Labor Supply in the Future:
Who Will Work?

Timo Boppart* Per Krusell† Jonna Olsson‡

September 2016

Full-length paper

Abstract

Evidence on hours worked per capita from historical time series as well as from a large cross-section of countries at different income levels strongly suggests that the income effect of increased labor productivity exceeds the substitution effect. To the extent that productivity keeps growing in the future, we should then expect people to want to work less and less. Given that labor-force participation is associated with natural indivisibilities, our perspective must furthermore imply that a smaller and smaller fraction of the population will be working. So who will then withdraw from the labor force, and when? This paper examines these questions in a frictionless setting where consumers/workers at any point in time differ in wealth and in productivity. It also addresses the normative issue: who should (given some welfare weights) work in the future?

*timo.boppart@iies.su.se, Institute for International Economic Studies, and CEPR

†per.krusell@iies.su.se, Institute for International Economic Studies, NBER, and CEPR

‡jonna.olsson@iies.su.se, Institute for International Economic Studies

1 Introduction

Some recent studies have reminded us of well-known patterns in hours worked. Bick, Fuchs-Schuendeln, and Lagakos (2016) show a striking pattern across countries at different degrees of development: in less productive/poorer countries, people work longer hours. Given that we can reasonably regard these observations as taken from institutional environments that, at least on average, have evolved to their present positions to reflect technology and needs, Bick et al. take these data to suggest that there is a fundamental feature of human preferences across consumption and leisure that makes us prefer to work less as we become more productive, everything else equal. If your wage rises, you have an incentive to work harder to take advantage of the higher wage, but on the other hand you are also richer with unchanged work hours and this suggests to take some time off. In this tradeoff, hence, Bick et al. argue that the latter effect—the income effect—dominates the former—the substitution effect. In our own past work (a subset of us: Boppart and Krusell (2016)) we come to the same conclusion, but now based on long time-series data, i.e., going back in history, from a smaller set of countries: as a given country becomes richer, its population works less and less. On an annual scale, the effect is barely visible but across long time periods and across large distances in space, it is impossible to ignore.

The fact that in our own, developed countries, our forefathers used to work harder has been in our awareness at least since Keynes's famous passage (see Keynes, 1930). Keynes envisioned a 15-hour work week for his grandchildren, as a result of continuously rising productivity and the choices they would make. So Keynes was on the mark regarding rising productivity but quantitatively off on hours, perhaps because he didn't actually do the back-of-the-envelope calculation. However, most countries have steadily moved hours worked per capita downward since then and although 15-hour workweeks still seem a bit distant, the trend is very clear, as pointed out above. The present paper is about a prediction for the future that follows by necessity, if past patterns are reproduced. The thought experiment is thus that productivity will continue to rise and the question is: what will happen to people's future labor supplies under the assumption that their consumption-leisure preferences are stable? The focus, however, is less on total hours than on the distribution of labor-market participation across the population. The reason is that we believe that labor-supply participation is associated with a fixed cost, or indivisibility: maybe it pays off to work 15 hours per week but not 1.5, and likely not 5 either. So our main question is: who will work in the future, given steadily increasing labor productivity? Will it be the highest earners who keep working, or will it be the lowest earners? Will it be the rich or the poor? And when do we expect labor-force withdrawal to occur for different groups, given current projections about future productivity growth?

We carry out our predictions using the preference class in Boppart and Krusell (2016), which is designed to be consistent with balanced-growth outcomes where hours change at a constant rate. In the data analyzed in that paper, hours declining at a constant rate is also argued to be good approximation conditional on constant productivity growth. This preference class is broader than the KPR class, i.e., that in King, Plosser, and Rebelo (1988), which is standard in the macroeconomic literature and which implies constant hours; under such preferences, the main question we ask here has a trivial answer since we would then expect hours to not change in the future, no matter what happens to productivity.

Our main focus is thus on the "Who?". Inequality in our economies both in income and in wealth have more and more become central features of macroeconomic analysis, though the inequality in hours choices have so far received less attention; an exception is Heathcote, Storesletten, and Violante (2014), and there are of course some others such as Erosa, Fuster, and Kambourov (2014). On one level the hours choice is about inequality in another consumption good, namely leisure, a good that certainly most people value greatly. Our modeling in this paper does not, perhaps, do justice to the current state of macroeconomic modeling with consumer heterogeneity, because it is based on frictionless markets. This, however, is a first and, we believe necessary, step toward a fuller analysis. We thus look at an entirely neoclassical growth model with labor-augmenting technological change where individual consumers face indivisibilities in individual hours choices: it is possible to work 0 hours or any amount above $\underline{h} > 0$ (subject, of course, also to an upper bound). We thus do not explicitly consider fixed costs, an assumption we make for the sake of tractability. Given our assumptions, a natural progression of working hours over time for a given household is thus to work fewer and fewer hours as wages grow and, eventually, to withdraw from the labor force. Thus, we are looking at a model with a mix of an intensive and an extensive margin. In our quantitative experiments, we calibrate our parameters so as to reproduce the historical patterns of declining hours supply per capita and then simply project forward. Households differ in productivity and wealth and these two features, and their distribution across the population, are the factors that influence the different paths for hours chosen for different people in our economy.

We present a list of theoretical results first. Some of these are basic some less obvious, but we cover many bases partly because labor supply under income effects exceeding substitution effects have not received much attention in the literature, especially in the case where the extensive margin is active—our focus here. We thus spend time working out results in static models, connecting planning solutions to equilibrium outcomes, discuss the role of convexification, and generally the effects of overall productivity growth on the patterns of working. A particularly striking result is that highly productive households (in relative terms) will withdraw sooner from the labor force, as will high-wealth households, where as in the typically used normative benchmark—that implied by a utilitarian social welfare function—high-skilled individuals withdraw last. We spend significant effort at characterizing the asymptotic behavior of our economy and conclude, for our benchmark specification, that there is convergence to a easy-to-calculate balanced growth path where participation falls at a constant rate. Finally, we provide a quantitative model where we calibrate the economy's initial wealth and skill distribution to cross-sectional U.S. data, solve, and project forward, with output in the form of a future path of wealth and earnings inequality.

The paper is organized as follows. We describe the basic model in Section 2 and the basic theory in Section 2.6, beginning with planning problems and moving toward competitive equilibria, and going from static via multiperiod to infinite-horizon settings. Section 4 contains the quantitative analysis and Section 5 looks at two relevant theoretical robustness exercises. Section 6 makes some concluding remarks.

2 A planning problem

To describe the economy, we formulate a planning problem. The objective of the planner is assumed to be a weighted sum of individual agents' utilities. We do not consider uncertainty in our analysis but it is straightforward to extend our framework and results to cover such cases. We will begin by stating a relatively general setting that we will later use in a quantitative exercise. This general setting can most straightforwardly be described as a neoclassical growth economy with heterogeneous agents, where heterogeneity appears in productivity and in the weights used in the planner's objective function on different households (in the decentralized economy, households differ in wealth in addition to productivity). Productivity here is a onedimensional stand-in for a variety of sources of heterogeneity influencing the hours choice (such as the "effort cost" of working, in addition to direct productivity differences); our intention in future work is to consider higher-dimensional heterogeneity, allowing us to capture aspects of aging, the size and composition of the household, and so on.

In our general setting, we allow capital accumulation and we consider a frictionless world where the only departure from a standard model—aside from the insistence on using utility functions where income effects exceed substitution effects—is a non-divisibility in hours worked. In the sequence of analyses in the coming sections, we will simplify this general model and consider cases without capital accumulation. It will become apparent at that point that the maximization of the planner's objective can be split up into a sequence of static problems where the only difference across periods will be the productivity of working. Hence, much of our analysis can be carried out in a static model where productivity is the exogenous variable that will be given a time interpretation.

2.1 Setup

Time is discrete, $t = 0, 1, \ldots T$, and we will potentially have $T \to \infty$. There is no population growth and there is a unit measure of heterogeneous individuals. Individuals are heterogeneous in $\omega > 0$, which determines their endowment of labor in efficiency units. At time t, the consumer's endowment will be assumed to be $\omega\gamma^t \geq 0$, where γ is the growth rate of (labor-augmenting) productivity, which is hence assumed to be constant over time; moreover, we focus on the case $\gamma > 1$.[1] Thus, a given household's productivity is constant in relative terms over time; again,

[1] Of course, the analysis can be carried out for any γ and, indeed, also for productivity paths that are not experiencing constant growth. One can, for example, consider the effects of a long-lasting (secular?) episode of stagnating productivity.

we could consider non-constancy and randomness and these extensions would, as far as we could tell, be straightforward.

We also assume that each household is associated with an $\iota \in [0,1]$, the parameter that indicates the weight in the planner's objective function. The joint density over ι and ω is given by $\Gamma(d\iota, d\omega)$ and the entire population size is normalized to one, i.e.,

$$\int_\iota \int_\omega \Gamma(d\iota, d\omega) = 1. \tag{1}$$

2.2 Preferences

Individuals have identical utility functions. These are additively separable preferences over time and the stationary momentary utility function is defined over consumption and hours worked, i.e.,

$$\mathcal{U}_0 = \sum_{t=0}^{T} \beta^t u(c_t, h_t). \tag{2}$$

We have $\beta < 1$. In what follows we will work with the following function:

$$u(c,h) = \frac{c^{1-\sigma} - 1}{1-\sigma} - \frac{\psi}{1+1/\theta} h^{1+1/\theta}, \tag{3}$$

where we assume $\sigma > 1$, $\theta > 0$, and $\psi > 0$. This is a special example from the class defined in Boppart and Krusell (2016): the so-called MaCurdy preferences (MaCurdy 1981). In this formulation, σ exceeding one means that the income effect exceeds the substitution effect; $\sigma = 1$ is a case of the King-Plosser-Rebelo preference class (King et al. 1988) where hours will not change over time as productivity grows. In our analysis below, we also discuss cases from the Boppart-Krusell class which feature a u which is not additive; these complicate the analysis and introduce some interesting aspects but, as we show in that context, the MaCurdy case is a robust one.

2.3 The indivisibility

The hours choice is constrained as follows: for each household,

$$h_t \in \{0, [\underline{h}, 1]\} = \mathcal{H}, \forall t. \tag{4}$$

As long as $\underline{h}$ is not equal to 0, this is a non-convex set; if it equals 1, the constraint implies that the consumer can only choose between working and not working. We use this constraint to capture the indivisibility and do not consider the deeper sources of the non-convexity. Such deeper sources could be various forms of fixed costs, such as commuting costs, or involve ways in which productivity per time unit becomes lower when less time is spent on the job. Our formulation has the main advantage that it is tractable.

2.4 Further heterogeneity within the household

Given the presences of an indivisibility, in some cases a household would, if it could, choose to randomize between 0 and $\underline{h}$. For sake of tractability, we will allow such a choice, but we will do it by assuming that each family—in line with some of the macroeconomic literature—has a continuum of identical members some of whom may end up working and some of whom may be assigned to leisure, while all enjoying the same level of consumption. So if the household would like to randomize with a probability of working equal to 1/3, we will simply have 1/3 of its members work for sure. This abstraction is not an essential feature of our analysis and we also examine cases where such convexification is not allowed. We do need notation for this phenomenon and this notation will be used extensively below: the fraction of members of a household working $h \in \mathcal{H}$ units of time is $\pi(h)$, with further indices and arguments for time, type, and so on whenever needed. Hence we need to assume, for all t and (ι, ω)

$$\pi_t(h; \iota, \omega) \in [0, 1], \tag{5}$$

$$\int_{h\in\mathcal{H}} \pi_t(h;\iota,\omega)dh = 1. \tag{6}$$

2.5 Resources

Output is produced by a neoclassical production function defined over capital k and an aggregate labor input n. Output can be consumed or invested. The depreciation rate is $\delta \geq 0$. Hence the resource constraint can be written as

$$F\left(k_t, \gamma^t n_t\right) + k_t(1-\delta) = \int_\iota \int_\omega c_t(\iota,\omega)\Gamma(d\iota, d\omega) + k_{t+1}, \tag{7}$$

where $n_t \equiv \int_\iota \int_\omega \omega \int_{h\in\mathcal{H}} \pi_t(h;\iota,\omega)hdh\Gamma(d\iota,d\omega)$ is the total labor input. Again, individual variables (like consumption) have an argument (ι,ω) whenever it is necessary to distinguish identities.

2.6 The planner's problem

We assume that the social planner maximizes a weighted sum of utilities:

$$\int_\iota \int_\omega \chi(\iota,\omega)\ \mathcal{U}_0(\iota,\omega)\Gamma(d\iota,d\omega), \tag{8}$$

where $\chi(\iota,\omega)$ denotes the Pareto weight on household (ι,ω). The expression $\mathcal{U}_0(\iota,\omega)$ is of course the present-value utility based on a set of choices. The choice sequence for the planner is $\left\{(c_t(\iota,\omega),\pi_t(h,\iota,\omega))_{\iota,\omega,h}, k_{t+1}\right\}_{t=0}^T$ for a given k_0. The planner faces the above-mentioned constraints (4), (5), (6), (7), and $k_{t+1} \geq 0$ for all t.

Hence, to summarize, let us compactly formulate our general planning problem as the maximization of

$$\int_\iota \int_\omega \sum_{t=0}^{T} \beta^t \left[\frac{c_t(\iota,\omega)^{1-\sigma}-1}{1-\sigma} - \int_{h\in\mathcal{H}} \pi_t(h;\iota,\omega)\frac{\psi}{1+1/\theta}h^{1+1/\theta}dh\right]\chi(\iota,\omega)\Gamma(d\iota,d\omega) \tag{9}$$

by choice of $\left\{(c_t(\iota,\omega),\pi_t(h;\iota,\omega))_{\iota,\omega,h}, k_{t+1}\right\}_{t=0}^T$ subject to

$$\pi_t(h;\iota,\omega) \in [0,1], \tag{10}$$

$$\int_{h\in\mathcal{H}} \pi_t(h;\iota,\omega)dh = 1, \tag{11}$$

and

$$F\left(k_t, \gamma^t n_t\right) + k_t(1-\delta) = \int_\omega c_t(\iota,\omega)\Gamma(d\iota, d\omega) + k_{t+1}, \tag{12}$$

where

$$n_t = \int_\iota \int_\omega \omega \int_{h\in\mathcal{H}} \pi_t(h;\iota,\omega) h\, dh\, \Gamma(d\iota, d\omega). \tag{13}$$

We will analyze this planning problem, with special emphasis on special cases of it, in the following sections. Most of our effort will move toward the analysis of decentralized equilibria, however, but the connection between solutions to the planner's problem and equilibria is an important one. In particular, we will comment on the mapping between planner weights and initial asset endowments in the decentralized model. This analysis will also allow us to comment on who *should* be working in the future—based, of course, on assumptions about welfare weight.

3 Qualitative analysis and competitive equilibria

What would normally follow next is a definition of a decentralized equilibrium for the environment described in the previous section. We will, however, not define our equilibrium in full generality yet but rather postpone it to a later section which discusses balanced growth (Section 3.6). Here, we will discuss very simple cases and we will, in fact, begin by studying the planning problem in some more detail, after which we will define equilibrium and discuss positive aspects.

The purpose of the present section is thus to provide qualitative insights from the studying the model. We will analyze it in steps. The first several steps involve assuming that there is no physical capital accumulation, i.e., that the production function is linear in labor: $F(k, \gamma^t n) = \gamma^t n$. The later addition of capital is not technically demanding and the main qualitative insights from the model without capital will survive; the quantitative analysis, however, requires capital.

It is clear from the planning problem above that in the absence of capital, the analysis can be divided up into a sequence of static problems. Thus, to the extent that one can map solutions to planning problems into decentralized market incomes, the properties of the latter can also be gleaned from studies of static economies. It is perhaps surprising that a static model is informative for our focus on how labor-force participation will evolve over time—how many, and who, will work at different future points in time as productivity grows. However, in our static model productivity is a key parameter—it will be represented by the shifting factor λ—and one can thus solve for labor-participation outcomes at every productivity level, for every household. The characterization of the static planning problem is therefore a key component of our analysis. As for the analysis of equilibria, the case with multiple periods has additional insights since individuals' wealth positions evolve dynamically and nontrivially. Hence, we also spend some time discussing stylized models with more than one period.

We use the static model for illustrating a number of important model features. The first feature, discussed in Section 3.1, is the role of the distribution of productivity across people for who will participate in the labor force. We begin the illustration by studying a planning problem with utilitarian weights, but we also make clear that planner weights are key for the distribution of hours worked across households. Utilitarian, i.e., equal, weights in the planner's problem are also of separate interest since they are often used in modern macroeconomic modeling to make normative insights. The idea here is that of a "veil of ignorance": ex ante, people are all the same, but some are born with high productivity/assets and others are less lucky. Who should then be working, and when, from this normative perspective? This analysis is contained in Section 3.1.1. In Section 3.1.2, we discuss the "representative-agent" special case briefly as well—here, all agents have the same productivity—and in Section 3.1.3 we look at the separate role of planner weights.

The importance of planner weights for who will work in the future leads us to the next topic: their relation to initial asset endowments in decentralized equilibria. In Section 3.2 we thus look at our first decentralized (static) economies and we provide a mapping to planning outcomes. In model's with non-convexities in choice sets, and the extensive margin of hours is one such example, this correspondence can be severed. However, for models like ours, as shown by Rogerson (1988), equilibria extended to allow lotteries that convexify the choice of hours again enable us to exploit the optimum-equilibrium connection. Indeed, in the quantitative analysis later in the paper, we use a decentralized setting with lotteries. Our illustrations here thus show a number of examples of mappings between the distribution of planner weights across households (in the planning problem) and the asset distribution (in the equilibrium with "transfers").

A third model feature concerns is aggregation: the present model does not admit a representative-agent reduced form. Thus, how are aggregate hours determined? And what is the response of hours worked to changes in the wage? These issues are discussed in Section 3.3.

The first model with rich dynamics is presented in Section 3.4, where a utilitarian optimum is computed and examined. The output of this model illustrates the lack of aggregation and how the nature of work changes over the longer run. The results here are also an important contrast to those coming from the quantitative equilibrium model in Section 4.

A central question in the paper is how the economy behaves in the "long run", assuming that productivity keeps growing. A relevant aspect of that question is under what conditions a balanced-growth path could materialize for the economy. To gain insight into such possibilities, we first look at the static planning problem again and examine it as the productivity shifting factor λ grows at a constant rate. This analysis, provided in Section 3.5, is non-trivial, as fundamental long-run features of the economy, such as its per-capita growth rate, can differ depending on the distribution of productivity and planner utility weights across households. In Section 3.6 we then turn to a truly dynamic model with capital and in its decentralized version. For this model, we provide conditions under which exact balanced growth obtains. This theorem is rather general but relies on some assumptions (on the initial asset distribution and on the productivity distribution) that will be made in the quantitative analysis. The proof of the theorem is constructive and gives important insights into how equilibria can be computed numerically.

We include some theoretical remarks after our quantitative model is presented. Section 5 thus makes two points, one about lotteries and one about non-additive utility. Lotteries are not always needed to decentralize optima. In Section 5.1 we thus illustrate this point and we also show that for every decentralized equilibrium with a nontrivial use of lotteries, there is a corresponding deterministic equilibrium, i.e., one not using lotteries, with the same ex-post consumption and hours distribution but with a different initial asset distribution. However, we also show that when lotteries are needed, the equilibrium without lotteries are quite a bit more complex to analyze.

In our benchmark setting, we use MaCurdy (1981)'s preference formulation, which features additive separability in consumption and hours. This formulation is convenient because the marginal utility of consumption does not depend on whether or not the households works, which means that in the case of lotteries, consumption will always be deterministic. In Section 5.2 we thus examine other, non-additive, utility functions where the income effect is stronger than the substitution effect: other members of the class of functions specified in Boppart and Krusell (2016).

3.1 Static planning problems

What role does the distribution of productivity across people play for who will participate in the labor force in the future? To work toward an answer, we will use a static model where overall labor productivity, represented by the shifting factor λ, is key.

3.1.1 Utilitarian weights

We begin by looking at the role of the distribution of productivity for the case of utilitarian planner weights: individuals are only heterogeneous in ω. In such cases, we will abuse notation and simply drop the ι argument in the various functions we use, such as Γ or π. To further simplify the exposition we first assume $\underline{h} = 1$ so that $\mathcal{H} = \{0, 1\}$, i.e., only the extensive margin is at play.

We can then write the planner's problem as the task of maximizing

$$\int_\omega \left[\frac{c(\omega)^{1-\sigma} - 1}{1-\sigma} - \pi(1;\omega) \frac{\psi}{1 + 1/\theta} \right] \Gamma(d\omega), \tag{14}$$

with respect to $(c(\omega), \pi(1;\omega))_\omega$, where $\pi(1;\omega)$ thus denotes the working share of the people with productivity ω, subject to the constraints

$$\pi(1;\omega) \in [0,1]\,, \tag{15}$$

and

$$n\lambda = \int_\omega c(\omega)\Gamma(d\omega), \tag{16}$$

where

$$n = \int_\omega \omega\pi(1;\omega)\Gamma(d\omega). \tag{17}$$

Here, λ denotes the current productivity level.

Because the preferences are additively separable in c and π and concave in c a planner with utilitarian weights will choose the same consumption level for all individuals, i.e., $c(\omega) = \bar{c}\ \forall\omega$. We thus have $\bar{c}^{-\sigma} = \mu$, where μ is the Lagrange multiplier on the resource constraint.

The planner chooses $\pi(1;\omega)$ depending on

$$\frac{\psi}{1+1/\theta} \lesseqgtr \mu\omega\lambda. \tag{18}$$

Since we are assuming heterogeneity between individuals in terms of productivity, this implies a threshold $\omega^* \equiv \frac{\psi}{1+1/\theta} \cdot \frac{1}{\mu\lambda}$ such that the planner chooses

$$\pi(1;\omega) = 0 \qquad \text{if } \omega < \omega^*$$

$$\pi(0;\omega) = 1 \qquad \text{if } \omega > \omega^*$$

Hence, the most productive people are assigned to work.

The resource constraint can then be rewritten as

$$\int_{\omega>\omega^*} \omega\Gamma(d\omega) = \mu^{-\frac{1}{\sigma}}\lambda^{-1}, \tag{19}$$

which, for a given distribution $\Gamma(\omega)$, implicitly defines the unknown marginal cost of resources μ (note that ω^* also depends on μ).[2]

To conclude, for every given level of λ, the planner will choose a cut-off ω^*. Individuals with a productivity below ω^* do not work, while individuals with higher productivity work. It is also easy to verify that the cut-off ω^* will be shifting up when λ grows, given that we have assumed $\sigma > 1$.

Finding 1. *With utilitarian weights and an extensive-margin choice only, the most productive people, and only they, work. As productivity grows, fewer and fewer people work.*

This finding obtains generally, i.e., for any productivity distribution, given utilitarian weights. It is a striking finding since many macroeconomic models with

[2]For some simple distributions $\Gamma(\omega)$ like a Pareto distribution (19) can be solved explicitly for μ.

heterogeneity, when they are used normatively (see Davila, Hong, Krusell, and Rios-Rull (2012) and the references therein), use precisely utilitarian weights.

Extending this problem to the case admitting interior hours choices too, i.e., $\mathcal{H} = \{0, [\underline{h}, 1]\}$ with $\underline{h} < 1$, is straightforward. In this case, the planner maximizes

$$\int_{\omega} \left[\frac{c(\omega)^{1-\sigma} - 1}{1-\sigma} - \int_{h \in \mathcal{H}} \pi(h;\omega) \frac{\psi}{1+1/\theta} h^{1+1/\theta} dh \right] \Gamma(d\omega) \tag{20}$$

with respect to $(c(\omega), \pi(h;\omega)\ \forall h \in \mathcal{H})_{\omega}$, where $\pi(h;\omega)$ denotes the share of the population with a given productivity level working h. The maximization is subject to

$$\pi(h;\omega) \in [0,1], \tag{21}$$

$$\int_{h \in \mathcal{H}} \pi(h;\omega) dh = 1, \tag{22}$$

and the resource constraint

$$n\lambda = \int_{\omega} c(\omega) \Gamma(d\omega), \tag{23}$$

where

$$n = \int_{\omega} \omega \int_{h \in \mathcal{H}} \pi(h;\omega) h dh \Gamma(d\omega). \tag{24}$$

If the hours choice is interior for an individual with productivity level ω then we have

$$\psi h(\omega)^{\frac{1}{\theta}} = \mu \omega \lambda, \tag{25}$$

where μ again is the Lagrange multiplier on the resource constraint, with $\mu = \bar{c}^{-\sigma}$. Note that (25) only holds with equality for ω with an interior solution. For other productivity levels there might be a mass point for h at 0, $\underline{h}$, or 1. We obtain $h(\omega) = 1$ for households with $\omega > \frac{\psi}{\mu\lambda}$. We have $h(\omega) = 0$ or $h(\omega) = \underline{h}$ for households with $\psi \underline{h}^{\frac{1}{\theta}} > \mu\omega\lambda$. So for all individuals with $\omega < \psi \frac{\underline{h}^{\frac{1}{\theta}}}{\mu\lambda}$ the planner chooses either $h(\omega) = 0$ or $h(\omega) = \underline{h}$ depending on

$$-\psi \frac{\underline{h}^{1+\frac{1}{\theta}}}{1+\frac{1}{\theta}} + \mu\omega\lambda\underline{h} \lesseqgtr 0. \tag{26}$$

This implies a threshold $\omega_1 \equiv \psi \frac{\underline{h}^{\frac{1}{\theta}}}{\mu\lambda\left(1+\frac{1}{\theta}\right)}$ such that the planner chooses $h = 0$ for all $\omega \leq \omega_1$ and $h = \underline{h}$ for all $\omega > \omega_1$.

To summarize, the planner solution will result in three break points:

$$\omega_1 = \frac{\underline{h}^{\frac{1}{\theta}}}{1+\frac{1}{\theta}} \cdot \frac{\psi}{\mu\lambda}$$
$$\omega_2 = \underline{h}^{\frac{1}{\theta}} \cdot \frac{\psi}{\mu\lambda}$$
$$\omega_3 = \frac{\psi}{\mu\lambda}$$

Individuals with $0 < \omega < \omega_1$ will not work at all. Individuals with $\omega_1 < \omega < \omega_2$ will work $\underline{h}$. Individuals with $\omega_2 < \omega < \omega_3$ will work $h = \left(\frac{\mu\omega\lambda}{\psi}\right)^{\theta}$. Finally, individuals with $\omega_3 < \omega$ will work highest possible hours, i.e., 1.

In analogy with (19) we can now write the resource constraint as

$$\underline{h} \cdot \int_{\omega_1}^{\omega_2} \omega\Gamma(d\omega) + \int_{\omega_2}^{\omega_3} \omega \left(\frac{\mu\omega\lambda}{\psi}\right)^{\theta} \Gamma(d\omega) + \int_{\omega_3}^{\infty} \omega\Gamma(d\omega) = \mu^{-\frac{1}{\sigma}}\lambda^{-1}. \tag{27}$$

For a given distribution $\Gamma(d\omega)$ this equation implicitly defines the equilibrium μ (note that ω_1, ω_2 and ω_3 also depend on μ).

To conclude, the conclusions we drew from the model with only an extensive margin extends to the case with $h \in \{0, [\underline{h}, 1]\}$. For every given level of λ, there will be cut-offs which will guide the working choice, with more productive agents working more or harder, and the cut-offs will be shifting up as everybody becomes more productive, i.e., as λ grows.

Finding 2. *With utilitarian weights and both an extensive and an intensive margin, work is monotonic in productivity: more productive people work more hours. As productivity grows, people work less and less in terms of the intensive as well as the extensive margin.*

3.1.2 Utilitarian weights and identical individuals

A special case of interest is that used in many macroeconomic analyses, where everyone in the economy has the same productivity. Hence, we now assume that individuals all have the same $\omega \equiv \overline{\omega} > 0$. To simplify the exposition, we restrict attention to an extensive-margin choice only.

The planner still solves the problem defined by equation (14) subject to the constraints (15) and (16) and has Pareto weights $\chi(\iota, \omega) = 1 \; \forall \iota, \omega$. And again since preferences are additively separable in c and h and concave in c the planner will choose the same consumption level $\overline{c}$ for all individuals, with $\overline{c}^{-\sigma} = \mu$. For the work choice, the planner will now need to use convexification, at least when λ is high enough. Hence, the choice is over the fraction of individuals who will work and it will be made in such a way that the marginal disutility from an additional individual working equals the marginal utility from consumption:

$$\frac{\psi}{1 + 1/\theta} = \lambda \overline{\omega} \mu, \tag{28}$$

where we assume that λ is large enough for this equation to hold with equality (for low enough values of λ, everybody will work). The resource constraint can then be rewritten as

$$\overline{\omega} \pi(1; \overline{\omega}) = \mu^{-\frac{1}{\sigma}} \lambda^{-1}. \tag{29}$$

Equations (28) and (29) now together define $\pi(1; \overline{\omega})$ and μ. The fraction of individuals working is given by

$$\pi(1; \overline{\omega}) = (\overline{\omega} \lambda)^{\frac{1-\sigma}{\sigma}} \left(\frac{\psi}{1 + 1/\theta} \right)^{-\frac{1}{\sigma}}. \tag{30}$$

To conclude, when everyone has the same productivity the utilitarian planner is indifferent about who is working in the economy: it only cares about the fraction of individuals working. The fraction working, $\pi(1; \overline{\omega})$, is decreasing in λ.

Finding 3. *With a representative-agent setting, we recover the Rogerson (1988) kind of outcome, where convexification is used actively at all points in time after a certain date. After that date, the fraction of people working every period is strictly increasing.*

We will return to the details of growth behavior later, e.g., for determining the asymptotic rate at which the labor force is declining as a function of the rate of productivity growth.

3.1.3 Heterogeneous weights

Now we turn back to the case where individuals are heterogeneous in ω. With the utilitarian planner weights we saw above that the most productive individuals were the ones who should be working. How does this conclusion change for other planner weights? The answer, on some level, is obvious: a higher planner weight will generate more utility for the household, which leads to more consumption and, at least weakly, less work. Let us, however, work out some examples for illustration.

Suppose, to focus the question, that we assume that the weights are identical for all individuals ι with identical ω, i.e., $\chi(\iota, \omega) = \chi(\omega)\ \forall \iota, \omega$. We also restrict attention to an extensive-margin choice only. Then we can write the planner's problem just in ω space. The planner thus maximizes

$$\int_{\omega} \left[\frac{c(\omega)^{1-\sigma} - 1}{1-\sigma} - \pi(1;\omega) \frac{\psi}{1+1/\theta} \right] \chi(\omega) \Gamma(d\omega) \tag{31}$$

subject to the same constraints as before: (15) and (16).

Generally, the work decision is guided by the following rule:

$$\omega \lambda \mu \lesseqgtr \chi(\omega) \frac{\psi}{1+1/\theta}. \tag{32}$$

It is clear from this equation that what matters for how the working assignment depends on productivity is $\omega/\chi(\omega)$. Thus, if this object is increasing (decreasing) in

ω, higher-productivity households will work more (less). The planner weight $\chi(\omega)$ is an abstract one at this point but will soon be given content when it is connected to initial-period asset holdings in the context of the market equilibrium. Turning to consumption, we obtain

$$c(\omega) = \left(\frac{\chi(\omega)}{\mu}\right)^{\frac{1}{\sigma}}, \tag{33}$$

i.e., whereas hours are increasing or decreasing in ω depending on the second derivative of $\chi(\omega)$, consumption is increasing in ω if and only if $\chi(\omega)$ is increasing, i.e., only the first derivative of χ matters.

To take a specific, but interesting example, suppose that the planner weights are $\chi(\omega) = \omega \ \forall \omega$. Then equation (32) becomes independent of ω:

$$\lambda\mu = \frac{\psi}{1 + 1/\theta}. \tag{34}$$

Assuming that λ is high enough that not all households work, this equation must hold with equality, since we must have some individuals working in the economy.

The consumption for any particular productivity level in this case is given by

$$c(\omega) = \left(\frac{\omega\lambda(1 + 1/\theta)}{\psi}\right)^{\frac{1}{\sigma}}. \tag{35}$$

Thus, the consumption plans are pinned down uniquely but there is indifference as to who works, subject to total production equaling total consumption.

Finding 4. *The planner weights χ allow much generality in terms of who is assigned to work. The key determinant of whether more productive (higher-ω) households work more or not is whether $\omega/\chi(\omega)$ is increasing or not. If $\chi(\omega)$ is linear, the planner is indifferent as to who works. Regardless, fewer people work the higher is aggregate productivity.*

Just like above, these results can easily be extended to the case where an intensive margin choice is added ($\underline{h} < 1$). The insights above then generalize.

3.2 Static competitive equilibria

In this section we use the static model to define competitive equilibria, solve for them, and illustrate the correspondence between solutions to the planner's problem and competitive allocations. Intuitively, the planner weights $\chi(\iota, \omega)$ will correspond, nonlinearly, to the initial asset endowments of households.

A static equilibrium is particularly simple in that there is no equilibrium interaction: all prices are exogenous and it is as if households act in isolation. With more periods, the market for savings is operative and nontrivially determines outcomes through an endogenous interest rate; we look at this case below.

The household's problem in a decentralized economy without convexification is naturally to maximize

$$u\big(c, h\big) = \frac{c^{1-\sigma} - 1}{1-\sigma} - \frac{\psi}{1+1/\theta} h^{1+1/\theta} \tag{36}$$

by choice of (c, h) subject to the constraints

$$c = h\omega\lambda + a(\iota, \omega) \tag{37}$$

$$h \in \{0, [\underline{h}, 1]\} \tag{38}$$

with $a(\iota, \omega)$ given. However, to obtain a tight link to the set of Pareto optima (as defined by solutions to the planning problem for different weight functions $\chi(\iota, \omega)$), we will instead use the following problem, where convexification is allowed: maximize

$$u\big(c, h\big) = \frac{c^{1-\sigma} - 1}{1-\sigma} - \frac{\psi}{1+1/\theta} \int_{h\in\mathcal{H}} \pi(h) h^{1+1/\theta} dh \tag{39}$$

by choice of $(c, (\pi(h))_{h\in\mathcal{H}})$ subject to the constraints

$$c = \lambda \int_{h\in\mathcal{H}} \pi(h) h\omega dh + a(\iota, \omega) \tag{40}$$

$$\int_h \pi(h) dh = 1, \quad \text{and} \quad \pi(h) \in [0, 1], \quad \forall h \in \mathcal{H}, \tag{41}$$

with $a(\iota, \omega)$ given. For either of these problems, the household's choice is independent of the behavior of other households. The market-clearing condition can be stated as requiring that total consumption must smaller or equal to total production. Trivially this will be satisfied with equality since utility is increasing in consumption, given that we assume

$$\int_{\omega}\int_{\iota} a(\iota, \omega)\Gamma(d\iota, d\omega) = 0, \tag{42}$$

i.e., that the sum of initial endowment claims across people equals zero. We do assume this.[3]

We see that, regardless of whether convexification is allowed or not, analysis of the equilibrium of the static economy boils down to analysis of the household's choice problem. Turning to the role of convexification, let us make some preliminary remarks for a simple case and then consider more general cases. To that end, let us assume that all initial asset endowments are zero, i.e., $a(\iota, \omega) = 0 \ \forall \iota, \omega$, and that only the extensive is operative ($\underline{h} = 1$). Then we see immediately that a decentralized economy without convexification generates the result that all agents work, no matter how low λ is, so long as it is positive. It is easy to show that this case is not generally optimal. In particular, agents with close enough endowment levels would benefit from a coin-flip (with some outcome probability) deal where consumption is equalized but work effort is determined by the coin flip. A less trivial question is whether equilibria where convexification is not allowed might nevertheless be constrained optimal, in the sense that there is no Pareto-improving allocation subject to the restriction that coin flips not be used in this alternative allocation. We will discuss this issue in Section 5.1. Here, let instead from now on focus on the cases where convexification is allowed and we will draw connections between the

[3]This of course is not restrictive; we could assume that there is a net amount of endowments, and then the market-clearing condition is that total consumption must equal total production plus total endowments.

endowment values and the planner utility weights on different households.

3.2.1 A general distribution of initial assets

Let us consider a more general initial asset allocation but let keep restricting attention to the extensive-margin choice for simplicity. Let us also assume that there is no asset inequality between households with the same ω; hence we can write $a(\iota, \omega) = a(\omega)$. Each household $\omega \in [0, \infty]$ now solves

$$\max_{c(\omega), \pi(1;\omega)} \left[\frac{c(\omega)^{1-\sigma} - 1}{1-\sigma} - \pi(1, \omega) \frac{\psi}{1 + 1/\theta} \right] \tag{43}$$

subject to

$$c = a(\omega) + \pi(1; \omega)\omega\lambda, \tag{44}$$

$$\pi(1; \omega) \in [0, 1], \qquad \forall \omega, \tag{45}$$

where $\pi(1; \omega)$ denotes the fraction of the family with productivity ω that is working. We will also assume for now, just to simplify the notation and discussion, that the combination of initial assets, $\omega\lambda$ and preference parameters are such that the constraint $\pi(1; \omega) \leq 1$ is not binding for any household.

The first-order conditions are

$$\omega\lambda c(\omega)^{-\sigma} = \frac{\psi}{(1 + 1/\theta)} - \eta(\omega) \tag{46}$$

$$\eta(\omega) \geq 0, \tag{47}$$

$$\eta(\omega)\pi(1; \omega) = 0, \tag{48}$$

where η is the multiplier on the constraint $\pi(1; \omega) \geq 0$: a family with $\eta > 0$ is thus in a corner with $\pi(1; \omega) = 0$. Of course, the solution to these conditions and the budget will make c and η be functions of ω but we do not need to indicate this by explicit function arguments here. When we refer to the solutions, we will use explicit arguments. So for those households having an interior solution for $\pi(1;\omega)$, consumption will be given by

$$c(\omega) = \left(\frac{\omega\lambda(1+1/\theta)}{\psi}\right)^{\frac{1}{\sigma}}. \tag{49}$$

The associated value for the probability of working for such a household will satisfy

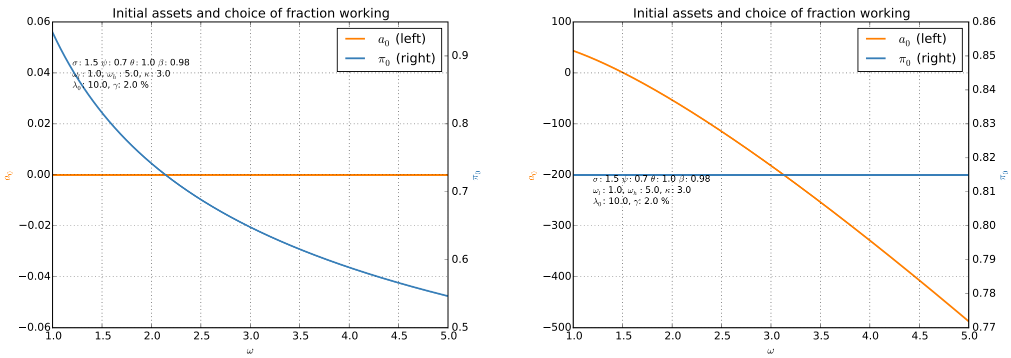
$$\pi(1;\omega) = \left(\left(\frac{(1+1/\theta)\omega\lambda}{\psi}\right)^{\frac{1}{\sigma}} - a(\omega)\right)(\omega\lambda)^{-1}, \tag{50}$$

and the condition for making this a solution to the maximization problem is that, given the parameter values and this household's ω (and, hence, its $a(\omega)$), this latter value be between 0 and 1. So looking at this expression, we see that there will be a range of ω values for which consumption does not directly depend on the asset level.

Now let us try to find a set of planner weights $\chi(\omega)$ that correspond to the asset levels in the equilibrium allocation just derived. Using the planner's condition for determining consumption levels, (33) in Section 3.1.3, we see that to match consumption levels, $\chi(\omega)$ has to equal (or, really, be proportional to) ω. Recall that, in that planning solution, the planner was indifferent as to who worked. For this range of ωs, then, the fraction working, $\pi(1;\omega)$, given in equation (50) is trivially consistent with the planner's solution. For the asset/endowment levels such that the household chooses to not work, consumption simply equals asset holdings: $c(\omega) = \mu a(\omega)$, which from the planner's first-order condition for consumption delivers $\chi(\omega) = a(\omega)^{\sigma}$. This value is also high enough relative to ω—recall that the working decision was determined by $\omega/\chi(\omega)$—that the planner indeed makes this household not work. This completes the mapping between a static competitive equilibrium with an arbitrary asset distribution to a planner allocation with a set of Pareto weights.

Finding 5. *The equilibrium allocation can be arrived at as a solution to the planner's problem—an application of the first welfare theorem. For households that choose nontrivial convexification in the equilibrium allocation, the planner is indifferent in* *terms of the randomization, but the household in the equilibrium allocation is not. For these households, higher asset levels do not result in higher consumption, nor in higher planner weights, but merely lower working probabilities.*

Two equilibrium allocations are illustrated in Figure 1: it shows the connection between initial assets and the choice of fraction working for two specific asset distributions and a specific choice of preference parameters. The first one has perfect asset equality, and hence work decisions differ by ω. The second equilibrium works out the asset distribution such that everybody makes the same work decision: more productive agents then need to be poorer.



(a) The choice of $\pi(1;\omega)$ if everyone is given initial assets $a(\omega) = 0\ \forall \iota, \omega$

(b) Distribution of initial assets so that every family chooses the same $\pi(1;\omega)$

Figure 1: Choice of $\pi(1;\omega)$ for different productivity levels; both examples assuming a truncated Pareto distribution of productivity

3.2.2 Connection to the planning problem with utilitarian weights

Using the same logic and equations as in the previous section, we can of course start from a planning allocation given some arbitrary weights and work out the associated asset levels. For a particular case, we find the answer instructive: the case of a utilitarian planner, where individuals are heterogeneous in terms of ω. The reason is that the utilitarian planner's objective often is mentioned as a good normative benchmark—and used, in practice, in many macroeconomic analyses to derive optimal policy—with the "behind-the-veil" argument. This argument supposes that there is a pre-stage, at which households are all identical but after which their skills are realized as draws from a common distribution. In the pre-stage, households would then, in an optimal allocation, arrange transfers to be set up ex post to compensate for the risk of receiving low skill draws. The implied optimal allocation is, then, exactly the utilitarian allocation we studied in Section 3.1.1. So by connecting this particular optimum to initial asset endowments in the decentralized equilibrium we exactly obtain the transfers needed to reach this "macroeconomic" optimum.

To simplify the exposition, we again assume that $\underline{h} = 1$. As we saw in 3.1.1 the planner solution is that everyone consumes $\overline{c}$ and individuals with efficiency above ω^* work, while those with efficiency below ω^* do not work.

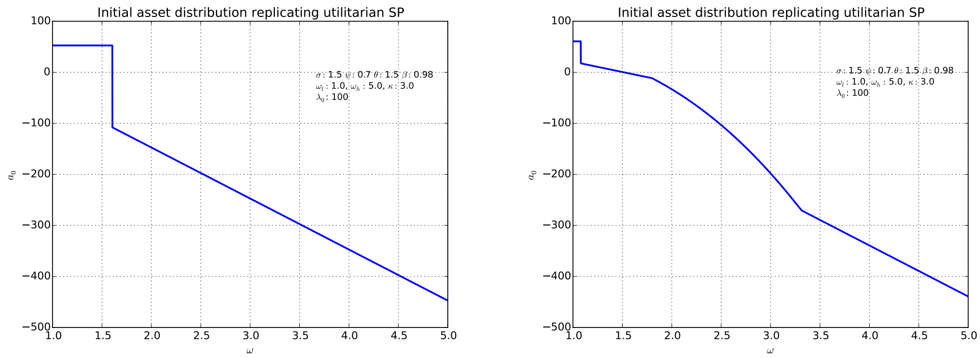
The asset distribution that gives rise to the same allocations of consumption and work in a decentralized economy as in the planner's solution is easy to work out. It is given by

$$a(\iota, \omega) = \overline{c} - h_{SP}(\omega)\omega\lambda, \tag{51}$$

where $h_{SP}(\omega)$ denotes the planner's choice of work allocation for that particular productivity level.

To give an example, we assume the individuals are distributed along the productivity dimension according to a truncated Pareto distribution with lower and upper bounds $\underline{\omega}$ and $\overline{\omega}$ and with shape parameter κ. Then Figure 2a shows the resulting asset distribution as a function of individual productivity. Figure 2b shows the corresponding graph if we choose $\underline{h} < 1$, so that also the intensive margin is in play.

The intuition behind why the individuals make the same working choice and therefore mechanically the same consumption choice as in the planning optimum is



(a) Only extensive margin; individuals with $\omega < \omega^*$ choose not to work, while the ones with higher productivity do

(b) $h \in \{0, [\underline{h}, 1]\}$; individuals work not at all, $\underline{h}$, at their interior solution h^*, or 1

Figure 2: Initial assets to replicate planner allocations, static model, utilitarian planner weights

straightforward. For the individuals with high productivity, $\omega > \omega^*$, the initial assets are negative, and hence they are forced to work; moreover, the more productive they are, the higher is the initial debt, so the net of this debt and their labor income cancel to deliver the same consumption allocation as for any other household. For the individuals with low productivity, they have received initial assets to be able to consume $\bar{c}$ without working. The disutility from working for them is larger than the increase in utility from the extra consumption for their productivity at that consumption level.

Finding 6. *To decentralize the utilitarian optimum, it is necessary to "tax" high-skill individuals by assigning them large initial debts; the higher are their skills, the larger is the debt. As a result, they will have to work. Low-skill individuals not assigned to work in the optimum, on the other hand, are recipients of transfers in the form of initial asset endowments.*

3.2.3 Connection to the planning problem with identical individuals

We finally turn back to the case when everyone in the economy has the same productivity. As we saw in Section 3.1.2, the utilitarian planner's solution is then that everyone consumes $\bar{c}$ and that a fraction of the households work. Moreover, the planner is indifferent as to who works. There are two different ways to think about the optimum in this homogeneous-agent utilitarian case. One is a lottery interpretation: the planner assigns an equal chance (or, rather, risk) of working to all households. The other interpretation is that some households are pre-assigned to work whereas others are not, without involving any lotteries. That is, the utilitarian planning outcome is set up to weight all households' utilities equally but this does not necessarily imply equal treatment in terms of outcomes, even ex ante.

Formally, the individual's problem is again to maximize (39) subject to (40) and (41). The first way to decentralize this allocation would be to give a fraction $\pi(1;\overline{\omega})$ of the individuals negative assets $a = \bar{c} - \lambda\overline{\omega}$ and a fraction $(1 - \pi(1;\overline{\omega}))$ positive assets $a = \bar{c}$. This would give individuals different utility, both in expectation and ex post, but the allocations of consumption and work would correspond to the planner's solution. This is reminiscent of the result in the dynamic model studied in (Krusell et al. 2008), where at any point in time different households have different assets but some work and some do not, only to later switch; an agent's asset position rather is an indication of how much an agent will have to work over the remaining lifetime.

The second way to decentralize this economy, one that gives all individuals the same expected utility, is to introduce a lottery. This treatment exactly follows (Rogerson 1988) and involves zero initial asset holdings for all agents. The ex-post utility of course will be different: the individuals who "win" the lottery do not have to work and will end up with higher utility than those who "lose" and have to work.

Finding 7. *It is possible to decentralize the utilitarian homogeneous-agent optimum in two very different ways: one involves equal ex-ante utilities and lotteries, whereas* *the other involves unequal ex-ante utilities and no lotteries.*

The fact that a given ex-post allocation can be decentralized with and without lotteries—by the reshuffling of initial asset endowments—is an important point that we discuss further in Section 5.1 below.

3.3 Aggregation

Our benchmark economy does not admit a representative agent. One immediate way to see this is from the fact the under the planner's solution aggregate outcomes depend on the welfare function as well as on the distribution of ω. Consequently, also in a decentralized economy the equilibrium outcome of the aggregate variables depends on the initial wealth distribution, and so on. In the following, we make some remarks on the aggregation issue. Furthermore, we comment on another related aspect: the implication of our framework for the Frisch elasticity of labor supply.

3.3.1 The determination of aggregate hours worked

Why does the model at hand not allow for a representative agent? In the specified economy, there are several reasons. First, we allow for heterogeneity in labor productivity, which in effect boils down to heterogeneity in the relative price between consumption and leisure, making simple aggregation not possible. Second, the MaCurdy preferences used are not in the Gorman class. Hence, the marginal propensity to decrease hours worked out of wealth differs between rich and poor households and the wealth distribution (or the distribution of the planner weights) starts to matter. Finally, by restricting the choice set of hours worked, with a positive $\underline{h}$ some agents are potentially constrained or forced to randomize between 0 and $\underline{h}$. This is another reason why the overall distribution starts to matter since it determines the fraction of constrained households. All these reasons imply that aggregate labor supply depends on the entire joint distribution of household characteristics (ω, a) (or, put in terms of the planning variables, on $(\chi(\iota, \omega), \omega)$).

It is important to stress that the lack of aggregation is a very general feature of models of endogenous labor supply. For instance, many of the commonly used preferences over consumption and leisure are not part of the Gorman class (with the important exception of the Cobb-Douglas and GHH—Greenwood, Hercowitz, and Huffman (1988)—case). Nevertheless, many macroeconomic models keep a focus on aggregates by simply assuming that there is one (type of) agent, i.e., by ruling out household heterogeneity in labor productivity or financial wealth. When it comes to the normative question, the literature usually focuses exclusively on the utilitarian welfare function. The discussion below shows that the choice of the welfare function is very important in our context and can completely overturn whether the most productive or the least productive drop out of work first.

3.3.2 The Frisch elasticity

Note that the benchmark MaCurdy preferences imply—as long as we have an interior solution—a constant Frisch elasticity of labor supply equal to θ. However, this is changed with a positive $\underline{h}$. With this discontinuity, we can obtain bunching at $\underline{h}$ so that for some households hours worked are completely inelastic with respect to the wage rate. However, at the same time we can have some households that randomize within the family which fraction, $\pi(1;\omega)$, has to work $\underline{h}$. In this case the objective function becomes

$$u(c,h) = \frac{c^{1-\sigma}}{1-\sigma} - \pi(1;\omega)\psi\frac{\underline{h}^{1+1/\theta}}{1+1/\theta}, \tag{52}$$

which is linear in the choice variable $\pi(1;\omega)$. Hence, for households that randomize, the Frisch elasticity tends to infinity. Since the fraction of households with a Frisch elasticity equal to θ, 0, and ∞ changes over time and depends on household heterogeneity, the "average" Frisch elasticity in the economy changes over time in a non-trivial way. However, we will show below in Section 3.5.1 that under some conditions, all the households either randomize or quit work altogether as time goes to infinity. In this case, we expect the Frisch elasticity of labor supply to increase over time at least eventually.

3.4 Multiperiod settings

So far we have only considered static models, since in the absence of capital, the analysis can then be divided up into a sequence of static problems. The connection between social optima and competitive equilibria are a little less trivial to work out with two or more periods, though they follow the same general logic. The reason is that the interest rate in the decentralized model is now endogenous, as is asset accumulation.

3.4.1 Decentralizing the utilitarian solution

We start by looking at the case of a utilitarian planner and heterogeneous agents. As discussed before, the planning problem is actually just a sequence of static problems and, hence, no new analysis needs to be considered here. Over time, we simply obtain a sequence of cutoff points in ω space, one for each time period. We denote those cutoff points $\{\omega_t^*\}_{t=0}^T$; each time period t, the individuals with productivity above (below) ω_t^* will (not) work. In terms of consumption, all individuals receive the same amount; this amount, which we denote $\overline{c_t}$, depends on time since work effort, and hence total output, will change over time both due to productivity growth and changes in total hours worked.

How do we find the initial asset distribution that replicates this planner solution? Assuming that we have T time periods, and restricting attention to $\underline{h} = 1$, the individual now solves the following problem:

$$\max_{\{c_t,\pi_t(1)\}_{t=0}^T} \sum_{t=0}^{T} \beta^t \left[\frac{c_t^{1-\sigma}-1}{1-\sigma} - \pi_t(1)\frac{\psi}{1+1/\theta} \right], \tag{53}$$

subject to

$$\sum_{t=0}^{T} p_t c_t(\iota,\omega) = a_0(\iota,\omega) + \omega \sum_{t=0}^{T} p_t \gamma^t \pi_t(1), \tag{54}$$

where $\frac{p_t}{p_{t+1}} = 1 + r_{t+1}$. Thus, given a planning allocation, if every household is given initial assets according to

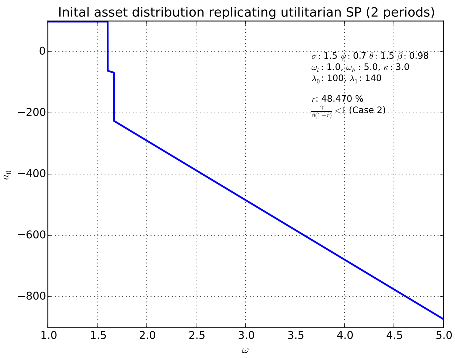
$$a_0(\iota,\omega) = \sum_{t=0}^{T} p_t \overline{c_t} - \omega \sum_{t=0}^{T} p_t \gamma^t \pi_t^{SP}(1;\omega), \tag{55}$$

where $\pi_t^{SP}(1;\omega)$ indicates the planner's working choice for this particular individual, every individual will choose consumption and working such that the planner's solution is decentralized. The interest rate for each period we can trivially find from the Euler equation of the individuals, which all look identical, since utility is additively separable in consumption and hours and they all consume the same amount.

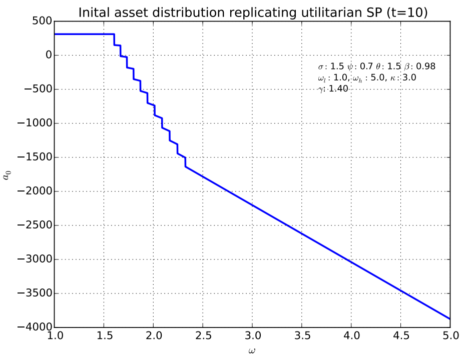
We illustrate these points assuming that the individuals are distributed along the productivity dimension according to a truncated Pareto distribution with lower and upper bounds $\underline{\omega}$ and $\overline{\omega}$ and with shape parameter κ. Figure 3 shows two examples of resulting asset distribution as a function of individual productivity, one where we assume two periods ($T = 1$), and one where we assume $T = 10$.

How can we be sure that the individuals will actually choose the same work sequence as dictated by the planning allocation, given $a_0(\iota,\omega)$? The answer is found in how interest rates clear asset markets over time: they do so in such a way that there are both borrowers and lenders. To obtain further insights into this process, we solve a simple two-period model in the next section as an illustration.

Finding 8. *The decentralization of the utilitarian optimum in a multi-period model implies an asset distribution where initial assets are inversely related to individual*



(a) Initial asset distribution for 2 periods



(b) Initial distribution if $T = 10$

Figure 3: The initial asset distributions needed to implement a utilitarian allocation

productivity levels, just like in the one-period model. In the case with a continuous productivity distribution and an extensive margin only, the initial asset distribution is a step function, with more steps the more periods there are.

Before moving on, let us point out that the above logic of how utilitarian optima are supported by competitive equilibria straightforwardly extend to the case when $\underline{h} < 1$, as in the one-period illustration earlier in Figure 2b.

3.4.2 A two-period model with lotteries

The implementation of the utilitarian optimum with a continuous ω distribution raised the issue of how work decisions are made over time. To explore this issue, let us now consider another special case, namely, that with zero initial asset holdings for all agents but different productivities. This case allows us to explain how the interest rates need to be set in order to clear asset markets and how they simultaneously guide the intertemporal working decisions of households. In the derivation of the key results here, it is important to use marginal calculations, which in general—and in the special case with zero initial assets for all agents—involves some households using convexification actively. We take a two-period model as an example but the principle extends to any number of periods.

We will not describe all the formal analysis here, but a key feature is that the household's problem (53) leads to qualitatively different conclusions depending on whether $\gamma/(1+r)$ is larger than, equal to, or smaller than β. The expression $\gamma/(1+r)$ gives the present value in consumption units of working in the future relative to working today: the wage rises by a factor γ and this amount needs to be discounted by $1 + r$. β, on the other hand, is the relative amount by which working in the future causes a lower utility loss, due to impatience. Hence, if these two quantities are equal, the household is indifferent as to the timing of work, whereas if the former is higher (lower), there is a strict preference to work late (early).

Moreover, there will be a monotonicity of work decisions in individual productivity levels ω. Suppose, for example, that $\gamma/(1 + r) > \beta$. Then there will be three cutoff points in ω space, and hence four regions. In the first, lowest region, people work both periods. In the next lowest region, people work in the second period and randomize in the first. In the next-to-highest region, people do not work in the first period but work in the second period. Finally, the most productive agents do not work in the first period and randomize in the second. In terms of consumption behavior, consumption is increasing in the skill level but rises faster in skill in the first and third region, since in the other two skill increases involve working less (randomizing to work with a lower probability).

Now consider the asset market in this two-period economy, still maintaining $\gamma/(1 + r) > \beta$. Here we can show that all agents will want to borrow in the first period, a key reason being that people postpone their income generation. But even the least skilled choose to borrow, even though they work full time in both periods, since the interest rate in this case is low; the verification of this assertion is straightforward.[4] We conclude that it must be, instead, that $\gamma/(1 + r) \leq \beta$.

[4]This result, along with others in this section, of course critically rely on $\sigma > 1$, which implies

If $\gamma/(1+r) < \beta$, there are again three cutoff points and four regions: the least skilled work in both periods, the next group works in the first and randomizes in the second, the second-highest skilled work in the first but not in the second period, and the most skilled randomize in the first period but do not work in the second. By looking carefully at consumption relative to income in the first period we see that it is possible to have both borrowers and savers and hence that asset markets clear. However, an equilibrium can also feature $\gamma/(1+r) = \beta$, where people are indifferent as to when to work. The equilibrium is generically unique.

Finding 9. *In a multi-period model with zero initial asset holdings, it must be that interest rates are such that working decisions are delayed; in a two-period model the interest rate must be low: $\gamma/(1+r) \leq \beta$. Higher-skilled individuals work less by withdrawing faster from the labor force.*

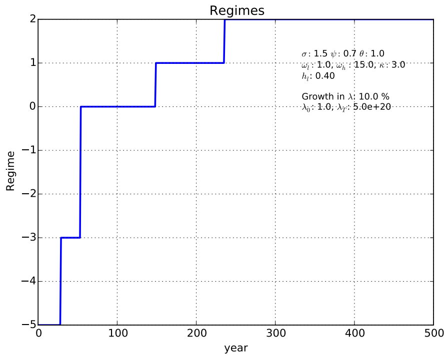
3.4.3 A utilitarian optimum in a long-horizon model: qualititative characteristics

In what follows, we will look at a much longer-horizon setting. In Sections 3.5 and 3.6 we go to the limit and study the asymptotic, and potentially balanced, long-run growth behavior. Here we simply extend the horizon to gain some insights on the transitional behavior. For this, we use the utilitarian planning problem and assume skill heterogeneity. As explained above, the planning problem is convenient since it can be studied purely as a sequence of static problems; the utilitarian case is an interesting normative benchmark.

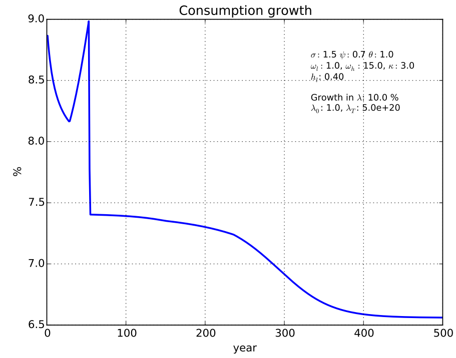
As an example, assume the individuals are distributed along the productivity dimension according to a truncated Pareto distribution with lower and upper bounds $\underline{\omega}$ and $\overline{\omega}$ and with shape parameter κ. Furthermore we assume $\underline{h} < 1$, i.e., that the intensive margin is also in play. Then for a certain parameter setting Figure 4

that the income effect of a wage increase dominates the substitution effect.

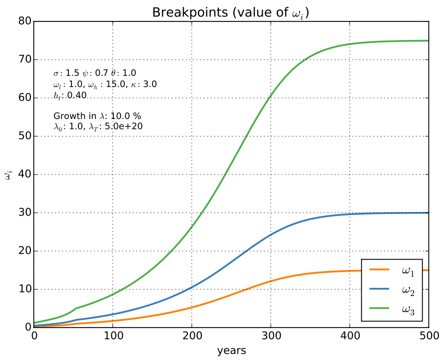
illustrates the resulting development for who is working, the growth of consumption, the cutoff points, and the mass of individuals in each of the four mutually exclusive working modes: working full time, on the intensive margin, at $\underline{h}$, or not at all.



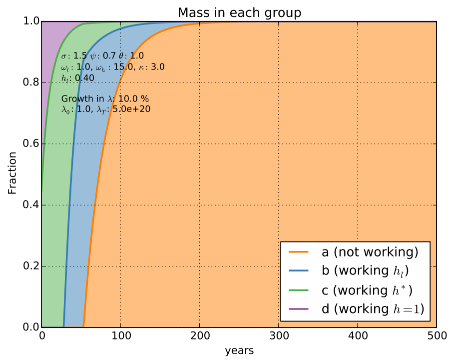
(a) Regimes



(b) Growth of consumption



(c) Values for cut-off points



(d) Mass of individuals by work group

Figure 4: The evolution over the long run for a utilitarian planner and heterogeneous individuals

The top left subgraph illustrates that the economy moves through "regimes" in terms of who is working. Each regime is a collection of the different active working modes of households in the population at that time. A number of such modes are possible and five modes are visited in the particular example displayed. The first regime (labeled -5) is that all households either work at the maximum number of hours (1) or an interior number, and after around 20 years, in the second regime (labeled -3), some households also work at the lowest amount $\underline{h}$. In the third regime (labeled 0), all the four modes of working coexist, and in the last two regimes (1 and 2) the full-time and intensive-hour households withdraw, in order, so that in the last regime there are only households working at the minimum amount or not at all.

Because the economy goes through distinct regimes in this case, the path for aggregates, such as consumption growth in the top right panel of the, is somewhat uneven (but continuous), because the nature of the economy's response to aggregate productivity change changes depending on people's working modes. This is an illustration of what we saw above in Section 3.3: there is no aggregation theorem in this class of economies and aggregate intra- as well as intertemporal substitution of labor, and hence consumption, depends nontrivially on the distribution of households.

In the utilitarian planning solution, again, the time at which a household makes its labor-force withdrawal is larger the higher is the household's level of productivity. The bottom left panel of the figure thus illustrate how the household productivity cutoffs ω_1 (between not working and working at $\underline{h}$), ω_2 (between $\underline{h}$ and an interior solution), and ω_3 (between an interior solution and the maximum) evolve over time. Because the aggregate wage path is monotonic in the example, there will not be any reversals, e.g., with households withdrawing and then coming back into the labor force later, which is reflected in the cutoff lines being monotonically increasing.

We see that, in our example, the transition to the final regime, where a small fraction of the households—the very most productive—work and provide output for the rest of the economy, is occurs in about 100 years. In the quantitative market equilibrium model examined in Section 4 below, the distribution of households over initial asset holdings and productivity is calibrated to the data. The data reveals, not surprisingly, that the most productive households also have high asset wealth, and both of these facts contribute to reversing the qualitative conclusion in the utilitarian optimum: a large fraction of the productive households will withdraw early from the labor force in the calibrated equilibrium model. This force will imply a slower transition of working patterns, since it is more difficult for the low-productivity households to provide consumption for the rest of the economy than it is for the high-productive ones, and in the interest of consumption smoothing this transition therefore becomes slower, given the same primitives (the quantitative results in Section 4 rely on a different calibration so in this sense the results there cannot be compared directly to those here).

Finding 10. *The utilitarian optimum allows consumption smoothing more easily since the most productive agents will work the longest and hence the transition to an economy where only a few households work can occur fast.*

3.5 Balanced growth without capital: planner's solutions

Under what conditions can we have a balanced growth path in this economy? Since working hours will move toward zero and the extensive margin will be active, the asymptotic behavior does not appear balanced, but we will show that a form of balanced growth can be shown to occur both in the planner's solution and in a market. The planning problem is easier to analyze and this is where we begin.

3.5.1 Asymptotically balanced growth

In the static planning problem above we saw that since preferences do not allow for aggregation, aggregate labor supply depends on the joint distribution of planner weights and labor efficiency units. Hence, in general, the dynamics of the economy will depend on the distribution. In this section we examine the asymptotic dynamics of the planner's solution. Note that the asymptotic behavior of the economy is also of interest when it comes to solving numerically for the decentralized equilibria, which in part motivates this analysis. So we ask whether the heterogeneity "washes out" in the long run or whether the asymptotic dynamics of aggregate prices and quantities even asymptotically depend on the distribution.

As we will see below the answer to that question depends on the distribution of planner weights and labor efficiency units. To prepare the analysis, consider once again the "representative-agent" economy without capital accumulation where all individuals have the same $\omega = \tilde{\omega}$ and utilitarian planner weights. In this case, the planner sets $h(\iota) = 1$ if $\tilde{\omega}\lambda < \psi^{-\frac{1}{\sigma-1}}$, an interior solution $h(\iota) = (\tilde{\omega}\lambda)^{-\frac{\sigma-1}{1/\theta+\sigma}} \psi^{-\frac{1}{1/\theta+\sigma}}$ if $\psi^{-\frac{1}{\sigma-1}} \leq \tilde{\omega}\lambda < \left(\psi \underline{h}^{1/\theta+\sigma}\right)^{-\frac{1}{\sigma-1}}$, and a fraction we define as s of the population working at $\underline{h}$ is set according to $s = (\tilde{\omega}\lambda)^{-\frac{\sigma-1}{\sigma}} \left(\frac{1+1/\theta}{\psi} \underline{h}^{-(\sigma+1/\theta)}\right)^{\frac{1}{\sigma}}$ if $\tilde{\omega}\lambda \geq \left(\psi \underline{h}^{1/\theta+\sigma}\right)^{-\frac{1}{\sigma-1}}$. Since λ goes to infinity, asymptotically, we will therefore be in the regime where only the extensive margin is active. Hence, asymptotically, as productivity grows at gross rate $\gamma > 1$, labor supply will shrink at gross rate $\gamma^{-\frac{\sigma-1}{\sigma}}$ (because the fraction s that works $\underline{h}$ shrinks at this rate) and consumption of all individual will grow at gross rate $\gamma^{\frac{1}{\sigma}}$.

Finding 11. *In the case where all households are identical and have identical planner weights, the economy will converge to a balanced growth path where the only working agents work at $\underline{h}$ and the size of this group goes to zero with the growth rate $\gamma^{-\frac{\sigma-1}{\sigma}}$; consumption per capita, on the other hand grows at rate $\gamma^{\frac{1}{\sigma}}$.*

We will see below that there are non-trivial distributions of planner weights and labor efficiency units the economy will converge asymptotically to the same point as just discussed, i.e., to representative-agent behavior. For some other distributions, however, the long-run outcome is qualitatively different and even the long-run growth rates depend on the distributional properties.

An important aspect is the question whether the distribution in the ratio $\frac{\omega}{\chi(\iota,\omega)}$ is bounded from above. If this is the case the planner would like to make everyone work less than $\underline{h}$ in finite time. Consequently, we know that asymptotically only the extensive margin is changing and we can describe the asymptotic growth rates of the economy in the following proposition, which is proved in the appendix, Section 7.1.

Proposition 1. *Suppose $\chi(\iota,\omega)$ is strictly positive for all individuals and we have a finite $\frac{\omega}{\chi(\iota,\omega)}$ for all individuals, and we have $\sigma > 1$ and $\gamma > 1$. Then, asymptotically all the individuals either work 0 or $\underline{h}$. The share working, s, is asymptotically zero and is asymptotically shrinking at constant rate, i.e.,*

$$\lim_{t\to\infty} \frac{s_{t+1}}{s_t} = \gamma^{-\frac{\sigma-1}{\sigma}} < 1. \tag{56}$$

Consumption of all the individuals and the total labor input in efficiency units, n, is growing asymptotically at the constant rate

$$\lim_{t\to\infty} \frac{n_{t+1}}{n_t} = \lim_{t\to\infty} \frac{C_{t+1}}{C_t} = \lim_{t\to\infty} \frac{c_{t+1}(\iota,\omega)}{c_t(\iota,\omega)} = \gamma^{\frac{1}{\sigma}} > 1. \tag{57}$$

So as long as the distribution of $\frac{\omega}{\chi(\iota,\omega)}$ is bounded, the asymptotic dynamics of the aggregate economy look identical to the one of a representative agent economy.

We illustrate the result just obtained with a truncated Pareto distribution in ω and a utilitarian welfare function. I.e., we suppose that the density in ω_i is given by the following truncated function:

$$\Gamma(\omega)d\omega = \kappa \frac{\omega_0^\kappa}{\omega^{\kappa+1}} \left(\frac{1}{1 - \left(\frac{\omega_0}{\bar{\omega}}\right)^\kappa} \right), \tag{58}$$

for $\omega_0 \leq \omega \leq \bar{\omega}$ (and zero otherwise). We have $\kappa > 0$. (With $\kappa \to -1$, this would nest a uniform distribution.) Figures 4 shows the resulting behavior where it can be seen how the consumption growth rate and the ω_3 thresold converge to their asymptotic level.

3.5.2 Exact balanced growth with infinite endowment support

With an unbounded $\frac{\omega}{\chi(\iota,\omega)}$, the working behavior need not even asymptotically be restricted to the extensive margin and the distribution can matter even for the asymptotic growth rates of the economy. Let us illustrate this point with the following example. So consider a planner with utilitarian weights. We assume a(n untruncated) Pareto distribution of household productivity levels ω such that the marginal distribution in ω can be expressed as

$$\Gamma(\omega)d\omega = \kappa\frac{\omega_0^\kappa}{\omega^{\kappa+1}}, \tag{59}$$

for $\omega \geq w_0$ (and zero otherwise). We have $\kappa > 1$. The following proposition now obtains.

Proposition 2. *A utilitarian planning solution given that the individual productivity distribution is Pareto, with the parametric form just described, leads to exact balanced growth where the fraction of people working grows at gross rate* $\gamma^{\frac{-\kappa(\sigma-1)}{1+\sigma(\kappa-1)}} < 1$ *and per-capita consumption grows at* $\gamma^{\frac{\kappa}{1+\sigma(\kappa-1)}} > 1$.

This is a striking example since it shows how in general even the long-run (asymptotic) growth *rates* of aggregate variables can depend on the distribution of productivities across the household distribution. I.e., this model is an endogenous-growth model. Similarly, planner weights too will matter in general. It also shows that even asymptotically the intensive margin of labor supply can change. These results require unbounded support, but the behavior of an economy with very large finite support will of course be very similar for a very long period of time even if its asymptotic behavior will be qualitatively different.

3.5.3 Exact balanced growth more generally

The asymptotic growth rates we constructed in Section 3.5.1 can also be reached in finite time provided some conditions hold. This insight is closely related to one that will be used in the next section as well in the context where there is capital accumulation (and where we look at market outcomes).

So consider a very simple: the representative-agent economy. In this case we will reach $\tilde{\omega}\lambda \geq \left(\psi \underline{h}^{1/\theta+\sigma}\right)^{-\frac{1}{\sigma-1}}$ in finite time and all the households work either $\underline{h}$ or 0. Furthermore, it can be shown that the fraction working, s, decreases at constant gross growth rate $\frac{s_{t+1}}{s_t} = \gamma^{-\frac{\sigma-1}{\sigma}} < 1$ and consequently income, marginal utility of consumption etc. all also grow at constant rates.

Similar results obtain for much more elaborate endowment distributions: when λ is high enough, all agents will be either at $\underline{h}$ or 0 hours of work. To obtain a balanced growth path where some agents are always at interior solutions does, however, rely on an unbounded distribution either of productivities (as in the previous example with a Pareto distribution) or relative planner weights.

3.6 Balanced growth in a market equilibrium with capital

We now move to another infinite-horizon economy which is decentralized and which has capital accumulation. The task in this section is thus to provide conditions under which this economy admits an exact balanced growth path. So suppose that households are heterogenous in their initial asset level a and their labor market efficiency ω. Households solve the following problem:

$$\max_{\left\{c_t, \{\pi_t(h)\}_{h\in\mathcal{H}}\right\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \left[\frac{c_t^{1-\sigma}}{1-\sigma} - \frac{\psi}{1+1/\theta}\int_{h\in\mathcal{H}} \pi_t(h) h^{1+1/\theta} dh\right] \tag{60}$$

subject to

$$\mathcal{H} = \{0, [\underline{h}, 1]\} \tag{61}$$

$$0 \leq \pi_t(h) \leq 1, \ \forall h \in \mathcal{H}, \tag{62}$$

$$\int_{h\in\mathcal{H}} \pi_t(h)\, dh = 1, \tag{63}$$

and

$$a(1+r_0)+\sum_{t=0}^{\infty}p_t\left[T_t+\omega w_t(1-\tau_t)\int_{h\in\mathcal{H}}\pi_t(h)h\,dh\right]=\sum_{t=0}^{\infty}p_tc_t. \tag{64}$$

We have $\frac{p_t}{p_{t+1}}=1+r_{t+1}$ and $p_0=1$. Here T is a governmental transfer and $\tau\in[0,1)$ is a proportional tax on labor income.

When we refer to decisions made as a result of the maximization problem just stated, we may use the additional argument (a,ω). In the following let us also define a household's wealth level in period t as follows:

$$a_{t+1}(a,\omega)=a_t(a,\omega)\left[1+r_t\right]+\omega w_t(1-\tau_t)\int_{h\in\mathcal{H}}\pi_t(h;a,\omega)h(a,\omega)dh+T_t-c_t(a,\omega), \tag{65}$$

where

$$a_0(a,\omega)=a. \tag{66}$$

In this economy an equilibrium is a sequence $\left\{\left(c_t(a,\omega),\{\pi_t(h;a,\omega)\}_{h\in\mathcal{H}},a_{t+1}(a,\omega)\right)_{a,\omega}\right\}_{t=0}^{\infty}$, a sequence of factor prices $\{r_t,w_t\}_{t=0}^{\infty}$, and an aggregate capital and labor input $\{k_t,n_t\}_{t=0}^{\infty}$, where $k_t=\int_a\int_\omega a_t(a,\omega)\Gamma(da,d\omega)$ and $n_t=\int_a\int_\omega\omega\int_{h\in\mathcal{H}}\pi_t(h;a,\omega)h(a,\omega)\,dh\Gamma(da,d\omega)$, such that:

1. the sequence $\left\{c_t(a,\omega),\{\pi_t(h;a,\omega)\}_{h\in\mathcal{H}}\right\}_{t=0}^{\infty}$ solves the household problem for all a,ω and for given factor prices;

2. the factor prices are consistent with perfect competition on the firm side, i.e., $r_t=F_1(k_t,\gamma^tn_t)-\delta$ and $w_t=\gamma^tF_2(k_t,\gamma^tn_t)$, $\forall t$;

3. and the budget of the government is balanced, i.e.,

$$T_t=\int_a\int_\omega\tau\omega w_t\int_{h\in\mathcal{H}}\pi_t(h;a,\omega)h(a,\omega)dh\Gamma(da,d\omega). \tag{67}$$

Let us define a balance-growth equilibrium as follows:

Definition 1. *A balanced-growth equilibrium is an equilibrium along which all the variables grow at constant rates.*

Note, of course, that different variables can grow at different rates, which may involve both increasing, decreasing, and constant variables.

In this economy, for a constant tax rate τ, there exists an exact balanced-growth equilibrium if $\sigma > 1$, $\gamma > 1$, $\omega > 0$, a is bounded from below, and an appropriate initial condition is met. We have:

Proposition 3. *Suppose the labor tax rate is constant and equal to τ. Furthermore suppose that a and ω are distributed such that*[5]

$$c_t(a,\omega) > \left(\frac{\omega w_t(1-\tau)(1+1/\theta)}{\psi \underline{h}^{\frac{1}{\theta}}}\right)^{\frac{1}{\sigma}}, \ \forall a, \omega. \tag{68}$$

Then, there exists a balanced growth path along which for all a, ω:

$$\frac{c_{t+1}(a,\omega)}{c_t(a,\omega)} = \frac{a_{t+1}(a,\omega)}{a_t(a,\omega)} = \frac{k_{t+1}}{k_t} = \frac{T_{t+1}}{T_t} = \gamma^{\frac{1}{\sigma}} > 1, \tag{69}$$

The fraction working $\underline{h}$ is for all a, ω either zero or $0 < s_t(a,\omega) < 1$, where

$$\frac{s_{t+1}(a,\omega)}{s_t(a,\omega)} = \frac{n_{t+1}}{n_t} = \gamma^{-\frac{\sigma-1}{\sigma}} < 1. \tag{70}$$

The interest and wage rates are given by

$$r_t = \frac{\gamma}{\beta} - 1 = F_1\left[\left(\frac{k}{n}\right)^{\star}, 1\right] - \delta, \ \forall t, \tag{71}$$

[5]Note that this condition is stated in terms of endogenous variables, i.e., $c_t(a,\omega)$. It is possible to state sufficient conditions on primitives such that this condition is met but for brevity we omit this statement. It is rather complex because of the lack of aggregation: the condition has to be stated indirectly as no closed-form solutions are available for factor prices in terms of primitives. The key conditions involve the minimum a and ω values not to be not too small and initial labor productivity being large enough.

where $\left(\frac{k}{n}\right)^\star$ *is the balanced-growth capital-labor ratio, and*

$$w_t = \gamma^t \left[F\left[\left(\frac{k}{n}\right)^\star, 1\right] - \left(\frac{\gamma}{\beta} - 1 + \delta\right)\left(\frac{k}{n}\right)^\star\right]. \tag{72}$$

Finally, the individual consumption level is given by

$$c_t(a,\omega) = s_t(a,\omega)\omega w_t(1-\tau) + a_t(a,\omega)\left[\frac{\gamma}{\beta} - \gamma^{\frac{1}{\sigma}}\right] + T_t, \tag{73}$$

where the fraction working $\underline{h}$ *of a particular family,* $s_t(\omega, a)$*, is given by*

$$s_t(a,\omega) = \max\left[(\omega w_t(1-\tau))^{-\left(1-\frac{1}{\sigma}\right)}\left(\frac{(1+1/\theta)}{\psi \underline{h}^{\frac{1}{\theta}}}\right)^{\frac{1}{\sigma}} - a_t(a,\omega)\frac{\frac{\gamma}{\beta} - \gamma^{\frac{1}{\sigma}}}{\omega w_t(1-\tau)} - \frac{T_t}{\omega w_t(1-\tau)}, 0\right]. \tag{74}$$

The proof of the proposition is available in the appendix, Section 7.1. The intuition why balanced growth obtains follows that presented in Sections 3.5.1 and 3.5.3.

Note that we can use this proposition to find out what the long-run interest rate will be, as it gives us long-run consumption growth. In solving numerically for transitional dynamics, a topic we turn to next, this insight is important.

4 The quantitative model

4.1 Calibration

The most important parameters to calibrate are the preference parameters σ and θ, since they not only control the changes in hours worked over time but also the cross-sectional implications of changes in (ω, a) on hours worked. For now we calibrate the model with $\sigma = 1.8$ and $\theta = 1.0$. As a discount rate β we use 0.96. Finally, ψ can be normalized such that average hours worked are 0.72 in 2013.

The other important element is the joint distribution of wealth and labor productivity. For this we rely on the information for the joint distribution of wealth, total income plus transfers, as well as labor earnings provided by Kuhn and Rios-Rull (2016). Kuhn and Rios-Rull (2016) provide the information for 100 groups of household split up into deciles in terms of earnings and wealth (and the corresponding density) for the U.S. in 2013. We drop observations that do not have strictly positive labor earnings. An important step is to further split up labor earnings into labor productivity and hours. For this we exploit the balanced-growth relationship between consumption, income, and wealth in a framework without labor invisibilities. With $\underline{h} = 0$ there exists a balanced-growth path along which the wealth of all households grows at rate $\gamma^{\frac{1+\theta}{1+\sigma\theta}}$ (see Boppart and Krusell (2016)). Along this balanced growth path the period budget constraint can be expressed as

$$c_t(a, \omega) = a_t(a, \omega)r_{t+1} + \omega w_t(1 - \tau_t)h_t(a, \omega) + T_t - (\gamma^{\frac{1+\theta}{1+\sigma\theta}} - 1)a_t(a, \omega). \tag{75}$$

Hence, we can impute consumption for each group for given preference and technology parameters from income, $a_t(a, \omega)r_{t+1} + \omega w_t(1 - \tau_t)h_t(a, \omega) + T_t$, and wealth $a_t(a, \omega)$. Given the information on consumption we can then calculate hours from the consumption-leisure first-order condition

$$c_t(a, \omega)^{-\sigma}\frac{w_t}{\psi}(1 - \tau_t)h_t(a, \omega)\omega = h_t(a, \omega)^{1+1/\theta}. \tag{76}$$

Finally given the information on hours we can infer the distribution of relative labor productivities ω. This gives us an initial distribution of households shown in figure 5.

For the aggregate production function we assume a Cobb-Douglas functional form with an output elasticity of capital $\alpha = 1/3$. We set the exogenous rate of technical change to $\gamma = 1.03$ and we assume a depreciation rate δ of 0.10.

4.2 Results

Figure 6 shows the results for some key variables: aggregate consumption, average hours, capital, and the gross interest rate. Consumption and capital grow at steady

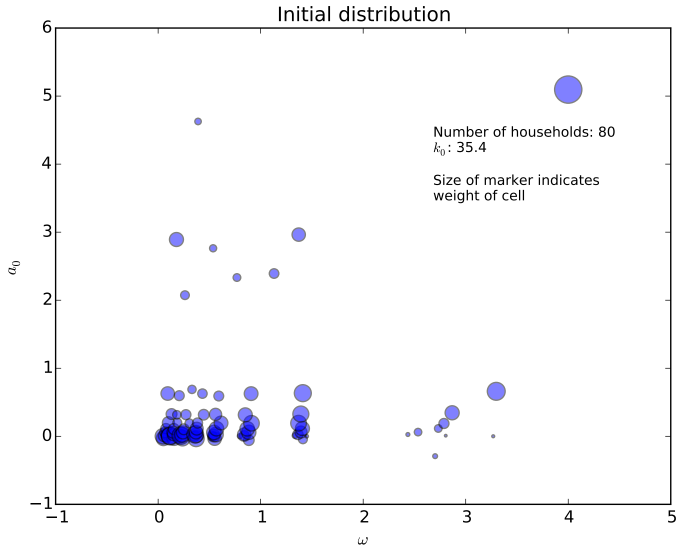
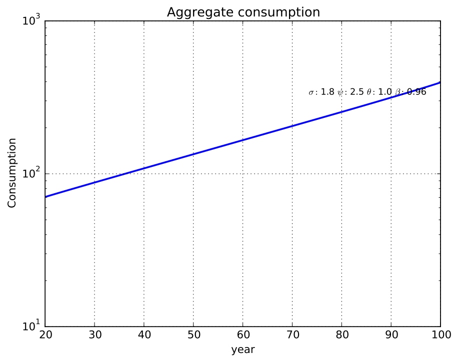


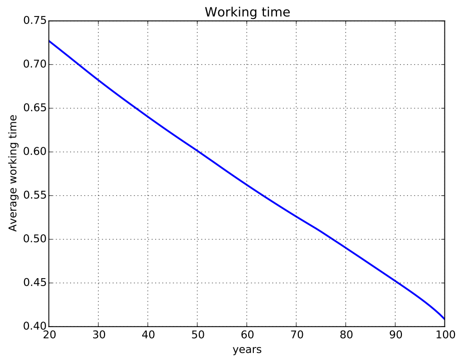
Figure 5: Initial distribution of households along the initial asset and productivity dimension, for source please refer to text.

rates, i.e., are approximately linear on a logarithmic scale suggesting that consumption grows at roughly 2.1% per year. Furthermore, the interest rate is rather flat over time. This is highlights that our theory produces rich and interesting individual dynamics whereas the economy is at the aggregate level approximately in line with balanced growth. Since the income effect dominates the substitution effect and wages grow, overall, hours worked fall monotonically over time by about 40% over the 80-year period considered. This decrease at the aggregate level is quantitatively in line with Boppart and Krusell (2016).

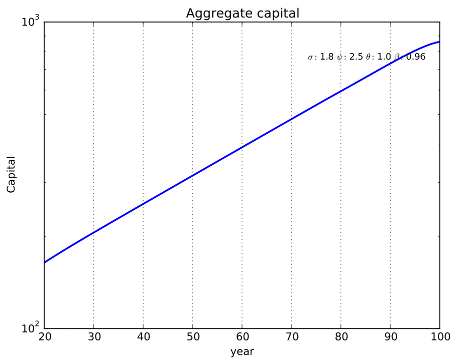
More interesting, however, and the main focus of this paper are the changes in hours worked in the cross-section of households. Figure 7 shows the dynamics of hours worked for five selected households. Wealthy and productive households are in the beginning of the plotted period already at an interior hours choice and consequently their hours worked fall monotonically as wages increase over time. Poorer



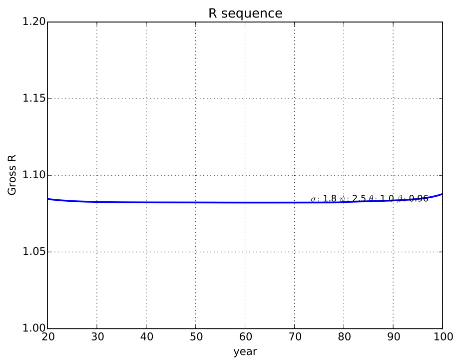
(a) Aggregate consumption growth



(b) Average hours



(c) Aggregate capital growth



(d) Gross interest rate

Figure 6: The evolution over the medium run for some key variables

and less productive households however are (at least initially) constrained by the full time budget and work $h = 1$. For the calibration it turns out that some very unproductive households will always be in this corner solution over the considered time period (see the orange line) whereas others (see the blue line) eventually switch to a situation with an interior hours choice. Towards the end of the period, high productive and wealthy households reach the $\underline{h}$ of 0.2, so they enter a period of constant hours at this lower level (see the red line). In the future these rich households will be the first ones to eventually decide to completely abstain from the labor market.

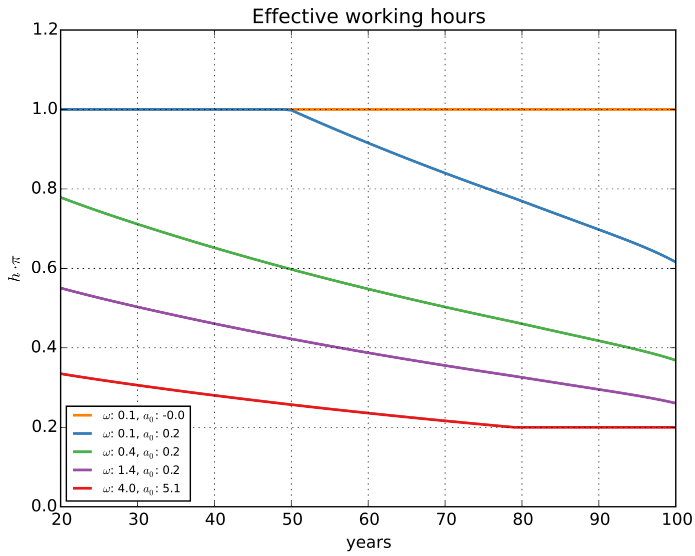


Figure 7: Trends in hours for different households

Figure 8 shows the working hours for the full distribution of households for three points in time. The size of the marker indicates hours worked in each time period. In the first time period households with low initial assets and low productivity are working full time (indicated by red color) while the households with more initial assets and/or higher productivity are working at an interior solution. The subsequent graphs show how working hours decline for different households, except for the least productive households with low initial assets who continue to work full time. Furthermore, we again see the households with low productivity are switching earlier on to an interior solution.

Figure 9 shows the resulting Gini coefficients for hours and earnings. As can be seen, the inequality in hours is increasing over time. The households working according to their interior solution are decreasing their labor supply over time, while the households working maximum hours keep on doing so; hence the difference in working hours is increasing. In the future however this past pattern will revert since the highly productive households will be stuck at $\underline{h}$ for a while whereas the low-productivity households (with high hours worked) still decrease their labor hours according to the interor solution.

Earnings inequality on the other hand first decreases, due to the fact that it is the households with the higher ω that decrease their labor supply, and hence the relative difference in earnings is decreasing. However, at some point the high ω households hit the lower bound of work and keep on working the same amount of hours, while the households who are still in the interior solution keep on decreasing their labor supply.

5 Robustness

Some theoretical extensions require discussion. One is to examine market equilibria where convexification is not allowed. Another is to go beyond the MaCurdy preference formulations. Each of those cases involve rather non-trivial analysis and we only briefly report our findings here. We refer interested readers to an online appendix available on each of our websites.

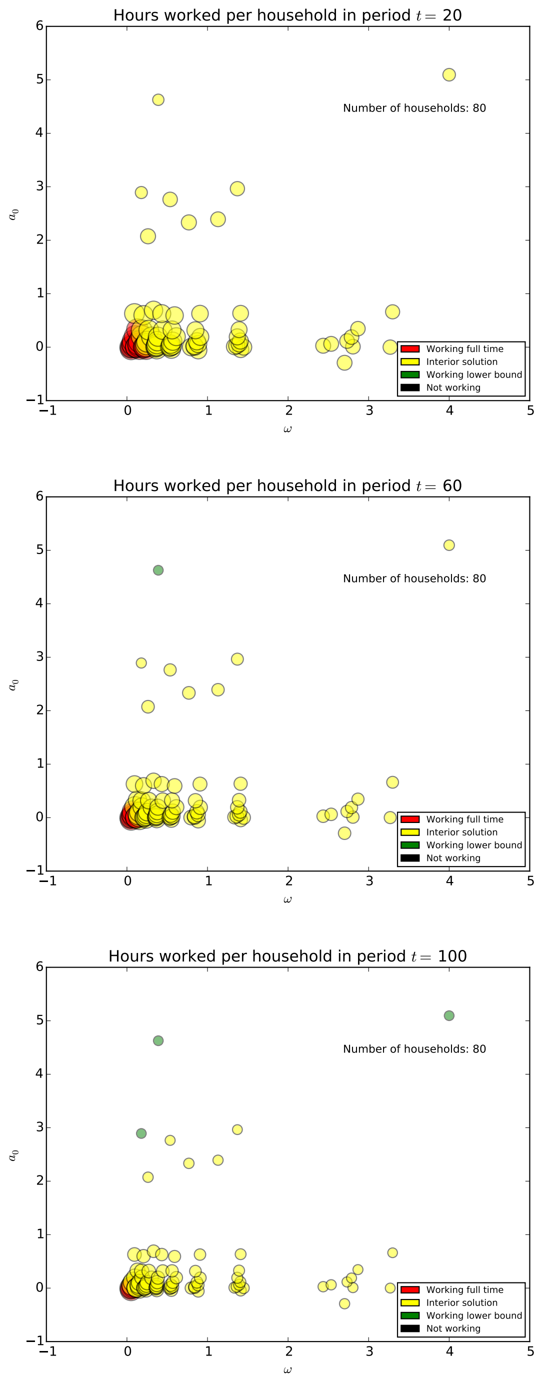


Figure 8: The working decisions for the individual household for three different points in time. Size of marker indicates hours worked.

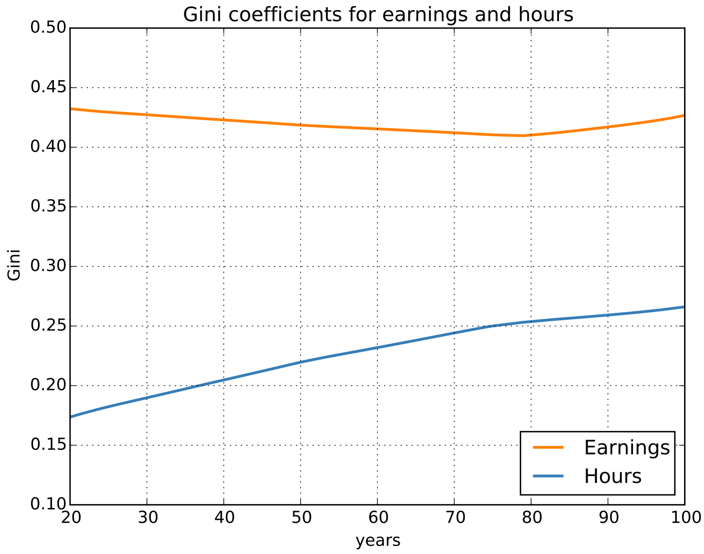


Figure 9: Gini coefficients for earnings and hours.

5.1 Equilibria without lotteries

Two issues arise when convexification is not allowed. One is positive: will equilibria have fundamentally different properties. Another is normative: is the equilibrium without lotteries constrained optimal, in the appropriate sense? We briefly visit these issues first in a static and then in a two-period model. Our aim is not to provide a fully general treatment but to make a few points. We will restrict attention to the special case where there is only an extensive-margin choice: $\underline{h} = 1$.

5.1.1 The static model

In a one-period model where there is no initial asset inequality, all households trivially choose to work. If a household has an asset endowment a then the decision whether or not to work is determined by $u(a, 0)$ versus $u(a + \omega, 1)$, which will define a downward-sloping cutoff line in (a, ω) space northwest of which the household does not work. Is this equilibrium constrained-optimal in the sense that there is no alternative no-convexification allocation in which no agent can be made better off without making another agent worse off? The answer is yes. Convexification is necessary in order to improve on the outcome. Constrained suboptimality can arise in models where markets are restricted not because agents are poor maximizers but because the restrictions on their ability to trade interact with endogenous equilibrium prices, but in the static model there is no endogenous equilibrium price.

Before proceeding, let us also remind the reader of the finding in Section 3.2.3: any lottery equilibrium allocation, defined in terms of its distribution of consumption and hours worked across the population, can be obtained without lotteries, so long as initial assets are redistributed appropriately. Conversely, we can map equilibria without lotteries here back to lottery equilibria.

5.1.2 A two-period model

In a multi-period model household behavior is harder to analyze. We use a two-period model as an illustration. The utility of a household with no initial assets can be rewritten as a function of ω, h_0, and h_1, using the lifetime budget constraint:

$$\omega^{1-\sigma} \frac{p^{\sigma}}{1-\sigma} \left(h_0 + h_1 \frac{\gamma}{1+r} \right)^{1-\sigma} - \frac{\psi}{1+\frac{1}{\theta}} \left(h_0 + \beta h_1 \right) \tag{77}$$

where we define p as the following:

$$p = 1 + \beta^{\frac{1}{\sigma}} (1+r)^{\frac{1-\sigma}{\sigma}}$$

Here, $h_0 = h_1 = 0$ is not an option, so three cases remain. In the case with lotteries, the timing of work in this model depended crucially on $\beta(1+r)/\gamma$: if it was high, the household preferred working in period 0 over working in period 1. Here, matters are less simple. There are, in general, three cutoff-values for ω determining how much and when to work. The qualitatively new case is a cutoff level for productivity such that an individual, who is rather productive and who will therefore only work in one period, chooses to switch from working in the first period to working in the second period instead. This level of productivity is defined by the solution to

$$\omega^{1-\sigma}\frac{p^{\sigma}}{1-\sigma}-\frac{\psi}{1+\frac{1}{\theta}}=\omega^{1-\sigma}\frac{p^{\sigma}}{1-\sigma}\left(\frac{\gamma}{1+r}\right)^{1-\gamma}-\beta\frac{\psi}{1+\frac{1}{\theta}}, \tag{78}$$

which simply sets the utility value of the two options to be equal. The resulting value, ω^*, satisfies

$$\omega_s^*=\left[p^{-\sigma}(1-\sigma)\cdot\frac{(1-\beta)\frac{\psi}{1+\frac{1}{\theta}}}{1-\left(\frac{\gamma}{1+r}\right)^{1-\sigma}}\right]^{\frac{1}{1-\sigma}}, \tag{79}$$

where we are interested in the case where $1+r>\gamma$. In this case, postponing work gives a lower present value of consumption. The key now is that if a household is productive enough, i.e., $\omega>\omega^*$, then working in the first period is preferable. The intuition for this result is that the disutility from working in the second rather than the first period is simply a function of β, independently of ω, whereas the resulting change in the utility of consumption depends on ω. When a household is productive enough, postponing work gives higher utility, since the marginal loss in present-value income is worth less in utility terms than the discounting gain from postponing work.

The analysis shows that, unlike in the case with lotteries, the timing of working can depend on the level of a household's productivity. It is also possible to show that there are indeed equilibrium interest rates such that $1+r>\gamma$ is met. Household's rankings of when to work, even under a constant interest rate, can by the same argument be very nontrivial with more periods. By the same token, as interest rates change by a small amount, the total number of hours worked in the population can change abruptly even if the distribution over household productivities is continuous.[6] This makes equilibrium construction much more challenging in multi-period models without convexification.

[6] Consider making $1+r$ move from just below to just above γ.

It would be interesting to examine constrained optimality in the two-period model, where there is a nontrivially determined interest rate. Our conjecture, that we have not verified yet, is that the equilibrium will not always be constrained-optimal.

5.2 Non-additive utility

In our benchmark setting, we use MaCurdy (1981)'s formulation of the utility function, which features additive separability in consumption and hours. This formulation is convenient because the marginal utility of consumption does not depend on whether or not the household works. However, there are many other non-additive utility functions where the income effect is stronger than the substitution effect. In Boppart and Krusell (2016), MaCurdy's formulation is a special case but there are other, interesting non-additive cases as well. We now discuss these briefly.

We focus on a homogeneous-agent case for simplicity, i.e., a version of Rogerson (1988), though not the special KPR case he considers. A fundamental difference in the non-additive case with lotteries is that consumption behavior of the working and non-working households are qualitatively different. A point in case is that where lotteries are needed, though we stress—for the same reason as explained in Section 3.2.3—any ex-post lottery equilibrium outcome can be generated as a deterministic equilibrium with appropriately chosen initial wealth levels of households. Thus consider a social planner who solves

$$\max_{c_0, c_1, \pi(1)} \pi(1)u(c_1, 1) + (1 - \pi(1))u(c_0, 0) \tag{80}$$

subject to $c_0 \geq 0$, $c_1 \geq 0$, $\pi(1) \in [0, 1]$, and

$$\pi c_1 + (1 - \pi(1))c_0 = \lambda\pi(1). \tag{81}$$

It is straightforward to solve this problem for a given function u in the desired class and then to examine the result of changing λ. We will, in particular, take a parametric class of u functions and find the asymptotic (gross) growth rates of the choice variables: g_π, g_1, and g_0, with the obvious notation. Recall that in the MaCurdy benchmark in this paper, these rates were—in the homogeneous-agent case—given by $\gamma^{\frac{1-\sigma}{\sigma}}$, $\gamma^{\frac{1}{\sigma}}$, and $\gamma^{\frac{1}{\sigma}}$, respectively: the consumption growth rates of working and non-working people are the same.

For the utility function, we use the following parametric form:

$$u(c,h) = \frac{c^{1-\sigma}\left[1 - a\left(hc^{\frac{\nu}{1-\nu}}\right)^b\right]^d - 1}{1-\sigma}, \tag{82}$$

where we have

$$a \equiv \frac{\psi(1-\sigma)}{1+\frac{1}{\theta}}, \quad b \equiv 1 + \frac{1}{\theta}, \quad \text{and} \quad d \equiv (1-\sigma)\kappa. \tag{83}$$

This form nests many well-known cases. We assume throughout that $\sigma > 1$ and $\kappa < 0$ and define $p \equiv \frac{\nu}{1-\nu}$ and $x = (1-\sigma)pb\kappa$ to save on notation. We will simply report the results here; for details, see our online appendix. It turns out that the asymptotic behavior is qualitatively different depending on the sign of $(1-\sigma)(1+pb\kappa)$, so that there are three distinct cases, as shown in Table 1. We see

	Sign of $(1-\sigma)(1+pb\kappa)$		
Growth rate	**Positive**	**Negative**	**=0**
g_π	$\gamma^{\frac{1-\sigma}{\sigma}}$	$\gamma^{-\frac{x}{\sigma}} = \gamma^{-\frac{(1-\sigma)pb\kappa}{\sigma}}$	$\gamma^{\frac{1-\sigma}{\sigma}} = \gamma^{-\frac{x}{\sigma}}$
g_0	$\gamma^{\frac{1}{\sigma}}$	$\gamma^{1-\frac{x}{\sigma}} = \gamma^{1-\frac{(1-\sigma)pb\kappa}{\sigma}}$	$\gamma^{\frac{1}{\sigma}} = \gamma^{1-\frac{x}{\sigma}}$
g_1	0	γ	$\gamma^{\frac{1}{1+pb}} = \gamma^{\frac{\kappa}{\kappa-1}}$

Table 1: Asymptotic growth rates

from the table that the benchmark case we consider is knife-edge in that $g_0 = g_1$. However, it is robust—in the sense that there is a large part of the parameter space with this property—in the characterization of g_π and g_0. For the working group, however, we see in the table that the asymptotic consumption growth rate can take on several values depending on parameters. This, however, does not have aggregate consequences in the case where $(1-\sigma)(1+pb\kappa) \geq 0$ because then the asymptotic share of total consumption going to this vanishing group is zero. We also see, by examining the middle case, $(1-\sigma)(1+pb\kappa) < 0$, that it is possible to have the vanishing group see its consumption growth rate rise at the pace of productivity, i.e., very fast, with a stable total share of overall production.

In sum, although qualitatively different cases are uncovered in the non-additive case, the MaCurdy formulation used as a benchmark in the paper is robust in terms of aggregate observables.

6 Concluding remarks

In this paper, we have studied a hypothetical future with ongoing productivity growth and a population of households that differ in labor productivity and wealth and that have preferences such that the income effect of productivity growth on labor supply exceeds the substitution effect. The set of results we obtain apply under the rather restrictive assumptions we make and we will now briefly comment on extensions that we believe would be important.

We focus on productivity differences in a stylized way; obviously, we would like to introduce a life-cycle and other "visible" characteristics of households, along with shocks. For example, in our dynastic setting it is not realistic to assume that the skill differences are permanent. It would be straightforward, however, to include shocks to productivity—we believe the extension would be simple—as long as one assumes there to be perfect insurance markets against these shocks. One particularly interesting hypothesis is one where skill differences become larger over time, i.e., that where the growing wage inequality continues.

It would be a more challenging, but very interesting, extension to have partially uninsurable shocks, as in the Bewley-Huggett-Aiyagari model. Such a model would clearly have consumers move back and forth between working and not working to some extent, but we believe that it would be feasible to analyze this kind of setting at least if lotteries are allowed. Here, one would expect the strong forces to withdraw early by high-skilled/rich households to be mitigated.

Moreover, clearly a transition path from the present to the future should take into account various other frictions. Unemployment is one. In this context, an often-discussed subject is how technological change (technology revolutions like IT) changes the demand for workers and therefore equilibrium employment. In our basic framework, employment in the long run will be determined by supply, not demand, but structural change such as technology shifts will clearly create room for worker demand to influence employment in the short to medium run.

It would be valuable to address the historical evolution of cross-sectional data on hours head on. Here, we believe that there are many unobservable components driving actual differences in hours worked, as evidenced by typically low R^2s in trying to explain this heterogeneity based on observables. This is, however, a continuation of the present project which we would like to pursue.

Finally, though our main aim with this project is to derive (and, eventually, evaluate) cross-sectional implications, there are potentially interesting implications for aggregate macroeconomic variables as well. We alluded to these already in Section 3.3: aggregates themselves as well as responses to shocks (e.g., multipliers) will change as the economy moves forward in the kind of model world we consider here. We have only began to scratch the surface of this topic here.

References

Bick, A., N. Fuchs-Schuendeln, and D. Lagakos (2016, January). How do average hours worked vary with development? cross-country evidence and implications. Working Paper 21874, National Bureau of Economic Research.

Boppart, T. and P. Krusell (2016, May). Labor supply in the past, present, and future: a balanced-growth perspective. Working Paper 22215, National Bureau of Economic Research.

Davila, J., J. H. Hong, P. Krusell, and J.-V. Rios-Rull (2012). Constrained efficiency in the neoclassical growth model with uninsurable idiosyncratic shocks. *Econometrica 80*(6), 2431–2467.

Erosa, A., L. Fuster, and G. Kambourov (2014, July). Towards a Micro-Founded Theory of Aggregate Labor Supply. Working Papers tecipa-516, University of Toronto, Department of Economics.

Heathcote, J., K. Storesletten, and G. L. Violante (2014). Consumption and labor supply with partial insurance: An analytical framework. *American Economic Review 104*(7), 2075–2126.

King, R. G., C. I. Plosser, and S. T. Rebelo (1988). Production, growth and business cycles. *Journal of Monetary Economics 21*(2), 195 – 232.

Krusell, P., T. Mukoyama, R. Rogerson, and A. Sahin (2008). Aggregate implications of indivisible labor, incomplete markets, and labor market frictions. *Journal of Monetary Economics 55*(5), 961 – 979. Carnegie-Rochester Conference Series on Public Policy: Labor Markets, Macroeconomic Fluctuations, and Monetary Policy November 9-10, 2007.

Kuhn, M. and J. Rios-Rull (2016). Update on the us earnings, income, and wealth distributional facts: A view from macroeconomics. *Federal Reserve Bank of Minneapolis Quarterly Review*.

MaCurdy, T. E. (1981). An empirical model of labor supply in a life-cycle setting. *The Journal of Political Economy*, 1059–1085.

Rogerson, R. (1988). Indivisible labor, lotteries and equilibrium. *Journal of Monetary Economics 21*(1), 3 – 16.

7 Appendix

7.1 Proofs

Proof of Proposition 1. *With productivity growth (i.e., $\gamma > 1$) and dominating income effects (i.e., $\sigma > 1$) the planner would like to decrease hours worked by all individuals. Then, under the assumption that $\frac{\omega}{\chi(\iota,\omega)}$ is finite, households either work zero or $\underline{h}$ in finite time and the share working $\underline{h}$ is asymptotically zero. An easy way to see this is to consider the planner's problem without the $\underline{h}$ constraint. In this case the first-order conditions are*

$$\mu_t = \chi(\iota,\omega)c_t(\iota,\omega)^{-\sigma} \tag{84}$$

and

$$\gamma^t\omega\mu_t = \chi(\iota,\omega)\psi\left(h_t(\iota,\omega)^\star\right)^{\frac{1}{\theta}} \tag{85}$$

where μ is the Lagrangian multiplier and $h_t(\iota,\omega)^\star$ the optimal hours choice while ignoring the $\underline{h}$ constraint. From this we see that in a given point in time consumption grows at the same rate for all individuals and also the desired hours choice $h_t(\iota,\omega)^\star$ will shrink at the same rate for all individuals (at a given point in time). Finally note *that under the assumption of a finite $\frac{\omega}{\chi(\iota,\omega)}$ for all individuals, $h_t(\iota,\omega)^\star$ converges to zero asymptotically. Hence, any positive $\underline{h}$ starts eventually binding.*

Now let us analyze the regime when all the $h_t(\iota,\omega)^\star$ are belwo $\underline{h}$. Under this regime we know that individuals with $\frac{\omega}{\chi(\iota,\omega)} \geq \frac{\psi \underline{h}^{\frac{1}{\theta}}}{\mu_t \gamma^t (1+1/\theta)}$ will be working $\underline{h}$ and all the other individuals do not work. Let us call the maximum of $\frac{\omega}{\chi(\iota,\omega)}$ in the population as $(\omega/\chi)^{max}$. Then asymptotically we must have

$$\lim_{t\to\infty} \mu_t \gamma^t = \frac{\psi \underline{h}^{\frac{1}{\theta}}}{(\omega/\chi)^{max}(1+1/\theta)}, \tag{86}$$

and s converges to zero. Why does $\mu_t\gamma^t$ have to converge to this level? First, note that $\mu_t\gamma^t > \frac{\psi \underline{h}^{\frac{1}{\theta}}}{(\omega/\chi)^{max}(1+1/\theta)}$ can be ruled out since it would imply no labor supply and consequently no consumption. Second, suppose $\mu_t\gamma^t$ stabilizes at a level strictly smaller than $\frac{\psi \underline{h}^{\frac{1}{\theta}}}{(\omega/\chi)^{max}(1+1/\theta)}$. Then, labor supply would be strictly positive and constant. This however cannot be an equilibrium since it would imply that all consumption levels grow at gross rate γ and we get a contradiction with a constant $\mu_t\gamma^t$.

Since $\mu_t\gamma^t$ stabilizes asymptotically we must have $1 = \frac{\mu_{t+1}}{\mu_t}\gamma = \left(\frac{C_{t+1}}{C_t}\right)^{-\sigma}\gamma = \gamma^{1-\sigma}\left(\frac{s_{t+1}}{s_t}\right)^{-\sigma}$, where s_t is the share still working and we used the resource constraint for the last equality. Solving this equation for the growth rates in consumption and the participation rate finally proves the statements in the proposition. ■

Proof of Proposition 2. *From the static planning analysis we know that all the households with $\omega \geq \frac{\psi}{\mu_t\gamma^t}$ work 1, household with $\frac{\psi}{\mu_t\gamma^t} > \omega > \frac{\psi \underline{h}^{1/\theta}}{\mu_t\gamma^t}$ work $\left(\frac{\mu_t\gamma^t\omega}{\psi}\right)^\theta$ hours, households with $\frac{\psi\underline{h}^{1/\theta}}{\mu_t\gamma^t} > \omega \geq \frac{\psi \underline{h}^{\frac{1}{\theta}}}{\mu_t\gamma^t(1+1/\theta)}$ work $\underline{h}$ hours, and households with $\omega < \frac{\psi \underline{h}^{\frac{1}{\theta}}}{\mu_t\gamma^t(1+1/\theta)}$ do not work at all. Then, aggregate consumption and output is given by*

$$C_t = \mu_t^{-\frac{1}{\sigma}} = \gamma^t \left[\int_{\frac{\psi \underline{h}^{\frac{1}{\theta}}}{\mu_t\gamma^t(1+1/\theta)}}^{\frac{\psi \underline{h}^{\frac{1}{\theta}}}{\mu_t\gamma^t}} \kappa \frac{\omega_0^\kappa}{\omega^\kappa} \underline{h} d\omega + \int_{\frac{\psi \underline{h}^{\frac{1}{\theta}}}{\mu_t\gamma^t}}^{\frac{\psi}{\mu_t\gamma^t}} \kappa \frac{\omega_0^\kappa}{\omega^{\kappa-\theta}} \left(\frac{\mu_t\gamma^t}{\psi}\right)^\theta d\omega + \int_{\frac{\psi}{\mu_t\gamma^t}}^{\infty} \kappa \frac{\omega_0^\kappa}{\omega^\kappa} d\omega \right]. \tag{87}$$

Solving the integrals allows us to express the equation as

$$\mu_t^{-\frac{1}{\sigma}} = \mu_t^{\kappa-1}\gamma^{\kappa t}\kappa\omega_0^{\kappa}\psi^{1-\kappa}\left[\frac{\underline{h}^{\frac{1-\kappa+\theta}{\theta}}\left((1+1/\theta)^{1-\kappa}-1\right)}{\kappa-1}+\frac{1-\underline{h}^{\frac{1-\kappa+\theta}{\theta}}}{1-\kappa+\theta}+\frac{1}{\kappa-1}\right]. \quad (88)$$

Note that the only endogenous variable in this equation is μ_t and we consequently have $\frac{\mu_{t+1}}{\mu_t} = \gamma^{\frac{-\sigma\kappa}{1+\sigma(\kappa-1)}} < 1$. We then obtain, for consumption growth,

$$\frac{C_{t+1}}{C_t} = \frac{c_{t+1}(\omega)}{c_t(\omega)} = \gamma^{\frac{\kappa}{1+\sigma(\kappa-1)}} > 1. \quad (89)$$

Finally, note that the labor market participation rate is given by 1 minus the cdf evaluated at $\frac{\psi\underline{h}^{\frac{1}{\theta}}}{\mu_t\gamma^t(1+1/\theta)}$. Under the Pareto distribution this gives us for the participation rate

$$s_t = \omega_0^{\kappa}\left(\frac{\mu_t\gamma^t(1+1/\theta)}{\psi\underline{h}^{\frac{1}{\theta}}}\right)^{\kappa}. \quad (90)$$

Consequently, the participation rate decreases at gross rate

$$\frac{s_{t+1}}{s_t} = \gamma^{\frac{-\kappa(\sigma-1)}{1+\sigma(\kappa-1)}} < 1. \quad (91)$$

But note that since we have a fat tail in the ω distribution some household will always be constraint at $h = 1$ and some households are in an interior solution and decrease their hours along the intensive margin. Hence, even asymptotically the intensive margin will matter. ■

Proof of Proposition 3. *We prove it by guessing and verifying the solution.*

Condition (68) makes sure that all households are "rich" enough such that it is optimal to randomize for each family member to work either $\underline{h}$ or 0, where the share working $s_t(a,\omega)$ is strictly small than one. Then, the household's problem can be simplified as

$$\max_{\{c_t,s_t\}_{t=0}^{\infty}}\sum_{t=0}^{\infty}\left[\frac{c_t^{1-\sigma}}{1-\sigma}-\frac{\psi}{1+1/\theta}s_t\underline{h}^{1+1/\theta}\right] \quad (92)$$

subject to

$$0 \leq s_t \leq 1, \quad (93)$$

and

$$a(1+r_0)+\sum_{t=0}^{\infty}p_t\left[T_t+\omega w_t(1-\tau_t)s_t\underline{h}\right]=\sum_{t=0}^{\infty}p_tc_t, \tag{94}$$

where s is the share working $\underline{h}$ for any given ω. We know that by assuming (68) the constraint $s_t \leq 1$ is never binding. Then, the associated Lagrangian can be written as

$$\begin{aligned}\mathscr{L} \;=\; & \sum_{t=0}^{\infty}\left[\frac{c_t(a,\omega)^{1-\sigma}}{1-\sigma}-\frac{\psi}{1+1/\theta}s_t(a,\omega)\underline{h}^{1+1/\theta}\right] \\ & +\lambda(a,\omega)\left[a(1+r_0)+\sum_{t=0}^{\infty}p_t\left[T_t-\omega w_t(1-\tau_t)s_t\underline{h}\right]-\sum_{t=0}^{\infty}p_tc_t\right]+\sum_{t=0}^{\infty}\kappa_t(a,\omega)\left[s_t(a,\omega)\right]\end{aligned}$$

The first-order conditions are:

$$\omega w_t(1-\tau)\underline{h}c_t(a,\omega)^{-\sigma}=\frac{\psi\underline{h}^{1+1/\theta}}{(1+1/\theta)}-\kappa_t(a,\omega) \tag{95}$$

$$\kappa_t(a,\omega)\geq 0, \tag{96}$$

$$s_t(a,\omega)\geq 0, \tag{97}$$

$$\kappa_t(a,\omega)s_t(a,\omega)=0, \tag{98}$$

$$\frac{c_{t+1}(a,\omega)}{c_t(a,\omega)}=\left[\beta(1+r_{t+1})\right]^{\frac{1}{\sigma}}, \tag{99}$$

and the period budget constraint

$$a_{t+1}(a,\omega)=a_t(a,\omega)\left[1+r_t\right]+\omega w_t(1-\tau)s_t(a,\omega)\underline{h}+T_t-c_t(a,\omega), \tag{100}$$

where κ_t is the multiplier attached to the non-negativity contraint on s_t.

Form the firm's optimization problem we get the standard first-order conditions that can be expressed as $r_t = F_1(k_t,n_t)-\delta$ and $w_t=\gamma^tF_2(k_t,n_t)$. The conjectured consumption growth rate implies $r_t=\frac{\gamma}{\beta}-1$ and pins down the balanced growth *capital-labor ratio* $\left(\frac{k}{n}\right)^{\star}$ *according to* (71)*. For the wage rate* (72) *then follows immediately from the firm's first-order condition and the fact that the production function exhibits constant returns to scale. Furthermore, note that since* w_t *grows at constant rate* γ *the first-order condition* (95) *implies that* $\kappa_t(a,\omega)$ *is constant along the balanced growth path for all households. Hence, the number of households that is not working at all, i.e.,* $s_t(a,\omega)=0$ *is constant over time. Furthermore, note that* (100) *is consistent with both* s*,* a *and* T *growing at the conjectured balanced growth rates. Note also that this is consistent with* k *and* n *growing at identical rates. Finally, for* $\kappa_t(a,\omega)=0$*, the first-order condition* (95) *is consistent with* (74)*.* ■