**Capital Taxation: Quantitative Explorations of the Inverse Euler Equation**

The MIT Faculty has made this article openly available. *Please share* how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1086/666747">http://dx.doi.org/10.1086/666747</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>University of Chicago Press, The</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Mon Mar 11 04:46:58 EDT 2019</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/82874">http://hdl.handle.net/1721.1/82874</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
Economies with private information provide a rationale for capital taxation. In this paper we ask what the welfare gains from following this prescription are. We develop a method to answer this question in standard general equilibrium models with idiosyncratic uncertainty and incomplete markets. We find that general equilibrium forces are important and greatly reduce the welfare gains. Once these effects are taken into account, the gains are relatively small in our benchmark calibration. These results do not imply that dynamic aspects of social insurance design are unimportant, but they do suggest that capital taxation may play a modest role.

I. Introduction

Advanced countries use the tax system as a tool for social insurance. Friedman (1962) proposed a “negative income tax” in which individuals with low enough income would owe a negative tax and thus collect...
transfers. In particular, he advocated a simple system in which a constant marginal tax on income is combined with a lump sum rebate—resulting in a linear tax function with a negative intercept (Sheshinski 1972). For a given distribution of before-tax income, such a tax reduces the dispersion in after-tax income, thus achieving a more equitable outcome. Mirrlees’s (1971) seminal work refined this idea by allowing for fully nonlinear taxation of income.

These contributions focus on the taxation of labor income to provide social insurance. Is there a similar role to be played by capital taxation? In an important paper, Atkinson and Stiglitz (1976) provided a negative answer using the same approach as Mirrlees. Under some conditions, they showed that the optimum taxes labor income only, leaving capital income untaxed. Their analysis rested on a model in which individuals work and save but do not face uncertainty about future earnings. A model that incorporates future uncertainty seems better suited for thinking about social insurance. Adding uncertainty turns out to make an important difference. Indeed, when individuals do face uncertainty, the optimal insurance arrangement calls for taxing capital positively: constrained efficient allocations satisfy an inverse Euler equation, as opposed to the agent’s standard intertemporal Euler equation, implying a positive capital tax (Diamond and Mirrlees 1977; Rogerson 1985; Ligon 1998; Golosov, Kocherlakota, and Tsyvinski 2003).

Hence, there is a theoretical case for positive capital taxation as part of an optimal social insurance system. The purpose of this paper is to evaluate the quantitative contribution of capital taxation toward welfare.

One can view our exercise as a counterpart to Lucas’s well-known Supply Side Economics (1990). Lucas computed the welfare gains from removing a linear tax on capital in a Ramsey model, where it is well known that capital should not be taxed (Judd 1985; Chamley 1986).¹ In a similar spirit, here, we use a dynamic version of Mirrlees’s model in which capital should be taxed and compute the welfare gains of taxing capital optimally, starting from a situation without capital taxes. Our focus, as was Lucas’s, is on the welfare implications of the capital tax itself, not those of the entire tax system.²

A key input of our model is the uncertainty of earnings. Workers are affected by various shocks that contribute toward this uncertainty, for example, unemployment risk, health shocks, or outright disability. More generally, individuals may not be able to perfectly predict the path of their future skill or earning opportunities. Many papers find it fruitful

---

¹ Thus, both Chamley and Judd reach a similar conclusion to Atkinson and Stiglitz, but in very different models. The common denominator is that both models lack idiosyncratic individual uncertainty, which is the focus of our paper.

² In Farhi and Werning (2010), we do solve for the fully optimal social insurance arrangement, including both labor and savings distortions.
to isolate a particular source of risk, such as unemployment (e.g., Hopenhayn and Nicolini 1997; Shimer and Werning 2008) or disability (e.g., Diamond and Mirrlees 1977; Golosov and Tsyvinski 2006), and study the optimal design of an insurance program that takes aim at this particular risk. However, for the purposes of evaluating the overall benefits from taxing capital, focusing on a particular risk would be a mistake. Instead, it is important to capture uncertainty more broadly, in a way that reflects estimates for overall earnings processes from panel data, not particular sources of risk. Such an approach is consistent with a large incomplete-markets literature that has been quite successful modeling consumption and wealth, taking as an input the estimated earnings process. We follow this tradition. Our method is flexible and allows us to study such model economies. In so doing, we bridge the gap between a more theoretical new dynamic public finance literature and this more applied incomplete-markets literature.

We lay down a Mirrleesian economy with neoclassical technology and idiosyncratic uncertainty. Agents consume and work, experiencing skill shocks that are private information. As a result, feasible allocations must be incentive compatible. In this setting, as long as utility is additively separable, constrained efficient allocations satisfy the inverse Euler equation. Because this condition is incompatible with the agents’ standard Euler equation, savings distortions are needed and take the form of a positive tax on capital. Starting from an equilibrium where agents save freely, introducing these optimal taxes increases welfare. We are interested in computing and understanding the determinants of these welfare gains. These measures answer the question posed earlier, regarding the importance of capital taxation as a component of social insurance.

Two approaches are possible to quantify the magnitude of these welfare gains, and we pursue both. The first approach requires modeling the stochastic process for individual consumption directly. The second models consumption indirectly. It uses as a starting point equilibrium allocations from a competitive economy in which agents are subject to idiosyncratic skill shocks and can save freely by accumulating a risk-free bond. The advantage of this approach is that direct empirical knowledge on the consumption process is more limited than that for income and wealth. We now describe our results following both approaches in turn.

We start with the first approach. We are able to obtain closed-form solutions when utility is logarithmic and consumption follows a geometric random walk. We use the simplicity of our closed-form solutions to derive some insights that illuminate the main determinants for welfare gains. In particular, our formulas highlight how the magnitude of optimal capital taxes as well as the efficiency gain from these taxes depend positively on the variance of consumption growth. As we review, the empirical literature only provides imperfect guidance for this statistic.
With a partial-equilibrium analysis, using a fixed interest rate, we find that the welfare gains span a wide range, going from 0 to 10 percent, reflecting the range of empirical evidence on the variance of consumption growth. This wide range of welfare gains leads us to consider a general equilibrium setting. We find that general equilibrium effects can dramatically reduce welfare gains, which now range from 0 to 0.25 percent.

We then turn to the second approach. We adopt an Aiyagari-Bewley incomplete-market model. Individuals face idiosyncratic labor income risk. They can save in a risk-free asset, but they cannot borrow. At a steady state equilibrium, the interest rate is constant and equal to the marginal product of capital. Although individual consumption fluctuates, the cross-sectional distribution of assets and consumption is invariant. Starting from this steady state, we introduce optimal savings distortions and compute the associated welfare gains, taking into account the resulting transitional dynamics. The solution is no longer closed form, but we show that it is sufficiently tractable to be computed numerically.

There are two important, and potentially offsetting, differences in this setting (second approach), relative to the geometric random-walk case (first approach) in which closed-form solutions were available. On the one hand, as is well known, agents are able to smooth consumption quite effectively in Aiyagari-Bewley models, minimizing the variance of consumption growth and reducing the potential welfare gains from savings distortions. On the other hand, equilibrium consumption in these models is not well described by a geometric random walk. Indeed, a steady state, with a stable cross-sectional distribution for consumption, requires a mean-reverting consumption process. We show that this tends to magnify potential welfare gains.

We compute the steady state equilibrium for a range of parameters for the income processes and the coefficient of relative risk aversion. For our baseline calibration, we find that welfare gains are relatively small, below 0.2 percent. Away from this baseline, we find that welfare gains increase with the coefficient of relative risk aversion and with the variance and persistence of the income process. However, unless one allows the planner to relax borrowing constraints, the welfare gains remain relatively modest.

Overall, our findings suggest that the welfare gains from savings distortions may be small. This does not imply that there are small welfare gains from optimal dynamic social insurance more generally, only that

---

3 Indeed, theoretically, the gains are potentially unbounded, establishing that it is incorrect to simply presume that the welfare gains are necessarily small, underscoring the need for a quantitative assessment that uncovers their main determinants.
the supporting role of savings distortions may be modest, especially when general equilibrium forces are taken into account.

On the methodological front, our strategy requires perturbing the consumption assignment and holding the labor assignment unchanged, while preserving incentive compatibility. The new allocation satisfies the inverse Euler equation and delivers the same utility while freeing up resources. The reduction in resources is our measure of welfare gains. By leaving the labor assignment unchanged, we focus on the gains from introducing savings distortions, without changing the incentive structure implicit in the labor assignment. In this way, we sidestep resolving the optimal trade-off between insurance and incentives.

There are several advantages to our approach. First, our exercise does not require specifying some components of the economy. In particular, no knowledge of individual labor assignment or the disutility of work function is required. In this way, the degree to which work effort responds to incentives is not needed. This robustness is important since empirical knowledge of these elasticities remains incomplete. Indeed, our welfare gains depend on only the original consumption assignment, the utility function for consumption, and technology. Second, our planning problem has the advantage of being tractable, even for rich specifications of uncertainty. In our view, having this flexibility is important for quantitative work. Finally, by focusing our attention on savings distortions in this way, the welfare gains we compute can be traced back to a few intuitive parameters, such as the variance of consumption growth, the coefficient of relative risk aversion, and the concavity of the production function.

Related literature.—The question we address is largely unexplored. One reason is that it is difficult to solve dynamic economies with private information, except for very particular cases, such as shocks that are independent and identically distributed (i.i.d.) over time or when shocks are binary and absorbing, as is the applications to unemployment and disability insurance (see below). The optimal taxation literature based on models with private information (see Golosov, Tsyvinski, and Werning [2006] and the references therein) often characterizes the constrained efficient allocations and taxes but rarely provides a quantitative analysis of the efficiency gains, especially in not focusing on the contribution of savings distortions. Recent progress has been made applying a “first-order approach” to characterize and compute constrained efficient allocations when shocks are persistent (Farhi and Werning 2010; Golosov, Troshkin, and Tsyvinski 2010). However, this approach remains computationally challenging and has been applied to date in partial equilibrium settings only. Moreover, the focus in these papers is on characterizing the optimum, not evaluating the contribution from capital taxes. Instead, our approach here zeros in on the contribution of capital
Overall, our results on the small welfare gains associated with savings distortions highlight the importance of focusing on labor distortions in these dynamic social insurance environments as pursued by these recent contributions.

Another difference is that much of the literature has proceeded by studying particular aspects of social insurance in isolation. In contrast, in our paper, for an evaluation of the benefits from savings distortions for social insurance, it is important to take a broader perspective. Golosov and Tsyvinski (2006) study disability insurance, and Shimer and Werning (2008) study unemployment insurance. In both cases, the nature of the stochastic process for shocks allows for a low-dimensional recursive formulation that is numerically tractable. Golosov and Tsyvinski (2006) provide a quantitative analysis of disability insurance. Disability is modeled as an absorbing negative skill shock. They calibrate their model and compute the welfare gains that can be reaped by moving from the most efficient allocation that satisfies free savings to the optimal allocation. They focus on logarithmic utility and report welfare gains of 0.5 percent. Shimer and Werning (2008) provide a quantitative analysis of unemployment insurance. They consider a sequential job search model, where a risk-averse, infinitely lived worker samples wage offers from a known distribution. Regarding savings distortions, they show that with constant absolute risk aversion (CARA) utility, allowing agents to save freely is optimal. With constant relative risk aversion (CRRA) utility, savings distortions are optimal, but the authors find that the efficiency gains they provide are minuscule. As with most quantitative exercises to date, both Golosov and Tsyvinski (2006) and Shimer and Werning (2008) are set in partial equilibrium settings with linear technologies.

There is a vast literature studying and calibrating incomplete-market Aiyagari-Bewley economies to consumption, income, and wealth data. These papers emphasize the role of consumers self-smoothing through the precautionary accumulation of risk-free assets. In most positive analyses, either government policy is ignored or else a simple transfer and tax system is included and calibrated to current policies. In some normative analyses, reforms of the transfer system, such as the income tax or social security, are evaluated numerically (e.g., Conesa and Krueger 2005). Our paper bridges the gap between this literature and the optimal tax literature.

The notion of efficiency used in the current paper is often termed constrained efficiency because it imposes the incentive-compatibility constraints that arise from the assumed asymmetry of information. Another, very different, notion has been developed for exogenously incomplete-market models (see Geanakoplos and Polemarchakis 1985). The idea is to take the available asset structure as given and ask whether individ-
uals can change their trading positions in a way that, while not individually optimal, collectively produces a Pareto improvement. This notion has been applied by Davila et al. (2005) in the Aiyagari-Bewley context.

II. A Two-Period Economy

This section presents a simple example economy with two periods and a linear technology. We use it to introduce some key concepts quickly and fix some basic ideas that are useful for the rest of the paper. The next sections generalize the model to an infinite horizon setting with general technologies.

In this example economy, there are two periods and a continuum of ex ante identical agents. They consume in both periods but work only in the second period. In the second period, they experience a labor productivity shock, the realization of which is independent across agents. The presence of uncertainty creates a role for insurance.

If productivity were publicly observable, then the first best would be attainable. The first best features perfect insurance in consumption, so that consumption is equalized across workers. It also requires workers with higher productivity to produce more.

Following Mirlees, we assume instead that productivity is private information and revealed to each worker at the beginning of the second period. With private information, the first best is no longer feasible because it is not incentive compatible. There is now a trade-off between insurance and incentives. The constrained efficient allocation, or second best, can be derived as the solution of a planning problem using the tools of mechanism design. The allocation now features distortions in labor and savings. It can be implemented using distortional nonlinear taxes on labor and capital. Roughly speaking, one can think of labor distortions as arising from the attempts to provide insurance while trading off incentives. It turns out that savings distortions reduce the cost of providing these incentives.

Why is this the case? The agent’s calculation, if he can save freely, is different for the usual precautionary saving motives. This discrepancy explains positive savings taxes. More technically, if agents were able to save freely, without savings distortions, then the standard Euler equation would hold. Instead, the constrained efficient allocation satisfies an inverse Euler equation, which is incompatible with the Euler equation. It is this conflict between the Euler and inverse Euler equations that requires the introduction of savings distortions.

How much do savings distortions contribute toward welfare? That is

\footnote{It is easy to allow for work in the first period, but this feature is not essential. The example economy is meant to be the simplest setting to introduce basic conceptual points.}
the question we focus on in this paper. That is, we wish to explore the quantitative importance of satisfying the inverse Euler equation, instead of the Euler equation.

A. Preferences

There are two periods $t = 0, 1$, and agents are ex ante identical. Consumption takes place in both periods, while work occurs only in period $t = 1$. Agents obtain utility

$$v = U(c_0) + \beta \mathbb{E}[U(c_1) - V(n_1; \theta)],$$

where $U$ is the utility function from consumption, $V$ is the disutility function from effective units of labor (hereafter, labor for short), and $\mathbb{E}$ is the expectations operator. Uncertainty is captured by an individual shock $\theta \in [\underline{\theta}, \bar{\theta}] = \Theta$ that affects the disutility of effective units of labor, where $\theta$ is an interval of $\mathbb{R}$. We will sometimes refer to $\theta$ as a skill or productivity shock, which is justified in the specification used by Mirrlees (1971) in which the disutility from labor $V(n; \theta)$ is equal to $v(n/\theta)$ for some convex function $v$ over work effort $n$, so that $\theta$ can be interpreted as productivity. To capture the idea that uncertainty is idiosyncratic, we assume that a version of the law of large number holds so that for any function $f$ on $\theta$, $\mathbb{E}[f]$ corresponds to the average of $f$ across agents.

The utility function $U$ is assumed to be increasing, concave, and continuously differentiable. We assume that the disutility function $V$ is continuously differentiable and that, for any $\theta \in \Theta$, the function $V(\cdot, \theta)$ is increasing and convex. We also assume the single-crossing property: $(\partial / \partial n_1)V(n_1; \theta) < 0$ is strictly decreasing in $\theta$, so that a high-shock $\theta$ indicates a low disutility from work.

B. Technology

We assume that technology is linear with a rate of return on savings equal to $q^{-1}$ and normalize labor productivity to 1. It is convenient to change variables and define an allocation by the triplet $\{u_0, u_1, n_1\}$, with $u_0 \equiv U(c_0), u_1(\theta) \equiv U(c_1(\theta))$. The cost of an allocation is

$$k_0 = c(u_0) + q \mathbb{E}[c(u_1(\theta)) - n_1(\theta)],$$

where $c \equiv U^{-1}$ is the inverse of the utility function.

C. Incentive Compatibility

The shock realizations are private information to the agent, so we must ensure that allocations are incentive compatible. By the revelation principle, we can consider, without loss of generality, a direct mechanism in which agents make a report $\hat{\theta}$ about their shock realization $\theta$ in period
\[ t = 1, \text{ and we assigned labor and utility from consumption as a function of this report. We require that they find truth-telling optimal:} \]

\[ u_0 + \beta(u_1(\theta) - V(n_1(\theta); \theta)) \geq u_0 + \beta(u_1(\hat{\theta}) - V(n_1(\hat{\theta}); \theta)) \]

\[ \text{for all } \theta, \hat{\theta} \in \Theta^2. \]  

(2)

As is well known (see, e.g., Milgrom and Segal 2002), the single-crossing property implies that incentive compatibility is equivalent to the condition that \( n_1 \) be nondecreasing, and

\[ u_0 + \beta(u_1(\theta) - V(n_1(\theta); \theta)) = u_0 + \beta(u_1(\hat{\theta}) - V(n_1(\hat{\theta}); \theta)) \]

\[ + \beta \int_{\hat{\theta}}^{\theta} \frac{\partial V}{\partial \hat{\theta}}(n_1(\hat{\theta}); \hat{\theta}) \, d\hat{\theta}. \]  

(3)

which can be thought of as an envelope condition, written in integral form.

\section{Perturbations}

Fix a utility level \( v \) and labor assignment \( \{n_1\} \). Then equations (1) and (3) pin down a unique value for the utility from consumption \( u_0 + \beta u_1(\theta) \) as a function of \( \theta \).\(^5\) In other words, equation (3) implies that providing incentives for a given labor assignment requires some given spread in \( u_0 + \beta u_1(\theta) \) across \( \theta \); expected utility \( v \) pins down the level with equation (1). Thus, for given utility assignments \( \{u_0, u_1\} \) that satisfy equations (1) and (3), all other feasible utility assignments \( \{\hat{u}_0, \hat{u}_1\} \) can be obtained by the following perturbation:

\[ \hat{u}_0 = u_0 - \beta \Delta \text{ and } \hat{u}_1(\theta) = u_1(\theta) + \Delta \text{ for all } \theta \in \Theta, \]

for some \( \Delta \in \mathbb{R} \).

\section{Efficiency and the Inverse Euler Equation}

Using this perturbation, it follows that a utility assignment \( \{u_0, u_1\} \) satisfying equations (1) and (3) minimizes the cost \( k_0 \) if and only if \( \Delta = 0 \) minimizes \( c(u_0 - \beta \Delta) + q \mathbb{E}[c(u_1(\theta) + \Delta)] \) over \( \Delta \), holding \( \{u_0, u_1\} \) fixed. The necessary and sufficient first-order condition is

\(^5\) This is where using eq. (3), rather than eq. (2), is convenient. This result relies on our assumption that types are continuous, so that \( \Theta \) is an interval. If instead the set of feasible types were finite, then incentive constraints could be slack for any given assignment of labor (a finite vector), creating room for some choice over \( u_0 + \beta u_1(\theta) \). However, as the number of types is increased so that the set of types becomes dense, approaching the continuum, then in the limit, \( v \) and the labor assignment once again determine a unique value for \( u_0 + \beta u_1(\theta) \) as a function of \( \theta \).
\[
\frac{1}{U'(c(u_0))} = \frac{1}{\beta q^{-1}} \mathbb{E} \left[ \frac{1}{U'(c(u_1(\theta)))} \right],
\] (4)

which is known as the *inverse Euler equation*. In contrast, when agents can save freely, the standard Euler equation
\[
U'(c(u_0)) = \beta q^{-1} \mathbb{E} [U'(c(u_1(\theta)))]
\] (5)
must hold. The standard Euler equation holds because agents perceive that the following perturbations are possible: \( \tilde{c}_0 = c_0 - q \varepsilon \) and \( \tilde{c}_1(\theta) = c_1(\theta) + \varepsilon \) for all \( \theta \in \Theta \). Note that these perturbations are risk free in terms of consumption. In contrast, the perturbations perceived by the planner are risk free in terms of utility.

If the inverse Euler equation holds, Jensen’s inequality implies that
\[
U'(c(u_0)) < \beta q^{-1} \mathbb{E} [U'(c(u_1(\theta)))]
\] (6)
as long as consumption at \( t + 1 \) is uncertain, conditional on information at \( t \). This inequality shows that equation (4) is incompatible with the standard Euler equation (eq. [5]). Thus, the optimum cannot be implemented by allowing agents to save freely at the technological rate of return.

F. An Intuition for the Optimality of Positive Savings Distortions

Start from a baseline allocation \( \{u_0, u_1\} \) with free savings, so that the Euler equation holds. Consider now perturbing this allocation so that \( \tilde{u}_0 = u_0 - \beta \Delta \) and \( \tilde{u}_1(\theta) = u_1(\theta) + \Delta \). The perturbed allocation is incentive compatibility and delivers the same level of utility. For some \( \Delta < 0 \), it satisfies the inverse Euler equation. The corresponding perturbed allocation has a lower cost than the baseline allocation.

Intuitively, although one is used to thinking of a risk-free asset as the natural one for borrowing and saving, our analysis suggests that, in economies with private information, the proper asset is one that does not upset incentive compatibility. Preserving incentive compatibility is possible with a hypothetical asset, with holdings \( \Delta \), that is risk free in terms of utility, instead of consumption. However, such an asset amounts to a risky asset in terms of its consumption payoff \( \Delta/U'(c_1(\theta)) \)—it must pay out more whenever the marginal utility of consumption is low (i.e., when consumption is high). From a finance perspective, this is the

---

6 If agents can save freely, then an allocation must satisfy the Euler equation, in addition to the incentive compatibility constraint. Indeed, for a given labor assignment \( \{n_t\} \) and utility level \( v \), the utility assignments \( \{u_0, u_1\} \) are uniquely determined by incentive compatibility, the requirement that the allocation deliver expected utility \( v \), and the Euler equation. That is, there exists a unique utility assignment that satisfies the Euler equation supporting \( \{n_t\} \) and \( v \). The converse is not generally true, although in practice we have found that this is not an issue for the allocations that we compute.
Fig. 1.—Efficiency gains as a function of \( n_1 \) for given \( v \). Upper curve imposes free savings. Lower curve implements the inverse Euler equation with optimal savings distortions.

definition of a risky return, with no expected premium to compensate. As a result, the optimum involves a negative holding of such an asset: it is attractive to borrow because repayment is higher when tomorrow’s consumption is higher. By borrowing today against good states of nature tomorrow, the planner increases today’s consumption relative to tomorrow’s, explaining why \( \Delta < 0 \) is optimal.7

Why does this imply a positive tax on capital? Consumers face a conventional risk-free saving technology, instead of the ideal but hypothetical risky asset described above (which is risk free in terms of utility but risky in terms of consumption). Thus, taxes are needed to induce agents to make the correct consumption and saving choice. In particular, we need a positive tax on capital to induce the agent to raise today’s consumption relative to tomorrow’s, to match the planner optimum.

G. Efficiency Gains from Savings Distortions

We have established that allowing agents to save freely is costly. The goal of this paper is to explore the magnitude of this cost. The lower curve in figure 1 represents the minimum cost \( k_0 \), for a given \( v \), as a function of the labor assignment \( \{n_1\} \). Along this curve, the inverse Euler equation holds. Similarly, the upper curve represents the cost of the allocation when agents can save freely. Along this curve, the Euler equation holds. The distance between these curves represents the cost of forgoing the

7 Borrowing here is relative to the solution with free borrowing and saving. The level of savings is not really pinned down without specifying some initial endowment or transfers from the government. At an optimum, it is also irrelevant whether the government can change the timing of transfers, due to the usual Ricardian equivalence arguments.
optimal taxation of savings and allowing, instead, agents to save freely. Indeed, the vertical distance represents precisely the added cost, for a given $v$ and $\{n_t\}$, of imposing the Euler equation instead of the inverse Euler equation.

We now consider three distinct allocations: (a) self-insurance, where agents can save freely with the technological return $q^{-1}$ and the output from their labor is untaxed; (b) the optimum for a planner that is subject to the incentive compatibility constraints and the constraint that agents can save freely; and (c) the constrained optimum, which solves a planning problem subject to the incentive compatibility constraints. Allocation $c$ can be implemented by a combination of nonlinear taxes on labor and capital, and allocation $b$ requires a nonlinear labor tax, while allocation $a$ features no taxes.

These allocations are displayed in figure 1. Point $b$ lies at the minimum of the upper curve, while $c$ lies at the minimum of the lower curve. Point $a$ is depicted on the far right of the upper curve, to convey the idea that we expect labor to be higher since it is not distorted by taxation.

Suppose the economy starts at self-insurance at point $a$ and moves toward the constrained efficient allocation, at point $c$. In the figure, this can be seen as a jump from the upper curve to lower curve, from point $a$ to $a'$, as well as a movement along the lower curve, from $a'$ to $c$. An alternative is to decompose the change from $a$ to $c$ as a movement along the upper curve to point $c'$ and then a vertical switch to the lower curve, from $c'$ to $c$.

The move from $a$ to $c$ combines the introduction of optimal labor and savings distortions. Therefore, it would be incorrect to attribute the entire gain between points $a$ and $c$ to the introduction of savings distortions. Instead, we propose interpreting the vertical distances between the two curves, either between $a$ and $a'$ or between $c'$ and $c$, as measuring the gains from savings distortions.

One might also consider the gains of moving from point $b$ to point $c$. This move represents the gains of introducing savings distortions, when labor distortions are fully optimized in both cases. Note that the vertical distance between $c$ to $c'$ provides an upper bound for the gains from $b$ to $c$, providing another rationale for focusing on the vertical distance between the two curves.

Furthermore, in practice, the distance between $a$ and $a'$ is greater than the distance between $c'$ and $c$ because it features greater uncertainty in consumption, which is responsible for the difference between the Euler and inverse Euler equations in the first place. Consumption inequality is greater at self-insurance because labor income is untaxed, resulting in greater after-tax income uncertainty.
TABLE 1

<table>
<thead>
<tr>
<th>Policy Experiment</th>
<th>Efficiency Gain (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a \rightarrow c$</td>
<td>3.08</td>
</tr>
<tr>
<td>$b \rightarrow c$</td>
<td>.33</td>
</tr>
<tr>
<td>$c' \rightarrow c$</td>
<td>.39</td>
</tr>
<tr>
<td>$a \rightarrow a'$</td>
<td>1.38</td>
</tr>
</tbody>
</table>

Note.—Gains are expressed as a percentage of aggregate consumption in autarky.

H. A Numerical Example

To illustrate this analysis, we have computed the three allocations for a parametrized version of this economy. In particular we adopt logarithmic utility $U(c) = \log c$ and $V(n; \theta) = \kappa(n/\theta)^{1/(1/\epsilon)}$ and set the Frisch elasticity to $\epsilon = 1/2$. With only two periods, it is natural to interpret a period as half a working lifetime (20 years). Correspondingly, we set $\beta = (0.96)^{20}$ and take the productivity distribution to be log normal with variance equal to $20 \times 0.0161$. This example is meant to provide a rough illustration of our concepts, not a full-blown calibration.

Table 1 displays our findings for this economy. The first row shows that the gains of moving from autarky to the constrained efficient allocation are relatively large, equal to 3.08 percent. Note, however, that at the constrained efficient allocation there are both savings and labor distortions. Thus, it would be incorrect to attribute these gains to savings distortions that implement the inverse Euler equation. In contrast, the other rows in the table provide measures of efficiency gains that may be attributed to savings distortions. In particular, the second row shows the gains from the optimum with free savings to the constrained optimum where agents face optimal savings distortions. The gains in this case are significantly smaller, 0.33 percent. Also shown are the vertical distances, $c'$ to $c$ and $a$ to $a'$, with gains of 0.39 and 1.38 percent, respectively. These numbers represent the gains, for a given labor allo-

---

8 The number 0.0161 corresponds to the variance of the innovations in the permanent component of income estimated by Storesletten, Telmer, and Yaron (2004b).

9 For these calculations, we compute the autarkic allocation as a competitive equilibrium with no taxes ($a$). This yields a utility assignment $\{u_0, u_t\}$, a labor assignment $\{n_t\}$, and utility level $v$. For this utility level and labor assignment, we can compute $a'$ by perturbing $a$ so that the inverse Euler equation holds. Allocation $b$ is solved by minimizing cost $k_0$, subject to the promise-keeping constraint (1), the Euler eq. (5), and the incentive compatibility constraints (3); we then verify that this allocation is attainable when agents can save freely. Allocation $c$ solves a relaxed version of that problem in which the Euler equation is dropped. Allocation $c'$ is obtained from $c$, by finding the perturbation that satisfied the Euler equation; we then verify that this allocation is feasible when agents can save freely. The efficiency gains are displayed as the reduction in cost $k_0$, scaled by the cost of consumption $\epsilon(u_0) + \varphi E[\epsilon(u_t(\theta))]$ of the autarkic allocation.
cation, of moving from the Euler to the inverse Euler equation. As such, they capture one notion of the gains from savings distortions. Also note that, as claimed earlier, $c'$ to $c$ is an upper bound on the distance between $b$ and $c$. In this example it is also the case that the distance between $a$ and $a'$ is an upper bound on the other two measures of efficiency gains from savings distortions, $b$ to $c$ and $c'$ to $c$.

I. Discussion

In this simple example economy, the allocations $a$, $a'$, $b$, $c$, and $c'$ are easily computed. In the more general infinite horizon setup with general technologies and arbitrary stochastic processes for productivity that we consider in the rest of the paper, this is no longer the case. In particular, computing numerically points $b$, $c$, and $c'$ is out of reach in most cases.

In the rest of the paper, we develop a general method to compute theoretically and quantitatively, for a given level of utility and a given labor assignment, the efficiency gains that can be realized by moving from a utility assignment that satisfies the Euler equation to one that satisfies the inverse Euler equation. That is, we develop a method to analyze theoretically and numerically the vertical distance between the two curves represented in figure 1 in very general economies with an infinite horizon, arbitrary stochastic processes for idiosyncratic productivity shocks, and general technologies. In the numerical application of our method, we consider an economy without taxes; thus, we focus on the analog of the move from $a$ to $a'$.

There are a number of advantages that come from approaching the efficiency gains from savings distortions through the vertical distance between the two curves in figure 1. First, we have explained how this vertical distance homes in on the benefit of savings distortions while also being informative of other, more encompassing, measures of efficiency gains that allow for changes in the labor assignment $\{n_1\}$.

Second, our measure of efficiency gains can be computed by a simple perturbation method. Indeed, consider the feasible allocation $\{u_0, u_1, n_1\}$ that satisfies the Euler equation. We can obtain the utility assignment $\{\tilde{u}_0, \tilde{u}_1\}$ that supports the same utility and labor assignment $\{n_1\}$ through a simple perturbation $\{u_0 - \beta \Delta, u_1 + \Delta\}$ of the baseline utility assignment $\{u_0, u_1\}$ while fixing the labor assignment $\{n_1\}$. Indeed, $\{\tilde{u}_0, \tilde{u}_1\}$ is the unique utility assignment within this class of perturbations that satisfies the inverse Euler equation. The existence of these simple perturbations greatly facilitates the computation of the efficiency gains from moving from $\{u_0, u_1, n_1\}$ and $\{\tilde{u}_0, \tilde{u}_1, n_1\}$.

Third, given the baseline utility assignment $\{u_0, u_1\}$, no knowledge of either the labor assignment $\{n_1\}$ or the disutility of work $V$ is needed in order to compute our measure of efficiency gains. In other words, one
does not need to take a stand on how elastic work effort is to changes in incentives. More generally, one does not need to take a stand on whether the problem is one of private information regarding skills or of moral hazard regarding effort, and so on. This robustness is a crucial advantage since current empirical knowledge of these characteristics and parameters is limited and controversial. The vertical distance between the two curves in figure 1 addresses precisely the question of whether the intertemporal allocation of consumption is efficient, without taking a stand on how correctly the trade-off between insurance and incentives has been resolved. Finally, as we will show, the magnitude of the efficiency gains is determined by some rather intuitive parameters: the relative variance of consumption changes, the coefficient of relative risk aversion, and the concavity of the production function.

III. Infinite Horizon

In this section, we lay down our general environment. We then describe a class of perturbations that preserve incentive compatibility. These perturbations serve as the basis for our method to compute efficiency gains.

A. The Environment

We cast our model within a general Mirrleesian dynamic economy. Our formulation is closest to Golosov et al. (2003). This paper obtains the inverse Euler equation in a general dynamic economy, where agents’ privately observed skills evolve as a stochastic process.

Preferences

Our economy is populated by a continuum of agent types indexed by $i \in I$ distributed according to the measure $\psi$. Preferences generalize those used in Section II and are summarized by the expected discounted utility

$$\sum_{t=0}^{\infty} \beta^t \mathbb{E}^i[U(c_t) - V(n_t^i; \theta)]$$

where $\mathbb{E}^i$ is the expectations operator for type $i$. Additive separability between consumption and leisure is a feature of preferences that we

---

10 This robustness property can be more formally described as follows. Consider the set $\Omega(\{u_0, u_t, n_t\}, v)$ of disutility functions $V$ with the following properties: $V$ is continuously differentiable; for any $\theta \in \Theta$, the function $V(\cdot, \theta)$ is increasing and convex; $V$ has the single-crossing property that $(\partial/\partial n_t) V(n_t; \theta)$ is strictly decreasing in $\theta$ and eq. (3) holds. The set of perturbed allocations $\{[\{u_0 - \beta \Delta, u_t + \Delta\} | \Delta \in \mathbb{R}] \}$ is the largest of allocations such that for all $V \in \Omega(\{u_0, u_t, n_t\}, v)$, incentive compatibility (3) and promise-keeping (1) hold.
adopt because it is required for the arguments leading to the inverse Euler equation.\textsuperscript{11}

Idiosyncratic uncertainty is captured by an individual-specific shock $\theta^i \in \Theta$, where as in Section II, $\Theta$ is an interval of the real line. These shocks affect the disutility of effective units of labor. We sometimes refer to them as skill or productivity shocks. The stochastic process for each individual $\theta^i$ is identically distributed within each type $i \in I$ and independently distributed across all agents. We denote the history up to period $t$ by $\theta^{i,t} \equiv (\theta^i_0, \theta^i_1, \ldots, \theta^i_t)$ and the probability measure on $\Theta^\infty$ corresponding to the law of the stochastic process $\theta^i_t$ for an agent of type $i$ by $\pi^i$.

Given any function $f$ on $\Theta^\infty$, we denote the integral $\int f(\theta^{i,\infty}) \, d\pi^i(\theta^{i,\infty})$, using the expectation notation $\mathbb{E}[f(\theta^{i,\infty})]$ or simply $\mathbb{E}[f]$. Similarly, we write $\mathbb{E}[f(\theta^{i,\infty})|\theta^{i,t-1}]$, or simply $\mathbb{E}_{t-1}[f]$, for the conditional expectation of $f$ given history $\theta^{i,t-1} \in \Theta^t$.

As in Section II, all uncertainty is idiosyncratic, and we assume that a version of the law of large number holds so that for any function $f$ on $\Theta^\infty$, $\mathbb{E}[f]$ corresponds to the average of $f$ across agents with type $i$.

To preview the use we will have for types, note that in our numerical implementation we will assume that skills follow a Markov process. We will consider allocations that result from a market equilibrium where agents save in a riskless asset. For this kind of economy, agent types are then initial asset holdings together with initial skill. The measure $\psi$ captures the joint distribution of these two variables.

It is convenient to change variables, translating consumption allocations into utility assignments $\{u_i(\theta^{i,\infty})\}$, where $u_i(\theta^{i,\infty}) \equiv U(c_i(\theta^{i,\infty}))$. This change of variable will make incentive constraints linear and render the planning problem, which we will introduce shortly, convex.

Information and Incentives

The shock realizations are private information to the agent. We invoke the revelation principle to derive the incentive constraints by considering a direct mechanism. Agents are allocated consumption and labor as a function of the entire history of reports. The agent’s strategy determines a report $\sigma^i(\theta^{i,\infty})$ for each period $t$ as a function of the history of shocks $\theta^{i,t}$. Define the history up to time $t$ of such reports to be $\sigma^{i,t}(\theta^{i,\infty}) = (\sigma^i_0(\theta^i_0), \sigma^i_1(\theta^i_1), \ldots, \sigma^i_t(\theta^i_t))$. The incentive compatibility constraint requires that truth-telling, $\sigma^{i,*}(\theta^{i,\infty}) = \theta^i$, be optimal, so that for all for all reporting strategies $[\sigma^i]$ and all $i \in I$.

\textsuperscript{11} The intertemporal additive separability of consumption also plays a role. However, the intertemporal additive separability of work effort is completely immaterial: we could replace $\sum_{t=0}^{\infty} \beta \mathbb{E}[V(n_t; \theta^i)]$ with some general disutility function $V(n_t)$. 
\[ v^i \equiv \sum_{i=0}^{\infty} \beta^i \mathbb{E}[u'(\theta^i)] - V(n_i(\theta^i); \theta^i)] \geq \sum_{i=0}^{\infty} \beta^i \mathbb{E}[u'(\sigma^i(\theta^i))] - V(n_i(\sigma^i(\theta^i)); \theta^i)]. \tag{7} \]

Technology

Let \( C_t \) and \( N_t \) represent labor and consumption for period \( t \), respectively. That is, letting \( \epsilon \equiv U^{-1} \) denote the inverse of the utility function,

\[ C_t = \int E'[\epsilon(u_t)] \, d\psi, \]

\[ N_t = \int E'[n_t] \, d\psi, \]

for \( t = 0, 1, \ldots \). In order to facilitate our efficiency gains calculations, it will prove convenient to index the resource constraints by \( \epsilon_t \), which represents the aggregate amount of resources that is being economized in every period. The resource constraints are then

\[ K_{t+1} + C_t + \epsilon_t \leq (1 - \delta)K_t + F(K_t, N_t), \quad t = 0, 1, \ldots, \tag{8} \]

where \( K_t \) denotes aggregate capital. The function \( F(K, N) \) is assumed to be homogenous of degree 1, concave, and continuously differentiable, increasing in \( K \) and \( N \).

Two cases are of particular interest. The first is the neoclassical growth model, where \( F(K, N) \) is strictly concave and satisfies Inada conditions \( F_K(0, N) = \infty \) and \( F_K(\infty, N) = 0 \). In this case, we also impose \( K_t \geq 0 \). The second case has linear technology \( F(K, N) = N + (q^{-1} - 1)K, \delta = 0, \) and \( 0 \leq q < 1 \). One interpretation is that output is linear in labor with productivity normalized to one, and a linear storage technology with a safe gross rate of return \( q^{-1} \) is available. Another interpretation is that this represents the economy-wide budget set for a partial equilibrium analysis, with constant interest rate \( 1 + r = q^{-1} \) and unit wage. Under either interpretation, we avoid corner solutions by allowing negative capital holdings, subject to \( K_{t+1} \geq -\sum_{i=1}^{\infty} q^i N_{t+r} \). This constraint allows borrowing up to the natural borrowing limit, equal to the present value of future labor income. With this borrowing limit, one can summarize the constraints on the economy by the single present-value condition

\[ \sum_{i=0}^{\infty} q^i C_t \leq \sum_{i=0}^{\infty} q^i (N_t - \epsilon_t) + \frac{1}{q} K_0. \]
Feasibility

An allocation \(\{u^i_n, n^i_n, K^i_n, e\}\) and utility profile \(\nu\) is feasible if conditions (7) and (8) hold. That is, feasible allocations must deliver utility \(\nu^i\) to agents of type \(i \in I\) and must be incentive compatible and resource feasible.

Free Savings

For the purposes of this paper, an important benchmark is the case in which agents can save, and perhaps also borrow, freely. Free borrowing and saving increases the choices available to agents, which adds further restrictions relative to the incentive compatibility constraints.

In this scenario, the government enforces labor and taxes as a function of the history of reports but does not control consumption directly. Disposable after-tax income is \(w_i n_i^i(\sigma^{i-i}(\theta^i)) - T^i_i(\sigma^{i-i}(\theta^i)).\) Agents face the following sequence of budget and borrowing constraints:

\[
c^i_i(\theta_i) + a_i^{i+1}(\theta_i) \leq w_i n_i^i(\sigma^{i-i}(\theta_i)) - T^i_i(\sigma^{i-i}(\theta_i)) + (1 + r) a_i^{i}(\theta_i), \quad (9a)
\]

\[
a_i^{i+1}(\theta_i) \geq a_i^{i}(\theta_i), \quad (9b)
\]

with \(a^i_0\) given. We allow the borrowing limits \(a_i^{i+1}(\theta_i)\) to be tighter than the natural borrowing limits.

Agents with type \(i\) maximize utility \(\sum_{t=0}^{\infty} \beta^t E[u_i^i(\sigma^{i-i}(\theta_i)) - V(n_i^i(\sigma^{i-i}(\theta_i)); \theta_i)]\) by choosing a reporting, consumption, and saving strategy \(\{\sigma_i^i, c_i, a_i^{i+1}\}\) subject to the sequence of constraints (9), taking \(a^i_0\) and \(\{n_i^i, T_i^i, w_i, r_i\}\) as given. A feasible allocation \(\{u_i^i, n_i^i, K_i, e\}\) is part of a free-savings equilibrium if there exist taxes \(\{T_i^i\}\), such that the optimum \(\{\sigma_i^i, c_i, a_i^{i+1}\}\) for an agent of type \(i\) with wages and interest rates given by \(w_i = F_n(K, N)\) and \(r_i = F_r(K, N) - \delta\) satisfies truth-telling \(\sigma_i^i(\theta_i) = \theta_i\) and generates the utility assignment \(u_i^i(\theta_i) = U(c_i(\theta_i)).\)

At a free-savings equilibrium, the incentive compatibility constraints (7) are satisfied. The consumption-savings choices of agents impose further restrictions. In particular, a necessary condition is the intertemporal Euler condition

\[
U''(c(u_i^i)) \geq \beta(1 + r_i^{i+1})E_i[u_i^{i+1}(c(u_i^{i+1}))], \quad (10)
\]

with equality if \(a_i^{i+1}(\theta_i) > a_i^{i+1}(\theta_i).\) Note that if the borrowing limits \(a_i^{i+1}(\theta_i)\) are equal to the natural borrowing limits, then the Euler equation (10) always holds with equality.

---

12 A special case of interest is where the dependence of the tax on any history of reports \(\tilde{\theta}^\sigma\) can be expressed through its effect on the history of labor \(n^\sigma(\tilde{\theta}^\sigma)\), i.e., when \(T_i(\tilde{\theta}^\sigma) = T_i^*(n^\sigma(\tilde{\theta}^\sigma))\) for some \(T_i^*\) function.

13 Note that individual asset holdings and taxes are not part of this definition because they are indeterminate due to Ricardian equivalence.
Efficiency

We say that the allocation \( \{u_i', n_i', K_i, e_i\} \) and utility profile \( \{v_i\} \) are dominated by the alternative \( \{\tilde{u}_i, \tilde{n}_i, \tilde{K}_i, \tilde{e}_i\} \) and \( \{\tilde{v}\} \), if \( \tilde{v} \geq v' \), \( \tilde{K}_0 \leq K_0 \), and \( e_i \leq \tilde{e}_i \) for all periods \( t \) and either \( \tilde{v} > v' \) for a set of agent types of positive measure \( \tilde{K}_0 < K_0 \) or \( e_i < \tilde{e}_i \) for some period \( t \). We say that a feasible allocation is efficient if it is not dominated by any feasible allocation. We say that an allocation is conditionally efficient if it is not dominated by a feasible allocation with the same labor allocation \( n_i = \tilde{n}_i \).

As explained in Section II, allocations that are part of a free-savings equilibrium are not conditionally efficient. Conditionally efficient allocations satisfy a first-order condition, the inverse Euler equation, which is inconsistent with the Euler equation. Being part of a free-savings equilibrium therefore acts as a constraint on the optimal provision of incentives and insurance. Efficiency gains can be reaped by departing from free savings.

B. Incentive Compatible Perturbations

In this section, we develop a class of perturbations of the allocation of consumption that preserve incentive compatibility. We then introduce a concept of efficiency, \( \Delta \)-efficiency, that corresponds to the optimal use of these perturbations. Our perturbation set is large enough to ensure that every \( \Delta \)-efficient allocation satisfies the inverse Euler equation. Moreover, we show that \( \Delta \)-efficiency and conditional efficiency are closely related concepts: \( \Delta \)-efficiency coincides with conditional efficiency on allocations that satisfy some mild regularity conditions.

A Class of Perturbations

For any period \( t \) and history \( \theta^{t,i} \), a feasible perturbation, of any baseline allocation, is to decrease utility at this node by \( \beta \Delta^i \) and compensate by increasing utility by \( \Delta^i \) in the next period for all realizations of \( \theta_{t+1}^{i} \). Total lifetime utility is unchanged. Moreover, since only parallel shifts in utility are involved, incentive compatibility of the new allocation is preserved. We can represent the new allocation as \( \tilde{u}(\theta^{t,i}) = u(\theta^{t,i}) - \beta \Delta^i, \tilde{u}_{t+1}(\theta^{t+1,i}) = u_{t+1}(\theta^{t+1,i}) + \Delta^i \), for all \( \theta_{t+1}^{i} \).

This perturbation changes the allocation in periods \( t \) and \( t+1 \) after history \( \theta^{t,i} \) only. The full set of variations generalizes this idea by allowing perturbations of this kind at all nodes:

\[
\tilde{u}(\theta) = u(\theta) + \Delta(\theta^{t,i}) - \beta \Delta(\theta^{t,i}),
\]

for all sequences of \( \{\Delta(\theta^{t,i})\} \) such that \( \tilde{u}(\theta^{t,i}) \in U(\mathbb{R}_+) \) and such that the limiting condition
for all reporting strategies \( \{ \sigma^i_t \} \). This condition rules out Ponzilike schemes in utility.\(^{14}\) By construction, the agent’s expected utility, for any strategy \( \{ \sigma^i_t \} \), is only changed by a constant \( \Delta_{-1}^i \):

\[
\sum_{i=0}^{\infty} \beta^i \mathbb{E}[\tilde{u}^i(\sigma^i(\theta^{i,i}))] = \sum_{i=0}^{\infty} \beta^i \mathbb{E}[u^i(\sigma^i(\theta^{i,i}))] + \Delta_{-1}^i. \tag{11}
\]

It follows directly from equation (7) that the baseline allocation \( \{ u^i_t \} \) is incentive compatible if and only if the new allocation \( \{ \tilde{u}^i_t \} \) is incentive compatible. Note that the value of the initial shifter \( \Delta_{-1}^i \) determines the lifetime utility of the new allocation relative to its baseline. Indeed, for any fixed infinite history \( \tilde{\theta}^{i,\infty} \), equation (11) implies that (by substituting the deterministic strategy \( \sigma^i_t(\theta^{i,i}) = \tilde{\theta}^i_t \))

\[
\sum_{i=0}^{\infty} \beta^i \tilde{u}^i(\tilde{\theta}^{i,i}) = \sum_{i=0}^{\infty} \beta^i u^i(\tilde{\theta}^{i,i}) + \Delta_{-1}^i \quad \forall \tilde{\theta}^{i,\infty} \in \Theta^i. \tag{12}
\]

Thus, ex post realized utility is the same along all possible realizations for the shocks.\(^{15}\)

Let \( T(\{ u^i_t \}, \Delta_{-1}^i) \) denote the set of utility allocations \( \{ \tilde{u}^i_t \} \) that can be generated by these perturbations starting from a baseline allocation \( \{ u^i_t \} \) for a given initial \( \Delta_{-1}^i \).\(^{16}\) This set is convex.

Below, we show that these perturbation are rich enough to deliver the inverse Euler equation. In this sense, they fully capture the characterization of optimality stressed by Golosov et al. (2003).

An allocation \( \{ u^i_n, n^i, \tilde{K}_n, \tilde{e}_i \} \) with utility profile \( v^i \) is \( \Delta \) efficient if it is feasible and not dominated by another feasible allocation \( \{ \tilde{u}^i_n, n^i, \tilde{K}_n, \tilde{e}_i \} \) such that \( \{ \tilde{u}^i \} \in T(\{ u^i_t \}, \Delta_{-1}^i) \).

Note that conditional efficiency implies \( \Delta \) efficiency since both concepts do not allow for changes in the labor allocation. Under mild regularity conditions, the converse is also true. More precisely, in Appendix A, we define the notion of regular utility and labor assignments \( \{ u^i_n, n^i \} \).\(^{17}\) We then show that \( \Delta \) efficiency coincides with conditional ef-

\(^{14}\) Note that the limiting condition is trivially satisfied for all variations with a finite horizon: sequences for \( \{ \Delta_i \} \) that are zero after some period \( T \), as was the case in the discussion of a perturbation at a single node and its successors.

\(^{15}\) The converse is nearly true: by taking appropriate expectations of eq. (12), one can deduce eq. (11), except for a technical caveat involving the possibility of inverting the order of the expectations operator and the infinite sum (which is always possible in a version with a finite horizon and \( \Theta \) finite). This caveat is the only difference between eqs. (11) and (12).

\(^{16}\) Our method involves recursive methods. For this reason, it is useful to allow for \( \Delta_{-1}^i \neq 0 \) in this definition, even though our planning problem in Sec. III.C imposes \( \Delta_{-1}^i = 0 \).

\(^{17}\) Regularity is a mild technical assumption that is necessary to derive an envelope condition, similar to that behind eq. (3), crucial for our proof.
ficiency on the class of allocations with regular utility and labor assignments. Indeed, given a regular utility and labor assignment \( \{ u_n, n \} \), the perturbations \( T(\{ u_n, \Delta_i \}) \) characterize all the utility assignments \( \{ \tilde{u}_n \} \) such that \( \{ \tilde{u}_n, n \} \) is regular and satisfies the incentive compatibility constraints (7).

Inverse Euler Equation

Building on Section II, we review briefly the inverse Euler equation, which is the optimality condition for any \( \Delta \)-efficient allocation.

**Proposition 1.** A set of necessary and sufficient conditions for an allocation to be \( \Delta \)-efficient is given by

\[
\frac{c'(u_i^t)}{\beta} \equiv \left[ \frac{1}{U'(c(u_i^t))} \right] \left[ \frac{1}{U'(c(u_i^{t+1}))} \right],
\]

where \( q_t = 1/(1 + r_t) \), and

\[
r_t = \frac{F_{\bar{K}}(K_{t+1}, N_t)}{K_t} - \delta
\]

is the technological rate of return.

A \( \Delta \)-efficient allocation that is not deterministic cannot allow agents to save freely at the technology’s rate of return since then equation (10) would hold as a necessary condition, which is incompatible with the planner’s optimality condition, equation (13).

C. Efficiency Gains from Optimal Savings Distortions

In this section we consider a baseline allocation and an improvement that yields a \( \Delta \)-efficient allocation. We define a metric for the efficiency gains from this improvement that is our measure of the gains from the introduction of optimal savings distortions.

If an allocation \( \{ u_n, n, K, e \} \) with corresponding utility profile \( \{ v \} \) is not \( \Delta \)-efficient, then we can always find an alternative allocation \( \{ \tilde{u}_n, n, \tilde{K}, \tilde{e} \} \) that leaves utility unchanged, so that \( \tilde{v} = v \) for all \( i \in I \), but economizes on resources: \( \tilde{K}_i \leq K_i \) and \( \tilde{e}_i \leq e_i \) with at least one strict inequality. In the rest of the paper, we restrict to cases in which \( e_i = 0 \), \( \tilde{K}_i = K_i \), and \( \tilde{e}_i = \lambda C_i \) for some \( \lambda > 0 \). We then take \( \lambda \) as our measure of efficiency gains between these allocations. This measure represents the resources that can be saved in all periods in proportion to aggregate consumption.

We now introduce a planning problem that uses this metric to compute the distance of any baseline allocation from the \( \Delta \)-efficient frontier. For any given baseline allocation \( \{ u_n, n, K, 0 \} \), which is feasible with \( e_i = 0 \) for all \( t \geq 0 \), we seek to maximize \( \lambda \) by finding an alternative allocation \( \{ \tilde{u}_n, n, \tilde{K}, \lambda C \} \) with \( \tilde{K}_0 = K_0 \).
\[
\tilde{K}_{t+1} + \int E[c(\tilde{\tilde{u}}_t)] \, d\psi + \tilde{\lambda}C_t \leq (1 - \delta) \tilde{K}_t + F(\tilde{K}_t, N_t),
\]  
\[t = 0, 1, \ldots\]

and

\[\{\tilde{\tilde{u}}_t\} \in \Upsilon(\{u_t\}, 0).\]

Let \(\tilde{C}_t \equiv \int E[c(\tilde{\tilde{u}}_t)] \, d\psi\) denote aggregate consumption under the optimized allocation. The optimal allocation \(\{\tilde{\tilde{u}}_t, n_t, \tilde{K}_t, \lambda C_t\}\) in this program is \(\Delta\)-efficient and saves an amount \(\tilde{\lambda}C_t\) of aggregate resources in every period. Our measure of the distance of the baseline allocation from the \(\Delta\)-efficient frontier is \(\tilde{\lambda}\). For future use, we denote the corresponding sequence of interest rates and intertemporal prices by \(1 + \tilde{r}_t = 1 + F_k(\tilde{K}_t, N_t) - \delta\) and \(\tilde{q}_t = 1/(1 + \tilde{r}_t)\).

In Appendix B, we explain why this planning problem is numerically tractable and detail a method to solve it. The basic idea is to introduce a relaxed planning problem that replaces the resource constraints with a single present value condition for a given sequence of intertemporal prices \(\{\tilde{q}_t\}\). The relaxed planning problem can then be further decomposed in a series of component-planning problems corresponding to the different types \(i \in I\).

In most situations, the baseline allocation admits a recursive representation for some endogenous state variable.\(^{18}\) This is the case whenever \(\theta_t\) is a Markov process and the baseline allocation depends on the history of shocks \(\theta_t^{\leq t-1}\) in a way that can be summarized by an endogenous state \(x_t\) with law of motion \(x_t = M(x_{t-1}, \theta_t)\) and given initial condition \(x_0\) (types then correspond to different initial values \(x_0\)). The endogenous state \(x_t\) is a function of the history of exogenous shocks \(\theta_t^{\leq t}\). Defining the state vector \(s_t = (x_t, \theta_t)\), there must exist a function \(\tilde{u}\) such that \(u(\theta_t^{\leq t}) = \tilde{u}(s_t)\) for all \(\theta_t^{\leq t}\). In this case, the component-planning problems can be boiled down to a simple Bellman equation with two state variables, \(s_t\) and \(\Delta_{t-1}\). This Bellman equation is mathematically isomorphic to solving an income-fluctuation problem, where \(\Delta_{t-1}\) plays the role of wealth and is amenable to numerical simulations.

\(^{18}\) The requirement that the baseline allocation be recursive in this way is hardly restrictive. Of course, the endogenous state and its law of motion depend on the particular economic model generating the baseline allocation. A leading example in this paper is the case of incomplete-markets Bewley economies in Huggett (1993) and Aiyagari (1994). In these models, described in more detail in Sec. V, each individual is subject to an exogenous Markov process for income or productivity and saves using a riskless asset. At a steady state, the interest rate on this asset is constant, so that the agent’s solution can be summarized by a stationary savings rule. The baseline allocation can then be summarized using asset wealth as an endogenous state, with law of motion \(M\) given by the agent’s optimal-saving rule. Another example are allocations generated by a dynamic contract. The state variable then includes the promised continuation utility (see Spear and Srivastava 1987) along with the exogenous state.
IV. Idiosyncratic and Aggregate Gains with Log Utility

In this section, we focus on the case of logarithmic utility. When utility is logarithmic, our parallel shifts in utility imply proportional shifts in consumption. We first show that the planning problem can be decomposed into an idiosyncratic planning problem and a simple aggregate-planning problem. Logarithmic utility also makes it possible to solve the idiosyncratic efficiency gains and the corresponding allocation in closed form when the baseline allocation is recursive and features constant consumption. In this case, the optimum in the planning problem and the efficiency gains $\hat{\lambda}$ can be solved out almost explicitly by combining the solution of the idiosyncratic problem with that of the aggregate-planning problem. We then illustrate our results in the simple benchmark case when the baseline allocation of consumption is a geometric random walk.

A. A Decomposition: Idiosyncratic and Aggregate

Idiosyncratic Efficiency Gains

The full planning problem maximizes over utility assignments and capital. It is useful to also consider a version of the problem that takes the baseline sequence of capital $\{K_t\}$ as given. Thus, define the idiosyncratic planning problem as maximizing subject to

\[
\mathbb{E}[c(\hat{u}_t)] \, d\psi + \hat{\lambda}C_t \leq C_t, \quad t = 0, 1, \ldots,
\]

and $\{\hat{u}_t\} \in T(\{u_t\}, 0)$. The idiosyncratic efficiency gains $\hat{\lambda}$ represent the constant proportional reduction in consumption that is possible without changing the aggregate sequence of capital. Of course, the total efficiency gains are larger than the idiosyncratic ones: $\hat{\lambda} \geq \hat{\lambda}'$.

The solution of the idiosyncratic planning problem improves over the baseline allocation by ensuring that the marginal rates of substitution corresponding to the inverse Euler equation $\mathbb{E}[c'(\hat{u}_{t+1})/c'(\hat{u}_t)]$ are equalized across types and histories in every period. These marginal rates of substitution, however, are not necessarily linked to any technological rate of transformation as in equation (14). The idiosyncratic efficiency gains thus correspond to the gains from equalizing the marginal rate of substitution across types and histories in every period, without changing the sequence of capital. The aggregate efficiency gains then capture the additional benefits from altering the aggregate allocation to equalize these marginal rates of substitution with the marginal rate of transformation in every period.
Aggregate Efficiency Gains

Given $\tilde{\lambda}' \in [0, 1)$, the aggregate-planning problem seeks to determine the aggregate allocation $\{\tilde{C}_t, N_t, \tilde{K}_t, \tilde{\lambda}C_t\}$ that maximizes $\tilde{\lambda}^A$, subject to $\tilde{K}_0 = K_0$,

$$\tilde{K}_{t+1} + \tilde{C}_t + (\tilde{\lambda}^A + \tilde{\lambda}')C_t \leq (1 - \delta)\tilde{K}_t + F(\tilde{K}_t, N_t), \quad t = 0, 1, \ldots,$$

and

$$\sum_{t=0}^{\infty} \beta^t U(\tilde{C}_t) = \sum_{t=0}^{\infty} \beta^t U(C_t(1 - \tilde{\lambda}')).$$

We refer to $\tilde{\lambda}^A$ as aggregate efficiency gains. Note that in this program we can always set $K_t = K$ and $\tilde{C}_t = (1 - \tilde{\lambda}')C_t$, which guarantees that $\tilde{\lambda}^A \geq 0$.

**Proposition 2.** Suppose the utility is logarithmic and consider a baseline allocation $\{u', n', K, 0\}$ such that $\sum_{t=0}^{\infty} \beta^t U(C_t)$ is well defined and finite. Then the total efficiency gains underlying the planning problem are given by the sum of the idiosyncratic and aggregate efficiency gains: $\lambda = \lambda' + \tilde{\lambda}^A$.

**Proof.** See Appendix C. QED

The proof of proposition 2 also establishes that the utility assignment $\{\hat{u}\}$ that solves the idiosyncratic planning problem and the utility assignment $\{\tilde{u}\}$ that solves the original planning problem are related by $\hat{u}_t = \tilde{u}_t + \delta_t$, where $\delta_t = U(\tilde{C}_t) - U(C_t(1 - \tilde{\lambda}'))$.

The analysis of the evolution of the aggregate allocation $\{\tilde{C}_t, \tilde{K}_t\}$ only requires the knowledge of $\tilde{\lambda}'$ but can otherwise be conducted separately from the analysis of the idiosyncratic problem. The aggregate-planning problem is simply that of a standard deterministic growth model, which, needless to say, is straightforward to solve. For example, suppose that the baseline allocation represents a steady state with constant aggregates, $C_t = C_{ss}$, $K_t = K_{ss}$, and $N_t = N_{ss}$. Then the optimized aggregate allocation $\{\tilde{C}_t, \tilde{K}_t\}$ converges to a steady state $\{\tilde{C}_{ss}, \tilde{K}_{ss}\}$ such that $1 - \delta + F(\tilde{K}_{ss}, N_{ss}) = 1/\beta$ and $\tilde{C}_{ss} = F(\tilde{K}_{ss}) - \delta\tilde{K}_{ss} - \tilde{\lambda}C_{ss}$. We will put that result to use in Section V.

The aggregate-planning problem is as a version of the problem that Lucas studied in his famous Supply Side Economics exercise (Lucas 1990). Both involve computing the welfare gains along the transition to a new steady state in the neoclassical growth model. There is, however, an important difference. Indeed, imagine that the Euler equation holds at the baseline allocation. Then whereas Lucas’s exercise involves computing a transition to a steady state with more capital, our exercise, by contrast, involves a transition to a steady state with less capital. The reason is that he considers removing a capital tax in a model where
capital should not be taxed, while we consider introducing optimal capital taxes in a model where capital should be taxed.

B. Example: Steady States with Geometric Random Walk

Although the main virtue of our approach is that we can flexibly apply it to various baseline allocations, in this section we begin with a simple and instructive case. We maintain the assumption of logarithmic utility throughout. We take the baseline allocation to be a geometric random walk: $c_{t+1} = \varepsilon_t c_t$ with $\varepsilon_t$ i.i.d. across agents and over time. In the language of Section III.C, the baseline allocation is recursive: $s_{t+1} = \varepsilon_t s_t$ with $\varepsilon_t$ i.i.d. and $c(s) = s$, so that $u(s) = U(s)$. Moreover, we assume that $\log(\varepsilon)$ is normally distributed with variance $\sigma_\varepsilon^2$ so that $\mathbb{E}[\varepsilon] = \mathbb{E}[\varepsilon^{-1}] = \exp(\sigma_\varepsilon^2)$. We also assume that the baseline allocation represents a steady state with constant aggregates $C_t = C_{ss}$, $K_t = K_{ss}$, and $N_t = N_{ss}$, which requires $\mathbb{E}[\varepsilon] = 1$. We define $r_{ss} = F_k(K_{ss}, N_{ss}) - \delta$ and $q_{ss} = 1/(1 + r_{ss})$. Moreover, we assume that the Euler equation holds at the baseline allocation, which requires $q_{ss} = \beta\mathbb{E}[\varepsilon^{-1}] = \beta \exp(\sigma_\varepsilon^2)$.

Although extremely stylized, a random walk is an important conceptual and empirical benchmark. First, most theories—starting with the simplest permanent income hypothesis—predict that consumption should be close to a random walk. Second, some authors have argued that the empirical evidence on income, which is a major determinant for consumption, and consumption itself show the importance of a highly persistent component (e.g., Storesletten, Telmer, and Yaron 2004a). For these reasons, a parsimonious statistical specification for consumption may favor a random walk. Indeed, one can construct an example economy in which a geometric random walk for consumption arises as a competitive equilibrium with incomplete markets.\footnote{Assume $V(n; \theta) = v(n/\theta)$ and that skills $\theta_i$ evolve as a geometric random walk. Individuals can only accumulate a riskless asset paying return $q^{-1}$, equal to the rate of return on the economy’s linear savings technology, which is assumed to satisfy $1 \geq \beta q^{-1}\mathbb{E}[\varepsilon^{-1}]$. There are no taxes. Finally, assume initial assets are zero. In equilibrium, agents hold zero assets and set $c_t = n_t = \theta_t \bar{n}$ for some constant $\bar{n}$.}

The advantage is that we obtain closed-form solutions for the optimized allocation, the intertemporal wedge, and the efficiency gains. The transparency of the exercise reveals important determinants for the magnitude of efficiency gains. A geometric random walk, however, is special for the following reason. If we apply the decomposition of Section IV.A, then the idiosyncratic efficiency gains are zero. The entirety of the efficiency gains is aggregate because, at the baseline allocation, the marginal rates of substitution $\mathbb{E}[c'(u'_{t+1})/(\beta c'(u'_t))]$ are already equalized across types and histories to $\beta^{-1}$. These marginal rates of substitution are not equalized, however, to the marginal rate of transformation...
Partial Equilibrium: Linear Technology

We first study the case in which the technology is linear with a rate of return \( q^{-1} > 1 \). Since the Euler equation must hold, we must have \( q = q_{ss} = \beta \exp (\sigma_s^2) \). Note that \( q < 1 \) imposes, for a given discount rate \( \beta \), an upper bound on the variance of the shocks \( \exp (\sigma_s^2) < \beta^{-1} \).

Since the idiosyncratic efficiency gains \( \lambda^{i} \) are zero, the solution of the planning problem can be derived by studying the aggregate-planning problem, which takes a remarkably simple form. The aggregate consumption sequence \( \{\hat{C}_t\} \) that solves the aggregate-planning problem is given by \( \hat{C}_t = C_{ss} \exp ((\beta/(1-\beta))\sigma_s^2) \exp (-ta_s^2) \). The efficiency gains \( \hat{\lambda} \) and the optimal utility assignment \( \{\hat{u}\} \) are then readily computed. We can also derive the intertemporal wedge \( \tau \) that measures the savings distortions at the optimal allocation \( U'(c(\hat{u}(s')))/U(c(s_i))) \).

**Proposition 3.** Suppose that utility is logarithmic and that the technology is linear. Suppose that the baseline allocation is a geometric random walk with constant aggregate consumption \( C_t = C_{ss} \), that the shocks \( e \) are lognormal with variance \( \sigma_e^2 \), and that the Euler equation holds at the baseline. Then the consumption assignment of the solution of the planning problem is given by \( \hat{c}(s^i) = \exp ((\beta/(1-\beta))\sigma_e^2) \exp (-ta_e^2) c(s_i) \). The intertemporal wedge at the optimal allocation is given by \( \tau = 1 - \exp (-\sigma_e^2) \). The efficiency gains are given by \( \hat{\lambda} = 1 - ([\beta^{-1} - \exp (\sigma_e^2)]/(\beta^{-1} - 1)) \exp ((\beta/(1-\beta))\sigma_e^2) \).

The optimized allocation has a lower drift than the baseline allocation. Intuitively, our perturbations based on parallel shifts in utility can be understood as allowing consumers to borrow and save with an artificial idiosyncratic asset, the payoff of which is correlated with their baseline idiosyncratic consumption process: they can increase their consumption today by reducing their consumption tomorrow in such a way that they reduce their consumption tomorrow more in states where consumption is high than in states where consumption is low. The desirable insurance properties of these perturbations make them attractive, leading to a front-loading of consumption. In other words, because our perturbations allow for better insurance, they reduce the benefits of engaging in precautionary savings by accumulating a buffer stock of risk-free assets. As a result, it is optimal to front-load consumption, by superimposing a downward drift \( \exp (-\sigma_e^2) \) on the baseline allocation, where the variance in the growth rate of consumption \( \sigma_e^2 \) indexes the strength of the precautionary savings motive at the baseline allocation.

In this example, the inverse Euler equation provides a rationale for a constant and positive wedge \( \tau = 1 - \exp (-\sigma_e^2) \) in the agent’s Euler
equation. This is in stark contrast to the Chamley-Judd benchmark result, where no such distortion is optimal in the long run, so that agents are allowed to save freely at the social rate of return.

The efficiency gains are increasing in $\sigma^2$. Note that when $\sigma^2 = 0$, there are no efficiency gains. For small values of $\sigma^2$, the wedge is given by $\tau \approx \sigma^2$. The formula for the efficiency gains then takes the form of a simple Ramsey formula $\hat{\lambda} \approx [\beta/(1 - \beta)^2]^{1/2}$. At the other extreme, as $\sigma^2 \to -\log(\beta)$ the efficiency gains converge to 100 percent. The reason is that then $q \to 1$, implying that the present value of the baseline consumption allocation goes to infinity; in contrast, the cost of the optimal allocation remains finite.

General Equilibrium: Concave Technologies

In this section, we maintain the assumption that utility is logarithmic. We also assume that the baseline allocation is a geometric random walk representing a steady state with constant aggregates—$C_t = C_{ss}$, $K_t = K_{ss}$, and $N_t = N_{ss}$—that the shocks are lognormally distributed, and that the Euler equation holds at the baseline allocation. We depart from the partial equilibrium assumption of a linear technology and consider instead the case of concave accumulation technologies. We argue that the efficiency effects may be greatly reduced. This point is not specific to the model or forces emphasized here. Indeed, a similar issue arises in the Ramsey literature: the quantitative effects of taxing capital greatly depend on the underlying technology. It is important to confront this issue to reach meaningful quantitative conclusions.

The point that general equilibrium considerations are important can be made most clearly from the following example. We consider the extreme case of a constant endowment: the economy has no savings technology, so that $C_t \leq N_t$ for $t = 0, 1, \ldots$ (Huggett 1993). Then the baseline allocation is $\Delta$ efficient. This result follows since one finds a sequence of intertemporal prices $\tilde{q}_t$ such that the inverse Euler equation (13) holds. Thus, in this exchange economy there are no efficiency gains from perturbing the allocation—no efficiency gains from savings.

---

20 Indeed, Stokey and Rebelo (1995) discuss the effects of capital taxation in representative agent endogenous growth models. They show that the effects on growth depend critically on a number of model specifications. They then argue in favor of specifications with very small growth effects, suggesting that a neoclassical model with exogenous growth may provide an accurate approximation.

21 A similar point is at the heart of Aiyagari’s (1994) paper, which quantified the effects on aggregate savings of uncertainty with incomplete markets. He showed that, for given interest rates, the effects could be enormous but that the effects were relatively moderate in the resulting equilibrium of the neoclassical growth model.
Certainly the fixed endowment case is an extreme example, but it serves to illustrate that general equilibrium considerations are extremely important.

Consider now a neoclassical production function $F(K, N)$. Applying the results in Section IV.A, we can decompose the efficiency gains into idiosyncratic and aggregate efficiency gains. We have already argued that the corresponding efficiency gains are equal to zero: $\lambda' = 0$. Therefore, just as in the partial equilibrium case, all the efficiency gains are aggregate. The aggregate-planning problem is a simple modification of the neoclassical growth model. The solution involves a transition to a steady state with an interest rate equal to $\bar{r}_{ss} = 1/\beta - 1$. The efficiency gains, which reflect the strength of the precautionary savings motive, depend on how much lower than $\bar{r}_{ss}$ the interest rate $r_{ss} = 1/(\beta \exp (\sigma^2_x))$ at the baseline allocation is. This, in turn depends on the variance of consumption growth $\sigma^2_x$.

Empirical Evidence

Suppose one accepts the random-walk specification of consumption as a useful approximation. What does the available empirical evidence say about the crucial parameter $\sigma^2_x$? Unfortunately, the direct empirical evidence on the variance of consumption growth is very scarce, due to the unavailability of good quality panel data for broad categories of consumption. Moreover, much of the variance of consumption growth in panel data may be measurement error or attributable to transitory taste shocks unrelated to the permanent changes we are interested in here.

There are a few papers that, somewhat tangentially, provide some direct evidence on the variance of consumption growth. We briefly review some of this recent work to provide a sense of what is currently available. Using PSID data, Storesletten et al. (2004a) find that the variance in the growth rate of the permanent component of food expenditure lies between 1 and 4 percent (see their table 3, p. 708). Blundell, Pistaferri, and Preston (2008) use PSID data but impute total consumption from food expenditure. Their estimates imply a variance of consumption growth of around 1 percent (see their n. 21). Krueger and Perri (2003) use the panel element in the Consumer Expenditure

---

22 This point holds more generally in an endowment economy with CRRA utility when the baseline allocation is a geometric random walk.

23 As already noted in Sec. IVA, this analysis is akin to the one in (Lucas 1990). But while his exercise involves computing a transition to a steady state with more capital, our exercise involves a transition to a steady state with less capital.

24 For the United States, the Panel Study of Income Dynamics (PSID) provides panel data on food expenditure, and it is the most widely used source in studies of consumption requiring panel data. However, recent work by Aguiar and Hurst (2005) shows that food expenditure is unlikely to be a good proxy for actual consumption.
Survey to estimate a statistical model of consumption. At face value, their estimates imply enormous amounts of mobility and a very large variance of consumption growth—around 6–7 percent—although most of this should be attributed to a transitory, not a permanent, component.\textsuperscript{25} In general, these studies reveal the enormous empirical challenges faced in understanding the statistical properties of household consumption dynamics from available panel data.

An interesting indirect source of information is the cohort study by Deaton and Paxson (1994). This paper finds that the cross-sectional inequality of consumption rises as the cohort ages. The rate of increase then provides indirect evidence for $\sigma^2$; their point estimate implies a value of $\sigma^2 = 0.0069$. However, recent work using a similar methodology finds much lower estimates (Heathcote, Storesletten, and Violante 2004; Slesnick and Ulker 2004).

Calibration

We now display the efficiency gains as a function of $\sigma^2$. We choose three possible discount factors $\beta = 0.96, 0.97, \text{and} 0.98$. For the linear technology model, this imposes $q = \beta \exp(\sigma^2)$. For the neoclassical growth model, we take $F(K, N) = K^\alpha N^{1-\alpha}$ with $\alpha = 1/3$. Figure 2 plots the efficiency gains as a function of $\exp(\sigma^2)$ for the linear technology model and for the neoclassical growth model. The figure uses an empirically relevant range for $\exp(\sigma^2)$.

Consider first the linear technology model. For the parameters under consideration, the efficiency gains range from minuscule (less than 0.1 percent) to very large (over 10 percent). The effect of the discount factor $\beta$ is nearly equivalent to increasing the variance of shocks; that is, moving from $\beta = .96$ to $\beta = .98$ has the same effect as doubling $\sigma^2$. To understand this, interpret the lower discounting not as a change in the actual subjective discount but as calibrating the model to a shorter period length. But then holding the variance of the innovation between periods constant implies an increase in uncertainty over any fixed length of time. What matters is the amount of uncertainty per unit of discounted time.

Consider now the neoclassical growth model. There again, there is considerable variation in the size of the efficiency gains, depending on the parameters. However, note that the efficiency gains are much smaller than in the linear technology model. Large differences in interest rates

\textsuperscript{25} They specify a Markov transition matrix with nine bins (corresponding to nine quantiles) for consumption. We thank Fabrizio Perri for providing us with their estimated matrix. Using this matrix, we computed that the conditional variance of consumption growth had an average across bins of 0.0646 (this is for the year 2000, the last in their sample, but the results are similar for other years).
Fig. 2.—Percentage efficiency gains when baseline consumption is a geometric random walk for consumption and the Euler equation holds as function of $\sigma^2$. \textit{Left}, gains in partial equilibrium; \textit{right}, gains in general equilibrium. Dotted line corresponds to $\beta = .96$; dashed line, to $\beta = .97$; and solid line, to $\beta = .98$.

$\tilde{r}_c - r_c$ are necessary to generate substantial efficiency gains: it takes a difference of around 2 percent to get efficiency gains that are bigger than 1 percent.

Lessons

Three lessons emerge from our simple exercise. First, efficiency gains are potentially far from trivial. Second, they are quite sensitive to two parameters of our exercise: the variance in the growth rate of consumption and the subjective discount factor. Third, general equilibrium forces can greatly reduce efficiency gains.

V. A Numerical Exploration

In the previous section, we provided a method to compute efficiency gains as a function of a baseline consumption process. We also explored
the potential gains for a simple random-walk consumption process. On the one hand, this had the advantage of allowing for easily interpretable closed-form solutions, delivering some important insights. On the other hand, the available empirical evidence reviewed above does not provide reliable estimates for the variance of consumption growth. Moreover, consumption may not be a random walk.

A different strategy is to use consumption processes obtained from an incomplete-market income-fluctuations model. This class of models has been relatively successful at matching data on income, consumption, and wealth. Consequently, in this section we take as our baseline the steady state equilibrium allocation of such a model. Since our goal is to explore the main determinants of the gains from optimal capital taxation, we simulate model economies for a wide range of parameters. There is no absolute consensus among the empirical literature, so this broad exploration also has the advantage of spanning the type of parameters used in various studies.

A. The Model Economy

We follow Aiyagari (1994), who considered a Bewley economy in which a continuum of agents each solve an income-fluctuations problem, saving in a risk-free asset. Efficiency labor is specified as a first-order autoregressive process in logarithms

\[ \log(n_t) = \rho \log(n_{t-1}) + (1 - \rho) \log(n_{ss}) + \varepsilon_t \]

where \( \varepsilon_t \) is an i.i.d. random variable assumed normally distributed with mean zero and standard deviation \( \sigma \). With a continuum of mass 1 of agents, the average efficiency labor supply is \( N_{ss} = n_{ss} \).

Labor income is given by the product \( w_{ss} n_t \) where \( w_{ss} \) is the steady state wage. Agents face the following sequence of budget constraints:

\[ a_{t+1} + c_t \leq (1 + r_{ss})a_t + w_{ss} n_t \]

for all \( t = 0, 1, \ldots \). In addition, borrowing is not allowed: \( a_t \geq 0 \).

The equilibrium steady state wage is given by the marginal product of labor \( w_{ss} = F'(K_{ss}, N_{ss}) \), and the interest rate is given by the net marginal product of capital, \( r_{ss} = F'(K_{ss}, N_{ss}) - \delta \).\(^{26}\) For any interest rate \( r_{ss} < \beta^{-1} - 1 \), agent optimization leads to an invariant cross-sectional distribution for \( s_t \), which we denote by \( \psi \). A steady state equilibrium requires average assets, under \( \psi \), to equal the capital stock \( K_{ss} \).

Individual consumption is a function of the state variable \( s_t = (a_t, n_t) \), which evolves as a Markov process. We take this as our baseline allocation with agents distinguished by their initial conditions \( i = s_0 \), distributed according to the invariant distribution \( \psi \).

\(^{26}\) For simplicity, this assumes no taxation. It is straightforward to introduce taxation. However, we conjecture that since taxation of labor income acts as insurance, it effectively reduces the variance of shocks to net income. Lower uncertainty will then only lower the efficiency gains we compute.
B. Calibration

We simulate the economy for the wide range of parameter values considered in Aiyagari (1994). To solve the planning problem, we use the theoretical results described in Appendix B. We show how to solve the general-equilibrium-planning problem by decentralizing it using intertemporal prices. A component planner then minimizes the present value cost, taking these intertemporal prices as given. One can solve this latter optimization using recursive methods. Appendix D contains some further details on the numerical implementation of this approach.

The discount factor is set to $\beta = .96$, the production function is Cobb-Douglas with a share of capital of 0.36, and capital depreciation is 0.08. The utility function is assumed CRRA, so that $U(c) = e^{1-\gamma}(1 - \sigma)$ with $\sigma \in \{1, 3, 5\}$. Aiyagari argues, on the basis of various sources of empirical evidence, for a baseline parametrization with a coefficient of autocorrelation of $\rho = 0.6$ and a standard deviation of labor income of 20 percent. We also consider different values for the coefficient of relative risk aversion, the autocorrelation coefficient $\rho \in \{0, 0.3, 0.6, 0.9\}$, and the standard deviation of log income $\text{SD}(\log(n_t)) = \sigma_e \in \{0.2, 0.4\}$.

These values are based on the following studies. Kydland (1984) finds that the standard deviation $\sigma_e$ of annual hours worked from PSID data is around 15 percent. Using data from the PSID and the National Longitudinal Surveys, Abowd and Card (1987, 1989) find that the standard deviation of percentage changes in real earnings and annual hours are about 40 and 35 percent, respectively. They report a first-order serial correlation coefficient $\rho$ of about 0.3, resulting in an estimate of $\sigma_e$ of 34 percent. Using PSID data, Heaton and Lucas (1996) estimate a range of 0.23–0.53 for $\rho$ and a range of 27–40 percent for $\sigma_e$.

Some recent studies, within a life cycle context, consider an alternative approach and estimate a process for log earnings indirectly, by using the increase in the observed cross-sectional inequality of earnings over time within a cohort. Two views have been articulated. The first view, dating back to Deaton and Paxson (1994) and developed most recently by Storesletten et al. (2004a, 2004b), posits that the increase in earning inequality over time within a cohort is due to large and persistent income shocks. This view leads to estimates of $\rho$ around or above 0.9 and a range of estimates for $\sigma_e$ between 0.3 and 0.6, on the high end of the range of parameter values explored by Aiyagari (1994). The second view, dating back to Lillard and Weiss (1979) and Hause (1980) and developed more recently by Guvenen (2007, 2009), argues that the

27 This literature typically fits an autoregressive moving-average model ARMA(1, 1) instead of an autoregressive of order 1, AR(1), to the income process. The presence of the moving-average component results in higher estimated values for the autoregressive coefficient.
increase in cross-sectional earnings inequality over time within a cohort is better explained by an alternative model in which agents face individual-specific income profiles. This view leads to lower values for \( \rho \) and \( \sigma_p \), around 0.8 and 25 percent, respectively. Since both approaches are based on a life cycle framework, the estimates are not directly relevant for our infinite-horizon setup.\(^{28}\)

C. Results

We find that the optimized allocation always features aggregates converging to new steady state values: \( \tilde{\tilde{C}} \rightarrow \tilde{C}_{ss}, \tilde{K} \rightarrow \tilde{K}_{ss}, \) and \( \tilde{q} \rightarrow \tilde{q}_{ss} \) as \( t \rightarrow \infty \). The new steady state interest rate \( \tilde{r}_{ss} = 1/\tilde{q}_{ss} - 1 \) is always higher than \( \tilde{r}_{ss} \geq 1/\beta - 1 \), as long as \( \sigma \geq 1 \), with equality when \( \sigma = 1 \). By contrast, the initial interest rate is \( r_{ss} < 1/\beta - 1 \).\(^{29}\) Figure 3 plots the path of ag-

\(^{28}\) Indeed, as we have shown, a key object for our analysis is the conditional variance of consumption growth, which depends on the conditional variance of permanent income growth. Life cycle models incorporate a retirement period, generating a longer horizon for consumption than for labor income, reducing the impact of permanent income shocks on consumption. Taking this into account, estimates for the persistence \( \rho \) and the standard deviation of labor income derived in the context of a life cycle model would therefore have to be adjusted downward in order to be used in our numerical exercises.

\(^{29}\) By contrast, when \( \sigma < 1 \), any steady state of a \( \Delta \)-efficient allocation features an interest rate lower than \( 1/\beta - 1 \). In the case of logarithmic utility functions, we provided a formal proof of this result in Sec. IV.A. In the more general CRRA utility function case, we rely solely on our numerical results.
aggregate consumption for one particular parameter case. Since the baseline allocation represents a steady state, its aggregate consumption is constant. Aggregate consumption for the optimized allocation is initially above this level but declines monotonically, eventually reaching a new, lower, steady state. For the same parameter values, figure 4 shows a typical sample path for individual income, cash in hand, consumption, and utility. Optimized consumption appears more persistent and displays a downward trend in the initial periods.

Tables 2–4 collect the results of our simulations. All tables report our measure of efficiency gains, \( \tilde{\lambda} \). In the logarithmic utility case (\( \sigma = 1 \)), table 2 includes the idiosyncratic and aggregate components \( \tilde{\lambda}^I \) and \( \tilde{\lambda}^A \). For references, the tables show the baseline and optimized steady state interest rates \( \tilde{r}_s \) and \( \tilde{r}_a \). The column showing \( \tilde{r}_s \) closely confirms Aiyagari’s table 2 (1994, 678).

The baseline’s interest rate \( \tilde{r}_s \) is decreasing in the size and persistence of the shocks, as well as in the coefficient of relative risk aversion. Precautionary saving motives are stronger and depress the equilibrium interest rate. In particular, \( \tilde{r}_s > 1/\beta - 1 \) when utility is logarithmic, and \( \tilde{r}_s > 1/\beta - 1 \) when \( \sigma > 1 \). Moreover, we find that \( \tilde{r}_s \) increases with \( \sigma \) and the size and the persistence of the labor income shocks. These comparative statics for \( \tilde{r}_s \) are precisely the reverse of those for \( \tilde{r}_a \), which is perhaps not surprising, given the reversal in the sign of the power co-

---

Footnote 30: The figure plots the cross-sectional average of a large Monte Carlo simulation. This explains why the path is not completely smooth, displaying wiggles due to some remaining sampling error.


table 2

Efficiency gains for replication of Aiyagari (1994) when \( \sigma = 1 \)

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>( r_\text{e} (%) )</th>
<th>( \tilde{r}_\text{e} (%) )</th>
<th>Idiosyncratic</th>
<th>Aggregate</th>
<th>Total</th>
<th>Borrowing</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4.14</td>
<td>4.17</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
</tr>
<tr>
<td>.3</td>
<td>4.13</td>
<td>4.17</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
</tr>
<tr>
<td>.6</td>
<td>4.09</td>
<td>4.17</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
<td>.0</td>
</tr>
<tr>
<td>.9</td>
<td>3.95</td>
<td>4.17</td>
<td>.2</td>
<td>.0</td>
<td>.2</td>
<td>.1</td>
</tr>
</tbody>
</table>

SD(\( \log (n_i - \log (n_{i-1})) \)) = .2

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>( r_\text{e} (%) )</th>
<th>( \tilde{r}_\text{e} (%) )</th>
<th>Idiosyncratic</th>
<th>Aggregate</th>
<th>Total</th>
<th>Borrowing</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4.06</td>
<td>4.17</td>
<td>.1</td>
<td>.0</td>
<td>.1</td>
<td>.1</td>
</tr>
<tr>
<td>.3</td>
<td>3.97</td>
<td>4.17</td>
<td>.2</td>
<td>.0</td>
<td>.2</td>
<td>.2</td>
</tr>
<tr>
<td>.6</td>
<td>3.79</td>
<td>4.17</td>
<td>.4</td>
<td>.0</td>
<td>.4</td>
<td>.3</td>
</tr>
<tr>
<td>.9</td>
<td>3.38</td>
<td>4.17</td>
<td>1.2</td>
<td>.1</td>
<td>1.3</td>
<td>1.1</td>
</tr>
</tbody>
</table>

SD(\( \log (n_i - \log (n_{i-1})) \)) = .4

Efficiency gains are increasing with the variance of the shocks and with their persistence. They also increase with the coefficient of relative risk aversion. Aiyagari argues, on the basis of various sources of empirical evidence, for a parameterization with a coefficient of autocorrelation of \( \rho = 0.6 \) and a standard deviation of labor income of 20 percent. For this baseline specification, we find that efficiency gains are small—below 0.2 percent for all three values for the coefficient of relative risk aversion.

In the logarithmic case, the efficiency gains we compute are less than 1.3 percent. Our decomposition along the lines of Section IV.A shows that idiosyncratic efficiency gains \( \tilde{\lambda}_I \) are more important than aggregate efficiency gains \( \tilde{\lambda}_A \). For the baseline parameterization, idiosyncratic efficiency gains are quite small. This could have perhaps been anticipated by our illustrative geometric random-walk example, where idiosyncratic gains are zero. Intuitively, efficiency gains from the idiosyncratic component require differences in the expected consumption growth rate across individuals. When individuals are able to smooth their consumption over time effectively, the remaining differences are small—as a result, so are the efficiency gains. Efficiency gains from the aggregate component are directly related to the difference between the equilibrium and optimal steady state capital. With logarithmic utility, this is equivalent to the difference between the equilibrium steady state interest rate and \( \delta^{-1} - 1 \), the interest rate that obtains with complete markets. Hence, our findings of low aggregate efficiency gains are directly related to Aiyagari’s (1994) main conclusion: for shocks that are not too large or for moderate risk aversion, precautionary savings are small in the
aggregate, in that steady state capital and interest rate are close in their complete-markets levels, as shown in our figure 2.

For the range of parameters that we consider, the efficiency gains range from less than 0.1 to 8.4 percent. However, efficiency gains larger than 1.3 percent are only reached for combinations of high values of relative risk aversion ($\sigma > 3$) and both large and highly persistent shocks—a standard deviation of labor income of 40 percent and a mean-reversion coefficient $\rho > 0.6$.

### D. The Role of Borrowing Constraints

The market arrangement in Aiyagari’s economy imposes borrowing constraints that limit the ability to trade consumption intertemporally. At the
baseline allocation, the Euler equation holds with equality $U'(c(u'_i)) = \beta (1 + r_w) E_t[U'(c(u'_{i+1}))]$ when assets and current income are high enough. However, for low levels of assets and income, the agent may be borrowing constrained, so that the Euler condition holds with strict inequality $U'(c(u'_i)) > \beta (1 + r_w) E_t[U'(c(u'_{i+1}))]$.

In contrast, since our planning problem does not impose arbitrary restrictions on the perturbations, it places no such limits on intertemporal reallocation of consumption. Thus, the perturbations can effectively undo limits to borrowing. As a result, part of the efficiency gains we compute can be attributed to the relaxation of borrowing constraints, not to the introduction of savings distortions.

To get an idea of the efficiency gains that are obtained from the relaxation of borrowing constraints, we performed the following exercise. The basic idea is to use the perturbations to construct a new allocation where the Euler equation always holds with equality. That is, the new allocation stops short of satisfying the inverse Euler equation, so it does not introduce positive intertemporal wedges. Instead, it removes the negative intertemporal wedges that were present due to borrowing constraints. We compute the resource savings from these perturbations as a simple measure of the efficiency gains due to the relaxation of borrowing constraints.

More precisely, we seek to determine the unique allocation $\{\tilde{u}_i^B, n_i^p, \tilde{K}_i^B, \tilde{\lambda}_i^B C_{i+1}\}$, where $B$ stands for borrowing, that satisfies the following constraints. First, we impose that the allocation be achievable through parallel perturbations of the baseline allocation and deliver the same utility $U(\{u_i\}, 0)$. Second, we impose that the resource constraints hold

$$\tilde{K}_i^B + \int \mathbb{E}_t[c(\tilde{u}_i^B)] d\psi + \tilde{\lambda}_i^B C_{i+1} \leq (1 - \delta) \tilde{K}_i^B + F(\tilde{K}_i^B, N_{i+1}), \quad t = 0, 1, \ldots,$$

and require that the initial capital be equal to the initial capital $\tilde{K}_0^B = K_0$ of the baseline allocation. Finally, we impose that the Euler equation hold for every agent in every period:

$$U'(c(\tilde{u}_i^B)) = \beta [1 - \delta + F(\tilde{K}_i^B, N_{i+1})] E_t[U'(c(\tilde{u}_{i+1}^B))].$$

The allocation $\{\tilde{u}_i^B, n_i^p, \tilde{K}_i^B, \tilde{\lambda}_i^B C_{i+1}\}$ can improve on the baseline allocation by insuring that the Euler equation holds for every agent in every period. Note that this allocation does not satisfy the inverse Euler equation and is therefore not $\Delta$ efficient. We adopt $\tilde{\lambda}_i^B$ as our measure of the efficiency gains deriving from the relaxation of borrowing constraints.

The rightmost column in each table reports the efficiency gains $\tilde{\lambda}_i^B$ that can be attributed to the relaxation of borrowing constraints. Perhaps the most important conclusion of our numerical exercise can be
drawn by comparing the total efficiency gains $\tilde{\lambda}$ to $\tilde{\lambda}^B$. We find that, across all specifications, most of the efficiency gains come from the relaxation of borrowing constraints. Indeed, the difference between $\tilde{\lambda}$ and $\tilde{\lambda}^B$ is less than 0.1 percent, except for the largest and most persistent shocks—a standard deviation of labor income growth of 40 percent and a mean-reversion coefficient $\rho = 0.9$. In other words, most of the efficiency comes from allowing agents to better smooth their consumption over time by alleviating borrowing constraints, rather than by optimally distorting savings as prescribed by the inverse Euler equation.

VI. Conclusions

This paper provided a method for evaluating the auxiliary role that savings distortions play in social insurance arrangements. We put it to use to evaluate the welfare importance of recent arguments for capital taxation on the basis of the inverse Euler equation. Our main finding is that, once general equilibrium effects are taken into account, these gains are small.

The method developed here is flexible enough to accommodate several extensions, and it may be of interest to investigate how these may affect the quantitative conclusions found here for our benchmark calibration. In a separate paper (Farhi and Werning 2008), we pursued such an extension using Epstein-Zin preferences that separate risk aversion from the intertemporal elasticity of substitution.

Our results do not imply that dynamic aspects of social insurance design are unimportant, only that the role for savings distortions is modest. More work is needed to characterize other aspects of efficient insurance arrangements. In Farhi and Werning (2010), we take a step in this direction, delivering an efficiency condition that provides a tight characterization of the optimal labor distortions.

Appendix A

Proofs for Section III.B

To simplify the arguments, we first assume that the horizon is finite with terminal period $T > 0$. We also assume that $V(n; \theta)$ is continuously differentiable with respect to $n$ and $\theta$. We also assume throughout that shocks are continuous.

Consider an assignment for utility and labor $\{u_i^n, n_i\}$. We define the value $U(\theta^t)$ conditional on history $\theta^t$ as follows:

$$U^t(\theta^{t+}) = \mathbb{E} \left[ \sum_{s=t}^T \beta^s [u_{i,v}(\theta^{s+}) - V(n_{i,v}(\theta^{s+}); \theta^{i,v})] \mid \theta^{i,v} \right].$$

Fix a history $\theta^t$ and consider a report $\hat{\theta}_i^t \in \Theta$. Define the strategy $\sigma^t_{\hat{\theta}_i}$ as follows: $\sigma^t_{\hat{\theta}_i}(\theta^{s+}) = \theta^{s+},$ except if $\theta^{s+} \succeq \theta^{s+}$, in which case
where the report can be different from the true shock. For clarity and brevity, we use the notation \( \hat{\theta}' \) for \( \sigma_{u_i}'(\theta, v_i, \theta_{i+1}, \theta_{i+2}, \ldots) \). We denote the continuation utility after history \( \theta^{i,t} \geq \theta^{i,t} \) under the strategy \( \sigma_{u_i}' \), by \( U'(\theta^{i,t}; \theta^{i,t}) \), or \( U'(\hat{\theta}^{i,t}; \hat{\theta}^{i,t}) \) for short. Similarly, we denote by \( \mathbb{E}[U'(\hat{\theta}^{i,t}; \hat{\theta}^{i,t})] \) or \( \mathbb{E}[U'(\hat{\theta}^{i,t})\hat{\theta}^{i,t}] \) the expectation of this continuation utility, conditional on the realized history \( \theta^{i,t} \) at date \( t \).

We say that the assignment for utility and labor \( \{u_i, n_i\} \) is regular if for all \( t \geq 0 \) and \( \theta^{i,t-1} \in \Theta^t \), the continuation utility \( U'(\theta^{i,t}; \theta_{i+1}) \) is absolutely continuous in the true shock \( \theta_i \) and differentiable with respect to the true shock \( \theta_i \), and the derivative, which we denote by

\[
-\frac{\partial}{\partial \theta_i} V(n(\theta'; \theta_i)) + \beta \frac{\partial}{\partial \theta_i} \mathbb{E}[U'(\hat{\theta}^{i,t+1})|\theta'; \theta_i],
\]

is bounded by a function \( b(\theta'; \theta_i) \), which is integrable with respect to \( \theta_i \).

Consider two regular, incentive compatible, assignments for utility and labor \( \{u_i, n_i\} \) and \( \{\hat{u}_i, n_i\} \) that share the same assignment for labor \( \{n_i\} \). Denote by \( \{U\} \) and \( \{U\} \) the corresponding continuation utilities. We now show that there exists \( \Delta_{-1} \in \mathbb{R} \) such that \( \{\hat{u}_i\} \in \mathcal{T}\{u_i\}, \Delta_{-1}\).

Note we can restate the result as follows: for all \( t \) with \( 0 \leq t \leq T-1 \) and for all \( \theta^{i,t-1} \in \Theta^t \), there exists \( \Delta(\theta^{i,t-1}) \in \mathbb{R} \) such that

\[
\hat{U}'(\theta^{i,t}) = U'(\theta^{i,t}) + \Delta(\theta^{i,t-1}).
\]

(A1)

Fix a history \( \theta^{i,t} \) and consider the strategies \( \sigma_{u_i}' \). Using theorem 4 in Milgrom and Segal (2002), incentive compatibility and regularity implies the following envelope condition:

\[
U'(\theta^{i,t-1}, \theta_i) = U'(\theta^{i,t-1}, \theta_i) + \int_{\theta_i}^{\theta_{i-1}} \left[ -\frac{\partial}{\partial \theta_i} V(n(\theta'; \theta_i)) + \beta \frac{\partial}{\partial \theta_i} \mathbb{E}[U'(\hat{\theta}^{i,t+1})|\theta'; \theta_i] \right] d\theta_i,
\]

(A2)

where \( \hat{\theta}^{i,s} = (\theta^{i,t-1}, \hat{\theta}_{i+1}, \ldots, \theta_{i+s}) \) for all \( s \geq 0 \), and \( \hat{\theta}^{i,s} = \theta^{i,t-1} \) for all \( s > 0 \).

The same expression holds for \( \{\hat{u}_i, n_i\} \). Using the fact that shocks are continuous and that \( \{u_i, n_i\} \) and \( \{\hat{u}_i, n_i\} \) share the same labor assignment, this implies that there exists \( \Delta(\theta^{i,t-1}) \in \mathbb{R} \) such that

\[
\hat{U}'(\theta^{i,t}) = U'(\theta^{i,t}) + \Delta(\theta^{i,t-1}) + \beta \int_{\theta_i}^{\theta_{i-1}} \frac{\partial}{\partial \theta_i} \mathbb{E}[\hat{U}'(\hat{\theta}^{i,t+1}) - U'(\hat{\theta}^{i,t+1})|\theta^{i,t-1}] d\theta_i.
\]

For \( t = T \), the last term on the right-hand side vanishes, and we have that for all \( \theta^{i,T} \in \Theta^{T+1} \), there exists \( \Delta(\theta^{i,T-1}) \in \mathbb{R} \) such that (A1) holds. We proceed by induction on \( t \). Suppose that \( t \geq 0 \) and that for all \( \theta^{i,t} \in \Theta^{t+1} \), there exists \( \Delta(\theta^{i,t}) \in \mathbb{R} \) such that (A1) holds. Then applying (A2) and using the fact that

\[
\frac{\partial}{\partial \theta_i} \mathbb{E}[\Delta(\hat{\theta}^{i,t})|\theta^{i,t-1}, \hat{\theta}^{i,t}] = 0,
\]
we immediately find that for all \( \theta_{t+1} \in \Theta \) there exists \( \Delta(\theta_{t+1}) \in \mathbb{R} \) such that
\[ \tilde{U}^i(\theta_{t+1}) = U^i(\theta_{t+1}) + \Delta(\theta_{t+1}). \]

The proof follows by induction.

When the horizon is infinite, the regularity conditions must be complemented with a limit condition. For example, expanding our envelope condition over two periods, we find
\[ \tilde{U}^i(\theta_{t+1}) = U^i(\theta_{t+1}) + \Delta(\theta_{t+1}) + \beta^2 \int_{\tilde{\theta}} \frac{\partial}{\partial \tilde{\theta}_{t+1}^i} \tilde{U}^i(\theta_{t+1}) \left( \tilde{U}^i_1(\tilde{\theta}_{t+1}^i) \right) d\tilde{\theta}_{t+1}^i. \]

Consider the sequence that is constructed by iterating our envelope condition for the utility and labor assignment \( \{u_t', n_t'\} \). The first element of the sequence is
\[ \beta \int_{\tilde{\theta}} \frac{\partial}{\partial \theta_{t+1}^i} \left( \tilde{U}^i(\theta_{t+1}) \right) d\theta_{t+1}^i. \]

Similarly, the second element of the sequence is
\[ \beta^2 \int_{\tilde{\theta}} \frac{\partial}{\partial \theta_{t+1}^i} \left( \tilde{U}^i(\theta_{t+1}) \right) d\theta_{t+1}^i. \]

Our proof then carries over, if the condition that the limit of these terms when the number of iterations goes to infinity is equal to zero is added to the requirements for regularity.

**Appendix B**

**Solving the Planning Problem: Relaxed Planning Problem and Component-Planning Problems**

**A. Relaxed Problem**

We now construct a relaxed planning problem that replaces the resource constraints with a single present value condition. This problem is indexed by a sequence of intertemporal prices or interest rates, which encodes the scarcity of aggregate resources in every period. The relaxed planning problem can be further decomposed in a series of component-planning problems corresponding to the different types \( i \in I \).

Given some positive intertemporal prices \( \{\tilde{Q}\} \) with the normalization that \( \sum_{i=0}^{\infty} \tilde{Q}C_i = 1 \), we replace the sequence of resource constraints (15) with the single present value condition
\[ \tilde{\lambda} = \sum_{i=0}^{\infty} \tilde{Q}_i \left( \tilde{F}(\tilde{K}_n, N) + (1 - \delta) \tilde{K}_i - \tilde{K}_{i+1} - \int \bar{E}(\bar{u}_t) d\theta \right), \quad (B1) \]
which is obtained by multiplying equation (15) by $\tilde{Q}_t$ and summing over $t = 0, 1, \ldots$. Formally, the relaxed planning problem seeks the allocation $\{\tilde{u}_t, n'_t, \tilde{K}_t, \tilde{\lambda}C_t\}$, where $\{\tilde{u}_t\} \in \mathcal{T}([u]_0, 0)$ that maximizes $\tilde{\lambda}$ given by equation (B1).

The connection with the original planning problem is as follows. Suppose that, given some $\{\tilde{Q}_t\}$, the optimal allocation $\{\tilde{u}_t, n'_t, \tilde{K}_t, \tilde{\lambda}C_t\}$ for the relaxed problem satisfies the resource constraints (15). Then, this allocation solves the original planning problem. This relaxed problem approach is adapted from Farhi and Werning (2007) and is related to the first welfare theorem proved in Atkeson and Lucas (1992).

The converse is also true. Indeed, the prices $\{\tilde{Q}_t\}$ are Lagrange multipliers, and Lagrangian necessity theorems guarantee the existence of prices $\{\tilde{Q}_t\}$ for the relaxed problem. The following lemma, which follows immediately from theorem 1, section 8.3, in Luenberger (1969), provides one such result.

**Lemma 1.** Suppose $\{\tilde{u}_t, n'_t, \tilde{K}_t, \tilde{\lambda}C_t\}$ solves the planning problem and the resource constraints (15) hold with equality. Then there exists a sequence of prices $\{\tilde{Q}_t\}$ such that this same allocation solves the relaxed planning problem.

The relaxed planning problem can be decomposed into a subproblem for capital and a series of component-planning problems for the utility assignment. The subproblem for capital maximizes the right-hand side of equation (B1) with respect to $\tilde{u}_0$. The first-order conditions, which are necessary and sufficient for an interior optimum, are

$$1 = \tilde{q}_t(F_t(K_{t+1}, N_{t+1}) + 1 - \delta), \quad t = 0, 1, \ldots,$$

where $\tilde{q}_t \equiv \tilde{Q}_{t+1}/\tilde{Q}_t$. This equation, together with the normalization that $\sum_{t=0}^{\infty} \tilde{Q}_tC_t = 1$, implies a one-to-one relationship between $\{K_{t+1}\}$ and $\{\tilde{Q}_t\}$.

The component-planning problem for type $i \in I$ maximizes the right-hand side of equation (B1) with respect to the utility assignment $\{\tilde{u}_t\} \in \mathcal{T}([u]_0, 0)$. The objective reduces to minimizing the present value of consumption:

$$\sum_{i=0}^{\infty} \tilde{Q}_t \mathbb{E}[c(\tilde{u}_t)].$$

(B2)

Each of these component-planning problems can be solved independently. Since both the objective and the constraints are convex, it follows immediately that, given $\{\tilde{q}_t\}$, the first-order conditions for optimality at an interior solution in the component-planning problem (B2) coincide exactly with the inverse Euler equation (13):

$$c'(\tilde{u}_t) = \frac{\tilde{q}_t}{\beta} \mathbb{E}[c'(\tilde{u}_{t+1})], \quad t = 0, 1, \ldots.$$

**B. A Bellman Equation**

In most situations of interest, the baseline allocation admits a recursive representation for some endogenous state variable. This is the case whenever $\{\theta_0^i\}$ is a Markov process and the baseline allocation depends on the history of shocks $\theta^i_{t-1}$ in a way that can be summarized by an endogenous state $x^i_t$, with law of motion $x^i_t = M(x^i_{t-1}, \theta^i_t)$ and given initial condition $x^i_0$. Note that the transition matrix $M$ is assumed to be independent of $i$. Types then correspond to different
initial value $x_i^0$. This difference is the only thing that distinguishes them. The
endogenous state $x'_i$ is a function of the history of exogenous shocks $\theta^{it}$. Defining
the state vector $s'_i = (x'_i, \theta'_i)$, there must exist a function $\tilde{u}$ such that $u(\theta^{it}) =
\tilde{u}(s'_i)$ for all $\theta^{it}$. In what follows, we drop the hat notation, and we stop indexing
the allocation by the type $i$ of the agents. We denote a baseline allocation by $u(s)$ and use the notation $c(s)$ for $c(u(s))$.

For such baseline allocations, we can reformulate the component-planning
problems recursively as follows. The idea is to take $s$, as an exogenous process
and keep track of the additional lifetime utility previously promised as an
endogenous state variable.

For any date $\tau$, define the continuation plans $\{u'_i\}_{i=0}^\infty$ with $u'_i(s) = u_{i+\tau}(s)$ and
the value function corresponding to the component-planning problem starting
at date $\tau$ with state $s$

$$K(\Delta_\tau; s; \tau) = \inf \n \sum_{i=0}^\infty \frac{\tilde{Q}_i}{\tilde{Q}_0} \mathbb{E}[c(\tilde{u}_i)|s_i = s]\text{ such that }\{\tilde{u}_i\} \in T([u'_i], \Delta_\tau).$$

This value function satisfies the Bellman equation

$$K(\Delta_\tau; s; \tau) = \min_\Delta [c(u(s) + \Delta - \beta \Delta) + \tilde{q}_i \mathbb{E}[K(\Delta, s'; \tau + 1)|s]] \quad (B3)$$

The optimization over $\Delta$ is one-dimensional and convex and delivers a policy
function $g(\Delta_\tau; s; \tau)$. Combining the necessary and sufficient first-order condition
for $\Delta$ with the envelope condition yields the inverse Euler equation in recursive
form

$$c'(u(s) + \Delta - \beta g(\Delta_\tau; s; \tau)) = \frac{\tilde{q}}{\beta} \mathbb{E}[c'(u(s')) + g(\Delta_\tau; s; \tau) - \beta g(g(\Delta_\tau; s; \tau), s'; \tau + 1)|s].$$

This condition can be used to compute $g(\cdot, \cdot; \tau)$ for given $g(\cdot, \cdot; \tau + 1)$. An
optimal plan for $\{\tilde{u}_i\}$ can then be generated from the sequence of policy functions
$\{g(\cdot, \cdot; \tau)\}$, by setting $\tilde{u}(s') = u(s) + \Delta(s') - \beta \tilde{u}(s')$ and using the recursion
$\Delta(s') = g(\Delta(s' - 1), s; \tau)$ with initial condition $\Delta_{-1} = 0$.

With a fixed discount factor $q$, the value function is independent of time,
$K(\Delta_\tau; s)$, and solves the stationary Bellman equation

$$K(\Delta_\tau, s) = \min_\Delta [c(u(s) + \Delta - \beta \Delta) + q \mathbb{E}[K(\Delta, s')|s]] \quad (B4)$$

This dynamic program admits an analogy with a consumer’s income-fluctuation
problem that is both convenient and enlightening. We transform variables by
changing signs and switch the minimization to a maximization. Let $\tilde{\Delta} = -\Delta$, $
\tilde{K}(\Delta_\tau; s) = -K(-\Delta_\tau; s)$, and $\tilde{U}(x) = -c(-x)$. Note that the pseudo utility function
$\tilde{U}$ is increasing and concave and satisfies Inada conditions at the extremes of
its domain. Reexpressing the Bellman equation (B4) using these transforma-

31 An important case is when the original utility function is CRRA $U(c) = c^\sigma/(1 - \sigma)$ for $\sigma > 0$ and $c \geq 0$. Then for $\sigma > 1$, the function $\tilde{U}(x)$ is proportional to a CRRA with coefficient of relative risk aversion $\tilde{\sigma} = \sigma/(\sigma - 1)$ and $x \in (0, \infty)$. For $\sigma < 1$, the pseudo utility $\tilde{U}$ is “quadratic-like” in that it is proportional to $-(-x)^{\rho}$ for some $\rho > 1$, and $x \in (-\infty, 0]$. 0.439
This reformulation can be read as the problem of when a consumer with a constant discount factor \( q \) facing a constant gross interest rate \( 1 + r = \beta^{-1} \), entering the period with pseudo financial wealth \( \tilde{\Delta} \), receives a pseudo labor income shock \(-u(s)\). The fictitious consumer must decide how much to save \( u(s) \); pseudo consumption is then \( x = -u(s) + \tilde{\Delta} - \beta \tilde{\Delta} \). The benefit of this analogy is that the income-fluctuations problem has been extensively studied and used; it is at the heart of most general equilibrium incomplete-market models (Aiyagari 1994).

With logarithmic utility, the Bellman equation can be simplified considerably. The idea is best seen through the analogy, noting that the pseudo utility function is exponential in this case. It is well known that for a consumer with CARA preferences, a one-unit increase in financial wealth, \( \Delta \), results in an increase in pseudo consumption, \( x \), of \( r/(1 + r) = 1 - \beta \) in parallel across all periods and states of nature. It is not hard to see that this implies that the value function takes the form \( \tilde{K}(\tilde{\Delta}; s) = e^{-(1-\beta)\tilde{\Delta}} \tilde{k}(s) \). These ideas are behind the following result.

**Proposition 4.** With logarithmic utility and constant discount \( q \), the value function in equation (B4) is given by

\[
\tilde{K}(\tilde{\Delta}, s) = e^{(1-\beta)\tilde{\Delta}} k(s),
\]

where function \( k(s) \) solves the Bellman equation

\[
k(s) = A e(s)^{1-\beta} \left( \mathbb{E}[k(s')|s] \right)^{\beta},
\]

where \( A = (q/\beta)^{\beta}/(1-\beta)^{1-\beta} \). The optimal policy for \( \Delta \) can be obtained from \( k(s) \) using

\[
\Delta = \Delta - \frac{1}{\beta} \log \left( \frac{(1-\beta)k(s)}{c(s)} \right).
\]

**Proof.** With logarithmic utility, the Bellman equation is

\[
K(s, \Delta) = \min_{\Delta} \left[ s \exp (\Delta - \beta \Delta) + q \mathbb{E}[K(s', \Delta)|s] \right]
\]

\[
= \min_{\Delta} \left[ s \exp ((1-\beta)\Delta + \beta(\Delta - \Delta)) + q \mathbb{E}[K(s', \Delta)|s] \right].
\]

Substituting that \( K(\Delta, s) = k(s) \exp ((1-\beta)\Delta) \) gives

\[
k(s) \exp ((1-\beta)\Delta)
\]

\[
= \min_{\Delta} \left[ s \exp ((1-\beta)\Delta + \beta(\Delta - \Delta)) + q \mathbb{E}[k(s') \exp ((1-\beta)\Delta)|s] \right]
\]

and canceling terms
\[ k(s) = \min_d \{ s \exp (\beta (\Delta - \Delta)) + q \mathbb{E}[k(s')] \exp ((1 - \beta)(\Delta - \Delta))[s] \} \]

\[ = \min_d \{ s \exp (-\beta d) + q \mathbb{E}[k(s')] \exp ((1 - \beta)d)[s] \} \]

\[ = \min_d \{ s \exp (-\beta d) + q \mathbb{E}[k(s')] \exp ((1 - \beta)d) \}, \]

where \( d \equiv \Delta - \Delta \). We can simplify this one-dimensional Bellman equation further. Define \( \hat{q}(s) \equiv q \mathbb{E}[k(s')] / s \) and

\[ M(\hat{q}) \equiv \min \{ \exp (-\beta d) + \hat{q} \exp ((1 - \beta)d) \}. \]

The first-order conditions gives

\[ \beta \exp (-\beta d) = \hat{q}(1 - \beta) \exp ((1 - \beta)d) \Rightarrow d = \log \left( \frac{\beta}{(1 - \beta) \hat{q}} \right). \] (B7)

Substituting back into the objective, we find that

\[ M(\hat{q}) = \frac{1}{1 - \beta} \exp (-\beta d) = \frac{1}{1 - \beta} \exp \left( -\beta \log \left( \frac{\beta}{(1 - \beta) \hat{q}} \right) \right) \]

\[ = \frac{1}{(1 - \beta)^{1 - \beta} \hat{q}^\beta} = B \hat{q}^\beta, \]

where \( B \) is a constant defined in the obvious way in terms of \( \beta \). The operator associated with the Bellman equation is then

\[ T[k](s) = sM \left( \frac{\mathbb{E}[k(s')][s]}{s} \right) = As^{1-\beta}(\mathbb{E}[k(s')] \hat{q}^\beta, \]

where \( A \equiv B \hat{q}^\beta = (q/\beta)^\beta / (1 - \beta)^{1 - \beta} \). Combining the Bellman \( k(s)/s = M(\hat{q}) = A(1-\beta)^{1-\beta} \hat{q}^\beta \) with equation (B7) yields the policy function as a function of \( K(s) \). QED

This solution is nearly closed form: one needs only compute \( k(s) \) using the recursion in equation (B5), which requires no optimization. No simplifications on the stochastic process for skills are required.

C. Idiosyncratic Gains with Log Utility

In this section, we assume log utility. The idiosyncratic planning problem lends itself to a similar analysis. In order to avoid repetitions, we only sketch the corresponding analysis. We can define a corresponding relaxed problem, given a sequence of prices \( \{\hat{Q}_t\} \). We can also prove an analogue of lemma 1. This relaxed planning problem can then be decomposed into a series of component-planning problems. When the baseline allocation is recursive, and given a sequence of prices \( \{\hat{Q}_t\} \), we can study the component-planning problems using a Bellman equation as in equation (B5). The corresponding first-order conditions also take the form of an inverse Euler equation

\[ c' (\hat{u}_t) = \frac{\hat{q}}{\beta} \mathbb{E} [c'(\hat{u}_{t+1})], \quad t = 0, 1, \ldots , \]

where \( \hat{q} \equiv \hat{Q}_{t+1} / \hat{Q}_t \).
D. A Complete Solution When the Baseline Is a Recursive Steady State with Log Utility

In this section, we assume log utility. Suppose that the baseline allocation is recursive with state $s$ and features constant aggregate consumption $C_s = C_s$. It is not necessary for the argument that aggregate capital $K_t$ and labor $N_t$ be constant at the baseline allocation. The idiosyncratic efficiency gains can then be easily computed. We can use $\hat{q} = \beta$ to compute the idiosyncratic allocation. Proposition 4 then allows us to compute the solution in closed form. This leads to

$$\hat{\lambda}' = 1 - \frac{(1 - \beta)\int k(s) d\psi(s)}{C_s},$$

where $k(s)$ solves equation (B5) with $q = \beta$. We can then apply proposition 2 and compute the efficiency gains $\lambda$ and the corresponding $\Delta$-efficient allocation by solving the aggregate-planning problem, a version of the neoclassical growth model. This method provides a complete solution to the planning problem with log utility when the baseline allocation is recursive and features constant aggregate consumption.

Appendix C

Proof of Proposition 2

Consider the aggregate allocation $\{\hat{C}_t, N_t, \hat{K}_t, \hat{\lambda}C_t\}$ that solves the aggregate-planning problem and the utility assignment $\{\hat{u}_t\} \in \mathcal{T}(\{u_t\}, 0)$ that solves the idiosyncratic planning problem as well as the corresponding idiosyncratic efficiency gains $\hat{\lambda}'$. Since the resource constraints hold with equality in the idiosyncratic planning problem, we know from lemma 1 that there exists a sequence of prices $\{\hat{Q}_t\}$ such that

$$c'(\hat{u}_t) = \frac{\hat{q}_t}{\beta} E_t[c'(\hat{u}_{t+1})], \quad t = 0, 1, \ldots, \quad (C1)$$

where $\hat{q}_t = \hat{q}_{t+1}/\hat{Q}_t$. Moreover, the sequence $\{\hat{q}_t\}$ is given by $\hat{q}_t = \beta C_t/C_{t+1}$. The aggregate allocation $\{\hat{C}_t, N_t, \hat{K}_t, \hat{\lambda}C_t\}$ satisfies the necessary and sufficient first-order conditions

$$U'(\hat{C}_t) = \beta(1 - \delta + F_k(\hat{K}_{t+1}, N_{t+1}))U'(\hat{C}_{t+1}), \quad t = 0, 1, \ldots. \quad (C2)$$

Define the following sequence $\{\delta_t\}$:

$$\delta_t = -U((1 - \hat{\lambda})C_t) + U(\hat{C}_t), \quad t = 0, 1, \ldots.$$ 

We have $\sum_{t=0}^\infty \beta^t \delta_t = 0$. With this choice of $\{\delta_t\}$, we then define a utility assignment $\{\tilde{u}_t\} \in \mathcal{T}(\{u_t\}, 0)$ as follows: $\tilde{u}_t = \hat{u}_t + \delta_t$. The allocation $\{\tilde{u}_t, n_t, \tilde{K}_t, \tilde{\lambda}C_t\}$ then satisfies all the constraints of the original planning problem. Moreover, using $c'(\tilde{u}_t) = c'(\delta_t)c'(\hat{u}_t)$ and equations (C1) and (C2), we find that

$$c'(\tilde{u}_t) = \frac{\tilde{q}_t}{\beta} E_t[c'(\tilde{u}_{t+1})], \quad t = 0, 1, \ldots,$$
where

\[ 1 = \tilde{q}(F(K^{t+1}, N^{t+1}) + 1 - \delta), \quad t = 0, 1, \ldots. \]

Hence, the allocation \( \{ \tilde{u}, n^t, \tilde{K}, \tilde{\lambda} C \} \) satisfies the sufficient first-order conditions in the planning problem. It therefore represents the optimum.

**Appendix D**

**Numerical Method**

To solve the planning problem, we use the result developed in Appendix B, Section A. To apply this result, we seek the appropriate sequence of discount factors \( \{ \tilde{q} \} \) as follows. For any given \( \{ \tilde{q} \} \), we solve the nonstationary Bellman equation (B3) using a policy iteration method with \( \Delta_{t-1} \) as the endogenous state variable and \( s_t \) as the exogenous state. Using the underlying policy function for consumption in equation (B3), and integrating in every period over \( \psi \), we compute an aggregate sequence of consumption \( \{ C \} \). Using the resource constraints, we can solve for \( \tilde{\lambda} \) and a sequence for capital \( \{ \tilde{K} \} \) that has \( \tilde{K}_0 = K_{0\infty}. \)

From Appendix B, Section A, if the condition \( 1/\tilde{q} = 1 + F(K, N) - \delta \) is met, this constitutes a solution. Otherwise, we take a new sequence of discount factors given by \( \tilde{q}_t = (1 + F(K, N) - \delta)^{-1} \) and iterate until convergence to a fixed point.

**References**


Farhi, Emmanuel, and Iván Werning. 2007. “Inequality and Social Discounting.”


