Diversity and Economic Performance in a Model with Progressive Taxation

Wei Wang[*]   Richard M. H. Suen[†]

This Version: 23rd July, 2018.

Abstract

Is a more heterogeneous population beneficial or harmful to long-term economic performance? This paper addresses this and other questions in a dynamic general equilibrium model where consumers have different labour productivity and time preference. We show how differences in the cross-sectional distribution of these characteristics can affect the economy via two channels. The first one involves changing the composition of the labour force; and the second one involves changing the cross-sectional distribution of marginal tax rate. We show how these channels are, respectively, determined by the shape of the labour supply function and the curvature 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, play in this matter? This paper addresses these questions using a dynamic general equilibrium model with heterogeneous consumers. In particular, the consumers are *ex ante* different in their labour productivity and time preference.[1] Our goal is to analyse how differences in the cross-sectional distribution of these characteristics affect long-term 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 has been very few theoretical research on this timely and important issue. This lack is somewhat surprising, given the prominence of heterogeneous-agent models in macroeconomics. The present study makes the first attempt to examine the issue of diversity using this type of model. Specifically, we adopt a similar deterministic neoclassical framework as in Sarte (1997), Li and Sarte (2004), Carroll and Young (2009, 2011) and Angyridis (2015). In these models, *ex ante* heterogeneity is the root cause 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 distribution of consumer types is typically taken as invariant in these previous studies. Thus, the effects of its changes are largely unexplored. The present study is intended to fill this gap.

The main points of this paper can be explained in terms of two types of effects, namely composition effects and general equilibrium effects. The former refers to changes in aggregate quantities due to changes in the composition of the underlying population, while the latter refers to changes

---

[1]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 has been reviewed in Frederick *et al.* (2002). We are agnostic about the source of consumer heterogeneity, 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]Implicitly, it is assumed that there is perfect consumption insurance so that individuals' choices are unaffected 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.

in individual-level quantities caused by the adjustment in equilibrium prices.[5] The exact nature of these effects depend on the type of heterogeneity considered. In the case of labour productivity heterogeneity, these effects primarily take place in the labour market. Specifically, any changes in the cross-sectional distribution of labour productivity will alter the composition of the labour force. This induces a shift in the aggregate labour supply function, thus triggering an adjustment in equilibrium wage rate (and interest rate), and in turn affects individuals' labour supply decision. Using general specifications of utility function, production function and progressive tax function, we derive conditions under which a more dispersed distribution of labour productivity will give rise to a higher level of aggregate labour supply and aggregate output in the steady state. Under these conditions, greater diversity will also benefit individual consumers by boosting their pre-tax income and consumption.

The case of time preference heterogeneity is more complicated due to the following reasons. Firstly, any changes in the cross-sectional distribution of time preference will not only affect the composition of the labour force, but also the cross-sectional distribution of pre-tax incomes; and these two changes often have conflicting effects on equilibrium prices. Secondly, the composition effect on aggregate labour supply is now more difficult to determine due to an income effect on individual labour supply. In light of these difficulties, the analysis of time preference heterogeneity is divided into three steps. First, in Section 4.1 we consider a simplified model in which labour supply is perfectly inelastic. This essentially shuts down the effect of time preference heterogeneity on aggregate labour supply. We then focus on how changing the distribution of time preference will affect the cross-sectional distribution of pre-tax income and marginal tax rates.[6] In the context of representative-agent models, the negative relation between marginal tax rate and capital accumulation is well understood: lowering the marginal tax rate can promote capital accumulation by raising the after-tax return on savings.[7] In this paper, we show that changing the distribution of time preference can have a similar effect on capital accumulation, even when there is no change in the tax function itself. The exact outcome of this is determined by the curvature of

---

[5] If we think of aggregate variables (e.g., aggregate consumption expenditure) as a weighted sum of the corresponding micro-level variables (e.g., household consumption expenditure), then the composition effect refers to changes in the weights while the general equilibrium effect concerns changes in the value of the micro-level variables.

[6] Since the tax function is assumed to be continuously differentiable, striclty increasing and strictly convex, there exists a one-to-one mapping between pre-tax income and marginal tax rate. Hence, the distribution of these two variables are isomorphic.

[7] 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.

the *marginal* tax function, which is an often overlooked feature of the tax function.[8] Specifically, if the marginal tax function is concave, then a mean-preserving but more dispersed distribution of time preference will lead to a lower average marginal tax rate and a higher level of capital accumulation. The opposite is true if the marginal tax function is convex. The intuition of these results can be explained as follows: Start with a homogeneous economy in which all consumers are *ex ante* identical and have the same pre-tax income. Suppose now a mean-preserving dispersion in time preference is introduced. This will create a non-degenerate distribution in pre-tax income and marginal tax rate. In particular, the relatively poor consumers in the 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.[9] Our main results in Section 4.1 generalise this comparison to two heterogeneous economies and provide the conditions under which greater diversity is beneficial or harmful to capital accumulation. Next, in Section 4.2 we resume the assumption of elastic labour supply but abstract away from the income effect. This is achieved by using the so-called "no-income-effect" utility function. In this case, the effect of time preference heterogeneity on aggregate labour supply can be easily characterised. In particular, if the marginal rate of substitution (MRS) between consumption and labour is a convex (or concave) function, then a mean-preserving spread in the distribution of time preference will lead to a downward (or upward) shift in the aggregate labour supply function. Finally, in Section 4.3 we provide some numerical examples to illustrate the composition effects and the general equilibrium effects in the full version of the model, where the income effect on labour supply is operative. Under some plausible parameter values, we find that a mean-preserving dispersion in time preference has only a mild positive effect on the capital-labour ratio (and hence the equilibrium prices), but a significant negative impact on aggregate labour supply. The latter is the result of a negative composition effect on the labour force.

The rest of the paper is organised as follows: Section 2 presents the baseline model. Section 3 analyses the effects of greater labour productivity heterogeneity. Section 4 focuses on the effects

---

[8]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 previously reported.

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

of time preference heterogeneity. Section 5 concludes.

2 The Baseline Model

2.1 Consumers

Time is discrete and is denoted by $t \in \{0, 1, 2, ...\}$. Consider an economy 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 is 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. The joint distribution of these characteristics across consumers is given by $H(\rho, \varepsilon)$, which is defined on 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 ρ and ε are, respectively, denoted by $H_1(\rho)$ and $H_2(\varepsilon)$.

In each time period, each consumer has one unit of time which can be divided between labour and leisure. Let $n_{i,t}$ and $l_{i,t}$ denote, respectively, the fraction of time spent on working and leisure activities 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). For every $n \in [0, 1]$, there exists $\underline{c}(n) \geq 0$ such that $U_c(c, n) \to \infty$ as $c \to \underline{c}(n)$.

Assumption A2 The marginal rate of substitution (MRS) 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.

The last part of Assumption A1 is similar in spirit to the Inada condition on the utility function. Specifically, $\underline{c}(n) \geq 0$ can be interpreted as a subsistence level of consumption (which may depend on n). When a consumer is close to subsistence, the marginal utility of consumption will become infinitely large. Thus, the consumer will never choose to consume at $\underline{c}(n)$. Assumption A2, on the other hand, ensures that consumption and leisure are both normal goods.[10] This assumption is equivalent to

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

for all (c,n). Both assumptions are satisfied by a large set of utility functions, including (i) the additively separable specification,

$$U(c,n) = \frac{c^{1-\sigma}-1}{1-\sigma} - A\frac{n^{1+\theta}}{1+\theta},$$

with $\sigma > 0$, $A > 0$ and $\theta > 0$; (ii) the "no-income-effect" utility function or GHH preferences, named after the work of Greenwood et al. (1988),

$$U(c,n) = \frac{\left(c - An^{1+\theta}\right)^{1-\sigma}-1}{1-\sigma}, \tag{3}$$

with $\sigma > 0$, $A > 0$ and $\theta > 0$; and (ii) the homothetic utility function,

$$U(c,n) = \frac{\left[c^\lambda (1-n)^{1-\lambda}\right]^{1-\sigma}-1}{1-\sigma}, \tag{4}$$

with $\sigma \geq 1$ and $\lambda \in (0,1)$.[11] For the additively separable and the homothetic specifications, the subsistence level of consumption in Assumption A1 is identical to zero, i.e., $\underline{c}(n) = 0$ for all n. For GHH preferences, we set $\underline{c}(n) \equiv An^{1+\theta}$ since the marginal utility of consumption $U_c(c,n) = \left(c - An^{1+\theta}\right)^{-\sigma}$ becomes infinite when c tends to $An^{1+\theta}$.

Next, we turn to consider the consumers' budget constraint. 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

---

[10] 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).

[11] By setting $\lambda = (1+\varsigma)^{-1}$ and $\sigma = 1 - \lambda^{-1}(1-\widetilde{\sigma})$, we can rewrite (4) as $U(c,n) = \lambda \left[c(1-n)^\varsigma\right]^{1-\widetilde{\sigma}} / (1-\widetilde{\sigma})$, with $\varsigma > 0$ and $\widetilde{\sigma} \geq 1$. This is essentially the utility function considered in King et al. (1988), except for a positive multiplicative constant. Thus, Assumptions A1 and A2 are also satisfied by this type of utility functions.

holdings at the beginning of time t. The consumer is in debt if this falls below zero. The interest income (or interest payment) associated with these assets is given by $r_t a_{i,t}$, 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 $\tau(\cdot)$.[12] The properties of the tax function are summarised below:

Assumption A3 The function $\tau : \mathbb{R}_+ \to \mathbb{R}_+$ is continuously differentiable and strictly increasing. 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. If $\tau(0) \leq 0$, then marginal rate progressivity 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 lump-sum transfer from the government. Conversely, a positive $\tau(0)$ can be viewed as a lump-sum tax appended to the progressive tax schedule. In this case, the average tax rate is non-monotonic in y.[13]

Consumer i's budget constraint at time t can be expressed as

$$c_{i,t} + a_{i,t+1} - a_{i,t} = y_{i,t} - \tau(y_{i,t}). \tag{5}$$

Taking prices and tax schedule as given, each consumer chooses 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 (5) and the initial amount of assets $a_0 > 0$.[14] 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 (5); the Euler equation for consumption

$$U_c(c_{i,t}, n_{i,t}) = \beta_i U_c(c_{i,t+1}, n_{i,t+1}) \left\{ 1 + \left[1 - \tau'(y_{i,t+1}) \right] r_{t+1} \right\}; \tag{6}$$

---

[12] This setup, which is commonly used in 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).

[13] The sign of $\tau(0)$ is immaterial for all of our theoretical results.

[14] 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.

the optimality condition for labour supply

$$\Psi\left(c_{i,t}, n_{i,t}\right) - w_t \varepsilon_i \left[1 - \tau'\left(y_{i,t}\right)\right] \begin{cases} \geq 0 & \text{if } n_{i,t} = 0, \\ = 0 & \text{if } n_{i,t} \in (0,1), \\ \leq 0 & \text{if } n_{i,t} = 1; \end{cases} \tag{7}$$

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 interest rate. The conditions in (7) take into account the possibility of corner solution for $n_{i,t}$. For instance, it is optimal to have $n_{i,t} = 0$ if the relative price of leisure, i.e., $w_t \varepsilon_i \left[1 - \tau'\left(y_{i,t}\right)\right]$, is less than or equal to the MRS at $n_{i,t} = 0$, i.e., $\Psi\left(c_{i,t}, 0\right)$.

2.2 Production and Government

On the supply side of the economy, there is a large number of identical firms. In each time 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, 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. Then the representative firm's problem is

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

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

The government collects taxes from the consumers and spends them entirely on "unproductive"

8

government purchases (G_t). This spending is called unproductive because it has no direct impact on the consumers' well-being and the production of goods. The government's budget constraint at any time t is given by

$$\int_0^1 \tau\left(y_{i,t}\right) di = G_t, \qquad \text{for all } t \geq 0. \tag{8}$$

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, $\{c_{i,t}, l_{i,t}, n_{i,t}, a_{i,t}\}_{t=0}^\infty$ solves consumer i's problem.

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

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

(iv) All markets clear in every time 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 stationary equilibria or steady states of this economy. These 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 pre-tax income, consumption, asset holdings and labour supply 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 completely determined by

$$\rho = r(k)\left\{1 - \tau'\left[y(k, \rho, \varepsilon)\right]\right\}, \tag{9}$$

$$c(k, \rho, \varepsilon) = y(k, \rho, \varepsilon) - \tau\left[y(k, \rho, \varepsilon)\right], \tag{10}$$

$$a(k, \rho, \varepsilon) = \frac{y(k, \rho, \varepsilon) - w(k)\varepsilon n(k, \rho, \varepsilon)}{r(k)}, \tag{11}$$

9

$$\Psi\left[c\left(k,\rho,\varepsilon\right),n\left(k,\rho,\varepsilon\right)\right] - \frac{w\left(k\right)}{r\left(k\right)}\varepsilon\rho \begin{cases} \geq 0 & \text{if } n\left(k,\rho,\varepsilon\right) = 0, \\ = 0 & \text{if } n\left(k,\rho,\varepsilon\right) \in \left(0,1\right), \\ \leq 0 & \text{if } n\left(k,\rho,\varepsilon\right) = 1. \end{cases} \tag{12}$$

Equation (9) is obtained by setting $U_c\left(c_{i,t},n_{i,t}\right) = U_c\left(c_{i,t+1},n_{i,t+1}\right)$ in the Euler equation of consumption.[15] The intuition behind this equation is as follows: In any stationary equilibrium, each consumer will maintain a constant level of marginal utility of consumption across time. This can be achieved if and only if the after-tax interest rate is equal to the consumer's rate of time preference. Such parity has two implications. Firstly, consumers with the same rate of time preference will face the same marginal tax rate and have the same level of pre-tax income, regardless of their labour productivity. In other words, $y\left(k,\rho,\varepsilon\right)$ is independent of ε. Secondly, for any ρ in $\left[\underline{\rho},\overline{\rho}\right]$, $y\left(k,\rho\right)$ is a strictly decreasing function in k. This is due to the following mechanism: Holding other things constant, an increase in k will lower the pre-tax interest rate and create an incentive for the consumer to substitute future consumption for current consumption. In order to maintain a constant marginal utility of consumption, it is necessary for the marginal tax rate to fall so as to maintain the equality in (9). Since $\tau'\left(\cdot\right)$ is strictly increasing, this leads to a lower level of pre-tax income for all ρ. In the sequel, we will refer to this as the *intertemporal smoothing effect*. Note that this effect arises only when the income tax schedule is nonlinear.[16] Equation (10) is obtained by setting $a_{i,t+1} = a_{i,t}$ in the sequential budget constraint. This, together with (9), implies that $c\left(k,\rho,\varepsilon\right)$ is also independent of ε. Equation (11) follows from the definition of pre-tax income. Equation (12) is obtained by substituting (9) into (7).

We now derive a single equation that can help determine the steady-state value of k. To start, define $\phi : \left[0,\overline{\tau}\right] \to \mathbb{R}_+ \cup \{+\infty\}$ as the inverse function of $\tau'\left(\cdot\right)$, i.e., $\phi\left[\tau'\left(y\right)\right] = y$ for all $y \geq 0$. Since the marginal tax function is continuous and strictly increasing, $\phi\left(\cdot\right)$ is a single-valued, continuous,

---

[15]Note that equation (9) is valid even if (i) there is *ex ante* heterogeneity in the utility function, i.e., $U^i\left(c,n\right) \neq U^j\left(c,n\right)$ for some $i \neq j$ in $\left[0,1\right]$, and (ii) there is no disutility from labour, i.e., $U\left(c,n_1\right) = U\left(c,n_2\right)$ for all $n_1 \neq n_2$ in $\left[0,1\right]$ and for all $c \geq 0$.

[16]If the income tax function is linear, i.e., $\tau'\left(y\right) = \widehat{\tau}$ for all $y \geq 0$, then the steady-state value of k is uniquely determined by $\left(1-\widehat{\tau}\right)r\left(k\right) = \underline{\rho}$. 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). By the same token, if we introduce another asset that offers a tax-free return, then only the most patient consumers will hold a strictly positive amount of this asset. Sarte (1997) shows that equation (9) plays a key role in obtaining a nondegenerate steady-state wealth distribution when consumers have different rates of time preference. The implications of the intertemporal smoothing effect, however, is less mentioned in existing studies.

strictly increasing function. Using (9) and the definition of pre-tax income, we can write

$$y(k, \rho) \equiv \phi \left[1 - \frac{\rho}{r(k)} \right] = w(k) \, \varepsilon n(k, \rho, \varepsilon) + r(k) \, a(k, \rho, \varepsilon). \tag{13}$$

Integrating both sides of (13) across all types of consumers yields

$$Y(k) \equiv \int_{\underline{\rho}}^{\overline{\rho}} \phi \left[1 - \frac{\rho}{r(k)} \right] dH_1(\rho) = [f(k) - \delta k] \, N(k), \tag{14}$$

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

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

and $f(k) \equiv F(k, 1)$ is the reduced-form production function. Equation (14) is essentially an accounting identity which states that the sum of all individuals' income equals national income (aggregate output minus depreciation of capital). We will refer to $Y(\cdot)$ as the *national income function*. A unique, non-trivial steady state exists if equation (14) has a single, strictly positive solution. The rest of this section is devoted to establishing the existence of such a solution.

The first step is to specify the range of plausible values of k. Since $\phi(\cdot)$ is only defined on $[0, \overline{\tau}]$, equations (13) and (14) are satisfied only if

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

for all $\rho \in \left[\underline{\rho}, \overline{\rho} \right]$. In other words, any k that solves (14) must satisfy

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

To ensure that this range is nonempty, it is necessary to have $\underline{\rho} > (1 - \overline{\tau}) \overline{\rho}$. By the strict concavity of $f(\cdot)$ and the Inada conditions on the production function, there exists a unique pair of values $k_{\max} > k_{\min} > 0$ such that

$$r(k_{\max}) \equiv f'(k_{\max}) - \delta = \overline{\rho} \quad \text{and} \quad r(k_{\min}) = \frac{\underline{\rho}}{1 - \overline{\tau}}. \tag{15}$$

Thus, any solution of equation (14) must be contained within the interval $[k_{\min}, k_{\max}]$.
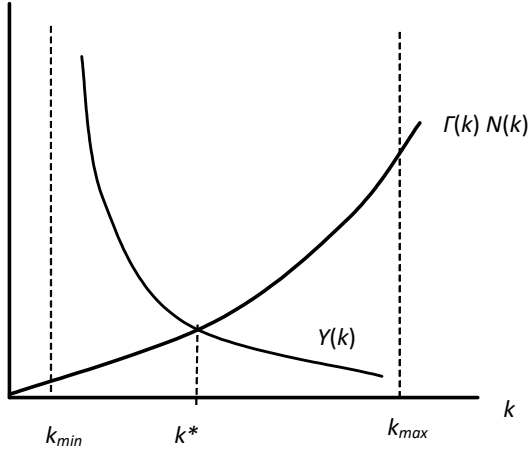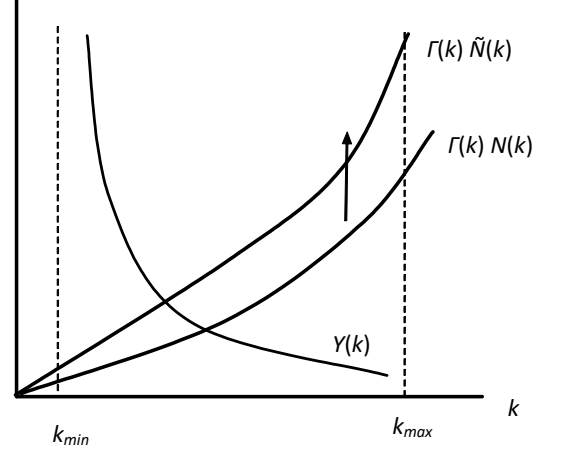
Figure 1a

Figure 1b

Lemma 1 provides a set of necessary and sufficient conditions under which a unique non-trivial steady state exists in the baseline model. All proofs can be found in the Appendix. A graphical illustration of the unique steady state is shown in Figure 1a.[17]

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*

$$N\left(k_{\max}\right)\left[f\left(k_{\max}\right) - \delta k_{\max}\right] > \int_{\underline{\rho}}^{\overline{\rho}} \phi \left(1 - \frac{\rho}{\overline{\rho}}\right) dH_1\left(\rho\right), \qquad (16)$$

and

$$N\left(k_{\min}\right)\left[f\left(k_{\min}\right) - \delta k_{\min}\right] < \int_{\underline{\rho}}^{\overline{\rho}} \phi \left[1 - \frac{\rho}{\underline{\rho}}\left(1 - \overline{\tau}\right)\right] dH_1\left(\rho\right). \qquad (17)$$

3   Heterogeneity in Labour Productivity

We now turn to the main subject of this paper, which is the economic consequences of greater consumer heterogeneity. In the current section, we focus on the effects of labour productivity heterogeneity. The effects of time preference heterogeneity will be examined in Section 4. In both sections, we assume that ε and ρ are statistically independent in the population so that $H\left(\rho, \varepsilon\right) = H_1\left(\rho\right) H_2\left(\varepsilon\right)$ for all $\left(\rho, \varepsilon\right)$. We then compare two economies that have the same fundamentals except for one of the marginal distributions.

---

[17]The inequalities in (16) and (17) are technical conditions which ensure that the two curves in Figure 1a cross at least once within $(k_{\min}, k_{\max})$. The function $\Gamma\left(\cdot\right)$ in Figures 1a and 1b is defined as $\Gamma\left(k\right) \equiv f\left(k\right) - \delta k$.

Two criteria will be used to compare different marginal distributions. The first one is the *Lorenz dominance criterion* (also known as Lorenz order or convex order), which is commonly used in studies of risk and inequality. Let $Q(\cdot)$ and $\widetilde{Q}(\cdot)$ be two distribution functions defined on the same support in $\mathbb{R}_+$ and have the same mean. Then $\widetilde{Q}(\cdot)$ is said to be "more unequal" than $Q(\cdot)$ under the Lorenz dominance criterion if

$$\int_0^x \widetilde{Q}(z)\,dz \leq \int_0^x Q(z)\,dz, \qquad \text{for all } x \geq 0. \tag{18}$$

The "more unequal" distribution $\widetilde{Q}(\cdot)$ is also called a mean-preserving spread of $Q(\cdot)$. It is well-known that (18) is satisfied if and only if

$$\int_0^\infty \xi(z)\,d\widetilde{Q}(z) \geq \int_0^\infty \xi(z)\,dQ(z)$$

for any convex function $\xi: \mathbb{R}_+ \to \mathbb{R}$, provided that the integrals exist. Intuitively, a "more unequal" distribution of consumer characteristic is one that exhibits greater cross-sectional variations, and thus represents a larger extent of consumer heterogeneity.

Another criterion that we use is the *starshaped order*. Recall that a function $\xi: \mathbb{R}_+ \to \mathbb{R}$ is starshaped if $\xi(0) \leq 0$ and $\xi(z)/z$ is non-decreasing in z. Then $\widetilde{Q}(\cdot)$ is said to be "more unequal" than $Q(\cdot)$ according to the starshaped order if

$$\int_0^\infty \xi(z)\,d\widetilde{Q}(z) \geq \int_0^\infty \xi(z)\,dQ(z) \tag{19}$$

for all bounded, continuous and starshaped function ξ.[18] The condition in (19) can be equivalently stated as[19]

$$\int_x^\infty z\,d\widetilde{Q}(z) \geq \int_x^\infty z\,dQ(z), \qquad \text{for all } x \geq 0. \tag{20}$$

If $Q(\cdot)$ and $\widetilde{Q}(\cdot)$ have the same mean (or aggregate), i.e., $\int_0^\infty z\,d\widetilde{Q}(z) = \int_0^\infty z\,dQ(z) \equiv S$, then (20) can be equivalently stated as

$$\frac{\int_x^\infty z\,d\widetilde{Q}(z)}{\int_0^\infty z\,d\widetilde{Q}(z)} \geq \frac{\int_x^\infty z\,dQ(z)}{\int_0^\infty z\,dQ(z)}, \qquad \text{for all } x \geq 0.$$

The expression on the right side of this inequality gives the fraction of the aggregate S that is concentrated in the interval $[x, \infty)$ under $Q(\cdot)$. The expression on the left can be similarly

---

[18] For more information on this type of ordering, see Shaked and Shanthikumar (2007, Section 4.A.6).

[19] For a formal proof of this statement, see Lemma A1 in the Appendix

interpreted. Thus, a mean-preserving but "more unequal" distribution under the starshaped order is one that is more concentrated on the top end of the support. The relation between these two types of order is as follows: If $Q(\cdot)$ and $\widetilde{Q}(\cdot)$ have the same mean and satisfy (20), then $\widetilde{Q}(\cdot)$ is also "more unequal" than $Q(\cdot)$ under the Lorenz order. But the converse of this is not true in general. Thus, the starshaped order is stronger than the Lorenz order. The rationale for using this stronger order will be explained below.

Consider two economies that have the same size of population, utility function $u(\cdot)$, production technology $F(\cdot)$, progressive tax schedule $\tau(\cdot)$ and distribution of time preference $H_1(\cdot)$ defined on $[\underline{\rho}, \overline{\rho}]$.[20] The only difference between them lies in the marginal distribution of labour productivity, which are denoted by $H_2(\varepsilon)$ and $\widetilde{H}_2(\varepsilon)$. Both of them are defined on $[\underline{\varepsilon}, \overline{\varepsilon}]$ and satisfy the assumption stated below. The second part of Assumption A5 ensures that a unique non-trivial steady state exists in both economies.

Assumption A5   (i) The average value of ε is identical under $H_2(\cdot)$ and $\widetilde{H}_2(\cdot)$. (ii) Conditions (16) and (17) are satisfied in both economies.

Notice that when k is held constant, changing the distribution of ε will have no effect on the individual-level variables defined by (9)-(12). Instead, changing this distribution will only affect the composition of aggregate labour supply. In other words, it will affect the solution of (14) but only through the function $N(\cdot)$.

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

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

Similarly, define $\widetilde{N}(\cdot)$ using $\widetilde{H}_2(\cdot)$. From Figure 1b, it is evident that if $N(k) \leq \widetilde{N}(k)$ 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 sufficient condition under which $N(k) \leq \widetilde{N}(k)$ for all $k \in (k_{\min}, k_{\max})$. This proposition is built upon the following intermediate result.

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.*

---

[20] 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 (15).

The intuition behind this result is simple: more productive workers have a higher opportunity cost of leisure, hence they choose to work more than less productive workers. This result holds whenever (i) labour income and interest income are taxed jointly, so that the marginal tax rate on these incomes are always the same, and (ii) the MRS between consumption and labour is strictly increasing in labour. Both assumptions are commonplace in existing studies.

Lemma 2 also implies that a one-percent increase in ε can potentially lead to a greater percentage increase in effective unit of labour, i.e., $\varepsilon n (k, \rho, \varepsilon)$. To see this formally, 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 Lemma 2 implies $\varepsilon_2 n (k, \rho, \varepsilon_2) > (1 + \Theta) \varepsilon_1 n (k, \rho, \varepsilon_1)$. Intuitively, this means an endogenous labour supply has the effect of amplifying the variations in labour productivity across consumers.

We now present a sufficient condition under which $N (k) \leq \widetilde{N} (k)$ is true for all $k \in (k_{\min}, k_{\max})$.

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*

$$\int_x^{\overline{\varepsilon}} \varepsilon dH_2 (\varepsilon) \leq \int_x^{\overline{\varepsilon}} \varepsilon d\widetilde{H}_2 (\varepsilon), \quad for\ all\ x \in [\underline{\varepsilon}, \overline{\varepsilon}]. \tag{21}$$

Proposition 3 is a direct application of the starshaped order mentioned earlier. To see this, first rewrite $N (k)$ and $\widetilde{N} (k)$ as

$$N (k) \equiv \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon \mathcal{N} (k, \varepsilon) dH_2 (\varepsilon) \quad \text{and} \quad \widetilde{N} (k) \equiv \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon \mathcal{N} (k, \varepsilon) d\widetilde{H}_2 (\varepsilon), \tag{22}$$

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

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

By Lemma 2, $\varepsilon \mathcal{N} (k, \varepsilon)$ is a bounded, continuous, starshaped function in ε for all $k \in (k_{\min}, k_{\max})$. Thus, we can interpret $N (k) \leq \widetilde{N} (k)$ as comparing the expected value of a starshaped function under two different distributions, and a sufficient condition for this is (21). 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.[21] For this reason, a stronger criterion (namely the starshaped order) is used in this comparison.

---

[21] The details of this point are available from the authors upon request.

We now consider the steady-state effects of an increase in labour productivity heterogeneity. Let k^* and $\widetilde{k}^*$ be the unique solution of (14) under $H_2(\cdot)$ and $\widetilde{H}_2(\cdot)$, respectively. Suppose Assumption A5 and (21) are satisfied so that $\widetilde{H}_2(\cdot)$ is a mean-preserving but more heterogeneous distribution than $H_2(\cdot)$ under the starshaped order. As explained earlier, this means $\widetilde{H}_2(\cdot)$ has a higher concentration at the top end of the labour productivity spectrum than $H_2(\cdot)$. By Proposition 3, the more heterogeneous economy will have a greater aggregate labour supply under any $k \in (k_{\min}, k_{\max})$. This leads to a lower steady-state value of capital-labour ratio in the more heterogeneous economy, i.e., $k^* \geq \widetilde{k}^*$ (see Figure 1b). By the intertemporal smoothing effect described in Section 2.3, a lower capital-labour ratio is associated with a higher pre-tax income and consumption for each consumer. Thus, according to our baseline model, greater heterogeneity in labour productivity is beneficial to all consumers. At the aggregate level, a more heterogeneous workforce is associated with a higher level of aggregate labour input and national income. These results are summarised in Proposition 4.[22]

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 (21). 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 $\rho \in \left[\underline{\rho}, \overline{\rho}\right]$.

4 Heterogeneity in Time Preference

Comparing to the previous section, the analysis of greater time preference heterogeneity is more challenging due to two reasons: Firstly, changing the distribution of ρ will not only shift the aggregate labour supply function $N(\cdot)$ on the right side of equation (14), but also the national income function $Y(\cdot)$ on the left. Because of this simultaneous movement, the overall results are often qualitatively ambiguous. Secondly, it is difficult to determine how $n(k, \rho, \varepsilon)$ changes with ρ in the presence of income effect on labour supply.[23] Without knowing this, we cannot ascertain qualitatively the effect of greater time preference heterogeneity on $N(\cdot)$.

Because of these complexities, theoretical results are available only under two additional conditions. In Section 4.1, we assume that individual labour supply is an exogenous constant. As a

---

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

[23]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 MRS between consumption and labour through the variable $c(k, \rho)$. The latter is what we refer to as the income effect.

result, aggregate labour input is independent of the distribution of ρ. This abstraction allows us to focus on the effects of time preference heterogeneity on $Y(\cdot)$ alone. As we will see below, these effects are entirely determined by the shape of the marginal tax function $\tau'(\cdot)$. This subsection thus highlights the role of progressive taxation in determining the impact of greater time preference heterogeneity. In Section 4.2, we resume the assumption of flexible labour supply but abstract away from the aforementioned income effect. This is achieved by using the "no-income-effect" utility function. In this case, the effects of greater time preference heterogeneity are jointly determined by the shape of the marginal tax function and the shape of the MRS between consumption and labour. Finally, in Section 4.3 we use numerical examples to illustrate the effects of time preference heterogeneity in the full version of the baseline model where the income effect is operative.

4.1 Exogenous Labour Model

In this subsection, the consumer's utility function is given by $U(c,n) \equiv u(c)$ for all $c \geq 0$ and $n \in [0,1]$, where $u : \mathbb{R}_+ \to \mathbb{R}$ is twice continuously differentiable, strictly increasing, strictly concave and satisfies $\lim_{c \to 0} u'(c) = \infty$.[24] Let $\widehat{\varepsilon} > 0$ be the average labour productivity in the population, i.e., $\widehat{\varepsilon} \equiv \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon dH_2(\varepsilon)$. Individual and aggregate labour supply are then given by $n_{i,t} = 1$ for all i and $N_t = \widehat{\varepsilon}$, respectively. 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 (9) and (10), but the labour supply conditions in (12) will be simplified to become $n(k,\rho,\varepsilon) = 1$, for all (k,ρ,ε). Equation (14) is now given by

$$\int_{\underline{\rho}}^{\overline{\rho}} \phi \left[1 - \frac{\rho}{r(k)} \right] dH_1(\rho) = [f(k) - \delta k]\widehat{\varepsilon}. \tag{23}$$

Note that any solution of (23) will only depend on the mean value of ε but not 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 (23) exists if and only if (16) and (17) are satisfied [with $N(k_{\max})$ and $N(k_{\min})$ replaced by $\widehat{\varepsilon}$].

We now compare two economies that are otherwise identical except for the distribution of ρ, denoted by $H_1(\cdot)$ and $\widetilde{H}_1(\cdot)$. Both are defined on $[\underline{\rho},\overline{\rho}]$ and satisfy Assumption A6. The first part of this assumption states that $\widetilde{H}_1(\cdot)$ is more heterogeneous than $H_1(\cdot)$ under the Lorenz dominance criterion.

---

[24]These assumptions will replace Assumptions A1 and A2 in the baseline model.

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.

Let $Y(\cdot)$ be the national income function defined using $H_1(\cdot)$ and k^* be the corresponding unique solution of (23). Their counterparts under $\widetilde{H}_1(\cdot)$ are denoted by $\widetilde{Y}(\cdot)$ and $\widetilde{k}^*$. A more heterogeneous population is said to be beneficial (or harmful) to long-term capital accumulation if $\widetilde{k}^* \geq k^*$ (or $\widetilde{k}^* \leq k^*$).[25] Proposition 5 states the conditions under which this is true in this model.

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 to compare an identical-agent (IA) economy, where all consumers have the same rate of time preference, to a heterogeneous-agent (HA) economy, where consumers have different rates of 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). To see the intuition behind these results, it is instructive to compare the distribution of marginal tax rates in these two economies.

Suppose for the moment that $H_1(\cdot)$ is a degenerate distribution at some point $\widehat{\rho}$ in $\left[\underline{\rho}, \overline{\rho}\right]$ and $\widetilde{H}_1(\cdot)$ is non-degenerate with mean $\widehat{\rho}$. In the IA economy, all consumers have the same pre-tax income $y(k^*, \widehat{\rho})$ and face the same marginal tax rate $\tau'[y(k^*, \widehat{\rho})]$. Introducing a mean-preserving spread in time preference will create a dispersion in 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.[26] If the marginal tax function is concave, then the average marginal tax rate will be lowered as a result. More specifically, if $\tau'(\cdot)$ is concave, then

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

---

[25] 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).

[26] This follows from the fact that $y(k, \rho)$ is strictly decreasing in ρ for all $k \in (k_{\min}, k_{\max})$. This property can be easily shown using the first part of (13).

for all $x \in \left[\underline{\rho}, \overline{\rho}\right]$. The expression on the right 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. Alternatively, if $\tau'(\cdot)$ is convex, then we have

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

for all $x \in \left[\underline{\rho}, \overline{\rho}\right]$. In this case, consumers in the HA economy face a higher marginal tax rate in general, which has a harmful impact on 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

$$\int_{\sigma(q)}^{\overline{\rho}} \tau'\left[y\left(k^*, \rho\right)\right] dH_1(\rho) \geq \int_{\widetilde{\sigma}(q)}^{\overline{\rho}} \tau'\left[y\left(\widetilde{k}^*, \rho\right)\right] d\widetilde{H}_1(\rho), \quad \text{for all } q \in [0, 1].$$

(ii) If the marginal tax function is convex, then

$$\int_{\underline{\rho}}^{\sigma(q)} \tau'\left[y\left(k^*, \rho\right)\right] dH_1(\rho) \leq \int_{\underline{\rho}}^{\widetilde{\sigma}(q)} \tau'\left[y\left(\widetilde{k}^*, \rho\right)\right] d\widetilde{H}_1(\rho), \quad \text{for all } q \in [0, 1].$$

We conclude this subsection by pointing out the relevance of concave marginal tax function in the existing literature. Two parametric forms of $\tau(\cdot)$ are typically 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 $\tau(y) = \zeta y^{1+\chi}$, where ζ and χ are two strictly positive parameters. It is straightforward to show that the corresponding 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),

$$\tau(y) = a_0 \left[y - \left(y^{-a_1} + a_2\right)^{-\frac{1}{a_1}}\right]. \tag{24}$$

This functional form has been used by Sarte (1997), Conesa and Krueger (2006), Erosa and Koreshkova (2007), and Carroll and Young (2011), among others. The second and third-order deriva-

tives of this function are given by

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

$$\tau'''\left(y\right) = \frac{\tau''\left(y\right)}{y}\left[a_1 - 1 - \left(2a_1 + 1\right)\left(\frac{a_2 y^{a_1}}{1 + a_2 y^{a_1}}\right)\right]. \tag{25}$$

In all existing applications, the parameters a_0, a_1 and a_2 are taken to be strictly positive which ensure that $\tau''\left(\cdot\right) > 0$. Gouveia and Strauss (1994) report estimates of a_1 ranging from 0.726 to 0.938 based on U.S. data. From (25), it is obvious that these values imply $\tau'''\left(\cdot\right) < 0$, i.e., a strictly concave marginal tax function.

4.2 Endogenous Labour Without Income Effect

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

In any stationary equilibrium, equations (9)-(11) will remain valid and the optimality condition for labour supply will be given by

$$v'\left[n\left(k, \rho, \varepsilon\right)\right] - \frac{w\left(k\right)}{r\left(k\right)}\varepsilon\rho \begin{cases} \geq 0 & \text{if } n\left(k, \rho, \varepsilon\right) = 0, \\ = 0 & \text{if } n\left(k, \rho, \varepsilon\right) \in \left(0, 1\right), \\ \leq 0 & \text{if } n\left(k, \rho, \varepsilon\right) = 1. \end{cases} \tag{26}$$

The two corner solutions can be ruled out by introducing some additional assumptions. The details are shown in Lemma 7.

Lemma 7 *Suppose Assumption A4 is satisfied. Then the following results hold for all $k \in \left(k_{\min}, k_{\max}\right)$ and for all $\left(\rho, \varepsilon\right) \in \left[\underline{\rho}, \overline{\rho}\right] \times \left[\underline{\varepsilon}, \overline{\varepsilon}\right].$*

(i) If $\lim_{n \to 0} v'\left(n\right) = 0$, then $n\left(k, \rho, \varepsilon\right) > 0$.

(ii) If $v'\left(1\right) > w\left(k_{\max}\right)\overline{\varepsilon}$, then $n\left(k, \rho, \varepsilon\right) < 1$.

The condition $\lim_{n \to 0} v'\left(n\right) = 0$ means that the marginal cost of labour is negligible when n is close to zero. But the marginal benefit of working is always strictly positive when $n > 0$, hence all consumers will choose to have $n > 0$. On the other hand, if $v'\left(1\right) > w\left(k_{\max}\right)\overline{\varepsilon}$ holds, then the

20

marginal cost of working at $n = 1$ will outweigh its marginal benefit under all possible steady-state wage rate and for all types of consumers. Thus, no one will find it optimal to choose $n = 1$.

Similar to the previous subsection, let $H_1(\cdot)$ and $\widetilde{H}_1(\cdot)$ be two distinct distributions of ρ that satisfy Assumption A6. When labour supply is flexible, changes in time preference heterogeneity will affect both the national income function $Y(\cdot)$ and the aggregate labour supply function $N(\cdot)$. The effects on $Y(\cdot)$ are the same as in Proposition 5.[27] The effects on $N(\cdot)$ are examined below.[28]

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

(i) If $v'(\cdot)$ is concave and satisfies $v'(1) > w(k_{\max})\,\overline{\varepsilon}$, then $n(k, \rho, \varepsilon)$ is convex in ρ and $N(k) \leq \widetilde{N}(k)$.

(ii) If $v'(\cdot)$ is convex and satisfies $\lim_{n \to 0} v'(n)$, then $n(k, \rho, \varepsilon)$ is concave in ρ and $N(k) \geq \widetilde{N}(k)$.

To explain these results, first consider the case when n is an interior solution. Such a solution is completely characterised by the first-order condition $v'(n) = \varpi$, where ϖ denotes the after-tax wage rate. According to (26), ϖ is determined by the steady-state capital-labour ratio and the consumer's own characteristics. For now we will ignore these details and express the individual labour supply function simply as $n(\varpi)$. An increasing $v'(\cdot)$ means that the marginal cost of labour is increasing. Thus, a consumer will choose to work longer hours if and only if he is compensated by a higher wage rate, i.e., $n(\varpi_2) \geq n(\varpi_1)$ iff $\varpi_2 \geq \varpi_1$. A concave $v'(\cdot)$ means that the marginal cost of labour is increasing in n but at a declining rate. Thus, when presented the same (absolute) increase in real wage, a high-wage earner will increase his labour supply more than a low-wage earner. Formally, this means for any $\Delta > 0$,

$$n(\varpi_2 + \Delta) - n(\varpi_2) \geq n(\varpi_1 + \Delta) - n(\varpi_1), \quad \text{whenever } \varpi_2 \geq \varpi_1. \tag{27}$$

Equation (27) is equivalent to saying that individual labour supply is a convex function in ϖ. Conversely, if $v'(\cdot)$ is convex, then the marginal cost of labour is increasing in n at an increasing

---

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

[28] For the specific funcitonal form in (3), we can write $v'(n) = A(1 + \theta) n^\theta$. This function is strictly concave (or strictly convex) if and only if $\theta < 1$ or $(\theta > 1)$. It also satisfies the condition $\lim_{n \to 0} v'(n) = 0$ whenever $\theta > 0$. Hence, the conditions in the second part of Proposition 8 are satisfied if $\theta > 1$. If, in addition, we use a Cobb-Douglas production function so that $f(k) = k^\alpha$ for some $\alpha \in (0, 1)$, then the conditions in the first part of Proposition 8 are satisfied if $\theta < 1$ and $A(1 + \theta) > (1 - \alpha)\left(\frac{\alpha}{\delta + \overline{\rho}}\right)^{\frac{\alpha}{1 - \alpha}} \overline{\varepsilon}$. 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.

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	$\uparrow$

rate. In this case, a high-wage earner will be more reluctant to increase his labour supply as ϖ increases. The inequality in (27) is now reversed which means $n(\varpi)$ is a concave function. When comparing across consumers with different rate of time preference, it suffice to note that the after-tax wage rate in (26) is linearly increasing in ρ. Thus, an increasing concave $v'(\cdot)$ will imply that $n(k,\rho,\varepsilon)$ is increasing and convex in ρ. A mean-preserving spread in ρ then leads to an increase in the average value of $\varepsilon n(k,\rho,\varepsilon)$ across all types of consumers, i.e., $N(k) \leq \widetilde{N}(k)$ for all plausible value of k.

The above arguments can be (partially) extended to allow for corner solutions in n. Let $\widehat{n}(\varpi)$ be the solution of the unconstrained problem, i.e., $v'[\widehat{n}(\varpi)] = \varpi$ for all ϖ. If $\widehat{n}(\varpi)$ is convex, then the composite function $\max\{\widehat{n}(\varpi),0\}$ is also convex but $\min\{\widehat{n}(\varpi),1\}$ is not. Thus, the first part of Proposition 8 is valid so long as the optimal labour supply is strictly less than one. This can be ensured by imposing the condition $v'(1) > w(k_{\max})\overline{\varepsilon}$. Likewise, if $\widehat{n}(\varpi)$ is concave, then $\min\{\widehat{n}(\varpi),1\}$ is also a concave function but $\max\{\widehat{n}(\varpi),0\}$ is not. Thus, we have included the condition $\lim_{n\to 0} v'(n) = 0$ in the second part of Proposition 8 to ensure that $n > 0$.

Based on the shape of $\tau'(\cdot)$ and $v'(\cdot)$, we can identify four possible scenarios. Table 1 summarises the overall effects of greater time preference heterogeneity in each of these cases. These 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 the result of two opposing forces: on one hand, an increase in time preference heterogeneity will lower the average marginal tax rate on asset return which encourages capital accumulation; on the other hand, such an increase will lead to an expansion in aggregate labour supply and suppress the capital-labour ratio. Which effect dominates is a

quantitative question. The other three cases in Table 1 can be interpreted in a similar fashion.

4.3 Numerical Examples

In the previous two sections, we have identified two channels through which greater time preference heterogeneity can affect the economy. The first one involves changing the cross-sectional distribution of marginal tax rates and the national income function, while the second one involves a composition effect on aggregate labour supply. In this section, we will use numerical examples to demonstrate these effects in the full version of the baseline model. There are two reasons why we resort to quantitative analysis here. Firstly, the presence of income effect on labour supply poses a serious challenge in characterising the shape of $n(k, \rho, \varepsilon)$ as a function of ρ. As a result, we cannot ascertain qualitatively the effects of greater time preference heterogeneity on $N(\cdot)$ as in Proposition 8. Secondly, as Table 1 suggests, the overall effects of greater time preference heterogeneity are often qualitatively ambiguous. The numerical examples presented below are intended to throw 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 working in the steady state. The resulting value of A is 47.28. The Frisch elasticity of labour supply is set to 0.40, based on the estimates by MaCurdy (1981) and Altonji (1986). The production function is assumed to take the Cobb-Douglas form, i.e., $F(K, N) = K^\alpha N^{1-\alpha}$, with $\alpha = 0.40$. We choose the value of δ so that the steady-state capital-output ratio matches the value observed in the United States over the period 1947-2016, which is 2.367.[29] The required value of δ is 5.3%. The progressive tax function is assumed to take the form in (24), with $a_0 = 0.258$ and $a_1 = 0.768$ as reported by Gouveia and Strauss (1994).[30] The value of a_2 is determined in two steps: First, we assume that government spending G accounts for 20.7% of aggregate output $F(K, N)$ in the steady state. This value is based on the share of government consumption expenditures in US GDP over the period

---

[29] We use the sum of private fixed assets and end-of-year stock of private inventories as our measure of aggregate capital stock. Data on private fixed assets and private inventories are obtained from the National Income and Product Accounts (NIPA).

[30] The same value of a_0 and a_1 are also used by Conesa and Krueger (2006) and Carroll and Young (2011).

Table 2 Targeted Statistics

Ratio of capital to output	2.367
Ratio of gov't spending to output	0.207
Ratio of national income to output	0.869
Average labour hours	0.333

Percentile of Income Distribution	Mean income (relative to lowest quintile)
$0 - 19.9$	1.0
$20 - 39.9$	2.0
$40 - 59.9$	3.2
$60 - 79.9$	5.5
$80 - 89.9$	8.2
$90 - 100$	28.5

1947-2016. We then solve the government's budget constraint in (8) for a_2.

In terms of consumer characteristics, we assume that labour productivity is uniformly distributed across consumers with $\underline{\varepsilon} = 1$ and $\bar{\varepsilon} = 100$. The distribution of ρ, on the other hand, is calibrated so that the distribution of pre-tax income in the benchmark numerical example matches certain features of its real-world counterpart.[31] Specifically, we divide the entire population into M income groups. All consumers within group j have the same level of pre-tax income $\bar{y}_j$, for $j \in \{1, 2, ..., M\}$. These incomes are ranked according to $\bar{y}_1 < ... < \bar{y}_M$. The population share of income group j is denoted by $\pi_j \in (0, 1)$, with $\sum_{j=1}^M \pi_j = 1$. Since there is an one-to-one mapping between pre-tax income and rate of time preference, all members within the same income group have the same ρ which can be determined by $\bar{y}_j = y(k^*, \rho_j)$, for all j. To construct a realistic income distribution, we first compute the relative income $\varpi_j \equiv \bar{y}_j / \bar{y}_1$ for six income groups based on the 2016 Survey of Consumer Finance (SCF) data reported in Bricker *et al.* (2017) Table 1.[32] These income groups include the bottom four quintiles, the 80-89.9 percentile and the top 10% of

---

[31] Since the pre-tax income function $y(k, \rho)$ is independent of ε, the cross-sectional distribution of labour productivity is irrelevant here.

[32] Specifically, we use the mean income data in 2016 to compute the relative incomes $\{\varpi_j\}_{j=1}^M$. We have also computed these values using data from five earlier rounds of SCF (2001, 2004, 2007, 2010 and 2013) as shown in Bricker *et al.* (2012) and Bricker *et al.* (2017). There are very little variations in these values, except for the top income group (ϖ_6). The numerical results are largely the same when we calibrate the model using these earlier data.

Table 3 Benchmark Parameter Values

A	Preference parameter	47.28
ψ	Frisch elasticity of labour supply	0.400
α	Share of capital income in aggregate output	0.400
δ	Depreciation rate of capital	0.053
a_0	Parameter in the progressive tax function	0.258
a_1	Parameter in the progressive tax function	0.768
a_2	Parameter in the progressive tax function	0.402
$\underline{\varepsilon}$	Minimum value of labour productivity	1
$\overline{\varepsilon}$	Maximum value of labour productivity	100
$\{\pi_j\}$	Population share of income groups	$[0.20, 0.20, 0.20,$ $0.20, 0.10, 0.10]$
$\{\rho_j\}$	Rate of time preference	$[0.0883, 0.0861, 0.0853,$ $0.0848, 0.0846, 0.0844]$

Table 4 Results of Numerical Examples

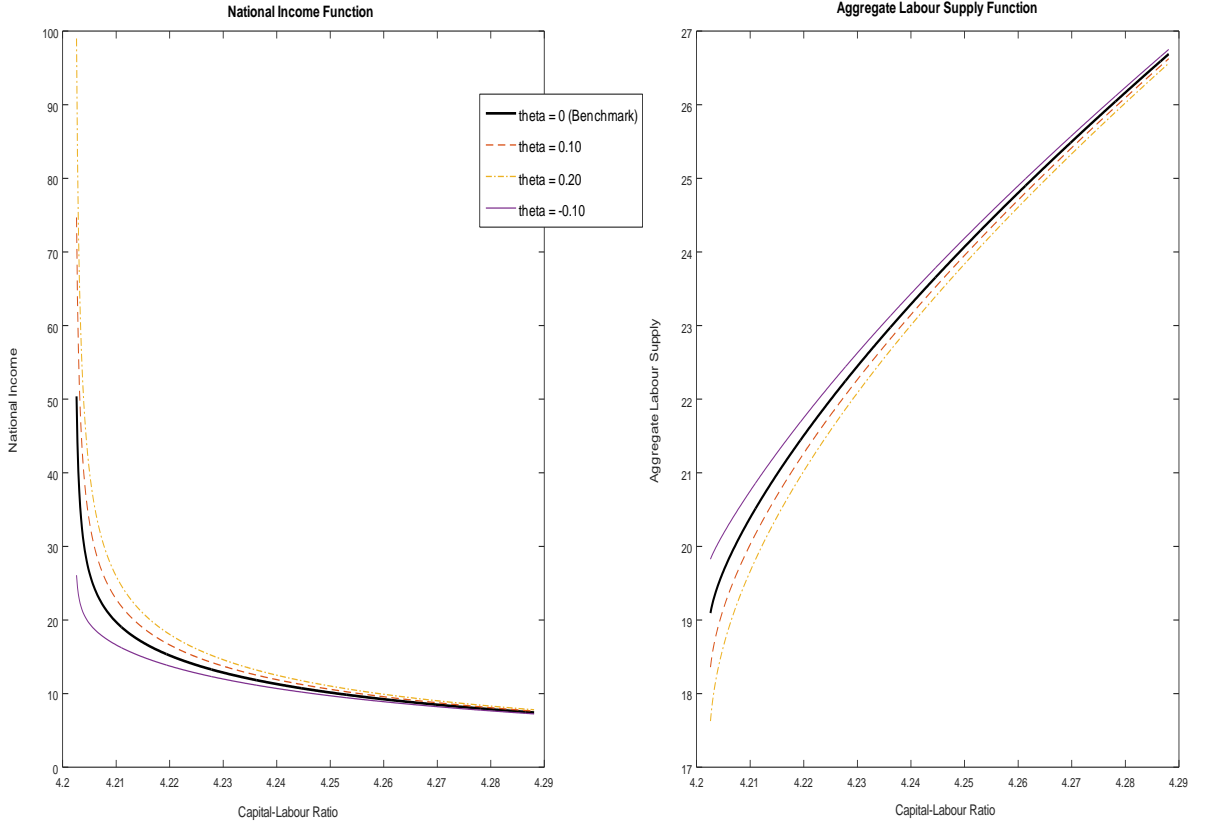|	Benchmark	% Changes from Benchmark			
	$\theta = 0$	$\theta = 0.05$	$\theta = 0.10$	$\theta = 0.20$	$\theta = -0.10$
k^*	4.204	0.02%	0.05%	0.09%	-0.04%
$N(k^*)$	19.475	-0.37%	-0.64%	-1.00%	1.57%
$Y(k^*)$	30.058	-0.37%	-0.62%	-0.97%	1.55%
$k^*N(k^*)$	81.873	-0.35%	-0.59%	-0.91%	1.53%
$(k^*)^{\alpha} \cdot N(k^*)$	34.589	-0.36%	-0.62%	-0.97%	1.55%
G^*	7.160	-0.42%	-0.72%	-1.12%	1.76%

Figure 2

the income distribution.[33] We then compute the value of $\overline{y}_1$ so that the ratio of national income $Y(k)$ to aggregate output $F(K, N)$ is 0.869 in the steady state. This matches the value observed in US data over the period 1947-2016.[34] Once $\overline{y}_1$ is known, we can solve for $\left\{\rho_j\right\}_{j=1}^{M}$ using $\overline{y}_j = \varpi_j \overline{y}_1 = y\left(k^*, \rho_j\right)$, for all j. Table 2 summarises all the targeted statistics mentioned above and Table 3 shows the benchmark parameter values. The resulting value of six key variables, namely the capital-labour ratio k^*, aggregate labour input $N(k^*)$, national income $Y(k^*)$, aggregate capital $k^* N(k^*)$, aggregate output $(k^*)^\alpha \cdot N(k^*)$ and government spending G^*, are reported in Table 4.[35]

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 distribution can be obtained by "hollowing out" the middle section and relocating the mass to the upper and lower

---

[33]Hence, we set $M = 6$, $\pi_j = 0.20$ for $j \in \{1, ..., 4\}$ and $\pi_5 = \pi_6 = 0.10$.

[34]We use net national product as our measure of national income in this calculation. Data on net national product are obtained from the NIPA.

[35]The MATLAB codes for generating these results are available from the authors' personal website.

ends. To put this in practice, first choose a value θ from the range $[0, \pi_3]$. Then define a new set of weights $\{\widetilde{\pi}_j\}$ on $\{\rho_j\}$ as follows: $\widetilde{\pi}_3 = \pi_3 - \theta$, $\widetilde{\pi}_1 = \pi_1 + \chi$, $\widetilde{\pi}_6 = \pi_6 + \theta - \chi$, and $\widetilde{\pi}_j = \pi_j$ for $j \in \{2, 4, 5\}$. The value of χ is chosen so that the mean value of ρ is unchanged, i.e.,

$$\sum_{j=1}^{6} \widetilde{\pi}_j \rho_j = \sum_{j=1}^{6} \pi_j \rho_j.$$

Using this procedure, we construct four alternative distributions of ρ with $\theta \in \{0.05, 0.10, 0.20, -0.10\}$. We then solve the baseline model under each of these distributions, while keeping all other parameters in Table 3 unchanged. In general, a larger value of θ represents a higher level of time preference heterogeneity. Thus, the distribution with $\theta = -0.10$ is actually *less* diverse than the benchmark distribution.

Figure 2 shows the national income function and aggregate labour supply function obtained under various values of θ, including the benchmark case $(\theta = 0)$.[36] Two results are immediate from these diagrams. Firstly, increasing the cross-sectional dispersion of ρ causes the national income function to shift upward. This pattern is consistent with the theoretical predictions in Proposition 5. Secondly, an increase in time preference heterogeneity will cause the aggregate labour supply function to shift down. In a robustness check below, we find the same pattern under two different values of ψ, namely 0.2 and 0.6. Thus, at least in this regard, the calibrated model behaves similarly as the "no-income-effect" model considered in Section 4.2 with a convex $v'(\cdot)$. The key variables obtained under these alternative distributions are shown in Table 4. Overall, these results suggest that increasing the cross-sectional dispersion in consumer's time preference has only a mild positive effect on k^*, but a more significant and negative impact on aggregate labour supply. The changes in other aggregate variables are largely driven by the changes in $N(k^*)$. We can further divide the effects on $N(k^*)$ into (i) a composition effect brought by the changes in the composition of the workforce [i.e., changes in $\{\pi_1, ..., \pi_6\}$], and (ii) a general equilibrium effect brought by the changes in k^* (which affect the level of individual labour supply). The contribution of these two effects are shown in Table 5. These results show that the negative impact of greater time preference on $N(k^*)$ is due to a dominating composition effect.

---

[36] The curves for $\theta = 0.05$ are almost indiscernible from those obtained from the benchmark case, hence they are omitted from these diagrams.

Table 5 Two Effects on Aggregate Labour Supply

	% Changes from Benchmark			
	$\theta = 0.05$	$\theta = 0.10$	$\theta = 0.20$	$\theta = -0.10$
Composition effect	-1.46%	-2.92%	-5.83%	2.92%
G.E. effect	0.94%	1.78%	3.22%	-2.53%

Changing the Elasticity of Labour Supply Table 6 reports the results obtained under two alternative values of ψ, which are 0.20 and 0.60. These results are both qualitatively and quantitatively similar to those reported in Table 4. This shows that our results are robust to changes in this parameter value.

Revenue-Equivalent Comparision We have also repeated the same exercise under the revenue-equivalent restriction. As before, we solve for the steady state under the alternative distributions of ρ (and $\psi = 0.40$). But this time, we re-calibrate the value of a_2 in each case so that (i) the government budget constraint is satisfied, and (ii) the value of G^* is the same as in the benchmark scenario. The results are shown in Table 7. Similar to those reported in Table 4 of the revised paper, an increase in time preference heterogeneity will have a mild positive impact on k^* and a negative impact on aggregate labour supply $N(k^*)$. This negative effect, however, is smaller under the revenue-equivalent restriction. This can be explained as follows: When there is an increase in time preference heterogeneity (say, from the benchmark case $\theta = 0$ to $\theta = 0.05$), the value of a_2 needs to be increased in order to maintain the same level of tax revenue. An increase in a_2 will raise both the level of $\tau(y)$ and the marginal tax rate $\tau'(y)$ for all $y > 0$. Formally, this means

$$\frac{\partial \tau(y; a_2)}{\partial a_2} > 0 \qquad \text{and} \qquad \frac{\partial}{\partial a_2} \left[\frac{\partial \tau(y; a_2)}{\partial y} \right] > 0, \qquad \text{for all } y > 0.$$

This will in turn lower the value of pre-tax income $y(k, \rho)$ and consumption $c(k, \rho)$ for all (k, ρ). The reduced consumption will then trigger a negative income effect on labour supply. To see this, first note that under the utility function

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

Table 6 Changing the Elasticity of Labour Supply

When $\psi = 0.20$

	Benchmark	% Changes from Benchmark			
	$\theta = 0$	$\theta = 0.05$	$\theta = 0.10$	$\theta = 0.20$	$\theta = -0.10$
k^*	4.204	0.02%	0.05%	0.09%	-0.04%
$N(k^*)$	18.254	-0.30%	-0.53%	-0.87%	1.18%
$Y(k^*)$	28.173	-0.29%	-0.52%	-0.84%	1.17%
$k^* N(k^*)$	76.739	-0.28%	-0.48%	-0.78%	1.14%
$(k^*)^\alpha \cdot N(k^*)$	32.420	-0.29%	-0.51%	-0.83%	1.16%
G^*	6.711	-0.34%	-0.60%	-0.97%	1.34%

When $\psi = 0.60$

k^*	4.204	0.02%	0.05%	0.09%	-0.04%
$N(k^*)$	20.526	-0.23%	-0.35%	-0.45%	1.27%
$Y(k^*)$	31.680	-0.22%	-0.34%	-0.42%	1.26%
$k^* N(k^*)$	86.290	-0.20%	-0.30%	-0.36%	1.23%
$(k^*)^\alpha \cdot N(k^*)$	36.455	-0.22%	-0.33%	-0.41%	1.25%
G^*	7.546	-0.26%	-0.40%	-0.52%	1.44%

Table 7 Additional Numerical Results (holding G^* constant)

	Benchmark	% Changes from Benchmark			
	$\theta = 0$	$\theta = 0.05$	$\theta = 0.10$	$\theta = 0.20$	$\theta = -0.10$
k^*	4.204	0.02%	0.04%	0.08%	-0.04%
$N(k^*)$	19.475	-0.06%	-0.11%	-0.20%	0.21%
$Y(k^*)$	30.058	-0.06%	-0.10%	-0.17%	0.20%
$k^* N(k^*)$	81.873	-0.04%	-0.07%	-0.12%	0.17%
$(k^*)^\alpha \cdot N(k^*)$	34.589	-0.05%	-0.10%	-0.17%	0.20%
a_2	0.402	0.405	0.408	0.412	0.391

the optimality condition for labour supply is given by

$$A\left[n\left(k,\rho,\varepsilon\right)\right]^{\frac{1}{\psi}} c\left(k,\rho\right) = \frac{w\left(k\right)}{r\left(k\right)}\rho\varepsilon$$

$$\Rightarrow n\left(k,\rho,\varepsilon\right) = \left[\frac{w\left(k\right)}{r\left(k\right)}\frac{\rho\varepsilon}{Ac\left(k,\rho\right)}\right]^{\psi}.$$

Thus, a decline in $c\left(k,\rho\right)$ will induce the workers to worker longer hours. This positive effect on labour supply will then (partially) counteract the negative composition effect reported in Table 5. Hence, the overall impact on aggregate labour supply is small under the revenue-equivalent restriction.

5 Conclusion

In this paper we analyse the long-run economic effects of diversity in a deterministic 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. Firstly, changing either the distribution of labour productivity or time preference will affect the composition of aggregate labour supply. The exact nature of this effect is determined by the shape of the individual labour supply function. Secondly, changing the distribution of time preference will also have an impact on the cross-sectional distribution of marginal tax rate. We show that the curvature of the marginal tax function holds the key in determining this effect. 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 equation (14) can be more succinctly expressed as $Y(k) = \Gamma(k) N(k)$. We will examine the 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, consider the national income function $Y(\cdot)$. Since $\phi(\cdot)$ is strictly increasing, $Y(\cdot)$ is strictly decreasing on $(k_{\min}, k_{\max})$ with

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

and

$$\lim_{k \to k_{\min}} Y(k) = \int_{\underline{\rho}}^{\overline{\rho}} \phi\left[1 - \frac{\rho}{\underline{\rho}}(1 - \overline{\tau})\right] dH_1(\rho). \tag{28}$$

Equation (28) follows from the facts that $\phi(\cdot)$ is a continuous function and $r(k)$ 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

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

Hence, the integral in (28) can be either convergent or divergent. For instance, if $H_1(\cdot)$ has a positive mass at $\underline{\rho}$, then $Y(k_{\min})$ is infinitely large and (17) 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 non-decreasing in k, for all (ρ, ε). Fix (ρ, ε) and suppose the contrary that $1 \geq n(k_2, \rho, \varepsilon) > n(k_1, \rho, \varepsilon) \geq 0$ 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

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

The third inequality follows from (12). The last step uses the facts that $w(\cdot)$ is strictly increasing

and $r(\cdot)$ is strictly decreasing. Since there is a contradiction, $n(k, \rho, \varepsilon)$ must be non-decreasing in k for all possible values of (ρ, ε).

The above results implies that $[Y(k) - \Gamma(k) N(k)]$ is strictly decreasing over the interval $(k_{\min}, k_{\max})$. If (16) and (17) are satisfied, then there exists a unique value k^* within this range that solves (14). 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 (16) and (17). 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 $1 \geq n(k, \rho, \varepsilon_2) > n(k, \rho, \varepsilon_1) \geq 0$ for some $\varepsilon_1 > \varepsilon_2$ in $[\underline{\varepsilon}, \overline{\varepsilon}]$. Since $\Psi(c, n)$ is strictly increasing in n, we have

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

which is a contradiction. Hence, $n(k, \rho, \varepsilon)$ must be non-decreasing in ε. If $n(k, \rho, \varepsilon)$ is an interior solution, then it is completely characterised by

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

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.[37]

---

[37] 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 $g\left(\cdot\right)$ defined on $\left[\underline{\varepsilon},\overline{\varepsilon}\right]$,*

$$\int_{\underline{\varepsilon}}^{\overline{\varepsilon}}\varepsilon g\left(\varepsilon\right)dH_2\left(\varepsilon\right)\leq\int_{\underline{\varepsilon}}^{\overline{\varepsilon}}\varepsilon g\left(\varepsilon\right)d\widetilde{H}_2\left(\varepsilon\right),\tag{29}$$

if and only if (21) *holds.*

Proof of Lemma A1 The proof of the "only if" part is obvious. Suppose (29) is valid for all bounded, non-decreasing functions defined on $\left[\underline{\varepsilon},\overline{\varepsilon}\right]$. For any $x\in\left[\underline{\varepsilon},\overline{\varepsilon}\right]$, define the indicator function $\mathcal{I}\left(\varepsilon;x\right)$ which equals one if $\varepsilon\geq x$ and zero otherwise. Since $\mathcal{I}\left(\varepsilon;x\right)$ is bounded and non-decreasing, it follows from (29) that

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

for any $x\in\left[\underline{\varepsilon},\overline{\varepsilon}\right]$. Next, consider the "if" part. Let $g\left(\cdot\right)$ be an arbitrary bounded, non-decreasing function defined on $\left[\underline{\varepsilon},\overline{\varepsilon}\right]$. Without loss of generality, we can assume $g\left(\underline{\varepsilon}\right)=0$. For any positive integer $m\geq 1$, partition the interval $\left[\underline{\varepsilon},\overline{\varepsilon}\right]$ into 2^m subintervals of equal length. The end-points of these intervals are given by

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

Next, define a step function $\eta_m\left(\cdot\right)$ so that $\eta_m\left(\varepsilon\right)=g\left(\widetilde{\varepsilon}_{i,m}\right)$ if $\varepsilon\in\left[\widetilde{\varepsilon}_{i,m},\widetilde{\varepsilon}_{i+1,m}\right)$, and $\eta_m\left(\varepsilon\right)=g\left(\overline{\varepsilon}\right)$ if $\varepsilon=\overline{\varepsilon}$. This function can also be written as

$$\eta_m\left(\varepsilon\right)=\sum_{i=1}^{m2^m}\lambda_{i,m}\mathcal{I}\left(\varepsilon;\widetilde{\varepsilon}_{i,m}\right),\tag{30}$$

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

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

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

By repeating the above procedure, we can construct a sequence of non-negative functions $\left\{\eta_m\left(\cdot\right)\right\}$ that converges pointwise to $g\left(\cdot\right)$. We now show that $\left\{\eta_m\left(\cdot\right)\right\}$ is a monotonically increasing sequence of functions, i.e., $\eta_m\left(\varepsilon\right)\leq\eta_{m+1}\left(\varepsilon\right)$ for any $\varepsilon\in\left[\underline{\varepsilon},\overline{\varepsilon}\right]$. Fix $\varepsilon\in\left[\underline{\varepsilon},\overline{\varepsilon}\right]$. Then there are

only two possible scenarios: either $\varepsilon < \left(\widetilde{\varepsilon}_{i,m} + \widetilde{\varepsilon}_{i+1,m}\right)/2$ or $\varepsilon \geq \left(\widetilde{\varepsilon}_{i,m} + \widetilde{\varepsilon}_{i+1,m}\right)/2$ for some $i \in \{1, ..., 2^m\}$. In the first scenario, we have $\eta_{m+1}\left(\varepsilon\right) = \eta_m\left(\varepsilon\right)$. In the second scenario,

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

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

$$\lim_{m\to\infty} \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon\eta_m\left(\varepsilon\right) dH_2\left(\varepsilon\right) = \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon g\left(\varepsilon\right) dH_2\left(\varepsilon\right), \tag{31}$$

$$\lim_{m\to\infty} \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon\eta_m\left(\varepsilon\right) d\widetilde{H}_2\left(\varepsilon\right) = \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} \varepsilon g\left(\varepsilon\right) d\widetilde{H}_2\left(\varepsilon\right). \tag{32}$$

Note that for each m, we have

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

where the first equality follows from (30). Suppose (21) is true for all $x \in [\underline{\varepsilon}, \overline{\varepsilon}]$. Since $\lambda_{i,m} \geq 0$ for all i, we have

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

Equation (29) then follows from (31) and (32). This completes the proof of Lemma A1. ∎

The results in Proposition 3 followed immediately by Lemma A1. Specifically, fix $k \in \left(k_{\min}, k_{\max}\right)$. By Lemma 2, $n\left(k, \rho, \varepsilon\right)$ is non-decreasing in ε for all $\rho \in \left[\underline{\rho}, \overline{\rho}\right]$. Thus, after integrating it over the distribution of ρ, the resulting function

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

is bounded and non-decreasing in ε. Then by Lemma A1, (21) is sufficient to ensure that

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

for any $k \in \left(k_{\min}, k_{\max}\right)$. This completes the proof of Proposition 3.

Proof of Proposition 4

Let k^* and $\widetilde{k}^*$ be the unique solution of (14) under $H_2(\cdot)$ and $\widetilde{H}_2(\cdot)$, respectively. As shown in the proof of Lemma 1, $N(\cdot)$ is non-decreasing while $Y(\cdot)$ is strictly decreasing. These results hold regardless of the distributions of ε. Hence, $\widetilde{N}(\cdot)$ and $\widetilde{Y}(\cdot)$ will have the same properties. Suppose the contrary that $k^* < \widetilde{k}^*$. Then we have

$$
\begin{aligned}
Y\left(\widetilde{k}^*\right) \;&<\; Y\left(k^*\right) = \Gamma\left(k^*\right) N\left(k^*\right) \\
&<\; \Gamma\left(\widetilde{k}^*\right) N\left(\widetilde{k}^*\right) \\
&\le\; \Gamma\left(\widetilde{k}^*\right) \widetilde{N}\left(\widetilde{k}^*\right) = Y\left(\widetilde{k}^*\right).
\end{aligned}
\tag{33}
$$

The first inequality uses the fact that $Y(\cdot)$ is strictly decreasing. The second inequality uses the fact that $\Gamma(\cdot)$ and $N(\cdot)$ are both strictly positive and increasing. The third inequality uses the result in Proposition 3. The last equality follows from the definition of $\widetilde{k}^*$. Since there is a contradiction in (33), it must be the case that $k^* \ge \widetilde{k}^*$. Since $y(k,\rho)$ and $c(k,\rho)$ are strictly decreasing in k for all $\rho \in [\underline{\rho}, \overline{\rho}]$, we have $y(k^*,\rho) \le y\left(\widetilde{k}^*,\rho\right)$ and $c(k^*,\rho) \le c\left(\widetilde{k}^*,\rho\right)$. Since $Y(\cdot)$ is strictly decreasing and $\Gamma(\cdot)$ is increasing, we have $Y(k^*) \le Y\left(\widetilde{k}^*\right)$ and $\Gamma(k^*) \ge \Gamma\left(\widetilde{k}^*\right)$. Taken together, these imply

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

This completes the proof of Proposition 4.

Proof of Proposition 5

The proof of Proposition 5 uses 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

$$
\tau'\left(\alpha y_1 + (1-\alpha) y_2\right) \gtrless \alpha \tau'(y_1) + (1-\alpha)\tau'(y_2)
$$

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

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

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

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

Suppose $\widetilde{H}_1\left(\cdot\right)$ is a mean-preserving spread of $H_1\left(\cdot\right)$ and $\tau'\left(\cdot\right)$ is convex so that $\phi\left(\cdot\right)$ 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 $H_1\left(\cdot\right)$ to $\widetilde{H}_1\left(\cdot\right)$ will shift the $Y\left(k\right)$ curve in Figure 1a to the left. Hence, we have $\widetilde{k}^* \leq k^*$. A similar argument can be used to establish the result in part (ii).

Proof of Proposition 6

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

$$\frac{\int_{\sigma(q)}^{\overline{\rho}} \rho dH_1\left(\rho\right)}{\int_{\sigma(q)}^{\overline{\rho}} dH_1\left(\rho\right)} \leq \frac{\int_{\widetilde{\sigma}(q)}^{\overline{\rho}} \rho d\widetilde{H}_1\left(\rho\right)}{\int_{\widetilde{\sigma}(q)}^{\overline{\rho}} d\widetilde{H}_1\left(\rho\right)} \quad \text{and} \quad \frac{\int_{\underline{\rho}}^{\sigma(q)} \rho dH_1\left(\rho\right)}{\int_{\underline{\rho}}^{\sigma(q)} dH_1\left(\rho\right)} \geq \frac{\int_{\underline{\rho}}^{\widetilde{\sigma}(q)} \rho d\widetilde{H}_1\left(\rho\right)}{\int_{\underline{\rho}}^{\widetilde{\sigma}(q)} d\widetilde{H}_1\left(\rho\right)},$$

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

$$\int_{\sigma(q)}^{\overline{\rho}} \rho dH_1\left(\rho\right) \leq \int_{\widetilde{\sigma}(q)}^{\overline{\rho}} \rho d\widetilde{H}_1\left(\rho\right), \tag{34}$$

$$\int_{\underline{\rho}}^{\sigma(q)} \rho dH_1\left(\rho\right) \geq \int_{\underline{\rho}}^{\widetilde{\sigma}(q)} \rho d\widetilde{H}_1\left(\rho\right), \tag{35}$$

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

Using (9) and (34), 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 (9) and (35).

Proof of Lemma 7

Part (i) The desired result follows immediately from the observation that $\lim_{n \to 0} v'\left(n\right) = 0$ is incompatible with $v'\left(0\right) \geq w\left(k\right)\varepsilon\rho/r\left(k\right) > 0$, for any $k \in \left[k_{\min}, k_{\max}\right]$ and for all $\left(\rho, \varepsilon\right) \in \left[\underline{\rho}, \overline{\rho}\right] \times \left[\underline{\varepsilon}, \overline{\varepsilon}\right]$.

Part (ii) 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 (26), we can write

$$
w\left(k_{\max}\right)\overline{\varepsilon} < v'\left(1\right) \leq \frac{w\left(k\right)}{r\left(k\right)}\varepsilon\rho. \tag{36}
$$

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 $\left(\rho, \varepsilon\right) \in \left[\underline{\rho}, \overline{\rho}\right] \times \left[\underline{\varepsilon}, \overline{\varepsilon}\right]$, we have

$$
\frac{w\left(k\right)}{r\left(k\right)}\varepsilon\rho < \frac{w\left(k_{\max}\right)}{r\left(k_{\max}\right)}\overline{\rho} \cdot \overline{\varepsilon} = w\left(k_{\max}\right)\overline{\varepsilon}. \tag{37}
$$

The last equality follows from the fact that $r\left(k_{\max}\right) = \overline{\rho}$. Since (36) and (37) are contradictory, it must be the case that $n\left(k, \rho, \varepsilon\right) < 1$ 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 $k \in (k_{\min}, k_{\max})$ and $\varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon}]$. Suppose $v'(1) > w(k_{\max}) \overline{\varepsilon}$ is satisfied. Then by Lemma 7, we have $1 > n(k, \rho, \varepsilon) \geq 0$, or equivalently,

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

for all $\rho \in [\underline{\rho}, \overline{\rho}]$. Suppose the contrary that there exists $\rho_1 < \rho_2$ in $[\underline{\rho}, \overline{\rho}]$ such that $n(k, \rho, \varepsilon)$ is strictly concave over the interval $[\rho_1, \rho_2]$. Then for any $\alpha \in (0, 1)$, we can write

$$n(k, \rho_\alpha, \varepsilon) > \alpha n(k, \rho_1, \varepsilon) + (1 - \alpha) n(k, \rho_2, \varepsilon) \geq 0,$$

where $\rho_\alpha = \alpha \rho_1 + (1 - \alpha) \rho_2$. Since $v'(\cdot)$ is strictly increasing and concave, we can get

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

This contradicts the hypothesis that $n(k, \rho_\alpha, \varepsilon) > 0$. Hence, $n(k, \rho, \varepsilon)$ must be convex in ρ over the entire range $[\underline{\rho}, \overline{\rho}]$. 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 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] Bricker, J., Kennickell, A.B., Moore, K.B., and Sabelhaus, J. (2012) "Changes in U.S. Family Finances from 2007 to 2010: Evidence from the Survey of Consumer Finances," *Federal Reserve Bulletin*, vol. 98, pp.1-80.

[8] Bricker, J., Dettling, L.J., Henriques, A., Hsu, J.W., Jacobs, L., Moore, K.B., Pack, S., Sabelhaus, J., Thompson, J., and Windle, R.A. (2017) "Changes in U.S. Family Finances from 2013 to 2016: Evidence from the Survey of Consumer Finances," *Federal Reserve Bulletin*, vol. 103, pp.1-42.

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

[10] 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.

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

[12] 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.

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

[14] 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.

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

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

[17] 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.

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

[19] 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.

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

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

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

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

[24] King, R.G., Plosser, C.I., and Rebelo, S.T. (1988) "Production, Growth and Business Cycles: I. The Basic Neoclassical Model," *Journal of Monetary Economics*, vol. 21, pp.195-232.

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

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

[27] 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.

[28] 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.

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

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

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

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

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