Submitted to Econometrica

CONSTRAINED EFFICIENCY IN THE NEOCLASSICAL GROWTH MODEL WITH UNINSURABLE IDIOSYNCRATIC SHOCKS[1]

JULIO DÁVILA, JAY H. HONG, PER KRUSELL[2,3] AND JOSÉ-VÍCTOR RÍOS-RULL[4]

We investigate the welfare properties of the one-sector neoclassical growth model with uninsurable idiosyncratic shocks. We focus on the notion of *constrained efficiency* used in the general equilibrium literature. Our characterization of constrained efficiency uses the first-order condition of a constrained planner's problem. This condition highlights the margins of relevance for whether capital is too high or too low: the factor composition of income of the (consumption-)poor. Using three calibrations commonly considered in the literature, we illustrate that there can be either over- or underaccumulation of capital in steady state and that the constrained optimum may or may not be consistent with a nondegenerate long-run distribution of wealth. For the calibration that roughly matches the income and wealth distribution, the constrained inefficiency of the market outcome is rather striking: it has much too low a steady-state capital stock.

KEYWORDS: Constrained Efficiency, Uninsurable Shocks.

1. INTRODUCTION

In this paper we investigate the welfare properties of the one-sector neoclassical growth model with uninsurable idiosyncratic shocks but precautionary savings. This kind of model was originally developed and analyzed by Bewley (1986), Imrohoroğlu (1989), Huggett (1993), and Aiyagari (1994), and it has become a standard workhorse for quantitatively based theoretical analysis of macroeconomics and inequality. The framework is mostly used for positive analysis, but in this paper we focus on some of its normative properties. In particular, we study the *constrained efficiency* of the market allocations. We define this concept following Diamond (1967), who in a similar context was interested in a notion of efficiency that did not allow the planner to directly overcome the friction implied by missing markets. Constrained inefficiency is sometimes also thought of in terms of a "pecuniary externality": the incomplete market structure itself induces outcomes that could be improved on, in the Pareto sense, if consumers

Dávila: Center for Operations Research and Econometrics, Université c. de Louvain, Belgium and Paris School of Economics, Centre d'Economie de la Sorbonne-CNRS, France. julio.davila@univ-paris1.fr

Hong: University of Rochester. jay.hong@rochester.edu

Krusell: IIES, CAERP, CEPR, NBER.Per.Krusell@iies.su.se

Ríos-Rull: University of Minnesota, Federal Reserve Bank of Minneapolis, CAERP, CEPR, NBER. vr0j@umn.edu

[1]We thank Tim Kehoe, Michael Magill, Iván Werning, and Martine Quinzii, as well as the editor and three anonymous referees, for very helpful comments.

[2]Corresponding author.

[3]Krusell thanks the National Science Foundation.

[4]Ríos-Rull thanks the National Science Foundation and the University of Pennsylvania Research Foundation for support. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

merely acted differently, i.e., if they used the same set of markets but departed from purely self-interested optimization.

We find that the laissez-faire equilibrium generally is not constrained efficient in our economy. In the baseline model, the key decision consumers make is how much to save, and the amount of savings influences equilibrium prices: the return to capital and the wage. Furthermore, prices influence how the market incompleteness affects consumers. With shocks to individual wages or employment, a lower wage and a higher rental rate make the uninsurable part of income smaller, suggesting that lower economy-wide saving may be welfare improving. Indeed, for a two-period economy we show that the laissez-faire equilibrium unambiguously has too much savings: if consumers just saved a little less, they would all face less uninsurable risk and all be better off. However, we also show that in economies with longer lives, an effect appears that works in the opposite direction. Idiosyncratic wage shocks generate wealth inequality, which propagates over time. Thus, since higher aggregate savings lower the return to wealth, an *increase* in savings can improve consumer welfare by reducing the wealth inequality induced by missing insurance markets.

A central part of the paper involves characterizing the precise conditions under which there is oversaving and undersaving. We pay particular attention to the infinite horizon version of our economy. We study the properties of constrained-optimal steady states and compare them to the laissez-faire outcomes. The analysis is based on a functional first-order condition that is a necessary condition of the constrained-efficiency planning problem. This first-order condition is, to our knowledge, new, and it is one of the key analytical tools put forth in this paper.

Our central finding is that whether there is over- or underaccumulation of capital depends very importantly on the *factor composition of the income of the poor agents*. A key factor behind whether the constrained optimum should have higher or lower capital than the laissez-faire equilibrium is the factor income of the (consumption-)poor, since these agents have a high weight due to the incompleteness of consumption insurance. One can show that the comparison between their *relative* labor income and their *relative* asset income guides how they are affected: if the poorest do better relative to the average for labor income than for asset income, then they are helped with a *larger* aggregate capital stock. If instead the consumption-poor agents have an income composition that is relatively stronger for assets than for labor income, the reverse result holds. In conclusion, in economies where the unlucky consumers do relatively better in terms of their labor income, there is too little aggregate capital accumulation.

We illustrate these properties quantitatively using a numerical model solution, where transition dynamics is taken into account. We study three cases: three commonly used versions of the standard model, each with a standard calibration. The first case we look at—the "unemployment economy"—is based on unemployment shocks, as in the version of Krusell and Smith (1998) with homogeneous preferences. For this economy, which has limited wealth dispersion, we find a modest amount of capital overaccumulation, as in the simplest version of the two-period model we study.

The second and third calibrations feature underaccumulation of capital. One of them uses a calibration like that in Castañeda, Díaz-Giménez, and Ríos-Rull (1998). This setup has realistic wealth dispersion due to large and persistent, though strongly mean-reverting, labor income shocks. Here, the consumption-poor are mainly wealth-poor, and hence the planner should increase the capital stock. Moreover, the discrepancy between the laissez-faire equilibrium and the constrained optimum is large. In the third calibration, the calibration used in Aiyagari (1994), the constrained-efficient steady state is only asymptotic: it involves ever-increasing wealth inequality, i.e., the underaccumulation is mainly counteracted by making the rich save more and more.

Diamond (1967) first raised the possibility of constrained inefficiency in a one-period, one-good stock market economy with multiplicative uncertainty concerning production, but he actually found constrained efficiency for the particular economy under study. However, examples of constrained inefficiency of competitive equilibria were soon provided in Hart (1975), Diamond (1980), Stiglitz (1982), Loong and Zeckhauser (1982), Newbery and Stiglitz (1984), and Greenwald and Stiglitz (1986). In particular, Stiglitz (1982) established that the constrained efficiency result shown in Diamond (1967) depended on his one-good assumption; with more goods, a reallocation of investments and portfolios would in general influence relative prices and, given the incompleteness of markets, could also improve consumer welfare. Geanakoplos, Magill, Quinzii, and Dreze (1990) later established generic (in terms of initial endowments) constrained inefficiency of competitive equilibria of a two-period stock market economy with many goods.[1] In contrast with the general equilibrium literature, we address the constrained inefficiency issue in the infinite-horizon workhorse macroeconomic model with uninsurable idiosyncratic shocks.[2] In our context, generic constrained inefficiency also comes through relative prices, though across production inputs rather than across consumption goods. Our examples moreover demonstrate that the inefficiency can be quite drastic quantitatively (for plausible model calibrations) as well as, we think, intuitively easily understood.

Section 2 describes the two-period model and analyzes constrained efficiency in this economy. It describes two cases: one without initial wealth heterogeneity, i.e., in the first period all consumers are identical, and one with initial wealth heterogeneity. The second of these cases is our way of illustrating how multi-period models, where wealth heterogeneity would be endogenous, would allow not only capital overaccumulation but also capital underaccumulation. Section 3 describes the model with an infinite horizon and describes laissez-faire equilibria. The associated constrained-efficiency planning problem is then described in Section 4 and the central first-order condition is derived. Section 5 carries out

[1] In Geanakoplos and Polemarchakis (1986) this result is established for an exchange economy. In that case the result is generic in initial endowments and utility functions.

[2] For a study of uninsurable idiosyncratic shocks from an incomplete markets, general equilibrium perspective, see Carvajal and Polemarchakis (2011).

the quantitative analysis for our different example calibrations of the infinite-horizon model. Before concluding in Section 7, Section 6 looks at two relevant extensions: a model with a labor-leisure choice and a model with "expenditure shocks" that do not involve relative prices. We discuss implementation, through explicit tax-transfer schemes, of the constrained optimum throughout the paper: in our two-period model, in Section 4.2, and in the quantitative section.

2. THE MECHANISMS: ILLUSTRATION USING A TWO-PERIOD MODEL

In the present section, we consider a two-period general equilibrium model of precautionary savings—a two-period version of Aiyagari (1994)—and use it to introduce and illustrate the notion of constrained inefficiency. With the simple model we are able to establish some useful qualitative results and to discuss some features of the model that will later be of quantitative importance in our numerically computed economies. We begin with an economy where consumers are identical in the first period (Section 2.1) and then consider initial wealth inequality (Section 2.2). Initial wealth heterogeneity is relevant since it will arise endogenously in a multiperiod setting, where the nature of inequality—whether it primarily reflects inequality in capital or labor income—turns out to be crucial for the qualitative as well as quantitative results. Much later in the paper (Section 6), we then revisit the two-period model for a brief discussion of two relevant extensions.

2.1. *Ex ante identical consumers*

Consider an economy with a continuum (measure 1) of two-period-lived, ex ante identical consumers. The consumers have time-additive, von Neumann-Morgenstern utility functions with a twice continuously differentiable, strictly increasing, and strictly concave period utility function u satisfying INADA conditions and a discount factor β. In the first period, period 1, each agent is endowed with ω units of output, which can be either consumed, c, or invested in an asset, a: capital. In period 2, consumers receive income from the capital they saved in period 1 and from working. The labor income of any given individual is random. In particular, the labor endowment can be either high or low, and it is independent across agents. We denote the period 2 labor endowments e_1 and e_2, with $0 < e_1 < e_2$; the probability that any agent's labor endowment is e_1 is π. Due to the independence of shocks across consumers, a law of large numbers operates so that also the *fraction* of agents with e_i is π_i; we sometimes use π to denote π_1. That is, there is no uncertainty about the period 2 labor endowment: the supply of labor is constant at $L = \pi e_1 + (1 - \pi)e_2$.

In the second period, output comes from production using capital and labor and a constant returns to scale (CRS) neoclassical production function f. Since all agents face the same maximization problem, and since this is a problem with a strictly concave objective and a linear constraint set, they will all make identical choices. Let the implied equilibrium choice of capital be K (per consumer, and in the aggregate, so that $a = K$). Then the output in period 2 is known to be $f(K, L)$. Output is produced by perfectly competitive firms in our equilibrium: they sell the output to consumers and rent the capital and the labor services from the same consumers at rates r and w, respectively. In equilibrium, thus, r and w will be set to equal the marginal products of the inputs; in particular, they will be deterministic. This means that in period 1, each consumer will see his capital income in period 2 as deterministic and equal to rK, whereas his labor income is random and equal to we.

It is a maintained assumption in our analysis that consumers can only save using capital; in particular, there is no pure insurance instrument available for reducing the idiosyncratic risk, so the only way of influencing the risk is through "precautionary savings."

Given the above, we have

Definition 1. A *competitive equilibrium* is a vector (K, r, w) such that (i) K solves

$$\max_{a\in[-we_1/r,\omega]} u(\omega - a) + \beta\left(\pi u(ra + we_1) + (1-\pi)u(ra + we_2)\right)$$

and (ii) $r = f_k(K, L)$ and $w = f_l(K, L)$, with $L = \pi e_1 + (1-\pi)e_2$.

It is straightforward to show that an equilibrium with $K \in (0, \omega)$ exists under suitable conditions on u and f.[3]

Can the market allocation be improved upon? The notion of constrained efficiency. In this economy, agents really make only one choice: the saving choice. Following the incomplete markets, general equilibrium literature, we discuss the efficiency properties of the equilibrium in terms of whether this one choice could be made in a better way: can it be made so as to improve on equilibrium utility? Formally, we call the equilibrium *constrained efficient* if there is no level of saving $\hat{K}$ such that, given competitive pricing of inputs in period 2, the utility of the consumer is higher than under the competitive equilibrium. That is, the equilibrium (K, r, w) we consider is efficient if there is no $\hat{K} \in [0, \omega]$ such that

$$u(\omega-\hat{K})+\beta\left(\pi u(f_k(\hat{K},L)\hat{K} + f_l(\hat{K},L)e_1) + (1-\pi)u(f_k(\hat{K},L)\hat{K} + f_l(\hat{K},L)e_2)\right) >$$

$$u(\omega-K)+\beta\left(\pi u(f_k(K,L)K + f_l(K,L)e_1) + (1-\pi)u(f_k(K,L)K + f_l(K,L)e_2)\right).$$

The question, thus, is whether a fictitious planner can improve on the allocation by simply commanding a different savings level for the representative consumer, while respecting all budget constraints of agents and letting firms operate freely under perfect competition. In particular, the fictitious planner is not

[3]In this section, we do not restrict borrowing unnecessarily; the consumer is allowed to borrow any amount as long as the debt can be paid back in all states of the world. Given that the utility function satisfies the INADA condition, we omit these issues in the present section.

allowed to "complete the markets" or in any way transfer goods between lucky and unlucky consumers: the only insurance asset is still capital. Thus, at least qualitatively, constrained inefficiency is a rather drastic form of market failure.

The market outcome is constrained inefficient: the formal argument. In this economy, whether or not it is possible to improve on the market allocation can be seen by considering the impact of a small variation dK of the aggregate capital. Differentiating the indirect utility, one obtains

$$dU = -u'(\omega - K)dK + \beta\left(\pi u'(rK + we_1)dC_1 + (1-\pi)u'(rK + we_2)dC_2\right),$$

where $dC_i = rdK + Kdr + e_i dw$, $i \in \{1, 2\}$.

The individual's first-order condition for savings reads

$$u'(\omega - K) = \beta\left(\pi u'(rK + we_1) + (1-\pi)u'(rK + we_2)\right) r.$$

This condition can be used to simplify the above expression, and it will lead many of the effects of increasing capital to vanish. We thus obtain

$$dU = \beta\Big(\left(\pi u'(rK + we_1) + (1-\pi)u'(rK + we_2)\right) K dr \\ + \left(\pi u'(rK + we_1)e_1 + (1-\pi)u'(rK + we_2)e_2\right) dw\Big),$$

so that we see that any effect of a marginal change of savings away from the competitive equilibrium has to operate through its effect on factor prices. The cancelations, of course, are just a result of the envelope theorem. As for how factor prices are affected by capital, we note that $dr = f_{KK}(K, L)dK$ and $dw = f_{KL}(K, L)dK$ so that

$$\begin{aligned} dU &= \beta\Big(\pi u'(rK + we_1)(Kf_{KK}(K, L) + e_1 f_{KL}(K, L)) \\ &\quad +(1-\pi)u'(rK + we_2)(Kf_{KK}(K, L) + e_2 f_{KL}(K, L))\Big)dK. \end{aligned}$$

Now note that because f is homogeneous of degree 1, $Kf_{KK}(K, L) + Lf_{KL}(K, L) = 0$ and therefore

$$dU = \beta\left(\pi u'(rK + we_1)\left(1 - \frac{e_1}{L}\right) + (1-\pi)u'(rK + we_2)\left(1 - \frac{e_2}{L}\right)\right) f_{KK} K dK.$$

Letting

$$\chi \equiv \frac{u'(f_k(K, L)k + f_l(K, L)e_1)}{u'(f_k(K, L)k + f_l(K, L)_2)} > 1,$$

this can be rewritten as

$$dU = \beta u'(f_k(K, L)k + f_l(K, L)e_2)\pi(\chi - 1)\left(1 - \frac{1}{L}\right) f_{kk}(K, L)KdK.$$

Thus, since $f_{kk} < 0$, equilibrium utility increases if $dK < 0$. We conclude, more generally, that the equilibrium is *constrained inefficient*. As is clear from the analysis, the key assumptions behind the result are that u is strictly concave and that f has a strictly decreasing marginal product of capital.

Specifically, as noted, the level of capital in the laissez-faire equilibrium is *too high*: a higher utility is obtained if all consumers save a little less in period 1. The intuitive reason for the overaccumulation of capital is as follows. More capital savings raises wages and lowers rental rates. The only source of market failure in this economy is the incomplete insurance. A small decrease in K from the equilibrium level thus lowers w and raises r, thereby scaling down the part of the consumer's income that is stochastic and scaling up the part that is deterministic: the amount of risk the consumer is exposed to is now smaller. Given that there is no direct insurance for this risk, this amounts to an improvement. The "distortion" on the agents' savings by moving savings away from the competitive equilibrium level for given prices is of a second-order magnitude, and thus the manipulation of prices so as to lower the de facto risk dominates.

Market incompleteness is of course key to our finding of constrained inefficiency: unlike in the complete markets case, prices are not optimally set here and agents' influence on prices should therefore be taken into account when making individual choices. An improvement on the competitive outcome thus requires taking an aggregate, "planning" perspective.

2.2. *The two-period model with initial wealth heterogeneity*

We now let ω differ across individuals as the economy starts; let its distribution be Γ. We will discuss the effects of altering total capital accumulation away from the laissez-faire level on consumers with different initial wealth holdings and then point to how the two-period model with initial wealth heterogeneity can be used to organize some of the findings for the quantitative infinite-horizon economy studied in the remainder of the paper.

Equilibrium and effects of altering aggregate saving. For the new setting, we have

DEFINITION 2. A *competitive equilibrium* is a vector $(a(\omega), K, r, w)$ such that (i) $a(\omega)$ solves

$$\max_{a\in[0,\omega]} u(\omega - a) + \beta \left(\pi u(ra + we_1) + (1-\pi)u(ra + we_2)\right);$$

(ii) $K = \int_\omega a(\omega)\Gamma(d\omega)$; and (iii) $r = f_k(K, L)$ and $w = f_l(K, L)$, with $L = \pi e_1 + (1-\pi)e_2$.

Does a decrease in capital accumulation—engineered, say, by making all agents save $\epsilon > 0$ less—increase utility here as well? Again, the key question is how the resulting price change (an increase in r and a decrease in w) would affect consumers' utility. Here, unlike in the case where consumers are equal ex ante, this question does not have an unambiguous answer: different consumers have different income compositions, and in particular those who are poor initially will rely more on labor income and therefore can be made worse off by a fall in w. It is straightforward to show that $a(\omega)$ is increasing, i.e., that higher initial wealth translates to higher capital income in the second period. An agent with net wealth ω, and second-period capital $a(\omega)$, will experience an effect on utility, $\frac{dU(\omega)}{dK}$, equal to

$$\beta \sum_{i=1}^{2} \pi_i u'(f_k(K,L)a(\omega) + f_l(K,L)e_i) \left(\frac{a(\omega)}{K} - \frac{e_i}{L} \right) f_{kk}(K,L)K.$$

To sign this expression, rewrite it as

$$\beta u'(f_k(K,L)a(\omega) + f_l(K,L)e_2) \\ \left[\pi(\chi(\omega) - 1) \left(\frac{a(\omega)}{K} - \frac{e_1}{L} \right) + \frac{a(\omega)}{K} - 1 \right] f_{kk}(K,L)K,$$

where $\chi(\omega)$ is defined analogously with χ above, i.e., as the ratio of utilities in the bad and good earnings states, a number greater than one. We see that this formula generalizes the one with ex ante homogeneous agents; the old term from the homogeneous case is changed slightly—$1 - \frac{e_1}{L}$ is replaced by $\frac{a(\omega)}{K} - \frac{1}{L}$—and a term $\frac{a(\omega)}{K} - 1$ is added. The old term has the same sign as before, unless $a(\omega)$ is sufficiently below average saving K; the new term has the same sign as the old term if and only if $a(\omega) > K$. Thus, we conclude that agents who start out with sufficiently below average wealth would prefer total saving to increase, not decrease, whereas the remainder of the population prefer lower aggregate saving.

Constrained efficiency and connections to the infinite-horizon model The analysis above shows that only with a very tight distribution of initial wealth would there be unanimity for a decrease in K relative to the laissez-faire equilibrium. Thus, with sufficient dispersion in initial wealth, it would not be possible to find a Pareto improvement by altering aggregate saving. However, the main point of considering initial wealth inequality here is that it provides a useful link to the analysis of the infinite-horizon model studied in the sections to follow. There, initial (as of time 0) wealth is identical across agents, but as a result of uninsurable earnings shocks, wealth levels will diverge over time; in a laissez-faire steady state, there is a non trivial joint distribution over asset levels and employment status. Thus, in that setting, as in the model studied in the previous section, there is a natural planner objective, namely, ex ante expected utility—which will be equal for all agents, though realized utility of course differs across consumers. Since ex ante expected utility amounts to a probability-weighted average, it can be thought of as a utilitarian objective: the planner is "behind the veil of ignorance." This means that an ex post desire for redistribution from the consumption-rich to the consumption-poor reflects the ex ante insurance aim.

Now turning back to our two-period model, based on the previous discussion we can think of it as the last two periods of a long-horizon model, which then means that the appropriate planner objective is the utilitarian one. Thus, whether more or less aggregate saving is called for in the second-to-last period is more readily answered: we only need to sum the effects on welfare across all consumers. Thus, we obtain a net effect, labeled Δ to be used later,

$$\Delta \equiv \beta f_{kk}(K,L)K \int_{\omega} \sum_{i=1}^{2} \pi_i u'(f_k(K,L)a(\omega) + f_l(K,L)e_i) \left[\frac{a(\omega)}{K} - \frac{e_i}{L}\right] \Gamma(d\omega). \tag{1}$$

This is simply a marginal utility weighted average of the individual effects of an increase in capital, which amounts to placing a higher weight on the low-wealth consumers. Now consider the case when the earnings shocks are small relative to wealth heterogeneity; for illustration, set $e_1 = e_2 = L$. Then the expression becomes

$$\beta f_{kk}(K,L)K \int_{\omega} u'(f_k(K,L)a(\omega) + f_l(K,L)L) \left[\frac{a(\omega)}{K} - 1\right] \Gamma(d\omega),$$

which is unambiguously positive: $f_{kk} < 0$ and the integral is negative, since higher weights are placed on the above-average wealth consumers—those with $a(\omega) > K$. Intuitively, higher capital leads to a lower return to capital, which helps the poor and hurts the rich; the utilitarian utility then supports it. This mechanism lies behind some of our quantitative results in what follows: the earnings process there leads to such highly dispersed wealth that, in steady state, it is desirable with a higher capital stock than under laissez-faire.

Note, finally, that with serially correlated earnings shocks—the quantitatively relevant case—there will be a positive correlation between earnings and asset wealth. As a shortcut description of this, the two-period model could have a π that decreases in ω. An inclusion of this effect would decrease the first term within brackets above, by decreasing the weight on the high ω terms. Thus, in the two-period model, a positive asset-earnings correlation weakens the arguments in favor of a lower aggregate capital stock. In the infinite-horizon model, however, the serial correlation of shocks will also influence the shape of the wealth distribution; moreover, this effect is not necessarily monotone.[4] In Section 6 below we also use the two-period model to look at extensions to two other cases of interest: idiosyncratic shocks unrelated to capital or labor income and valued leisure.

[4] See, e.g., Krusell and Smith (1997).

2.3. *The constrained optimum and implementation*

On the basis of the above discussion, we define the constrained optimum by the solution to

$$\max_{a(\omega)} \int_{\omega} \left[u(\omega - a(\omega)) + \beta \sum_{i=1}^{2} \pi_i u(f_k(K,L)a(\omega) + f_l(K,L)e_i) \right] \Gamma(d\omega),$$

where $K = \int_{\omega} a(\omega)\Gamma(d\omega)$ and $L = \pi e_1 + (1-\pi)e_2$. Here, thus, the planner chooses a function $a(\omega)$, i.e., saving for all consumers, in order to maximize a utilitarian objective. Again, the utilitarian objective may seem unmotivated in the two-period model, but the idea, elaborated on above, is that this two-period model represents the last two periods of a longer-horizon problem of which at time zero all consumers were equal.[5]

Given our INADA conditions on utility, we obtain interior solutions and a typical first-order condition then reads

$$u'(c(\omega)) = \beta \left(\pi u'(c_1(\omega)) + (1-\pi)u'(c_2(\omega))\right) f_k(K,L) + \Delta,$$

where $c(\omega) \equiv \omega - a(\omega)$ and $c_i(\omega) \equiv f_k(K,L)a(\omega) + f_l(K,L)e_i$, $i = 1,2$, and Δ equals the expression in (1). Notice here that Δ does not depend on ω: the added benefit (if $\Delta > 0$), or cost (if $\Delta < 0$), of saving is the same for all consumers.

The constrained optimum can be interpreted as attained through a direct mandate for each consumer, issued by the planner, how to save. It can also be obtained through explicit tax incentives. The key feature of the tax system is to allow the first-order condition above to be satisfied for all consumers. This will be accomplished by a tax wedge on the saving decision, which we can define as proportional capital income, accompanied by a lump-sum transfer so that the net transfer to the agent is zero. The tax wedge will need to depend on individual histories, here represented by individual beginning of period wealth: $\tau(\omega)$ will appear in the Euler equation as

$$u'(c(\omega)) \equiv \beta \left(\pi u'(c_1(\omega)) + (1-\pi)u'(c_2(\omega))\right) f_k(K,L)(1-\tau(\omega)),$$

so that, now again rewriting using $\chi(\omega)$ to denote the ratio of marginal utilities across states 1 and 2 (a number greater than 1), we obtain by also imposing the constrained-optimal first-order condition that

$$\tau(\omega) = -\frac{f_{kk}(K,L)K}{f_k(K,L)} \int_{\tilde{\omega}} \frac{u'(c_2(\tilde{\omega}))}{u'(c_2(\omega))} \frac{\pi(\chi(\tilde{\omega}) - 1)\left(\frac{a(\tilde{\omega})}{K} - \frac{e_1}{L}\right) + \frac{a(\tilde{\omega})}{K} - 1}{\pi(\chi(\omega) - 1) + 1} \Gamma(d\tilde{\omega}). \tag{2}$$

[5] Formally, one could specify an ex ante stage where all consumers would be identical and a lottery mechanism that would deliver the ω distribution as an outcome. Such a model would formally justify the above objective but not add anything of essence.

The lump-sum transfer must equal $\tau(\omega)a(\omega)f_k(K,L)$.[6]

It is instructive to revisit our special cases here. First, if there is no wealth heterogeneity, we obtain

$$\tau = -\frac{f_{kk}(K,L)K}{f_k(K,L)} \frac{\pi(\chi-1)\left(1-\frac{e_1}{L}\right)}{\pi(\chi-1)+1} > 0,$$

and, second, if there is wealth heterogeneity but there are no income shocks,

$$\tau(\omega) = -\frac{f_{kk}(K,L)K}{f_k(K,L)} \int_\omega \frac{u'(c_2(\tilde{\omega}))}{u'(c_2(\omega))} \left(\frac{a(\tilde{\omega})}{K} - 1\right) \Gamma(d\tilde{\omega}) < 0.$$

Thus, the optimum in the first case is attained with a proportional tax on capital income. In the second case, the optimum demands a subsidy, which moreover is higher for consumers with high initial wealth. Richer consumers necessitate a higher subsidy (and lump-sum tax) because the societal payoff Δ from increasing saving beyond what would be in the self-interest of the consumer, as noted above, does not depend on this consumer's wealth, and thus it is larger for richer consumers, since their private marginal utility is lower. The second example illustrates that, in general, "simple" uniform (across types) policies do not suffice for achieving the constrained optimum.

3. THE INFINITE-HORIZON ECONOMY

We now study the infinite-horizon version of the above model. As above, we look at a continuum of agents subject to idiosyncratic shocks $e_i \in E$, where $E \equiv \{e_1, \cdots, e_i, \cdots, e_I\}$, that are i.i.d. across agents and that follow a Markov process with transition matrix $\pi_{i,j}$. Agents have standard preferences: an expected discounted sum of a strictly increasing and strictly concave utility function, i.e., $\mathbb{E}_0 \; \{\sum_t \; \beta^t \; u(c_t)\}$. Agents do not have access to state-contingent contracts but can only accumulate assets in the form of real capital; we denote it a. There is a lower bound on asset holding: $\underline{a}$.[7] As in the previous section, these assets are rented by competitive firms each period and used for production purposes according to a CRS neoclassical production function f that uses capital and efficient units of labor. Capital accumulation is assumed to follow a geometric structure: a fraction δ of the capital stock depreciates from one period to the next.[8]

The nature of the budget constraint that agents face is thus $c + a' = a(1 + r) + ew$, where we use primes to denote the next period's values. Individual

[6]Of course, the timing of the transfer is immaterial when there are no borrowing constraints; only the present value matters.

[7]This lower bound may arise from the existence of a solvency constraint that requires that agents are always able to pay back their debt or from an explicit borrowing constraint.

[8]Note that the assumptions in the previous section can be thought of as assuming 100% depreciation, or that alternatively f was defined to include undepreciated capital.

agents are indexed by the pair $\{e, a\}$ that describes their labor endowment and wealth, respectively. The state of the economy can be summarized by probability measure x over the Borel sets of compact set $S = E \times A$. In this context, aggregate amounts of factors of production and their rental prices are $K = \int_S a dx$, $L = \int_S e dx$, $r = f_K(K, L) - \delta$, and $w = f_L(K, L)$, respectively. We use the notation $r(x)$ and $w(x)$, though on occasion we use $r(K)$ and $w(K)$, since L is constant due to the law of large numbers.

In this economy the aggregate state variable is the distribution of agents over labor earnings and wealth, x, which agents have to know in order to compute prices.[9] We write $x' = H(x)$ to describe the law of motion of the distribution. Then, the agent's problem is

$$(3) \quad v(x, e, a) \quad = \quad \max_{\substack{c \geq 0 \\ a' \in A}} u(c) + \beta \sum_{e'} \pi_{e,e'} \quad v(x', e', a') \qquad \text{s.t.}$$

$$(4) \qquad c + a' = a\,[1 + r\,(x)] + e\,w\,(x) \quad \text{and} \quad x' = H(x),$$

with solution $a' = h(x, e, a)$. An important feature of this problem is the requirement that the agent's assets lie in compact set A.

We now turn to the construction of an aggregate law of motion of the economy. Using decision rule h and transition matrix π, we construct an individual transition process. Let $B \in S$ be a Borel set. Define Q by

$$(5) \qquad Q(x, e, a, B; h) = \sum_{e' \in B_e} \pi_{ee'} \ \chi_{h(x,e,a) \in B_a},$$

where χ is the indicator function. It is easy to see that Q is indeed a transition function. We now define the updating operator $T(x, Q)$ that yields tomorrow's distribution given today's:

$$(6) \qquad x'(B) \ = \ T(x, Q)(B) = \int_S Q(x, e, a, B; h) \ dx.$$

An equilibrium requires that agents' expectations are correct. Formally,

Definition 3. A *recursive competitive equilibrium* is a pair of functions h and H such that h solves problem (3) given H and that $H(x) = T(x, Q(.; h))$.

A *steady state* for this economy is a distribution $\widetilde{x}$ such that $\widetilde{x} = T(\widetilde{x}, Q)$. Steady states have the property that the interest rate is lower than the rate of time preference, or that the aggregate capital stock is higher than that of an economy with perfect markets or no shocks (for a discussion and a proof of this result, see Huggett (1997)). The interpretation of this result is one of *precautionary saving*: savings play dual roles here, by allowing not just intertemporal smoothing but also some (limited) amount of smoothing across states.

[9] While prices today can be known just from today's aggregate capital, future prices cannot be known from today's aggregate capital because decision rules are not linear. Hence, the distribution is the appropriate state variable. See Krusell and Smith (1997), Krusell and Smith (1998), or Ríos-Rull (1998) for more elaborate discussions.

3.1. *Characterization: first-order conditions and steady-state capital*

The consumer's first-order condition for savings is the recursive functional equation

$$u'\left(x,e,a,h(x,e,a)\right) \geq \beta[1+r(H(x))] \sum_{e'} \pi_{e,e'} u'\left(H(x),e',h(x,e,a),h[H(x),e',h(x,e,a)]\right),$$

with equality if $a' > \underline{a}$, which can be rewritten compactly as $u'(c) \geq \beta[1 + r(H(x))]\sum_{e'} \pi_{e,e'} u'(c)$.

Steady states can be readily found as a fixed point of an aggregate steady-state capital demand function, which depends on the interest rate which in turn is given by the marginal productivity of capital (see below). Let $h^m(e,a;r)$ be the decision rule implied by a constant interest rate r (and associated wage w). It solves

$$u'\left[a(1+r)+ew-h^m(e,a;r)\right] \geq \beta \sum_{e'} \pi_{e,e'} u'\left[h^m(e,a;r)\,(1+r)+ew-h^m[e',h^m(e,a;r);r]\right], \tag{7}$$

with equality if $h^m > \underline{a}$.

Let the stationary aggregate capital implied by $h^m(.,.;r)$ be $K(r)$, a continuous function of r (Ríos-Rull (1998)). A steady state is therefore characterized by a $\bar{K}$ and a rate of return $\bar{r}$ such that given $\bar{r}$, aggregate capital is $\bar{K}$, i.e., $\bar{K} = K(\bar{r})$, and $\bar{r}$ is the marginal productivity of capital implied by $\bar{K}$, i.e., $\bar{r} = f_k(\bar{K}, L)$.[10]

4. CONSTRAINED-OPTIMAL ALLOCATIONS IN THE INFINITE-HORIZON ECONOMY

Our quantitative focus will be on the characterization of constrained-optimal steady states (Section 4.3 below), i.e., long-run outcomes of a well-defined infinite-horizon planning problem. The initial condition in this planning problem is one of complete equality across consumers—the first period of our two-period model above. This initial condition implies that the objective function of the planner here is "utilitarian," i.e., weighs all consumers' utils equally. One could, of course, consider arbitrary weighting schemes; our choice of equal weights is motivated by our focus on insurance, as opposed to redistribution. Also as before, the planner is constrained to only consider allocations with zero net transfers across consumers.

[10]In a steady state, where the discount rate exceeds the interest rate, there is an upper bound to the assets that agents hold (see Huggett (1997)).

Our main characterization result—the first-order condition below—is derived using a variational approach and is thus based on a sequential formulation of the constrained-efficiency problem. However, for descriptive purposes, we use a recursive formulation of this planning problem in the main text. It reads

$$\Omega(x) \quad = \quad \max_{y(e,a)\in A} \int_S u\left[a(1+\ r\left(K\right)) + e\ w\left(K\right) - y(e,a)\right] \ dx \ + \ \beta\,\Omega(x') \tag{8}$$

$$\text{s.t.} \qquad x' = T(x, Q(.;y)), \quad K = \int_S a dx. \tag{9}$$

We use the function h^* to denote the implied decision rule for $y(e,a)$ at x: $a' = h^*(x,e,a)$.

4.1. *The first-order condition*

For convenience, we will now assume that the distribution x admits a density.[11] Our key analytical characterization in this paper is the first-order condition for the planner:

PROPOSITION 1 *If the distribution x admits a density, the first-order necessary condition of problem (8) can be stated as the following functional equation in the decision rule h^*: for all $\{e,a\} \in S$,*

$$\begin{aligned} u'\left(a\ [1 + r\left(K\right)] + ew\left(K\right) - h^*(x,e,a)\right) \geq \\ \beta\left[1 + r\left(K'\right)\right] \sum_{e'} \pi_{e,e'} u'\left(h^*(x,e,a)\ [1 + r\left(K'\right)] + \right. \\ \left. e'w\left(K'\right) - h^*(x',e',h^*(x,e,a)]\right) + \beta f_{KK}(K',L)K' \\ \int_S u'\left(a'\ [1 + r\left(K'\right)] + e'w\left(K'\right) - h^*(x',e',a')\right)\left[\frac{a'}{K'} - \frac{e'}{L}\right] dx', \end{aligned} \tag{10}$$

where the inequality becomes equality if $h^*(x,e,a) > \underline{a}$.

PROOF: See Appendix A. *Q.E.D.*

For later reference, we will define Y as the implied law of motion of the distribution: $x' = Y(x) = T(x, Q(.;h^*))$. Omitting some arguments, we can write the first-order condition compactly as

$$u'(c) \geq \ \beta\ (1 + r') \ \sum_{e'} \pi_{e,e'} \ \ u'(c') \ \ + \ \Delta, \tag{11}$$

with $\Delta = \beta f'_{KK} K' \int_S u'(c') \left[\frac{a'}{K'} - \frac{e'}{L}\right] dx'$. This equation is the guide for individual savings at different values for (e,a). It is the dynamic generalization of

[11] That our characterization is possible to carry out without this assumption was kindly pointed out to us by Iván Werning.

the formula from our two-period model: equation (1) in particular displays the two-period version of the extra term, Δ, in the Euler equation. The Δ in the two-period case and the one here are only different in the following two ways: (i) the probability weighting here (implicit in the dx' expression) involves more values (if earnings risk is serially correlated), since the probabilities of different outcomes next period depend on the earnings states last period; and (ii) the joint distribution over assets and earnings, over which the integral is defined, is determined more nontrivially: it depends on the entire history of shocks.

Clearly, the sign of the extra term, and its magnitude, depend on the distribution of agents across assets and earnings, in addition to primitives such as the curvature of the utility function. To evaluate these, we focus on a steady-state version of this equation and thereafter use a calibrated version of it in order to compare how different the long-run outcomes are between the laissez-faire allocation and the constrained optimum.

4.2. *Implementation of the constrained optimum*

The constrained planning allocation can be implemented with tax-transfer schemes. In the two-period model, sufficient instruments involve savings wedges, along with lump-sum transfers, that depend on initial wealth, as shown in equation (2). In models with more periods, the same kind of instrument can be used, the only difference being that the savings wedges (and the lump-sum transfers) now need to be history dependent.

To this aim, consider the first-order condition for the constrained optimum, i.e., condition (10). Using e^t to denote the complete history of productivity realizations for an agent, we can define the tax wedge $\tau(e^t)$, a tax on the gross return from saving decided on in period t. Using the consumer's Euler equation for this decision then reads

$$u'\left(c^*(x,e,a)\right) = \beta\left[1 + r\left(K'\right)\right]\left[1 - \tau(e^t)\right]\sum_{e'}\pi_{e,e'}u'\left(c^*(x',e',h^*(x,e,a))\right), \tag{12}$$

where we have used c^* to denote the consumption function associated with the savings function h^*.[12] This equation awkwardly mixes recursive and sequential notation, but it does so for a reason. Using (10), this equation can be used to solve out directly for taxes, state by state and period by period, as a function of optimal consumption-savings plans, as determined by the constrained planning problem (where (10) is a key condition). Thus, (12) shows that the optimal tax wedge will inherit a recursive structure. In particular, one can use this equation to see that the tax rate will depend on (x,e,a): this quantity summarizes all the necessary information in e^t.[13] In the quantitative section below, looking at

[12] For cases where (10) holds with inequality there is, of course, a range of tax rates that could be used.

[13] Of course, using this dependence explicitly in the formulation of the consumer's problem would not be possible without somehow making clear that these tax rates are given, and not possible to influence by the saving choice.

steady states, we will illustrate numerically how the tax wedges computed by the above formula depend on e and a.

4.3. *The constrained-efficient steady state*

A steady state for the planner is a decision rule $\bar{h}^*$ and an associated distribution $\bar{x}^*$ such that $\bar{x}^* = Y(\bar{x}^*)$ and $\bar{h}^*(e,a) = h^*(\bar{x}^*, e, a)$. With $\bar{K}^* \equiv \int_S a \, d\bar{x}^*$, a steady state satisfies

$$\begin{aligned}(13) \qquad & u'\left(a\left[1+r\left(\bar{K}^*\right)\right] + ew\left(\bar{K}^*\right) - \bar{h}^*\left(e,a\right)\right) \geq \beta[1+r(\bar{K}^*)] \\ & \sum_{e'} \pi_{e,e'} u'\left(\bar{h}^*(e,a)\left[1+r\left(\bar{K}^*\right)\right] + e'w\left(\bar{K}^*\right) - \bar{h}^*(e', \bar{h}^*(e,a))\right) + \Delta,\end{aligned}$$

where

$$\Delta = K f_{kk} \int_S u'\left(a'\left[1+r\left(\bar{K}^*\right)\right] + e'w\left(K^*\right) - \bar{h}^*(e',a')\right)\left[\frac{a'}{K^*} - \frac{e'}{L}\right] d\bar{x}^*.$$

This is a functional equation that can be solved using standard numerical methods. Note in this context that if Δ is positive, there is no guarantee that there exists an upper bound to individual asset holdings. However, an upper bound can be imposed, and in the numerical simulations one would then verify whether or not it is violated.

The object of study in our steady-state analysis is a "modified Golden Rule," i.e., a long-run outcome that is optimal from the perspective of taking discounting into account. In other words, our constrained-optimal steady state is *not* derived simply from maximizing steady-state utility without regard to initial conditions and the costs or benefits of reaching that steady state. Instead, it answers the question: if the allocation that is ex ante constrained optimal has the property that there is convergence to a steady state, what are the properties of the implied steady-state distribution? Next, we answer this question quantitatively, and we also show what kind of tax system would implement the allocation.

5. QUANTITATIVE ANALYSIS BASED ON CALIBRATIONS TO U.S. DATA

In an economy calibrated to actual data, is capital accumulation too high or too low from a constrained-efficiency perspective? In this section we show that the answer is sensitive to the calibration entertained, and we explain the intuitive reasons for these differences. We consider three calibrations, all of which are "standard" calibrations in the literature, emphasizing different kinds of wage-income risk. The first one is an economy where labor income is a two-state process, interpretable as employment/unemployment. This economy turns out to have the property of the simplest of the two-period models in Section 2: it has too much capital. The second economy emphasizes asymmetric wage risk of the sort that generates a realistic (i.e., very high) wealth dispersion. In this economy, capital accumulation is much too low. In the third economy—an economy with individual wage risk calibrated to be symmetric, as in Aiyagari (1994), thus generating much less wealth dispersion—capital accumulation is also too low. However, this economy has a new feature: the constrained-efficient steady state is only asymptotic, i.e., there is ever-increasing wealth dispersion. Thus, standard calibrations can give rise to a constrained-efficient policy that implies (i) that capital accumulation should be counteracted or promoted and (ii) that wealth inequality should be contained or made to grow without bound.

We display our results in Sections 5.1–5.3. We then explain and compare these outcomes in Section 5.4 in terms of the properties of Δ, the key ingredient in the constrained planner's first-order condition and other features of the different economies.

5.1. *The market economy has too much capital: the unemployment economy*

In this economy, preferences are of the constant relative risk aversion (CRRA) form, $\sum_t \beta^t \frac{c_t^{1-\sigma}-1}{1-\sigma}$, with the period set to be one year. Production occurs through a standard neoclassical production function $F(K_t, L_t) = K_t^{\theta} L_t^{1-\theta}$. Calibration of the above parameters and the rate of capital depreciation is rather standard: the steady state of the laissez-faire economy is targeted to a real interest rate of 4%, a capital-output ratio of slightly below 3, and a labor share of 0.64, which is accomplished assuming that $\beta = 0.887$, $\delta = 0.08$, and $\theta = 0.36$, and the intertemporal elasticity of substitution is set to 0.5.

In the unemployment economy, the idiosyncratic shock captures unemployment risk rather than wage risk: it can take only two values, with a very low value of unemployment, as in Krusell and Smith (1998) or Castañeda, Díaz-Giménez, and Ríos-Rull (1998). This calibration interprets all the earnings volatility and dispersion as coming from shocks. We set the unemployment rate to 5% and the average duration of unemployment to as high as 2.6 years, thus targeting "long-term unemployment," which exhibits long duration and is also viewed as highly cyclical. By construction, our calibration thus makes unemployment a relatively severe shock, indeed more severe in terms of income losses than what it appears to be for the average unemployed. The parameters involved are $e = [0.01 \;\; 1.00]$ for the labor endowment (unemployed, employed) and $\Pi_{1,\cdot} = [0.62 \;\; 0.38]$ and $\Pi_{2,\cdot} = [0.02 \;\; 0.98]$ for the transition matrix. Table I displays the results.

As can be seen in Table I, the constrained optimum requires a *lower* long-run level of capital than what is generated in the laissez-faire economy. Thus, the prescription here is to move toward the first-best level of capital. Note also that the implied effect of constrained-optimal policy on inequality is minor. This supports the notion that the key determinant is the factor composition of the income of the poor. In this model economy, the poor are unemployed and their labor income is essentially zero. This makes the model economy de facto capital intensive.

5.2. *The market economy has much too little capital: high wealth dispersion*

It is not immediate how to generate equilibrium wealth dispersion of a magnitude similar to that in the data. The literature offers several possibilities; here, we follow the approach in Diaz, Pijoan-Mas, and Ríos-Rull (2003), which is a simplified version of that in Castañeda, Díaz-Giménez, and Ríos-Rull (2003). The general idea is (i) to include very large wage outcomes and that (ii) from the very high wage outcomes there is a significant risk of a large fall in the wage. The combination of these features makes the highest earners have a significant demand for precautionary saving.[14] Thus, a properly chosen three-state Markov chain allows us to generate a laissez-faire steady state with inequality measures for earnings and wealth quite close to those in the U.S. data. Let the three levels of earnings be $\{1.00, 5.29, 46.55\}$; let there be no direct transitions between the highest and lowest state; let the probability of staying be 0.992, 0.98, and 0.917 in the low, middle, and high states, respectively, and let the transition probability from middle to low be 0.009. Then the stationary distribution is $\pi^\star = \{0.498, 0.443, 0.059\}$. This process has an earnings Gini index of 0.60, there is strong persistence, and there is a nontrivial risk of dropping from the highest state to the middle state. The results for the economy calibrated in the manner just described are shown in Table II.

The first notable finding is that the market economy has large precautionary savings in steady state: aggregate wealth is 2.33 times larger than in the economy without shocks. As a result of the additional capital, output is 35.5% higher. We also see that the wealth Gini index of the market economy is quite large, 0.861, slightly larger than the 0.803 of the U.S. economy whereas the share of wealth held by the richest 5% is 54.55 while in the data it is 57.80%.

Second, the constrained optimum requires an even higher level of capital in the long run: it is as much as 8.5 times that in the deterministic economy and even 3.65 times that in the laissez-faire allocation. These differences are very large. The wealth distribution in the market and the constrained-optimal allocations are very similar. This fact ultimately derives from the constraint on the planner, which is precisely that no direct redistribution be made across agents: all redistributions occur through price changes.

To summarize, the constrained-efficient steady state of the model economy has much more capital than does the laissez-faire steady state (which itself has much more capital than does the first-best steady state, due to precautionary savings). At the same time, the distribution of wealth as measured by the shares owned by the various groups is very similar. Thus, the laissez-faire precautionary savings economy is associated with *much too little* capital: the pecuniary externality from savings in this economy is not only positive but also large.

[14] Wealth heterogeneity can also, for example, result from discount factor heterogeneity, as modeled in Krusell and Smith (1998). Such an assumption embodies the idea that "the wealth-poor are poor because they chose to become poor"; in the present calibration, the poorest consumers have simply been unlucky with their earnings realizations.

For this economy, we also illustrate the tax implementation discussed in Section 4.2. Figure 1 plots different subsidy wedges (net-of-tax returns), all positive here since there is capital underaccumulation, for different individuals. Luckier consumers, as measured either by wealth or by current labor productivity, obtain larger wedges, and the wedges are over 14 percentage points for the luckiest consumers. The pre-subsidy real interest rate, which is indicated by the dotted line in Figure 1, is negative here—a little below -2%—and some of the least lucky consumers apparently indeed do not obtain much more even after the subsidy.

5.3. *Ever-increasing wealth dispersion: the original Aiyagari (1994) calibration*

Using an AR(1) process in the logarithm of labor income with normally distributed (symmetric) shocks, Aiyagari (1994) selects the persistence and volatility parameters based on Kydland (1984), who uses the Panel Study of Income Dynamics (PSID), and on Abowd and Card (1987) and Abowd and Card (1989), who use both PSID and National Longitudinal Surveys (NLS) data. Aiyagari then approximates this process using a seven-state Markov chain following the procedures described in Tauchen (1986). We follow the same procedure, although we reduce the Markov chain to three states. We take our benchmark to have an autocorrelation of 0.6 and a coefficient of variation of 0.2.[15] We specify the other parameters so that the economy with complete markets (the standard neoclassical growth model) satisfies standard properties. The interest rate is set to 4.167%: β is 0.96. Our only departure from Aiyagari (1994) is to set the intertemporal elasticity of substitution, $\frac{1}{\sigma}$, to be equal to 0.5.[16] The capital share is equal to 0.36 and the capital-output ratio is set to slightly under 3.[17]

The asymptotic steady state of the market outcome in the Aiyagari economy has 2.03% more assets and 0.70% more assets than its full insurance counterpart, with an interest rate of 4.011% instead of 4.167%. The implied coefficient of variation of wealth is 0.718 and its Gini index of wealth is 0.388, far below the dispersion observed in the data.

The interesting feature of this economy, however, is that the constrained-efficient allocation has ever-increasing wealth inequality. For this economy, first, it is possible to verify numerically that no exact steady state exists.[18] Second,

---

[15] The specific parameters for the three-state process are given by $e \in \{e_1, e_2, e_3\} = \{0.78, 1.00, 1.27\}$, $\pi_{e'|e_1} = \{0.66, 0.27, 0.07\}$, $\pi_{e'|e_1} = \{0.28, 0.44, 0.28\}$, and $\pi_{e'|e_1} = \{0.07, 0.27, 0.66\}$. The resulting stationary distribution is $\pi^\star = \{0.337, 0.326, 0.337\}$.

[16] Aiyagari (1994) considers the values 1, 0.33, and 0.2. Though we do not report these results, we have found that our Aiyagari economy does not change its qualitative features if one changed the intertemporal elasticity of substitution to 1 or 3.

[17] It is set to 2.959 so that the depreciation rate of capital, δ, equals 0.08.

[18] As for the previous economies, a steady state is computed by (i) guessing on a value for aggregate capital and on a value for Δ, (ii) solving the constrained-efficient first-order condition, (iii) simulating an individual's history long enough that the distribution across assets and earnings is obtained, and (iv) computing the implied aggregate capital stock and Δ. To verify that a steady state does not exist, one can search exhaustively for values of capital and Δ.

to solve for a constrained-efficient outcome, i.e., for the transition toward an asymptotic steady state from some given initial distribution of assets and earnings, we use techniques similar to those developed in Krusell and Smith (1997) and Krusell and Smith (1998); for details, see the Appendix B. For simplicity, we consider as the initial condition the steady-state income-wealth distribution of the market allocation. We see from Figure 2, which describes how different moments of the distribution of capital evolve, that the constrained-efficient outcome features increasing wealth concentration over time, with otherwise well-behaved aggregates.[19]

5.4. *Comparisons*

The previous sections illustrate that it is nontrivial to predict the nature of the inefficiency in this class of models. The key object is the nature of Δ: the average of $a/K - e/L$ in the population, weighted by marginal utilities. The marginal utilities are declining in consumption, thus giving higher weight to the consumption unlucky. The effect Δ is fundamentally endogenous, and it turns out that it is negative for the unemployment economy but positive for the two other economies. Hence, given that a higher Δ in the constrained-efficient Euler equation means a stronger incentive to save, the unemployment economy has too much and the other economies too little capital. Why is Δ smaller in the unemployment economy? The determinants of Δ were discussed in Section 2; the sign of $a/K - e/L$ for the agents with the lowest consumption is particularly important, since those agents receive the highest weight. In the unemployment economy, the wealth distribution is rather tight and the poorest agents are wage poor more than wealth poor—$a/K - e/L$ is positive for them, and $\Delta < 0$.

In the economy with a realistic wealth distribution, $a/K - e/L$ is negative for the consumption poorest, since this economy has a very skewed wealth distribution (the wage dispersion is high but is less skewed). Moreover, due to the highly unequal wealth distribution, Δ is quantitatively large, leading to a large discrepancy between the constrained optimum and the laissez-faire outcome. As we saw in Table II, capital is in fact so high in the constrained optimum that the real interest rate is negative. In contrast, the Aiyagari (1994) economy, which also has $\Delta > 0$, has only a modest discrepancy between the savings in the constrained optimum and the laissez-faire; in particular, the constrained-efficient real interest rate here is positive. Now compare these two economies from the perspective of the "lucky" high-asset consumers; how can they be induced to save more, so as to raise aggregate saving beyond the laissez-faire level? In an economy with a negative real interest rate, assets must not be kept too high, since then there is a net loss to the agent in the period budget constraint. In fact, assets must clearly be bounded in our calibrated economy with large wealth inequality—they cannot exceed $-w/r$ times the highest individual productivity level, which is fi-

[19] Appendix B illustrates transition paths for an economy with a steady state.

nite. Thus, the long-run wealth distribution will necessarily keep a finite support. When the interest rate is positive, however, one can see how the support may need to be infinite, and wealth dispersion may need to expand indefinitely. The Euler equation of the consumer has an additional term on the right-hand side, a term that induces more asset accumulation. That term will be more and more important relative to consumption the higher is consumption, and hence an element of "increasing returns to asset accumulation" has been introduced on the level of the individual. It follows that marginal propensities to save can become increasing in wealth. Whether this effect is strong or not depends on Δ and on the interest rate. A higher interest rate will make $a_{t+1} - a_t = we_t - c_t + ra_t$ greater; it will also be increasing in asset holdings, unless the marginal propensity to consume out of asset holdings is above the interest rate, an undesirable feature from the perspective of consumption insurance. Thus, ever-increasing wealth for the luckiest is an outcome that is consistent with achieving constrained efficiency.

The economy that delivers a more realistic wealth distribution—with high dispersion—features saving rates that are, relative to the other economies studied, higher for the rich and lower for the poor. Overall, however, in steady state saving remains too low. This raises two questions. One is how the higher, constrained-optimum level of capital is raised—by making all the consumers save more or by differentiated saving responses among the rich and the poor? Interestingly, the answer is that the constrained optimum dictates that the rich are made to save even more but that the poor are made to lower their saving. This result is apparent from comparing the steady-state decision rules in laissez-faire to those in the constrained optimum and can be understood from looking at the Euler equations of the poor and the rich. Here, the added marginal utility benefit from saving, Δ, has a larger effect on the rich, as the other terms of their Euler equation are small, i.e., their marginal utilities are low. Wealth dispersion, however, barely changes because even though the rich save more, the returns to being lucky fall considerably with the higher capital stock. The second question that arises as a result is whether the underlying reason for the dispersion in saving rates among the rich and the poor matters for our undersaving result. It may indeed matter, and we discuss this issue briefly in the concluding section of the paper.

6. EXTENSIONS

In this section, we very briefly look at two extensions of relevance. The first of these is one with valued leisure, a setting commonly used in the recent literature; the second looks at other kinds of shocks. The two-period model is used for illustration, and we abstract from initial wealth heterogeneity.

Suppose consumers have period utility $u(c, 1-l)$, where l is labor supplied to the market. For simplicity, let labor supply be a choice variable only in the second period; in the first period, it is fixed at some exogenous value $\bar{l}$. This economy thus incorporates another "insurance channel" through the hours choice: in response to the income shock, the consumer can choose to work more.[20] An equilibrium here is defined as follows:

DEFINITION 4. A *competitive equilibrium* is a vector (K, r, w, L_1, L_2) such that (i) (K, L_1, L_2) solves

$$\max_{a,l_1,l_2} u(\omega - a, 1-\bar{l}) + \beta\left(\pi u(ra + we_1 l_1, 1 - l_1) + (1-\pi)u(ra + we_2 l_2, 1 - l_2)\right)$$

and (ii) $r = f_k(K, L)$ and $w = f_l(K, L)$, with $L = \pi e_1 L_1 + (1-\pi)e_2 L_2$.

Similarly, along the above lines, we can derive the first-order condition for the constrained optimum. For capital, we obtain

$$u'(c^*, 1-\bar{l}) = \beta\left(\pi u'(c_1^*, 1 - l_1^*) + (1-\pi)u'(c_2^*, 1 - l_2^*)\right) + \Delta^k, \tag{14}$$

where

$$\Delta^k = \beta u'(c_2^*, 1 - l_2^*)\pi(\chi - 1)\left(1 - \frac{e_1 l_1^*}{L}\right) f_{kk}(k^*, l^*)k^*,$$

with $c^* = \omega - k^*$, $c_i = f_k(k^*, l^*)k^* + f_l(k^*, l^*)e_i$, $i \in \{1, 2\}$, and $\chi = u'(c_1^*, 1 - l_1^*)/u'(c_2^*, 1 - l_2^*)$. For labor, we obtain, for $i \in \{1, 2\}$,

$$u_{1-l}(c_i^*, 1 - l_i^*) = u'(c_i^*, 1 - l_i^*) f_l(k^*, l^*)e_i + \Delta^l, \tag{15}$$

with

$$\Delta^l = u'(c_2^*, 1 - l_2^*)\pi(\chi - 1)\left(\frac{e_1 l_1^*}{L} - 1\right) f_{ll}(k^*, l^*)l^*.$$

Here, clearly, Δ^k and Δ^l represent deviations from the laissez-faire first-order conditions. Somewhat more stringent conditions are needed to sign these terms than in the case without valued leisure, but standard calibrations would imply (i) that $u'(c_1^*, 1 - l_1^*) > u'(c_2^*, 1 - l_2^*)$ (recall that $e_2 > e_1$), so that $\chi > 1$, and (ii) that $e_1 l_1 < e_2 l_2$. Thus, we would have, as before, that $\Delta^k < 0$, whereas $\Delta^l > 0$: in the absence of wealth heterogeneity, capital is still overaccumulated in equilibrium, and moreover consumers work too little. Additional work effort would lower wages, which would reduce the risky part of consumers' income. For specific parametric examples, with a constant elasticity of substitution between c and $1 - l$ or the case where there are no wealth effects on labor supply, it is easy to show that these properties are met.

Turning to the second and last extension, consider shocks that do not influence either labor or capital income. First, consider a "health expenditure shock" that appears as a negative income shock: second-period income is now $ra + w - e$, where e is either high or low. In the absence of other shocks, the laissez-faire equilibrium

[20] For an analysis of this kind of mechanism, see, e.g., Domeij and Floden (2006), Pijoan-Mas (2006), or Heathcote, Storesletten, and Violante (2009).

is now constrained efficient. A change in savings away from equilibrium would change r and w as before, but $\Delta = 0$ would result, since the sum of capital and labor income would not be influenced by the price changes on the margin. With more than two periods, however, health expenditure shocks would lead to a constrained-inefficient equilibrium: now savings would be too low, since health shocks do generate capital income dispersion in periods after the second period, and the risk implied by this dispersion would be counteracted by higher aggregate saving (through a reduction in r). Second, consider multiplicative taste shocks: period 2 utility is now $u(c)e$, where e is again either high or low. As for health expenditure shocks, the laissez-faire equilibrium would be constrained efficient in the two-period model: price changes induced by altered aggregate savings would not influence consumers' income on the margin. In longer-horizon models, there would be non trivial effects, since the taste shocks—which can now also be interpreted as shocks to discounting—would induce differences in saving. Thus, changes in aggregate saving, again through prices, will have different impacts on different consumers and will be a valuable instrument for the planner. It would be interesting to study this case, as well as the interaction of shocks, in a long-horizon model in much more detail.

7. CONCLUSION

In this paper, we have argued that the laissez-faire competitive equilibrium of the one-sector neoclassical growth model with uninsurable idiosyncratic wage shocks, which plays a prominent role in the recent macroeconomic literature, may offer significant scope for welfare improvements. Obviously, welfare improvements could be achieved by simply adding insurance markets—or, equivalently, by using tax-transfer schemes that effectively distribute from the lucky to the unlucky—but what we argue here is that improvements can be made without altering the market structure, and without forcing any transfers between consumers. That is, we argue that the equilibrium is *constrained inefficient*, in the sense of Diamond (1967): if consumers merely departed somewhat from their individual optimization, equilibrium prices would be altered and everybody's welfare would rise. The reason is that this model has a "pecuniary externality": with incomplete markets, the price mechanism does not fully work. In particular, the uninsurable risk agents face can be made more or less hurtful by altering prices.

In the standard setting we look at, the risk appears through prices. The direct risk in the model is wage risk, but the insurability of this risk also induces wealth differences, and the propagation of this wealth inequality critically depends on the interest rate. Therefore, whether it is beneficial to raise wages or interest rates depends on the details of the model calibration. We derive the constrained planner's first-order condition, and it reveals whether savings should be increased or decreased relative to laissez-faire. The key here is whether the consumption-poor consumers—who define the "unlucky" in this economy—have lower labor earnings or asset holdings relative to the average consumer. We illustrate with three standard calibrations from the literature and show that these calibrations give rise to very different qualitative as well as quantitative results. For the calibration that arguably best fits the data on inequality, the constrained optimum calls for a rather large increase in overall savings relative to the laissez-faire outcome.

It is important to relate our study to existing tax policy analyses using similar models. Given that our focus is on individual uninsurable risks, there is no direct connection to representative agent treatments such as the classic zero-capital tax result in Chamley (1986). The most closely related studies are arguably Aiyagari (1995) and Aiyagari and McGrattan (1998). Aiyagari (1995) argues for a tax on capital using a model where consumer utility also depends on endogenously chosen government expenditures. Since an Euler equation has to hold for government expenditures, and there is no aggregate uncertainty, the government has no precautionary savings motive, and hence it is optimal to tax savings so as to bring the economy to the first-best level of aggregate capital: the interest rate has to equal the time discount rate. Aiyagari's tax policy, unlike ours, also involves net transfers across consumers. the study of Aiyagari and McGrattan (1998) is also different from our study in that it does not insist on zero net transfers across consumers; moreover, it has government expenditures, though modeled to be a constant percentage of output. Debt has a variety of effects; among others, it lowers the capital stock, and it requires financing, which under the lump-sum tax scheme considered by the authors hurts the poor disproportionately. Overall, the authors find debt to have a small effect on welfare compared to actual U.S. debt policy. Other related studies include those on unemployment insurance (see, e.g., the recent overview in Mukoyama (2010)), progressive taxation, social security, and so on. In all these studies, the precise restrictions placed on government policy are crucial and dictate what constitutes optimal policy. Our restriction—that no net transfers be made across consumers and only the existing market structure be used—is different in that it focuses more narrowly on the question of what "pecuniary externalitities" operate in the missing-markets economy. In any applied studies of the effect of tax-transfer changes, one of the channels is the one we focus sharply and uniquely on here, though other channels will typically be present in those studies as well.

Our findings raise several questions worthy of further research. One regards the source of idiosyncratic risks. The standard model under study here looks at wage/unemployment shocks, but consumers are subject to other risks as well. In Section 6 we briefly consider "expenditure shocks," which do not operate through prices, and find constrained efficiency. Relatedly, it would be valuable to go beyond the consumption-saving choice and look at other decisions such as the labor-leisure choice and the education choice, both of which might be associated with constrained inefficiency. We look at the labor-leisure choice in the context of the two-period model (in Section 6) and find that consumers "underwork": taking the effect of wages on the de facto wage risk into account, everybody's utility would increase with slightly higher work effort than in a laissez-faire equilibrium.

Our analysis here just scratches the surface of an interesting problem; a more ambitious analysis of this issue is under way by Athreya, Tam, and Young. Another interesting possibility is that idiosyncratic, uninsurable income may lead to constrained inefficiency even in the absence of capital accumulation. A very recent paper, Farinha Luz and Werquin (2011), shows that the Huggett model (Huggett (1993), which does not have capital accumulation) can be constrained inefficient too: utility can go up for all agents if a stricter borrowing constraint is "self-imposed."

An equally important issue concerns the calibration. Our calibration where wealth dispersion is as large as in the data follows Castañeda, Díaz-Giménez, and Ríos-Rull (2003), which relies on risks of large (but infrequent) hikes and drops in earnings. There are alternative ways of explaining the wide dispersion in wealth that we observe in most economies, and it would be interesting to consider constrained efficiency in such settings as well. One involves preference heterogeneity, where wealth inequality derives chiefly from persistent shocks to patience. A complementary assumption is that poor agents receive larger transfers, such as those implied by unemployment insurance or food stamp programs, leading them to save less.[21] A model might give results that are quite different from those we obtain here, since they suggest that wealth inequality is not all a result of incomplete risk sharing. In other words, the planner would be more willing to let those who choose to become poor (rich) stay poor (rich).

Similarly, other forms of heterogeneity in preferences (such as risk aversion) or in individuals' abilities or opportunities (e.g., possibly making it harder for some to participate in asset markets than it is for others) would also be valuable to examine from the perspective of constrained efficiency. The hope is that ultimately, microeconomic studies allow us to better distinguish which elements of individual heterogeneity are key and which are not.

Finally, it is interesting to examine constrained efficiency in contexts where there are explicit reasons for market incompleteness, such as private information. This feature has several applications already, two of which are particularly noteworthy. One is the work of Golosov and Tsyvinski (2007), which finds, as we do, that there can be over- or underaccumulation of capital relative to the constrained-efficient outcome depending on the details of the calibration (in their case, the nature of the unobservable skill process). In contrast, in the context of the Diamond-Dybvig model of banking, Farhi, Golosov, and Tsyvinski (2009) find interest rates in the "shadow market"—where depositors can trade among them without being observed by banks—to always be too high (this model, however, does not have capital). Doubtlessly, many interesting applications are yet to appear in this new literature.

[21]These two assumptions are key in Krusell and Smith (1997) and Krusell and Smith (1998).

REFERENCES

Abowd, J. M., and D. Card (1987): "Intertemporal Labor Supply and Long Term Employment Contracts," *American Economic Review*, 77, 50–68.

——— (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57, 411–445.

Aiyagari, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, 109, 659–684.

——— (1995): "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting," *Journal of Political Economy*, 103, 1158–1175.

Aiyagari, S. R., and E. R. McGrattan (1998): "The optimum quantity of debt," *Journal of Monetary Economics*, 42, 447–469.

Bewley, T. (1986): "Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers," in *Contributions to Mathematical Economics in Honor of Gérard Debreu*, ed. by W. Hildenbrand, and A. Mas-Colell. North Holland, Amsterdam.

Carvajal, A., and H. M. Polemarchakis (2011): "Idiosyncratic Risk and Financial Policy," *Journal of Economic Theory*, 146, 1569–1597.

Castañeda, A., J. Díaz-Giménez, and J.-V. Ríos-Rull (1998): "Exploring the Income Distribution Business Cycle Dynamics," *Journal of Monetary Economics*, 42, 90–130.

Castañeda, A., J. Díaz-Giménez, and J. V. Ríos-Rull (2003): "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 111, 818–857.

Chamley, C. (1986): "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, 54, 607–622.

Diamond, P. (1980): "Efficiency with Uncertain Supply," *The Review of Economic Studies*, 47, 645–651.

Diamond, P. A. (1967): "The Role of a Stock Market in a General Equilibrium Model with Technological Uncertainty," *American Economic Review*, 57, 759–776.

Diaz, A., J. Pijoan-Mas, and J.-V. Ríos-Rull (2003): "Precautionary Savings and Wealth Distribution under Habit Formation Preferences," *Journal of Monetary Economics*, 50, 1257–1291.

Domeij, D., and M. Floden (2006): "The Labor-Supply Elasticity and Borrowing Constraints: Why Estimates Are Biased," *Review of Economic Dynamics*, 9, 242–262.

Farhi, E., M. Golosov, and A. Tsyvinski (2009): "A Theory of Liquidity and Regulation of Financial Intermediation," *Review of Economic Studies*, 76, 973–992.

Farinha Luz, V., and N. Werquin (2011): "Constrained Inefficiency in a Huggett Model," Working Paper, Yale University.

Geanakoplos, J., M. Magill, M. Quinzii, and J. Dreze (1990): "Generic Inefficiency of Stock Market Equilibrium When Markets Are Incomplete," *Journal of Mathematical Economics*, 19, 113–151.

Geanakoplos, J., and H. M. Polemarchakis (1986): "Existence, Regularity, and Constrained Suboptimality of Competitive Allocations when the Asset Market is Incomplete," in *Essays in Honor of Kenneth Arrow* Vol. 3 *Uncertainty, Information, and Communication*, ed. by W. Heller, R. Starr, and D. Starrett, pp. 65–95. Cambridge: Cambridge University Press.

Golosov, M., and A. Tsyvinski (2007): "Optimal Taxation with Endogenous Insurance Markets," *Quarterly Journal of Economics*, 122, 487–534.

Greenwald, B. C., and J. E. Stiglitz (1986): "Externalities in Economies with Imperfect Information and Incomplete Markets," *Quarterly Journal of Economics*, 101, 229–264.

Hart, O. D. (1975): "On the Optimality of Equilibrium When the Market Structure is Incomplete," *Journal of Economic Theory*, 11, 418–443.

Heathcote, J., K. Storesletten, and G. L. Violante (2009): "Consumption and Labor Supply with Partial Insurance: An Analytical Framework," Working Paper 15257, National Bureau of Economic Research.

Huggett, M. (1993): "The Risk-Free Rate in Heterogeneous-Agent, Incomplete-Insurance Economies," *Journal of Economic Dynamics and Control*, 17, 953–970.

——— (1997): "The One-Sector Growth Model with Idiosyncratic Shocks: Steady States and Dynamics," *Journal of Monetary Economics*, 39, 385–403.

Imrohoroğlu, A. (1989): "Cost of Business Cycles with Indivisibilities and Liquidity Constraints," *Journal of Political Economy*, 97, 1364–1383.

Krusell, P., and A. A. Smith (1997): "Income and Wealth Heterogeneity, Portfolio Choice, and Equilibrium Asset Returns," *Macroeconomic Dynamics*, 1, 387–422.

——— (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867–896.

Kydland, F. E. (1984): "Labor-Force Heterogeneity and the Business Cycle," *Carnegie-Rochester Conference Series on Public Policy*, 21, 173–209.

Loong, L. H., and R. Zeckhauser (1982): "Pecuniary Externalities Do Matter When Contingent Claims Markets Are Incomplete," *Quarterly Journal of Economics*, 97, 171–179.

Mukoyama, T. (2010): "Understanding the Welfare Effects of Unemployment Insurance Policy in General Equilibrium," Working Paper, University of Virginia.

Newbery, D. M. G., and J. E. Stiglitz (1984): "Pareto Inferior Trade," *Review of Economic Studies*, 51, 1–12.

Pijoan-Mas, J. (2006): "Precautionary Savings or Working Longer Hours?," *Review of Economic Dynamics*, 9, 326–352

Ríos-Rull, J.-V. (1998): "Computation of Equilibria in Heterogeneous-Agent Models," in *Computational Methods for the Study of Dynamic Economies*, ed. by R. Marimon, and A. Scott, chap. 9. New York: Oxford University Press.

Stiglitz, J. E. (1982): "The Inefficiency of the Stock Market Equilibrium," *Review of Economic Studies*, 49, 241–261.

Tauchen, G. (1986): "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions," *Economics Letters*, 20, 177–181.

TABLE I

Steady states for the unemployment economy

	First Best	Market Outcome	Constrained Optimum
Aggregate assets	2.959	3.359	3.279
Output	1.000	1.047	1.038
Capital-output ratio	2.959	3.209	3.160
Interest rate	4.167%	3.219%	3.392%
Coeff. of variation of wealth	0.0	0.203	0.200
Gini index for wealth	0.0	0.108	0.105

TABLE II

Steady states for the high wealth dispersion economy

	First Best	Market Outcome	Constrained Optimum
Aggregate assets	1.736	4.017	15.668
Output	1.000	1.353	2.208
Capital-output ratio	1.736	2.970	7.096
Real interest rate	12.740%	4.123%	-2.927%
Coeff. of variation of wealth	0.0	2.562	2.501
Gini index for wealth	0.0	0.861	0.864
Percentage of wealth of the top 5%	0.0	54.55	51.26

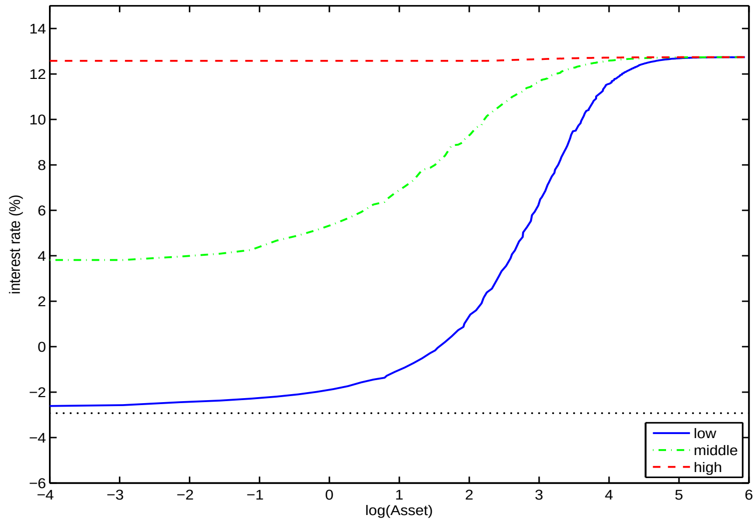

FIGURE 1.— Net-of-tax returns to savings as a function of e and a.

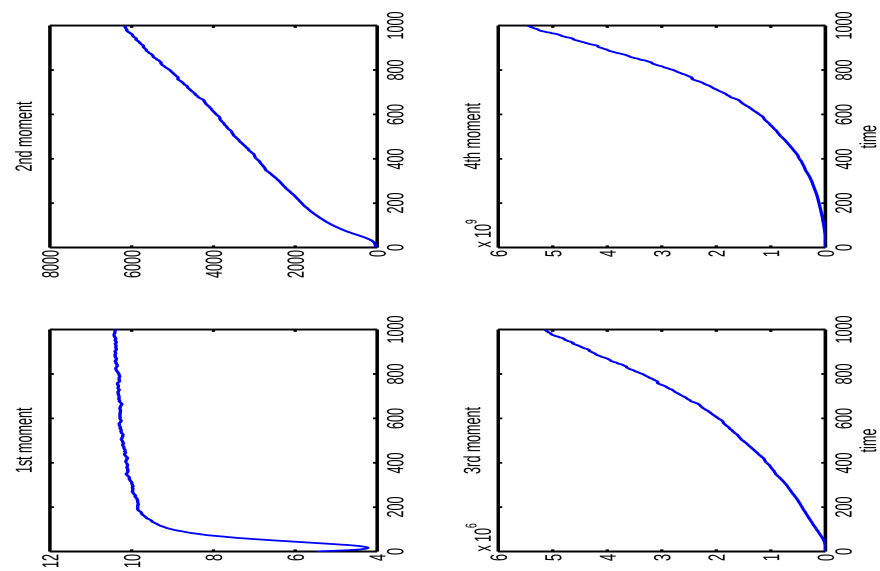

FIGURE 2.— Transition from the market steady state to the constrained optimum.

APPENDIX A: PROOFS

PROPOSITION 1 *If the distribution x admits a density, the first-order condition of the solution to problem (8) is, for all $(a,e) \in S$,*

$$
\begin{aligned}
(16)\qquad & u'\left(a\,[1+r(x)] + e\,w(x) - h^*(x,e,a)\right) \geq \\
& \beta\,\left[1+r'(x')\right] \sum_{e'} \pi_{ee'}\, u'\left(h^*(x,e,a)\,\left[1+r'(x')\right] + e'\,w'(x') - a''\right) \\
& + \beta \sum_{e'} \int_S \pi_{\widetilde{e}e'}\, u'\left(h^*(x,\widetilde{e},\widetilde{a})\,\left[1+r'(x')\right] + e'\,w'(x') - a''\right) \\
& \qquad \left[h^*(x,\widetilde{a},\widetilde{a})\; f_{KK}(K'(x')) + e'\,f_{LK}(K'(x'))\right]\, dx(\widetilde{a},\widetilde{e}),
\end{aligned}
$$

where we use a'' as an abbreviation for $h^*(x',e',h^*(x,e,a))$, and where $x' = Y(x) \equiv T(x,y)$, and again the inequality becomes equality if $h^*(x,e,a) > \underline{a}$.

PROOF: The sequence of policy rules h_t by which, given a distribution of savings and labor at period t with density x_t, a planner would instruct an agent endowed with labor e and capital a to save $h_t(x_t,e,a)$ must solve

$$
\begin{aligned}
&\max_{h_t} \sum_t \beta^{t-1} \sum_e \int u\left(c_t\right) x_t(a,e)da \\
&\text{s.t. } c_t + h_t(x_t,e,a) = a[1+f_K(K(x_t))] + e f_L(K(x_t))
\end{aligned}
$$

given x_1, where

$$K(x_t) = \sum_e \int a x_t(a,e)da$$

is the aggregate capital at period t. The updating operator for the sequence of distribution densities x_t of savings a and labor e is

$$x'(a',e') = \sum_e \pi_{ee'} \frac{x((h^*)^{-1}(x,e,a'),e)}{\frac{d}{da}h^*(x,e,(h^*)^{-1}(x,e,a'))}$$

at any period t.[22] Therefore, the planner's optimal policy rule h^* that instructs an agent with e and a today to save the amount $h^*(x,e,a)$, given a distribution of labor and savings today with density x and the savings a'' tomorrow, must maximize

$$
\begin{aligned}
&\sum_e \int u\Big(a[1+f_K(K(x))] + e f_L(K(x)) - a'\Big) x(a,e)da \\
&\qquad + \beta \sum_{e'} \int u\Big(a'[1+f_K(K(x'))] + e' f_L(K(x')) - a''\Big) x'(a',e')da',
\end{aligned}
$$

[22] Here, $(h^*)^{-1}(x,e,a')$ denotes value at a' of the inverse of $h^*(x,e,\cdot)$, for given e and x. In effect, note that assuming the policy rule $h^*(x,e,\cdot)$ is increasing for all e, then the mass of agents endowed with labor e' and capital $\tilde{a}'$ in any given interval $(0,a')$ next period is

$$\int_0^{a'} x'(\tilde{a}',e')d\tilde{a}' = \sum_e \pi_{ee'} \int_0^{(h^*)^{-1}(x,e,a')} x(a,e)da$$

from which the updating operator follows taking the derivative on both sides with respect to a', the upper limit of the integral on the left-hand side.

with

$$x'(a',e') = \sum_e \pi_{ee'} \frac{x((h^*)^{-1}(x,e,a'))}{\frac{d}{da}h^*(x,e,(h^*)^{-1}(a',e,x))},$$

with respect to a'. Merging the sums over e and the integration with respect to a, rearranging terms, and making the necessary changes of variables, h^* should thus maximize

$$\sum_e \int \Big[u\Big(a[1+f_K(K(x))] + ef_L(K(x)) - h^*(x,e,a)\Big) \\ + \beta \sum_{e'} \pi_{ee'} u\Big(h^*(x,e,a)[1+f_K(K'(x'))] + e'f_L(K'(x')) - a''\Big)\Big] \cdot x(a,e)da,$$

where, using the same changes of variables,

$$K'(x') = \sum_{e'} \int a' \sum_e \pi_{ee'} \frac{x((h^*)^{-1}(x,e,a'),e)}{\frac{d}{da}h^*(x,e,(h^*)^{-1}(a',e,x))}\, da' = \sum_e \int h^*(x,e,a)x(a,e)\, da.$$

For any variation $\delta_{h^*_{e_0}}$ of the optimal policy rule $h^*(x,e_0,\cdot)$ that determines the savings of the agents endowed with a given level of labor e_0 and given a distribution x, and for any $\varepsilon \neq 0$, the policy rule

$$h^\varepsilon(x,e,a) = h^*(x,e,a) + \varepsilon \chi_{e=e_0} \delta_{h^*_{e_0}}(a),$$

where $\chi_{e=e_0}$ is the indicator function for $e = e_0$, should be suboptimal. Therefore, the derivative with respect to ε at 0 of

$$\psi(\varepsilon) = \sum_e \int \Big[u\Big(a[1+f_K(K(x))] + ef_L(K(x)) - y^\varepsilon(a,e,x)\Big) \\ + \beta \sum_{e'} \pi_{ee'} u\Big(h^\varepsilon(x,e,a)[1+f_K(K'(T(x,h^\varepsilon)))] + e'f_L(K'(T(x,h^\varepsilon))) - a''\Big)\Big] \cdot x(a,e)da$$

must be 0. That is to say,

$$\frac{d}{d\varepsilon}\psi(0) = \int \Big[-u'\Big(a[1+f_K(K(x))] + e_0 f_L(K(x)) - h^*(x,e_0,a)\Big)\delta_{h^*_{e_0}}(a) \\ + \beta \sum_{e'} \pi_{e_0 e'} u'\Big(h^*(x,e_0,a)[1+f_K(K'(x'))] + e'f_L(K'(x')) - a''\Big) \\ \cdot \delta_{h^*_{e_0}}(a)[1+f_K(K'(x'))]\Big] x(a,e_0)da \\ + \beta \sum_{ee'} \pi_{ee'} \int \Big[u'\Big(h^*(x,e,a)[1+f_K(K'(x'))] + e'f_L(K'(x')) - a''\Big) \\ \cdot [h^*(x,e,a)f_{KK}(K'(x')) + e'f_{LK}(K'(x'))] \int \delta_{h^*_{e_0}}(\tilde{a})x(\tilde{a},e_0)d\tilde{a}\Big] x(a,e)da = 0.$$

Since this must hold for any variation, it must hold in particular for $\chi_{a \geq a_0}$, the indicator function of $a \geq a_0$, for any a_0. Therefore, rearranging terms and substituting $\chi_{a \geq a_0}$ for $\delta_{h^*_{e_0}}$,

$$\int_{a_0}^{+\infty} \Big[- u'\Big(a[1+f_K(K(x))] + e_0 f_L(K(x)) - h^*(x,e_0,a)\Big)$$
$$+ \beta \sum_{e'} \pi_{e_0 e'} u'\Big(h^*(x,e_0,a)[1+f_K(K'(x'))] + e' f_L(K'(x')) - a''\Big)$$
$$\cdot [1 + f_K(K'(x'))]\Big] x(a,e_0) da$$
$$+ \beta \sum_{ee'} \pi_{ee'} \int \Big[u'\Big(h^*(x,e,a)[1+f_K(K'(x'))] + e' f_L(K'(x')) - a''\Big)$$
$$\cdot [h^*(x,e,a) f_{KK}(K'(x')) + e' f_{LK}(K'(x'))] \int_{a_0}^{+\infty} x(\tilde{a},e_0) d\tilde{a}\Big] x(a,e) da = 0.$$

The right-hand side of the last expression is therefore a constant function of a_0 equal to 0. As a consequence, its derivative with respect to a_0 must be 0, i.e., for all a_0 and all e_0,

$$- u'\Big(a_0[1+f_K(K(x))] + e_0 f_L(K(x)) - h^*(x,e_0,a_0)\Big)$$
$$+ \beta \sum_{e'} \pi_{e_0 e'} u'\Big(h^*(x,e_0,a_0)[1+f_K(K'(x'))] + e' f_L(K'(x')) - a''\Big) \cdot [1 + f_K(K'(x'))]$$
$$+ \beta \sum_{ee'} \pi_{ee'} \int u'\Big(h^*(x,e,a) f_K(K'(x')) + e' f_L(K'(x')) - a''\Big)$$
$$\cdot [h^*(x,e,a) f_{KK}(K'(x')) + e' f_{LK}(K'(x'))] x(a,e) da = 0,$$

i.e., for all (a,e)

$$u'\Big(a[1+r(x)] + e w(x) - h^*(x,e,a)\Big) =$$
$$\beta [1 + r'(x')] \sum_{e'} \pi_{ee'} u'\Big(h^*(x,e,a)[1+r'(x')] + e' w'(x') - a''\Big)$$
$$+ \beta \sum_{\tilde{e}e'} \pi_{\tilde{e}e'} \int u'\Big(h^*(x,\tilde{e},\tilde{a})[1+r'(x')] + e' w'(x') - a''\Big)$$
$$\cdot [h^*(x,\tilde{e},\tilde{a}) f_{KK}(K'(x')) + e' f_{LK}(K'(x'))] x(\tilde{a},\tilde{e}) d\tilde{a}.$$

Q.E.D. *Q.E.D.*

APPENDIX B: APPROXIMATION METHODS

We now briefly describe how we solve the problem of the planner outside the steady state. Given $V(m)$ for $m \in I\!R^{n_m}$ and x, we can solve the following problem:

$$\max_{y(e,a)} \int u(a[1 + r(K)] + e\, w(K) - y(e,a))\, dx + \beta \;\; V(m') \tag{17}$$

$$\text{subject to} \qquad m' = \varphi(x,y). \tag{18}$$

We select a set of moments $m \in I\!R^{n_m}$ of the distribution x; the moments are defined by $m^i = \int_S a^i \, dx$, and typically we simply choose the first three moments. We then approximate the function V as a quadratic function of the logarithms of the chosen moments of the distribution, with some coefficient restrictions. A typical functional form that we have chosen is

$$V(m) = \alpha_0 + \alpha_1 \log(m^1) + \alpha_2 \log(m^2) + \alpha_3 \log(m^3) + \alpha_4 \log^2(m^1).$$

To solve the planning problem and to obtain the coefficients α, we proceed as follows:

1. Make an initial guess on V, labeled V^0. That is, guess an $\alpha^0 \in I\!R^{n_\alpha}$.
2. Choose an initial distribution x_0 and calculate its moments m_0. Then generate a sample for 10,000 individuals.
3. Construct a time series of distributions, x_{t+1}, decision rules, y_t, moments, m_{t+1}, and current returns, R_t, with the following iterative procedure for T periods (we choose $T = 3,000$). Given x_t and its moments m_t,
   (a) solve problem (17) to obtain decision rules y_t:
      i. choose the grid points for asset holdings (here it is important to assign more points on the lower asset range to better approximate decisions of agents with lower asset holdings),
      ii. with problem (17) taking the form

$$\max_{y_t(e,a)} \int_{S_t} u\left(a[1+r_t(m_t^1)] + e\, w_t(m_t^1) - y_t(e,a)\right)\, dx_t + \beta\, V^0(m_{t+1})$$

$$\text{subject to} \qquad m_{t+1}^i = \int_{S_t} y_t^i\, dx_t,$$

      iii. and find $y_t(e,a)$ satisfying the associated first-order condition for all $\{e,a\} \in S$, i.e.,

$$u'\left(a[1+r_t(m_t^1)] + e\, w_t(m_t^1) - y_t(e,a)\right) \geq \beta \sum_{i=1}^{n_m} V_i^0(m_{t+1})\, i\, [y_t(e,a)]^{i-1};$$

   (b) update the distribution and calculate its moments x_{t+1} and m_{t+1}:

$$\begin{aligned} x_{t+1} &= T(x_t, Q(., y_t)) \\ m_{t+1}^i &= \int_{S_t} y_t^i\, dx_t; \end{aligned}$$

   (c) and, finally, use the obtained distribution and decision to compute current returns R_t using

$$R_t = \int_{S_t} u\left[a(1+\ r\,(x_t)) + e\, w\,(x_t) - y_t(e,a)\right]\ dx_t,$$

   going back to 3a until $t = T$.
4. Use current returns and moments as well as V^0 to calculate a new set of values for all t, using

$$V_t' \equiv R_t + \beta V^0(m_t).$$

5. Define $V^1(m)$ by running a regression of the set of values $\{V_t'\}$ on the moments $\{m_t\}$ to obtain the new coefficients α^1.
6. Compare V^0 with V^1. If these functions are not sufficiently similar, update V^0 using V^1 and go back to 2; otherwise, stop.[23]

We follow this procedure for various combinations of the first three moments of x. All of these imply ever-increasing inequality, with the first moment remaining stationary. Table III shows time series data for aggregate assets and for the coefficient of variation of assets over time, for various specifications of function V. The top panel illustrates that the first moment settles down and becomes stationary after a few hundred periods, whereas the bottom panel shows that inequality keeps increasing over time.

[23]The accuracy of the obtained solution can be judged by the errors in the regression, once a fixed point in coefficients is found.

TABLE III

TIME SERIES: SIMULATED DATA FOR THE AGGREGATE ECONOMY

time	0	200	400	600	800	1000	1200	1400	1600	1800	2000
						Aggregate Asset					
A	5.59	9.59	9.97	10.08	10.13	10.18	10.24	10.27	10.30	10.31	10.28
B	5.59	9.71	10.06	10.16	10.22	10.26	10.31	10.34	10.37	10.39	10.35
C	5.59	9.07	9.72	9.95	10.04	10.11	10.20	10.25	10.29	10.32	10.30
D	5.59	6.77	9.28	9.76	9.97	10.10	10.21	10.26	10.31	10.33	10.31
						Coeff. of Variation					
A	0.72	4.12	5.10	5.74	6.33	6.88	7.36	7.87	8.37	8.87	9.39
B	0.72	4.86	6.26	7.37	8.35	9.25	10.13	10.96	11.70	12.35	12.93
C	0.72	3.27	4.39	5.20	5.91	6.66	7.40	8.12	8.77	9.31	9.84
D	0.72	2.10	3.75	4.85	5.79	6.70	7.51	8.19	8.81	9.39	10.00

Specifications for V:

A : $\alpha_0 + \alpha_1 \log(m^1) + \alpha_2 \log(m^2) + \alpha_3 \log(m^3) + \alpha_4 \log^2(m^1)$

B : $\alpha_0 + \alpha_1 \log(m^1) + \alpha_2 \log(m^2) + \alpha_3 \log^2(m^1)$

C : $\alpha_0 + \alpha_1 \log(m^1) + \alpha_2 \log(m^2) + \alpha_3 \log^2(m^1) + \alpha_4 \log^2(m^2)$

D : $\alpha_0 + \alpha_1 \log(m^1) + \alpha_2 \log(m^2) + \alpha_3 \log^2(m^1) + \alpha_4 \log(m^1) \log(m^2)$

We also applied this procedure to the economies that admit steady states, and here we do find convergence to the steady states of the planner economy: all four moments settle down relatively quickly, starting from an initial condition given by the steady state of the laissez-faire economy. Figure 3 illustrates.

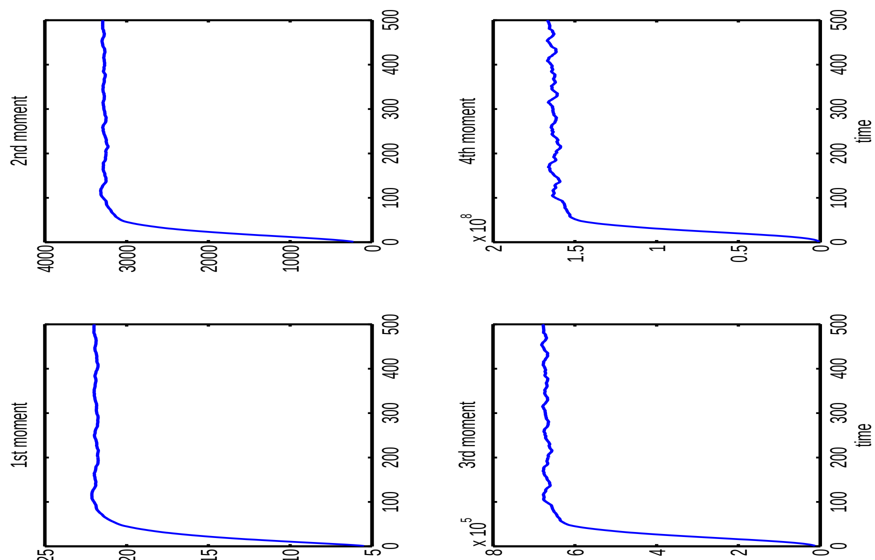

FIGURE 3.— Transition from the market-economy steady state to the constrained optimum: high wealth dispersion model.