{z})\right) - v\left(\frac{y(\theta^{t-1}, \hat{\theta}, z^{t-1}, \underline{z})}{\theta_t z_t}\right) + \beta \sum_{z_{t+1}} \pi(z_{t+1}) \hat{w}(\theta^{t-1}, \hat{\theta}, z^{t-1}, \underline{z}, z_{t+1} | \theta).$$

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

[10] It is also possible to show that for all histories (θ^{t-1}, z^t) and all t, $E[y(\theta^t, z^t)^{1-\gamma} | \theta^{t-1}, z^t] / \eta(z_t)^{\gamma-1}$ is independent of z_t. When there are more than two types per period, these two restrictions need not imply the result. However, if only neighborhood constraints bind at the optimum, that is, if the solution of the relaxed problem solves the complete problem–see Kapicka [2010], Farhi and Werning [2010]—, then, we can generalize our result for many types.

Proposition 4 *Assume that $\rho \neq 1$ and that types are not constant over time, then if Assumption A is valid the optimal allocation is* ***not*** *memoryless.*

Proposition 4 establishes that, if an allocation is such that the same incentive compatibility constraints binds in both aggregate states of the world in period t following a given history (θ^{t-1}, z^{t-1}) then the allocation will display dependence on z^t in $t+s$, $s > 0$. It will thus violate the condition in the definition of a memoryless allocation that it be independent of past aggregate shocks for all t, z^t, θ^t.

What this proposition shows is that absence of memory is a very knife-edge case when there is idiosyncratic risk.

Finally note that when idiosyncratic shocks are i.i.d., then at any node z^t it is always the downward constraint that binds, which means that Assumption A is guaranteed to be satisfied. Hence, the following corollary is an immediate consequence of Proposition 4.

Corollary 1 *Assume that $\rho \neq 1$ and idiosyncratic shocks are i.i.d., then the optimal allocation is not memoryless.*

Proposition 4 and its corollary provide conditions under which an allocation displays memory. If we are also interested in separability, however, this will not help. Indeed, because separability is only necessary for lack of memory, we cannot infer from Proposition 4 whether the allocations associated with $\rho \neq 1$ are separable or not. For that we shall rely on our numeric exercises. Still, the fact that separability is necessary for lack of memory implies that, the calibration of this type of models will require a large state space, if one wants to consider non-separable allocations.

To conclude this section, the next Proposition shows that separability always characterizes the period T allocation.

Proposition 5 *The optimal allocation is separable in period T, 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*

$$c(\theta^T, z^T) = \tilde{c}(\theta^T, z^{T-1})\eta(z_T) \text{ and } y(\theta^T, z^T) = \tilde{y}(\theta^T, z^{T-1})\eta(z_T).$$

4 Macro Consequences of Private Memory

Modern macroeconomics has strongly relied on a representative agent setting to advance our understanding of business cycles, the effect of fiscal policies, etc. Notwithstanding its usefulness, the use of a representative agent is known to be based on strong assumptions about preferences, endowment processes, or market structure. For example, if individuals are identical with respect to their preferences and endowment processes an obvious representation obtains. Alternatively if markets are complete a standing representative agents is justified even if individuals are heterogeneous. Under complete markets, the welfare theorems allow one to focus on efficient allocations, which are relatively easy to characterize, bypassing the need for solving for the equilibrium.

In contrast, in our economy agents are heterogeneous and markets are necessarily incomplete for insurance is restricted by the informational structure of the problem. The allocations derived herein are, however, *constrained* efficient. As we shall see, qualitative predictions for macro-aggregates still obtains in our model.

The underlying idea behind the use of a normative prescriptions as a positive theory is that many different and sometimes informal institutions can implement optimal allocations. One cannot 'a priori' rule out the possibility that current institutions are producing constrained efficient outcomes. When focus is shifted from specific policies to allocations one may 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–e.g, Townsend [1994], Ligon et al. [2002] etc. An alternative approach, 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, 2008, 2009] who have used the pricing kernel associated with a dynamic Mirrlees economy to price the excess return of equity over risk free bonds and the conditional return on forward exchange rate. We focus on the business cycle behavior of the labor wedge.

Chari et al. [2007] emphasize the role played by the so-called labor wedge in generating the observed pattern of 'choices' for the representative agent through the business cycles.[11] This behavior of labor wedges is at odds with our intuition regarding tax smoothing borrowed from **?**. Using a dynamic Mirrlees setting, however, we show that it need not be inefficient. Whether one may use this result to take the real business cycle literature idea that fluctuations are optimal responses to real technological shock one step further and ask whether state-varying labor wedges are optimal (constrained efficient, in this case) responses to real shocks in a world where asymmetric information precludes full insurance, is a question that goes beyond our reach in this paper.

We see our exercises as a first, acknowledgedly small, step in the agenda of integrating the literature that studies the optimal adjustments of fiscal policies over the business cycle in the Ramsey Taxation environment (for example, **?**, **?** with the recent literature extending the insights first shown by Mirrlees [1971] to dynamic economies, the so called New Dynamic Public Finance literature.

4.1 The Labor Wedge

Chari et al. [2007] have shown how equilibrium allocations for disaggregated model economies may in many circumstances be viewed as equilibrium allocations for a representative agent economy with suitably defined time-varying wedges. This observation is used to identify which particular wedges are more important to generate the observed behavior of macroeconomic aggregates for a representative agent along the business cycle; a procedure they have named Business Cycle Accounting. We borrow from them the idea of defining a labor wedge for the representative agent despite recognizing that the economy need not be one for which aggregation of individual choices is possible.

We start by defining the aggregate variables we are going to use. Aggregate consumption is $C(z^t) \equiv \sum \mu(\theta^t)c(\theta^t, z^t)$, while output is $Y(z^t) \equiv \sum \mu(\theta^t)y(\theta^t, z^t)$. Because there is no aggregate savings in our economy, $C(z^t) = Y(z^t)$ for all z^t. To derive the economy's productivity, $W(z^t)$, we need first to define aggregate hours,

$$L(z^t) \equiv \sum_{\theta^t} \mu(\theta^t) \frac{y(\theta^t, z^t)}{\theta_t z_t}$$

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

[11]Many potential explanations for these wedges are considered in Shimer [2009].

We shall endow the representative agent with the same preferences as individuals. Hence,[12]

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

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,

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

where we have used the fact that $C(z) = Y(z)$.[13]

Separability and The Labor Wedge We have seen that, in the pure heterogeneity case or if $u(c) = \ln c$, 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

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

while the estimated productivity is

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

Deriving the wedge is now a simple task

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

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

[12] 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].

[13] Alternatively we could have defined productivity as $z_t E[\theta_t]$. This definition obtains if we use an income-weighted average in our definition of aggregate hours. This definition of aggregate productivity is invariant to policies.

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 when allocation is separable 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 could generate such variation through composition effects, but this is not the case, in our model.

An important consequence of this observation is that simply allowing for heterogeneity will not do. We need partially insured private risks to generate variation in wedges. Moreover, separability also obtains when $\rho = 1$, according to Proposition 3. The fact that for risk aversion equal to one wedges are invariant through the business cycle raises the possibility that if it is to vary for $\rho \neq 1$ it may correlate with different signs with the cycle depending on whether $\rho > 1$ of $\rho < 1$ —see our discussion in Section 3.

5 Numeric Exercises

In this section we run numerical exercises that help illustrate some of the properties that we have derived in Section 3. This exercise is similar to those in Golosov et al. [2006]. There a two period Mirrlees economy with government spending shocks that are public information is studied. The main difference between their exercises and ours is the focus. For example, because they are not interested in investigating the role of memory of aggregate shocks, they do not consider aggregate shocks in the first period. More importantly, they do not explore the macro-consequences of their model. Kocherlakota [2005] also provides a numerical example that is similar to ours in many ways but focuses on the impact of individual shock persistence and the size of the public shocks on capital taxation. He mentions neither labor tax nor the possibility of persistent effects of aggregate shocks.

In our exercises we fix $\gamma = 2$, $\underline{\theta} = 1$, $\bar{\theta} = 2$, $\mu_1(\underline{\theta}) = \mu_1(\bar{\theta}) = .5$, $\pi(\underline{z}) = \pi(\bar{z}) = .5$, $\underline{z} = 1$ and $\bar{z} = 2$. Our measure of persistence, $\alpha \in [0, 1]$, is the value of the non-unit eigenvalue of the transition matrix for idiosyncratic shocks.[14] We will only consider symmetric transitions, $\mu(\bar{\theta}, \bar{\theta}|\bar{\theta}) = \mu(\underline{\theta}, \underline{\theta}|\underline{\theta})$. In this case, the probability of maintaining the same productivity in the next period, i. e., $\mu(\theta, \theta|\theta)$ for all θ is directly associated with our mea-

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

sure in this two type context. In particular, as we vary $\mu(\theta, \theta|\theta)$ in $[0.5, 1]$, our persistence measure varies from 0 to 1; 0 obtains in the i.i.d. case and 1 in the pure heterogeneity case.

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 pure heterogeneity 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 in both periods. In any state, 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 in a static Mirrlees economy, and indicates that only downward constraints bind at the optimum for both states.[15]

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 high productivity individuals. The logic is straightforward: 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. In our example it happens so dramatically for very low levels of persistence that the temporary 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

[15] In a dynamic Mirrlees economy with a continuum of types Farhi and Werning [2010] show that this result of no distortions on the top is only valid it the support of the distribution of types is fixed, a condition that is satisfied in our model.

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

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.

In the third set of figures we show how *backloading* varies with persistence and across states of nature. To define backloading we borrow from the concept of seed values in Fernandes and Phelan [2000]'s auxiliary problem. Assume that the individual has learnt her 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 value 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 our two period example. 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.

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 aggregate 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 aggregate 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 smaller for $\rho = 1/2$.

The last set of figures deals with the macro consequences of these micro facts. The single most important aspect is that labor wedges vary with output in the first period, provided that persistence is not full and $\rho \neq 1$. 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-cylclical. When, however, $\rho = 5$, labor wedges are countercyclical, in agreement with the stylized facts about business cycles.

In accordance with our results in Section 4, labor wedges do not vary with current output in the second period. It does vary with past aggregate shock, however.

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 need not suggest inefficiencies in a second best world.

References

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

Andrew Atkeson and Robert E. Lucas. On efficient distribution with private information. *Review of Economic Studies*, 59:427–453, 1992.

Orasio Attanasio and Nicola Pavoni. Risk sharing in private information models, with asset accumulation: Explaining the excess smoothness of consumption. Working papers, University College London, 2008.

[17]There are some possibilities, though. 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. One must be careful in using simplifying assumptions that eliminate the interesting aspects of the problem, as we have seen to be the case with log preferences.

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

Varadarajan V. Chari, Patrick J. Kehoe, and Ellen R McGrattan. New kenesian models: Not yet useful for policy analysis. *American Economic Journal: Macroeconomics*, 1(1):242–266, 2009.

Gabrielle Demange. Sharing aggregate risk under moral hazard. Technical Report 2008-27, Paris School of Economics, 2008.

Emmanuel Farhi and Iván Werning. Insurance and taxation over the life cycle. MIT mimeo, 2010.

Ana Fernandes and Christopher Phelan. A recursive formulation for repeated agency with history dependence. *Journal of Economic Theory*, 91:223–247, 2000.

Mikhail Golosov, Narayana Kocherlakota, and Aleh Tsyvinski. Optimal indirect and capital taxation. *Review of Economic Studies*, 70(3):569–587, 2003.

Mikhail Golosov, Aleh Tsyvinski, and Iván Werning. New dynamic public finance: A user's guide. *NBER Macroeconomics Annual 2006*, 2006.

Edward Green. Contractual arrangements for intertemporal trade. In Edward Prescott and Neil Wallace, editors, *Lending and the Smoothing of Uninsurable Income*. Minneapolis: University of Minnesota Press, 1987.

Bengt Holmström. Moral hazard in teams. *Bell Journal of Economics*, 10(0):74–91, 1982.

Marek Kapicka. Efficient allocations in dynamic private information economies with persistent shocks: A first order approach. UCSB mimeo., May 2010.

Narayana Kocherlakota. Zero expected wealth taxes: A mirrlees approach to dynamic optimal taxation. *Econometrica*, 73:1587–1621, 2005.

Narayana Kocherlakota and Luigi Pistaferri. Household heterogeneity and real exchange rates. *Economic Journal*, 117:C1–C25, 2007.

Narayana Kocherlakota and Luigi Pistaferri. Inequality and real exchange rates. *Journal of The European Economic Association*, 6, 2008.

Narayana Kocherlakota and Luigi Pistaferri. Asset pricing implications of pareto optimality with private information. *Journal of Political Economy*, 0:forthcoming, 2009.

Jonathan Levin. Multilateral contracting and the employment relationship. *Quarterly Journal of Economics*, 117(3):1075–1103, 2002.

Ethan Ligon, Jonathan Thomas, and Tim Worrall. Informal insurance arrangements with limited commitment: Theory and evidence from village economies. *Review of Economic Studies*, 69:209–244, 2002.

James A. Mirrlees. An exploration in the theory of optimal income taxation. *Review of Economic Studies*, 38:175–208, 1971.

Christopher Phelan. Incentives and aggregate shocks. *The Review of Economic Studies*, 61: 681–700, 1994.

Cristopher Phelan. On the long-run implications of repeated moral hazard. *Journal of Economic Theory*, 79:174–191, 1998.

William P. Rogerson. Repeated moral hazard. *Econometrica*, 53(1):69–76, 1985.

Florian Scheuer. Pareto-optimal taxation with aggregate uncertainty and financial markets. MIT working paper, 2009.

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

Christopher Sleet. Endogenously incomplete markets: Macroeconomic implications. Mimeo. Tepper Business School, 2006.

Stephen E. Spear and Sanjay Srivastava. On repeated moral hazard with discounting. *Review of Economic Studies*, 54:599–617, 1987.

Jonathan Thomas and Tim Worrall. Income fluctuation and asymmetric information: An example of a repeated principal-agent problem. *Journal of Economic Theory*, 51:367–390, 1990.

Robert Townsend. Risk and insurance in village india. *Econometrica*, 62(3):539–591, 1994.

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

Iván Werning. Optimal fiscal policy with redistribution. *Quarterly Journal of Economics*, 122:925–967, 2007.

Robert Wilson. The theory of syndicates. *Econometrica*, 36:119–132, 1968.

A Appendix

A.1 General Proofs

Proof. of Lemma 1 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 $\varepsilon(\Delta)$. That is, $u(\hat{c}(\bar{\theta}^t, z^t)) = u(c(\bar{\theta}^t, z^t)) - \varepsilon(\Delta)$. For simplicity, consider $\mu(\theta^t) = \mu(\bar{\theta})$. For small enough Δ, the two are related by $\varepsilon(0) = 0$ and

$$\varepsilon'(0) = \frac{u'(c(\bar{\theta}^t, z^t))}{u'(c(\theta^t, z^t))}. \tag{19}$$

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, fix some $z^{t+1} \succ z^t$, and for every $\theta^{t+1} \succ \theta^t$,

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

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

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

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

$$-\sum_{\theta^{t+1}\succ\theta^t}\frac{\mu(\theta^{t+1}|\theta^t)}{u'(c(\theta^{t+1},z^{t+1}))}\Delta+\sum_{\bar{\theta}^{t+1}\succ\bar{\theta}^t}\frac{\mu(\bar{\theta}^{t+1}|\bar{\theta}^t)}{u'(c(\bar{\theta}^{t+1},z^{t+1}))}\varepsilon'(0)\Delta. \tag{20}$$

Substituting (19) in (20) 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. ■

Proof. of Proposition 1 For simplicity we assume the existence of only two types, namely, $\bar{\theta}$ and $\underline{\theta}$. The argument is easily extended. If $\mu\left(\theta',\theta|\theta'\right)=1$, for $\theta=\theta'$, 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

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

and

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

for $\theta,\hat{\theta}=\bar{\theta},\underline{\theta}$.

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 $\phi\left(\hat{\theta}|\theta\right)$ and $\phi\left(\theta|\hat{\theta}\right)$. The first order conditions associated with $c_t(\theta,z^t)$ and $y_t(\theta,z^t)$ are

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

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

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

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

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

■

Proof. of Proposition 2 Define functions $\tilde{c}(\theta^t,z^t)$ and $\eta(z_t)$ such that $c(\theta^t,z^t)=\tilde{c}(\theta^t,z^t)\eta(z_t)$. Note that this is without loss. If an allocation is not separable in t then, there must be a pair $\theta^t,\hat{\theta}^t$ such that $\tilde{c}(\theta^t,z^t)/\tilde{c}(\hat{\theta}^t,z^t)$ is a function of z_t. Using Lemma 1, however, gives

$$\frac{\tilde{c}(\theta^t,z^t)^\rho}{\tilde{c}(\hat{\theta}^t,z^t)^\rho}=\frac{\mathbb{E}\left[c(\theta^t,\theta',z^{t+1})^\rho|\theta^t,z^{t+1}\right]}{\mathbb{E}\left[c(\hat{\theta}^t,\theta',z^{t+1})^\rho|\hat{\theta}^t,z^{t+1}\right]}. \tag{21}$$

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, which shows that the allocation depends on z^t. ■

Proof. of Proposition 3 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

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

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, $\tilde{c}(\theta^t, z^t) = \tilde{c}(\theta^t), \tilde{y}(\theta^t, z^t) = \tilde{y}(\theta^t)$. ■

Proof. of Proposition 5 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 $c_T(.) = \tilde{c}_T(.)\, z_T^{\frac{\gamma}{\gamma+\rho-1}}$ and $y_T(.) = \tilde{y}_T(.)\, z_T^{\frac{\gamma}{\gamma+\rho-1}}$, 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

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

And the resource constraints turns into

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

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. ■

A.2 Proof of Proposition 4

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

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

and

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

Let us define for every θ^t the set $\Gamma^{\sigma}(\theta^t)$ through

$$\Gamma^{\sigma}(\theta^t) \equiv \{(\tilde{\theta}^t, z^t) \in \Theta^t \times Z^t; \sigma_t(\tilde{\theta}^t, z^t) = \theta^t\}.$$

Then we can note that

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

if $(\theta^t, z^t) \in \Gamma^{\sigma}(\tilde{\theta}^t)$, while

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

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

$$\max \sum_{t=1}^{T} \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t)\mu_t\left(\theta^t\right)\left[\mathfrak{u}(\theta^t, z^t) - \mathfrak{v}(\theta^t, z^t)\right]$$

subject to

$$\sum_{t=1}^{T} \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t)\mu_t\left(\theta^t\right)\left[\mathfrak{u}(\theta^t, z^t) - \mathfrak{v}(\theta^t, z^t)\right] \geq$$
$$\sum_{t=1}^{T} \beta^{t-1} \sum_{\theta^t} \sum_{z^t} \pi_t(z^t)\mu_t\left(\theta^t\right)\left[\mathfrak{u}(\theta^t, z^t|\sigma) - \mathfrak{v}(\theta^t, z^t|\sigma)\right]$$

for all $\sigma \in \Sigma$,

$$\sum_{\theta^t} \mu_t\left(\theta^t\right)\left[C(\mathfrak{u}(\theta^t, z^t)) - \Psi(\mathfrak{v}(\theta^t, z^t))\right] \leq 0$$

for all t, z^t.

This is a concave program (linear objective, and concave restrictions). Before we do so, however, we shall reduce the number of incentive compatibility constraints. That is we will exploit the fact that an allocation is incentive compatible if and only if it is immune to one node deviations. This is equivalent to proving that an allocation is incentive compatible if and only if it satisfies

$$\begin{aligned} u\left(c(\theta^t, z^t)\right) - v\left(\frac{y(\theta^t, z^t)}{\theta_t z_t}\right) + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) w(\theta^t, z^{t+1}) \quad &\geq \\ u\left(c(\theta^{t-1}, \bar{\theta}, z^t)\right) - v\left(\frac{y(\theta^{t-1}, \bar{\theta}, z^t)}{\theta_t z_t}\right) + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{t+1}) \hat{w}(\theta^{t-1}, \bar{\theta}, z^{t+1}) \quad &. \end{aligned} \tag{22}$$

for all $t \leq T$, (θ^t, z^t), and $(\theta^{t-1}, \bar{\theta}, z^t)$. For completeness at the final node define $w(\theta^{t+1}, z^{T+1}) = \hat{w}(\theta^{t+1}, z^{T+1} | \bar{\theta}) = 0$ for all $\theta^{T+1}, \bar{\theta}, z^{T+1}$.

Lemma 2 *An allocation is incentive compatible if and only if it satisfies inequalities (22).*

Proof. To show that (5) implies (22) is immediate. Just note that if there was a one period one state deviation that yielded more expected utility following any history that occurs with positive probability (since we are not considering the pure heterogeneity 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 truth-telling. This contradicts the assumption that (3) is satisfied.

To prove the converse we assume that (5) is violated, i.e., there is a strategy σ that yields more expected utility than the truthful strategy. What we will show is that we will be able to define a strategy that only violates one constraint of the type (22). Next, define τ, z^τ as the last node for which strategy σ recommends lying in that specific period. Let $\sigma_{\tau-1}(\theta^{\tau-1}) = \tilde{\theta}^{\tau-1}$ and note that we can assume, without loss that

$$\begin{aligned} u(c_\tau(\sigma(\theta^\tau, z^\tau), z^\tau) - v\left(\frac{y_t(\sigma(\theta^\tau, z^\tau), z^\tau)}{\theta_\tau z_\tau}\right) + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{\tau+1}) \hat{w}(\sigma(\theta^\tau, z^\tau), z^{\tau+1}) \quad &> \\ u(c_t(\tilde{\theta}^{\tau-1}, \theta_\tau, z^\tau) - v\left(\frac{y_t(\tilde{\theta}^{\tau-1}, \theta_\tau, z^\tau)}{\theta_\tau z_\tau}\right) + \beta \sum_{z^{t+1} \succ z^t} \pi(z^{\tau+1}) w(\tilde{\theta}^{\tau-1}, \theta_\tau, z^{\tau+1}) & \end{aligned}$$

if $\tau < T$ or

$$u(c_\tau(\sigma(\theta^\tau, z^\tau), z^\tau) - v\left(\frac{y_t(\sigma(\theta^\tau, z^\tau), z^\tau)}{\theta_\tau z_\tau}\right) \quad > \\ u(c_t(\tilde{\theta}^{\tau-1}, \theta_\tau z^\tau) - v\left(\frac{y_t(\tilde{\theta}^{\tau-1}, \theta_\tau, z^\tau)}{\theta_\tau z_\tau}\right)$$

if $\tau = T$.

To understand why this is without loss just note that if this condition is violated, then the strategy of following σ until the $z^{\tau-1} \prec z^\tau$ and telling the truth from τ, z^τ on is weakly better than following σ, so we may as well focus on this latter strategy.

Now consider the strategy of always announcing truthfully unless state $(z^\tau, \tilde{\theta}^{\tau-1}, \theta_\tau)$ realizes, in which case, one announces $\sigma(\tilde{\theta}^{\tau-1}, \theta_\tau, z^\tau)$. This is a one state deviation strategy that improves on truth-telling, thus violating (22). ■

Lemma 3 *If the allocation is separable in c it is also separable in y.*

Proof. Define as $\lambda(z^t)$ the multiplier related to the resource constraint in aggregate state z^t. Consider first order condition with respect to $c(\theta^t, z^t)$,

$$\mu(\theta^t|\theta^{t-1})\lambda(z^t) - c(\theta^t, z^t)^\rho \mu(\theta^t|\theta^{t-1}) = \sum_{\sigma:(\tilde{\theta}^t, z^t)\in\Gamma^\sigma(\theta^t)} \delta^\sigma(\theta^t, z^t|\tilde{\theta}^t, z^t)/\pi(z^t) \\ - \sum_{\sigma:(\theta^t, z^t)\in\Gamma^\sigma(\tilde{\theta}^t)} \delta^\sigma(\tilde{\theta}^t, z^t|\theta^t, z^t)/\pi(z^t), \qquad (23)$$

where $\delta^\sigma(\theta^t, z^t|\tilde{\theta}^t, z^t)$ is the multiplier associated with the constraint that an agent following strategy σ will not find it in his interest to announce θ^t at the node $(\tilde{\theta}^t, z^t)$.

Under separability on consumption, the left hand side is

$$\eta(z_t)\left[\mu(\theta^t|\theta^{t-1})\mathbb{E}[\tilde{c}(\theta^t, z^{t-1})^\rho] - \tilde{c}(\theta^t, z^{t-1})^\rho \mu(\theta^t|\theta^{t-1})\right]$$

which is separable in z^t.

Using the fact that we can restrict ourselves to one period deviation strategies and because we have only two types, only one multiplier δ is positive. Take, then the type which allocation is not envied at z^t. Then, the left hand side of (23) is

$$\delta\left(\tilde{\theta}^t, z^t|\theta^t\right)/\pi(z^t) = \eta(z_t)\tilde{\delta}\left(\tilde{\theta}^t, z^{t-1}|\theta^t\right)/\pi(z^t)$$

with $\tilde{\delta}(\tilde{\theta}^t, z^{t-1}|\theta^t)/\pi(z^t)$ independent of z_t. Since the first order condition with respect to $y(\theta^t, z^t)$ is

$$\theta_t^{-\gamma}\left\{\mu(\theta^t|\theta^{t-1})\lambda(z^t) + \delta(\tilde{\theta}^t, z^t|\theta^t)/\pi(z^t)\right\} = \mu(\theta^t|\theta^{t-1})z_t^{\gamma}y(\theta^t, z^t)^{\gamma-1}$$

Then, the right hand side of the expression above is also independent of z_t. Therefore,

$$\frac{y(\theta^t, z^{t-1}, z)^{\gamma-1}}{y(\theta^t, z^{t-1}, \hat{z})^{\gamma-1}} = \frac{\eta(z)\hat{z}^{\gamma}}{\eta(\hat{z})z^{\gamma}}.$$

A similar procedure applies to $\tilde{\theta}$ to show that both are separable. ■

Lemma 4 *If the allocation is separable then* $\eta(z) = z^{\frac{\gamma}{\gamma+\rho-1}}$.

Proof. Note that adding the first order conditions with respect to y and using $\lambda(z^t) = \mathbb{E}[c(\theta^t, z^{t-1})^{\rho}\theta^{t-1}]\eta(z_t)^{\rho}$ we get

$$\frac{\mathbb{E}[\theta_t^{-\gamma}|\theta_{t-1}]\mathbb{E}[c(\theta^t, z^{t-1})^{\rho}|\theta^{t-1}]}{\mathbb{E}[y(\theta^t, z^{t-1})^{1-\gamma}|\theta^{t-1}]}\eta(z_t)^{\rho} = \eta(z_t)^{1-\gamma}z_t^{\gamma},$$

which implies $\eta(z) = z^{\frac{\gamma}{\gamma+\rho-1}}$. ■

Proof. of Proposition 4: Assume that the optimal allocation does not display memory. This implies that utility promises in $\tau+1$ are independent of z_{τ}. Define $\tilde{c}(\theta^t, z^t) = c(\theta^t, z^t)/\eta(z_t)$ and $\tilde{y}(\theta^t, z^t) = y(\theta^t, z^t)/\eta(z_t)$, with $\eta(z) = z^{\frac{\gamma}{\gamma+\rho-1}}$. Note that due to Lemma 3, above, if the allocation is separable then both $\tilde{c}(\theta^t, z^t)$ and $\tilde{y}(\theta^t, z^t)$ are not functions of z_t. Note that an allocation is incentive compatible if and only if

$$\frac{\tilde{c}\left(\theta^t, z^t\right)^{1-\rho}}{1-\rho} - \frac{\tilde{y}\left(\theta^t, z^t\right)^{\gamma}}{\gamma\theta_t^{\gamma}} - \left[\frac{\tilde{c}\left(\theta^{t-1}, \theta, z^t\right)^{1-\rho}}{1-\rho} - \frac{\tilde{y}\left(\theta^{t-1}, \theta, z^t\right)^{\gamma}}{\gamma\theta_t^{\gamma}}\right] \geq$$
$$\beta\left\{\sum_{z^{t+1}\succ z^t}\pi(z^{t+1})\left[\hat{w}(\theta^{t-1}, \theta, z_{t+1}) - w(\theta^t, z_{t+1})\right]\right\}\eta(z_t)^{\rho-1} \tag{24}$$

Recall that, under assumption A, for some $t < T$ there is a history (θ^{t-1}, z^{t-1}) such that

$$u\left(c(\theta^{t-1}, \theta, z^{t-1}, \bar{z})\right) - v\left(\frac{y(\theta^{t-1}, \theta, z^{t-1}, \bar{z})}{\theta_t z_t}\right) + \beta\sum_{z_{t+1}}\pi(z_{t+1})w(\theta^{t-1}, \theta, z^{t-1}, \bar{z}, z_{t+1}) =$$
$$u\left(c(\theta^{t-1}, \hat{\theta}, z^{t-1}, \bar{z})\right) - v\left(\frac{y(\theta^{t-1}, \hat{\theta}, z^{t-1}, \bar{z})}{\theta_t z_t}\right) + \beta\sum_{z_{t+1}}\pi(z_{t+1})w(\theta^{t-1}, \hat{\theta}, z^{t-1}, \bar{z}, z_{t+1}),$$

and

$$u\left(c(\theta^{t-1},\theta,z^{t-1},\underline{z})\right)-v\left(\frac{y(\theta^{t-1},\theta,z^{t-1},\underline{z})}{\theta_t z_t}\right)+\beta\sum_{z_{t+1}}\pi(z_{t+1})w(\theta^{t-1},\theta,z^{t-1},\underline{z},z_{t+1})=$$

$$u\left(c(\theta^{t-1},\hat{\theta},z^{t-1},\underline{z})\right)-v\left(\frac{y(\theta^{t-1},\hat{\theta},z^{t-1},\underline{z})}{\theta_t z_t}\right)+\beta\sum_{z_{t+1}}\pi(z_{t+1})w(\theta^{t-1},\hat{\theta},z^{t-1},\underline{z},z_{t+1}),$$

which means that (24) holds as an equality for some θ^t at both aggregate states. The term in curly brackets in the right hand side of expression (24) 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 $(c(\theta^t,z^t),y(\theta^t,z^t))$ is not separable. However, from Proposition 2, this cannot be the case. ■

Figure 1: Memory - $\rho = 5$

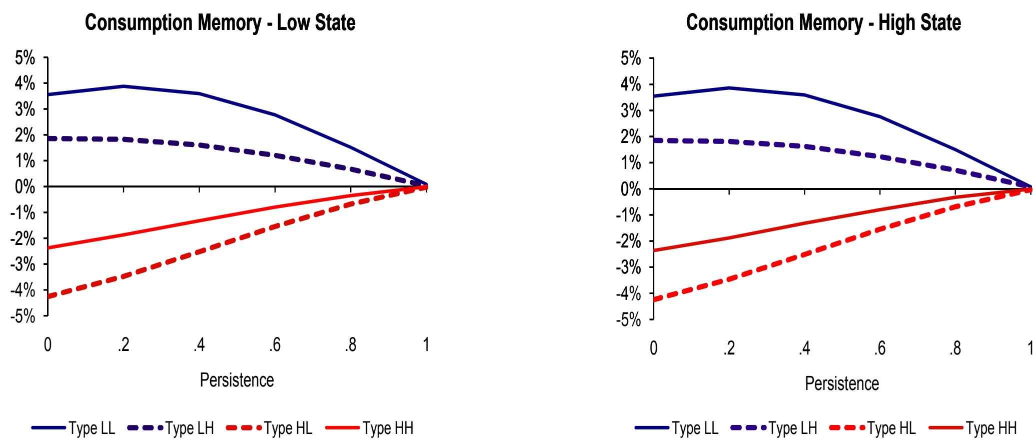


Figure 2: Memory - $\rho = .5$

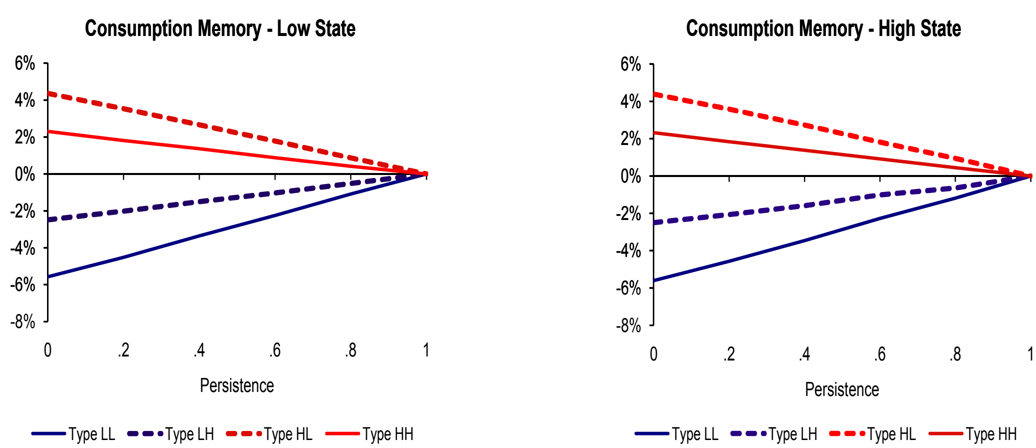
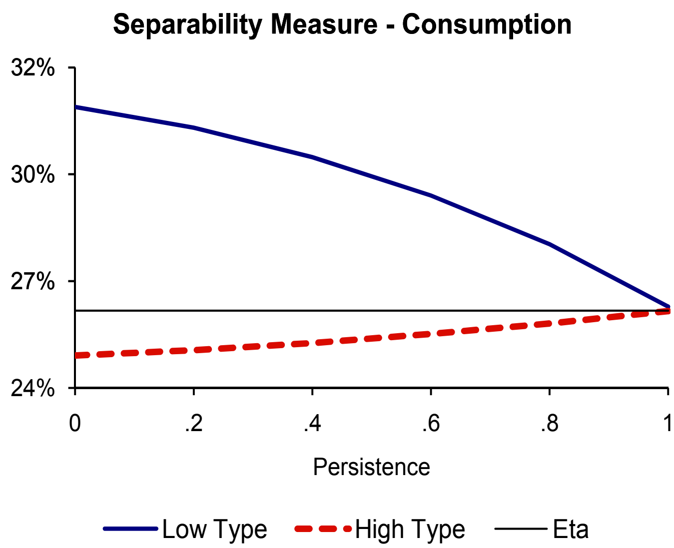


Figure 3: Separability in Consumption

Separability Measure $\rho = 5$



Separability Measure $\rho = .5$

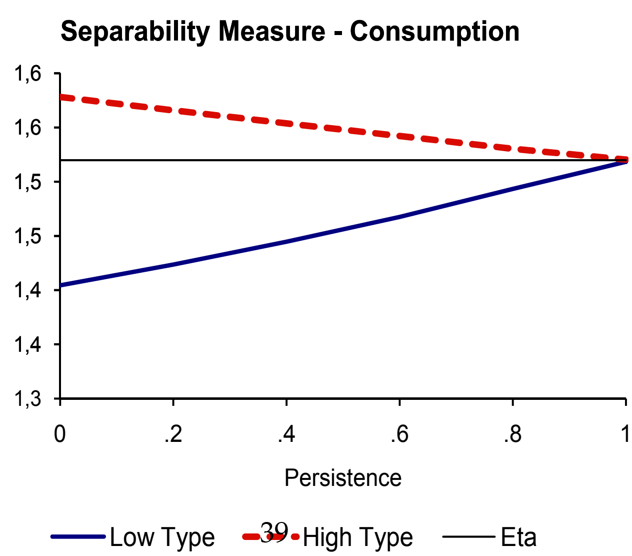


Figure 4: Backloading - $\rho = 5$

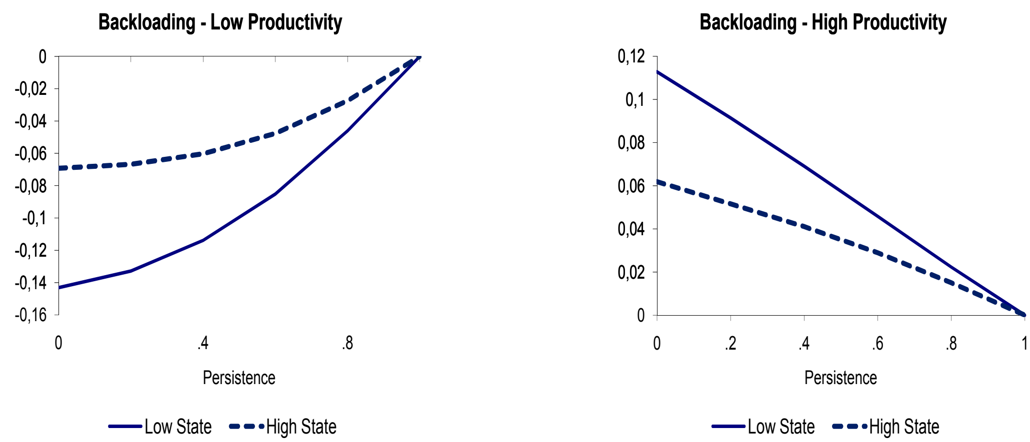


Figure 5: Backloading - $\rho = .5$

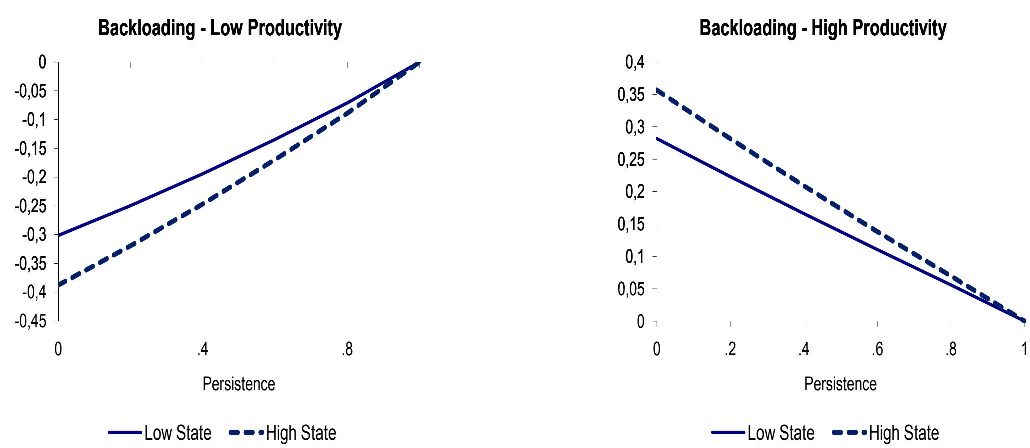


Figure 6: 1st Period Consumption and Output - $\rho = 5$

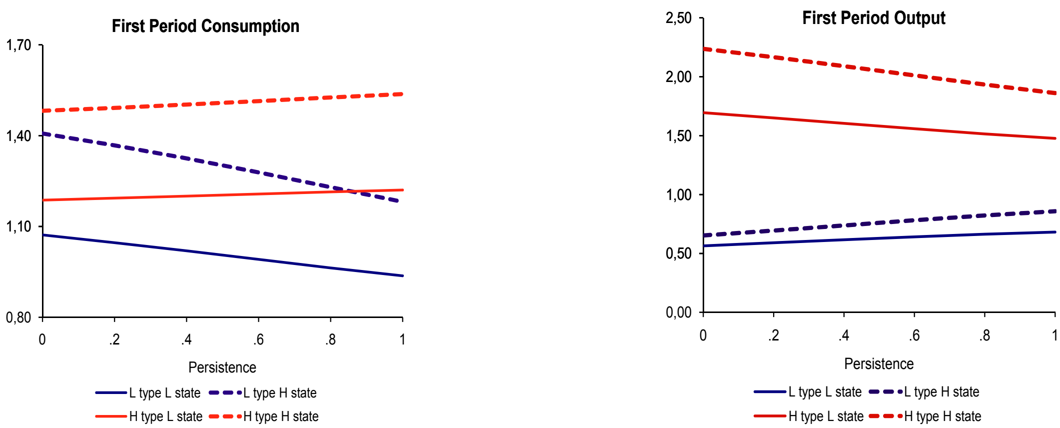


Figure 7: 1st Period Consumption and Output - $\rho = .5$

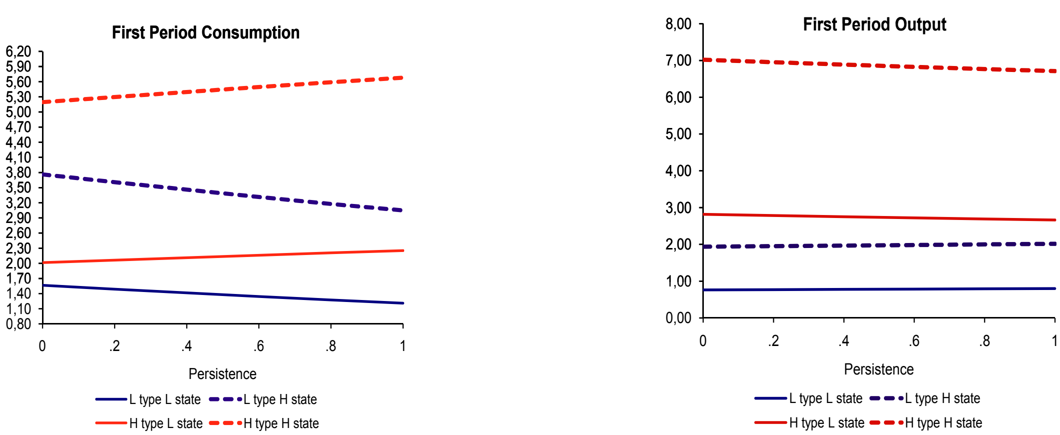


Figure 8: Labor Supply and Mg Tax Rates - $\rho = 5$

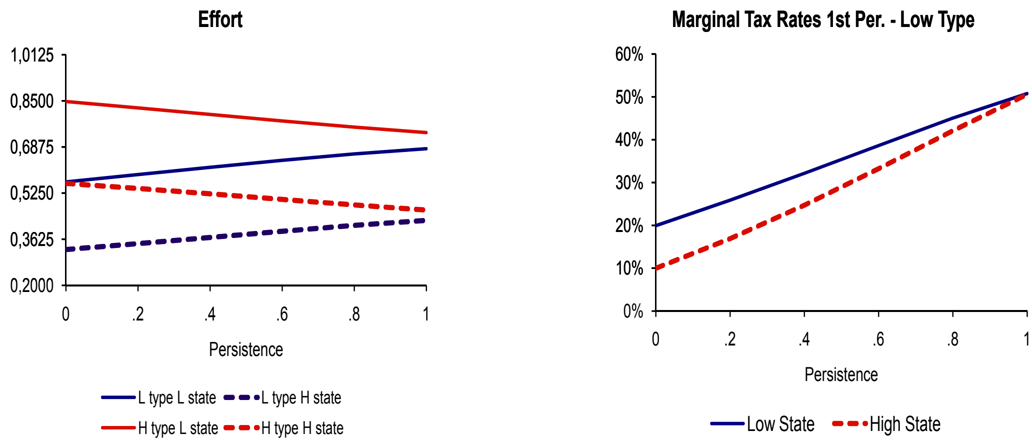


Figure 9: Labor Supply and Mg Tax Rates - $\rho = .5$

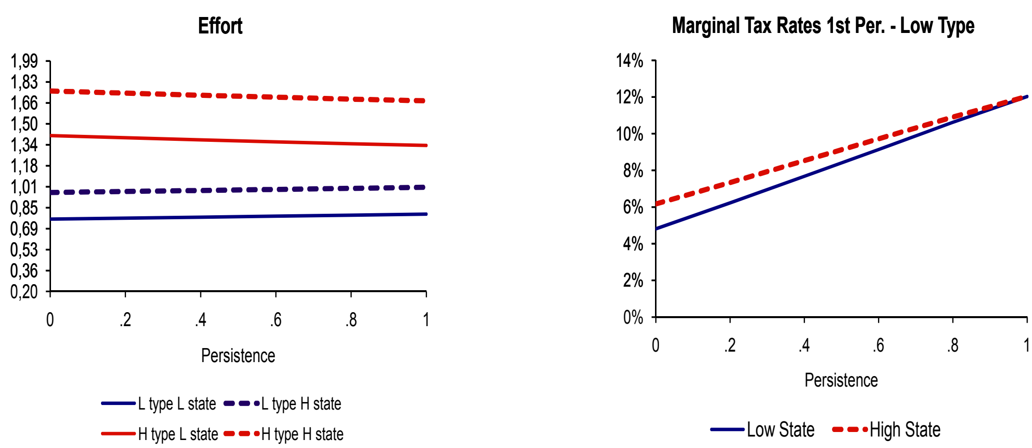
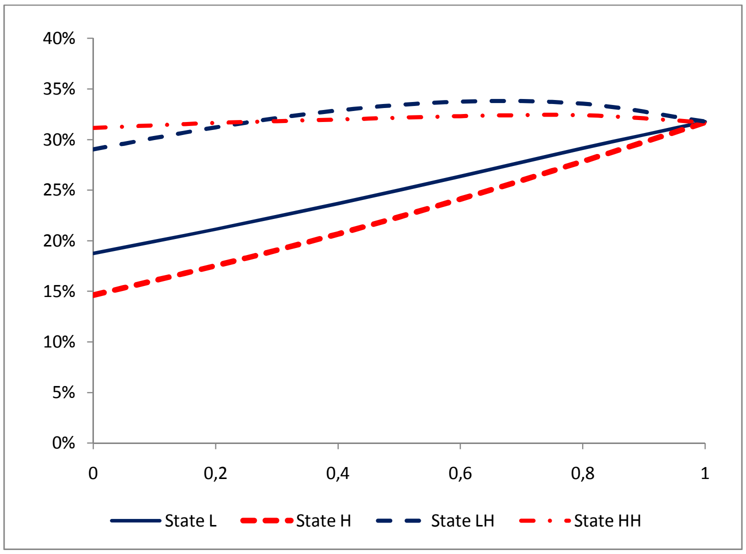
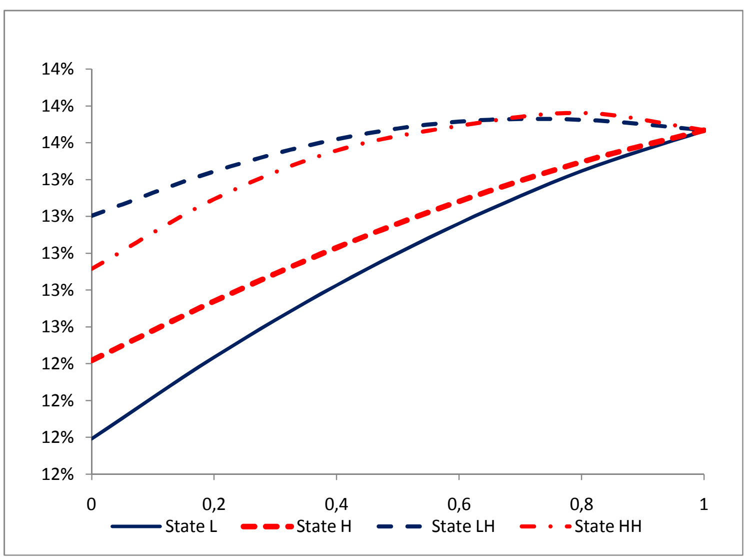


Figure 10: Labor Wedge



Labor Wedge $\rho = 5$



Labor Wedge $\rho = .5$