Diversity and Economic Performance in a Model with Progressive Taxation

Wei Wang and Richard M. H. Suen

Southwestern University of Finance and Economics, University of Leicester

27 February 2017

Diversity and Economic Performance in a Model with Progressive Taxation

Wei Wang* Richard M. H. Suen†

First Version: 30th October, 2015.

This Version: 27th February, 2017.

Abstract

This paper examines the relation between diversity and long-term economic performance in a dynamic general equilibrium model where consumers differ *ex ante* in time preference and labour productivity. We show that changing the distribution of these characteristics will affect the steady state by distorting the composition of aggregate labour supply. The exact nature of this effect depends on the shape of the individual labour supply function. Changing the distribution of time preference will also affect the distribution of marginal tax rates across individuals. The aggregate outcome of this is determined by the concavity or convexity of the marginal tax function.

Keywords: Consumer Heterogeneity, Progressive Taxation, Endogenous Labour Supply.

JEL classification: D31, E62.

*Research Institute of Economics and Management, Southwestern University of Finance and Economics, Gezhi Building 1205, Liulin Campus, Chengdu, Sichuan, P. R. China, 611130. Email: *sieglindwang@hotmail.com*

†Corresponding Author: School of Business, Economics Division, University of Leicester, Leicester LE1 7RH, United Kingdom. Phone: +44 116 252 2880. Email: *mhs15@le.ac.uk*

1 Introduction

Is a more heterogeneous population beneficial or harmful to long-term economic performance? What role does redistributive policy, such as progressive taxation, plays in this matter? In this paper, we address these questions in a dynamic general equilibrium model where consumers differ *ex ante* in time preference and labour productivity.[1] Our analysis focuses on how diversity in these consumer characteristics will affect long-run economic outcomes.

The economic implications of diversity have long been a subject of empirical research.[2] Several recent studies have provided evidence on the positive effect of ethnic and cultural diversity on productivity and economic growth (e.g., Ottaviano and Peri, 2006; Ager and Brückner, 2013; Trax *et al.*, 2015; Alesina *et al.*, 2016).[3] In contrast, there have been very few theoretical research on this timely and important issue. This lack is somewhat surprising, given the widespread use of heterogeneous-agent models in macroeconomics. The present study provides the first attempt to apply this kind of model to analyse the economic effects of diversity. More specifically, we adopt a similar deterministic framework as in Sarte (1997), Li and Sarte (2004), Carroll and Young (2009, 2011) and Angyridis (2015). In this type of model, *ex ante* heterogeneity is the root of income and wealth inequality.[4] Progressive taxation comes into play by distorting prices and incentives, which in turn influences how *ex ante* heterogeneity translates into *ex post* economic inequality. The present study adds to this line of research in two ways: First, we examine how changes in the distribution of consumer characteristics will affect the steady state of the model economy. In particular, our model takes into account heterogeneity in both time preference and labour productivity, and their effects are considered separately.[5] This type of analysis has not been previously undertaken. Second, unlike most of the existing studies, we do not confine ourselves to specific parameterised form of the fundamentals (i.e., utility function, production technology and progressive tax schedule). Instead, most of our results (except those in Section 4.3) are obtained

[1] We are agnostic about the origin of these differences, which can be due to racial, cultural, physiological or other reasons. Throughout this paper, we will treat the terms "diversity" and "*ex ante* heterogeneity" as synonymous.

[2] For extensive survey of this literature, see Alesina and La Ferrara (2005) and Alesina *et al.* (2016).

[3] The analysis in Ottaviano and Peri (2006), Ager and Brückner (2013) and Trax *et al.* (2015) are based on micro-level data from developed countries, such as Germany and the United States. Alesina *et al.* (2016), on the other hand, conduct cross-country comparisons using aggregate level data from 195 countries. Other cross-country studies, such as Easterly and Levine (1997) and Collier and Gunning (1999), focus on African countries and find a negative relation between ethnic diversity and economic growth.

[4] The implicit assumption is that there is perfect consumption insurance so that individuals' choices are not affected by idiosyncratic risks. Keane and Wolpin (1997) and Huggett *et al.* (2011) argue that predetermined differences in consumer characteristics are more important than idiosyncratic risks in explaining the dispersion in lifetime wealth and lifetime utility.

[5] Time preference heterogeneity has been previously considered in Sarte (1997), Li and Sarte (2001), Carroll and Young (2011), Suen (2014) and Angyridis (2015) among others. The empirical evidence on this type of heterogeneity is reviewed in Frederick *et al.* (2002).

based on some generic properties of these fundamentals.

Our main findings can be summarised as follows: In terms of labour productivity heterogeneity, the effects of greater diversity are rather straightforward. When prices are held constant, changing the distribution of labour productivity will only affect the composition of aggregate labour supply. Individuals' choices, including their labour supply decisions, are affected only indirectly through the general equilibrium effect on wage rate and interest rate. Within this model, we are able to derive a necessary and sufficient condition under which an increase in labour productivity heterogeneity will lead to an expansion in aggregate labour supply in the steady state. In such a scenario, greater diversity will benefit the consumers by raising their income and consumption.

The effects of time preference heterogeneity, by contrast, are more intricate due to the presence of two often conflicting forces. First, changing the distribution of time preference will affect aggregate economic outcomes by distorting the distribution of marginal tax rates across individuals. In the context of representative-agent models, the negative relation between marginal tax rate and capital accumulation is well understood: a decrease in marginal tax rate raises the return of savings which in turn promotes capital accumulation.[6] One novelty of the present study is to show that in a heterogeneous-agent economy, changing the composition of the underlying population can influence the effective marginal tax rate, even when there is no change in the tax schedule *per se*. Interestingly, the outcome of this mechanism is determined by the concavity and convexity of the *marginal* tax function, which is an often overlooked feature of the tax schedule.[7] If the marginal tax function is concave, then a more heterogeneous population will have a lower average marginal tax rate and a higher level of capital accumulation. The opposite is true when it is convex. The intuition of this can be seen by considering the following example: Start with a homogeneous economy in which all consumers are *ex ante* identical, receive the same amount of before-tax income and face the same progressive tax schedule. Suppose now a mean-preserving dispersion in consumer characteristics is introduced. Such dispersion will lead to a non-degenerate distribution in before-tax income and marginal tax rate. In particular, the relatively poor consumers in the

[6] Empirical evidence on this is scant, however, mainly because of the difficulty in measuring marginal tax rate. For this reason, many studies focus on the relation between average tax rate and economic growth. One exception is Padovano and Galli (2001) which construct country-wide point estimates of effective marginal tax rate for 23 OECD countries over the period 1951-1990 and show that this measure is negatively correlated with economic growth. The question of how the distribution or dispersion of marginal tax rates would affect aggregate economic outcomes, however, remains unexplored.

[7] If a tax function $\tau(\cdot)$ is thrice differentiable, then the corresponding marginal tax function is concave (or convex) if and only if the third-order derivative $\tau'''(\cdot)$ is negative (or positive). It is important to note that almost all the existing quantitative studies on progressive taxation have adopted a specification which implies a concave marginal tax function (see Section 4.1 for details). But the relation between this and the distribution of marginal tax rates has not been fully explained until now.

heterogeneous economy will pay a lower marginal tax rate than in the homogeneous world, and the relatively rich will pay a higher rate. The shape of the marginal tax function matters when it comes to aggregation. If the marginal tax function is concave, then the decrease in marginal tax rate among the poor will outweigh the increase among the rich. As a result, the heterogeneous economy will have a lower average marginal tax rate than the homogeneous economy.[8] Our main results in Section 4.1 generalise this comparison to any two heterogeneous economies which are otherwise identical except for the distribution of time preference. Second, changing this distribution will also affect the formation of aggregate labour supply, as in the case of labour productivity heterogeneity. When there is no income effect on labour supply, the relation between time preference heterogeneity and aggregate labour depends crucially on the curvature of the marginal rate of substitution between consumption and labour. It remains a challenge to verify whether this type of results will hold when the income effect is operational. In Section 4.3, we provide some numerical examples that can offer some insights on this issue.

The rest of the paper is organised as follows: Section 2 describes the baseline model. Section 3 presents the results regarding an increase in labour productivity heterogeneity. Section 4 analyses the effects of greater time preference heterogeneity. Section 5 concludes.

2 The Baseline Model

2.1 Consumers

Time is discrete and denoted by $t \in \{0, 1, 2, ...\}$. The economy under study is inhabited by a continuum of infinitely-lived consumers with different rate of time preference and labour productivity. The size of population is constant over time and normalised to one. Let $\rho_i > 0$ be the rate of time preference of the ith consumer, $i \in [0, 1]$, and $\varepsilon_i > 0$ his labour productivity. Both are predetermined and constant over time. These characteristics are cross-sectionally distributed according to the function $H(\rho, \varepsilon)$, which is defined over the support $[\underline{\rho}, \overline{\rho}] \times [\underline{\varepsilon}, \overline{\varepsilon}]$ with $\overline{\rho} > \underline{\rho} > 0$ and $\overline{\varepsilon} > \underline{\varepsilon} > 0$. This distribution can be either discrete or continuous (or mixed). The marginal distribution of time preference and labour productivity are denoted by $H_1(\rho)$ and $H_2(\varepsilon)$, respectively.

In each time period, each consumer is endowed with one unit of time which can be divided between work and leisure. Let $n_{i,t}$ and $l_{i,t}$ denote, respectively, the fraction of time spent on work

[8] The effects under a convex marginal tax function are similar but in opposite directions.

and leisure by the ith consumer at time t. These variables are subject to the following constraints:

$$n_{i,t} \in [0,1], \quad l_{i,t} \in [0,1], \quad \text{and} \quad n_{i,t} + l_{i,t} = 1. \tag{1}$$

There is a single commodity in this economy which can be used for consumption and investment. Let $c_{i,t}$ be the consumption of the ith consumer at time t. All consumers have preferences over sequences of consumption and labour hours which can be represented by

$$\sum_{t=0}^{\infty} \beta_i^t U(c_{i,t}, n_{i,t}), \tag{2}$$

where $\beta_i \equiv (1+\rho_i)^{-1}$ is the subjective discount factor of the ith consumer and $U(\cdot)$ is the (per-period) utility function. The latter is identical for all consumers and has the following properties.

Assumption A1 The utility function $U : \mathbb{R}_+ \times [0,1] \to \mathbb{R}$ is twice continuously differentiable, strictly increasing in c, strictly decreasing in n and jointly strictly concave in (c,n). It also satisfies the conditions: $\lim_{c\to 0} U_c(c,n) = \infty$ for all $n \in [0,1]$ and $\lim_{n\to 0} U_n(c,n) = 0$ for all $c \geq 0$.

Assumption A2 The marginal rate of substitution between consumption and labour, denoted by $\Psi(c,n) \equiv -U_n(c,n)/U_c(c,n)$, is non-decreasing in c and strictly increasing in n.

Assumption A2 implies that both consumption and leisure are normal goods.[9] This assumption can be equivalently stated as

$$U_{cn}(c,n) \leq \frac{U_n(c,n)}{U_c(c,n)} U_{cc}(c,n) \qquad \text{and} \qquad U_{nn}(c,n) \leq \frac{U_n(c,n)}{U_c(c,n)} U_{cn}(c,n),$$

for all (c,n). Assumptions A1 and A2 are satisfied by most of the commonly used utility functions in quantitative studies. Two examples are the class of additively separable utility functions, i.e., $U(c,n) = u(c) - v(n)$, and the so-called "no-income-effect" utility function, i.e., $U(c,n) \equiv u(c - v(n))$.[10] In both cases, $u(\cdot)$ is strictly increasing and strictly concave, while $v(\cdot)$ is strictly increasing and strictly convex. For the "no-income-effect" utility, the marginal rate of substitution between consumption and labour is given by $\Psi(c,n) = v'(n)$, which is independent of c.

[9] This means, holding other things constant, an increase in non-wage income in the current period will lead to an increase in current consumption and a decrease in current labour supply. This normality assumption is commonly used in existing studies. See for instance, Nourry (2001) and Datta *et al.* (2002).

[10] The latter is also often referred to as GHH preferences, named after the study by Greenwood *et al.* (1988).

Next, we turn to the budget constraint for an individual consumer. Let w_t be the wage rate for an effective unit of labour at time t. Then consumer i's labour income at time t is given by $w_t \varepsilon_i n_{i,t}$. Consumers can save and borrow through a single risk-free asset. Let $a_{i,t}$ denote consumer i's asset holdings at the beginning of time t. The consumer is in debt if this variable takes a negative value. The interest income (or interest payment) associated with these assets is $r_t a_{i,t}$, where r_t is the interest rate. The sum of labour income and interest income, denoted by $y_{i,t} \equiv w_t \varepsilon_i n_{i,t} + r_t a_{i,t}$, is subject to a progressive tax schedule.[11] This is represented by a function $\tau(\cdot)$ which satisfies the following properties.

Assumption A3 The tax function $\tau : \mathbb{R}_+ \to \mathbb{R}_+$ is continuously differentiable and strictly increasing with $\tau(0) \leq 0$. The marginal tax rate is zero at the origin, i.e., $\tau'(0) = 0$, strictly increasing for all $y \geq 0$ and satisfies $\lim_{y \to \infty} \tau'(y) = \overline{\tau} \leq 1$.

The assumption of an increasing marginal tax rate is often referred to as marginal rate progressivity. This, together with $\tau(0) \leq 0$, is equivalent to average rate progressivity, i.e., average tax rate $\tau(y)/y$ is increasing in y. A negative value of $\tau(0)$ can be interpreted as a fixed lump-sum transfer from the government. In this case, $\tau(y)$ is the net tax payment for a consumer with taxable income y.

Consumer i's budget constraint at time t is then given by

$$c_{i,t} + a_{i,t+1} - a_{i,t} = y_{i,t} - \tau(y_{i,t}). \tag{3}$$

Taking prices and tax schedule as given, each consumer's problem is to choose a sequence of consumption, leisure, labour and asset holdings so as to maximise his lifetime utility in (2), subject to the time-use constraints in (1), the sequential budget constraint in (3) and the initial amount of assets $a_0 > 0$.[12] There is no other restriction on borrowing except the no-Ponzi-scheme condition, which is implied by the transversality condition stated below. The solution of the consumer's problem is completely characterised by the sequential budget constraint in (3); the Euler equation

[11] This setup, which is commonly used in the existing studies, implicitly assumes that interests paid on loans are tax deductible. This assumption is adopted mainly for analytical convenience. In most countries, interests paid on personal loans are in general not deductible from taxes. In the United States, for instance, taxpayers can claim deductions on interests paid on student loans and residential mortgages but not on other types of loans (such as credit card debts).

[12] The current framework can be easily extended to allow for heterogeneity in initial wealth. But since we focus on steady-state analysis, this type of heterogeneity is irrelevant for our results.

for consumption

$$U_c\left(c_{i,t}, n_{i,t}\right) = \beta_i U_c\left(c_{i,t+1}, n_{i,t+1}\right)\left\{1 + \left[1 - \tau'\left(y_{i,t+1}\right)\right] r_{t+1}\right\}; \tag{4}$$

the optimality condition for labour supply

$$\Psi\left(c_{i,t}, n_{i,t}\right) \equiv -\frac{U_n\left(c_{i,t}, n_{i,t}\right)}{U_c\left(c_{i,t}, n_{i,t}\right)} \leq w_t \varepsilon_i \left[1 - \tau'\left(y_{i,t}\right)\right], \tag{5}$$

which holds with equality if $n_{i,t} \in (0, 1)$; and the transversality condition

$$\lim_{T\to\infty}\left\{\left[\prod_{t=1}^{T}\left(1+\varphi_{i,t}\right)\right]^{-1} a_{i,T+1}\right\} = 0,$$

where $\varphi_{i,t} \equiv \left[1 - \tau'\left(y_{i,t}\right)\right] r_t$ is the after-tax return from asset holdings. The condition in (5) takes into account the possibility that a consumer may choose to have zero leisure hours in certain time period, i.e., $n_{i,t} = 1$ for some t.[13] This happens when the relative price of leisure, i.e., $w_t \varepsilon_i \left[1 - \tau'\left(y_{i,t}\right)\right]$, is greater than or equal to the marginal rate of substitution at $n_{i,t} = 1$, i.e., $\Psi\left(c_{i,t}, 1\right)$.

2.2 Production and Government

On the supply side of the economy, there is a large number of identical firms. In each period, each firm hires labour and rents physical capital from the competitive factor markets, and produces output using a neoclassical production function: $Y_t = F\left(K_t, N_t\right)$, where Y_t denotes output at time t, K_t and N_t denote capital input and labour input, respectively. The properties of the production function are summarised below.

Assumption A4 The production function $F : \mathbb{R}_+^2 \to \mathbb{R}_+$ is twice continuously differentiable, strictly increasing and strictly concave in (K, N). It also exhibits constant returns to scale (CRTS) in the two inputs and satisfies the Inada conditions.

Since the production function exhibits CRTS in its inputs, we can focus on the profit-maximisation problem of a single representative firm. Let R_t be the rental price of physical capital at time t.

[13] The other corner solution $n_{i,t} = 0$ can be ruled out by the condition $\lim_{n\to 0} U_n\left(c, n\right) = 0$, for all $c \geq 0$, stated in Assumption A1.

Then the representative firm's problem is

$$\max_{K_t,N_t} \{F(K_t, N_t) - w_t N_t - R_t K_t\},$$

and the first-order conditions are $R_t = F_K(k_t, 1)$ and $w_t = F_N(k_t, 1)$, where $k_t \equiv K_t/N_t$ is the capital-labour ratio at time t.

Tax revenues collected by the government are entirely spent on "unproductive" government purchases (G_t). This spending is called unproductive because it has no direct effect on consumers' utility and firms' production. The government's budget constraint in every period t is given by

$$\int_0^1 \tau(y_{i,t})\, di = G_t, \qquad \text{for all } t \geq 0. \tag{6}$$

2.3 Competitive Equilibrium

Given a progressive tax schedule, a competitive equilibrium consists of sequences of allocations $\{c_{i,t}, l_{i,t}, n_{i,t}, a_{i,t}\}_{t=0}^{\infty}$ for each $i \in [0,1]$, aggregate inputs $\{K_t, N_t\}_{t=0}^{\infty}$, prices $\{w_t, r_t, R_t\}_{t=0}^{\infty}$ and government spending $\{G_t\}_{t=0}^{\infty}$ such that

(i) Given prices and the tax function, $\{c_{i,t}, l_{i,t}, n_{i,t}, a_{i,t}\}_{t=0}^{\infty}$ solves consumer i's problem.

(ii) Given prices, $\{K_t, N_t\}_{t=0}^{\infty}$ solves the representative firm's problem in every period.

(iii) The government's budget is balanced in every period.

(iv) All markets clear in every period, so that

$$K_t = \int_0^1 a_{i,t} di, \qquad \text{and} \qquad N_t = \int_0^1 \varepsilon_i n_{i,t} di, \qquad \text{for all } t \geq 0.$$

We confine our attention to the stationary equilibria or steady states of this economy, which can be characterised as follows: For any non-trivial steady state with capital-labour ratio $k > 0$, let $w(k)$ and $r(k)$ be the corresponding wage rate and interest rate. To highlight the dependence of individual choices on (ρ, ε), we use $y(k, \rho, \varepsilon)$, $c(k, \rho, \varepsilon)$, $a(k, \rho, \varepsilon)$ and $n(k, \rho, \varepsilon)$ to denote, respectively, the level of before-tax income, consumption, asset holdings and labour of a type-(ρ, ε) consumer in this steady state (the subscript i will be omitted from this point on). These individual-level variables are determined by

$$\rho = r(k) \left\{1 - \tau'[y(k, \rho, \varepsilon)]\right\}, \tag{7}$$

$$c(k,\rho,\varepsilon) \equiv y(k,\rho,\varepsilon) - \tau\left[y(k,\rho,\varepsilon)\right], \tag{8}$$

$$a(k,\rho,\varepsilon) = \frac{y(k,\rho,\varepsilon) - w(k)\,\varepsilon n(k,\rho,\varepsilon)}{r(k)}, \tag{9}$$

$$1 \geq n(k,\rho,\varepsilon), \qquad \Psi\left[c(k,\rho,\varepsilon), n(k,\rho,\varepsilon)\right] \leq \frac{w(k)}{r(k)}\varepsilon\rho, \tag{10}$$

$$\left[1 - n(k,\rho,\varepsilon)\right]\left\{\frac{w(k)}{r(k)}\varepsilon\rho - \Psi\left[c(k,\rho,\varepsilon), n(k,\rho,\varepsilon)\right]\right\} = 0. \tag{11}$$

Equation (7) is obtained by setting $U_c(c_{i,t}, n_{i,t}) = U_c(c_{i,t+1}, n_{i,t+1})$ in the Euler equation of consumption.[14] The intuition of this equation is as follows: In any stationary equilibrium, each consumer has a tendency to perfectly smooth their marginal utility of consumption over time. To achieve this, the after-tax return from asset holdings must be equated to the consumer's rate of time preference. This has two important implications. Firstly, consumers with the same rate of time preference will face the same marginal tax rate and have the same level of before-tax income, regardless of their labour productivity. In other words, $y(k,\rho,\varepsilon)$ is independent of ε. Secondly, for any given ρ in $\left[\underline{\rho}, \overline{\rho}\right]$, $y(k,\rho)$ is a strictly decreasing function in k. This is due to the following mechanism: Holding other things constant, an increase in k will encourage the consumer to substitute future consumption for current consumption by lowering the before-tax return from asset holdings. In order to maintain a constant marginal utility of consumption, it is necessary for the marginal tax rate to fall so that the after-tax return is again equal to ρ. Since $\tau'(\cdot)$ is strictly increasing, this means before-tax income $y(k,\rho)$ will have to fall after an increase in k. In subsequent discussions, we will refer to this as the intertemporal smoothing effect. Note that this effect arises only when the income tax schedule is nonlinear.[15] Equation (8) is obtained by setting $a_{i,t+1} = a_{i,t}$ in the sequential budget constraint. This, together with (7), implies that $c(k,\rho,\varepsilon)$ is also independent of ε. Equation (9) follows from the definition of before-tax income. Equations (10) and (11) are the complementary slackness conditions for labour. In particular, the second inequality in (10) is obtained by substituting (7) into (5).

[14] Note that equation (7) remains valid even if (i) we allow for *ex ante* heterogeneity in the utility function, i.e., $U^i(c,n) \neq U^j(c,n)$ for some $i \neq j$ in $[0,1]$, and (ii) there is no disutility from labour, i.e., $U(c,n_1) = U(c,n_2)$ for all $n_1 \neq n_2$ in $[0,1]$ and for all $c \geq 0$.

[15] If the income tax function is linear, i.e., $\tau'(y) = \widehat{\tau}$ for all $y \geq 0$, then the steady-state value of k is uniquely determined by $(1-\widehat{\tau})\,r(k) = \underline{\rho}$. In this case, only those consumers with the lowest rate of time preference (i.e., the most patient consumers) will hold a strictly positive amount of assets. All other consumers will either have zero wealth (if they are not allowed to borrow) or exhaust the borrowing limit (if an *ad hoc* borrowing constraint is in place) as in the model of Becker (1980). Sarte (1997) shows that equation (7) plays a key role in obtaining a nondegenerate steady-state wealth distribution in the presence of time preference heterogeneity. The implications of the intertemporal smoothing effect, however, is less mentioned in the existing literature.

The next step is to derive a single equation that can help determine the value of k. Define $\phi : [0, \overline{\tau}] \rightarrow \mathbb{R}_{+} \cup \{+\infty\}$ as the inverse of the marginal tax function, i.e., $\phi[\tau'(y)] = y$ for all $y \geq 0$. Since $\tau'(\cdot)$ is continuous and strictly increasing, its inverse is a single-valued, continuous, strictly increasing function. Using (7) and the definition of before-tax income, we can write

$$y(k, \rho) \equiv \phi\left[1 - \frac{\rho}{r(k)}\right] = w(k)\,\varepsilon n(k, \rho, \varepsilon) + r(k)\,a(k, \rho, \varepsilon). \tag{12}$$

Integrating both sides of (12) across all types of consumers yields

$$Y(k) \equiv \int_{\underline{\rho}}^{\overline{\rho}} \phi\left[1 - \frac{\rho}{r(k)}\right] dH_1(\rho) = [f(k) - \delta k]\,N(k), \tag{13}$$

where $H_1(\rho)$ is the marginal distribution of ρ; $N(k)$ is the aggregate labour supply function defined by

$$N(k) \equiv \int_{\underline{\rho}}^{\overline{\rho}} \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon n(k, \rho, \varepsilon)\, dH(\rho, \varepsilon),$$

and $f(k) \equiv F(k, 1)$ is the reduced-form production function. Equation (13) is essentially an accounting identity which states that the sum of all individuals' income equals national income (defined as aggregate output minus depreciation of capital). In the sequel, we will refer to $Y(\cdot)$ as the national income function. A unique, non-trivial steady state exists if (13) has a single, strictly positive solution. The rest of this section is devoted to establishing the existence of a unique, non-trivial steady state.

We begin by specifying the range of plausible values of k. Since $\phi(\cdot)$ is only defined on $[0, \overline{\tau}]$, equations (12) and (13) are satisfied only if

$$1 - \overline{\tau} \geq 1 - \frac{\rho}{r(k)} \geq 0,$$

for all $\rho \in [\underline{\rho}, \overline{\rho}]$. In other words, any k that solves (13) must satisfy

$$\frac{\underline{\rho}}{1 - \overline{\tau}} \geq r(k) \geq \overline{\rho}. \tag{14}$$

To ensure that this range is nonempty, it is necessary to impose the condition $\underline{\rho} > (1 - \overline{\tau})\,\overline{\rho}$. By the strict concavity of $f(\cdot)$ and the Inada conditions on the production function, there exists a

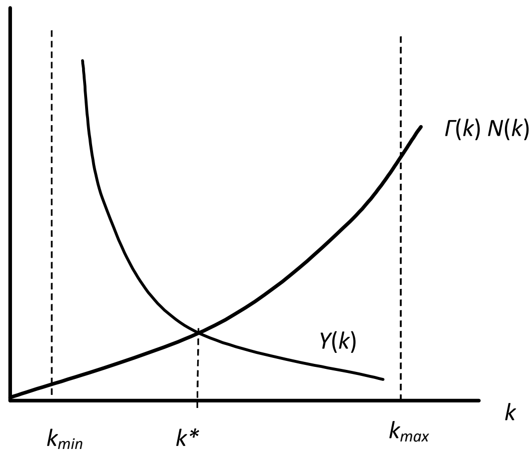

Figure 1a

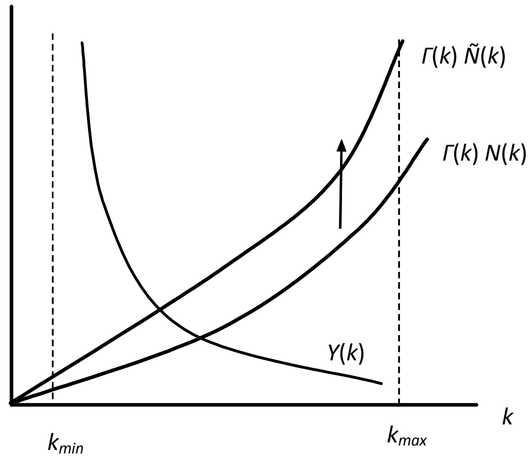

Figure 1b

unique pair of values $k_{\max} > k_{\min} > 0$ such that

$$r(k_{\max}) \equiv f'(k_{\max}) - \delta = \overline{\rho} \qquad \text{and} \qquad r(k_{\min}) = \frac{\underline{\rho}}{1-\overline{\tau}}.$$

It follows that any solution of (13) must be contained within the interval $[k_{\min}, k_{\max}]$. Lemma 1 provides a set of necessary and sufficient conditions for the existence of a unique steady state. The proof of this lemma and other theoretical results can be found in the Appendix. A graphical illustration of the unique steady state is provided in Figure 1a.[16]

Lemma 1 *Suppose Assumptions A1-A4 and $\underline{\rho} > (1-\overline{\tau})\overline{\rho}$ are satisfied. Then a unique steady state with capital-labour ratio $k^* \in (k_{\min}, k_{\max})$ exists if and only if*

$$N(k_{\max})[f(k_{\max}) - \delta k_{\max}] > \int_{\underline{\rho}}^{\overline{\rho}} \phi\left(1 - \frac{\rho}{\overline{\rho}}\right) dH_1(\rho), \tag{15}$$

and

$$N(k_{\min})[f(k_{\min}) - \delta k_{\min}] < \int_{\underline{\rho}}^{\overline{\rho}} \phi\left[1 - \frac{\rho}{\underline{\rho}}(1-\overline{\tau})\right] dH_1(\rho). \tag{16}$$

[16] Conditions (15) and (16) are some technical conditions which ensure that the two curves in Figure 1a will cross at least once within the range $(k_{\min}, k_{\max})$. The function $\Gamma(\cdot)$ in these diagrams is defined as $\Gamma(k) \equiv f(k) - \delta k$, which also appears in the proof of Lemma 1.

3 Heterogeneity in Labour Productivity

In this section we examine how greater heterogeneity in labour productivity will affect the steady state of the baseline model. Our results are based on a comparison between two economies which are otherwise identical except for the marginal distribution of ε. To facilitate this comparison, we assume that the two characteristics, ρ and ε, are statistically independent in the population so that $H(\rho, \varepsilon) = H_1(\rho) H_2(\varepsilon)$ for all (ρ, ε).

Consider two economies which have the same size of population, utility function $u(\cdot)$, production technology $F(\cdot)$, progressive tax schedule $\tau(\cdot)$ and marginal distribution of time preference $H_1(\cdot)$ defined over $[\underline{\rho}, \overline{\rho}]$.[17] The only difference between them lies in the distribution of labour productivity, which are denoted by $H_2(\varepsilon)$ and $\widetilde{H}_2(\varepsilon)$. Both are defined over $[\underline{\varepsilon}, \overline{\varepsilon}]$ and satisfy the assumption below.

Assumption A5 (i) The average value of ε is identical under $H_2(\cdot)$ and $\widetilde{H}_2(\cdot)$. (ii) Conditions (15) and (16) are satisfied in both economies.

The second part of Assumption A5 ensures that both economies have a unique steady state. Notice that when k is held constant, changing the distribution of labour productivity will have no effect on the individual-level variables defined by (7)-(11). From (13), it is evident that the distribution of ε will affect the steady-state capital-labour ratio only through the aggregate labour supply function. Intuitively, what this means is that when prices are kept constant, changing the distribution of ε will only affect the composition of aggregate labour supply. Individual-level variables are affected only indirectly through the general equilibrium effect on $w(k)$ and $r(k)$.

Let $N(\cdot)$ be the aggregate labour supply function defined under $H_2(\cdot)$, i.e.,

$$N(k) \equiv \int_{\underline{\rho}}^{\overline{\rho}} \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon n(k, \rho, \varepsilon)\, dH_2(\varepsilon)\, dH_1(\rho),$$

and similarly define $\widetilde{N}(k)$ using $\widetilde{H}_2(\cdot)$. From Figure 1b, it is evident that if $\widetilde{N}(k) \geq N(k)$ for all $k \in [k_{\min}, k_{\max}]$, then the economy with $H_2(\cdot)$ will have a higher steady-state capital-labour ratio than the one with $\widetilde{H}_2(\cdot)$. The opposite is true if the ordering of $N(\cdot)$ and $\widetilde{N}(\cdot)$ is reversed. Proposition 3 provides a necessary and sufficient condition under which $\widetilde{N}(k) \geq N(k)$ for all $k \in [k_{\min}, k_{\max}]$. The implications of this on other variables of interest are summarised in Proposition 4. Before presenting these results, we first establish a basic property of individual

[17] This implies that both economies have the same range of plausible values of steady-state capital-labour ratio, $[k_{\min}, k_{\max}]$, as defined in (14).

labour supply function which plays a key role in the analysis.

Lemma 2 *Suppose Assumptions A1-A4 and* $\underline{\rho} > (1 - \overline{\tau})\,\overline{\rho}$ *are satisfied. Then for any* $k \in [k_{\min}, k_{\max}]$ *and* $\rho \in [\underline{\rho}, \overline{\rho}]$, $n(k, \rho, \varepsilon)$ *is a non-decreasing function in* ε. *If, in addition,* $n(k, \rho, \varepsilon)$ *is an interior solution, then it is strictly increasing in* ε.

The intuition of this result is clear: since more productive workers have a higher opportunity cost of leisure, they tend to work more than less productive ones. This result holds in general when (i) the marginal tax rate on labour income and interest income are identical, and (ii) the marginal rate of substitution between consumption and labour is strictly increasing in labour. Both assumptions are satisfied in most of the existing studies.

Lemma 2 also has the implication that a one-percent increase in ε can, in some cases, lead to a greater percentage increase in effective unit of labour, i.e., $\varepsilon n(k, \rho, \varepsilon)$. Specifically, let $\varepsilon_2 = (1 + \Theta)\,\varepsilon_1$, for some $\Theta > 0$, and suppose $n(k, \rho, \varepsilon_1)$ and $n(k, \rho, \varepsilon_2)$ are both interior solutions. Then we have $\varepsilon_2 n(k, \rho, \varepsilon_2) > (1 + \Theta)\,\varepsilon_1 n(k, \rho, \varepsilon_1)$. Thus, in the current setting an endogenous labour supply has the effect of amplifying the variations in labour productivity across consumers.

We now present the main results of this section.

Proposition 3 *Suppose Assumptions A1-A4 and* $\underline{\rho} > (1 - \overline{\tau})\,\overline{\rho}$ *are satisfied. Then* $N(k) \leq \widetilde{N}(k)$ *for all* $k \in [k_{\min}, k_{\max}]$ *if and only if*

$$\int_x^{\overline{\varepsilon}} \varepsilon dH_2(\varepsilon) \leq \int_x^{\overline{\varepsilon}} \varepsilon d\widetilde{H}_2(\varepsilon), \qquad \textit{for all } x \in [\underline{\varepsilon}, \overline{\varepsilon}]. \tag{17}$$

To explain the main ideas behind this proposition, first rewrite $N(k)$ and $\widetilde{N}(k)$ as

$$N(k) \equiv \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon \mathcal{N}(k, \varepsilon)\, dH_2(\varepsilon) \qquad \text{and} \qquad \widetilde{N}(k) \equiv \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon \mathcal{N}(k, \varepsilon)\, d\widetilde{H}_2(\varepsilon), \tag{18}$$

where $\mathcal{N}(k, \varepsilon)$ is the average labour supply across all consumers with the same level of ε, i.e.,

$$\mathcal{N}(k, \varepsilon) \equiv \int_{\underline{\rho}}^{\overline{\rho}} n(k, \rho, \varepsilon)\, dH_1(\rho).$$

We can now interpret $N(k) \leq \widetilde{N}(k)$ as comparing the expected value of $\varepsilon\mathcal{N}(k, \varepsilon)$ under $H_2(\cdot)$ and $\widetilde{H}_2(\cdot)$.[18] By Lemma 2, $\varepsilon\mathcal{N}(k, \varepsilon)$ is a bounded, continuous, *starshaped* function in ε and (17)

[18] If $\varepsilon\mathcal{N}(k, \varepsilon)$ is convex in ε for any given $k \in [k_{\min}, k_{\max}]$, then $N(k) \leq \widetilde{N}(k)$ if and only if $\widetilde{H}_2(\cdot)$ is more unequal than $H_2(\cdot)$ under the Lorenz dominance criterion. The function $\varepsilon\mathcal{N}(k, \varepsilon)$, however, is not convex in general. The details of this point are available from the authors upon request. This is the main reason why we opt for a stronger stochastic ordering, namely the starshaped ordering (defined below).

is a necessary and sufficient condition for ranking the expected value of this type of functions.[19] In the statistics literature, this type of ordering is known as the *starshaped ordering* of probability distributions. A detailed discussion of this type of ordering can be found in Shaked and Shanthikumar (2007, Section 4.A.6). If two distributions $H_2(\cdot)$ and $\widetilde{H}_2(\cdot)$ have the same mean and satisfy (17), then $\widetilde{H}_2(\cdot)$ is also more *unequal* than $H_2(\cdot)$ under the standard Lorenz dominance criterion [or equivalently, $\widetilde{H}_2(\cdot)$ is a mean-preserving spread of $H_2(\cdot)$]. Thus, Lorenz ordering is implied by starshaped ordering. When applied to the current context, this means $\widetilde{H}_2(\cdot)$ is more dispersed than $H_2(\cdot)$, and thus represents a more heterogeneous population.

Equipped with the result in Proposition 3, we can now compare the steady states under $H_2(\cdot)$ and $\widetilde{H}_2(\cdot)$. Let k^* and $\widetilde{k}^*$ be the unique solution of (13) under $H_2(\cdot)$ and $\widetilde{H}_2(\cdot)$, respectively. In terms of aggregate variables, a more dispersed distribution of labour productivity is associated with a lower value of capital-labour ratio but a higher value of aggregate labour and national income. At the individual level, greater heterogeneity will benefit the consumers by raising their before-tax income and consumption. These results are summarised in Proposition 4.

Proposition 4 *Suppose Assumptions A1-A5 and $\underline{\rho} > (1 - \overline{\tau})\overline{\rho}$ are satisfied. Suppose $\widetilde{H}_2(\cdot)$ is more heterogeneous than $H_2(\cdot)$ according to (17). Then we have*

(i) $k^* \geq \widetilde{k}^*$, $N(k^*) \leq \widetilde{N}\left(\widetilde{k}^*\right)$ *and* $Y(k^*) \leq Y\left(\widetilde{k}^*\right)$.

(ii) $y(k^*, \rho) \leq y\left(\widetilde{k}^*, \rho\right)$ *and* $c(k^*, \rho) \leq c\left(\widetilde{k}^*, \rho\right)$ *for all* $\rho \in \left[\underline{\rho}, \overline{\rho}\right]$.

These results can be explained as follows: Holding other factors (especially k) constant, an increase in labour productivity heterogeneity will lead to an expansion in aggregate labour supply. This lowers the capital-labour ratio which then triggers the aforementioned intertemporal smoothing effect. In particular, before-tax income will increase across all types of consumers, followed by national income and individual consumption.[20]

4 Heterogeneity in Time Preference

We now examine the steady-state outcomes of greater time preference heterogeneity. Following the same approach as in Section 3, we compare two economies which are otherwise identical except

[19] Recall that a real-valued function $\xi(\cdot)$ defined on $[0, \infty)$ is called starshaped if $\xi(0) \leq 0$ and $\xi(x)/x$ is nondecreasing.

[20] The effects on aggregate capital $K \equiv kN(k)$ and aggregate output $N(k) f(k)$, however, are ambiguous, due to the opposing effects on k and $N(k)$.

for the marginal distribution of ρ. Throughout this analysis, we will maintain the assumption that ρ and ε are statistically independent across the population.

Comparing to the previous section, the task we face here is more complicated due to the following reasons: Firstly, since national income is formed by summing individual's before-tax income $y(k, \rho)$ over the distribution of ρ, changing this distribution will have a direct effect on the function $Y(\cdot)$. This effect is not present in the case of labour productivity heterogeneity as $y(k, \rho)$ is independent of ε. Secondly, it is now more difficult to characterise the shape of $n(k, \rho, \varepsilon)$ in ρ due to the presence of an income effect on labour supply. More specifically, changes in ρ will affect individual labour supply in two ways: (i) by changing the after-tax wage rate through the variable $y(k, \rho)$, and (ii) by distorting the marginal rate of substitution between consumption and labour through the variable $c(k, \rho)$. The latter is what we refer to as the income effect.

Because of these added complexities, the theoretical analysis in this section is based on two simplified versions of the baseline model. First, in Section 4.1 we assume that individual labour supply is an exogenous constant. As a result, aggregate labour input is independent of the distribution of ρ and other parts of the economy. This abstraction allows us to focus on the effects of time preference heterogeneity on $Y(k)$ alone. As we will see below, these effects are entirely determined by the shape of the marginal tax function $\tau'(\cdot)$. This part of the analysis thus highlights the role of the progressive tax schedule in determining the impact of greater time preference heterogeneity. Next, in Section 4.2 we reintroduce a flexible labour supply but abstract away from the aforementioned income effect. This is achieved by adopting the so-called "no-income-effect" preferences. In this case, the effects of greater time preference heterogeneity are jointly determined by the shape of the marginal tax function and that of the marginal rate of substitution between consumption and labour. Finally, in Section 4.3 we provide some simple numerical examples to illustrate the workings of the baseline model where the income effect is operational.

4.1 Exogenous Labour Model

The baseline model can be significantly simplified once we remove the assumption of flexible labour supply. Instead of imposing Assumptions A1 and A2, we now set $U(c, n) \equiv u(c)$ for all $c \geq 0$ and $n \in [0, 1]$, where $u : \mathbb{R}_+ \to \mathbb{R}$ is a twice continuously differentiable, strictly increasing and strictly concave function that satisfies $\lim_{c \to 0} u'(c) = \infty$. Individual and aggregate labour supply are then given by $n_{i,t} = 1$ for all i and $N_t = \widehat{\varepsilon} > 0$, where $\widehat{\varepsilon}$ is the average labour productivity across consumers. The rest of the economy is the same as in the baseline model.

In any stationary equilibrium, $y(k,\rho)$ and $c(k,\rho)$ are again determined by (7) and (8) but the Kuhn-Tucker conditions in (10)-(11) are now superseded by $n(k,\rho,\varepsilon) = 1$, for all (k,ρ,ε). The steady-state value of k is then determined by

$$\int_{\underline{\rho}}^{\overline{\rho}} \phi \left[1 - \frac{\rho}{r(k)}\right] dH_1(\rho) = [f(k) - \delta k]\,\widehat{\varepsilon}. \tag{19}$$

Note that any solution of (19) depends only on the mean value of ε but not any other moment. Thus, there is no loss of generality in assuming that $H_2(\varepsilon)$ is a degenerate distribution at $\widehat{\varepsilon}$. Using the same line of argument as in the proof of Lemma 1, one can show that a unique solution of (19) exists if and only if (15) and (16) are satisfied [with $N(k_{\max})$ and $N(k_{\min})$ replaced by $\widehat{\varepsilon}$].

We now compare two economies which are otherwise identical except for the distribution of ρ, which are denoted by $H_1(\cdot)$ and $\widetilde{H}_1(\cdot)$. Both are defined over $[\underline{\rho}, \overline{\rho}]$ and satisfy the assumption below.

Assumption A6 (i) $\widetilde{H}_1(\cdot)$ is a mean-preserving spread of $H_1(\cdot)$. (ii) A unique steady state exists in both economies.

The first part of Assumption A6 implies that $\widetilde{H}_1(\cdot)$ is more dispersed than $H_1(\cdot)$, and thus represents a more heterogeneous population. Let $Y(\cdot)$ and $\widetilde{Y}(\cdot)$ be the national income function defined using $H_1(\cdot)$ and $\widetilde{H}_1(\cdot)$, respectively. Similarly, let k^* and $\widetilde{k}^*$ be the unique solution of (19) under $H_1(\cdot)$ and $\widetilde{H}_1(\cdot)$. Then a more heterogeneous population is said to be beneficial (or harmful) to long-term capital accumulation if $\widetilde{k}^* \geq k^*$ (or $\widetilde{k}^* \leq k^*$).[21] The main results of this subsection are summarised below.

Proposition 5 *Suppose Assumptions A3, A4, A6 and $\underline{\rho} > (1-\overline{\tau})\,\overline{\rho}$ are satisfied.*

(i) If the marginal tax function is concave, then $Y(k) \leq \widetilde{Y}(k)$ for all $k \in (k_{\min}, k_{\max})$ and a more heterogeneous population is beneficial to long-term capital accumulation.

(ii) If the marginal tax function is convex, then $Y(k) \geq \widetilde{Y}(k)$ for all $k \in (k_{\min}, k_{\max})$ and a more heterogeneous population is harmful to long-term capital accumulation.

One interesting special case is when $H_1(\cdot)$ is also a degenerate distribution, say at some point $\widehat{\rho}$ in $[\underline{\rho}, \overline{\rho}]$. In this case, we are comparing an identical-agent (IA) economy, in which all consumers are

[21] Since aggregate labour is an exogenous constant, aggregate capital, aggregate output and national income are all increasing in k. Thus, Proposition 5 is equivalent to saying that a more heterogeneous population is beneficial (or harmful) to aggregate output and national income if the marginal tax function is concave (or convex).

identical, to a heterogeneous-agent (HA) economy, where consumers have different time preference. Proposition 5 then implies that the HA economy will have a higher (or lower) level of long-run capital accumulation than the IA economy if the marginal tax function is concave (or convex). The intuition of these results can be understood by comparing the distribution of marginal tax rates in these two economies.

In the IA economy, all consumers have the same before-tax income $y(k^*, \widehat{\rho})$ and face the same marginal tax rate $\tau'[y(k^*, \widehat{\rho})]$. A mean-preserving spread in the rate of time preference will lead to a dispersion in both of these variables. In particular, it will lower the marginal tax rate for those with ρ greater than $\widehat{\rho}$ and raise the marginal tax rate for the others.[22] If the marginal tax function is concave, then the average marginal tax rate will be lowered as a result. Specifically, if $\tau'(\cdot)$ is concave, then we have

$$\tau'[y(k^*, \widehat{\rho})] \geq \frac{1}{1-\widetilde{H}_1(x)} \int_x^{\overline{\rho}} \tau'\left[y\left(\widetilde{k}^*, \rho\right)\right] d\widetilde{H}_1(\rho),$$

for all $x \in [\underline{\rho}, \overline{\rho}]$. The expression on the right side is the average marginal tax rate faced by those with $\rho \geq x$ in the HA economy. The lower average marginal tax rate then contributes to a higher level of capital accumulation in the HA economy. On the contrary, if $\tau'(\cdot)$ is concave, then we have

$$\tau'[y(k^*, \widehat{\rho})] \leq \frac{1}{\widetilde{H}_1(x)} \int_{\underline{\rho}}^{x} \tau'\left[y\left(\widetilde{k}^*, \rho\right)\right] d\widetilde{H}_1(\rho),$$

for all $x \in [\underline{\rho}, \overline{\rho}]$. In this case, consumers in the HA economy will in general face a higher marginal tax rate, which is detrimental to capital accumulation.

Our next proposition generalises this comparison to any two HA economies that satisfy Assumption A6. For any $q \in [0,1]$, define $\sigma(q)$ as the qth quantile of $H_1(\cdot)$, i.e., $\sigma(q) \equiv \sup\{\rho : H_1(\rho) \leq q\}$. Similarly, define $\widetilde{\sigma}(q)$ as the qth quantile of $\widetilde{H}_1(\cdot)$.

Proposition 6 *Suppose Assumptions A3, A4, A6 and $\underline{\rho} > (1-\overline{\tau})\overline{\rho}$ are satisfied.*

(i) If the marginal tax function is concave, then for any $q \in [0,1]$

$$\int_{\sigma(q)}^{\overline{\rho}} \tau'[y(k^*, \rho)] dH_1(\rho) \geq \int_{\widetilde{\sigma}(q)}^{\overline{\rho}} \tau'\left[y\left(\widetilde{k}^*, \rho\right)\right] d\widetilde{H}_1(\rho).$$

[22] This is because $y(k, \rho)$ is a strictly decreasing function in ρ under any given value of k. This property can be easily shown using the first part of (12).

(ii) If the marginal tax function is convex, then for any $q \in [0,1]$

$$\int_{\underline{\rho}}^{\sigma(q)} \tau' \left[y\left(k^*, \rho\right)\right] dH_1(\rho) \leq \int_{\underline{\rho}}^{\widetilde{\sigma}(q)} \tau' \left[y\left(\widetilde{k^*}, \rho\right)\right] d\widetilde{H}_1(\rho).$$

We conclude this subsection by pointing out the prevalence of concave marginal tax function in the existing literature. Two parametric forms of $\tau(\cdot)$ are commonly used in quantitative studies. The first one is the isoelastic form adopted by Guo and Lansing (1998), Li and Sarte (2004) and Angyridis (2015). This can be expressed as $\tau(y) = \zeta y^{1+\chi}$, with $\zeta > 0$ and $\chi > 0$. The parameter χ is often interpreted as a measure of tax progressivity. It is straightforward to show that the marginal tax function is concave (or convex) when $\chi \leq 1$ (or $\chi \geq 1$). Using U.S. tax returns data, Li and Sarte (2004) estimate that the value of χ was 0.88 in 1985 and 0.75 in 1991. Both imply a strictly concave marginal tax function. Another commonly used tax function is the one proposed and estimated by Gouveia and Strauss (1994),

$$\tau(y) = a_0 \left[y - \left(y^{-a_1} + a_2\right)^{-\frac{1}{a_1}} \right]. \tag{20}$$

This functional form was adopted by Sarte (1997), Conesa and Krueger (2006), Erosa and Koreshkova (2007), Carroll and Young (2011) among others. The second and third-order derivatives of this function are given by

$$\tau''(y) = a_0 a_2 (1+a_1)\left(1 + a_2 y^{a_1}\right)^{-\left(2+\frac{1}{a_1}\right)} y^{a_1 - 1},$$

$$\tau'''(y) = \frac{\tau''(y)}{y}\left[a_1 - 1 - (2a_1 + 1)\left(\frac{a_2 y^{a_1}}{1 + a_2 y^{a_1}}\right)\right]. \tag{21}$$

In all existing applications, the parameters a_0, a_1 and a_2 are taken to be strictly positive which ensure that $\tau''(y) > 0$. Gouveia and Strauss (1994) report estimates of a_1 ranging from 0.726 to 0.938 based on U.S. data. From (21), it is obvious that these values of a_1 imply $\tau'''(\cdot) < 0$, i.e., a strictly concave marginal tax function.

4.2 Endogenous Labour Without Income Effect

Suppose now the utility function is given by $U(c,n) = u[c - v(n)]$, where $u : \mathbb{R}_+ \to \mathbb{R}$ and $v : [0,1] \to \mathbb{R}_+$ are both twice continuously differentiable and strictly increasing. The former is also strictly concave and satisfies $\lim_{x\to 0} u'(x) = \infty$, while the latter is strictly convex and satisfies $\lim_{n\to 0} v'(n) = 0$. The rest of the economy is the same as in Section 2.

In any stationary equilibrium, equations (7)-(9) will remain valid while the optimality condition for labour supply can be simplified to become

$$v'\left[n\left(k,\rho,\varepsilon\right)\right] \leq \frac{w\left(k\right)}{r\left(k\right)}\varepsilon\rho, \tag{22}$$

with equality holds if $n\left(k,\rho,\varepsilon\right) < 1$. It is now straightforward to specify the conditions under which $n\left(k,\rho,\varepsilon\right)$ is an interior solution for all types of consumers.

Lemma 7 *Suppose Assumption A4 is satisfied. If* $v'\left(1\right) > w\left(k_{\max}\right)\overline{\varepsilon}$, *then* $n\left(k,\rho,\varepsilon\right) < 1$ *for all* $k \in \left[k_{\min}, k_{\max}\right]$ *and for all* $\left(\rho,\varepsilon\right) \in \left[\underline{\rho},\overline{\rho}\right] \times \left[\underline{\varepsilon},\overline{\varepsilon}\right]$.

Let $H_1\left(\cdot\right)$ and $\widetilde{H}_1\left(\cdot\right)$ be two distinct distributions of ρ defined over $\left[\underline{\rho},\overline{\rho}\right]$. As in Section 4.1, we assume that $\widetilde{H}_1\left(\cdot\right)$ is a mean-preserving spread of $H_1\left(\cdot\right)$, and a unique steady state exists under both distributions. When labour supply is flexible, an increase in time preference heterogeneity will not only affect the national income function $Y\left(\cdot\right)$, but also the aggregate labour supply function $N\left(\cdot\right)$. The effects on $Y\left(\cdot\right)$ are the same as in Proposition 5.[23] The effects on $N\left(\cdot\right)$ are examined below.

Proposition 8 *Suppose Assumptions A3, A4, A6,* $\underline{\rho} > \left(1-\overline{\tau}\right)\overline{\rho}$ *and* $v'\left(1\right) > w\left(k_{\max}\right)\overline{\varepsilon}$ *are satisfied. Then the following results hold for any* $k \in \left[k_{\min}, k_{\max}\right]$ *and for any* $\varepsilon \in \left[\underline{\varepsilon},\overline{\varepsilon}\right]$.

(i) If $v'\left(\cdot\right)$ *is a concave function, then* $n\left(k,\rho,\varepsilon\right)$ *is convex in* ρ *and* $N\left(k\right) \leq \widetilde{N}\left(k\right)$.

(ii) If $v'\left(\cdot\right)$ *is a convex function, then* $n\left(k,\rho,\varepsilon\right)$ *is concave in* ρ *and* $N\left(k\right) \geq \widetilde{N}\left(k\right)$.

To understand these results, first consider the labour supply decision of a single consumer. Suppose the condition for interior solution is satisfied. Then optimality is attained when the marginal rate of substitution between consumption and labour, i.e., $v'\left(n\right)$, equals the after-tax wage rate. The latter is determined by the steady-state capital-labour ratio, as well as the consumer's own characteristics. For the sake of this discussion, we will denote the after-tax wage rate simply as ϖ and individual labour supply as $n\left(\varpi\right)$. Since $v'\left(n\right)$ is increasing, individual labour supply is increasing in ϖ, i.e., $n\left(\varpi_2\right) \geq n\left(\varpi_1\right)$ whenever $\varpi_2 \geq \varpi_1$. The curvature of $v'\left(n\right)$ then determines whether a high-wage earner will respond more to the same increase in ϖ than a low-wage earner. Specifically, if $v'\left(n\right)$ is concave, then a high-wage earner will have a larger response to the same

[23] In particular, for any $k \in \left(k_{\min}, k_{\max}\right)$, $Y\left(k\right)$ is less (or greater) than $\widetilde{Y}\left(k\right)$ if the marginal tax function is concave (or convex). This result is independent of the assumptions on labour supply.

Table 1: Main Results of Section 4.2

$\tau'(\cdot)$	$v'(\cdot)$	$Y(k^*)$	k^*
concave	concave	$\uparrow$	Ambiguous
convex	convex	$\downarrow$	Ambiguous
convex	concave	Ambiguous	$\downarrow$
concave	convex	Ambiguous	$\uparrow$

increase in after-tax wage rate, i.e.,

$$n(\varpi_2 + \Delta) - n(\varpi_2) \geq n(\varpi_1 + \Delta) - n(\varpi_1).$$

This is equivalent to saying that individual labour supply is a convex function in ϖ. When comparing across consumers with different rate of time preference, first note that the after-tax wage rate in (22) is linearly increasing in ρ. Thus, consumers with a higher value of ρ will have a higher after-tax wage rate, and by the above reasoning, individual labour supply is a convex function in ρ when $v'(\cdot)$ is concave.[24] A mean-preserving spread in ρ will then lead to an increase in the average value of $\varepsilon n(k, \rho, \varepsilon)$ across all types of consumers, hence $N(k) \leq \widetilde{N}(k)$ for any given k.[25]

Based on the results in Propositions 5 and 8, we can identify four possible scenarios depending on the shape of $\tau'(\cdot)$ and $v'(\cdot)$. Table 1 summarises the overall effects of greater time preference heterogeneity in each of these cases. These effects can be easily seen with the aid of Figure 1a, hence the proof is omitted. For instance, when both $\tau'(\cdot)$ and $v'(\cdot)$ are concave, an increase in time preference heterogeneity will shift both the national income function and the aggregate labour supply function up, according to Propositions 5 and 8. This will lead to an unambiguous increase in national income, but an ambiguous effect on the capital-labour ratio. The latter is due to the presence of two opposing forces: on one hand, an increase in time preference heterogeneity will lower the average marginal tax rate on asset return which is beneficial to capital accumulation; on the other hand, such an increase will induce an expansion in aggregate labour supply and suppress

[24] It is possible to establish a partial converse of this result. Specifically, if $n(k, \rho, \varepsilon)$ is convex (or concave) in ρ over the interval $\left[\underline{\rho}, \overline{\rho}\right]$, then $v'(\cdot)$ is concave (or convex) between $n\left(k, \underline{\rho}, \varepsilon\right)$ and $n(k, \overline{\rho}, \varepsilon)$ [see the proof of Proposition 8 for further details]. This suggests that the condition of a concave (or convex) $v'(\cdot)$ is rather tight.

[25] Studies that use the "no-income-effect" preferences typically assume that $v(\cdot)$ has an isoelastic form, e.g., $v(n) \equiv An^{1+\theta}$, with $A > 0$ and $\theta > 0$. The first derivative of this function is strictly concave (or strictly convex) if and only if $\theta < 1$ (or $\theta > 1$). A less-than-unity value of θ seems to be more common in the existing literature. For instance, Greenwood *et al.* (1988), Jaimovich and Rebelo (2009) and Correia (2010) have used values ranging from 0.4 to 0.8.

the capital-labour ratio. Which effect will prevail is a quantitative question. The other three cases in Table 1 can be interpreted in a similar way.

4.3 Numerical Examples

In Sections 4.1 and 4.2, we have identified two mechanisms through which greater heterogeneity in time preference can affect the macroeconomy. The first one manifests itself by affecting the cross-sectional distribution of marginal tax rates and the national income function, while the second one operates through the aggregate labour supply function. In this subsection, we will use some numerical examples to illustrate the working of these forces in the full version of the baseline model. There are two specific reasons why we resort to quantitative analysis here. First, the presence of income effect poses a serious challenge in characterising the shape of $n(k, \rho, \varepsilon)$ as a function in ρ. Without knowing this, we cannot ascertain qualitatively the effect of greater time preference heterogeneity on $N(\cdot)$ as in Proposition 8. Second, as the results in Table 1 indicate, the overall effects of greater heterogeneity in ρ are often ambiguous due to simultaneous changes in $Y(\cdot)$ and $N(\cdot)$. The use of numerical calculations can help shed some light on these issues.

Consider a parameterised version of the baseline model with the following specifics: One period in the model is a year. The consumer's period utility function is given by

$$U(c, n) = \ln c - A\frac{n^{1+1/\psi}}{1+1/\psi},$$

where A is a positive-valued parameter and ψ is the Frisch elasticity of labour supply. The value of A is calibrated so that, on average, consumers spend about one-third of their time on work in the steady state. The resulting value of A is 54.30. The Frisch elasticity of labour supply is set to 0.40, which is consistent with the estimates obtained by MaCurdy (1981) and Altonji (1986). The production function is assumed to take the Cobb-Douglas form, i.e., $F(K, L) = K^{\alpha}L^{1-\alpha}$, with $\alpha = 0.40$. We choose the value of δ so that the steady-state capital-output ratio matches its empirical counterpart in the United States over the period 1953-2009, which is 2.427.[26] This requires $\delta = 8.74\%$. The progressive tax function is assumed to take the form in (20), with $a_1 = 0.768$ and $a_2 = 0.031$. These values are taken from Gouveia and Strauss (1994). The value of a_0 is calibrated so that the average tax rate, defined as the ratio between total tax revenues and total taxable income, is 14.18%. This matches the average tax rate in the United

[26] To calculate this value, we use the sum of private fixed assets and the end-of-year stock of private inventories as our measure of aggregate capital. Data on private fixed assets and private inventories are obtained from the Bureau of Economic Analysis website.

States over the period 1986-2014.[27] The required value of a_0 is 0.323.[28] Under this choice of (a_0, a_1, a_2), the implied marginal tax function is strictly concave, i.e., $\tau'''(\cdot) < 0$. In terms of consumer characteristics, we assume that both ρ and ε are uniformly distributed across consumers with $\underline{\rho} = 0.0571$, $\overline{\rho} = 0.0753$, $\underline{\varepsilon} = 1$ and $\overline{\varepsilon} = 100$. The implied maximum and minimum value of β are 0.946 and 0.930, respectively. Similar range of values of (β, ε) have also appeared in Carroll and Young (2011).[29] The benchmark parameter values are summarised in Table 2. The resulting value of five key variables, namely the capital-output ratio k^*, aggregate labour $N(k^*)$, national income $Y(k^*)$, aggregate capital $k^* N(k^*)$ and aggregate output $(k^*)^\alpha \cdot N(k^*)$, are reported in Table 3.[30]

The next step is to construct some alternative distributions of ρ with different degrees of time preference heterogeneity. Intuitively, a mean-preserving spread of the benchmark uniform distribution can be obtained by "hollowing out" the middle section and relocating the mass to the upper and lower ends. To put this in practice, first define $\Delta \equiv \left(\overline{\rho} - \underline{\rho}\right)$, $\rho_1 \equiv \underline{\rho} + \Delta/3$ and $\rho_2 \equiv \underline{\rho} + 2\Delta/3$. Using ρ_1 and ρ_2 we can partition the distribution of ρ into three groups of equal mass. In the examples presented below, we focus on distributions of ρ with density

$$\widetilde{h}_1(\rho) = \begin{cases} \chi_1/\Delta & \text{for } \rho \in \left[\underline{\rho}, \rho_1\right), \\ \theta/\Delta & \text{for } \rho \in [\rho_1, \rho_2), \\ \chi_2/\Delta & \text{for } \rho \in [\rho_2, \overline{\rho}], \end{cases}$$

for some positive-valued parameters θ, χ_1 and χ_2. The benchmark uniform distribution corresponds to the case when $\theta = \chi_1 = \chi_2 = 1$. A mean-preserving spread can be obtained by lowering θ and solving (χ_1, χ_2) from

$$\int_{\underline{\rho}}^{\overline{\rho}} \widetilde{h}_1(\rho)\, d\rho = 1 \qquad \text{and} \qquad \int_{\underline{\rho}}^{\overline{\rho}} \rho \widetilde{h}_1(\rho)\, d\rho = \frac{\overline{\rho} + \underline{\rho}}{2}.$$

We then solve the baseline model under four alternative distributions with $\theta \in \{0, 0.5, 0.95, 1.5\}$,

[27] Data on this measure of average tax rate are readily available from Figure A on page 31 of *Statistics of Income — 2014 Individual Income Tax Returns*, Internal Revenue Service, Washington, D.C.

[28] The marginal tax function associated with (20) is given by $\tau'(y) = a_0 \left[1 - (1 + a_2 y^{a_1})^{-(1+1/a_1)}\right]$. As y tends to infinity, the marginal tax rate becomes a_0 in the limit. Hence, we have $\overline{\tau} = a_0 = 0.323$.

[29] Unlike Carroll and Young (2011), we do not calibrate the distribution of consumer characteristics to match the empirical distribution of income and wealth. The main purpose of the numerical examples is to demonstrate how changes in the distribution of ρ will affect the functions $Y(\cdot)$ and $N(\cdot)$ in the baseline model, rather than to mimic the observed patterns of inequality. For this reason, we opt for the most parsimonious distribution, which is the uniform distribution. One important consideration behind the choice of $\left\{\underline{\rho}, \overline{\rho}, \underline{\varepsilon}, \overline{\varepsilon}\right\}$ is the existence of steady state, i.e., to ensure that the conditions in (15) and (16) are satisfied.

[30] The MATLAB codes for generating the reported results are available on the authors' personal website.

Table 2 Benchmark Parameter Values

A	Preference parameter	54.30
ψ	Frisch elasticity of labour supply	0.400
α	Share of capital income in total output	0.400
δ	Depreciation rate of capital	0.0874
a_0	Parameter in the progressive tax function	0.323
a_1	Parameter in the progressive tax function	0.768
a_2	Parameter in the progressive tax function	0.031
$\underline{\rho}$	Minimum value of rate of time preference	0.0571
$\overline{\rho}$	Maximum value of rate of time preference	0.0753
$\underline{\varepsilon}$	Minimum value of labour productivity	1.0
$\overline{\varepsilon}$	Maximum value of labour productivity	100.0

Table 3 Results of Numerical Examples

	Benchmark	% Changes from Benchmark			
	$\theta = 1.0$	$\theta = 0$	$\theta = 0.5$	$\theta = 0.95$	$\theta = 1.5$
k^*	4.384	0.27%	0.15%	0.02%	-0.17%
$N(k^*)$	19.478	9.72%	4.83%	0.48%	-4.74%
$Y(k^*)$	27.719	9.79%	4.86%	0.48%	-4.78%
$k^* N(k^*)$	85.397	10.02%	4.98%	0.49%	-4.90%
$(k^*)^{\alpha} \cdot N(k^*)$	35.180	9.84%	4.89%	0.49%	-4.80%

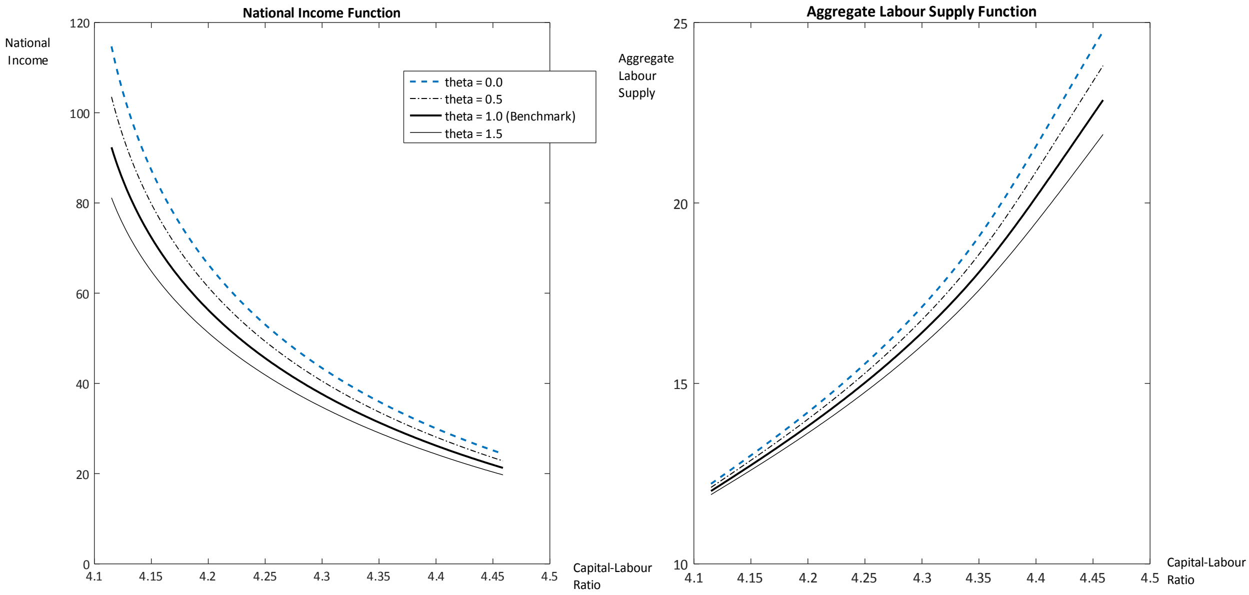

Figure 2

while keeping all other parameters and the distribution of ε unchanged. In general, a lower value of θ indicates a higher level of heterogeneity. Thus, the distribution with $\theta = 1.50$ is actually *less* diverse than the benchmark uniform distribution.

Figure 2 shows the national income function $Y(\cdot)$ and the aggregate labour supply function $N(\cdot)$ obtained under various values of θ, including the benchmark case.[31] Two results are immediate from these diagrams. First, increasing the cross-sectional dispersion of ρ will shift up the national income function over the entire range of $[k_{\min}, k_{\max}]$. This pattern is consistent with the theoretical predictions in Proposition 5. Second, an increase in time preference heterogeneity will also shift up the aggregate labour supply function. In a robustness check, we find the same pattern under different values of ψ within the range of $[0, 1]$.[32] Thus, at least in this regard, the baseline model considered in these numerical examples is similar to the "no-income-effect" model with a concave $v'(\cdot)$. The value of k^* and other aggregate variables under these alternative distributions are reported in Table 3. Although these results are only suggestive in nature, they do show some interesting patterns. For instance, increasing the dispersion of time preference tends to have a

[31] The curves for $\theta = 0.95$ are almost indiscernible from those obtained from the baseline case, hence they are omitted from these diagrams.

[32] These results are not shown here due to space constraint, but they are reported in the working paper version available on the authors' personal website.

very mild positive effect on k^*. This suggests that the effect captured by Proposition 5 is slightly stronger than the one described in Proposition 8. The changes in other aggregate variables then largely reflect the changes in $N(k^*)$. In particular, a more dispersed distribution of ρ is associated with a higher level of aggregate labour input. These results also point to an important difference between an exogenous-labour model and a flexible-labour model. In the former, an increase in time preference heterogeneity will have no effect on aggregate labour supply but a significantly positive effect on k^* [due to the upward shift in $Y(\cdot)$]. In the latter, such an increase will only have a negligible effect on k^* but a significant positive effect on aggregate labour supply.

5 Conclusion

In this paper we analyse the long-run economic effects of diversity in a neoclassical model with *ex ante* heterogeneous consumers, flexible labour supply and progressive taxation. Our results highlight two important channels through which consumer heterogeneity can affect the steady state. First, changing either the distribution of labour productivity or time preference will affect the formation of aggregate labour supply. The exact nature of this effect is determined by the shape of the individual labour supply function, which is often difficult to determine qualitatively. Second, changing the distribution of time preference will have an effect on the distribution of marginal tax rates across individuals. It is shown that the concavity or convexity of the marginal tax function holds the key in determining these effects. In this analysis, we assume that time preference and labour productivity are independent of each other. This assumption is adopted mainly for analytical convenience. As pointed out by Carroll and Young (2009), such a model may fail to capture the observed patterns of correlation between different types of income. One possible direction of future research is to analyse the effects of diversity without imposing the independence assumption. The model considered here also does not take into account the political institutions that contribute to the progressive tax system or other redistributive policies. As discussed in Alesina and La Ferrara (2005), these institutions play a crucial role in resolving the conflicting interests within a diverse population, and this will in turn determine the economic effects of diversity. One exciting and important direction of future research is to introduce some political elements (such as a voting mechanism) into our baseline model and analyse the effects of diversity in a politico-economic equilibrium.

Appendix

Proof of Lemma 1

Define $\Gamma(k) \equiv f(k) - \delta k$ over the interval $[k_{\min}, k_{\max}]$. Then the steady-state condition in (13) can be more succinctly expressed as $Y(k) = \Gamma(k) N(k)$. We will establish some basic properties of each of these functions, starting with $\Gamma(\cdot)$. Since $f(\cdot)$ is strictly increasing and strictly concave, there exists a unique value $k_{GR} > 0$ such that $\Gamma'(k) \gtreqless 0$ if and only if $k \lesseqgtr k_{GR}$. Since $\Gamma'(k_{\max}) = f'(k_{\max}) - \delta = \overline{\rho} > 0$, we have $k_{\max} < k_{GR}$ which means $\Gamma(\cdot)$ is strictly increasing over $[k_{\min}, k_{\max}]$ with $\Gamma(k_{\min}) > 0$. Next, we turn to the national income function $Y(\cdot)$. Since $\phi(\cdot)$ is strictly increasing, $Y(\cdot)$ is strictly decreasing on $(k_{\min}, k_{\max})$ with

$$Y(k_{\max}) = \int_{\underline{\rho}}^{\overline{\rho}} \phi\left(1 - \frac{\rho}{\overline{\rho}}\right) dH_1(\rho) > 0,$$

$$\lim_{k \to k_{\min}} Y(k) = \int_{\underline{\rho}}^{\overline{\rho}} \phi\left[1 - \frac{\rho}{\underline{\rho}}(1 - \overline{\tau})\right] dH_1(\rho). \tag{23}$$

Equation (23) follows from the fact that $r(k)$ approaches $\underline{\rho}/(1 - \overline{\tau})$ as k tends to $k_{\min}$. Note that the limiting condition $\lim_{y \to \infty} \tau'(y) = \overline{\tau}$ implies

$$\lim_{\rho \to \underline{\rho}} \phi\left[1 - \frac{\rho}{\underline{\rho}}(1 - \overline{\tau})\right] = +\infty.$$

Thus, the integral in (23) is potentially divergent. For instance, if $H_1(\cdot)$ has a positive mass at $\underline{\rho}$, then $Y(k_{\min})$ is infinitely large and the condition in (16) is automatically satisfied.

Finally, we will show that the aggregate labour supply function $N(\cdot)$ is non-decreasing. It suffice to show that $n(k, \rho, \varepsilon)$ is a non-decreasing function in k, for any (ρ, ε). Fix (ρ, ε) and suppose the contrary that $1 \geq n(k_2, \rho, \varepsilon) > n(k_1, \rho, \varepsilon) > 0$ for some $k_1 > k_2$ in $[k_{\min}, k_{\max}]$. Since $c(k, \rho)$ is strictly decreasing in k, we have $c(k_1, \rho) < c(k_2, \rho)$. By Assumption A2, $\Psi(c, n)$ is non-decreasing in c and strictly increasing in n. Hence, we have

$$\begin{aligned}
\Psi[c(k_1, \rho), n(k_1, \rho, \varepsilon)] &\leq \Psi[c(k_2, \rho), n(k_1, \rho, \varepsilon)] \\
&< \Psi[c(k_2, \rho), n(k_2, \rho, \varepsilon)] \\
&\leq \frac{w(k_2)}{r(k_2)} \varepsilon\rho < \frac{w(k_1)}{r(k_1)} \varepsilon\rho.
\end{aligned}$$

The third inequality follows from (10) and the last one uses the facts that $w(\cdot)$ is strictly increasing and $r(\cdot)$ is strictly decreasing. The above condition implies that $n(k_1, \rho, \varepsilon) = 1$ which contradicts $1 \geq n(k_2, \rho, \varepsilon) > n(k_1, \rho, \varepsilon) > 0$. Hence, $n(k, \rho, \varepsilon)$ is a non-decreasing function in k for all possible values of (ρ, ε).

In sum, we have shown that $[Y(k) - \Gamma(k) N(k)]$ is strictly decreasing over the interval $(k_{\min}, k_{\max})$. If (15) and (16) are satisfied, then there exists a unique value k^* within this range that solves (13). Conversely, if this equation has a unique interior solution, then the two curves in Figure 1a must cross once over the interval $(k_{\min}, k_{\max})$, which implies (15) and (16). This completes the proof of Lemma 1.

Proof of Lemma 2

Fix $k \in [k_{\min}, k_{\max}]$ and $\rho \in [\underline{\rho}, \overline{\rho}]$. Suppose the contrary that $1 \geq n(k, \rho, \varepsilon_2) > n(k, \rho, \varepsilon_1) > 0$ for some $\varepsilon_1 > \varepsilon_2$ in $[\underline{\varepsilon}, \overline{\varepsilon}]$. Since $\Psi(c, n)$ is strictly increasing in n, we have

$$\Psi[c(k, \rho), n(k, \rho, \varepsilon_1)] < \Psi[c(k, \rho), n(k, \rho, \varepsilon_2)] \leq \frac{w(k)}{r(k)}\rho\varepsilon_2 < \frac{w(k)}{r(k)}\rho\varepsilon_1,$$

where the second inequality follows from (10). The above condition implies $n(k, \rho, \varepsilon_1) = 1$, which contradicts the hypothesis of $1 \geq n(k, \rho, \varepsilon_2) > n(k, \rho, \varepsilon_1) > 0$. Hence, $n(k, \rho, \varepsilon)$ is a non-decreasing function in ε for all $k \in [k_{\min}, k_{\max}]$ and for all $\rho \in [\underline{\rho}, \overline{\rho}]$.

If $n(k, \rho, \varepsilon)$ is an interior solution, then it is completely characterised

$$\Psi[c(k, \rho), n(k, \rho, \varepsilon)] = \frac{w(k)}{r(k)}\rho\varepsilon.$$

By Assumptions A1-A2 and the implicit function theorem, $n(k, \rho, \varepsilon)$ is continuously differentiable in ε. Straightforward differentiation then yields

$$\frac{\partial \Psi}{\partial n}\frac{\partial n(k, \rho, \varepsilon)}{\partial \varepsilon} = \frac{w(k)}{r(k)}\rho > 0.$$

Since $\partial\Psi/\partial n > 0$, the desired result follows. This completes the proof of Lemma 2.

Proof of Proposition 3

We first establish an intermediate result.[33]

[33] The proof of Lemma A1 has been outlined in Shaked and Shanthikumar (2007, p.204-205). We include a more detailed proof here for the sake of clarity and completeness.

Lemma A1 *For any bounded, non-decreasing function* $g(\cdot)$ *defined on* $[\underline{\varepsilon}, \overline{\varepsilon}]$,

$$\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon g(\varepsilon)\, dH_2(\varepsilon) \leq \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon g(\varepsilon)\, d\widetilde{H}_2(\varepsilon), \tag{24}$$

if and only if (17) holds.

Proof of Lemma A1 The proof of the "only if" part is obvious. Suppose (24) is valid for all bounded, non-decreasing functions defined on $[\underline{\varepsilon}, \overline{\varepsilon}]$. For any $x \in [\underline{\varepsilon}, \overline{\varepsilon}]$, define the indicator function $\mathcal{I}(\varepsilon; x)$ which equals one if $\varepsilon \geq x$ and zero otherwise. Since $\mathcal{I}(\varepsilon; x)$ is bounded and non-decreasing, it follows from (24) that

$$\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon \mathcal{I}(\varepsilon; x)\, dH_2(\varepsilon) = \int_{x}^{\overline{\varepsilon}} \varepsilon dH_2(\varepsilon) \leq \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon \mathcal{I}(\varepsilon; x)\, d\widetilde{H}_2(\varepsilon) = \int_{x}^{\overline{\varepsilon}} \varepsilon d\widetilde{H}_2(\varepsilon),$$

for any $x \in [\underline{\varepsilon}, \overline{\varepsilon}]$. Next, consider the "if" part. Let $g(\cdot)$ be an arbitrary bounded, non-decreasing function defined on $[\underline{\varepsilon}, \overline{\varepsilon}]$. Without loss of generality, we can assume $g(\underline{\varepsilon}) = 0$. For any positive integer $m \geq 1$, partition the interval $[\underline{\varepsilon}, \overline{\varepsilon}]$ into 2^m subintervals of equal length. Specifically, define a set of end-points $\{\widetilde{\varepsilon}_{i,m}\}$ according to

$$\widetilde{\varepsilon}_{i,m} = \underline{\varepsilon} + \frac{i-1}{2^m}(\overline{\varepsilon} - \underline{\varepsilon}), \qquad \text{for } i = 1, ..., 2^m + 1.$$

Define a function $\eta_m(\cdot)$ as follows

$$\eta_m(\varepsilon) = \begin{cases} g(\widetilde{\varepsilon}_{i,m}) & \text{if } \varepsilon \in [\widetilde{\varepsilon}_{i,m}, \widetilde{\varepsilon}_{i+1,m}), \\ g(\overline{\varepsilon}) & \text{if } \varepsilon = \overline{\varepsilon}. \end{cases}$$

This function can be expressed as a linear combination of simple functions, i.e.,

$$\eta_m(\varepsilon) = \sum_{i=1}^{m2^m} \lambda_{i,m} \mathcal{I}(\varepsilon; \widetilde{\varepsilon}_{i,m}), \tag{25}$$

where $\mathcal{I}(\varepsilon; \widetilde{\varepsilon}_{i,m}) = 1$ if $\varepsilon \geq \widetilde{\varepsilon}_{i,m}$ and zero otherwise. The coefficient $\lambda_{i,m}$ is given by

$$\lambda_{i,m} = \begin{cases} g(\underline{\varepsilon}) & \text{for } i = 1, \\ g(\widetilde{\varepsilon}_{i+1,m}) - g(\widetilde{\varepsilon}_{i,m}) & \text{for } i = 2, ..., 2^m. \end{cases}$$

Since $g(\cdot)$ is non-decreasing and non-negative, we have $\lambda_{i,m} \geq 0$ for all i. Hence, $\eta_m(\varepsilon) \geq 0$ for all $\varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon}]$.

By repeating the above procedure, we can construct a sequence of non-negative functions $\{\eta_m(\cdot)\}$ that converges pointwise to $g(\cdot)$. We now show that $\{\eta_m(\cdot)\}$ is a monotonically increasing sequence of functions, i.e., $\eta_m(\varepsilon) \leq \eta_{m+1}(\varepsilon)$ for any $\varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon}]$. Fix $\varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon}]$. Then there are only two possible scenarios: either $\varepsilon < (\widetilde{\varepsilon}_{i,m} + \widetilde{\varepsilon}_{i+1,m})/2$ or $\varepsilon \geq (\widetilde{\varepsilon}_{i,m} + \widetilde{\varepsilon}_{i+1,m})/2$ for some $i \in \{1, ..., 2^m\}$. In the first scenario, we have $\eta_{m+1}(\varepsilon) = \eta_m(\varepsilon)$. In the second scenario,

$$\eta_{m+1}(\varepsilon) = \eta_m\left(\frac{\widetilde{\varepsilon}_{i,m} + \widetilde{\varepsilon}_{i+1,m}}{2}\right) \geq g(\widetilde{\varepsilon}_{i,m}) = \eta_m(\varepsilon).$$

Hence, $\{\eta_m(\cdot)\}$ is a monotonically increasing sequence of non-negative functions. By the monotone convergence theorem, we have

$$\lim_{m\to\infty} \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon \eta_m(\varepsilon)\, dH_2(\varepsilon) = \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon g(\varepsilon)\, dH_2(\varepsilon), \tag{26}$$

$$\lim_{m\to\infty} \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon \eta_m(\varepsilon)\, d\widetilde{H}_2(\varepsilon) = \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon g(\varepsilon)\, d\widetilde{H}_2(\varepsilon). \tag{27}$$

Note that for each m, we have

$$\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon \eta_m(\varepsilon)\, dH_2(\varepsilon) = \sum_{i=1}^{m2^m} \lambda_{i,m} \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon \mathcal{I}(\varepsilon; \widetilde{\varepsilon}_{i,m})\, dH_2(\varepsilon) = \sum_{i=1}^{m2^m} \lambda_{i,m} \int_{\widetilde{\varepsilon}_{i,m}}^{\overline{\varepsilon}} \varepsilon dH_2(\varepsilon),$$

where the first equality follows from (25). Suppose (17) is true for all $x \in [\underline{\varepsilon}, \overline{\varepsilon}]$. Since $\lambda_{i,m} \geq 0$ for all i, we have

$$\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon \eta_m(\varepsilon)\, dH_2(\varepsilon) \leq \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon \eta_m(\varepsilon)\, d\widetilde{H}_2(\varepsilon).$$

Equation (24) then follows from (26) and (27). This completes the proof of Lemma A1.

Fix $k \in [k_{\min}, k_{\max}]$. By Lemma 2, $n(k, \rho, \varepsilon)$ is non-decreasing in ε for all $\rho \in [\underline{\rho}, \overline{\rho}]$. Thus, after integrating it over the distribution of ρ, the resulting function

$$\mathcal{N}(k, \varepsilon) \equiv \int_{\underline{\rho}}^{\overline{\rho}} n(k, \rho, \varepsilon)\, dH_1(\rho)$$

is bounded and non-decreasing in ε. Then by Lemma A1,

$$N(k) \equiv \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon \mathcal{N}(k, \varepsilon)\, dH_2(\varepsilon) \leq \widetilde{N}(k) \equiv \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon \mathcal{N}(k, \varepsilon)\, d\widetilde{H}_2(\varepsilon)$$

if and only if (17) holds. This completes the proof of Proposition 3.

Proof of Proposition 4

Let k^* and $\widetilde{k}^*$ be the unique solution of (13) under $H_2(\cdot)$ and $\widetilde{H}_2(\cdot)$, respectively. As shown in the proof of Lemma 1, $n(k,\rho,\varepsilon)$ is non-decreasing in k for all (ρ,ε). Hence, $N(\cdot)$ and $\widetilde{N}(\cdot)$ are both non-decreasing functions. Suppose the contrary that $k^* < \widetilde{k}^*$. Then we have

$$Y\left(\widetilde{k}^*\right) < Y(k^*) = \Gamma(k^*) N(k^*) < \Gamma\left(\widetilde{k}^*\right) N\left(\widetilde{k}^*\right) \leq \Gamma\left(\widetilde{k}^*\right) \widetilde{N}\left(\widetilde{k}^*\right). \tag{28}$$

The first inequality uses the fact that $Y(\cdot)$ is a strictly decreasing function as shown in the proof of Lemma 1. The second inequality uses the fact that $\Gamma(\cdot)$ and $N(\cdot)$ are both increasing functions. The last inequality uses the result in Proposition 3. Condition (28), however, contradicts the hypothesis that $Y\left(\widetilde{k}^*\right) = \Gamma\left(\widetilde{k}^*\right)\widetilde{N}\left(\widetilde{k}^*\right)$. Hence, it must be the case that $k^* \geq \widetilde{k}^*$. Since $y(k,\rho)$ and $c(k,\rho)$ are strictly decreasing in k for all $\rho \in [\underline{\rho},\overline{\rho}]$, we have $y(k^*,\rho) \leq y\left(\widetilde{k}^*,\rho\right)$ and $c(k^*,\rho) \leq c\left(\widetilde{k}^*,\rho\right)$. Similarly, we have $Y(k^*) \leq Y\left(\widetilde{k}^*\right)$ as $Y(\cdot)$ is strictly decreasing. Finally, the implications on aggregate labour input can be seen as follows:

$$N(k^*) = \frac{Y(k^*)}{\Gamma(k^*)} \leq \frac{Y\left(\widetilde{k}^*\right)}{\Gamma\left(\widetilde{k}^*\right)} = \widetilde{N}\left(\widetilde{k}^*\right).$$

This completes the proof of Proposition 4.

Proof of Proposition 5

The proof is built upon the following intermediate result:

Lemma A2 *Suppose Assumption A3 is satisfied. Then $\phi(\cdot)$ is a convex (or concave) function if and only if $\tau'(\cdot)$ is concave (or convex).*

Proof of Lemma A2 Pick any two positive real numbers y_1 and y_2, and any $\alpha \in (0,1)$. Then

$$\tau'(\alpha y_1 + (1-\alpha) y_2) \gtreqless \alpha\tau'(y_1) + (1-\alpha)\tau'(y_2)$$

$$\Leftrightarrow \phi\left[\tau'(\alpha y_1 + (1-\alpha) y_2)\right] \gtreqless \phi\left[\alpha\tau'(y_1) + (1-\alpha)\tau'(y_2)\right]$$

$$\Leftrightarrow \alpha y_1 + (1-\alpha) y_2 \gtreqless \phi\left[\alpha\tau'(y_1) + (1-\alpha)\tau'(y_2)\right]$$

$$\Leftrightarrow \alpha\phi\left[\tau'\left(y_1\right)\right] + \left(1-\alpha\right)\phi\left[\tau'\left(y_2\right)\right] \gtreqless \phi\left[\alpha\tau'\left(y_1\right) + \left(1-\alpha\right)\tau'\left(y_2\right)\right].$$

The second line uses the fact that $\phi(\cdot)$ is strictly increasing. The third and fourth lines follow from the identity $\phi\left[\tau'\left(y\right)\right] = y$. Hence, $\phi(\cdot)$ is a convex (or concave) function if and only if $\tau'(\cdot)$ is concave (or convex). This completes the proof of Lemma A2.

Suppose $\widetilde{H}_1(\cdot)$ is a mean-preserving spread of $H_1(\cdot)$ and $\tau'(\cdot)$ is convex so that $\phi(\cdot)$ is concave. Then we can write

$$Y\left(k\right) \equiv \int_{\underline{\rho}}^{\overline{\rho}} \phi\left[1 - \frac{\rho}{r\left(k\right)}\right] dH_1\left(\rho\right) \geq \widetilde{Y}\left(k\right) \equiv \int_{\underline{\rho}}^{\overline{\rho}} \phi\left[1 - \frac{\rho}{r\left(k\right)}\right] d\widetilde{H}_1\left(\rho\right),$$

for any $k \in \left(k_{\min}, k_{\max}\right)$. In other words, changing the distribution of time preference from $H_1(\cdot)$ to $\widetilde{H}_1(\cdot)$ will shift the $Y(k)$ curve in Figure 1a to the left. Thus, we have $\widetilde{k}^* \leq k^*$. A similar argument can be used to establish the results in part (ii).

Proof of Proposition 6

For any $q \in [0,1]$, define $\sigma(q) \equiv \sup\left\{\rho : H_1(\rho) \leq q\right\}$ and $\widetilde{\sigma}(q) \equiv \sup\left\{\rho : \widetilde{H}_1(\rho) \leq q\right\}$. According to (3.A.41) and (3.A.42) in Shaked and Shanthikumar (2007, p.118), $\widetilde{H}_1(\cdot)$ is a mean-preserving spread of $H_1(\cdot)$ if and only if

$$\frac{\int_{\sigma(q)}^{\overline{\rho}} \rho dH_1\left(\rho\right)}{\int_{\sigma(q)}^{\overline{\rho}} dH_1\left(\rho\right)} \leq \frac{\int_{\widetilde{\sigma}(q)}^{\overline{\rho}} \rho d\widetilde{H}_1\left(\rho\right)}{\int_{\widetilde{\sigma}(q)}^{\overline{\rho}} d\widetilde{H}_1\left(\rho\right)}, \qquad \text{for any } q \in [0,1];$$

$$\frac{\int_{\underline{\rho}}^{\sigma(q)} \rho dH_1\left(\rho\right)}{\int_{\underline{\rho}}^{\sigma(q)} dH_1\left(\rho\right)} \geq \frac{\int_{\underline{\rho}}^{\widetilde{\sigma}(q)} \rho d\widetilde{H}_1\left(\rho\right)}{\int_{\underline{\rho}}^{\widetilde{\sigma}(q)} d\widetilde{H}_1\left(\rho\right)}, \qquad \text{for any } q \in [0,1].$$

Since $\int_{\theta(q)}^{\overline{\rho}} dH_1\left(\rho\right) = \int_{\widetilde{\theta}(q)}^{\overline{\rho}} d\widetilde{H}_1\left(\rho\right) = 1 - q$ and $\int_{\underline{\rho}}^{\sigma(q)} dH_1\left(\rho\right) = \int_{\underline{\rho}}^{\widetilde{\sigma}(q)} d\widetilde{H}_1\left(\rho\right) = q$, these conditions can be more succinctly expressed as

$$\int_{\sigma(q)}^{\overline{\rho}} \rho dH_1\left(\rho\right) \leq \int_{\widetilde{\sigma}(q)}^{\overline{\rho}} \rho d\widetilde{H}_1\left(\rho\right), \tag{29}$$

$$\int_{\underline{\rho}}^{\sigma(q)} \rho dH_1\left(\rho\right) \geq \int_{\underline{\rho}}^{\widetilde{\sigma}(q)} \rho d\widetilde{H}_1\left(\rho\right), \tag{30}$$

for all $q \in [0,1]$.

Using (7) and (29), we can write

$$\begin{aligned}\int_{\sigma(q)}^{\overline{\rho}} \tau'\left[y\left(k,\rho\right)\right] dH_1\left(\rho\right) &= \int_{\sigma(q)}^{\overline{\rho}} dH_1\left(\rho\right) - \frac{1}{r\left(k\right)} \int_{\sigma(q)}^{\overline{\rho}} \rho dH_1\left(\rho\right) \\ &\geq \int_{\widetilde{\sigma}(q)}^{\overline{\rho}} d\widetilde{H}_1\left(\rho\right) - \frac{1}{r\left(k\right)} \int_{\widetilde{\sigma}(q)}^{\overline{\rho}} \rho d\widetilde{H}_1\left(\rho\right) = \int_{\widetilde{\sigma}(q)}^{\overline{\rho}} \tau'\left[y\left(k,\rho\right)\right] d\widetilde{H}_1\left(\rho\right),\end{aligned}$$

for any $k \in \left(k_{\min}, k_{\max}\right)$. If the marginal tax function is concave so that $\widetilde{k}^* \geq k^*$, then we have

$$\begin{aligned}\int_{\sigma(q)}^{\overline{\rho}} \tau'\left[y\left(k^*,\rho\right)\right] dH_1\left(\rho\right) &\geq \int_{\widetilde{\sigma}(q)}^{\overline{\rho}} \tau'\left[y\left(k^*,\rho\right)\right] d\widetilde{H}_1\left(\rho\right) \\ &\geq \int_{\widetilde{\sigma}(q)}^{\overline{\rho}} \tau'\left[y\left(\widetilde{k}^*,\rho\right)\right] d\widetilde{H}_1\left(\rho\right).\end{aligned}$$

The second inequality uses the fact that $\tau'\left(\cdot\right)$ is strictly increasing and $y\left(k,\rho\right)$ is strictly decreasing in k. The results in part (ii) can be similarly obtained by using (7) and (30).

Proof of Lemma 7

Suppose the contrary that $n\left(k,\rho,\varepsilon\right) = 1$ for some $k \in \left[k_{\min}, k_{\max}\right]$ and for some $\left(\rho,\varepsilon\right) \in \left[\underline{\rho},\overline{\rho}\right] \times \left[\underline{\varepsilon},\overline{\varepsilon}\right]$. Then using (22), we can write

$$w\left(k_{\max}\right)\overline{\varepsilon} < v'\left(1\right) \leq \frac{w\left(k\right)}{r\left(k\right)}\varepsilon\rho.$$

By Assumption A4, $w\left(\cdot\right)$ is strictly increasing and $r\left(\cdot\right)$ is strictly decreasing. Thus, for any $k \in \left[k_{\min}, k_{\max}\right]$ and $\left(\rho,\varepsilon\right) \in \left[\underline{\rho},\overline{\rho}\right] \times \left[\underline{\varepsilon},\overline{\varepsilon}\right]$,

$$\frac{w\left(k\right)}{r\left(k\right)}\varepsilon\rho < \frac{w\left(k_{\max}\right)}{r\left(k_{\max}\right)}\overline{\rho} \cdot \overline{\varepsilon} = w\left(k_{\max}\right)\overline{\varepsilon}.$$

The last equality follows from the fact that $r\left(k_{\max}\right) = \overline{\rho}$. Note that these two expressions are contradictory. Hence, we have $n\left(k,\rho,\varepsilon\right) < 1$ for all $k \in \left(k_{\min}, k_{\max}\right)$ and for all $\left(\rho,\varepsilon\right) \in \left[\underline{\rho},\overline{\rho}\right] \times \left[\underline{\varepsilon},\overline{\varepsilon}\right]$. This completes the proof of Lemma 7.

Proof of Proposition 8

Fix $k \in \left[k_{\min}, k_{\max}\right]$ and $\varepsilon \in \left[\underline{\varepsilon},\overline{\varepsilon}\right]$. Suppose $v'\left(1\right) > w\left(k_{\max}\right)\overline{\varepsilon}$ is satisfied. Then by Lemma 7, we have

$$v'\left[n\left(k,\rho,\varepsilon\right)\right] = \frac{w\left(k\right)}{r\left(k\right)}\varepsilon\rho,$$

for all $\rho \in \left[\underline{\rho}, \overline{\rho}\right]$. Pick any ρ_1 and ρ_2 in $\left[\underline{\rho}, \overline{\rho}\right]$ and define $\rho_\alpha = \alpha\rho_1 + (1-\alpha)\rho_2$ for any $\alpha \in (0,1)$. Suppose $v'(\cdot)$ is concave, then we have

$$\begin{aligned} v'\left[n\left(k,\rho_\alpha,\varepsilon\right)\right] &= \frac{w(k)}{r(k)}\varepsilon\rho_\alpha = \alpha v'\left[n\left(k,\rho_1,\varepsilon\right)\right] + (1-\alpha)v'\left[n\left(k,\rho_2,\varepsilon\right)\right] \\ &\leq v'\left[\alpha n\left(k,\rho_1,\varepsilon\right) + (1-\alpha)n\left(k,\rho_2,\varepsilon\right)\right]. \end{aligned}$$

Since $v'(\cdot)$ is strictly increasing, the last inequality implies $n(k,\rho,\varepsilon)$ is convex in ρ, i.e.,

$$n\left(k,\rho_\alpha,\varepsilon\right) \leq \alpha n\left(k,\rho_1,\varepsilon\right) + (1-\alpha)n\left(k,\rho_2,\varepsilon\right).$$

A "partial" converse can be obtained by reversing these steps, i.e., if $n(k,\rho,\varepsilon)$ is convex in ρ over the entire interval $\left[\underline{\rho}, \overline{\rho}\right]$, then $v'(n)$ is concave between $n\left(k,\underline{\rho},\varepsilon\right)$ and $n\left(k,\overline{\rho},\varepsilon\right)$. This is just a partial converse because the range between $n\left(k,\underline{\rho},\varepsilon\right)$ and $n\left(k,\overline{\rho},\varepsilon\right)$ may not cover the entire domain of $v'(\cdot)$.

Since convexity is preserved by integration, this means $\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon n(k,\rho,\varepsilon)\, dH_2(\varepsilon)$ is also a convex function in ρ. Finally, since $\widetilde{H}_1(\cdot)$ is a mean-preserving spread of $H_1(\cdot)$, we have

$$N(k) \equiv \int_{\underline{\rho}}^{\overline{\rho}} \left[\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon n(k,\rho,\varepsilon)\, dH_2(\varepsilon)\right] dH_1(\rho) \leq \int_{\underline{\rho}}^{\overline{\rho}} \left[\int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon n(k,\rho,\varepsilon)\, dH_2(\varepsilon)\right] d\widetilde{H}_1(\rho) \equiv \widetilde{N}(k),$$

for all $k \in [k_{\min}, k_{\max}]$. The results in part (ii) can be established by using the same line of argument. This completes the proof of Proposition 8.

References

[1] Ager, P., and Brückner, M. (2013). "Cultural Diversity and Economic Growth: Evidence from the U.S. during the Age of Mass Migration." *European Economic Review*, vol. 64, pp.76-97.

[2] Alesina, A., and La Ferrara, E. (2005). "Ethnic Diversity and Economic Performance." *Journal of Economic Literature*, vol. 43, pp.762-800.

[3] Alesina, A., Harnoss, J., and Rapoport, H. (2016). "Birthplace Diversity and Economic Prosperity." *Journal of Economic Growth,* vol. 21, pp.101-138.

[4] Altonji, J. (1986). "Intertemporal Substitution in Labour Supply: Evidence from Micro Data," *Journal of Political Economy*, vol. 94, pp.s176-s215.

[5] Angyridis, C. (2015). "Endogenous Growth with Public Capital and Progressive Taxation." *Macroeconomic Dynamics*, vol. 19, pp.1220-1239.

[6] Becker, R.A. (1980). "On the Long-Run Steady State in a Simple Dynamic Model of Equilibrium with Heterogeneous Households," *Quarterly Journal of Economics*, vol. 95, pp.375-382.

[7] Carroll, D.R., and Young, E.R. (2009). "The Stationary Wealth Distribution under Progressive Taxation." *Review of Economic Dynamics*, vol. 12, pp.469-478.

[8] Carroll, D.R., and Young, E.R. (2011). "The Long Run Effects of Changes in Tax Progressivity." *Journal of Economic Dynamics & Control*, vol. 35, pp.1451-1473.

[9] Collier, P., and Gunning, J.W. (1999). "Explaining African Economic Performance." *Journal of Economic Literature*, vol. 37, pp.64-111.

[10] Conesa, J.C., and Krueger, D. (2006). "On the Optimal Progressivity of the Income Tax Code." *Journal of Monetary Economics*, vol. 53, pp.1425–1450.

[11] Correia, I. (2010). "Consumption Taxes and Redistribution," *American Economic Review*, vol. 100, pp.1673-1694.

[12] Datta, M., Mirman, L.J., and Reffett, K.L. (2002). "Existence and Uniqueness of Equilibrium in Distorted Dynamic Economies with Capital and Labor," *Journal of Economic Theory*, vol. 103, pp.377-410.

[13] Easterly, W., and Levine, R. (1997). "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics*, vol. 112, pp.1203-1250.

[14] Erosa, A., and Koreshkova, T. (2007). "Progressive Taxation in a Dynastic Model of Human Capital." *Journal of Monetary Economics,* vol. 54, pp.667-685.

[15] Frederick, S., Loewenstein, G., and O'Donoghue, T. (2002). "Time Discounting and Time Preference: A Critical Review." *Journal of Economic Literature,* vol. 40, pp.351-401.

[16] Gouveia, M., and Strauss, R.P. (1994). "Effective Federal Individual Tax Functions: An Exploratory Empirical Analysis." *National Tax Journal*, vol. 47, pp.317-339.

[17] Greenwood, J., Hercowitz, Z., and Huffman, G.W. (1988). "Investment, Capacity Utilization, and the Real Business Cycle," *American Economic Review*, vol. 78, pp.402-417.

[18] Guo, J.T., and Lansing, K.J. (1998). "Indeterminacy and Stabilization Policy." *Journal of Economic Theory*, vol. 82, pp.481-490.

[19] Huggett, M., Ventura, G., and Yaron, A. (2011). "Sources of Lifetime Inequality." *American Economic Review,* vol. 101, pp.2923-2954.

[20] Jaimovich, N., and Rebelo, S. (2009). "Can News about the Future Drive the Business Cycle?" *American Economic Review*, vol. 99, pp.1097-1118.

[21] Keane, M.P., and Wolpin, K.I. (1997). "The Career Decisions of Young Men." *Journal of Political Economy,* vol. 105, pp.473-522.

[22] Li, W., and Sarte, P.D. (2004). "Progressive Taxation and Long-Run Growth." *American Economic Review*, vol. 94, pp.1705-1716.

[23] MaCurdy, T.E. (1981). "An Empirical Model of Labor Supply in a Life Cycle Setting," *Journal of Political Economy*, vol. 89, pp.1059-1085.

[24] Nourry, C. (2001). "Stability of Equilibria in the Overlapping Generations Model with Endogenous Labor Supply," *Journal of Economic Dynamics & Control*, vol. 25, pp.1647-1663.

[25] Ottaviano, G., and Peri, G. (2006). "The Economic Value of Cultural Diversity: Evidence from U.S. Cities." *Journal of Economic Geography*, vol. 6, pp.9-44.

[26] Padovano, F., and Galli, E. (2001). "Tax Rates and Economic Growth in the OECD Countries (1950-1990)," *Economic Inquiry*, vol. 39, pp.44-57.

[27] Sarte, P.D. (1997). "Progressive Taxation and Income Inequality in Dynamic Competitive Equilibrium." *Journal of Public Economics,* vol. 66, pp.145-171.

[28] Shaked, M., and Shanthikumar, J.G. (2007). *Stochastic Orders.* Springer.

[29] Suen, R.M.H. (2014). "Time Preference and the Distributions of Wealth and Income," *Economic Inquiry*, vol. 52, pp.364-381.

[30] Trax, M., Brunow, S., and Suedekum, J. (2015). "Cultural Diversity and Plant-Level Productivity." *Regional Science and Urban Economics*, vol. 53, pp.85-96.