Essays on the Size and the Redistributive Properties of Pay-As-You-Go Pension Systems

Christophe Hachon

Université Paris I Panthéon Sorbonne

Thèse pour le doctorat de Sciences Economiques
soutenue par

Christophe Hachon

le 4 Décembre 2008.

Essais sur la Taille et le Caractère Redistributif des Systèmes de Retraite par Répartition

---

Directeur de thèse

Monsieur Antoine d'Autume - Professeur Université Paris I

Composition du jury

Monsieur Pascal Belan - Professeur Université Nantes (rapporteur)
Monsieur Helmuth Cremer - Professeur Université Toulouse I
Monsieur Nicolas Drouhin - Maître de Conférences Ecole Normale Supérieure de Cachan
Madame Anne Lavigne - Professeur Université Orléans (rapporteur)
Monsieur Thomas Seegmuller - Chargé de Recherche CNRS

Remerciements

> "Imagination is more important than knowledge. For knowledge is limited, whereas imagination embraces the entire world, stimulating progress, giving birth to evolution."
>
> Albert Einstein

La rédaction d'une thèse est un exercice à la fois très stimulant mais aussi très périlleux. En effet, le cursus universitaire est construit sur la base de l'apprentissage de connaissances établies. La rédaction d'une thèse, quant à elle, requiert d'autres qualités assez sensiblement différentes de celles nécessaires jusqu'alors. Il s'agit d'abord de se situer sur la frontière des connaissances en fonction des ses affinités avec les différents points de la littérature. Ensuite, il faut non-seulement savoir décrypter la littérature existante, mais aussi repousser de façon marginale (au moins pour le moment) cette même frontière. La thèse implique donc que l'on doive développer de nouvelles qualités. Ce passage peut être naturel pour certains, ou peut nécessiter une remise en cause importante pour d'autres. J'appartiens plutôt à cette dernière catégorie de personnes, ce qui explique pourquoi ma première année de thèse a été relativement compliquée, et la liste de personnes à remercier assez longue.

Mes premiers remerciements s'adressent à Antoine d'Autume qui a su me rassurer alors que je commençais à m'inquiéter du temps nécessaire à l'apprentissage des méthodes de la recherche. Plus qu'un "directeur" de thèse, il a été un "accompagnateur" pour mon travail. Ce rôle est d'autant plus important que le facteur psychologique est primordial pour la réussite d'un projet comme celui-ci. Je tiens également à remercier Nicolas Drouhin qui a su développer mon intérêt et mes compétences pour la macro-économie. J'ai pu bénéficier de discussions très stimulantes et d'encouragements permanents de sa part. D'autre part, ce parcours universitaire n'aurait pu être le même sans la participation de Richard Racine. Il est à l'origine de ma réussite au concours d'entrée à l'ENS de Cachan, ce qui m'a permis de m'épanouir intellectuellement dans les meilleures conditions qui soient.

Mes remerciements s'adressent également aux acteurs des quelques séminaires auxquels j'ai participés car ils m'ont permis d'améliorer de manière significative la qualité de mon travail. Je remercie plus particulièrement J.O. Hairault, T. Krieger, T. Seegmuller et B. Wigniolle.

Mais la réussite d'une thèse repose également sur l'environnement familial et amical.

Je remercie en premier lieu mon épouse, Florelle, qui non-seulement a supporté les changements erratiques de mon humeur, mais qui m'a également encouragé lors des passages à vide. Elle m'a consacré du temps et ses compétences pour les relectures de mes articles en anglais. Malheureusement, mon niveau et son manque de temps ne lui ont permis que de corriger les erreurs de base pour rendre mes articles compréhensibles par tous. Mes remerciements s'adressent aussi à ma mère qui a su trouver les mots pour m'encourager et me soutenir dans les moments de doute et de remise en question. Enfin, je remercie ma famille proche (la famille Rafik, mes beaux-parents,...) pour leur soutien.

L'environnement amical a aussi joué un rôle très important lors de la rédaction de cette thèse. Je remercie Reynald-Alexandre Laurent sans qui ce travail n'aurait pu aboutir. Nos conversations stimulantes ont eu un impact significatif sur ma curiosité intellectuelle, qualité indispensable d'un chercheur. Je remercie également Sumudu Kankanamge pour ses encouragements et le temps qu'il a pu me consacrer. Mes pensées s'adressent aussi aux membres du laboratoire Eurêqua (Marie-Pierre, Natacha, Thomas, Victor,...). Enfin, je remercie mes amis auxerrois et egriselliens (Isabelle, Jack, Romain, Sebastien) qui ont cru en moi lors des baisses de moral.

Pour résumer, la rédaction d'une thèse est un processus complexe dans lequel interviennent ses compétences propres, celles des personnes qui encadrent le travail, et le soutien d'un nombre important d'acteurs. C'est l'ensemble de ces ingrédients qui ont concouru à l'aboutissement de ce travail.

Contents

General Introduction

> "(...) organisation of social insurance should be treated as one part only of a comprehensive policy of social progress. Social insurance fully developed may provide income security; it is an attack upon Want. (...) The main feature of the plan for social security is a scheme for social insurance against interruption and destruction of earning power (...)"
>
> William H. Beveridge
>
> *Social Insurance and Allied Services*, 1942

Depuis la fin de la seconde guerre mondiale, la *protection sociale* est devenue une dimension importante de l'intervention de l'Etat. Cette protection sociale consiste à mettre en place un ensemble de transferts sociaux permettant de couvrir les agents contre les aléas de la vie (chômage, maladie, accidents, vieillesse,...), mais aussi de permettre aux plus démunis de satisfaire leurs besoins essentiels (exemple : le Revenu Minimum d'Insertion en France). Ces deux composantes de la protection sociale renvoient respectivement aux notions d'assurance et de redistribution. L'Etat met en place un mécanisme assurantiel tout d'abord, i.e. une *redistribution horizontale* des ressources; en transférant des indemnités aux agents dont l'un des risques mentionnés précédemment se serait réalisé. Il s'agit d'éviter que les agents ne subissent des changements trop brutaux de leurs ressources. L'objectif n'est pas seulement de procurer aux individus un outil assurantiel qui n'aurait pu être proposé par une entreprise du secteur privé. Il est aussi de contraindre les agents à participer à cette activité. L'Etat est alors une institution bienveillante qui pallie à la possible myopie des agents. De plus, l'Etat met en place des transferts permettant aux agents les plus pauvres de satisfaire leurs besoins essentiels. On retrouve dans cette pratique des éléments de la philosophie Rawlsienne (1971) selon laquelle le critère des choix sociaux repose sur la maximisation du bien-être de la personne la plus défavorisée. Ces transferts assurent une *redistribution verticale* des ressources, i.e. une redistribution entre les catégories sociales de la population. Toutefois, cette redistribution verticale est limitée puisqu'elle ne concerne que les individus dont les besoins essentiels ne sont pas satisfaits. Notons dès à présent que ces deux notions de redistribution (horizontale et verticale) ne sont pas disjointes. En effet, les outils liés à la redistribution horizontale peuvent également être le support d'une redistribution verticale. Nous détaillerons ceci un peu plus loin.

Le développement des systèmes de protection sociale s'inscrit dans un contexte historique très particulier[1]: situation post-seconde guerre mondiale et post-crise des années 1930; allongement de la durée de vie; amélioration de la qualité de la vie du fait des progrès médicaux; développement politique du paradigme keynesien,... Le tableau 1 illustre l'accroissement très significatif des dépenses sociales[2] depuis la fin de la seconde guerre mondiale. Si elles représentaient moins de 5% du PIB en 1930, cette part s'échelonnait de 13 à 33% du PIB en 1995 dans la plupart des pays développés. Ce phénomène traduit le passage progressif d'une *solidarité subjective* (en son âme et conscience) vers une *solidarité objective* (organisée par l'Etat). L'Etat se substitue alors aux institutions privées et aux structures de marché pour promouvoir la solidarité entre les individus (Merrien 2007)[3]. Toutefois, ce développement de la protection sociale n'a pas été homogène. Ainsi, les Etats-Unis et le Royaume-Uni semblent avoir opté pour un système dans lequel les dépenses sociales sont limitées. Dans ce cadre, les fonctions assurantielles sont pour une large part déléguées aux assurances privées et l'intervention de l'Etat se cantonne essentiellement à aider les plus démunis[4]. Des pays comme l'Allemagne, la France ou l'Italie ont quant à eux opté pour des systèmes de protection sociale assez développés mais dans lesquels le mécanisme assurantiel joue un rôle prépondérant. Ces systèmes n'assurent qu'une faible redistribution entre les catégories sociales[5]. Enfin, les pays scandinaves, dont la Suède, semblent avoir opté pour une protection sociale très développée puisque les dépenses sociales représentent près du tiers de la richesse nationale. Pour ces pays, la protection sociale est autant un instrument assurantiel qu'un outil de redistribution[6].

Parmi les risques mentionnés précédemment, il en est un qui a pris une place prépondérante : le risque vieillesse. En effet, les dépenses de pensions représentent 48.1% du total des prestations sociales dans l'union européenne en 2000 (source : eurostat). Plus spécifique-

[1]Nous ne mentionnons ici que quelques facteurs ayant influencé le développement de la protection sociale. Nous sommes conscients qu'une étude plus approfondie serait nécessaire, mais celle-ci sort du cadre de cette thèse. Il s'agirait notamment de prendre chaque risque, et de voir dans quelle mesure la couverture de ce risque répondait à une attente sociétale.

[2]Ici définies comme la somme des transferts minimum de survie, des allocations chômage, des systèmes de retraite publics, des dépenses publiques de santé et des aides au logement.

[3]Cette conception d'une substituabilité entre solidarité publique et privée a notamment été remise en cause par Lindert (2004).

[4]Dans la typologie de Esping-Andersen (2007), il s'agit du *modèle libéral*. Cet auteur définit trois types de modèles sociaux : le modèle libéral, le modèle conservateur et le modèle socio-démocrate.

[5]Il s'agit du *modèle conservateur* ou *corporatiste* (Esping-Andersen 2007).

[6]Il s'agit du *modèle socio-démocrate* (Esping-Andersen 2007).

Table 1: **Transferts sociaux dans les pays de l'OCDE, 1880-1995, en % du PIB à prix courants**, *Source* : Lindert (1994), tableau 1.2 p.12.

Pays							OCDE ancien[a]			OCDE nouveau[b]		
	1880	1890	1900	1910	1920	1930	1960	1970	1980	1980	1990	1995
Allemagne	0.5	0.53	0.59	..[c]	..	4.82	18.1	19.53	25.66	20.42	19.85	24.92
Etats-Unis	0.29	0.45	0.55	0.56	0.7	0.56	7.26	10.38	15.03	11.43	11.68	13.67
France	0.46	0.54	0.57	0.81	0.64	1.05	13.42	16.68	22.55	22.95	23.7	26.93
Italie	0[d]	0	0	0	0	0.08	13.1	16.94	21.24	17.1	21.34	23.71
Pays-Bas	0.29	0.3	0.39	0.39	0.99	1.03	11.7	22.45	28.34	26.94	27.59	25.7
Royaume-Uni	0.86	0.83	1	1.38	1.39	2.24	10.21	13.2	16.42	16.94	18.05	22.52
Suède	0.72	0.85	0.85	1.03	1.14	2.59	10.83	16.76	25.94	29.78	32.18	33.01

[a] Ancienne base de données de l'OCDE.
[b] Nouvelle base de données de l'OCDE.
[c] .. : Positif mais chiffre non-disponible.
[d] Absence de transferts sociaux.

ment, ce chiffre est de 45.7% pour l'Allemagne, 46.7% pour la France, 46.8% pour le Royaume-Uni et 60.5% pour l'Italie. De plus, la part de la richesse nationale qui est consacrée aux pensions est de 13.1% pour l'Allemagne, 13.2% pour la France, 11.8% pour le Royaume-Uni et 14.7% pour l'Italie. Ces chiffres témoignent de l'importance actuelle des transferts sociaux liés aux retraites. En outre, l'organisation de ces transferts peut prendre différentes formes. En effet, les systèmes de retraite peuvent adopter soit une structure par répartition, soit une structure par capitalisation[7]. Les systèmes de retraite par capitalisation reposent sur le principe d'un *transfert inter-temporel* de ressources. Ainsi, le système par capitalisation perçoit les cotisations versées pendant la période d'activité par les agents. L'organisme place cette somme, et la redistribue corrigée par le rendement du placement, à cette même génération lorsque ceux-ci seront à la retraite. Dans ce cas, les agents cotisent pour eux-mêmes. Chaque génération subit alors un risque plus ou moins aléatoire[8] lié au rendement de cet investissement. Les systèmes de retraite par répartition reposent quant à eux sur le principe de la *solidarité inter-générationnelle*. Les prélèvements effectués sur la population active de la période t sont redistribués aux retraités de cette même période. Si les systèmes de retraite par capitalisation peuvent être organisés ou bien par l'Etat, ou bien par des entreprises du secteur privé, le système par répartition quant à lui ne peut être organisé que par l'Etat. En effet, il suppose qu'il existe un contrat entre les générations, et que le versement de cotisations correspond à l'engagement des générations à venir (certaines qui ne sont pas encore nées d'ailleurs) de cotiser également à ce système. Au moins deux formes de risques sont propres à ce système. Il existe d'abord un risque politique puisqu'il n'est pas certain que le contrat inter-générationnel ne soit pas rompu si le montant des cotisations devient trop important. Ce point pose notamment problème actuellement du fait du poids fiscal nécessaire au financement des retraites par répartition[9]. Le deuxième risque porte sur l'évolution de la masse cotisante, i.e. sur le montant collecté par l'Etat une fois l'agent à la retraite[10].

Les systèmes de retraite par capitalisation et les systèmes par répartition semblent donc pouvoir être distingués (i) par les mécanismes de transferts qu'ils opèrent, mais aussi (ii) par

[7]Belan et Pestieau (1998) définissent un ensemble plus complet de critères permettant de classer les systèmes de retraite.

[8]Selon la composition de cet investissement : actions, obligations,...

[9]Les réformes actuelles peuvent alors être vues comme un moyen d'éviter la rupture de ce contrat.

[10]Ce montant des cotisations dépend de la taille de la population active occupée, mais aussi du taux de croissance des salaires.

les risques qui leur sont associés. Dans ce travail nous allons essentiellement nous intéresser au cas des systèmes de retraite par répartition, et nous allons considérer implicitement que le système par capitalisation est parfaitement substituable au mécanisme de l'épargne privée[11]. Le deuxième point que nous n'exploiterons pas ici concerne les différents risques propres à chacun des systèmes. Ainsi, dans tout notre travail, nous supposerons que l'environnement macro-économique et que le rendement de l'épargne privée sont certains.

Mais il existe un autre risque partagé par les deux types de systèmes de retraite : le risque de survie. En effet, il est possible que les agents cotisent pour leur retraite sans pouvoir complètement bénéficier des fruits de leur épargne (pour le système par capitalisation), ou bien des fruits du contrat inter-générationnel (système par répartition). La distribution de la probabilité de survie joue alors un rôle central dans cette analyse. Ainsi, la rectangularisation de cette fonction, qui a été observée dans la plupart des pays développés, correspond à une réduction de l'incertitude portant sur la durée de vie. Dans ce travail nous n'exploiterons pas vraiment cette dimension. Lorsque nous introduirons la durée de vie des agents dans notre modélisation, nous supposerons qu'elle est certaine (d'Autume 2003). Toutefois, il est intéressant de noter que la représentation analytique de cette hypothèse est identique à celle des modèles à générations imbriquées à la Yaari (1965). Dans ces modèles, il est supposé qu'il existe un événement aléatoire tel que les agents peuvent ou bien ne vivre qu'une période, ou bien vivre complètement leurs deux périodes de vie (Chakraborty 2004, Drouhin 1997, 2001a, 2001b).

Si dans cette thèse nous n'exploitons pas vraiment la notion de risque, nous allons en revanche exploiter l'autre dimension des systèmes de retraite par répartition à savoir leur aspect redistributif. Plus spécifiquement, nous utiliserons les propriétés liées à leur redistribution inter-générationnelle, mais aussi celles liées à la redistribution verticale des ressources.

La propriété de redistribution inter-générationnelle des ressources est liée au mode de fonctionnement même des systèmes de retraite par répartition. Elle implique que la population active à une date donnée cotise en vue de financer les pensions des retraités vivant à cette même date. Si la population est homogène, alors le rendement de ce système va dépendre du taux de croissance des recettes fiscales qui lui sont affectées, i.e. du taux de croissance des salaires et du taux de croissance de la population (Aaron 1966). En univers certain, ce rendement doit être comparé au coût d'opportunité de ces cotisations,

[11] Voir Drouhin (1997) pour une discussion sur ce point.

i.e. au rendement de l'épargne. Empiriquement, il est possible d'observer que le rendement moyen de l'épargne est très nettement supérieur à celui des systèmes de retraite par répartition (Dutta *et al.* 2000). Cette situation correspond au cas théorique d'une économie en sous-accumulation. Ce résultat est la source d'une vaste littérature qui s'est intéressée à l'existence même des systèmes de retraite par répartition, compte tenu du fait que la richesse de sécurité sociale de toute nouvelle génération est négative. En effet, la richesse de sécurité sociale correspond à la valeur actualisée de l'ensemble des taxes payées, et de l'ensemble des pensions reçues du système de retraite par répartition. Or, si le taux d'actualisation est plus grand que le taux de rendement du système de retraite par répartition, alors la valeur actualisée des bénéfices reçus durant la période de retraite ne permet pas de compenser le coût lié aux cotisations. Autrement dit, le coût d'opportunité lié à la participation au système de retraite par répartition est trop important.

Mais les systèmes de retraite par répartition peuvent également être le support d'une redistribution verticale des ressources, ou bien à travers son système fiscal, ou bien à travers les pensions versées aux agents. Pour simplifier, nous supposerons que la seule source d'hétérogénéité entre nos agents est constituée par leur niveau de qualification. De plus, nous ne nous intéresserons pas spécifiquement au système fiscal puisque nous supposerons que le financement des transferts s'effectue grâce à un taux de taxe uniforme pour toute la population. En revanche, nous exploiterons plus précisément la notion de redistribution verticale à travers le système de pensions. En vue de bien comprendre notre définition de la redistribution verticale au sein des systèmes de retraite par répartition, il est préférable de définir préalablement ce que l'on entend par "redistribution verticale pure".

i. La Redistribution Verticale Pure

Le principe de la redistribution verticale pure consiste à modifier la répartition initiale des revenus et des patrimoines au sein de la population de façon à limiter les inégalités économiques. Prenons le cas simple où la seule source de revenus des ménages, sans l'intervention de l'Etat, est composée du salaire. Un des moyens à la disposition de l'Etat en vue de corriger cette distribution brute de la richesse est de mettre en place une politique fiscale et de transferts . Ainsi, si l'Etat applique une taxe proportionnelle identique sur les salaires de tous les agents et qu'il redistribue le montant collecté de manière uniforme, alors il assure une redistribution verticale des ressources (Meltzer et Richard 1981). Le rendement fiscal de la taxe va évidemment dépendre de la distortion qu'elle engendre sur les comportements économiques, notamment en terme d'offre de travail. La redistribution des ressources peut alors être mesurée par la contribution nette de chaque agent au système de transferts. La contribution nette se définit comme la valeur actualisée de l'ensemble des taxes payées, et de l'ensemble des transferts reçus de l'Etat[12]. Une contribution nette positive signifie que le montant versé excède le montant reçu du système de transferts. Inversement, une contribution nette négative signifie qu'un agent a plus bénéficié qu'il n'a payé au système de transferts.

La notion de redistribution verticale est également liée à celle de progressivité d'un système de transferts. Cette progressivité peut porter soit sur le système fiscal utilisé, soit sur les versements, ou bien encore sur la redistribution globale du système. Si le taux de taxe appliqué est une fonction croissante des ressources sur lesquelles il s'applique alors un système fiscal est progressif. Des versements sont progressifs s'ils dépendent négativement de la richesse brute de l'agent. Finalement, nous dirons qu'un système de transferts est progressif si les contributions nettes des agents sont une fonction croissante de la richesse brute de l'agent[13]. Dans cette thèse, nous ne détaillerons que la progressivité des systèmes de versements ainsi que la progressivité globale d'un système de transferts. Ceci est lié à notre représentation très simplifiée du système fiscal puisqu'il va s'agir d'appliquer le même

[12]Notons que la richesse de sécurité sociale est une mesure de la contribution nette appliquée aux systèmes de retraite par répartition.

[13]Cette définition s'appuie largement sur la définition de la progressivité donnée par Coronado *et al.* (2000). Dans leur article, la progressivité est mesurée par le coefficient de Gini des revenus. Nelissen (1999) quant à lui utilise l'indice de Theil. Cet indice est obtenu par $T = \sum(X_i/Y).\ln(X_i/\bar{X})$, avec X_i le revenu de la personne i, Y la somme des revenus ($\sum X_i$) et $\bar{X}$ le revenu moyen. Nous aurons $T = 0$ si tous les agents de l'économie perçoivent le même revenu, et $T = \ln(X_j/\bar{X})$ si tout le revenu est détenu seulement par le groupe d'agents j. Notre définition est donc plus restrictive que celle que ces deux auteurs en donnent puisqu'elle implique que les contributions nettes sont des fonctions strictement croissantes du salaire initial de l'agent. Ceci implique nécessairement que l'indice de Gini et l'indice de Theil seront plus faibles après l'intervention de l'Etat. Tandis que les indices utilisés par Coronado *et al.* (2000) et par Nelissen (1999) correspondent à la progressivité moyenne induite par le système de retraite, notre définition définit une progressivité au sens stricte. Une version encore plus restrictive serait de définir la progressivité comme la situation dans laquelle les contributions nettes augmentent plus que proportionnellement au revenu. Notre définition est donc plus souple que celle-ci mais plus restrictive que celle retenue dans la littérature empirique. Notons que Borck (2007) mentionne la progressivité et la régressivité des systèmes de retraite sans définir explicitement ce qu'il entend par là. Cependant, il semble que notre définition corresponde le mieux à ce que ce que l'auteur voulait dire.

taux de taxe sur les salaires de tous les agents.

ii. Les Retraites Comme Outil de Redistribution Verticale ?

Si les systèmes de retraite par répartition sont un outil de redistribution horizontale, ils peuvent aussi être le support d'une redistribution verticale des ressources. Toutefois, du fait du décalage temporel entre la perception des cotisations et le versement de la pension, il peut être pertinent de distinguer deux instruments permettant de mesurer la redistribution verticale d'un système de retraite par répartition (Legros 1994).

Tout d'abord, les systèmes de retraite peuvent assurer une redistribution des ressources mesurées par unité de temps. Dans ce cas, un système de retraite assure une *redistribution instantanée* des ressources si le taux de remplacement[14] des pensions est une fonction décroissante du niveau de salaire de l'agent. Prenons l'exemple suivant (Casamatta *et al.* 2000) : soit une économie dans laquelle les agents vivent deux périodes. Ils travaillent durant la première et cotisent pour le système de retraite, et perçoivent une pension durant leur deuxième période de vie. Les individus diffèrent par leur niveau de salaire. Le taux de remplacement (β) s'écrit:

$$\beta(w) = \frac{\nu(\lambda w + (1-\lambda)\bar{w})}{w}$$

avec ν le taux de remplacement du salaire moyen de l'économie, w le salaire de l'agent, et $\bar{w}$ le salaire moyen de l'économie. λ mesure le degré d'indexation des pensions sur le salaire d'activité de l'agent. Si $\lambda = 1$, alors les pensions ne dépendent que du salaire d'activité et le système n'assure pas de redistribution instantanée des ressources puisque le taux de remplacement est le même quel que soit le niveau de salaire. Nous dirons alors que le système est *Bismarckien*[15]. Inversement, si $\lambda = 0$, alors les pensions versées sont les mêmes pour tous et le taux de remplacement est une fonction strictement décroissante du niveau de salaire. Dans ce cas simple, les pensions assurent une redistribution instantanée maximale des ressources. Comme dans Casamatta *et al.* (2000), nous définissons

[14]Le taux de remplacement est le rapport entre le montant de la pension reçue par unité de temps et le salaire d'activité de l'agent.

[15]Cette définition revient à simplifier le système Bismarckien à une seule composante. En faisant ainsi, nous occultons les autres dimensions de ce modèle social dont son mode d'organisation. La littérature sur la protection sociale distingue traditionnellement le système Bismarckien du système Beveridgien (voir Merrien (2007) entre autres).

un tel système comme étant *Beveridgien*[16]. Enfin, si $\lambda \in (0,1)$, alors nous dirons que le système de retraite est mixte dans le sens où il est composé à la fois d'éléments Bismarckiens, mais aussi d'éléments Beveridgiens. Notre définition du système Beveridgien se rapproche plus de sa conception scandinave que de sa conception anglo-saxone. Dans le cas scandinave (modèle socio-démocrate), la redistributivité et les montants distribués sont importants. En revanche, dans le cas anglo-saxon, le système de transfert est beaucoup moins développé et verse des pensions forfaitaires minimales. Cette modélisation n'est pas sans conséquences. En effet, elle ne permet pas de rendre compte, par exemple, de l'existence simultanée d'un minimum vieillesse et d'un système Bismarckien de pensions au-delà de ce "plancher de survie". Plus généralement, de par les simplifications qu'elle opère, cette modélisation ne permet pas de rendre compte de la complexité de tous les systèmes de retraite par répartition. Une telle volonté conduirait à construire une équation différente pour chaque pays pour les pensions. En revanche, cette modélisation constitue une première approximation permettant de concevoir et de représenter simplement leur redistributivité instantanée.

Les systèmes de retraite par répartition peuvent donc être le support d'une redistribution verticale instantanée des ressources. Toutefois, cette redistribution instantanée n'assure pas nécessairement une *redistribution de long-terme* des ressources (Legros 1994). Celle-ci peut être mesurée par la contribution nette d'un groupe d'agents au système de retraite. La valeur des cotisations versées au système de retraite va dépendre du niveau de salaire et de la durée d'activité, tandis que la valeur des pensions obtenues par l'agent va dépendre du montant des pensions par unité de temps et de la durée pendant laquelle il bénéficie du système de retraite. Ainsi, le temps passé à s'éduquer et le temps de loisir auront un impact négatif sur le montant cotisé. Concernant la valeur totale des pensions perçues, l'âge de départ en retraite influence négativement cette valeur, tandis que l'espérance de vie, en augmentant la période de perception des pensions, a une influence positive sur cette valeur. Par conséquent, la redistribution instantanée n'assure pas nécessairement une redistribution de long-terme des ressources. Pour illustrer ceci, reprenons le cadre très simple que nous avions décrit précédemment et introduisons le fait que plus les agents ont un salaire élevé et plus ils vivent longtemps[17]. Si $T(w)$ représente la durée de vie (supposée certaine) d'un agent ayant un salaire w (avec $T'(w) > 0$), alors la contribution

[16]De même, les autres dimensions du système Beveridgien ne sont prises en compte.

[17]Des éléments empiriques sur ce point seront fournis ci-après.

nette de cet agent au système de retraite peut s'écrire:

$$CN(w) = w\tau - \frac{\nu(\lambda w + (1-\lambda)\bar{w})T(w)}{1+\delta}$$

avec δ le taux d'actualisation. Les agents cotisent une fraction de leur salaire et reçoivent une pension aussi longtemps qu'ils restent en vie. Un système de retraite sera dit progressif à long-terme si les contributions nettes des agents sont des fonctions strictement croissantes du niveau de salaire. Autrement dit, la progressivité de long-terme implique que plus un agent dispose d'un niveau de salaire important et moins il bénéficie du système de transferts. Or, dans notre exemple, le fait que la durée de vie soit corrélée avec le niveau de salaire implique que la *progressivité instantanée* (mesurée par le lien entre taux de remplacement et niveau de salaire) n'assure pas nécessairement qu'un système de retraite soit *progressif à long-terme*. Dans cette thèse, lorsque nous distinguerons ces deux notions, nous mettrons l'accent sur l'impact de l'introduction de différences d'espérances de vie en omettant les effets liés à l'âge de départ en retraite ou bien ceux liés au temps d'éducation. Nous ne faisons pas ce choix en supposant que ces deux variables n'auraient pas d'impacts significatifs sur nos résultats, mais bien plutôt dans un souci de clarté des résultats analytiques et de mise en évidence des résultats de base, i.e. en ne prenant en compte que l'un des critères affectant la redistribution de long-terme.

Les systèmes de retraite par répartition assurent donc une redistribution inter-générationnelle, mais aussi une redistribution intra-générationnelle des ressources. L'ensemble de ces aspects redistributifs des systèmes de retraite par répartition semblent donc être suffisamment complexe pour mériter un examen plus approfondi.

Organisation et Contenu de cette Thèse

L'objectif de cette thèse consiste donc à étudier l'impact des redistributions opérées par les systèmes de retraite par répartition. Plus précisément, nous discutons trois types d'arguments dans notre travail, chacun d'eux constitue une partie de cette thèse.

La première question porte sur l'existence même des systèmes de retraite par répartition. Nous développons alors des arguments positifs et normatifs. Les deux parties suivantes se distinguent très clairement par le fait que la deuxième partie ne s'intéresse qu'à la redistributivité instantanée des systèmes de retraite par répartition, tandis que la troisième et dernière partie cherche à exploiter les différences existantes entre la redistributivité instantanée et la redistributivité de long-terme des systèmes de retraite par répartition.

Dans une première partie, nous étudions la détermination endogène des systèmes de retraite par répartition. L'analyse d'une telle évolution ne saurait être faite sans préciser le rôle prépondérant qu'a joué et que continue d'exercer la démographie sur cette variable. Pour cela, nous considérons tout d'abord que ce système de transferts n'opère qu'une redistribution inter-générationnelle des ressources. La prise en compte d'une redistribution intra-générationnelle nous a conduit à élargir notre analyse au cas de la détermination de transferts verticaux purs tout d'abord, mais aussi de la détermination de systèmes de transferts assurant à la fois des transferts verticaux purs et des transferts inter-générationnels. L'idée essentielle est que la redistribution inter-générationnelle n'est qu'un élément de la redistribution opérée par les Etats.

i. Le Cas de la Redistribution Horizontale Pure

Dans le premier chapitre, nous essayons d'étudier la façon dont est déterminée la taille d'un système de retraite par répartition lorsque les agents ne diffèrent que par leur âge, i.e. le système de retraite n'opère de redistribution qu'entre les générations. Notre objectif est de rendre compte de l'impact du vieillissement démographique sur cette variable. Ce vieillissement se caractérise par une hausse de l'espérance de vie ainsi que par une diminution du taux de fécondité. Par exemple, en France, l'espérance de vie à la naissance est passée de 66.5 ans en 1950 à 79 ans en 2000, tandis que dans le même temps le taux de fécondité est passé de 2.73 à 1.8 entre ces deux dates (Nyce et Schieber 2005). Cependant, l'impact de ce vieillissement démographique sur la taille du système de sécurité sociale n'est pas trivial. En vue de souligner ceci, nous utilisons deux arguments : un premier argument qui étudie l'impact de ce changement démographique dans le cadre d'une analyse positive, et un second argument dans le cadre d'une analyse normative.

Tout d'abord, nous utilisons un modèle à générations imbriquées en temps continu, en petite économie ouverte, dans lequel les agents ne diffèrent que par leur âge et dont la durée de vie est certaine. Chaque agent vote à chaque période pour la taille du système de sécurité sociale en ne considérant que ses cotisations et ses bénéfices futurs. Ainsi, la valeur actualisée des cotisations d'une personne dont l'âge est proche de l'âge de départ en retraite sera faible tandis que les bénéfices associés seront élevés. Autrement dit, la richesse de sécurité sociale est une fonction croissante de l'âge de l'agent. Dans ce cas, un vieillissement de la population implique une augmentation de l'âge médian des votants ce qui exerce une influence positive sur le taux de taxe du système de retraite. En revanche, ce vieillissement implique que le nombre de bénéficiaires de ces transferts augmente tandis que le nombre de cotisants se réduit, ce qui diminue le montant des pensions par retraité et par unité de temps. Ceci exerce une influence négative sur la taille du système de retraite. Au final, l'impact net du vieillissement démographique sur le taux de cotisation au système de retraite par répartition est ambigu. A fortiori, son impact sur le taux de remplacement l'est également. Ce modèle illustre un résultat standard de la littérature en économie politique portant sur l'ambiguïté de la relation entre la taille des systèmes de retraites et le vieillissement de la population (Uebelmesser 2004, Persson and Tabellini 2000, Galasso et Profeta 2004, 2007).

Dans la seconde section de ce chapitre, nous utilisons un modèle à générations imbriquées en économie fermée dans lequel les agents ne vivent qu'une fraction de leur seconde période de vie. Nous reprenons donc une structure à la Diamond à la différence près que le facteur d'actualisation n'est plus uniquement un facteur d'escompte psychologique pur, mais il dépend également d'un élément objectif que constitue la durée de vie (Drouhin 2001b). Nous montrons alors que pour atteindre l'optimum social, un vieillissement de la population peut induire une évolution en cloche du taux de remplacement du système de retraite par répartition. En revanche, le taux de taxe reste une fonction croissante de ce changement démographique. Ces résultats qualitatifs concernant la relation entre vieillissement et taille du système de retraite par répartition, sont les mêmes que ceux obtenus par Atkinson (2000) bien que celui-ci ait utilisé une structure complètement différente. En effet, celui-ci se place dans le cadre d'une petite économie ouverte dans laquelle les individus diffèrent par leur niveau de salaire et le gouvernement décide de la taille des systèmes de retraite étant donné son aversion aux inégalités. Il alors est intéressant de noter que ces deux modèles normatifs, dont la structure de base est très différente, conduisent au même résultat qualitatif.

Ce premier chapitre montre que l'on peut expliquer l'évolution récente des systèmes de retraite par répartition. En effet, il est possible d'observer que le taux de remplacement moyen et le taux de taxe ont augmenté dans la plupart des pays développés. Par exemple, pour la France, le taux de remplacement moyen est passé de 0.5 en 1975 à 0.65 en

1995 (Nyce et Schieber 2005). Dans le même temps, le taux de taxe sur les salaires utilisé pour financer le système de retraite est passé de 8.5 à 19.8%. Ce chiffre a même atteint environ 25% en 2007[18]. En revanche, depuis les années 1980, le poids fiscal des systèmes de retraites est tel que leur générosité a largement été remise en question via des réformes structurelles importantes (Galasso et Profeta 2004).

ii. Le Cas de la Redistribution Verticale Pure

La prise en compte de la solidarité intra-générationnelle nous a tout d'abord conduit à nous intéresser au cas le plus extrême de celle-ci, ce que nous faisons **dans notre deuxième chapitre**. Ainsi, nous nous sommes tout d'abord placés dans le cas où les prélèvements et les versements sont effectués à la même période et en faveur de la même génération d'agents. Dans ce chapitre, nous développons un modèle d'économie politique à la Meltzer et Richard (1981). Le modèle de base de ces auteurs est que les agent vivent une période durant laquelle ils décident de leur offre de travail. Ces agents diffèrent par leur niveau de productivité et votent sur la taille d'un système de transferts qui applique un taux de taxe identique sur les salaires et qui redistribue le montant ainsi collecté de façon forfaitaire. L'environnement de cette économie est purement statique, i.e. la structure de la population ainsi que le niveau des salaires par unité de temps sont donnés. Le modèle montre que l'identité de l'électeur médian est la même que celle de l'individu ayant le niveau de productivité médian. Le système de transferts a un impact distorsif sur l'offre de travail. D'autres auteurs, tels que Persson et Tabellini (1994) ou bien Alesina et Rodrik (1994), ont effectué un premier pas en faveur de l'introduction d'un environnement dynamique dans lequel les agents votent sur la taille d'une redistribution verticale des ressources. Leur objectif est d'expliquer la corrélation négative qu'ils observent entre taux de croissance et redistribution intra-générationnelle. Si leur modèle admet bien une dynamique d'accumulation, en revanche la structure de la population est statique et elle n'est pas modifiée par les résultats des politiques de redistribution. L'ensemble de ces modèles a comme propriété que le résultat du vote sur les transferts est stationnaire. Tout changement dans le résultat du vote ne peut provenir que d'une modification de l'un des paramètres pertinents du modèle telle que la distribution des salaires. Notre chapitre complète donc cette littérature sur les deux

[18]Source : Observatoire des Retraites. Ce taux de taxe est sensiblement équivalent pour les cadres et non-cadres.

points suivants : (1) nous endogénéisons la structure de la population et (2) nous étudions la dynamique des transferts verticaux purs. Pour effectuer notre analyse, nous utilisons un modèle à la Meltzer et Richard (1981) dans lequel nous supposerons pour simplifier que l'offre de travail est déterminée de façon exogène et dans lequel il n'existe que deux types d'agents : des agents éduqués et des non-éduqués. Contrairement à Meltzer et Richard (1981), nous supposons que les agents choisissent par leur décision d'éducation, leur position sociale. Les agents votent à chaque période sur la taille des transferts verticaux purs. Compte tenu de la simplicité de notre modèle, les agents non-éduqués seront favorables à des transferts très importants tandis que les agents éduqués souhaiteront limiter, autant que faire se peut, l'importance de ces transferts. Ainsi, si la majorité de la population est éduquée, alors la taille du système de transferts sera faible. Inversement, un système de transferts très développé sera mis en place si la population non-éduquée est majoritaire. L'anticipation de la taille des transferts, et donc du résultat du vote de la date t, a un impact distorsif sur les décisions d'éducation des agents effectuées à la date $t-1$. En effet, la redistribution a un impact négatif sur les décisions d'éducation. Ce chapitre complète la littérature existante puisqu'il permet d'étudier la dynamique conjointe de la structure de la population et de la redistribution des ressources. Nous montrons que différents scénarios peuvent apparaître dans lesquels les propriétés de prophéties auto-réalisatrices exercent un rôle déterminant. Notre modèle autorise l'apparition de tels phénomènes puisque les anticipations sur le résultat du vote ont une influence sur la structure de la population, et donc sur le résultat effectif du vote. Nous montrons qu'une économie peut converger à long-terme soit vers un système de transferts très développé ou bien vers un système de transferts peu développé. Cependant, la principale originalité de ce chapitre concerne l'influence des prophéties auto-réalisatrices sur la dynamique de l'économie. Ainsi, le passage vers une faible redistribution peut être accéléré par les anticipations des agents. Par exemple, si l'incitation à devenir éduqué est suffisante[19], et si les agents anticipent que le taux de taxe sera faible, alors nombre d'agents vont s'éduquer, et la population qualifiée sera effectivement majoritaire à la période suivante. Inversement, si les agents anticipent que le taux de taxe sera élevé, alors peu d'agents décideront de s'éduquer et le taux de taxe élevé sera effectivement choisi. Enfin, il est possible de montrer qu'un système de transferts peut ne pas être stationnaire à long-terme, dans le sens où une économie peut passer d'un système de transferts à un autre en fonction des anticipations des agents. Ce chapitre il-

[19] Nous expliquerons que cela provient du différentiel de salaire qui est une variable dynamique.

lustre le fait que la dynamique des systèmes redistributifs est loin d'être triviale puisqu'elle ne dépend pas seulement d'éléments objectifs tels que la distribution des salaires, mais elle dépend aussi d'éléments subjectifs liés aux croyances et aux anticipations des agents. Ce résultat complète la littérature initiée par Alesina et Angeletos (2005) concernant le rôle des croyances pour la détermination des transferts verticaux. Ces derniers montrent que les croyances dans le degré *fairness* d'une économie influence le degré souhaité de redistribution verticale.

iii. Le Cas de la Redistribution Horizontale et Verticale

Dans le chapitre précédent, nous avons étudié la détermination positive d'un système de transferts assurant une redistribution verticale des ressources. La demande de transferts était alors déterminée de façon endogène puisque la structure des votants dépendait des choix d'occupation en matière d'éducation. L'idée **de ce troisième chapitre** est que les pensions de retraites ne sont qu'un élément d'une politique redistributive plus globale qui inclut également des éléments de redistribution verticale pure intra-générationnelle. Nous supposons dans ce chapitre que l'Etat prélève une taxe sur les salaires en vue de financer un système de retraite par répartition mais aussi un système de transferts intra-générationnels. De plus, nous utilisons une économie à la Yaari (1965) dans laquelle les agents vivent leur deux premières périodes de vie de façon certaine, mais leur survie jusqu'en troisième période dépend de la réalisation d'un événement aléatoire. Lors de leur première période de vie les agents choisissent de s'éduquer ou non en fonction du différentiel de salaire qui prévaudra à la période suivante. Nous supposons donc implicitement (pour éviter les propriétés d'indétermination obtenues dans le chapitre précédent) que la taille des systèmes de transferts n'a pas d'impact direct sur les décisions d'éducation. De plus, le différentiel de salaire entre les agents éduqués et non-éduqués est supposé être une fonction croissante du niveau d'accumulation du capital (Acemoglu 2002, Krusell *et al.* 2000). Une fois devenus travailleurs (deuxième période de vie), les agents participent à la vie politique. De plus, ils décident à la fin de leur seconde période de vie, du montant qu'ils épargnent en vue de financer leur consommation de troisième période de vie, i.e. la période de retraite. Le montant de l'épargne agrégée dépend négativement de la taille du système redistributif. La réduction de l'accumulation du capital qui en résulte aura une influence négative sur le différentiel de salaire entre les agents éduqués et non-éduqués, et donc sur les décisions d'éducation.

Chaque génération de travailleurs est de taille plus importante que la génération de retraités vivant à cette même période du fait du risque de survie existant à la fin de la deuxième période de vie. Lorsque les agents travaillent, ils prennent comme donné le taux de taxe dont ils bénéficieront lorsqu'ils seront vieux. Ainsi, la seule incitation pour les travailleurs à soutenir un tel système redistributif est qu'il redistribue verticalement une partie des sommes collectées aux jeunes de cette même période. Contrairement au chapitre précédent, les agents non-éduqués ne vont pas nécessairement voter en faveur de la redistribution la plus large, puisqu'une partie des ressources publiques est affectée au financement des retraites. Plus précisément, ils voteront en faveur du taux de taxe le plus grand si le différentiel entre le salaire de cette catégorie de la population et le salaire moyen de l'économie est suffisamment grand. Les vieux quant à eux votent à toutes les périodes pour le système de transferts le plus large puisqu'ils en bénéficient sans en payer le coût.

Nous montrons qu'il est possible qu'apparaisse une dynamique en trois temps pour l'accumulation du capital mais aussi pour la taille du système de redistribution. Tout d'abord, tant que le niveau d'accumulation du capital est faible, et que le différentiel de salaire entre les agents éduqués et non-éduqués est également faible, peu d'agents décident de s'éduquer. Par conséquent, le salaire moyen de l'économie n'est pas encore substantiellement différent du salaire des agents non-éduqués. Pour ces derniers, le coût du système lié à la taxe sur les salaires est supérieur à ses bénéfices. Les éduqués et les non éduqués forment donc une coalition et votent en faveur d'un système redistributif peu développé. Puis, une fois que l'accumulation du capital est plus importante et que le salaire moyen augmente, les non-éduqués et les vieux forment une coalition et votent en faveur d'un système redistributif important. Enfin, l'accumulation du capital, et donc le différentiel de salaire, est telle qu'une très grande partie des agents décide de s'éduquer. Cette fraction de la population est si large qu'elle obtient la majorité et un système de transferts plus faible est adopté.

Ce chapitre complète la littérature mentionnée dans le chapitre 2 puisque nous incluons à la fois des éléments de redistribution intra-générationnelle pure et des éléments de redistribution inter-générationnelle. Cette modélisation n'est pas nouvelle en soi puisqu'elle a notamment été utilisée par Razin *et al.* (2002a, 2002b, 2004) ou bien par Galasso et Profeta (2007). En revanche, l'élément nouveau de ce chapitre concerne la dynamique endogène de la structure de la population. Ainsi, dans notre modèle, la dynamique des systèmes de retraite doit être entendue comme un élément d'une dynamique plus globale des systèmes redistributifs; ces deux dynamiques étant influencées par l'évolution de la structure de la population. Ce dernier élément constitue également un nouvel angle de réflexion en ce qui concerne l'analyse positive des systèmes de retraite (Galasso et Profeta 2002).

Dans la deuxième partie de cette thèse nous exploitons les effets macro-économiques en terme d'accumulation du capital, de bien-être, de distribution des salaires et de richesse, d'une modification de la redistributivité instantanée des systèmes de retraite par répartition. La redistributivité instantanée dans cette partie est mesurée par le degré d'indexation des pensions sur les salaires, i.e. par le paramètre λ que nous avons défini précédemment.

Cette question de l'impact de la redistributivité instantanée est importante, d'une part, parce les pays ont adopté des structures très hétérogènes et que l'on ne s'est pas interrogé sur les conséquences à long-terme de ces choix et des éventuels coûts de la transition vers un système plus efficace à long-terme; d'autre part, parce qu'elle n'est pratiquement pas évoquée dans la littérature économique.

Le tableau 2 montre que les taux de remplacement selon les niveaux de salaire sont très hétérogènes en fonction des pays. Ce tableau reporte les taux de remplacement pour les agents dont le salaire est égal à la moitié du salaire moyen de l'économie, est égal au salaire moyen, et est égal au double du salaire moyen[20]. Comme nous l'avions mentionné ci-dessus, une forte redistributivité instantanée implique que le taux de remplacement est une fonction largement décroissante du niveau de salaire. Ainsi, Casamatta *et al.* (2000) en concluent que le Canada, les Pays-Bas et la Nouvelle-Zélande ont des systèmes de retraite Beveridgiens. La France, l'Allemagne et l'Italie ont quant à eux des systèmes plutôt Bismarckiens. Enfin, le Japon, le Royaume-Uni et les Etats-Unis ont opté pour des structures mixtes.

L'un des premiers articles s'intéressant aux conséquences macro-économiques mentionnées ci-dessus est celui de Sommacal (2006). Sommacal utilise une modèle à générations imbriquées dans lequel l'offre de travail de première période est endogène. Les agents diffèrent par leur niveau de productivité. Pour simplifier l'auteur suppose qu'il n'existe que deux groupes d'agents : des éduqués et des non-éduqués, dont la taille respective est

[20]Sauf pour l'Italie pour laquelle ces ratios sont calculés pour des agents dont le salaire ne représente que 25% du salaire moyen de l'économie; pour des agents disposant du salaire moyen; et pour les agents dont le salaire est trois fois supérieur au salaire moyen.

	Taux de remplacement		
	$1/2 \times \bar{w}$	$\bar{w}$	$2 \times \bar{w}$
France	84	84	73
Allemagne	76	72	75
Italie	$103(1/4 \times \bar{w})$	90	$84\ (3 \times \bar{w})$
Japon	77	56	43
RU	72	50	35
Etats-Unis	65	55	32
Canada	76	44	25
Pays-Bas	73	43	25
Nouvelle-Zélande	75	38	19

Table 2: Taux de remplacement des systèmes de retraite publics. *Source* : Disney et Johnson (2001).

donnée. La redistributivité instantanée n'a donc pas d'impact sur les choix en matière d'éducation des agents. Les agents cotisent pour un système de retraite en payant une taxe proportionnelle à leur salaire. De plus, les pensions versées aux agents dépendent pour partie du montant cotisé, tandis que l'autre part versée est forfaitaire. En utilisant la méthode de résolution numérique Dynare (voir Juillard 1996), l'auteur montre qu'un accroissement de la part forfaitaire des pensions a un impact négatif sur le niveau de production de l'économie, ainsi que sur le niveau de richesse et l'offre de travail de tous les agents de l'économie. Gorski *et al.* (2007) se sont quant à eux intéressés aux conséquences en terme de choix d'éducation de la redistribution instantanée des systèmes de retraite. Ces auteurs utilisent une structure à générations imbriquées dans laquelle les agents choisissent de s'éduquer ou non compte tenu de leur coût idiosyncratique associé à la période d'apprentissage. Le principal résultat de cet article est qu'un accroissement de la part Bismarckienne des systèmes de retraite par répartition conduit à une réduction de la part de la population qui décide de s'éduquer. L'idée défendue est que les agents éduqués vivent plus longtemps que les non-éduqués, et donc bénéficient de la partie Beveridgienne pendant une plus longue période. Par conséquent, réduire cette part revient à pénaliser la population éduquée, ce qui incite les agents à moins s'éduquer[21].

[21]Notons néanmoins que nous suspectons que ce résultat ne repose que sur l'hypothèse très restrictive de leur article selon laquelle le salaire de l'individu seuil net du coût de l'éducation doit être le même que

CSP	Probabilité de décéder entre 35 et 65 ans (en %)	Espérance de vie à 35 ans (en années)
Cadres, Professions Libérales	13	44.5
Agriculteurs exploitants	15.5	43
Professions Intermédiaires	17	42
Artisans, Commerçants et Chefs d'entreprise	18.5	41.5
Employés	23	40
Ouvriers	26	38

Table 3: Mortalité des hommes suivant la CSP sur la période 1982-1996. *Source* : Mesrine (1999).

Le quatrième et le cinquième chapitre visent à compléter cette littérature en mettant en évidence le rôle clé que peuvent jouer les inégalités face à la mort lorsque l'on étudie l'impact du caractère redistributif des systèmes de retraite par répartition. L'intuition est la suivante : augmenter la redistribution d'un système de retraite (une baisse de λ) incite les agents les plus éduqués à épargner plus, et les moins éduqués à réduire leur épargne. Or, puisque les agents les plus éduqués sont également ceux dont la durée de vie est la plus longue, la hausse de leur épargne fait plus que compenser la baisse de celle des agents les moins éduqués. Mais, avant de détailler les mécanismes de nos chapitres, rappelons dans quelle mesure il existe des inégalités sociales face à la mort.

Il semble qu'il y ait un large consensus autour de l'idée que le niveau de salaire, le niveau d'éducation ou la catégorie socio-professionnelle ont une influence significative sur l'espérance de vie des individus. De très nombreuses études existent à ce sujet (voir Adams *et al.* (2003) pour une revue de la littérature). Prenons l'exemple de la France. Mesrine (1999) établit qu'il semble exister un lien très fort entre la catégorie socio-professionnelle (CSP) et l'espérance de vie (tableau 3). Ainsi, un ouvrier a une espérance de vie à 35 ans de plus de six ans inférieure à celle d'un cadre.

l'agent soit éduqué ou non. En effet, la non-ambiguïté de leur résultat pose problème puisqu'en l'absence d'inégalités face à la mort, une plus faible redistribution devrait conduire les agents à plus s'éduquer. Or, leur modèle ne met pas en évidence l'existence d'un seuil d'inégalités face à la mort à partir duquel leur résultat est vérifié.

Dans cette partie (et dans le reste de cette thèse), nous allons considérer que les agents éduqués et non éduqués diffèrent par leur espérance de vie, mais que le niveau et le différentiel d'espérance de vie sont donnés. **Dans le quatrième chapitre**, nous examinons l'impact d'un accroissement de la redistributivité des systèmes de retraites dans le cadre d'un modèle à générations imbriquées. En vue d'obtenir des résultats analytiques, nous supposons que les agents offrent leur travail de façon inélastique durant leur première période de vie et sont à la retraite durant leur deuxième période de vie. Cette hypothèse n'est pas sans conséquences sur les résultats qualitatifs de notre modèle mais cela va nous permettre d'obtenir un premier socle théorique permettant de mettre en évidence les mécanismes de base. L'endogénéisation de l'offre de travail au sens large (offre de travail intensive et choix d'âge de départ en retraite) est laissée pour les recherches à venir. De plus, les agents diffèrent par leur niveau de productivité. Cette dernière a une influence positive sur leur espérance de vie. En première période, les agents décident du montant qu'ils souhaitent épargner pour leurs vieux jours sachant que le marché du crédit est parfait, i.e. qu'il n'existe pas de contraintes de liquidité. Cette décision d'épargne tient compte de l'existence d'un système de retraite par répartition, et notamment des pensions qui lui sont associées. Nous montrons que les effets liés à la redistributivité du système de retraite diffèrent selon le mode d'organisation adopté. Si le système de retraite admet une structure à générosité-définie, i.e. si le taux de taxe est la variable d'ajustement, alors un accroissement de la redistributivité du système de retraite augmente le montant de l'épargne agrégée car la hausse de l'épargne des qualifiés fait plus que compenser la diminution de l'épargne des agents peu qualifiés. L'accumulation du capital est alors plus importante et les salaires augmentent. A long-terme, si cet effet est suffisamment grand, il est possible que la richesse, et donc le bien-être, de tous les agents de l'économie augmente. Toutefois, si le taux de remplacement (ν) est la variable d'ajustement, alors seuls les agents non-éduqués bénéficient de cette politique redistributive. La différence avec le cas précédent tient en ce que la hausse de la redistributivité permet d'augmenter le taux de remplacement ν, ce qui bénéficie essentiellement aux agents dont la durée de vie est la plus longue, i.e. les plus éduqués. Ceux-ci sont donc incités à réduire leur épargne, ce qui réduit l'accumulation du capital et donc le niveau des salaires.

Dans le cinquième chapitre, nous étudions également l'impact d'une modification de la redistribution instantanée d'un système de retraite par répartition mais dans un cadre plus général. Nous utilisons notamment une fonction de production dont la substituabilité entre le travail et le capital dépend du niveau de qualification du travail (hypothèse de "capital-skill complementarity"). Les agents diffèrent toujours par leur niveau de productivité et par leur espérance de vie. Le système de retraite est organisé de façon à ce que le taux de remplacement du salaire moyen de l'économie reste inchangé. Nous montrons analytiquement qu'un accroissement de la redistributivité instantanée a un impact positif sur l'accumulation du capital et donc sur les inégalités de salaire. Puis, nous calibrons notre modèle et nous montrons qu'il est possible qu'une telle politique bénéficie à tous les agents de l'économie à condition que les inégalités d'espérance de vie soient suffisamment importantes. Ce chapitre est complémentaire au précédent en ce sens qu'il généralise la démonstration de certains de ses résultats, et donne une nouvelle implication de ceux-ci en terme d'inégalités de salaire. Enfin, il permet de mettre plus facilement en évidence le rôle clé que jouent les inégalités face à la mort dans notre modèle.

Dans la troisième partie de cette thèse, nous étudions la notion de redistribution de long-terme des systèmes de retraite par répartition. Plus précisément, nous montrons en quoi la prise en compte des inégalités face à la mort modifie les propriétés de la progressivité de long-terme des systèmes de retraite. Dans ce cas, comme nous l'avons mentionné précédemment, la progressivité instantanée n'assure pas nécessairement une progressivité de long-terme. La progressivité instantanée est mesurée par le degré d'indexation des pensions sur les salaires. Plus cette indexation est importante (α tend vers 1), plus la redistributivité instantanée est faible. La progressivité de long-terme, quant à elle, est mesurée par la contribution nette au système de retraite. Notre objectif est d'étudier cette progressivité de long-terme pour chaque groupe d'agents caractérisé par un niveau de salaire. Par conséquent, les contributions nettes sont définies comme la différence entre le montant total payé par un groupe d'agents et ce que ce même groupe reçoit du système de retraite pour une période donnée (Drouhin 2001a). On montre que le taux d'actualisation utilisé (δ) est égal au taux de croissance de la population. Cette formulation fournit une information sur la base d'un groupe d'agents. Mais il est possible de normaliser ces contributions en fonction de la taille de chaque groupe pour obtenir la contribution nette individualisée au système de retraite. En effet, l'information précédente pouvait être biaisée par la taille des groupes d'agents. La normalisation fournit une information supplémentaire qui permet de déterminer, à l'échelle d'un individu, qui bénéficie réellement du système de retraite, i.e. les personnes dont la contribution nette est négative; et la valeur de cette contribution nette. Mais la principale information porte sur l'identité de l'agent qui bénéficie le plus du système de retraite.

Puisque nous souhaitons mettre l'accent sur le rôle clé joué par les inégalités face à la mort, nous allons considérer que les agents diffèrent par leur niveau de salaire et que ce dernier a une influence positive sur l'espérance de vie des agents. De plus, nous allons considérer que la durée d'activité est identique pour tous les agents et qu'ils commencent à travailler à la même date. La prise en compte du temps d'éducation et de l'âge de départ en retraite, tous deux étant des fonctions croissantes du niveau de salaire, ajouterait d'autres facteurs affectant la redistributivité de long-terme des systèmes de retraite. Tandis que la durée de l'éducation réduirait la progressivité du système de retraite, l'âge de départ en retraite aurait l'effet inverse.

Les inégalités face à la mort semblent être significatives comme nous l'avons montré précédemment. Elles sont donc susceptibles de modifier considérablement la redistributivité de long-terme des systèmes de retraite par répartition. Mitchell et Zeldes (1996) mentionnent, sans le montrer, que les inégalités face à la mort impliquent que le système de retraite américain est bien moins progressif qu'il n'y paraît. Coronado *et al.* (2000) et Liebman (2001) confirment cette intuition en utilisant des modèles de micro-simulation calibrés sur l'économie américaine. Nelissen (1999) obtient les mêmes résultats qualitatifs pour la Hollande. La littérature théorique a ensuite exploité ces résultats. Le modèle de Gorski *et al.* (2007), que nous avons déjà présenté, en est un exemple. De plus, Borck (2007) a montré, dans le cadre d'un modèle d'économie politique à la Casamatta *et al.* (2000), que si le niveau de salaire était corrélé positivement avec l'espérance de vie, alors la coalition votant en faveur de l'instauration d'un système de retraite pouvait être constituée des agents les plus pauvres et des agents les plus riches. Les agents les plus pauvres votent en faveur du système de retraite car celui-ci permet une redistribution instantanée des ressources; tandis que les agents les plus riches soutiennent le système de retraite du fait de sa régressivité de long-terme.

La littérature peut donc être décomposée en deux parties. La première partie de la littérature cherche à montrer empiriquement que les inégalités face à la mort modifient la progressivité de long-terme des systèmes de retraite. La deuxième partie de cette littérature cherche à exploiter de telles propriétés que ce soit pour les décisions politiques, ou bien pour les décisions d'éducation. Cette partie va compléter cette littérature sur deux points.

Tout d'abord, la classification que nous avons donnée montre qu'il manque un élément à cette littérature, à savoir l'*étude théorique* de l'impact des inégalités face à la mort sur la redistributivité de long-terme des systèmes de retraite par répartition. Ce point permettra de distinguer simplement la progressivité instantanée de celle de long-terme, et de montrer que les résultats de Borck (2007) par exemple, découlent directement de cette distinction et de quelques propriétés élémentaires de la progressivité des systèmes de retraite. Le deuxième apport de cette partie est de montrer que du fait de la possible régressivité des systèmes de retraite, une hausse de la générosité moyenne d'un système de retraite par répartition (hausse de ν) a un impact ambigu sur la part de la population éduquée.

Dans le sixième chapitre, nous analysons les conditions dans lesquelles il y a progressivité instantanée et progressivité de long-terme des systèmes de retraite par répartition. Pour cela, nous utilisons le cadre d'une petite économie ouverte dans laquelle les agents vivent deux périodes et se distinguent par leur niveau de salaire. Chaque agent est doté aléatoirement à sa naissance d'un niveau de productivité. Nous supposerons que la durée de vie des agents dépend positivement de leur niveau de salaire. Durant leur première période de vie, les agents travaillent et offrent leur travail de façon inélastique, tandis que durant leur deuxième période de vie les agents sont à la retraite et perçoivent une pension de la part du Gouvernement. Ce dernier finance ces pensions en appliquant une taxe proportionnelle sur les salaires. Cette taxe est identique pour tous les agents. Les pensions par unité de temps versées aux agents de l'économie dépendent de la générosité moyenne de l'économie (ν), ainsi que du degré d'indexation sur les salaires (λ). Ce degré d'indexation est une mesure de la progressivité instantanée du système de retraite. Moins les pensions sont indexées sur les salaires, et plus le système est redistributif. La progressivité de long-terme, quant à elle, sera mesurée par la contribution nette d'un groupe d'agents, de salaire identique, au système de retraite. En vue de souligner le rôle clé que jouent les inégalités face à la mort pour l'étude de la redistributivité de long-terme, nous effectuons tout d'abord notre analyse en considérant que tous les agents ont la même durée de vie. On montre alors que la redistributivité instantanée assure une redistribution de long-terme des ressources[22]. Le système est progressif à long-terme dans le sens défini précédemment, à savoir que les contributions nette individualisées sont une fonction croissante du niveau de

[22] Ainsi, le modèle de Casamatta *et al.* (2000) peut être lu comme un cas particulier du modèle de Borck (2007) dans lequel l'espérance de vie est la même dans toute la population.

salaire. Puis, nous introduisons les inégalités face à la mort. Nous étudions tout d'abord les deux cas extrêmes concernant la redistributivité instantanée. Le cas où tous les agents ont la même pension (cas Beveridgien pur), et le cas où les pensions ne dépendent que du salaire d'activité (cas Bismarckien pur). Dans le cas Beveridgien, on montre que les agents les plus pauvres continuent de bénéficier d'une contribution nette négative mais le système peut ne plus être parfaitement progressif à long-terme. En effet, notre définition de la progressivité de long-terme implique que l'agent le plus pauvre doit bénéficier le plus du système de transferts, ce qui n'est plus assuré ici. Dans le cas Bismarckien, on montre que le système de retraite est régressif, dans le sens où il redistribue les ressources des agents les plus pauvres vers les agents les plus riches, et que plus le niveau de salaire est important et plus on bénéficie des transferts. Ce résultat est assez intuitif puisque le caractère Bismarckien du système de retraite implique qu'il n'y a pas de redistribution instantanée des ressources, tandis que les inégalités face à la mort impliquent une redistribution des ressources vers les agents dont la durée de vie est la plus grande. Dans le cas où le système de retraite est mixte ($\lambda \in (0, 1)$), il est possible de montrer que les résultats précédents peuvent être généralisés aux cas où les systèmes sont très largement Beveridgiens, ou très largement Bismarckiens. Dans le cas intermédiaire il n'est pas possible d'obtenir de résultats analytiques clairs, c'est pourquoi nous fournissons un petit exercice de calibration sur le cas français qui n'a qu'une vertus illustrative et ne saurait représenter l'entière complexité du système de retraite français. On montre que les propriétés de régressivité du système français apparaissent pour des niveaux de corrélations très significatifs entre l'espérance de vie et le niveau de salaire. De manière générale, le différentiel d'espérance de vie réduit considérablement les propriétés redistributives du système de retraite. Enfin, il est possible d'avoir une configuration dans laquelle les agents les plus pauvres et les agents les plus riches bénéficient le plus du système de retraite. C'est cette propriété qui a notamment été utilisée par Borck (2007).

Dans le septième et dernier chapitre, nous exploitons nos résultats du chapitre précédent pour étudier l'impact de la générosité moyenne du système de retraite (ν) sur la part de la population éduquée. Ce chapitre se distingue de l'article de Gorski *et al.* (2007) en terme d'objectif, puisque ceux-ci cherchent à analyser l'impact de la redistributivité instantanée sur la part de la population éduquée. De plus, notre modèle nous semble être plus micro-fondé puisque dans leur article les agents ne prennent pas en compte l'impact direct des inégalités face à la mort sur l'utilité des agents. Enfin, notre modèle ne repose pas sur l'hypothèse restrictive que pour l'agent seuil, i.e. l'agent qui est indifférent entre s'éduquer ou non, les salaires nets du coût de l'éducation de première période de vie doivent être égaux, que cet agent décide de s'éduquer ou non[23].

Docquier et Paddison (2003) montrent qu'un accroissement de la générosité des systèmes de retraite a un impact négatif sur la part de la population éduquée. Deux canaux de transmission peuvent expliquer ceci : (1) la hausse de la générosité diminue le montant épargné et donc l'accumulation du capital, ce qui augmente le taux d'intérêt. Ceci réduit la valeur actualisée des rendements de l'éducation. (2) Dans leur modèle il n'y a pas d'inégalités face à la mort, donc la redistributivité instantanée assure la redistributivité de long-terme. *Ceteris paribus*, une hausse de la taille du système redistributif réduit les inégalités de richesse et donc l'incitation à s'éduquer. Le Garrec (2005) obtient les mêmes résultats qualitatifs. Ce chapitre va essentiellement se concentrer sur le deuxième canal de transmission entre la générosité moyenne du système de retraite et la part de la population éduquée, i.e. quant au rôle joué par la progressivité du système de retraite. Compte tenu des résultats du chapitre précédent, nous discutons de l'impact des inégalités face à la mort sur la progressivité du système de retraite.

Dans ce chapitre, nous analysons le cas d'une petite économie ouverte dans laquelle la population est composée d'agents éduqués et d'agents non-éduqués. Les agents décident de s'éduquer ou non en fonction de la comparaison des niveaux d'utilité indirecte retirée dans chacune des situations, sachant que la décision d'éducation implique un coût en terme de loisir. Cette utilité indirecte dépend des niveaux de richesse obtenus par les agents. En modifiant la générosité moyenne du système de retraite (ν), nous modifions les niveaux de richesse obtenus par les agents éduqués et les non-éduqués, ce qui a un impact sur la part de la population éduquée. Nous montrons que tant que le système de retraite est progressif, alors un accroissement de ν diminue la part de la population éduquée comme dans Docquier et Paddison (2003) et Le Garrec (2005). Un système de retraite est progressif si la redistributivité instantanée est suffisamment grande pour compenser la régressivité induite par les inégalités face à la mort. En revanche, nous montrons qu'il est possible qu'une hausse de ν ait un impact positif sur la part de la population éduquée si le système

[23]Notre structure plus générale montre même que leur résultat n'a que peu de chances d'apparaître pour une calibration standard du modèle. Ceci contraste très sensiblement avec leurs résultats qui ne sont pas ambigus compte tenu de leurs hypothèses restrictives.

de retraite est régressif à long-terme, i.e. si les inégalités face à la mort font plus que compenser la redistributivité instantanée.

Part I

The Endogenous Determination of the size of Redistributive Policies

Chapter 1

The Impact of the Ageing of the Population on the Size of Pension Systems

1.1 Introduction

For a few years, some countries have considered solutions which could be adopted to solve the problem of the fiscal burden of pension systems. One of the solutions is to increase the size of the working population by fertility policies or by immigration policies. These policies aim at limiting the impact of the ageing population on growth. A first question is: are these policies relevant whatever the structure of pension systems? A pension system can have a defined-benefit or a defined-contribution structure. We define a "defined-benefit" pension system as a pension system in which the tax rate is the adjustment variable while the replacement rate is given. We call a "defined-contribution" pension system a pension system in which the tax rate (the fiscal burden) is fixed while the replacement rate (the generosity of pension systems) is the adjustment variable.

Section 3 of this chapter analyzes the impact of an ageing population on the decentralized value of capital accumulation for defined-benefit and defined-contribution pension systems. We notably show that the fertility policies are relevant only for defined-benefit pension systems and for a high level of life expectancy. We also show that life expectancy does not necessarily have a positive impact on capital accumulation because of the fiscal burden of pension systems.

Moreover, some countries have begun to question the generosity of their pension systems because of the fiscal burden they imply for workers. A part of the theoretical economy literature tried to explain this phenomenon. We give the main arguments of this literature in section 2 of this chapter. For example, it has been explained in the political economy literature that there can be an inverted U-shaped relationship between life expectancy and the equilibrium tax rate[1] (Uebelmesser 2004, Persson and Tabellini 2000). Indeed, an ageing population implies that the age of the median voter increases. However, because of the large number of elderly people, the return of pension systems is small, which has a negative impact on the chosen tax rate. Consequently, there are two opposite effects and we cannot conclude as for the impact of life expectancy on the tax rate.

Another natural argument would concern the optimality of such an evolution. That is what we do in section 3. As in Atkinson[2] (2000), we show that it can be optimal for a

[1] The Browning's paper (1975) also studied the impact of the ageing population in a political economy environment.

[2] Atkinson (2000) uses a small open economy model in which Governments differ by their inequalities aversion.

country to have an inverted U-shaped relationship between life expectancy and the generosity of pension systems (the replacement rate), while the tax rate remains an increasing function of life expectancy.

1.2 The Political Economy Approach

In this section, we show, in a political economy environment, that an ageing population has an ambiguous impact on the size of the pension system.

We assume a small open economy in which agents only differ by their age. At each date ϵ a new generation is born. The size of each cohort is constant over time and normalized to 1.

1.2.1 Consumers

Each agent born at date t works until he reaches the age $t+R$. Consequently, R denotes the age of retirement, which is assumed to be fixed exogenously and constant over time. Then, an agent is in retirement until he dies. To simplify, we assume that every agent knows with certainty his date of death, denoted by $t+T$, with $T > R > 0$. At date ϵ, let Υ_t^ϵ denote the age of an agent born at period t. We have: $\Upsilon_t^\epsilon = \epsilon - t$. As long as an agent is young ($\Upsilon_t^\epsilon < R$), he supplies his labor inelastically. Given our assumption of a small open economy, the wage level is exogenously given, constant over time, and equal to w whatever the age of individuals. Agents pay a proportional tax τ on their wages. The revenues of this tax are used to finance a Pay-As-You-Go (PAYG) pension system. Each retired agent receives a pension $\tilde{p}$ per unit of time. Denoting by $c_t(\mu)$, the consumption level at date μ of an agent born at date t, the budget constraint at date ϵ can be written:

$$\begin{aligned}\int_{\epsilon}^{t+T} c_t(\mu)e^{-r(\mu-\epsilon)}d\mu &= \int_{\epsilon}^{t+R} we^{-r(\mu-\epsilon)}d\mu + k_t^\epsilon + SSW_t^\epsilon \text{ if } 0 < \Upsilon_t^\epsilon < R \\ &= k_t^\epsilon + SSW_t^\epsilon \text{ if } T > \Upsilon_t^\epsilon > R\end{aligned}$$

with t the date of birth of the agent, ϵ the current date and k_t^ϵ the accumulated wealth of an agent born at date t evaluated at date ϵ. We assume that $k_t^t = 0$, i.e. when an agent is born, he has no wealth endowment. Furthermore, because there is no altruism in our model, we have $k_t^{t+T} = 0$. r denotes the discount rate, here represented by the interest rate which is assumed to be strictly positive.

SSW_t^ϵ denotes the social security wealth of an agent born at date t, assessed at date ϵ. It corresponds to the discounted value of costs and benefits of the social security system. The social security wealth can be written:

$$\begin{aligned} SSW_t^\epsilon &= \int_{t+R}^{t+T} \tilde{p} e^{-r(\mu-\epsilon)} d\mu - \int_{\epsilon}^{t+R} w\tau e^{-r(\mu-\epsilon)} d\mu \text{ if } 0 < \Upsilon_t^\epsilon < R \\ &= \int_{\epsilon}^{t+T} \tilde{p} e^{-r(\mu-\epsilon)} d\mu \text{ if } T > \Upsilon_t^\epsilon > R \end{aligned}$$

In the first row, the first component on the right-hand-side (RHS) is the discounted value of pensions, whereas the second component is the discounted value of taxes used to finance the social security system. The younger an agent is (the higher the distance between ϵ and $t+R$ is), the higher the cost of the social security system is.

The utility level at date ϵ of an agent born at date t can be written:

$$U_t^\epsilon = \int_{\epsilon}^{t+T} u(c_t(\mu)) e^{-\rho(\mu-\epsilon)} d\mu \tag{1.1}$$

with $u'() > 0$ and $u''() < 0$. $\rho > 0$ denotes the psychological discount rate of agents. U_t^ϵ is a strictly increasing function with respect to the consumption flow. Thus, the utility level of an agent is maximized if the resources of the agent are maximized, i.e. if the SSW[3] of the agent is maximized.

1.2.2 Government

The Government levies a tax τ on wages to finance a PAYG pension system. This tax rate is determined by a majority voting equilibrium. In that case, the median voter theorem applies iff[4] preferences with respect to τ are single-picked. The median voter is such that half of the population would prefer a lower tax rate, while the other half of the population would prefer a higher tax rate. Before determining this political equilibrium, we have to find the relationship between the tax rate and the pension level thanks to the budget constraint of the Government, which has the following form:

$$\int_{\epsilon-R}^{\epsilon} \tau w d\mu = \int_{\epsilon-T}^{\epsilon-R} \tilde{p} d\mu + \chi \int_{\epsilon-R}^{\epsilon} \frac{\tau^2}{2} w d\mu \tag{1.2}$$

[3]Social Security Wealth.

[4]In the rest of this PhD Thesis, "iff" denotes "if and only if".

with $\chi > 0$. The second term in the RHS of this equation is a cost associated with the size of the pension system[5]. It can represent two things: (1) a cost associated with the management of the pension system. The higher the tax rate is, the higher the costs are. (2) A distorsive impact of the pension system on individual choices. These distortions are not explicitly taken into account in our model. χ denotes the size of the cost of pension systems.

It implies:

$$\tilde{p} = \left(\tau - \chi\frac{\tau^2}{2}\right) w\frac{R}{T-R} \tag{1.3}$$

We assume that $\chi > 1$. It means that the marginal impact on pensions per agent per unit of time, of an increase in the tax rate remains positive as long as $0 \leq \tau < 1/\chi < 1$, and negative otherwise ($\tau > 1/\chi$).

1.2.3 The Political Equilibrium

Replacing $\tilde{p}$ in the budget constraints of agents with its expression given by equation (1.3), the social security wealth of an agent born at date t assessed at date ϵ can be written:

$$\begin{aligned} SSW_t^\epsilon &= \int_{t+R}^{t+T} \left(\tau - \chi\frac{\tau^2}{2}\right) w\frac{R}{T-R} e^{-r(\mu-\epsilon)}d\mu - \int_{\epsilon}^{t+R} w\tau e^{-r(\mu-\epsilon)}d\mu \text{ if } 0 < \Upsilon_t^\epsilon < R \\ &= \int_{\epsilon}^{t+T} \left(\tau - \chi\frac{\tau^2}{2}\right) w\frac{R}{T-R} e^{-r(\mu-\epsilon)}d\mu \text{ if } T > \Upsilon_t^\epsilon > R \end{aligned}$$

Agents maximize their social security wealth with respect to τ, under the constraint $\tau \in [0,1]$. Let us first consider the case of agents who are retired at date ϵ, i.e. agents for whom $T > \Upsilon_t^\epsilon > R$. The SSW of retired agents is a strictly concave function of τ. Retired agents choose the tax rate which maximizes the level of pensions, i.e. $\tau = 1/\chi$.

For workers, the FOC[6] has the following form:

$$(1-\chi\tau)\frac{R}{T-R}\left[e^{-rR} - e^{-rT}\right] = \left[e^{-r\Upsilon_t^\epsilon} - e^{-rR}\right] \tag{1.4}$$

We define the LHS (Left-Hand-Side) of this equation as the function $f(\tau, T)$, and the RHS as the function $g(\Upsilon_t^\epsilon)$. $f()$ is such that $f(1/\chi, T) = 0$, and $f_1(\tau, T) < 0$[7]. Furthermore, we have $g_1() < 0$.

[5]It is only used to avoid corner solutions. Perotti (1993) uses the same assumption.

[6]First-Order-Condition.

[7]$z_j()$ denotes the derivative of the function $z()$ with respect to its jth argument.

The LHS of this equation is the marginal benefit from a marginal increase in the tax rate. It depends on the old age dependency ratio (the inverse of $R/(T-R)$). A higher old age dependency ratio reduces the marginal benefit from a tax on the pension system because the marginal revenues of this tax are shared among a larger retired population. The element between brackets in the LHS represents the period during which agents perceive the pension. The more T differs from R, the longer the period during which agents will benefit from the pension system.

The RHS is the marginal cost associated with a marginal increase in the tax rate. It measures the period during which agents will pay for the pension system. The older the agents are, the less they pay for the pension system.

Proposition 1.1 *(i) At each date ϵ, an agent of age Υ_t^ϵ votes for a strictly positive (null) tax rate if and only if $f(0,T) > g(\Upsilon_t^\epsilon)$ ($f(0,T) \leq g(\Upsilon_t^\epsilon)$). (ii) There exists a threshold value $\breve{\Upsilon} \in (0,R)$ such that every agent younger (older) than $\breve{\Upsilon}$ votes for a null (positive) tax rate. $\breve{\Upsilon}$ is defined by $f(0,T) = g(\breve{\Upsilon})$. (iii) Moreover, as long as $\Upsilon_t^\epsilon < R$, then agents choose a tax rate smaller than $1/\chi$. (iv) An increase in Υ_t^ϵ has a positive impact on the chosen tax rate as long as $\Upsilon_t^\epsilon \in (\breve{\Upsilon}, R)$.*

Proof: (i) As $f_1() < 0$, the LHS and the RHS have at most one intersection point. Given the fact that $f(1/\chi, T) = 0$, these two curves intersect iff the LHS is strictly superior to the RHS for low values of τ. (ii) It can easily be proved that $f(0,T) > g(R)$. For $\Upsilon = 0$, the condition is $T(1-e^{-rR}) - R(1-e^{-rT}) > 0$. For $r = 0$ this expression is null. Thus it is sufficient to prove that the derivative of this expression with respect to r is strictly positive for $r > 0$, an we obtain $re^{-rR}(T - Re^{-r(T-R)}) > 0$. (iii) Obvious given that $f(1/\chi, T) = 0$. (iv) If the two curves intersect then the RHS is a decreasing function of Υ_t^ϵ which has a positive impact on the chosen tax rate.□

This proposition gives the properties which have to be respected by functions $f()$ and $g()$ for each agent to choose an interior solution for the tax rate. Furthermore, the older the agents are, the higher the tax rate they vote for. Indeed, old workers will pay for the pension system for only a few years but they will benefit from it for a period $T-R$. Conversely, young agents ($\Upsilon \in (0, \breve{\Upsilon})$) prefer savings to the pension system because of the differential of return. Indeed, the return on savings is the interest rate which is assumed to be strictly positive, whereas the return on the pension system is null because we assume that the population and the wages are constant. In the specific case in which $r = 0$ then $\breve{\Upsilon} = 0$ because new agents are indifferent between savings and the contribution to the pension system.

Right now, we can determine the political equilibrium of the economy. As mentioned above, the political equilibrium is determined by the preferences of the median voter. Here, the median voter theorem applies since preferences of agents are single-picked. Moreover, since the chosen tax rate is an increasing function of the age of agents, the median voter is such that every agent younger than him chooses a lower tax rate, while every agent older than him chooses a higher tax rate.

At each date ϵ, the age of the median voter is determined by:

$$\int_{\epsilon-T}^{\epsilon-\Upsilon_{median}} d\mu = \frac{1}{2}\int_{\epsilon-T}^{\epsilon} d\mu \tag{1.5}$$

which implies that:

$$\Upsilon_{median} = \frac{1}{2}T \tag{1.6}$$

The higher the length of life is, the older the median voter is. This result illustrates the debate on the development of a *gerontocracy* in developed countries. Because of the increase in life expectancy in these countries, the median voter becomes older and political decisions have to take into account this group of agents.

Proposition 1.2 *The political equilibrium of this economy (τ^{median}) is such that:*

- *If $\Upsilon_{median} < R$ then $\tau \geq 0$ iff $\Upsilon_{median} \geq \breve{\Upsilon}$*

$$(1-\chi\tau^{median})\frac{R}{T-R}\left[e^{-rR} - e^{-rT}\right] = \left[e^{-r\Upsilon_{median}} - e^{-rR}\right] \tag{1.7}$$

  and $\tau^{median} = 0$ iff $\Upsilon_{median} < \breve{\Upsilon}$

- If $R < \Upsilon_{median} < T$ then $\tau^{median} = 1/\chi$.

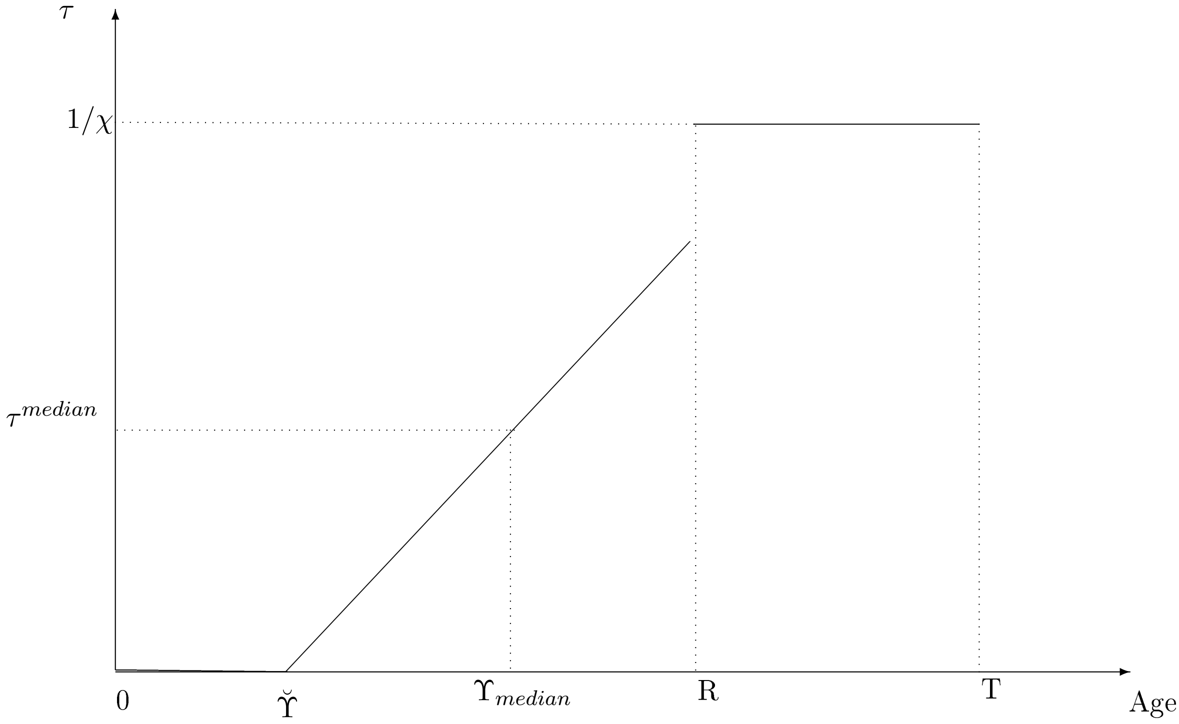

Figure 1.1: Case in which $\breve{\Upsilon} < \Upsilon_{median} < R$

This proposition can be illustrated graphically (see figure 1.1). The next step is to consider an increase in life expectancy and to study its impact on the equilibrium tax rate τ^{median}.

We only consider the interior solution ($\breve{\Upsilon} < \Upsilon_{median} < R$ and equation (1.7) is satisfied) because of its realistic form. Differentiating equation (1.7) with respect to τ^{median} and T, knowing that Υ_{median} is an increasing function of T, we have:

$$\underbrace{\chi \frac{R}{T-R}\left[e^{-rR}-e^{-rT}\right]}_{>0} d\tau^{median} =$$

$$\left\{ \underbrace{-(1-\chi\tau^{median})\frac{R}{(T-R)^2}\left[e^{-rR}-e^{-rT}\right]}_{A<0} \underbrace{+(1-\chi\tau^{median})\frac{R}{T-R}re^{-rT}}_{B>0} \underbrace{+\frac{r}{2}e^{-\frac{r}{2}T}}_{C>0} \right\} dT \tag{1.8}$$

Element A is negative. It represents the impact of an increase in life expectancy on the old-age dependency ratio. Indeed, an increase in T increases this ratio, which reduces the expected benefits from the pension system. Element B is positive. It represents the fact that an increase in T, increases the period during which agents benefit from the pension system. Element C is positive because the increase in T has a positive impact on the age of the median voter, which has a positive impact on the size of the pension system. Finally, the global impact of an increase in T on the tax rate is ambiguous.

Proposition 1.3 *An increase in life expectancy has an ambiguous impact on the tax rate at the political equilibrium (τ^{median}).*

Finally, we have to consider the impact of an increase in life expectancy (T) on pensions per unit of time. Differentiating equation (1.3) with respect to T, we have:

$$\frac{d\tilde{p}}{dT} = \underbrace{\left(1-\chi\tau^{median}(T)\right) w\frac{R}{T-R}\tau_1^{median}(T)}_{A} \underbrace{-\left(\tau^{median}-\chi\frac{(\tau^{median})^2}{2}\right) w\frac{R}{(T-R)^2}}_{B<0} \tag{1.9}$$

An increase in T has two kinds of effects on the level of pensions per unit of time. First, it has an impact on the equilibrium tax rate (element A). This effect can be positive or negative as mentioned above. Secondly, for a given level of the tax rate, an increase in T increases the share of the population who benefit from the pension system. It implies that the pension per retired agent has to be lower. The global impact of an increase in T on $\tilde{p}$

is ambiguous.

Proposition 1.4 *An increase in life expectancy (T) has an ambiguous impact on the pension level per unit of time ($\tilde{p}$).*

An interesting result of this political equilibrium model is that an increase in life expectancy does not necessarily have a positive impact on the tax rate. Moreover, even if the tax rate increases, the pension level per unit of time can decrease.

1.3 The Optimal Level of Social Security Benefits

In this section we use an overlapping generations model in a closed economy with capital accumulation *à la* Diamond. First, we study the impact of the ageing of the population on the decentralized equilibrium. Then, we analyze the relationship between the ageing of the population and the decentralization of the optimum.

1.3.1 The Model

We consider a closed economy in which at each period two generations overlap. The length of each period is normalized to 1. The size of the generation born at the beginning of the period t is denoted by N_t. We assume a constant growth rate of the population such that: $N_t = (1+n)N_{t-1}$. We assume that $n \geq 0$. We consider consumers, firms and the Government successively.

Consumers

There is no uncertainty. Each agent can live his whole first period of life. However, he lives only a fraction T of his second period of life with certainty (d'Autume 2003). His utility function has the following form[8]:

$$U_t = u(c_t) + \beta T u\left(\frac{d_{t+1}}{T}\right) \tag{1.10}$$

The utility level of an agent born at the beginning of the period t depends on the level of consumption of the two periods of life (c_t and d_{t+1}). β is the psychological discount factor.

[8]A justification of this utility function is given in the appendix of this chapter.

T has two effects on the utility level of an agent. First, an increase in T implies that each agent values his future consumption more. However, it also decreases the consumption flow per unit of time which has a negative impact on the utility level.

To simplify the formal analysis we assume in the rest of this chapter that $u(x) = ln(x)$. It simplifies the proofs without altering our message[9]. We assume that each agent supplies his work inelastically during his first period of life. The wage level is denoted by w_t. At each period, the Government levies a tax τ_t to finance a PAYG pension system. Finally, the net wage of each young agent is $w_t(1-\tau_t)$. When an agent is old (second period of life) he receives a pension as long as he is still alive, i.e. during a fraction T of his second period of life. Pensions depend on the current wage level of the economy. The pension level per unit of time of each retired agent of the period $t+1$ has the following form: $\nu_{t+1}w_{t+1}$, with ν_{t+1} the replacement of the pension system. The budget constraints of an agent born at period t are:

$$c_t = w_t(1 - \tau_t) - S_t \tag{1.11}$$

$$d_{t+1} = S_t R_{t+1} + \nu_{t+1} w_{t+1} T \tag{1.12}$$

with S_t the saving level, and R_{t+1} the interest factor. Each agent maximizes (1.10) with respect to S_t given the budget constraints (1.11) and (1.12). At equilibrium we obtain:

$$S_t = \frac{\beta T}{1 + \beta T} w_t(1 - \tau_t) - \nu_{t+1} T \frac{w_{t+1}}{R_{t+1}(1 + \beta T)} \tag{1.13}$$

An increase in the net wage of the first period of life has a positive impact on savings. *Ceteris paribus*, an increase in the level of pensions (in ν) has a negative impact on savings. Indeed, if pensions are higher, agents will consume a part of this increase in their first period of life. Consequently, they save less. Finally, for $\nu = \tau = 0$, i.e. without any pension system, an increase in the length of life (T) has a positive impact on savings because each agent values his second period utility level more.

Firms

In our economy, there is only one good which is used as *numeraire*. To simplify the formal analysis we assume that the production function of a representative firm is cobb-

[9]Using Maple, it can be shown that we obtain the same qualitative results for $u(x) = \frac{x^{1-\eta}}{1-\eta}$, for $\eta > 0$ but $\eta \neq 1$.

douglas:

$$Y_t = AK_t^{\alpha}L_t^{1-\alpha} \tag{1.14}$$

with $0 < \alpha < 1$, A the technology level ($A > 0$), K_t the physical capital level, and L_t the quantity of labour used to produce the final good. We assume a perfect competition on the final good market and on the inputs markets. It implies that the interest factor and wages are equal to their respective marginal productivity level:

$$w_t = A(1-\alpha)K_t^{\alpha}L_t^{-\alpha} = A(1-\alpha)k_t^{\alpha} \tag{1.15}$$

$$R_t = A\alpha K_t^{\alpha-1}L_t^{1-\alpha} = A\alpha k_t^{\alpha-1} \tag{1.16}$$

with k_t denoting the level of capital per worker ($k_t = K_t/L_t$).

Government

At each period t, the Government levies a tax τ_t on wages to finance an unfunded pension system. Each retired agent receives a pension indexed on the current wage level of the economy ($\nu_t w_t$). In this chapter, ν denotes the replacement rate of the pension system. These pensions are paid as long as agents are alive, i.e. during a fraction T of their second period of life. Moreover, we assume that there is no debt in our economy. Then, the budget constraint of the Government at each period is:

$$N_t\tau_t w_t = \nu_t w_t T N_{t-1} \tag{1.17}$$

which implies that:

$$\tau_t = \nu_t \frac{T}{1+n} \tag{1.18}$$

As in d'Autume (2003) we obtain that the tax rate is the product between the replacement rate and the old-age dependency ratio. Our qualitative results depend on the structure of pension systems. To simplify the study of the equilibrium, we assume that a pension system can adopt either a defined-benefit or a defined-contribution pension system. We define a defined-benefit pension system as a system in which the tax rate is the adjustment variable whereas the replacement rate (λ) is given and constant over time:

$$\tau = \nu \frac{T}{1+n} \tag{1.19}$$

Conversely, a defined-contribution pension system is such that the generosity of the pension system (ν) is the adjustment variable whereas the tax rate (τ) is given and constant over time:

$$\nu = \tau \frac{1+n}{T} \tag{1.20}$$

It implies that following an ageing of the population (a decrease in n and/or an increase in T), if the pension system has a defined-benefit structure, the tax rate increases. Conversely, if the pension system has a defined-contribution structure, the replacement rate decreases .

1.3.2 The Macroeconomic Equilibrium and its Properties

The macroeconomic equilibrium of this economy is such that the two following conditions are satisfied:

$$L_t = N_t \ , \forall t \tag{1.21}$$

$$K_{t+1} = N_t S_t \ , \forall t \tag{1.22}$$

The dynamics of the economy is summarized by these two equations. Let us now consider the case of defined-benefit and defined-contribution pension systems successively.

Defined-Benefit Pension Systems

The dynamics of the economy is obtained using equations (1.13), (1.15), (1.16), (1.19), (1.21) and (1.22). At steady state we obtain:

$$k^* = \left[\frac{\beta T A(1-\alpha)\left(1 - \nu \frac{T}{1+n}\right)}{(1+n)(1+\beta T) + \frac{1-\alpha}{\alpha}\nu T} \right]^{\frac{1}{1-\alpha}} \tag{1.23}$$

We now study the properties of the capital per worker at steady state. Let us first consider the net impact of an increase in the growth rate of the population (n).

Proposition 1.5 *If the inequality below is satisfied then there exists a threshold value of life expectancy ($\hat{T}$) such that $\partial k^*/\partial n < 0$ for $T \in (0, \hat{T})$, and $\partial k^*/\partial n > 0$ for $T \in (\hat{T}, 1)$.*

$$2(1+\beta)\frac{\nu}{1+n} + \frac{1-\alpha}{\alpha} \times \frac{\nu^2}{(1+n)^2} > (1+\beta) \tag{1.24}$$

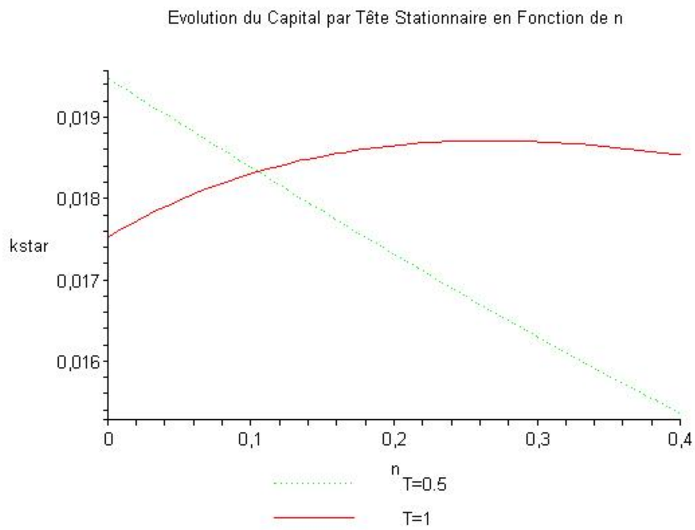

Figure 1.2: k in function of n for different values of T, for $\beta = 0.6$, $\alpha = 0.33$, $\nu = 0.5$, and $A = 1$.

Proof: Differentiating equation (1.23) with respect to n, we obtain an expression of the form: $\kappa_1 T \frac{eT^2+fT-g}{(-)^2} = \kappa_1 T \frac{P(T)}{(-)^2}$. It is obvious that $P(T \rightarrow 0) < 0$. The condition ensures that $P(T \rightarrow 1) > 0$. See Figure 1.□

The growth rate of the population (n) has two effects on capital accumulation. (1) A dilution effect: a higher growth rate of the population implies that it is more difficult to endow each additional worker with capital. This effect is standard in overlapping-generations models with capital accumulation. However, (2) an increase in n reduces the tax rate because of its negative impact on the old-age dependency ratio. It increases the net wage of the first period of life and then capital accumulation.

The main result of this proposition is that in defined-benefit pension systems the effect of the fertility rate depends on the value taken by life expectancy. It emphasizes the relevance of fertility policies when life expectancy is high.

Let us now consider the impact of an increase in life expectancy on capital accumulation. In standard theoretical literature (Drouhin 1997), it has been shown that an increase

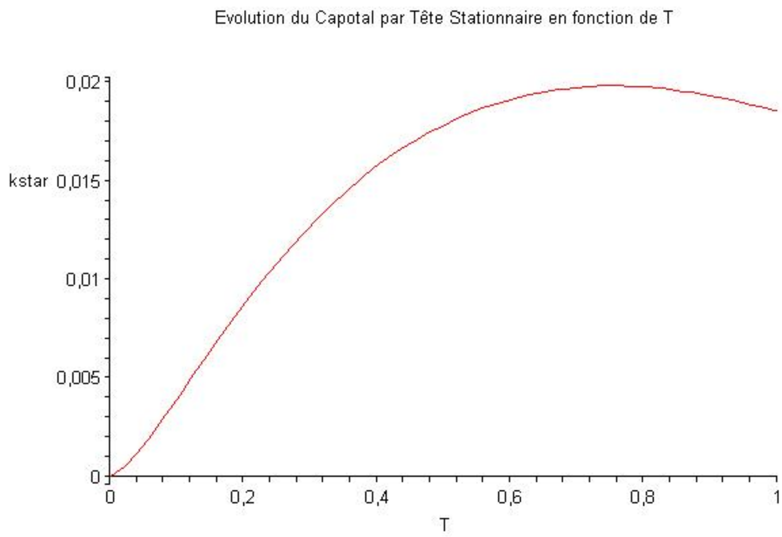

Figure 1.3: k in function of T for $n = 0.1$.

in life expectancy has a positive impact on capital accumulation. The following proposition shows that this result can be questioned for defined-benefit pension systems.

Proposition 1.6 *If*[10] $\nu > \frac{1+n}{2}$ *then there exists a threshold value* $\tilde{T}$ *such that* $\partial k^*/\partial T > 0$ *for* $T \in (0, \tilde{T})$*, and* $\partial k^*/\partial T < 0$ *for* $T \in (\tilde{T}, 1)$.

Proof: Differentiating equation (1.23) with respect to T, we obtain an expression of the form: $\frac{eT^2+fT+g}{(-)^2}$. It is obvious that that $P(0) > 0$. The condition ensures that $P(1) < 0$. See figure 2.□

An increase in the length of life (T) has two effects on capital accumulation. (1) A higher length of life implies that agents value the utility of their second period of life more. They save more to finance their second period consumption. (2) A higher length of life has a positive impact on the tax rate which reduces the disposable income of the first period of life and thus the savings of agents. As long as life expectancy is not too high, the effect on the tax rate is low and an increase in life expectancy has a positive impact on capital accumulation. However, a high life expectancy implies that the tax rate of the pension

[10]It is a sufficient condition.

system has to be high. This positive impact on the tax rate reduces savings.

Defined-Contribution Pension Systems

Let us now consider the case of defined-contribution pension systems. The steady level of capital accumulation is:

$$k_{dc}^{*} = \left[\frac{\beta TA(1-\alpha)\,(1-\tau)}{(1+n)(1+\beta T) + \frac{1-\alpha}{\alpha}(1+n)\tau} \right]^{\frac{1}{1-\alpha}} \tag{1.25}$$

We consider the impact of an increase in n and in T successively.

Proposition 1.7 *An increase in the growth rate of the population (n) has a negative impact on k_{dc}^{*}.*

Proof: Obvious given equation (1.25).□

An increase in n has two negative impacts on the steady state value of capital per worker. (1) It has a dilution effect. It is the same as the one mentioned above. (2) It increases the level of pensions received by agents when they are retired. It implies that they save less for their second period consumption. Consequently, the capital level per worker is a decreasing function of n.

Proposition 1.8 *An increase in the length of life (T) has a positive impact on k_{dc}^{*}.*

Proof: Obvious given equation (1.25).

An increase in the length of life has two positive effects on capital accumulation. (1) Agents save more because they value their second period consumption in their utility function more. (2) The level of pensions per unit of time decreases because of the defined-contribution structure. Agents save more to compensate for the decrease in their pension.

Remark: We can already emphasize the fact that the impact of demographic variables completely depends on the structure of pension systems. The adoption of defined-contribution pension systems can be a solution to avoid the negative impact of the ageing of the population on capital accumulation.

Previously, we emphasized the impact of demographic variables on capital accumulation on the equilibrium of our economy. Let us now consider the impact of these variables on the decentralization of the optimum.

1.3.3 The Optimum and its Decentralization

Because there are no externalities in our model, we only need one instrument for the decentralization of the optimum. Let us first determine the optimum (de la Croix and Michel 2002). Then, we define the value that has to be taken by the policy instruments to decentralize the optimum. The optimum is a sequence $(k_t, c_t, d_t)_{t=0..\infty}$ such that the discounted sum of the utility functions is maximized:

$$Max_{c_t,d_t,k_t} \sum_{i=-1}^{\infty} \gamma^i \left(u(c_i) + \beta T u\left(\frac{d_{i+1}}{T}\right)\right)$$
$$s.t. c_t + \frac{d_{t+1}}{1+n} = f(k_t) - (1+n)k_{t+1}$$
$$k_0 \text{ and } c_{-1} \text{ given}$$

This program is standard and we obtain that at steady state the level of capital per worker is:

$$k^{MGR} = \left(\frac{\gamma \alpha A}{1+n}\right)^{\frac{1}{1-\alpha}} \tag{1.26}$$

Lemma 1.1 *Without pension system ($\tau = \nu = 0$), (i) $k > k^{MGR}$ if and only if $\frac{\beta T}{1+\beta T}(1-\alpha) > \alpha\gamma$. (ii) If $\frac{\beta}{1+\beta}(1-\alpha) > \alpha\gamma$ then there exists a threshold value T_S such that $k < k^{MGR}$ (>) for $T < T_S$ (>).*

Proof: Using equation (1.23) or (1.25) with $\tau = \nu = 0$ and comparing it with equation (1.26), we obtain the result.□

The equilibrium value of the capital per worker is too large compared to its optimal value if life expectancy is sufficiently high and if the discount factor is small. It implies that if life expectancy is too short then savings has to be encouraged and a negative tax rate is chosen. However, once life expectancy becomes high ($T > T_S$) then savings is too large and it is optimal to implement an unfunded pension system.

Moreover, because of the form of the utility function, all the optimal sequence can be decentralized with a constant political instrument (de la Croix and Michel 2002). Comparing k and k^{MGR}, we obtain that the replacement rate which decentralizes the optimum is:

$$\nu^{opt} = \frac{1+n}{1-\alpha} \times \frac{\beta T(1-\alpha) - \alpha\gamma(1+\beta T)}{T(\gamma+\beta T)} \tag{1.27}$$

Proposition 1.9 *If the inequality below is satisfied, then there exists a threshold value $\check{T}$ such that $\partial\nu^{opt}/\partial T > 0$ for $T \in (0, \check{T})$, and $\partial\nu^{opt}/\partial T < 0$ for $T \in (\check{T}, 1)$.*

$$\alpha\gamma(2\beta+\gamma) < \beta^2(1-\alpha-\alpha\gamma) \tag{1.28}$$

Proof: Differentiating equation (1.27) with respect to T, we obtain an expression having the form: $\frac{1+n}{1-\alpha}\frac{P(T)}{(-)^2}$. We have $P(0) > 0$, but $P(1) < 0$ if and only if inequality (1.28) is satisfied.□

Inequality (1.28) is satisfied if γ is sufficiently small, i.e. if at each period t the weight of the welfare of future generations is small. It can easily be shown that $\check{T} > T_S$. The main mechanism is the following. As long as life expectancy is short, the optimal replacement rate is negative. However, as life expectancy becomes higher, agents save more and ν^{opt} increases and becomes positive once $T > T_S$. Finally, for large values of T, the fiscal burden of pension systems is so strong that the replacement rate has to decrease to decentralize the optimum.

We obtain that to decentralize the optimum, it could be appropriate to have an inverted U-shaped relationship between life expectancy and the replacement rate of pension systems.

Let us now consider the relationship between life expectancy and the optimal tax rate.

Corollary 1.1 *If $\gamma \leq 1$, then τ^{opt} is an increasing function of life expectancy (T).*

Proof: $\tau^{opt} = \nu^{opt} \frac{T}{1+n}$, and using equation (1.27) , we obtain the result.□

This result explains why it can be optimal for countries to use a non-monotonous generosity for pension systems whereas the fiscal burden of these systems goes on increasing with life expectancy. The recent evolutions of pension systems are compatible with our results.

1.4 Conclusion

This chapter presents the impact of the ageing of the population on the size of pension systems.

First, we have have shown in a political economy model that an ageing of the population does not necessarily have a positive impact on the tax rate and the replacement rate of pension systems.

Then, we have shown in a model *à la* Diamond that the impact of demographic variables completely depends on the structure of pension systems (defined-benefit or defined-contribution). Moreover, we have explained why it can be optimal for countries to have an inverted U-shaped relationship between life expectancy and the generosity of pension systems, whereas the tax rate is an increasing function of life expectancy. An interesting extension would be to do the same exercise with an endogenous retirement age.

1.5 APPENDIX

The basic model with uncertain lifetime is the Yaari's model (1965). This model can be included in overlapping generations models in the following way. Let us assume that at each period t a young generation is born. Each agent is sure to live his whole first period of life. However, each agent has only a probability ζ to live his second period of life. The utility level of an agent who lives two periods is:

$$U_t^V = u(c) + \beta u(d)$$

with c and d the first and second period consumption level. β denotes the pure psychological discount rate. Conversely, the utility level if an agent dies is:

$$U_t^D = u(c)$$

The expected utility level is[11]:

$$EU_t = \zeta U_t^V + (1-\zeta) U_t^D = u(c) + \beta\zeta u(d)$$

Then, we obtain the same utility function as in Chakraborty (2004) or Drouhin (1997), to name but just a few.

Another way to represent life expectancy is to consider that there is no uncertainty. In that case, the utility level of an agent born at period t can be written:

$$U_t = u(c) + \beta T u\left(\frac{d}{T}\right)$$

with $1+T$ the length of life of this agent. This function is notably used by Andersen (2005, 2008) or d'Autume (2003). It implies that if the length of life increases, then agents value their second period utility level more. However, it also decreases the consumption flow per unit of time.

This utility function is the result of the initial equation:

$$U_t = u(c_t) + \beta \int_0^T u(\tilde{d}_{t+1}) d\mu$$

with $\tilde{d}_{t+1}$ the second period consumption per unit of time. Note that there is a discount rate between the two periods of life but not within each of them.

[11]Let us note that the expected utility model has notably been questioned by Bommier (2006) and by Drouhin (2001b, 2006).

Finally, some authors combine the probabilistic and the non-probabilistic representation of the length of life. For example, Ponthière (2006) uses the following utility function:

$$U_t = u(c) + \beta\zeta T u\left(\frac{d}{T}\right)$$

In that case, ζ denotes the probability that an agent survives at the end of his first period of life. As for T, it denotes the length of the second period of life that agents live with certainty if they survive at the end of the first period of life.

Chapter 2

The Dynamics of the Welfare State with Endogenous Educational Choices and Pure Vertical Transfers

2.1 Introduction

In this chapter we explore the main role played by indeterminacy for the study of the dynamics of the welfare state and for the dynamics of the structure of the population in a political economy model. By "structure of the population" we mean that agents can be educated or uneducated[1]. Furthermore, as in Meltzer and Richard (1981), the welfare state only makes pure vertical transfers. Uneducated agents vote for high vertical transfers, whereas educated agents vote for low vertical transfers. We show that indeterminacy has an impact on the period from which a country chooses a smaller welfare state. Moreover, we show that, because of indeterminacy, countries can switch from a large welfare state to a smaller one, depending on the expectations of agents concerning the size of pure vertical transfers. The intuition of the indeterminacy property is the following: if the initial wage differential is low, then only a few agents decide to be educated and the majority of people (uneducated agents) vote for high vertical transfers. Then, because of the knowledge accumulation, the wage differential increases (Acemoglu 2002) and more agents decide to educate themselves. However, the educational decision also depends on the size of vertical transfers. Consequently, for a reasonably high value of the wage differential[2], if agents expect high vertical transfers, most of them remain uneducated, and usually, uneducated agents keep the majority. But if agents expect small vertical transfers, then most of them decide to educate themselves and educated agents usually get the majority, and they vote for a small welfare state. Finally, once the wage differential is very high, whatever the size of vertical transfers, the share of the educated population is so high that agents know that a small welfare state will be chosen in the next period.

Consequently, this model allows us to study the dynamics of the welfare state and to emphasize some properties related to indeterminacy. Moreover, our model suggests that the evolution of the welfare state, and notably the crisis of the welfare state, is the result of new political pressures because of the endogenous dynamics of the structure of the population. To our knowledge, the literature on the welfare state has still not considered this problem previously.

Some papers have studied the endogenous determination of redistribution using political equilibrium models (Meltzer and Richard 1981, Persson and Tabellini 2000). In these models, agents differ by their wage level which are exogenously determined and the envi-

[1]We do not consider a continuum of agents but only two groups of agents.
[2]We define more precisely in the model what the term "reasonably high" means.

ronment is static. Moreover, the labor supply of agents is endogenous. Agents vote on the size of pure vertical transfers. These transfers are financed through a tax rate on wages, and are used to finance a flat benefit, given that these social transfers have a distorsive impact on labor supply. In these models, the tax rate is determined by the agent having a wage corresponding to the median wage. Belletini and Ceroni (2007) use the same kind of model, but they assume an exogenous labor supply and agents can face a liquidity constraint. In that case, the median voter is not necessarily the agent having the median wage. Compared to these models, we add a dynamic environment. Furthermore, as in Belletini and Ceroni (2007), we do not consider the case of an endogenous labor supply to keep the model tractable.

Some other papers have studied the linkages between growth and redistribution (Alesina and Rodrik 1994, Persson and Tabellini 1994). In these second kind of papers, the tax rate bears on the return of savings, and thus the fiscal system has a distorsive impact on capital accumulation. These transfers are used to finance flat benefits for Persson and Tabellini (1994), and to finance Government spendings on productive activities for Alesina and Rodrik (1994). The common feature of these two papers is to show that more inequalities lead to a higher demand for redistribution, which increases the tax rate on capital accumulation. It implies that the growth rate of the economy decreases. These models can be considered as the first models which include both a dynamic environment and a demand for redistribution. However, in their models, the structure of the population is constant over time and they only consider the dynamics of the income distribution. Perotti (1993) incorporates an endogenous determination of the structure of the population *à la* Galor and Zeira (1993), into a model in which the size of the social security system is endogenously determined. However, he does not study the dynamics of the structure of the population[3].

Our paper completes this literature on two significant points. (1) Because of the static environment in these papers, there is not really a dynamics of the welfare state. All changes in the size of the welfare state come from modifications in some relevant parameters such as wage inequalities. Consequently, their models cannot provide indeterminacy properties, while our model seems to show that it can dramatically change the dynamics of the welfare state. (2) Because the structure of the population is exogenous, these models do not

[3]Only Bisin and Verdier (2000, 2002) incorporate a vote on public spending and an endogenous structure of the population in a cultural transmission model.

study the impact of the increase in the share of the educated population on the demand for redistribution. However, the dynamics of modern economies also has an impact on the structure of the population, which changes the political support for intra-generational transfers.

Acemoglu (2002) reported two stylized facts concerning the evolution of the structure of the population. (1) There has been a large increase in the supply of skills in the US economy over the past sixty years. (2) There has been an increase in the relative return (the wage differential) of education. Firstly, an increase in the skill premium has a positive impact on the share of the educated population. Moreover, following an increase in the share of the educated population, the stock of knowledge in the economy raises, which increases the marginal productivity of skilled agents more than that of unskilled agents. Our chapter explores these two causalities.

In this chapter, we assume an overlapping generations model in which agents live for two periods. In the first period of their life, agents decide to educate themselves or not, while in the second one agents work and supply one unit of labor inelastically. At each period t, agents vote on the size of the welfare state which can take two values to simplify the model: a high and a low tax rate. The Government pays flat benefits to workers. Thus, educated agents (with high wages) vote for the low tax rate, while uneducated agents (with low wages) vote for the high tax rate. It implies that as long as uneducated (educated) agents have the majority the high (low) tax rate is chosen. Consequently, it is the structure of the population which determines the size of the welfare state in our model. That is why we try to explain the co-evolution of the structure of the population and of the size of the welfare state. The accumulation variable in our model is knowledge capital. The share of the educated population has a positive impact on the stock of knowledge. *Ceteris paribus*, a higher knowledge capital level has a positive impact on the share of the educated population because of the increase in the wage differential between the educated and the uneducated population. Conversely, the higher the tax rate (redistribution) is, the less agents decide to educate themselves. As the tax rate of period t has an influence on the educational choices of agents born at period $t-1$, there appears indeterminacy because of what the agents expect from the result of the vote.

We distinguish four relevant cases. In the first case, the economy necessarily converges towards a small welfare state in the long run. A large welfare state only corresponds to a transitional equilibrium. However, the existence of indeterminacy has an influence on the period from which educated agents get the majority. In the second case, whatever the initial value of the knowledge capital, the economy converges towards a large welfare state. In the third case, the economy can, in the long run, adopt either a high or a low welfare state, depending on the initial value of the knowledge capital and on the beliefs of agents. In the fourth case, we show that an economy can switch from a small welfare state to a large one according to the beliefs of agents concerning the size of intra-generational transfers.

This chapter is organized as follows. Section 2 presents the basic structure of the model. In section 3, we study the equilibrium and the dynamics of the economy. Section 4 provides some concluding remarks.

2.2 The Model

We assume an overlapping generations model in which agents live for two periods[4]. During their first period of life, agents decide to educate themselves or not. It is a binary choice. If an agent decides not to educate himself, then he spends the time of his first period of life on leisure[5]. The utility level of leisure is normalized to 0. Conversely, if an agent decides to educate himself, then he has to bear a psychological cost related to the learning process. In this chapter, we assume that it represents the time which cannot be spent for leisure. When an agent is born, he is randomly endowed with an educational cost (θ). This cost reduces the utility level of agents who decide to educate themselves ($-v(\theta)$).

θ necessarily belongs to the interval $\Omega_\theta = [\underline{\theta}, \bar{\theta}]$. The density function of the variable θ is denoted by $f(\theta)$. $f(\theta)$ is such that $f(\theta) > 0$, $\forall \theta \in \Omega_\theta$, $\int_{\underline{\theta}}^{\underline{\theta}} f(\theta)d\theta = 0$ and $\int_{\underline{\theta}}^{\bar{\theta}} f(\theta)d\theta = 1$. $f(\theta)$ also denotes the fraction of the population endowed with a cost θ. $F()$ denotes the cumulative distribution function of this variable.

We make the following assumption concerning the function $v(\theta)$:

Assumption 1: $v(\bar{\theta}) > 0$, $v(\theta)$ is of class C^1 and $v'(\theta) > 0$, $\forall \theta \in \Omega_\theta$.

The first part of this assumption only means that there exists a fraction of the pop-

[4]The size of generations does not matter in our model because there are no inter-generational transfers. Consequently, the size of each generation is assumed to be constant over time, and is denoted by N.

[5]The length of each period is normalized to 1. We assume that there only exists a psychological cost in acquiring knowledge.

ulation for whom the time spent on education is a cost. Otherwise, every agent would choose to be educated. It also implies that our model does not exclude the case in which education provides more utility than leisure for a part of the population. For these people, education is a way to blossom out. It is the case iff $v(\underline{\theta}) < 0$.

The last part of this assumption implies that a longer time spent on the learning process reduces leisure time, and thus the utility level.

During their second period of life, every agent works and supplies one unit of labor inelastically. At period t, the wage level of an educated agent is denoted by w_t^e, while that of an uneducated agent is denoted by w_t^u. At each period t, each worker pays a proportional tax (τ_t) on his wage to finance the welfare state. The Government uses these fiscal revenues to finance flat transfers (b_t). All agents benefit from these transfers.

Moreover, we assume that there only exists one good in the economy which is used as a *numeraire*.

2.2.1 Consumers

The budget constraint at period t of an agent born at period $t-1$ is:

$$W_t^i = w_t^i(1-\tau_t) + b_t \tag{2.1}$$

with $i \in \{e, u\}$. At period $t-1$, an agent expects the wealth which he will receive when he works. The expectation bears on the wage level $(w_t^{i,a})$ and on the size of the welfare state $(\tau_t^a,\ b_t^a)$. Indeed, our model includes a dynamic of wages, and at each period, the size of the welfare state is endogenously chosen by the population.

Then, at period $t-1$ we have:

$$W_t^{i,a} = w_t^{i,a}(1-\tau_t^a) + b_t^a \tag{2.2}$$

Each agent is assumed to consume his wealth when he works. The utility level of an agent born at period $t-1$ can be written[6]:

$$U_t^i = -(1-I)v(\theta) + \ln\left(W_t^i\right) \tag{2.3}$$

[6] We use this utility function because of the tractability of our model.

with $I = 0$ if an agent decides to educate himself, and $I = 1$ if he decides to remain uneducated. At period $t-1$ an agent decides to educate himself iff: $U_t^{u,a} < U_t^{e,a}$. This is equivalent to:

$$\ln\left(\frac{W_t^{e,a}}{W_t^{u,a}}\right) > v(\theta) \tag{2.4}$$

The use of the log-utility function implies that educational choices depend on the wealth inequality ratio. An increase in this ratio corresponds to an increase in the opportunity cost of remaining uneducated.

The distribution of the cost θ among the population is such that, given the expectation of τ_t^a, agents for whom θ is small decide to educate themselves, whereas agents endowed with a high θ decide to remain uneducated. We denote this threshold by $\tilde{\theta}_t$.

2.2.2 Knowledge Capital

In our economy, the accumulation variable is knowledge capital. h_t denotes the knowledge capital, i.e. the available stock of knowledge at period t for each agent. In doing so, we assume that agents (educated and uneducated) only differ by their ability to transform knowledge capital into productivity, i.e. that educated and uneducated agents differ by their ability to transform the stock of knowledge into a useful information for the production process. The dynamics of the knowledge capital is assumed to have the following form:

$$h_{t+1} = \Psi(h_t, \tilde{\theta}_t) \tag{2.5}$$

with[7] $\Psi_1() > 0$ and $\Psi_2() > 0$. It implies that the current stock of knowledge has a positive impact on the future stock of knowledge[8]. Moreover, the share of the educated population at period t has a positive impact on the knowledge capital of period $t+1$ because these people do research.

The timing of the model is the following. Agents decide at period $t-1$ to educate themselves or not. Education consists in the acquisition of methods to transform and improve the stock of knowledge into useful information for the production process. As for uneducated agents, their ability level is not sufficient to use the stock of knowledge efficiently. Conversely, educated agents acquire methods to better use the available information and to improve it. At period t, agents use the methods which they acquired when they were

[7] g_i denotes the derivative of $g()$ with respect to its *ith* argument. g_{ij} denotes the derivative of g_i with respect to its *jth* argument.

[8] This is a standard assumption in the human capital literature.

young, given the current knowledge capital level. Moreover, educated agents develop new knowledge by research activities, which will only be available at period $t+1$. It implies that there need time before using this new knowledge.

The previous assumptions imply that when an agent is born at period $t-1$, he observes the current knowledge capital h_{t-1} and the share of the educated population $\tilde{\theta}_{t-1}$. Consequently, he rightly expects the value h_t. The only problem concerns the expectation of the size of the welfare state of period t.

2.2.3 Firms

We assume that the technology of firms has the following form:

$$Y_t = f(h_t, L_t^e, L_t^u) \tag{2.6}$$

with $f_i() > 0$, $\forall i \in \{1, 2, 3\}$. It means that the stock of knowledge and both kinds of labor have a positive impact on the production level. More specifically, firms take the available stock of knowledge of the current period (h_t) as given. h_t is a positive externality for each firm. However, firms choose the quantity of labor which they decide to use in the production process. We make two further assumptions on the properties of the production function: $f_{ii}() < 0$, and $f_{ij}() > 0$ if $i \neq j$. These assumptions notably imply a decreasing marginal productivity of both kinds of labor. Assuming a perfect competition on the final good market and on the inputs markets, we have:

$$w_t^e = f_2(h_t, L_t^e, L_t^u) > 0 \tag{2.7}$$

$$w_t^u = f_3(h_t, L_t^e, L_t^u) > 0 \tag{2.8}$$

Finally, we make the following assumption on the wages of agents:

Assumption 2: $w_t^e > w_t^u$ and $\dfrac{f_{21}}{f_2}h > \dfrac{f_{31}}{f_3}h$.

The first part of this assumption implies that the wage level of educated agents is higher than the wage level of uneducated agents. Thus, it is implicitly assumed that the marginal productivity of unskilled workers is always finite, even if the quantity of unskilled workers is small. Formally, it means that $f_2(h, N, 0) > f_3(h, N, 0)$, $\forall h$, i.e. even if all agents decide to educate themselves and if firms employ them, then the scarcity of unskilled workers is not sufficient for their wage level to become higher than that of skilled workers.

The second part of this assumption implies that the elasticity of the wage of skilled agents with respect to knowledge capital is higher than the elasticity of the wage of unskilled agents with respect to knowledge capital. It means that, *ceteris paribus*, the increase in the stock of knowledge increases the wage inequality ratio w^e/w^u. This assumption is based on the empirical fact that the increase in the stock of knowledge has increased the wage inequality ratio (Acemoglu 2002). In Acemoglu's paper (2002), the stock of knowledge is represented by a technological bias in favor of skilled workers.

2.2.4 Government

The Government budget constraint can be written:

$$N\tau_t\bar{w}_t = Nb_t \tag{2.9}$$

knowing that τ_t is endogenously determined at each period by a voting procedure. The political equilibrium is detailed below.

2.3 The Dynamic Equilibrium

The dynamic equilibrium of this economy is the sequence $\{\theta_t, \tau_t, h_t\}_{t=0,..,+\infty}$ which satisfies the first order conditions of firms ((2.7) and (2.8)), the educational choices of agents (2.4), and the Government budget constraint (2.9). This sequence is such that the input markets and the output one are at equilibrium at each period. Thus, at each period we have:

$$L_t^e = NF(\tilde{\theta}_t) \tag{2.10}$$

$$L_t^u = N(1 - F(\tilde{\theta}_t)) \tag{2.11}$$

Before studying the dynamics of the economy, we have to specify some properties of wages at equilibrium.

2.3.1 Some Properties of Wages at Equilibrium

Given equations (2.10) and (2.11) the wages of educated and uneducated agents can be written:

$$w^e(h_t, \tilde{\theta}_t) = f_2(h_t, NF(\tilde{\theta}_t), N(1 - F(\tilde{\theta}_t))) \tag{2.12}$$

$$w^u(h_t, \tilde{\theta}_t) = f_3(h_t, NF(\tilde{\theta}_t), N(1 - F(\tilde{\theta}_t))) \tag{2.13}$$

Given our assumptions on the production function, it can easily be verified that the stock of knowledge capital (h_t) has a positive impact on both w^e and w^u ($w_1^e() > 0$ and $w_1^u() > 0$). However, it can also be shown that the share of the educated population ($F(\tilde{\theta}_t)$) has a negative impact on the wage level of educated agents ($w_2^e() < 0$), and a positive one on the wages of uneducated agents ($w_2^u() > 0$). This result comes from the decreasing marginal productivity of labor.

The average wage of the economy can be written:

$$\bar{w}_t = F(\tilde{\theta}_t)w^e(h_t, \tilde{\theta}_t) + (1 - F(\tilde{\theta}_t))w^u(h_t, \tilde{\theta}_t) \equiv \bar{w}(h_t, \tilde{\theta}_t) \tag{2.14}$$

The function $\bar{w}(h_t, \tilde{\theta}_t)$ is such that $\bar{w}_1() > 0$. An increase in the knowledge capital level has a positive impact on the average wage of the economy because of the increase in the wages of educated and uneducated agents (with $\tilde{\theta}_t$ constant). However, the share of the educated population has an ambiguous impact on the average wage of the economy. Differentiating equation (2.14) with respect to $\tilde{\theta}_t$ we have:

$$\frac{\partial \bar{w}(h_t, \tilde{\theta}_t)}{\partial \tilde{\theta}_t} = \underbrace{f(\tilde{\theta}_t)(w^e - w^u)}_{A>0} + \underbrace{F(\tilde{\theta}_t)\frac{\partial w^e(h_t, \tilde{\theta}_t)}{\partial \tilde{\theta}_t}}_{B<0} + \underbrace{(1 - F(\tilde{\theta}_t))\frac{\partial w^u(h_t, \tilde{\theta}_t)}{\partial \tilde{\theta}_t}}_{C>0} \tag{2.15}$$

Elements A and C are both positive. Element A denotes the increase in the average wage because of the increase in the share of the population having the higher wage level. Element C denotes the increase in the wage level of uneducated agents because of the bigger scarcity of this input. Finally, element B is negative because of the smaller scarcity of the educated population.

Given this preliminary analysis, and given equation (2.5), we have to determine the share of the educated population at each period. However, this choice depends on the result of the voting procedure. Consequently, we first study the determination of the share of the educated population and the size of transfers from the Government. Finally, we consider the dynamics of our economy, i.e. the sequence $\{\theta_t, \tau_t, h_t\}_{t=0,..,+\infty}$.

2.3.2 The Share of the Educated Population

At period $t-1$, every agent for whom the following inequality is (not) satisfied, decides to educate himself (to remain uneducated):

$$\ln\left(\frac{w_t^e(1-\tau_t^a)+\tau_t^a\bar{w}_t}{w_t^u(1-\tau_t^a)+\tau_t^a\bar{w}_t}\right) > v(\theta) \tag{2.16}$$

If it exists, the threshold value $\tilde{\theta}_t$ is defined by:

$$LHS(h_t,\tilde{\theta}_t,\tau_t^a) \equiv \ln\left(\frac{w^e(h_t,\tilde{\theta}_t)(1-\tau_t^a)+\tau_t^a\bar{w}(h_t,\tilde{\theta}_t)}{w^u(h_t,\tilde{\theta}_t)(1-\tau_t^a)+\tau_t^a\bar{w}(h_t,\tilde{\theta}_t)}\right) = v(\tilde{\theta}_t) \tag{2.17}$$

Lemma 2.1 *(i) There exists a unique threshold value $\tilde{\theta}_t$ such that $\underline{\theta} < \tilde{\theta}_t < \bar{c}$ iff $LHS(h_t,\underline{\theta},\tau_t^a) > v(\underline{\theta})$ and $LHS(h_t,\bar{\theta},\tau_t^a) < v(\bar{\theta})$. (ii) This threshold is such that $\partial\tilde{\theta}_t/\partial\tau_t^a < 0$ and $\partial\tilde{\theta}_t/\partial h_t > 0$.*

Proof: See the appendix of this chapter.□

The first element of the second part of lemma 2.1 shows that the size of the welfare state has a negative impact on the share of the educated population. Indeed, because the Government use flat transfers, they redistribute resources in favor of unskilled agents. Consequently, an increase in the size of the welfare state increases the opportunity cost to educate oneself.

The second element of the part (ii) of the lemma shows that, for a given tax rate, more agents decide to educate themselves if the knowledge capital level increases. Indeed, as the knowledge capital increases, the wage inequality ratio (w^e/w^u) increases because of the larger marginal impact of the knowledge capital on the wages of educated agents than on the wages of uneducated agents. Thus, the stock of knowledge increases the cost of remaining uneducated, and more agents decide to educate themselves.

Given lemma 2.1, we can define a function $\tilde{\theta}()$ such that:

$$\tilde{\theta}_t = \tilde{\theta}(h_t,\tau_t^a) \tag{2.18}$$

with $\tilde{\theta}_1() > 0$ and $\tilde{\theta}_2() < 0$.

2.3.3 The Political Equilibrium

In our economy, the Government is only an institution which applies the decision taken by the majority of the population. Each group of agents (educated and uneducated agents) is homogenous, thus every educated agent votes for the same tax rate. As for uneducated agents, they vote for the same tax rate which is different from that of educated agents.

We assume that agents have to choose between two tax rates, $\bar{\tau}$ and $\underline{\tau}$, such that: $1 > \bar{\tau} > \underline{\tau} > 0$. Agents choose the tax rate which maximizes their wealth level. The wealth level of an agent of type i can be written:

$$W_t^i = w_t^i(1 - \tau_t) + \tau_t \bar{w}_t \tag{2.19}$$

Given assumption 2, the wealth of educated agents is a decreasing function of τ_t, while the wealth of uneducated agents is an increasing function of τ_t. It implies that educated agents vote for the lower tax rate $\underline{\tau}$, while uneducated agents vote for the higher tax rate $\bar{\tau}$.

The political equilibrium is:

$$\tau_t = \begin{cases} \underline{\tau} & \text{if } F(\tilde{\theta}_t) > 1/2 \\ \bar{\tau} & \text{if } F(\tilde{\theta}_t) < 1/2 \end{cases}$$

We conclude from this analysis that there exists a threshold $\hat{\theta}_t$ such that:

$$F(\hat{\theta}_t) = 1/2 \tag{2.20}$$

As long as $\tilde{\theta}(h_t, \tau_t^a) < \hat{\theta}_t$, uneducated agents have the majority and the higher tax rate $(\bar{\tau})$ is chosen. Conversely, once $\tilde{\theta}(h_t, \tau_t^a) > \hat{\theta}_t$, educated agents have the majority and the lower tax rate $\underline{\tau}$ is chosen. It implies that the expectations of agents concerning the tax rate has an influence on the result of the vote. We define two threshold values for the knowledge capital level: h_a and h_b, such that:

$$F(\tilde{\theta}(h_a, \underline{\tau})) = 1/2 \tag{2.21}$$

and

$$F(\tilde{\theta}(h_b, \bar{\tau})) = 1/2 \tag{2.22}$$

Given that $\tilde{\theta}_1() > 0$ and $\tilde{\theta}_2() < 0$, it is straightforward to show that $h_b > h_a$. More specifically, differentiating equation (2.20) with respect to h_t and τ_t^a we obtain:

$$\frac{dh}{d\tau^a} = \frac{\tilde{\theta}_2(h, \tau^a)}{\tilde{\theta}_1(h, \tau^a)}$$

This equation means that the gap between h_a and h_b is all the more small as the ratio between the marginal impact of the tax rate and the marginal impact of the stock of knowledge is low. It means that the differential between the elasticity of the wages of educated agents and the elasticity of the wages of uneducated agents with respect to the stock of knowledge plays a significant role. If this differential is high, then a small increase in h provides a strong incentive to educate oneself. Consequently, if τ highly increases, then it could be sufficient that h slightly increases for educated agents to obtain the majority.

Moreover, if the distance between h_a and h_b is small, then h_a and h_b take low (high) values if a large fraction of the population has small (big) educational costs. It could represent some other specificities of education which are not introduced in this chapter such as low monetary costs of education. It means that the educated population gets the majority for low values of the stock of knowledge.

Lemma 2.2 sums up the result of the vote for different values of h_t.

Lemma 2.2 *If $h_t < h_a$, then no agent born at period $t-1$ expects that $\underline{\tau}$ can be adopted and $\tau_t = \bar{\tau}$. For $h_t > h_b$, no agent expects the tax rate $\bar{\tau}$ to be adopted and thus $\tau_t = \underline{\tau}$. For $h_a < h_t < h_b$, then we have self-fulfilling prophecies.*

For $h_t < h_a$, the knowledge capital level is so small that agents know at period $t-1$ that uneducated agents will have the majority at period t. Conversely, for $h_t > h_b$, the knowledge capital level is so high that agents know at period $t-1$ that educated agents will have the majority at period t. For the intermediate case ($h_a < h_t < h_b$), if agents expect that the low (high) tax rate will be chosen, then a large (small) share of the population will educate itself and educated agents will (not) actually have the majority.

2.3.4 The Dynamics

Given equations (2.5), the stock of knowledge of period t only depends on the share of the educated population of period $t-1$ and on the knowledge capital level h_{t-1}. Assuming that both can be observed without cost, agents know at period $t-1$ the value of the knowledge capital of period t (h_t). However, agents born at period $t-1$ expect the result of the vote of period t. This expectation has an impact on the share of the educated population at period t, and thus it has an impact on the knowledge capital of period $t+1$.

With equation (2.18), equation (2.5) can be rewritten:

$$h_{t+1} = \Psi(h_t, \tilde{\theta}(h_t, \tau_t^a)) \equiv \Phi(h_t, \tau_t^a) \tag{2.23}$$

Given the properties of the functions $\Psi()$, we have $\Phi_1() > 0$, i.e. a higher knowledge capital level at period t has a direct positive impact on h_{t+1}, and an indirect impact through the share of the educated population of period t. Given Lemma 2.2, the dynamics of our economy is described by:

$$h_{t+1} = \begin{cases} \Phi(h_t, \bar{\tau}) \text{ if } h_t < h_a, \text{ or if } h_a < h_t < h_b \text{ and } \tau_t^a = \bar{\tau} \\ \Phi(h_t, \underline{\tau}) \text{ if } h_t > h_b, \text{ or if } h_a < h_t < h_b \text{ and } \tau_t^a = \underline{\tau} \end{cases}$$

h_{t+1} is directly influenced by choices made at period $t-1$. Indeed, the share of the educated population of period t is determined at period $t-1$. However, this decision depends on the expected tax rate of period t. Consequently, for low (high) values of knowledge capital, agents know that the high (low) tax rate will be chosen at period t, and only a few (most) people decide to educate themselves, which has a negative (positive) impact on h_{t+1}. For $h_a < h_t < h_b$, if agents expect the tax rate $\bar{\tau}$ to be chosen at period t, then only a few people decide to educate themselves, which implies that uneducated agents will actually have the majority. At period $t-1$, the expectation of a high tax rate at period t has a negative impact on knowledge capital of period $t+1$.

Assumption 3: $\Phi_{11}(h_t, \tau_t^a) < 0$ and $\Phi(0, \tau_t^a) \geq 0$.

The first part of this assumption means that the function $\Phi()$ is a strictly concave function of h_t, which is a sufficient condition for the stability of the steady state. The second part of this assumption implies that there exists a unique non-trivial steady state for knowledge capital. Let us call a *potential steady state*, a steady state which should be obtained if an economy kept the same tax rate τ at every period. $\underline{h}$ and $\bar{h}$ denote the steady state knowledge capital levels for the tax rates $\underline{\tau}$ and $\bar{\tau}$ respectively. Given assumption 4 and the properties of the functions $\Phi()$ and $\tilde{\theta}()$, it is straightforward to show that $\underline{h}$ and $\bar{h}$

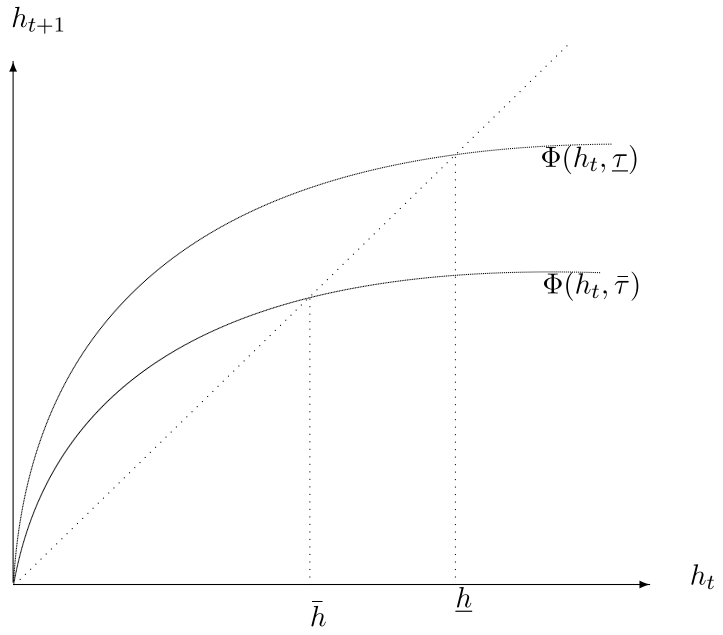

Figure 2.1: Potential steady states

are stable, and that $\underline{h} > \bar{h}$. This result comes from the negative impact of redistributive policies on the share of the educated population (see figure 2.1).

The dynamics of our economy is not trivial and we have to consider some cases in function of the respective positions of $\bar{h}$, $\underline{h}$, h_a and h_b[9]. We consider the case in which the initial value of knowledge capital h_0 is given. The first generation of agents born at period -1, decide to educate themselves in function of h_0 and of τ_0^a.

Case 1: $\bar{h} > h_b$ and $\underline{h} > h_b$

Figure 2.2 illustrates that case, which occurs as long as the gap between h_a and h_b is not too high and if a large fraction of the population has small educational costs. In that case, the economy cannot converge towards the low steady state $\bar{h}$, because at that point, educated agents have the majority and the tax rate $\underline{\tau}$ is chosen. In that case, a large welfare state is only a temporary equilibrium. Let us consider that the initial value of knowledge capital is small ($h_0 < h_a$). Figure 2.2 gives two examples of trajectories. h_0 is so small that agents born at period -1 know that the high tax rate will be chosen

[9] We do not study all cases but only the ones which we consider the most relevant.

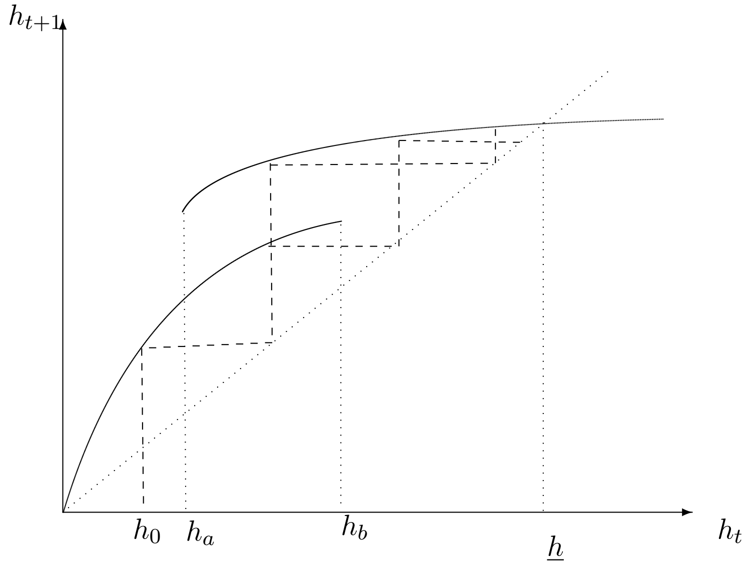

Figure 2.2: Illustration of case 1

at period 0. However, given the value h_1, agents born at period 0 can have two kinds of expectations. They can either expect that τ_1 will be high and in that case the knowledge capital follows the low trajectory; or they can expect that τ_1 will be low and the knowledge capital follows the high trajectory. In the long-run, whatever the initial value of h, the economy will have a small welfare state $\underline{\tau}$ and knowledge capital will converge towards $\underline{h}$ but self-fulfilling prophecies can change the dynamics of the economy and the date from which a small welfare state is adopted[10].

Case 2: $h_a > \bar{h}$ and $h_a > \underline{h}$

Figure 2.3 illustrates that case, which occurs as long as the gap between h_a and h_b is not too high and if the educational costs remain high for a significant part of the population. In that case, if $h_0 < \bar{h}$, then the economy converges towards $\bar{h}$ and a large welfare state is chosen at each period. In the long run, if $h_0 > h_b$ the economy converges towards $\bar{h}$ and a large welfare state. However, the economy adopts, at least temporarily, a low tax rate.

[10]As in Bisin and Verdier (2000, 2002), the beliefs of agents matter to study the dynamics of our economy. However, we assume that there is no coordination problem in the economy, i.e. agents share the same beliefs concerning the future policy.

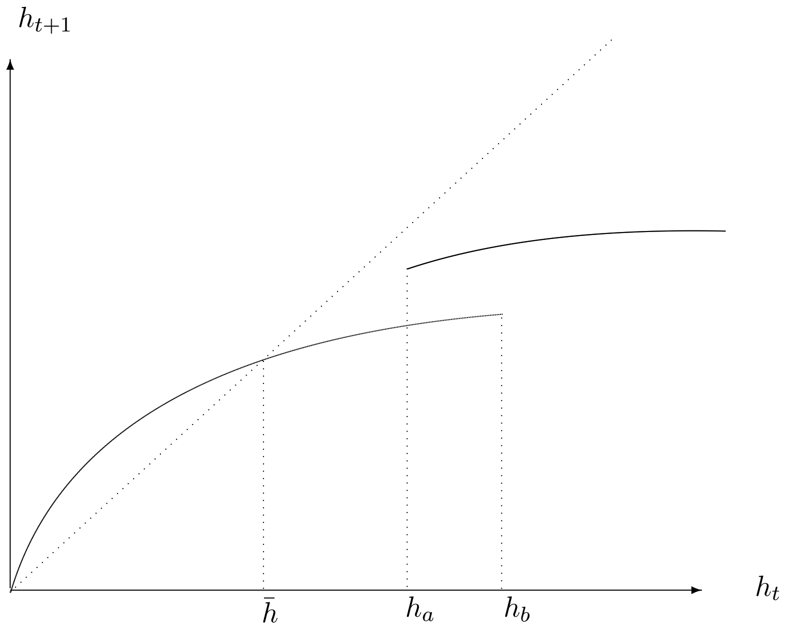

Figure 2.3: Illustration of case 2

As in the previous case, self-fulfilling prophecies only change the transitional dynamics but not its behavior at steady state.

Case 3: $\underline{h} > h_b > h_a > \bar{h}$

Figure 2.4 illustrates this case, which occurs as long as the gap between h_a and h_b is not too high and if a large fraction of the population has small educational costs. In that case, the economy can converge towards two kinds of steady state. If $\bar{h} < h_0 < h_a$, agents born at period -1 expect that a high tax rate will be chosen at period 0, consequently, only a few people educate themselves. Less and less agents decide to educate themselves at every following periods and the economy converges towards the low knowledge capital steady state $\bar{h}$ with a large social security system ($\bar{\tau}$). Conversely, if $\underline{h} > h_0 > h_b$, agents born at period -1 know that a low tax rate will be chosen at period 0. Consequently, a large part of them decide to educate themselves. More and more agents decide to educate themselves at every following periods. The economy converges towards the high knowledge capital steady state $\underline{h}$ with a small welfare state ($\underline{\tau}$).

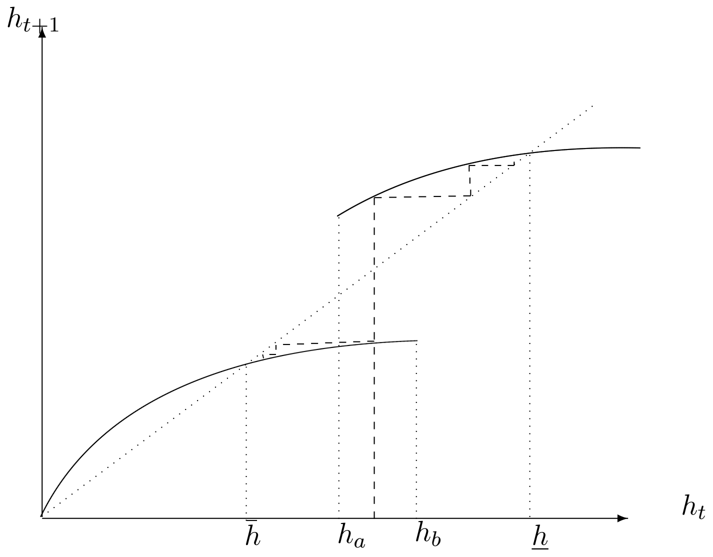

Figure 2.4: Illustration of case 3

Finally, if $h_a < h_0 < h_b$ the economy can converge either towards $\bar{h}$, or towards $\underline{h}$. It depends on the beliefs of agents. If the agents initially expect a low tax rate to be chosen, then the economy will converge towards $\underline{h}$ and a low tax rate will be chosen at steady state. Conversely, if agents expect a high tax rate to be chosen at steady state, $\bar{h}$ is chosen at steady state.

Case 4: $h_b > \underline{h} > \bar{h} > h_a$

Figure 2.5 illustrates that case, which occurs as long as the gap between h_a and h_b is high, i.e. if the share of the educated population depends more on the redistributive policies than on the elasticity of wages with respect to the stock of knowledge[11]. In this specific context the economy can switch from one steady state to the other according to the beliefs of agents concerning the size of the welfare state. For example, let us assume that h_0 is sufficiently small for agents to expect a large welfare state to be adopted at period

[11]It means that the gap between the elasticity of the wages of educated agents and the elasticity of the wages of uneducated agents with respect to h is small. In that case the technological bias introduced by the stock of knowledge is small.

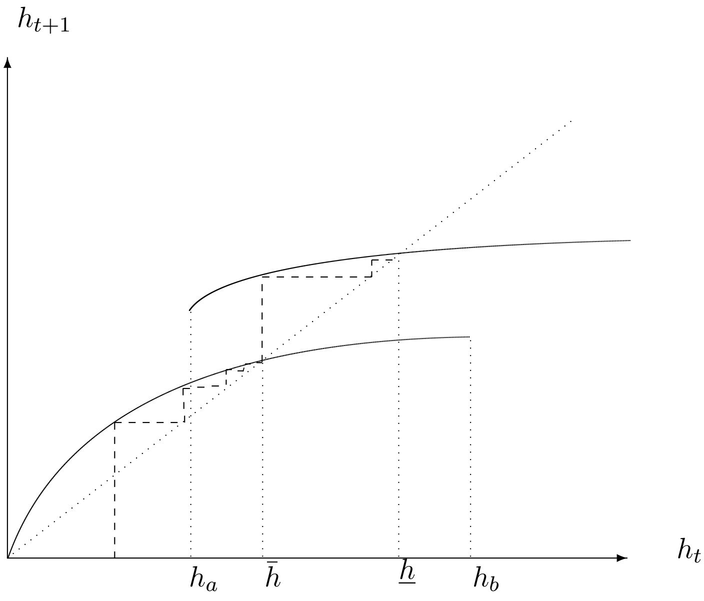

Figure 2.5: Illustration of case 4

0. If this belief endures for every following generation, knowledge capital will converge towards $\bar{h}$ which is associated with a large social security system. But, once at steady state, if the beliefs of agents change and if young agents expect that the welfare state will not be generous, then a large part of the population decides to educate itself and $\underline{\tau}$ will actually be chosen. If this belief endures, the economy will converge towards $\underline{h}$ with a less redistributive welfare state. In the same way, once the economy reaches this new steady state, if some generations of agents believe that a large welfare state will be chosen, the economy will come back to the steady state $\bar{h}$.

2.4 Concluding Remarks

In this chapter we study a basic economy *à la* Meltzer and Richard (1981) in a dynamic framework, with an endogenous structure of the population. Our model emphasizes the significant role played by self-fulfilling prophecies.

We study four relevant cases. In the first case, the economy converges towards a small welfare state. But, because of self-fulfilling prophecies, an economy can switch from a large towards a small welfare state more or less quickly. In the second case the economy converges towards a large welfare state. In the third case, the economy can, in the long run, adopt either a large or a small welfare state, depending on the initial value of the knowledge capital and on the beliefs of agents. In the fourth case, we show that an economy can switch from one steady state to the other according to the beliefs of agents concerning the size of the welfare state.

Using our model, the analysis of the crisis of the welfare state does not seem to be trivial. Indeed, the structure of the population could explain why a larger share of the population questions the size of the welfare state. However, it could be a transitory situation until a larger share of the population expects larger transfers as in the fourth case of this chapter. Consequently, the size of the welfare state could follow a cyclical dynamics, which seems to be a new result in the political economy literature.

Our model could be extended at least in three ways: (i) it would be interesting to introduce pension systems. We try to do that in the next chapter. (ii) We could endogenize fertility choices as the structure of the population has a significant impact on this variable. (iii) We could endogenize the educational costs using a voting procedure on public educational expenditures.

2.5 APPENDIX

This appendix provides the proof of lemma 2.1.

(i) The RHS is a strictly increasing function of $\tilde{\theta}_t$ by assumption (see assumption 1). There only remains to proove that the LHS is a decreasing function of $\tilde{\theta}_t$.

Differentiating the ratio of the LHS of equation (2.17), we obtain an expression of the following form:

$$\frac{\frac{\partial w^e}{\partial \tilde{\theta}}(1-\tau^a)(w^u(1-\tau^a)+\tau^a\bar{w}) - \frac{\partial w^u}{\partial \tilde{\theta}}(1-\tau^a)(w^e(1-\tau^a)+\tau^a\bar{w}) + \tau^a(1-\tau^a)(w^u-w^e)\frac{\partial \bar{w}}{\partial \tilde{\theta}}}{\left[w^u(h,\tilde{\theta})(1-\tau^a)+\tau^a\bar{w}(h,\tilde{\theta})\right]^2}$$

Using equation (2.15) this expression becomes:

$$\underbrace{\frac{(1-\tau^a)w^u\frac{\partial w^e}{\partial \tilde{\theta}} - \frac{\partial w_u}{\partial \tilde{\theta}}(1-\tau^a)(w^e(1-\tau^a)+\tau^a\bar{w})}{\left[w^u(h,\tilde{\theta})(1-\tau^a)+\tau^a\bar{w}(h,\tilde{\theta})\right]^2}}_{<0}$$

$$\underbrace{\frac{+\tau^a(1-\tau^a)(w^u-w^e)[f(\tilde{\theta})(w^e-w^u)+(1-F(\tilde{\theta}))\frac{\partial w^u}{\partial \tilde{\theta}}]}{\left[w^u(h,\tilde{\theta})(1-\tau^a)+\tau^a\bar{w}(h,\tilde{\theta})\right]^2}}_{<0}$$

(ii) The LHS is a strictly decreasing function of τ_t^a.

Differentiating the ratio of the LHS of equation (2.17) with respect to h_t, we obtain an expression of the following form:

$$\frac{\frac{\partial w^e}{\partial h}(1-\tau^a)(w^u(1-\tau^a)+\tau^a\bar{w})-\frac{\partial w^u}{\partial h}(1-\tau^a)(w^e(1-\tau^a)+\tau^a\bar{w})+\frac{\partial \bar{w}}{\partial h}\tau^a(1-\tau^a)(w^u-w^e)}{\left[w^u(h,\tilde{\theta})(1-\tau^a)+\tau^a\bar{w}(h,\tilde{\theta})\right]^2}$$

Using equation (2.14) we obtain:

$$\frac{(1-\tau^a)\left(\frac{\partial w^e}{\partial h}w^u-\frac{\partial w^u}{\partial h}w^e\right)}{\left[w^u(h,\tilde{\theta})(1-\tau^a)+\tau^a\bar{w}(h,\tilde{\theta})\right]^2}>0$$

The numerator is positive under assumption 2 and if $\tau^a < 1$. As we assume that τ cannot be equal to 1 then this condition is satisfied.

□

Chapter 3

The Dynamics of the Welfare State with Vertical and Horizontal Transfers

3.1 Introduction

Since the end of the second world war, the size of social security systems has increased without any Government questioning this expansion (Nyce and Schieber 2005). However, for a few years, an increasing part of public opinion has begun to consider a privatization of social security systems in most developed countries. A wide literature studies the macroeconomic consequences of such a privatization especially concerning pensions[1]. However, those authors do not explain this political change. In this chapter, we argue that it can be the last step of a global dynamics. We assume that the pension system is only one element of the redistributive policy of a country. Each agent votes on the size of the whole redistributive policy. We explain why it is possible for an economy to finance its redistributive policies by using a high tax rate only temporarily. The main intuition is that the increase in the share of the educated population has a negative impact on the demand for vertical redistribution. Consequently, there is a new political pressure in favor of a smaller welfare state.

In this chapter, social security is only one part of the welfare state. Indeed, the Welfare State also provides pure vertical transfers. Intra-generational transfers are such that a share of revenues from payroll taxes are used to finance a flat transfer in favor of workers. The rest of the revenues are used to finance a social security system which provides inter-generational transfers through a Pay-As-You-Go (PAYG) pension system.

A wide part of the political economy literature has studied these two kinds of transfers[2]. But only a few papers consider these two dimensions together[3]. And when they do so, they use a static framework, which implies that the identity of the median voter is only influenced by exogenous parameters[4]. In this chapter, we use a more complete dynamics. The structure of the working population depends on the development level of the economy. Indeed, we assume that the capital level has a stronger impact on the wages of educated

[1]See Belan and Pestieau (1998), Docquier and Paddison (2003), or Casarico and Devillanova (2007). These results are questioned notably by Groezen *et al.* (2007), Lambrecht *et al.* (2005) or Le Garrec (2005).

[2]See Meltzer and Richard (1981) or Persson and Tabellini (2000) for intra-generational transfers; and Galasso and Profeta (2002) for inter-generational transfers.

[3]Razin *et al.* (2002a,2002b,2004) and Galasso and Profeta (2007) do so in order to detail the effects of the ageing of the population.

[4]Except Bisin and Verdier (2001) who study the dynamics of the Welfare State in a cultural transmission model.

agents than on those of uneducated agents. It implies that a higher level of capital has a positive impact on the share of the educated population. However, educated agents prefer a small welfare state. Thus, if an economy reaches a high level of capital, it is possible that the size of the welfare state decreases. This mechanism is the main intuition of our chapter.

Let us now detail our model. We use an overlapping generations model with a closed economy in which agents live for three periods. They can be young, workers or retired. When they are young, agents decide to educate themselves or not. This choice only depends on the comparison between the wage differential (between the wage of the educated population and that of the uneducated population[5]) and their idiosyncratic leisure cost of education. The wage differential depends on the capital level because the capital has a stronger impact on the wages of educated agents than on those of uneducated agents. Indeed, Krusell *and al.* (2000) show that the increase in wages which have been observed in the United-States can be explained by a capital-skill complementarity technology. In this chapter, we do not use a capital-skill complementarity technology but a technology in which the capital level has an external effect on the marginal productivity of workers. This assumption enables us to obtain clear analytical results. In each period, the population votes on the size of the welfare state. They can choose either a high tax rate or a low tax rate. Because agents take the next period tax rate as given, workers only support the redistributive policies because of their intra-generational component. Educated agents prefer a small welfare state because they neither benefit from vertical transfers, nor from the pension system. Conversely, old agents vote for the highest welfare state because they benefit from the pension system without bearing its cost. Finally, uneducated agents support the redistributive policy if and only if the pure vertical transfers are sufficiently high. As long as the capital level is low, the average wage of the economy is low and uneducated agents prefer a small welfare state. Indeed, vertical transfers are not sufficient for uneducated agents to benefit from the redistributive policy. However, once the capital accumulation becomes significant, uneducated agents vote for the highest welfare state.

In this context we show that an economy can converge towards two kinds of equilibrium. The first one is such that the capital level is low. It implies that uneducated and old agents have the majority and vote for a large welfare state. But the economy can also

[5]When agents are young, they do not consider transfers from the Government because we assume that they are myopic. It is explained with more details in the next section.

converge towards another steady state in which the capital level is high. In that case a large part of the population decides to educate itself. Consequently, educated agents have the majority and vote for a small welfare state. In this model high tax rates can correspond to transitional equilibria. It is possible to observe three stages in the dynamics of the welfare state: first a low tax rate, then a high tax rate and finally a low tax rate.

Our model is a first attempt to understand the dynamics of the welfare state better and to explain, in a new framework, the recent political pressures in favor of a privatization of social security systems.

This chapter is organized as follows. In section 2 we detail our model and our main assumptions. In section 3 we study the dynamics of our model and its properties. Section 4 includes some concluding remarks.

3.2 The Model

We assume a closed economy in which generations overlap. In each period, three generations coexist: young, working and retired agents. The size of each new generation is assumed to be constant and is normalized to 1 without loss of generality. The length of each period is normalized to one. We assume that agents live their two first periods of life. However, every agent has a probability T to live their third period of life[6].

When young, the agents decide to educate themselves or not. It is the only economic decision which they can take. When an agent is born, he is randomly endowed with an educational cost θ. θ denotes the psychological cost that each agent bears if he decides to educate himself (Le Garrec 2005). Here, the cost endowment has an impact on the necessary time for education, i.e. on the time which cannot be spent on leisure. For the sake of simplicity, we assume that it is the only cost of education. Thus, we assume that uneducated agents do not begin to work earlier. $f(\theta)$ and $F(\theta)$ denote the density function and the cumulative distribution function of this variable. But it also represents the share of the population with a cost θ. θ is defined over an interval $\Omega_\theta = [\underline{\theta}, \bar{\theta}]$. $f(\theta)$ is such that $f(\theta) > 0$, $\forall \theta \in \Omega_\theta$, $F(\underline{\theta}) = 0$ and $F(\bar{\theta}) = 1$. A low θ implies that an agent has a high ability to educate himself, i.e. the agent does not need much time to educate himself. Conversely, a high θ implies that the agent has a high cost to educate himself. Each agent is assumed to choose between being educated or not. If he chooses to educate himself then

[6] It is a model *à la* Yaari (1965).

he bears the psychological cost mentioned above. But if he decides to remain uneducated then he bears no psychological cost.

In this economy agents are assumed to be myopic. Indeed, when they are young they do not take into account their third period utility and the transfers from the Government[7][8]. It implies that when they decide to educate themselves or not, they consider that they consume their wage when they work. Nevertheless, once they begin to work, agents are assumed to take into account their last period utility and all their resources.

When agents work, they save a share of their wage for their old age consumption.

Consumers

The preferences of a representative agent born at the beginning of the period $t-1$ are the following[9]:

$$U_{t-1}^{i}(\theta) = \ln(c_t^i) + \tilde{\Lambda}\beta T\left(d_{t+1}^{i}\right) - (1-I)v(\theta) \tag{3.1}$$

with $I = 1$ if the agent decides not to educate himself ($i = u$), and $I = 0$ in the opposite case ($i = e$); and $\beta \in (0,1)$ the psychological discount factor. c_t^i and d_{t+1}^i denote the consumption when the agent is adult and when he is old respectively. The first two parts of this equation are standard and imply that an increase in life expectancy (T) increases the weight of the second period consumption. The third term ($v(\theta)$) in the utility function denotes the psychological cost of education. We make the following assumption concerning the function $v(\theta)$:

[7]This assumption highly simplifies the analytical results. It is almost the same assumption as in Cremer *et al.* (2007) or in Feldstein (1985), except that agents are only myopic when they are young, i.e. when they decide or not to educate themselves. In this way we emphasize the link between the educational choices and the capital level. It is a useful result to obtain a tractable model without indeterminacy. The previous chapter explores this extension.

[8]Feldstein (1985) and Feldstein and Liebman (2002) use the same assumption for the representation of the myopia of agents. The myopia concerns the future utility and the transfers from the Government. For example in his model, agents decide, when they work, to save or not for their old age consumption. Their utility function when they are young is: $U_t = u(w_t(1-\tau_t) - S_t) + \tilde{\Lambda}v(S_tR + \alpha b_{t+1})$, with b the pension received by agents when they are old. In his model, $\tilde{\Lambda}$ measures the myopia of agents concerning their second period of life utility function. α measures the myopia of agents concerning the transfers from the Government. For the specific case in which $\tilde{\Lambda} = \alpha = 0$, then agents are completely myopic.

[9]We assume quasi-linear preferences in order to obtain a tractable model.

Assumption 1: $v(\bar{\theta}) > 0$, $v(\theta)$ is of class C^1 and $v'(\theta) > 0$, $\forall\theta \in \Omega_\theta$.

We do not exclude the case in which $v(\underline{\theta}) < 0$. It implies that for a fraction of the population there could be a psychological benefit from being educated. This benefit is all the more large as θ tends towards $\underline{\theta}$. Finally, $\tilde{\Lambda}$ denotes the myopia of agents. Given our previous assumptions we have that $\tilde{\Lambda} = 0$ when an agent is young and $\tilde{\Lambda} = 1$ when he works. The budget constraints of agents are:

$$c_t^i = w_t^i(1 - \tilde{\Lambda}\tau_t) + \tilde{\Lambda}G_t^i - S_t^i \tag{3.2}$$

$$d_{t+1}^i = S_t^i \tilde{R}_{t+1} + \tilde{\Lambda}p_{t+1}^i \tag{3.3}$$

with $i = e$ ($i = u$) if the agent is (un)educated. w^i denotes the wage earned by a worker of type i, and G_t^i is the Government transfer for the working population. The labor supply of agents is assumed to be inelastic and is equal to 1 during his second period of life. Each agent retires at the end of this period. τ_t is the tax rate used to finance the welfare state. S_t is the savings of the agent. $\tilde{R}_{t+1}$ is the return on savings, and p_{t+1} is the pension of the agent. We assume a perfect annuity market, such that $\tilde{R}_{t+1} = R_{t+1}/T$.

When an agent is young, he considers that $c_t^i = w_t^i$. Consequently, an agent born at period $t-1$ decides to educate himself if $U_{t-1}^e > U_{t-1}^u$. This condition can be written extensively as:

$$\ln(w_t^e) - v(\theta) > \ln(w_t^u) \tag{3.4}$$

or:

$$\ln\left(\frac{w_t^e}{w_t^u}\right) > v(\theta) \tag{3.5}$$

We define $\tilde{\theta}$ as the educational cost which is such that the expected utility if the agent decides to educate himself exactly compensates the expected utility if the agent makes the opposite choice. At each period t, $\tilde{\theta}$ is defined by:

$$v(\tilde{\theta}_t) = \ln\left(\frac{w_t^e}{w_t^u}\right) \tag{3.6}$$

For $\theta < \tilde{\theta}_t$ the agent decides to educate himself. Conversely, agents for whom the cost is higher than $\tilde{\theta}_t$ decide not to educate themselves. $\tilde{\theta}_t$ determines the share of the educated population in period t which is defined as: $F(\tilde{\theta}_t)$.

At the beginning of the period t, an agent born in period $t-1$ becomes a worker. He is not myopic any more and $\tilde{\Lambda} = 1$. He saves a part of his wage for his old age consumption. Given his choice of the previous period (he can be educated ($i = e$) or not ($i = u$)), the agent maximizes (3.1) with respect to S_t^i subject to the budget constraints (3.2) and (3.3). It is straightforward to show that[10]:

$$S_t^i = \frac{\beta R_{t+1}\left(w_t^i(1-\tau_t) + G_t^i\right) - 1}{\beta R_{t+1}} \tag{3.7}$$

Firms

The number of firms is normalized to 1. We assume a perfect competitive economy on the final good market and on the inputs markets. The production function of a representative firm is assumed to have the following form[11]:

$$Y_t = XK_t + A_t L_t^e + B_t L_t^u \tag{3.8}$$

with K_t the physical capital, L_t^e the skilled labor demand and L_t^u the unskilled labor demand. $X > 0$ denotes the marginal productivity of capital at the firm level, which is constant by assumption. Each firm takes A_t and B_t as given. Furthermore, we assume that[12]:

$$A_t = A(K_t) \tag{3.9}$$

and that:

$$B_t = B(K_t) \tag{3.10}$$

with $A(K_t) > 0$ and $B(K_t) \geq 0$. It means that the marginal productivity of each kind of labor depends on the average level of capital in the economy. Each firm takes A_t and B_t as given. Furthermore, we make the following assumption:

---

[10] S_t^i is independent of the pension of the last period of life because of the quasi-linearity of the utility function. Moreover, it can easily be shown that the total savings is positive. As the saving function is an increasing function of the wage level, it does not necessarily imply that uneducated agents save. However, because of our assumption of perfect annuity markets, it is not really a problem.

[11] We use this analytical form because we can study the dynamics of our economy more easily.

[12] It implies that the capital level has an external effect on the wages of educated agents and on the marginal product of capital.

Assumption 2: (i) $A(K_t) > B(K_t), \forall K_t > 0$. (ii) $K_t A'(K_t)/A(K_t) > K_t B'(K_t)/B(K_t) > 0$.

This assumption implies (i) that the marginal productivity of the educated population is higher than that of the uneducated population if there is a positive quantity of capital in the economy. (ii) Furthermore, an increase in the stock of physical capital (K_t) has a positive impact on the knowledge of the educated and uneducated population, which increases their marginal productivity. However, the elasticity of the marginal productivity of the educated population with respect to the capital level is higher than the elasticity of the marginal productivity of the uneducated population with respect to the capital level.

We assume a complete depreciation of the capital. As we have a perfect competitive market the prices of inputs are:

$$R_t = X \tag{3.11}$$

$$w_t^e = A(K_t) \tag{3.12}$$

$$w_t^u = B(K_t) \tag{3.13}$$

Because of assumption 2, the wage ratio between the wages of educated and uneducated agents: A_t/B_t, is an increasing function of K_t. A higher physical capital level has a positive impact on wage inequalities. It seems to be a reasonable assumption given the empirical evidence on this point (Acemoglu 2002).

We can determine the share of the educated population of the period t. Using equation (3.6) we obtain:

$$\ln\left(\frac{A(K_t)}{B(K_t)}\right) = v(\tilde{\theta}_t) \tag{3.14}$$

This equation defines an implicit relationship between $\tilde{\theta}_t$ and K_t. And finally, under assumption 2:

$$\tilde{\theta}_t \equiv \tilde{\theta}(K_t) \text{ with } \theta'(K_t) > 0 \tag{3.15}$$

It implies that a higher physical capital level has a positive impact on the share of the educated population ($F(\tilde{\theta}_t)$).

Using equations (3.12), (3.13), and (3.15) we can define the average wage of the economy as:

$$\bar{w}(K_t) = A(K_t)F(\tilde{\theta}_t) + B(K_t)(1 - F(\tilde{\theta}_t)) \tag{3.16}$$

Government

We assume that the Government takes its decisions in order to be reelected. Here, we only consider redistributive transfers of the Government. These transfers are intra-generational through flat transfers, and inter-generational through a PAYG pension system. We assume that agents vote on the size of the welfare state as a whole. To simplify, the tax rate used to finance the welfare state can take two values called τ^1 and τ^2, with $1 > \tau^1 > \tau^2 \geq 0$. A fraction ($1 - \alpha$, with $1 > \alpha > 0$) of current tax revenues is used for intra-generational transfers (G^i) of the same period, and a fraction (α) of these revenues for inter-generational transfers (p^i).

The budget constraint of the Government for inter-generational transfers can be written[13]:

$$\alpha \tau^j \bar{w}_t = \left[p_t^{e,j} F(\tilde{\theta}_{t-1}) + p_t^{u,j} (1 - F(\tilde{\theta}_{t-1})) \right] T \tag{3.17}$$

with $j = 2$ ($= 1$) if the tax rate τ^2 (τ^1) is chosen. $p^{i,j}$ denotes the pension received by agent of type i if the tax rate τ^j is chosen. Because of the form of the utility function, we do not have to specify the form of inter-generational transfers to study the dynamics of the economy. The Left-Hand-Side (LHS) of this expression is the product between the tax rate and the average wage of the economy. The Right-Hand-Side (RHS) is the average pension paid to the old agents who have survived at the end of period $t - 1$. $p^{e,j}$ denotes the pension paid to educated agents if the tax rate τ^j is chosen.

Concerning intra-generational transfers, it is assumed that every agent receives the same amount[14]. Each agent receives:

$$G_t^i = (1 - \alpha) \tau^j \bar{w}_t \tag{3.18}$$

$\forall i \in \{e, u\}$ and for a given $j \in \{1, 2\}$.

At each period, the population votes on the size of the current welfare state taking as given the tax rate of the next period[15]. We assume that a tax rate is chosen if more than

[13]We do not have to specify $p^{i,j}$ because it does not play any role in our dynamics because of our log-linear utility function.

[14]Intra-generational transfers are Beveridgian.

[15]We do not use the Markovian equilibria of Krusell and Rios-Rull (1996) and of Krusell, Quadrini and Rios-Rull (1997).

half of the population support this tax rate. Consequently, the most important agent is the median voter. We analyze the identity of this agent below.

Let us now determine the tax rate chosen by each group of the population (young agents, uneducated workers, educated workers and old agents).

We assume that young agents do not participate to the voting procedure because they neither benefit nor pay for the welfare state. Old agents vote for the maximum tax rate because they benefit from the pension system without paying for it. The size of the old population is T.

Let us now consider the behavior of educated workers. They take as given the tax rate which will be used for their pension but they choose the current tax rate[16]. Using equations (3.1), (3.2) and (3.3) the welfare level of educated agents is a strictly decreasing function of τ_t (for $K_t > 0$) because:

$$-w_t^e + (1-\alpha)\bar{w}_t < 0$$

Then an educated worker always chooses the lower tax rate τ^2.

For uneducated agents, the decision is a little bit more complicated. Indeed, their welfare level is an increasing function of τ_t iff:

$$-w_t^u + (1-\alpha)\bar{w}_t \geq 0$$

It is obvious that if $\alpha = 0$ ($\alpha = 1$) then they choose the higher (lower) tax rate. This condition implies that uneducated agents choose the lower tax rate if:

$$\alpha \geq 1 - \frac{w_t^u}{\bar{w}_t}$$

and the higher tax rate if the inequality sign is "less than". This condition implies the following lemma:

[16] In this chapter there exists a second kind of myopia because agents do not take into account the impact of the current tax rate on capital accumulation. To sum up, the timing of the model is the following: (1) agents decide to educate themselves or not with a first kind of myopia. (2) They vote on the size of the welfare state without considering the impact of this choice on capital accumulation. (3) Given the policy and the educational choice agents choose the amount to save.

Lemma 3.1 *(i) If* $\alpha \leq 1 - \frac{w^u(0)}{\bar{w}(0)}$ *then uneducated agents always vote for the higher tax rate* τ^1. *(ii) If* $\alpha \geq 1 - \frac{w^u(0)}{\bar{w}(0)}$ *and if* $\alpha \leq 1 - \frac{w^u(+\infty)}{\bar{w}(+\infty)}$, *then there exists a threshold* $\breve{K}$ *such that if* $K_t < \breve{K}$ *then uneducated agents vote for the lower tax rate* (τ_2). *But if* $K_t > \breve{K}$ *they vote for the higher tax rate* (τ_1).

Proof: The RHS of the previous inequality is an increasing function of K_t. If the RHS is strictly higher (smaller) than the left-hand-side (LHS) for K_t which tends towards 0 ($+\infty$), then there exists a threshold value $\breve{K}$ such that the previous inequality is satisfied (reversed) for $K_t < \breve{K}$ ($>$).□

Lemma 3.1 means that as long as $K_t < \breve{K}$, the average wage of the economy is not sufficiently high for uneducated agents to vote for the high tax rate[17]. $\breve{K}$ is all the more high as a large share of tax revenues is dedicated to pensions. In the rest of this chapter we only consider the case in which $\breve{K} > 0$.

To sum up, two cases have to be distinguished.

(i) In the first one, for $K_t < \breve{K}$, all workers vote for the lower tax rate τ^2. Only the T old agents vote for the higher tax rate τ^1. Since old agents cannot have the majority, the tax rate τ^2 is chosen by the Government.

(ii) In the second case, for $K_t > \breve{K}$, educated workers vote for τ^2, whereas uneducated workers and old agents vote for τ^1. The Government chooses the tax rate τ^2 if educated agents have the majority, i.e.:

$$\tau_t = \tau^1 \text{ if } F(\tilde{\theta}_t) < 1/2 \times (1+T)$$

and:

$$\tau_t = \tau^2 \text{ if } F(\tilde{\theta}_t) > 1/2 \times (1+T)$$

We define $\hat{\theta}$ as the threshold value for θ such that:

$$F(\hat{\theta}) = \frac{(1+T)}{2}$$

[17]For $\alpha = 0$, i.e. there is no pension system, then $\breve{K} = 0$.

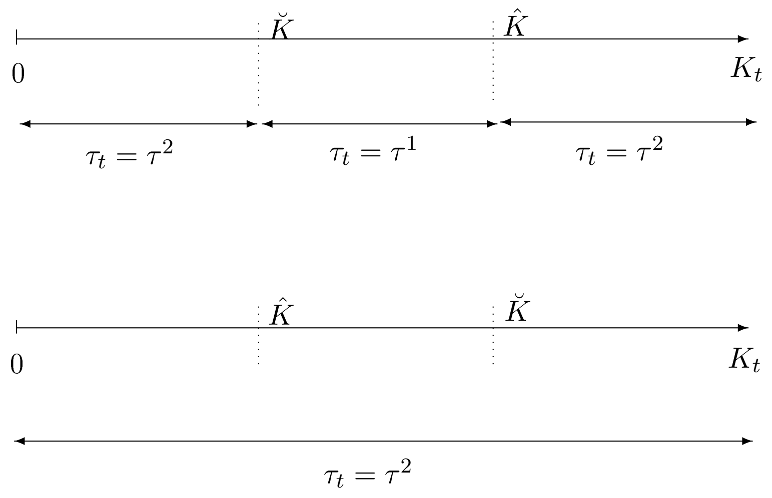

Figure 3.1: Illustration of lemma 2

Using equation (3.15), we can determine implicitly the threshold value of the capital level ($\hat{K}$) such that the previous equation is satisfied:

$$F(\theta(\hat{K})) = \frac{(1+T)}{2} \tag{3.19}$$

In the rest of this paper, we assume that $\hat{K}$ is finite and positive. If $K_t < \hat{K}$, the educated population does not have the majority. Conversely, if $K_t > \hat{K}$, educated agents represent more than half of the total voting population.

A higher T implies that there are more old agents. The share of the educated population has to be higher for them to obtain the majority. It explains why $\hat{K}$ depends positively on T.

The comparison between $\hat{K}$ and $\breve{K}$ has the following implication:

Lemma 3.2 *If $\hat{K} < \breve{K}$ then τ^1 is never chosen whatever the value taken by K_t. If $\hat{K} > \breve{K}$ then if $K_t < \breve{K}$ or $K_t > \hat{K}$ the tax rate τ^2 is chosen; but if $\breve{K} < K_t < \hat{K}$ the tax rate τ^1 is chosen.*

Proof: The result comes directly from lemma 3.1 and from the computation of $\hat{K}$. Figure 3.1 illustrates this lemma. □

In the first case ($\hat{K} < \breve{K}$), as long as $K_t < \breve{K}$, uneducated agents choose the lower tax rate and only old agents vote for the larger social security system. Once $K_t > \breve{K}$ educated agents have the majority and choose the lower tax rate, which implies that the tax rate τ^1 is never chosen. In the second case ($\hat{K} > \breve{K}$), as long as $K_t < \breve{K}$, uneducated agents choose τ^2, and then a majority vote for τ_2. For $\breve{K} < K_t < \hat{K}$, uneducated agents vote for the higher tax rate τ^1, and their coalition with old agents has the majority. It implies that in this interval τ^1 is chosen. However, for $K_t > \hat{K}$, educated agents have the majority and the tax rate τ^2 is chosen.

Let us now study the dynamics of the economy.

3.3 The Dynamics and its Properties

At the equilibrium of each period all markets clear. Then we have: $L_t^e = F(\tilde{\theta}_t)$ and $L_t^u = 1 - F(\tilde{\theta}_t)$.

The dynamics of the capital accumulation summarizes the dynamics of our economy. Capital market clears if:

$$K_{t+1} = S_t^{e,j} L_t^e + S_t^{u,j} L_t^u \tag{3.20}$$

e and u denote the educated and uneducated population respectively. j belongs to the set $\{1,2\}$. $j = 1$ $(= 2)$ if the tax rate τ^1 (τ_2) is chosen when agents work[18]. The saving function of each group of agents is:

$$S_t^{i,j} = \frac{\beta R_{t+1}\left(w_t^i(1-\tau^j) + \overline{G}_t^j\right) - 1}{\beta R_{t+1}} \tag{3.21}$$

with $i \in \{e,u\}$, $j \in \{1,2\}$. Using equations (3.11), (3.18) and (3.21) in equation (3.20), we obtain:

$$K_{t+1} = (1-\alpha\tau^j)\bar{w}(K_t) - \frac{1}{\beta X} \tag{3.22}$$

For the sake of simplicity for graphical illustrations we assume that the function $\bar{w}(k_t)$ is concave and that $\lim_{K\to+\infty} \bar{w}(K) = 0$. Furthermore, we have a unique steady state iff

[18]Let us recall that saving does not depend on pensions because of the quasi-linearity of our utility function.

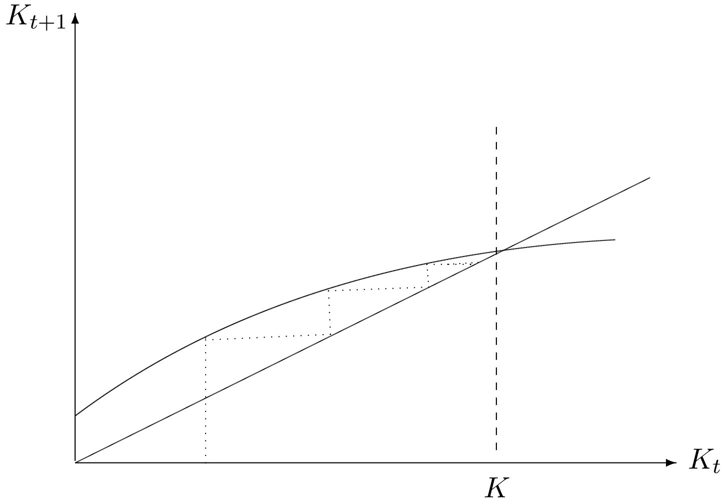

Figure 3.2: Dynamics of K_t for a single regime

$(1-\alpha\tau^1)\bar{w}(0) > \frac{1}{\beta X}$.

Let us call a "*potential steady state*" a steady state which should be obtained if an economy kept the same tax rate τ^j at every period. These steady states have the following properties:

Lemma 3.3 *There exists two potential steady states: K^1 for $\tau = \tau^1$, and K^2 for $\tau = \tau^2$. These steady states are stable, not trivial, and are such that $K^1 < K^2$.*

Proof: Because of the concavity of the RHS and because $\lim_{K\to+\infty} \bar{w}'(K) = 0$, and given that $(1-\alpha\tau^1)\bar{w}(0) > \frac{1}{\beta X}$, then there exists one and only one intersection point between the RHS and the LHS. Moreover, $K^1 < K^2$ because $\tau^1 > \tau^2$. Finally, differentiating equation (3.22) we obtain $dK_{t+1}/dK_t = RHS'(K_t)/LHS'(K_{t+1})$. The RHS of this equation is inferior to 1 in the intersection point (the steady state) because of the linearity of the LHS and the concavity of the RHS. It demonstrates the stability of the steady state (See figure 3.2).$\square$

Because $\tau^1 > \tau^2$, the steady state value of the physical capital is higher in case 2 than in case 1. Indeed, a high value of τ has a negative impact on savings because it decreases

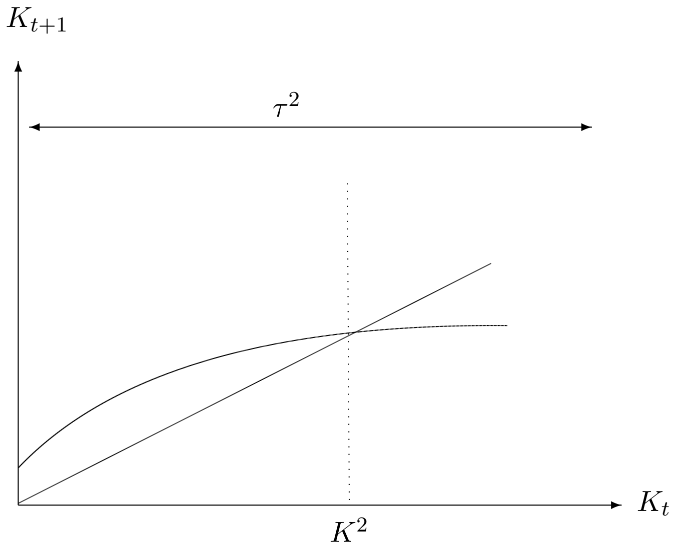

Figure 3.3: Case where $\hat{K} < \breve{K}$

the average net wage.

The dynamic behavior of the economy depends on the value taken by $\hat{K}$. It is summarized in the following proposition:

Proposition 3.1 *If $\hat{K} < \breve{K}$ then the tax rate τ^1 is never chosen and the dynamics of the economy is completely described by equation (3.22) with j=2, and K_t tends towards K^2. If $\hat{K} > \breve{K}$ then the level of capital can either tend towards K^1 or towards K^2. Moreover, the economy can adopt temporarily one of the two tax rates.*

Proof: The following figures illustrate this proposition.□

Let us first consider the case where $\hat{K} < \breve{K}$ (see figure 3.3). That case occurs if almost all taxes are dedicated to the pension system, because the average wage of the economy has to be high for uneducated agents to vote for the higher tax rate.

In this specific case the tax rate τ^1 is never chosen. Indeed, as long as K_t is lower than $\breve{K}$ only old agents vote for the higher tax rate. For $K_t > \breve{K}$ uneducated agents vote for

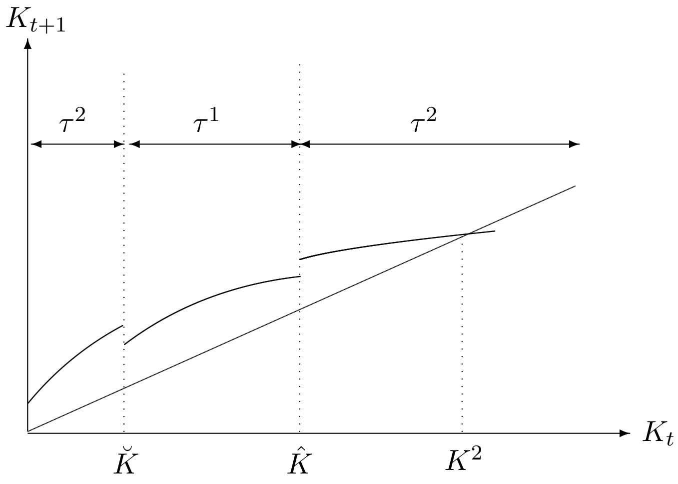

Figure 3.4: Case where $\hat{K} < K_1$

the higher tax rate but their coalition with old agents does not have the majority. Consequently, whatever the value of K_t the dynamics of the economy is described by equation (3.22) with $\tau_t = \tau^2$, $\forall t$.

Right now we assume that $\hat{K} > \breve{K}$.

If $\hat{K} < K^1$ (see figure 3.4), then the capital tends towards K^2. That case occurs if the share of tax revenues dedicated to pensions is not too high and if a large fraction of the population has low educational costs ($\hat{K}$ small). That case shows that in a finite time the economy will adopt the tax rate τ^2. More precisely, if $K_0 < \breve{K}$, then in the initial period the tax rate τ^2 is chosen because even uneducated agents do not vote for the larger welfare state. Then for $\breve{K} < K_t < \hat{K}$, uneducated and old agents have the majority and vote for the tax rate τ^1. This occurs as long as $K_t < \hat{K}$. But in a finite time, agents save more because of the lower tax rate. In every following period, agents know that the educated population has the majority and that the tax rate remains at the level τ^2. In that case, high tax rates are only transitional equilibria.

If $K^2 > \hat{K} > K^1 > \breve{K}$ (see figure 3.5), then the steady state value of the capital crucially depends on the initial value K_0. If $\breve{K} < K_0 < \hat{K}$,the value of K is sufficiently low for uneducated and old agents to represent the majority. Then agents save less. The economy

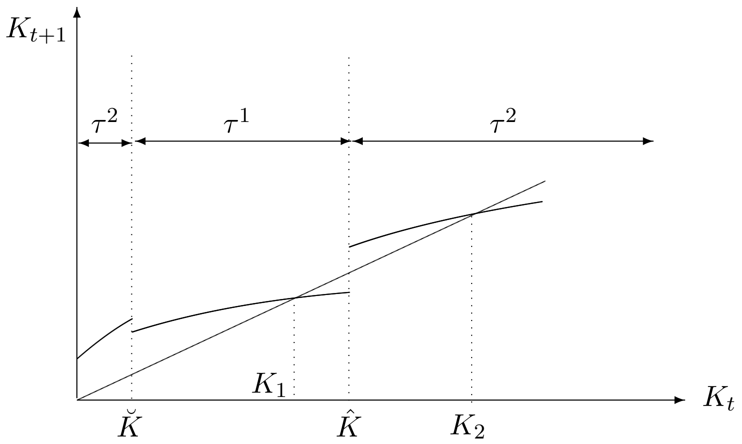

Figure 3.5: Case where $K_1 < \hat{K} < K_2$

behaves in the same way at every period and the capital tends toward K^1. Conversely, if $K_0 > \hat{K}$, the capital tends toward K_2 because every agent expects the educated population to represent the majority. That case occurs if the economy can reach a steady state before the educated population gets the majority.

Finally, if $\breve{K} < K^1 < K^2 < \hat{K}$ (see figure 3.6), then the tax rate τ^2 cannot be chosen at steady state. Indeed the capital tends towards K^1 whatever its initial value. In a finite time the coalition of old and uneducated agents obtain the majority and vote for the higher tax rate. It decreases savings and consequently the steady state value of capital. That case occurs if educational costs are too high for a large fraction of the population. Indeed, the educated population becomes majoritarian for a too high value of the stock of capital.

3.4 Concluding Remarks

In this chapter, we show that an economy can converge towards two kinds of equilibria. The first one is such that the capital level is low, which implies that uneducated and old agents have the majority and vote for a large welfare state. But the economy can also converge towards another steady state in which the capital level is high. In that case a large fraction of agents decides to educate itself. Consequently educated agents have the majority and vote for a small welfare state. In this model, high tax rates can be transitional equilibria. It is possible to observe three stages in the dynamics of the welfare state: first

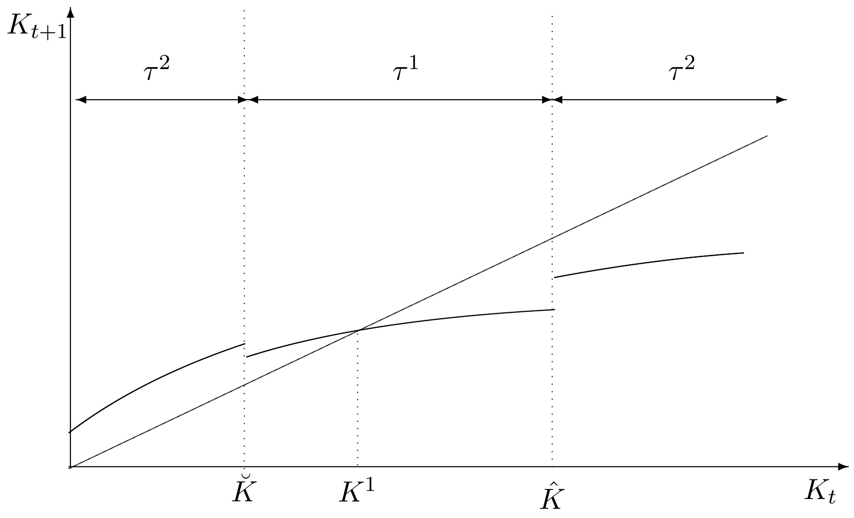

Figure 3.6: Case where $\breve{K} < K^1 < K^2 < \hat{K}$

a low tax rate, then a high tax rate and finally a low tax rate.

This chapter is a first step in order to better understand the recent reforms of social security systems as the result of new political pressures. We emphasize the role of the structure of the working population but other factors can play an important role, notably the development of financial markets. In this new context, agents probably find that these instruments are more efficient than the Government to finance their old age consumption.

Part II

The Macroeconomic Impacts of the Instantaneous Redistributivity of Pension Systems

Chapter 4

Redistribution, Pension Systems and Capital Accumulation

4.1 Introduction

Pension systems can be classified according to three dimensions (Feldstein and Liebman 2002, Villa 2004). Firstly, they can adopt either a Pay-As-You-Go (PAYG) or a fully-funded structure[1]. The size of unfunded pension systems is large in most industrialized countries. For example, the payroll tax rate used to finance them ranges from 12.4% for the United-States to 29.6% for Italy (Nyce and Schieber 2005, pp.236).

Secondly, pension systems can have either a defined-benefit or a defined-contribution structure. A pension system has a defined-benefit structure if it is the tax rate which adjusts itself to changes in the economic and demographic environment. Conversely, it has a defined-contribution organization if it is the replacement rate which adjusts itself. Most countries have chosen a defined-benefit pension system (Nyce and Schieber 2005). However, because of the increase in life expectancy, the fiscal burden of this structure has increased strongly. Consequently, some countries, such as Italy, have adopted a defined-contribution pension system[2].

Thirdly, pension systems can be more Beveridgian or more Bismarckian. A pension system is purely Beveridgian if every agent receives the same pension. Conversely, a pension system is purely Bismarckian if pensions depend completely on the wages of agents. A pension system is mixed if it has a Beveridgian and a Bismarckian component. The more a pension system is Beveridgian, the higher intra-generational transfers are. Countries highly differ by this intra-generational component. France, Germany and Italy have a Bismarckian structure. Canada, the Netherlands and New-Zeland are essentially Beveridgian. Finally, Japan, the United-Kingdom and the United States have mixed pension systems (Sommacal 2006, Casamatta *et al.* 2000).

Theoretical literature has explored the impact on the economic activity of the size of PAYG pension systems[3] organized either with a Beveridgian or with a Bismarckian structure[4]. The usual result is that Bismarckian systems provide more incentives to accumulate

[1]In this chapter we only consider PAYG pension systems. In a portofolio approach Dutta *et al.* (2000) show that a mixed pension system is desirable because risks of funded and unfunded pension systems differ.

[2]In this chapter we do not explain the switch from a defined-benefit to a defined-contribution pension system. A model with a representative agent and an increasing life expectancy would be more appropriate for this study.

[3]See Belan and Pestieau (1998), Breyer and Straub (1993) or Homburg (1990) among others for the analysis of the transition from unfunded pension systems towards fully-funded pension systems.

[4]See Docquier and Paddison (2003), or Casarico and Devillanova (2007). These results are questioned

human and physical capital and then induce a higher growth rate than Beveridgian pension systems. But in fact, pension systems are usually a combination of these two elements. Only a few authors have studied the impacts of a change along the third axis mentioned above[5]. However, it is a central issue given the wide dispersion of countries along this third axis.

The main idea of this chapter is that PAYG pension systems can adopt a structure which combines Bismarckian and Beveridgian components, and we study the impacts of a policy which increases the Beveridgian component of pension systems. We show that these effects are different depending on whether the pension system has a defined-benefit or a defined-contribution structure. Last but not least, the inequalities of length of life play an important role in the qualitative and quantitative results of this chapter.

There is a growing empirical literature which analyzes these inequalities. Mesrine (1999) studies the inequalities of length of life according to socio-professional groups in France[6]. The most striking feature of his paper is that a worker has a probability to die between 35 and 65 years old almost twice higher than that of an executive manager. Furthermore, their life expectancy at 35 is 38 and 44 respectively. The same qualitative results are observed in the United-States (Panis and Lillard 1995, Deaton and Paxson 2000).

Finally, Robert-Bobbée and Cadot (2007) show that this inequality is also observed for elderly people. For agents who are 86, the ones with highest education level can expect to live 20% longer than the ones with lowest education level.

Only a few papers have explored the economic impacts of these health inequalities. Mitchell and Zeldes (1996, pp.365) emphasized that these health inequalities have implications on the redistributive properties of pension systems but they do not provide any empirical or analytical analysis. Drouhin (2001a) showed with a small open economy that a Bismarckian PAYG pension system induces transfers from agents with a short life expectancy to agents with a long life expectancy. His model is a first step in order to study the impacts of the inequalities of length of life but it uses only a Bismarckian structure and there are no general equilibrium effects in his model. The political economy literature

notably by Groezen *et al.* (2007), Lambrecht *et al.* (2005) or Le Garrec (2005).

[5]Except Sommacal (2006) with an endogenous labor supply model with a defined-contribution pension system.

[6]These inequalities also depend on other factors like sex or the geographical localization. For example, in France the life expectancy of women is 84.1, whereas that of men is only 77.2 (INSEE, 2006). Moreover, Rican and Salem (1999) show that there are strong disparities according to the localization of people in France.

has recently become interested in the implications of the link between life expectancy and wages[7].

In this chapter we study the macroeconomic impact of a policy which modifies the redistributive properties of an unfunded pension system[8]. In order to obtain clear qualitative results for every macroeconomic variable we first give an analytical resolution of our model. Then, because the impact is ambiguous for some variables, we calibrate our model on French data and we numerically solve our model. We work on French data because the French pension system is highly Bismarckian (Casamatta *et al.* 2000) and because the efficiency of such a system is widely questioned. However, we also show that our numerical results do not depend on this specific case. Using an overlapping generations model with a closed economy and heterogenous agents, we show that a weaker link between contributions and benefits has an impact on the level of capital per capita if and only if there are inequalities of length of life. We also show that this redistributive policy has positive implications for every agent of the economy if the system has a defined-benefit structure. The tax rate and inequalities decrease, whereas the wealth of each agent increases. However, with a defined-contribution pension system, this policy has a negative impact on every macroeconomic variable except on the wealth of the poorest agents.

Gorski *et al.* (2007) also emphasized the role of the mortality differential to analyze the impact on educational choices of a change towards a more Beveridgian pension system. They find that this impact is positive. In this chapter, we analyze the impact of this policy on physical capital accumulation.

This chapter is organized as follows. Section 2 presents the main elements of our model. In section 3 we detail the dynamics of the economy and its properties. The implications in terms of utility and inequalities are studied in section 4. In section 5 we calibrate and solve our model. Finally, section 6 includes some concluding remarks.

---

[7]Borck (2007) shows that the size of a pension system can be determined by a coalition of elderly, very poor and very rich agents. Poor agents benefit from the Beveridgian part of the pension system, whereas rich agents benefit the longest time from the pension system.

[8]In this chapter the term "redistributivity" means that we change the Bismarckian structure of pension systems. A decrease in the redistributivity means that there is a stronger link between wages and pensions per unit of time.

4.2 The Model

At each period t, it is assumed that two generations overlap: the young and the old. Their respective size are N_t and N_{t-1}. The population grows at a constant rate $n > -1$, such that $N_t = (1+n)N_{t-1}$. Each member of one generation receives at the beginning of his life a productivity endowment a. This productivity takes its values in the interval $\Omega_a = [a_-, a_+]$. The density function and the cumulative distribution function of a are denoted by $f(a)$ and $F(a)$ respectively. These functions are such that: $\int_{\Omega_a} f(a)da = 1$, $F(a_-) = 0$ and $F(a_+) = 1$. Furthermore, $\bar{a}$ denotes the average productivity of the economy:

$$\bar{a} = \int_{\Omega_a} af(a)da \tag{4.1}$$

The density function $f(a)$ is assumed to be independent of time and of the level of capital.

Each agent lives completely his first period of life[9] but only a fraction $T(a) \in (0,1)$ of his second period of life[10]. We assume that $T'(a) > 0$. The higher the productivity is, the longer the length of life is. In doing so we assume that the length of life depends positively on the productivity level of each agent. In our model, the wage level is an increasing function of the productivity level. Consequently, the assumption on $T(a)$ uses the empirical evidence that the wage level is a significant variable to explain the mortality differential between agents (Adams *et al.* 2003)[11]. Borck (2007) uses the same assumption in a political economy framework.

The average length of life is denoted by $\bar{T}$ and is determined by:

$$\bar{T} = \int_{\Omega_a} T(a)f(a)da \tag{4.2}$$

The link between productivity and length of life is measured by the covariance:

$$COV_{T(a),a} = \int_{\Omega_a} T(a)af(a)da - \bar{T}\bar{a} \tag{4.3}$$

[9] The length of each period is normalized to 1.

[10] There is no uncertainty in this economy to simplify our model. However our model can also be interpreted as a model with uncertain lifetime. In that case $T(a)$ is the probability that an agent survives at the end of his first period of life as in Drouhin (2001a). It also implies that there exists a perfect annuity market for each group of agents.

[11] See also Mesrine (1999), Pannis and Lillard (1995), or Deaton and Paxson (2000).

This covariance is positive because of our assumption on the sign of $T'(a) > 0$ (See appendix E). The stronger the link between $T(a)$ and a is, the bigger this covariance is. Conversely, if $T'(a) = 0$, i.e. if the length of life is the same for every agent, then this covariance is null.

4.2.1 Consumers

The utility of consumers depends on their consumption flows of their two periods of life. For an agent born in period t endowed with a productivity level a, $c_t(a)$ and $d_{t+1}(a)/T(a)$ denote the first period and the second period consumption flows respectively. Their utility function is intertemporally separable and has the following form[12]:

$$U_t(a) = u(c_t(a)) + \beta T(a) u\left(\frac{d_{t+1}(a)}{T(a)}\right) \tag{4.4}$$

where $\beta \in (0, 1)$ represents the pure time preference factor for the present, and the $T(a)$ in front of their second period utility implies that the longer the length of life is, the more consumers value their utility of this period[13].

Each agent offers inelastically his work during his first period of life and obtains a wage $w(a)$[14]. This wage is taxed at a rate τ, and the revenues of this tax are used to finance a PAYG pension system. When an agent becomes old he receives a pension p(a). For an agent born in period t, his budget constraints are:

$$c_t(a) = w_t(a)(1 - \tau) - S_t(a) \tag{4.5}$$

$$d_{t+1}(a) = R_{t+1} S_t(a) + p_{t+1}(a) \tag{4.6}$$

with R_{t+1} the interest factor and $S_t(a)$ the saving function.

We also assume that the utility function has the following form : $u(x) = \ln(x)$. It simplifies the analytical expressions[15]. Using all these assumptions, the saving function is

[12]See the appendix of chapter 1 for a justification.

[13]In this chapter we do not represent fertility choices even if the life expectancy differential and wage inequalities have an impact on these choices.

[14]In doing so we do not model the burden of income taxation on labor supply.

[15]It notably simplifies the conditions that will be obtained and the aggregation of the saving functions. Our qualitative results do not depend on this assumption.

the following[16]:

$$S_t(a) = \frac{\beta T(a) w_t(a)(1-\tau)}{(1+\beta T(a))} - \frac{p_{t+1}(a)}{(1+\beta T(a))R_{t+1}} \tag{4.7}$$

4.2.2 Firms

We assume a perfect competition on the final good market and on the inputs markets. The production function of firms is[17]:

$$Y_t = AK_t^{\alpha}\left(N_t \int_{\Omega_a} af(a)da\right)^{1-\alpha} \tag{4.8}$$

with $0 < \alpha < 1$, K_t the physical capital level, and $A > 0$ the level of the technology. As there is perfect competition on each market, firms take wages and interest factors as given. Profit maximisation implies the following expressions for prices given that the final good is the *numéraire*:

$$R_t = A\alpha K_t^{\alpha-1}\left(\int_{\Omega_a} af(a)N_t da\right)^{1-\alpha} \equiv A\alpha k_t^{\alpha-1}\bar{a}^{1-\alpha} \tag{4.9}$$

$$w_t = A(1-\alpha)K_t^{\alpha}\left(\int_{\Omega_a} af(a)N_t da\right)^{-\alpha} \equiv A(1-\alpha)\frac{k_t^{\alpha}}{\bar{a}^{\alpha}} \tag{4.10}$$

with $k_t \equiv K_t/N_t$, the capital level per young agent. w_t is the wage per efficiency unit of work. For agents with a productivity level a, their wage is:

$$w_t(a) = w_t a = A(1-\alpha)\frac{k_t^{\alpha}}{\bar{a}^{\alpha}}a \tag{4.11}$$

It implies that relative wages are independent of the level of capital, whereas absolute differences of wages depend on it.

In the rest of this chapter, $\bar{w}_t$ will denote the average wage of the economy at period t. It has the following expression:

$$\bar{w}_t = \int_{\Omega_a} w_t(a)f(a)da = A(1-\alpha)k_t^{\alpha}\bar{a}^{1-\alpha} \tag{4.12}$$

[16]In appendix H we show that, at equilibrium, agents having high wages save, whereas agents having low wages can resort to borrowing. It emphasizes the main role played by the absence of liquidity constraints in our model.

[17]Our results do not depend on the form of the production function but it clarifies our analysis with simple analytical results.

4.2.3 The Pension System

We assume a PAYG pension system. The revenues of this system come from a proportional tax on wages : τ. It is used to provide a pension for elderly people. Their pension depends on the wages of young agents having the same productivity as theirs, and on the average wage of the economy. Their respective weighting is λ and $(1-\lambda)$. The first part of this pension represents the Bismarckian component, whereas the second part represents the Beveridgian component of this system (Casamatta *et al.* 2000). λ measures the indexation of pensions on wages of agents. **The smaller λ is, the more this pension system is redistributive.**[18].

Consumers receive only a fraction ν (with $0 < \nu \leq 1$) of this weighted average, and only during their second period of life $T(a)$. ν denotes the average replacement rate of the pension system. The pension of an agent endowed with a productivity level a is:

$$p_{t+1}(a) = \nu\left(\lambda w_{t+1}(a) + (1-\lambda)\bar{w}_{t+1}\right)T(a) \tag{4.13}$$

With equations (4.11) and (4.12) we obtain:

$$p_{t+1}(a) = \nu A(1-\alpha)\frac{k_{t+1}^{\alpha}}{\bar{a}^{\alpha}}(\lambda a + (1-\lambda)\bar{a})T(a) \tag{4.14}$$

We also assume that the Government does not use debt. It implies that for every period we have:

$$\int_{\Omega_a} \tau w_{t+1}(a) f(a) N_{t+1} da = \int_{\Omega_a} p_{t+1}(a) f(a) N_t da \tag{4.15}$$

We show in appendix A that we obtain the following expression:

$$\tau = \nu\frac{\bar{T}}{1+n}\left(1+\lambda\rho^2\right) \tag{4.16}$$

It defines the tax rate in function of the parameters of the model. We say that it characterizes a defined-benefit organization. ρ^2 denotes the covariation coefficient between life expectancy and the productivity level, i.e. $\rho^2 = \frac{COV_{T(a),a}}{\bar{T}\bar{a}}$.

[18]In this chapter the term "redistributivity" only concerns the instantaneous redistribution of pension systems and not the long run redistribution of pension systems. The long run redistribution, which is the discounted difference between tax paid and amount received, can be very different because of the life expectancy differential.

Proposition 4.1 *With a defined-benefit PAYG pension system τ is an increasing function of λ*[19] *if and only if $COV_{T(a),a} > 0$, i.e. if there are inequalities of length of life.*

This result is very intuitive. Indeed, the richer agents are, the longer their length of life is. Therefore, an increase in λ (i.e. a decrease in the redistributivity of the pension system) increases the indexation of pensions on their wages. It implies that the pension of rich agents increases. Moreover, they benefit from these pensions for a longer period of time than other agents. Consequently, the tax rate has to increase to finance these additional expenditures.

We have to note that this result depends only on the budget constraint of the Government and not on the preferences of consumers.

Let us now assume that we have a defined-contribution PAYG pension system (τ is exogenous). It is the replacement rate ν which adjusts itself in order to maintain the Government budget constraint at equilibrium:

$$\nu = \tau \frac{(1+n)}{\bar{T}\,(1+\lambda\rho^2)} \tag{4.17}$$

Proposition 4.2 *With a defined-contribution PAYG pension system, the replacement rate (ν) is a decreasing function of λ if and only if $COV_{T(a),a} > 0$.*

The intuition is the same as before. A smaller indexation on wages (a smaller λ) benefits to poor agents who live for a shorter period of time than rich ones. Then, for a given replacement rate, expenditures are lower. Finally, Government can increase the replacement rate for every agent.

Corollary 4.1 *With a defined-benefit (defined-contribution) PAYG pension system, the tax rate (replacement rate) is independent of the redistributivity of the pension system if and only if there are not any inequalities of length of life.*

Without inequalities of length of life a variation in λ does not affect the total amount of pensions which are paid.

[19] This proposition can partly explain why Bismarckian pension systems are bigger than Beveridgian ones.

4.3 The Dynamics and its Properties

The dynamics of this economy is represented through the equation of capital accumulation. Furthermore, because the marginal return of capital in the production function is decreasing, the economy converges towards a steady state equilibrium such that the capital level per worker is constant. The dynamics is the following:

$$K_{t+1} = \int_{\Omega_a} S_t(a) f(a) N_t da \tag{4.18}$$

It is straightforward to show that we finally obtain:

$$k_{t+1}\left[1+n+\nu\frac{1-\alpha}{\alpha}\int_{\Omega_a}\frac{\lambda\frac{a}{\bar{a}}+(1-\lambda)}{1+\beta T(a)}T(a)f(a)da\right] = \frac{\beta A(1-\alpha)(1-\tau)k_t^{\alpha}}{\bar{a}^{\alpha}}\int_{\Omega_a}\frac{T(a)a}{1+\beta T(a)}f(a)da \tag{4.19}$$

The RHS of this equation is a strictly concave function of k_t. Consequently, there is a unique non-trivial steady state which has the following form:

$$(k^*)^{1-\alpha} = \frac{\frac{\beta A(1-\alpha)(1-\tau)}{\bar{a}^{\alpha}}\int_{\Omega_a}\frac{T(a)a}{1+\beta T(a)}f(a)da}{1+n+\nu\frac{1-\alpha}{\alpha}\int_{\Omega_a}\frac{\lambda\frac{a}{\bar{a}}+(1-\lambda)}{1+\beta T(a)}T(a)f(a)da} \tag{4.20}$$

Proposition 4.3 *With a defined-benefit PAYG pension system, a decrease in λ has a positive impact on k^*.*

Proof: The numerator of equation (4.20) is a decreasing function of λ because only τ depends positively on λ. Moreover, we know that $T(a)/(1+\beta T(a))$ is an increasing function of a. It implies that $T(a)/(1+\beta T(a)) < T(\bar{a})/(1+\beta T(\bar{a}))$ $(>)$, $\forall a < \bar{a}$ $(>)$. Then, $(a-\bar{a})T(a)/(1+\beta T(a)) > T(\bar{a})/(1+\beta T(\bar{a}))(a-\bar{a})$, $\forall a$. The denominator is an increasing function of λ if the following condition is satisfied : $\int_{\Omega_a}\frac{a-\bar{a}}{1+\beta T(a)}T(a)f(a)da \geq 0$. We know that $\frac{a-\bar{a}}{1+\beta T(a)}T(a)f(a) \geq \frac{a-\bar{a}}{1+\beta T(\bar{a})}T(\bar{a})f(a)$, $\forall a \in \Omega_a$. Integrating the two sides of this equation on the interval Ω_a, the RHS is equal to zero and the condition mentioned above is satisfied.□

Two kinds of effects play a role when we analyse the effects of a decrease in λ. The former concerns the impact on the tax rate. Indeed, we have showed in proposition 1 that the tax rate is an increasing function of λ. If λ falls, the tax rate decreases for every consumer, which has a positive effect on savings without ambiguity. The latter concerns the impact on the pension received by each agent. If λ decreases, consumers with a productivity lower than $\bar{a}$ receive a greater pension, whereas consumers with a productivity higher than $\bar{a}$ receive a smaller pension. The first group of agents saves less and the second one saves more. Proposition (2) shows that the net effect on savings is positive. Indeed, agents for whom the pension decreases have a longer length of life than the others. Consequently, the increase in the savings of rich agents overcompensates the decrease in the savings of poor agents.

Proposition 4.4 *With a defined-contribution PAYG pension system, a decrease in λ has a positive impact on k^* if and only if:*

$$\frac{\int_{\Omega_a} \frac{1}{1+\beta T(a)} T(a)f(a)da}{\int_{\Omega_a} T(a)f(a)da} \leq \frac{\int_{\Omega_a} \frac{a-\bar{a}}{1+\beta T(a)} T(a)f(a)da}{\int_{\Omega_a} (a-\bar{a})T(a)f(a)da} \tag{4.21}$$

Proof: τ is fixed because it is a defined-contribution pension system. It is ν which adjusts itself and only the last term of the denominator depends on λ. The condition ensures that the derivative of this term with respect to λ is positive.□

In a defined-contribution PAYG pension system, we have showed in proposition 1 (bis) that ν is a decreasing function of λ. Then, following an increase in the redistributivity of the pension system (a decrease in λ), the Government increases the replacement rate. It has a positive impact on the pension of every consumer *ceteris paribus*, and thus a negative effect on savings. But the decrease in λ has a positive (negative) impact on the savings of agents endowed with a productivity higher (smaller) than $\bar{a}$. The condition of the proposition ensures that the positive effect is higher than the two negative ones.

Proposition 4.5 *(i) If there are no inequalities of length of life then k^* does not depend* *on λ[20]. (ii) This result remains true for every homothetic preference*[21].

Proof: See appendix B.□

We have showed with proposition 1 and 1 (bis) that if $T(a) = T$ for all a, then the tax rate (replacement rate) is independent of λ. The only effects concern the increase in the savings of agents endowed with a productivity higher than $\bar{a}$, and the decrease in the savings of agents endowed with a productivity lower than $\bar{a}$. These last two effects exactly compensate.

4.4 Wealth, Consumption and Redistribution

This section has two main objectives. The first one is to study the evolution of the wealth, of the consumption and of the utility of an agent if the degree of redistribution of the pension system increases (λ decreases). The second one is to study the evolution of inequalities of consumption and of welfare if λ decreases.

These analytical results are obtained at steady state to simplify the exposition. Every derivative is thus a comparison between steady states.

4.4.1 Wealth, Welfare and Redistribution

The wealth of an agent born in period t endowed with a productivity level a, has the following form:

$$W_t(a) = w_t(a)(1-\tau) + \frac{p_{t+1}(a)}{R_{t+1}} \tag{4.22}$$

We want to know if the wealth of each consumer increases when the redistribution of the pension system is higher (λ decreases).

Proposition 4.6 *With a defined-benefit pension system, if λ decreases then the wealth of agents endowed with a productivity smaller than $\bar{a}$ increases, whereas the impact on the*

[20] Casaricco and Devillanova (2008) obtain this qualitative result in an economy with capital accumulation and educational choices.

[21] This result is questioned for non-homothetic preferences. This result depends on the linearity of the saving function with respect to wages. For an analysis of the concavity of the consumption function with respect to the wage level, see Drouhin (2001b).

wealth of other agents is ambiguous. The net effect is positive for every agent if[22]*:*

$$-\frac{\frac{dk}{d\lambda}}{k} \geq \frac{\frac{a+}{\bar{a}} - 1}{\lambda\frac{a+}{\bar{a}} + 1 - \lambda} \tag{4.23}$$

Proof: See appendix C.□

Proposition 1 has showed that the tax rate is an increasing function of λ. Furthermore, we have showed with proposition 2 that k^* is a decreasing function of λ. Then the net wage of the first period of life is higher when the redistributivity of the pension system increases. More generally, wages per efficiency unit of work increase.

Moreover, a decrease in λ reduces the indexation of pensions on wages. Consequently, it has a positive impact on the pensions of agents endowed with a productivity smaller than $\bar{a}$ and a negative effect on pensions of agents endowed with a productivity higher than $\bar{a}$. The condition in the proposition ensures that for rich agents ($a > \bar{a}$) all positive effects overcompensate the decrease in the indexation of pensions on wages.

Proposition 4.7 *With a defined-contribution pension system, if λ decreases then:*

- *If the condition of proposition 2 (bis) is true then the wealth of agents endowed with a productivity smaller than $\bar{a}$ increases, whereas the impact on the wealth of other agents is ambiguous. The net effect is positive for every agent if*[23]*:*

$$-\frac{\nu\frac{dk}{d\lambda} + k\frac{d\nu}{d\lambda}}{k} \geq \frac{\frac{a+}{\bar{a}} - 1}{\lambda\frac{a+}{\bar{a}} + 1 - \lambda} \tag{4.24}$$

- *Otherwise, the net impact is ambiguous for every consumer.*

Proof: See appendix C.□

If the condition of proposition 2 (bis) is true then a decrease in λ has a positive impact on k^*. Furthermore, it affects the pension of agents differently depending on whether consumers have a productivity higher or lower than $\bar{a}$. The effects are the same as before except that τ is fixed exogenously. Every agent benefits from the increase in ν and more

[22]It is a sufficient condition.
[23]It is a sufficient condition.

particulary agents with a long life expectancy. That is why the condition is less restrictive than that of proposition 4. Nevertheless, if the effect on k^* is negative then the impact on the wealth is ambiguous for every consumer.

The utility of an agent depends on the level of consumption of the two periods of his life. Using the budget constraints of consumers we obtain:

$$c_t(a) = \frac{W_t(a)}{1+\beta T(a)} \tag{4.25}$$

and

$$d_{t+1}(a) = \beta T(a) R_{t+1} \frac{W_t(a)}{1+\beta T(a)} \tag{4.26}$$

The consumption level of his first period of life depends on λ only through the wealth level, whereas the consumption level of his second period of life depends on the wealth level and on the interest factor. The utility level is an increasing function of the redistributivity of the pension system if and only if:

$$-(1+\beta T(a)) \frac{dW(a)/d\lambda}{W(a)} > -(1-\alpha)\beta T(a) \frac{dk/d\lambda}{k} \tag{4.27}$$

The LHS represents the evolution of the wealth of an agent and the RHS the evolution of the interest factor. Indeed, a change in λ affects k^* and thus the interest factor. Let us consider the case of a defined-benefit pension system. A decrease in λ has a positive impact on the wealth of every consumer ($dW(a)/d\lambda < 0$). But at the same time it reduces the interest factor ($dk^*/d\lambda < 0$). The net effect on utility is thus ambiguous. More precisely, the net effect can be negative for agents with a long life expectancy because they save a large part of their wealth and are strongly affected by the decrease in the interest factor.

4.4.2 Inequalities and Redistribution

To study inequalities, two groups of agents are used: the poorest endowed with a productivity level a_- and the richest endowed with a productivity level a_+[24]. The main objective is to study welfare inequalities, but the relative inequalities of wealth have to be studied first.

[24]We do not use here the Gini coefficient for analytical convenience. See section 5 for an estimation of the Gini coefficient in our model.

Proposition 4.8 *With a defined-benefit pension system, the relative inequality of wealth $W(a_-)/W(a_+)$ is an increasing function of the redistributivity of the pension system (a decrease in λ) if*[25]*:*

$$\frac{T(a_-)}{a_-} \geq \frac{T(a_+)}{a_+} \tag{4.28}$$

Proof: See appendix D.□

The direct impact of a decrease in λ is to reduce the pensions of rich agents ($a > \bar{a}$) and to increase these of poor agents ($a < \bar{a}$). It increases the ratio $W(a_-)/W(a_+)$. Moreover, a decrease in λ has a positive effect on net wages because of its positive impact on capital *per capita* and because of its negative impact on the tax rate. This effect benefits essentially to the richest. Finally, a decrease in λ has a positive impact on w_t/R_t. The richest are the ones who essentially benefit from this effect because they live for a longer period of time. The condition of the proposition ensures that the redistributive effect dominates every other.

Proposition 4.9 *With a defined-contribution pension system, if the condition of proposition 2 (bis) is true then the relative inequality of wealth $W(a_-)/W(a_+)$ is an increasing function of the redistributivity of the pension system (a decrease in λ) if*[26] *:*

$$\frac{T(a_-)}{a_-} > \frac{T(a_+)}{a_+} \times \frac{\lambda\frac{a_+}{\bar{a}} + 1 - \lambda}{\lambda\frac{a_-}{\bar{a}} + 1 - \lambda} \tag{4.29}$$

Proof: See appendix D.□

The interpretation is the same as before except that τ is fixed exogenously and that λ has a negative impact on ν. The increase in the replacement rate benefits essentially to agents with a long length of life, i.e. to rich agents. Condition (4.29) is therefore more restrictive than condition (4.28) and cannot be true for a λ which tends towards 1.

The study of welfare inequalities can now be done. These inequalities can be measured as the difference between the utility of the richest ($U(a_+)$) and the utility of the poorest ($U(a_-)$). Analytically it has the following form:

[25]It is a sufficient condition

[26]It is a sufficient condition

$$U(a_+) - U(a_-) = (1 + \beta T(a_+))\ln(W(a_+)) - (1 + \beta T(a_-))\ln(W(a_-)) +$$
$$\beta(T(a_+) - T(a_-))(\alpha - 1)\ln(k) + cste \quad (4.30)$$

If the redistributivity of the pension system increases (λ decreases), the previous differential decreases if and only if:

$$-(1+\beta T(a_+))\frac{dW(a_+)/d\lambda}{W(a_+)} - \beta(T(a_+) - T(a_-))(\alpha-1)\frac{dk/d\lambda}{k} < -(1+\beta T(a_-))\frac{dW(a_-)/d\lambda}{W(a_-)} \quad (4.31)$$

This equation is useful because it details the different channels through which the redistributivity has an impact on the utility differential. Let us study the case of a defined-benefit pension system. First, let us assume that the condition of proposition 5 is true. Then we have showed that the wealth ratio ($W(a_-)/W(a_+)$) is an increasing function of the redistributivity of the pension system, i.e.:

$$-\frac{dW(a_+)/d\lambda}{W(a_+)} < -\frac{dW(a_-)/d\lambda}{W(a_-)} \quad (4.32)$$

Condition (4.31) is more restrictive if, for the moment, we neglect the impact on the interest rate. Indeed, the richest can benefit from their wealth for a longer period of time. Then the decrease in the wealth inequalities does not necessarily imply a decrease in the utility differential. Nevertheless, the LHS also shows that the decrease in the interest rate affects more strongly the richest who save more because of their high length of life. This last effect reduces the utility differential.

4.5 Calibration and Results

We choose to calibrate our model on French data because the French pension system is clearly Bismarckian. As it will be mentioned later, Hairault and Langot (2008) find that λ in the French pension system is 0.885. Then we can consider the opportunity to switch towards a more Beveridgian pension system.

The availability of data thanks to the study of Hairault and Langot (2008) is also a main factor which has influenced our choice to consider the French case[27].

First of all we have to define an interval for the set Ω_a. We assume that it is: $\Omega_a = [0.08, 1]$. The ratio a_+/a_- is 12.5. It implies that the wage inequality ratio between the

[27] Appendix G sums up our calibration and the main statistics.

poorest and the richest is 12.5. Piketty (2002), studying the distribution of wages in France, finds a ratio of 5 between the wages of the first and of the last decile. The gap between this empirical fact and our calibration can be explained by the fact that we use the two extreme values of a *continuum* and as a consequence wage inequalities are greater. We could even say that it underestimates the reality. We choose this interval for Ω_a because once it is combined with the density function of a, our model matches the Gini coefficient of the wage distribution calculated by Hairault and Langot (2008) on French data.

The density function of productivity levels ($f(a)$) has to respect the essential property: mode<median<mean (Lambert 2001, pp.23). This property is a common feature of most industrialized countries. It implies that the wage distribution among the population is asymetric. The most common income level is less than the median wage. And because of strong wage inequalities the median wage is less than the average wage of the economy.

$f(a) = b - ca$, with $b, c \in \mathbb{R}$ is the simplest way to represent it. b and c have to be fixed such that: $f(a) > 0$, $\forall a$ and $\int_{\Omega_a} f(a)da = 1$. Furthermore the Gini index has to tend towards 0.32 in order to match the estimation on French data used in Hairault and Langot (2008).

Lambert (2001) shows that the Gini index can be calculated as:

$$G = -1 + 2\int_{a_-}^{a_+} \frac{aF(a)f(a)}{\bar{a}} da \tag{4.33}$$

The following density function respects these properties:

$$f(a) = 2.1129 - 1.9a \tag{4.34}$$

Moreover we can check that the mean is higher than the median because $\int_{a_-}^{\bar{a}} f(a)da > 0.5$.

The second important function that we have to specify is $T(a)$. To simplify and because of the lack of information we assume that this function has the form: $T(a) = a$. We obtain that $\bar{T} = 0.4167$ and that $COV_{T(a),a} = 0.05533$. It implies that the average length of life of the population is 77 years old. It is slightly lower than the average life expectancy observed in France which is 80 years old (World Bank)[28].

The initial value of λ is fixed at 0.885. It is the estimation obtained by Hairault and Langot (2008) on French data. It implies that the French pension system is highly

[28] Appendix F shows that it has no impact on our qualitative results.

Bismarckian. The growth rate of the population is $n = 0.3$. It corresponds to an annual growth rate of the population of 0.65% calculated by Charpin (1999) on French data. The technology parameter A is normalized to 1.

Finally, the last two parameters are common to a wide economic literature which uses calibration to solve overlapping-generations models. The length of each period is 40 years. The elasticity of the production function with respect to capital is $\alpha = 0.33$. It also represents the share of capital in total output. The pure time preference factor is $\beta = 0.6$ (d'Autume 2003), i.e. an annual psychologic discount rate of 1.3%.

We analyse the effects of a decrease in λ, i.e. an increase in the Beveridgian part of the pension system[29]. We distinguish between the long run effects and the transitional dynamics for defined-benefit and for defined-contribution pension systems.

4.5.1 The Long Run Effects

With a defined-benefit pension system it is the tax rate which adjusts itself and the average replacement rate (ν) is fixed at 0.757 which is the value obtained by Hairault and Langot (2008) on French data. The annual interest rate obtained is approximately 4.4%[30].

Qualitatively, we observe the expected results. Indeed an increase in the redistributivity of the pension system (a decrease in λ) has a negative impact on the tax rate, and a positive one on the steady state capital per worker, on the GDP per capita and on the wealth level. Welfare inequalities decrease.

Quantitatively, diminishing arbitrarily λ from 0.885 to 0.785, i.e. a decrease of 11.3%, we find a decrease in the tax rate of 2.49%. The steady state level of capital per worker and the GDP per capita increase of 2.7% and of 0.88% respectively. Welfare inequalities decrease of 1.18%. Finally, the Gini coefficient of wealth[31] decreases which means that wealth inequalities decrease. Table (1) sums up the main results.

[29] Appendix F provides a sensitivity analysis.

[30] The annual interest rate is obtained by $R^{1/40} - 1$, with R the interest factor obtained using equation (4.9).

[31] Using the same methodology as for the distribution of wages, the Gini coefficient of Wealth is obtained using the formula of Lambert (2001, pp.33) : $G = -1 + 2\int_{a_-}^{a_+} \frac{W(a)F(a)f(a)}{\overline{W}(a)}da.$

$\Delta\lambda = -11.3\%$	Defined-Benefit	Defined-Contribution
$\Delta\tau^a$	-2.5%	-
$\Delta\nu$	-	+2.55%
Δk^*	+2.7%	-0.15%
$\Delta W(a_+)$	+1.14%	-0.55%
$\Delta W(a_-)$	+2.68%	+.053%
ΔRIW^b	+1.53%	+1.08%
$\Delta GDPpc^c$	+0.88%	-0.05%
$\Delta dUtil^d$	-1.18%	-0.85%
IGb^e	0.3383	0.3351
IGa	0.3364	0.3338

[a]Here we report a change in % and not in $\%pts$.
[b]$RIW = W(a_-)/W(a_+)$.
[c]GDPpc means GDP per capita.
[d]$dUtil = U(a_+) - U(a_-)$.
[e]IGb (IGa) denotes the Gini coefficient before (after) the change in λ.

Table 4.1: Macroeconomic impact of a more redistributive pension system

We now study the case of a defined-contribution pension system. The tax rate is fixed exogenously at 0.23. It is the value calculated by Hairault and Langot (2008), and it is near the tax rate reported by Nyce and Schieber (2005). We also study the impact of an arbitrarily decrease in λ. The annual interest rate is approximately 3.9%.

Qualitatively the results show an increase in the replacement rate. Furthermore, the net effect on savings is negative since the steady state capital per young decreases. This last effect implies a decrease in the wealth of the richest, whereas the net effect remains positive for the poorest because of the redistributive effect.

Quantitatively, diminishing λ of 11.3% (from 0.885 to 0.785), we find an increase in the replacement rate of 2.55%. The steady state level of capital per young and the GDP per capita decrease of 0.15% and of 0.05% respectively. The utility inequalities decrease of 0.85%. As before we also observe a decrease in the Gini coefficient of wealth, i.e. a decrease in wealth inequalities. Table (1) sums up the main results.

Two conclusions can right now be stressed: (i) the net impact is greater for a defined-benefit pension system than for a defined-contribution pension system because in the first case every effect has the same sign. (ii) For a defined-contribution pension system the only positive impact of the redistributivity is to reduce inequalities.

4.5.2 The Transitional Dynamics

The main objective of this part is to study the short run effects of an unexpected decrease in λ of 11.3%. We assume that the economy is initially at its steady state. λ is assumed to remain constant during the first two periods and then to decrease to 0.785. Agents born in period 2 do not expect this change and thus do not adjust their savings. But, for every following generation the assumption of perfect foresight implies that they exactly adjust their savings in order to maximize their utility. Because of the unpredictability of the change in λ, the capital per worker remains constant until period 3 and adjusts only during the following periods.

With a defined-benefit pension system the tax rate becomes 0.31 from period 3 (0.3 initially). Agents born in period 1 are not affected by this change and are used as a reference. The capital per young adjusts progressively to its new steady state value. The utility of the richest decreases substantially for agents born in period 2 because they do not sufficiently save for their second period of life. But the utility of the poorest increases until it reaches a new steady state value which is higher.

Utility inequalities decrease strongly right the second generation and then stabilize themselves after a very small increase because of the adjustment of the savings of the richest. Figures 1-4 sum up the main results.

For defined-contribution pension systems the simulation is the same. Qualitative results show a quick adjustment of the variables towards their new steady state value. Only the utility levels of the consumers born in period 2 describe a different trajectory. The utility of the richest and that of the poorest decrease and increase respectively. Figures 5-8 sum up the main results.

Remark : Qualitative and quantitative results are very different according to the nature of the pension system (defined-benefit or defined-contribution). It has to be taken into account in order to study the impact of a change in the redistributive properties of a pension system.

4.6 Conclusion

The increase in the redistributivity of a defined-benefit pension system can : (i) decrease the taxe rate of the pension system; (ii) increase the capital per capita; (iii) increase the wealth and the welfare of every agent; (iv) reduce the inequalities of wealth and of welfare. However, if the pension system has a defined-contribution structure, then the only positive effect is that it increases the wealth and the utility of the poorest agents.

Therefore, the knowledge of the nature of a pension system (defined-benefit or defined-contribution) and the taking into account of the life expectancy differential are both important in order to determine the qualitative and quantitative impacts of a more redistributive pension system.

The first extension of this chapter would be to introduce labor supply in order to take into account the distorsive impact of our redistributive policy.

Another application of this chapter would be to study the impact of redistributive policies on educational choices. In the case of a capital-skill complementarity, and given the mechanism we described above, it is possible that a more redistributive pension system implies that a larger share of the population decides to educate herself. Another extension would be to clarify theoretically the debate on the inequalities of contribution to pension systems. Indeed, the inequalities of length of life imply that pension systems are far less progressive than they seem. That is what we do in the last chapter of this thesis.

4.7 APPENDIX

Appendix A

Computation of the expression of τ:

$$(1+n)\int_{\Omega_a}\tau w_{t+1}(a)f(a)da=\int_{\Omega_a}p_{t+1}(a)f(a)da \tag{4.35}$$

Furthermore, we know that:

$$p_{t+1}(a)=\nu A(1-\alpha)\frac{k_{t+1}^{\alpha}}{\bar{a}^{\alpha}}(\lambda a+(1-\lambda)\bar{a})T(a) \tag{4.36}$$

Computing the RHS we obtain the following expression:

$$RHS=\nu A(1-\alpha)\frac{k_{t+1}^{\alpha}}{\bar{a}^{\alpha}}\left(\lambda\int_{\Omega_a}T(a)af(a)da+(1-\lambda)\bar{a}\int_{\Omega_a}T(a)f(a)da\right) \tag{4.37}$$

Equation (4.3) implies: $\int_{\Omega_a}T(a)af(a)da=COV_{T(a),a}+\bar{a}\bar{T}$. The second part of the expression between brackets is the average length of life. Finally we have:

$$RHS=\nu A(1-\alpha)\frac{k_{t+1}^{\alpha}}{\bar{a}^{\alpha}}(\lambda COV_{T(a),a}+\bar{a}\bar{T}) \tag{4.38}$$

We recognize in the LHS the average wage: $\displaystyle\int_{\Omega_a}w_{t+1}(a)f(a)da$. Then equalizing the LHS and the RHS we obtain equation (4.16).□

Appendix B

(i) The study of equation (4.20) shows that τ and the denominator become independent of λ if $T(a)=a$, $\forall a$.□

(ii) Let us consider the case of homothetic preferences which have the following form:

$$U_t(a)=U\left(c_t(a),\frac{d_{t+1}(a)}{T(a)}\right) \tag{4.39}$$

The intertemporal budget constraint of this agent is:

$$c_t(a)+\frac{d_{t+1}(a)}{R_{t+1}}=w_t(a)(1-\tau)+\frac{p_{t+1}(a)}{R_{t+1}}\equiv W_t(a) \tag{4.40}$$

Given the preferences, the solution for consumers is:

$$c_t(a) = \xi(T(a), R_{t+1})W_t(a) \tag{4.41}$$

And finally:

$$S_t(a) = w_t(a)(1-\tau) - c_t(a) = \xi_1(T(a), R_{t+1})w_t(a) - \xi_2(T(a), R_{t+1})\frac{p_{t+1}(a)}{R_{t+1}} \tag{4.42}$$

Therefore savings is a linear function of the wage and of the pension. Assuming that the length of life is the same for every agent ($T(a) = T$, $\forall a$) then the capital market equilibrium can be written:

$$(1+n)k_{t+1} = \int_{\Omega_a} S_t(a)f(a)da \tag{4.43}$$

or,

$$(1+n)k_{t+1} = \xi_1(T, R_{t+1})\bar{w}_t - \xi_2(T, R_{t+1})\frac{\nu T\bar{w}_{t+1}}{R_{t+1}} \tag{4.44}$$

λ does not appear in this expression.$\square$

Appendix C

Proof of proposition 4:

The derivative of equation (4.22) with respect to λ gives the following expression:

$$\frac{dW_t(a)}{d\lambda} =$$

$$aA\frac{1-\alpha}{\bar{a}^\alpha}\left(\alpha k^{\alpha-1}\frac{dk}{d\lambda}(1-\tau) - \frac{d\tau}{d\lambda}k^\alpha\right) + \nu\frac{1-\alpha}{\alpha}T(a)\left[\left(\lambda\frac{a}{\bar{a}} + 1 - \lambda\right)\frac{dk}{d\lambda} + k(\frac{a}{\bar{a}} - 1)\right]$$

We know that $d\tau/d\lambda > 0$ and that $dk^*/d\lambda < 0$. Finally, the previous expression is negative if the second part of the equation is negative, i.e. if:

$$\frac{\frac{dk}{d\lambda}}{k} \leq \frac{1-\frac{a}{\bar{a}}}{\lambda\frac{a}{\bar{a}} + 1 - \lambda}$$

However, as the RHS is a decreasing function of a, then it is sufficient for this inequality to be true for $a = a_+$.

Remark: This inequality is always true for $a < \bar{a}$.$\square$

Proof of proposition 4 (bis):

The methodology is the same as before except that τ is fixed exogenously and that ν is a decreasing function of λ.□

Appendix D

Proof of proposition 5:

Equation (4.22) can be written :

$$W_t(a) = A(1-\alpha)\frac{k_t^\alpha}{\bar{a}^\alpha}a(1-\tau) + \nu\frac{1-\alpha}{\alpha}k_{t+1}\left(\lambda\frac{a}{\bar{a}} + (1-\lambda)\right)T(a) \tag{4.45}$$

or, in the steady state:

$$W_t(a) = k\left[A(1-\alpha)\frac{k^{\alpha-1}}{\bar{a}^\alpha}a(1-\tau) + \nu\frac{1-\alpha}{\alpha}\left(\lambda\frac{a}{\bar{a}} + (1-\lambda)\right)T(a)\right] \tag{4.46}$$

With equation (4.20), the LHS between brackets can be written:

$$a\frac{1+n+\nu\frac{1-\alpha}{\alpha}\int_{\Omega_a}\frac{\lambda\frac{a}{\bar{a}}+(1-\lambda)}{1+\beta T(a)}T(a)f(a)da}{\beta\int_{\Omega_a}\frac{T(a)a}{1+\beta T(a)}f(a)da} \equiv af(\lambda)$$

Equation (4.46) becomes:

$$W_t(a) = k\left[af(\lambda) + \nu\frac{1-\alpha}{\alpha}\left(\lambda\frac{a}{\bar{a}} + (1-\lambda)\right)T(a)\right] \tag{4.47}$$

The relative wealth inequalities can be written:

$$\frac{W_t(a_-)}{W_t(a_+)} = \frac{a_-f(\lambda) + \nu\frac{1-\alpha}{\alpha}\left(\lambda\frac{a_-}{\bar{a}} + (1-\lambda)\right)T(a_-)}{a_+f(\lambda) + \nu\frac{1-\alpha}{\alpha}\left(\lambda\frac{a_+}{\bar{a}} + (1-\lambda)\right)T(a_+)} \tag{4.48}$$

The result of the proposition is obtained if the derivative of this expression with respect to λ is negative. It is true if and only if:

$$[f(\lambda) - \lambda f'(\lambda)]\left[a_+\left(\frac{a_-}{\bar{a}} - 1\right)T(a_-) - a_-\left(\frac{a_+}{\bar{a}} - 1\right)T(a_+)\right] <$$
$$f'(\lambda)[a_+T(a_-) - a_-T(a+)] + \nu\frac{1-\alpha}{\alpha}T(a_+)T(a_-)\left(\frac{a_+ - a_-}{\bar{a}}\right)$$

The LHS has two components. The second is obviously negative. It is straightforward to show that $1 > \lambda\frac{f'(\lambda)}{f(\lambda)}$ and then that the left hand side is negative.

It only remains to show that the RHS is positive. It is true under the condition of the proposition.□

Proof of proposition 5 (bis):

The relative wealth inequalities can be written:

$$\frac{W_t(a_-)}{W_t(a_+)} = \frac{A(1-\alpha)\frac{k^{\alpha-1}}{\bar{a}^{\alpha}}a_-(1-\tau) + \nu\frac{1-\alpha}{\alpha}\left(\lambda\frac{a_-}{\bar{a}} + (1-\lambda)\right)T(a_-)}{A(1-\alpha)\frac{k^{\alpha-1}}{\bar{a}^{\alpha}}a_+(1-\tau) + \nu\frac{1-\alpha}{\alpha}\left(\lambda\frac{a_+}{\bar{a}} + (1-\lambda)\right)T(a_+)} \tag{4.49}$$

The derivative of this expression with respect to λ is negative if and only if:

$$\begin{aligned}\tfrac{1-\alpha}{\alpha}c\left[(\alpha-1)k^{\alpha-2}\tfrac{dk}{d\lambda} - \tfrac{d\nu}{d\lambda}k^{\alpha-1}\right]\left[a_-T(a_+)(\lambda\tfrac{a_+}{\bar{a}}+1-\lambda) - a_+T(a_-)(\lambda\tfrac{a_-}{\bar{a}}+1-\lambda)\right] \\ +\tfrac{1-\alpha}{\alpha}ck^{\alpha-1}\left[a_+T(a_-)(\tfrac{a_-}{\bar{a}}-1) - a_-T(a_+)(\tfrac{a_+}{\bar{a}}-1)\right] \\ +\nu\tfrac{1-\alpha}{\alpha}T(a_-)T(a_+)\left(\tfrac{a_- - a_+}{\bar{a}}\right) < 0\end{aligned}$$

with $c = A(1-\alpha)\frac{1-\tau}{\bar{a}^{\alpha}} > 0$.

The last two terms are strictly negative. Then under the condition $dk/d\lambda < 0$, and knowing that $d\nu/d\lambda < 0$, the sign of the first term depends only on the sign of the condition mentioned in the proposition.□

Appendix E

The covariance can also be written as: $\int_{\Omega_a}(a-\bar{a})(T(a)-\bar{T})f(a)da$. But as $\int_{\Omega_a}(a-\bar{a})f(a)da = 0$, we can write that: $\int_{\Omega_a}(a-\bar{a})(T(a)-\bar{T})f(a)da = \int_{\Omega_a}(a-\bar{a})(T(a)-X)f(a)da$, with X a constant, whatever the value of X. So it is particulary true for $X = T(\bar{a})$. Then we can write that: $\int_{\Omega_a}(a-\bar{a})(T(a)-\bar{T})f(a)da = \int_{\Omega_a}(a-\bar{a})(T(a)-T(\bar{a}))f(a)da$. The RHS is positive as it is an integral on a product of terms with the same sign because $T'(a) > 0$.□

Appendix F

In this appendix we try to determine if our qualitative results depend on an initial condition, on the form taken by $T(a)$ or on values taken by our parameters, notably by the average replacement rate (ν) or the tax rate (τ). In doing so we extend our results to other countries than France.

For defined-benefit pension systems

Firstly, let us consider the impact of a decrease in λ in function of its initial value. A simple numeric exercise, using our calibration, shows that our qualitative results remains true whatever the initial value of λ and whatever the percentage of change in λ. It implies that a decrease in λ has always a positive impact on capital per capita and on wealth of every agent. It also always has a negative impact on wealth inequalities, on the Gini coefficient and on the utility differential ($dUtil$).

Secondly, we do the same exercise but with the new function $T(a) = a^{0.75}$. The form of this function implies that the average life expectancy of agents in our model is 80 years old, what matches the observed life expectancy in most industrialized countries. We find the same qualitative results. As previously, our results do not depend on the initial value taken by λ.

Thirdly, we solve our model for different values of ν ($\nu \in \{0.757, 0.6, 0.4\}$)[32]. Whatever the function $T(a)$ which is chosen, our qualitative results are unchanged.

For defined-contribution pension systems

We find a monotonous relationship between macroeconomic variables and λ. It implies that the impact of λ on macroeconomic variables has the same sign as this reported in Table 1 whatever its initial value.

As in the defined-benefit case, the use of the functional form $T(a) = a^{0.75}$ has no impact on our qualitative results. λ still has a monotonous impact on macroeconomic variables.

Finally, we check that our qualitative results remain unchanged for $\tau \in \{0.1, 0.23, 0.3\}$.

Appendix G

In this appendix, we sum up our calibration of the functions and of the parameters of our model. Furthermore, we detail some important statistics.

[32] $\nu = 0.4$ seems to be the lowest replacement rate in among industrialized countries. See Nyce and Schieber (2005, pp.236).

Parameter	Meaning	Value	Source(s)
α	$R_t K_t / Y_t$[a]	0.33	Sommacal (2006) among others
β	Actualization factor	0.6	APDR=1.3%[b], d'Autume (2003),
			Heer and Maussner (2005)
A	The technology level	1	Normalization
n	Population's growth rate	0.3	AGR=0.65%[c], Charpin (1999)
ν	Average Replacement rate[d]	0.757	Hairault and Langot (2008)
τ	Tax rate[e]	0.23	Hairault and Langot (2008)
λ_I	Initial value of λ[f]	0.885	Hairault and Langot (2008)

[a]The share of income spent on capital.
[b]Annual psychological discount rate.
[c]AGR=annual growth rate.
[d]For defined-benefit pension systems.
[e]For defined-contribution pension systems.
[f]We use this value as a reference. We analyse the effects of a decrease in λ knowing that λ is initially λ_I.

Table 4.2: Basic Calibration of the model

The Basic Calibration

The length of each period is 40 years. Table 2 sums up the basic parameters which we use for the numerical resolution of our model.

The calibration of functions and their main statistics

Firstly, we calibrate the interval Ω_a. We use:

$$\Omega_a = [0.08, 1]$$

The ratio a_+/a_- is lower than the one found in Acemoglu (2002) but higher than the one of Piketty (2002). The corresponding density function is:

$$f(a) = 2.1129 - 1.9a \tag{4.50}$$

These two components respect the two main properties:

- mode<median<mean, Source: Lambert (2001)
- $IG_w = 0.32$[33] in France, Source: Hairault and Langot (2008), INSEE (1999)

In our model, we have:

$$\bar{a} = 0.4167$$
$$a_{median} = 0.378$$
$$a_{mode} = a_{-} = 0.08$$
$$Var(a) = 0.05533$$

Secondly, we calibrate the function $T(a)$:

$$T(a) = a$$

It implies that the distribution of the length of life has the same properties as the distribution of the variable a. Furthermore, we have:

$$COV_{T(a),a} = Var(a) = 0.05533$$

Knowing that the length of each period is 40 years, the average length of life[34] is 77 years. It is lower than the figure for France which is around 80 years[35] (Source: INSEE or World Bank). The standard deviation is:

$$\sigma_{T(a)} = 0.24$$

which corresponds to a standard deviation of almost 9.4 years[36].

Appendix H

In this appendix we show that agents having a high productivity level have a positive savings, whereas agents having a low productivity level can resort to borrowing at equilibrium.

The saving function of an agent endowed with a productivity level a is given by function (4.7). Saving is positive iff:

---

[33] IG_w denotes the Gini coefficient of wages.

[34] The life expectancy of each individual is $(1 + T(a)) * 40$.

[35] Appendix F shows that it has no impact on our qualitative results.

[36] The standard deviation for the function $T(a) = a^{0.75}$ is lower than 9 years.

$$\beta T(a)w_t(a)(1-\tau) > \frac{p_{t+1}(a)}{R_{t+1}}$$

Using equations (4.9), (4.11), (4.12) and (4.14) we obtain:

$$(1-\tau)\beta A\alpha\bar{a}^{1-\alpha}\frac{a}{\nu(\lambda a+(1-\lambda)\bar{a})}k_t^{\alpha} > k_{t+1}$$

Then, substituting k_{t+1} by its expression (4.19), we obtain that agent endowed with a productivity level a save iff:

$$a > \frac{\nu(1-\lambda)(1-\alpha)\int_{\Omega_a}\frac{T(a)a}{1+\beta T(a)}f(a)da}{\alpha(1+n)+\nu(1-\lambda)(1-\alpha)\int_{\Omega_a}\frac{T(a)}{1+\beta T(a)}f(a)da} \equiv \tilde{a}$$

$\tilde{a}$ defines the threshold value of productivity, such that agents endowed with a productivity level higher than $\tilde{a}$ save, whereas agents endowed with a smaller productivity than $\tilde{a}$ resort to borrowing.

Proposition 4.10 *The threshold value $\tilde{a}$ is such that $a_+ > \tilde{a}$. However, we can have either $a_- > \tilde{a}$, or $\tilde{a} > a_-$.*

Proof: For the proof of the first part of the proposition, it is sufficient to note that:

$$a_+ > a_+\frac{\nu(1-\lambda)(1-\alpha)\int_{\Omega_a}\frac{T(a)}{1+\beta T(a)}f(a)da}{\alpha(1+n)+\nu(1-\lambda)(1-\alpha)\int_{\Omega_a}\frac{T(a)}{1+\beta T(a)}f(a)da} > \frac{\nu(1-\lambda)(1-\alpha)\int_{\Omega_a}\frac{T(a)a}{1+\beta T(a)}f(a)da}{\alpha(1+n)+\nu(1-\lambda)(1-\alpha)\int_{\Omega_a}\frac{T(a)}{1+\beta T(a)}f(a)da}$$

However, we cannot conclude as for the sign of: $a_- \gtrless \tilde{a}$.□

This result implies that we can determine the impact of a change in the size or in the structure on the share of the population who resorts to borrowing.

Proposition 4.11 *If $\tilde{a} \in (a_-, \bar{a})$, then (i) for defined-benefit pension systems, an increase in the replacement rate of the pension system or an increase in their Beveridgian component have both a positive impact on the share of the population who resorts to borrowing. (ii) For defined-contribution pension systems, an increase in the tax rate of the pension system or an increase in their Beveridgian component have both a positive impact on the share of the population who resorts to borrowing.*

Proof: (i) In the case of defined-benefit pension systems, the replacement rate is given. Then, let us define $\tilde{a}$ as a function of ν and of λ. It can easily be shown that $\partial\tilde{a}(\nu,\lambda)/\partial\nu > 0$ and that $\partial\tilde{a}(\nu,\lambda)/\partial\lambda < 0$. (ii) For defined-contribution pension systems, the replacement rate is a function such that: $\partial\nu(\tau,\lambda)/\partial\tau > 0$ and $\partial\nu(\tau,\lambda)/\partial\lambda < 0$. The same method as before can be used.$\square$

This proposition emphasizes the main role played by the absence of liquidity constraints in our model to study the impact of an increase in the Beveridgian part of pension systems. Indeed, the impact on the welfare of agents could be very different if we had considered this kind of constraint. Even if the system redistributes more resources in favor of poor agents, as a larger share of poor agents becomes constrained, the net impact on the welfare of these new constrained agents is not so clear.

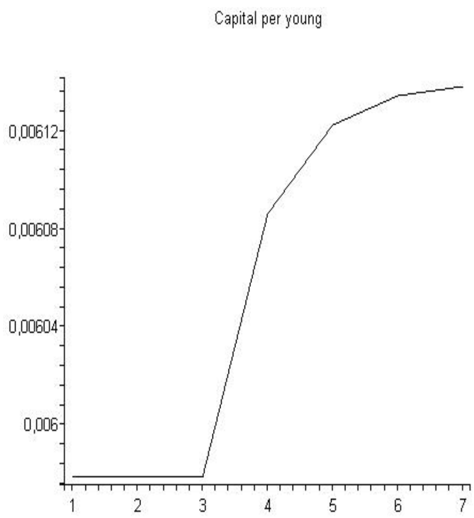

Figure 4.1: Capital per young (k_t) for defined-benefit pension systems. Periods are reported on the abscissa.

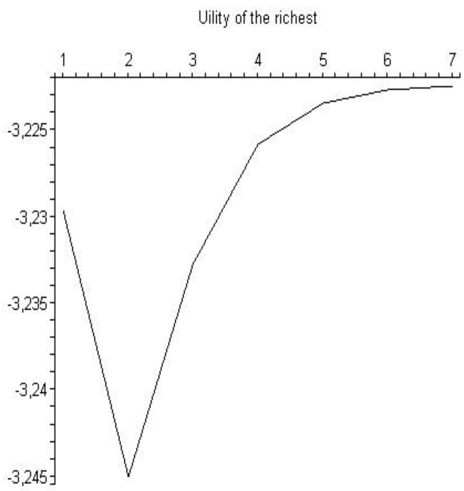

Figure 4.2: Utility of the richest ($U_t(a_+)$) for defined-benefit pension systems. For example $U_1(a_+)$ is the utility of the richest born in period 1.

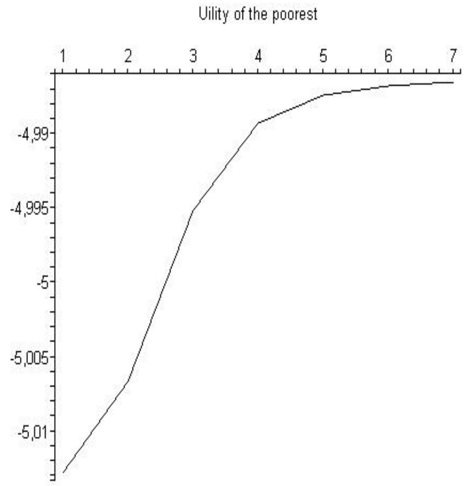

Figure 4.3: Utility of the poorest ($U_t(a_-)$) for defined-benefit pension systems. For example $U_1(a_-)$ is the utility of the poorest born in period 1.

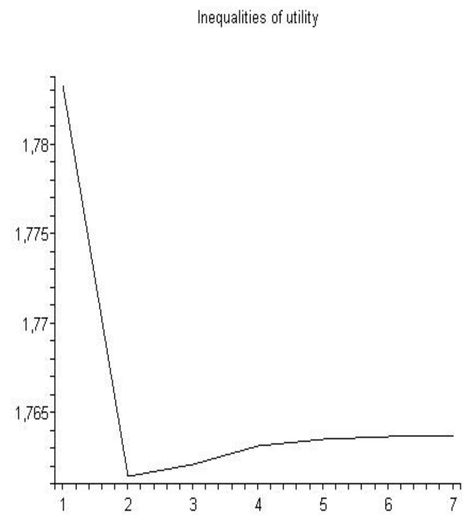

Figure 4.4: Utility differential for defined-benefit pension systems

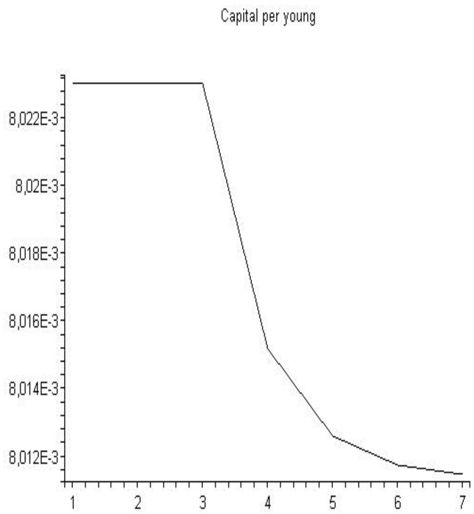

Figure 4.5: Capital per young (k_t) for defined-contribution pension systems

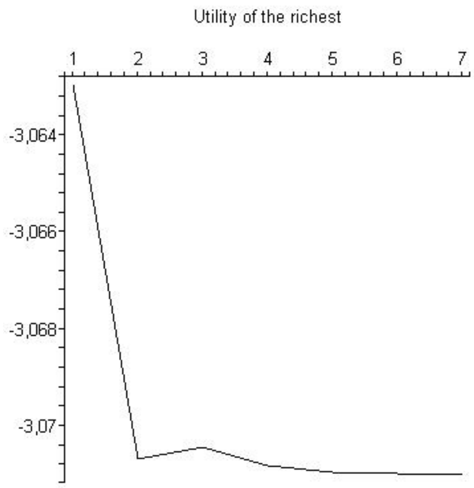

Figure 4.6: Utility of the richest ($U_t(a_+)$) for defined-contribution pension systems

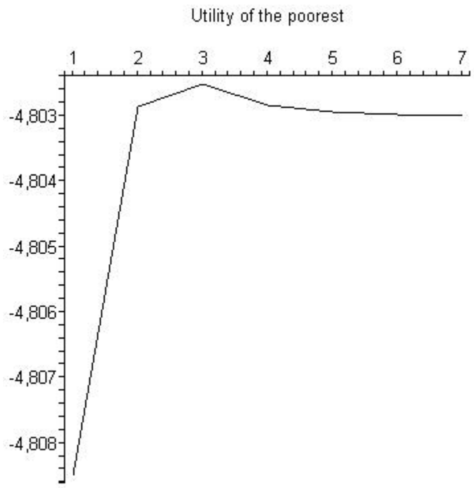

Figure 4.7: Utility of the poorest ($U_t(a_-)$) for defined-contribution pension systems

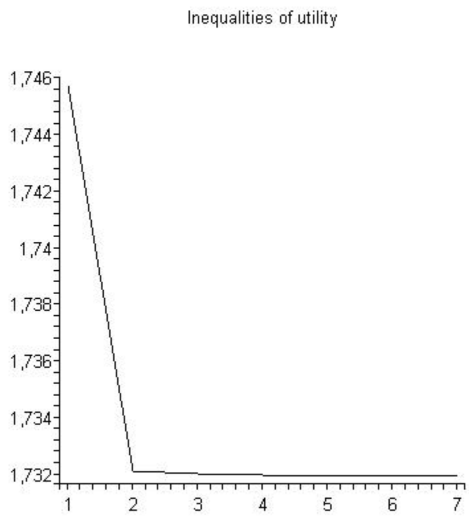

Figure 4.8: Utility Differential for defined-contribution pension systems

Chapter 5

Redistribution, Pension Systems and Capital Accumulation: An extension

5.1 Introduction

Most of the economic literature related to pension systems has studied the macroeconomic impact of an increasing size of pension systems[1]. However, only a few papers deal with the impact of the structure of pension systems. In this chapter, the term "structure" means that a pension system can be Beveridgian, Bismarckian or a mix of the two. As in Casamatta *et al.* (2000), a pension system has a pure Beveridgian structure if each agent receives the same pension. Conversely, if pensions completely depend on wages then the pension system is Bismarckian. The structure of pension systems determines their redistributive properties[2]. The more the pension system is Beveridgian, the more it redistributes resources among the population. France, Germany and Italy have a Bismarckian structure. Canada, the Netherlands and New-Zeland are essentially Beveridgian. Finally, Japan, the United-Kingdom and the United States have mixed pension systems (Sommacal 2006, Casamatta *et al.* 2000). Countries are different from one another because of this intra-generational component, so it can be relevant to consider the macroeconomic impacts of a policy which changes the structure of pension systems.

In this chapter, we analyze the impact of a policy which increases the Beveridgian part of pension systems. We find that the life expectancy differential play a significant role in the study of the impact of the structure of pension systems.

There is a growing empirical literature which analyzes the life expectancy differential[3]. Mesrine (1999) studies the inequalities of length of life according to socio-professional groups in France[4]. The most striking feature of his paper is that a worker has a probability to die between 35 and 65 almost twice higher than that of an executive manager. Furthermore, their life expectancy at 35 is 38 and 44 respectively. The same qualitative results are observed in the United-States (Panis and Lillard 1995, Deaton and Paxson 2000). Finally, Robert-Bobbée and Cadot (2007) show that this inequality is also observed for elderly people. For agents who are 86, the ones with highest education level can expect to

[1]See Docquier and Paddison (2003), or Casarico and Devillanova (2007). These results are notably questioned by Groezen *et al.* (2007), Lambrecht *et al.* (2005) or Le Garrec (2005).

[2]Here and in the rest of this chapter, the term "redistribution" means "instantaneous redistribution", i.e. the redistribution of wealth through the indexation of pensions on wages.

[3]See Attanasio and Emmerson (2001), Bommier *et al.* (2003) or Adams *et al.* (2003) for a survey.

[4]These inequalities also depend on other factors like sex or geographical localization. For example, in France the life expectancy of women is 84.1, while that of men is only 77.2 (INSEE 2006). Moreover, Rican and Salem (1999) show that there are strong disparities according to the localization of people in France.

live 20% longer than the ones with lowest education level.

In this chapter, we consider an overlapping generations economy in which agents live for two periods. We assume that there is a fixed fraction of skilled agents in the population[5]. As the empirical literature suggests, skilled agents have a longer life expectancy than unskilled ones. Both offer their labor inelastically when they are young[6] (first period of life), and both retire at the very beginning of their second period of life. The Government levies a tax rate on wages in order to finance a Pay-As-You-Go (PAYG) pension system. This pension system has a mixed structure, i.e. it has a Beveridgian and a Bismarckian component. Finally, we assume that firms use a capital-skill complementarity technology. It means that the elasticity of substitution between capital and unskilled labor is higher than that between capital and skilled labor (Krusell *et al.* 2000). This assumption has been empirically observed[7] and it implies that an increase in the capital per capita level increases wage inequalities (Duffy *at al.* 2004).

We show analytically that if the Beveridgian part of pension systems increases, then it has a positive impact on capital per capita. Given the technology of firms, it means that this redistributive policy has a positive impact on wages and on wage inequalities. The impact on other macroeconomic variables cannot be determined *a priori*. It explains why we calibrate our model. If the life expectancy differential is sufficiently high, we show that a more redistributive pension system increases the wealth and the welfare of each agent of the economy. Moreover, such a policy decreases the tax rate of the pension system.

In the previous chapter, we made the same exercise with a continuum of agents endowed with different productivity levels. This chapter differs from this in two ways. Firstly, we use more general utility and production functions in our basic framework. Thus, we show that the structure of pension systems has an impact on wage inequalities. Secondly, because there are only two groups of agents in this chapter, it is easier for us to emphasize that the life expectancy differential is what matters to determine the macroeconomic impact of our redistributive policy.

---

[5]It means that we assume that the structure of pension systems has no impact on occupational choices.

[6]In doing so we do not model the burden of income taxation on labor supply. We make the same assumption as in Feldstein (1985) given that:

> "The primary cost of providing social security benefits is the welfare loss that results from reductions in private saving" (Feldstein 1985, pp.303).

[7]See Griliches (1969), Fallon and Layard (1975), Krusell *et al.* (2000) or Duffy *et al.* (2004).

Two kinds of papers can be related to ours. In the first kind of papers, authors study the macroeconomic impact of a change in the structure of pension systems. Sommacal (2006) studies the macroeconomic impacts of a more redistributive unfunded pension system. He uses a defined-contribution pension system[8] with endogenous labor supply and imperfect substituability between two kinds of labor: skilled and unskilled labor. He finds that an increasing redistributivity of pension systems decreases output but also the wealth of each agent of the economy. Compared to this study, we make three different assumptions. We assume an exogenous labor supply, a defined-benefit pension system and a capital-skill complementarity technology. We show that conclusions are completely different in this new framework.

The second kind of papers studies the impacts of a change in the size of pension systems when firms use a capital-skill complementarity technology. Casarico and Devillanova (2007) find that an increase in the size of pension systems has a negative impact on capital accumulation, on the share of the educated population, on output and on wage inequalities.

This chapter is organized as follows. In section 2 we present our model and our main assumptions. In section 3 we study the dynamics of our model and its properties. In section 4 we solve our model numerically. Section 5 includes some concluding remarks.

5.2 The Model

We assume an overlapping generations economy in which agents live for two periods. In the first period of their life, agents work. In the second one, they are retired. In order to have a tractable model which includes a capital-skill complementarity technology, we assume that there are only two groups of agents: skilled and unskilled agents. We assume that there are eN_t skilled agents and $(1-e)N_t$ unskilled agents at each period t, with $1 > e > 0$. There is no uncertainty. Skilled and unskilled agents only live a fraction φ and σ respectively, of their second period of life. It is the same assumption as in Gorski *et al.* (2007). We denote by ϵ the mortality differential between these two groups of agents. It means that $\epsilon = \varphi - \sigma$. We assume that $\epsilon \geq 0$. For $\epsilon = 0$ there are no inequalities of length of life, and for $\epsilon = \varphi$ uneducated agents live only their first period of life. Figure 5.1 represents the life cycle of an agent.

[8] A pension system has a defined-benefit structure if it is the tax rate which adjusts itself to changes in the economic and demographic environment. Conversely, it has a defined-contribution organization if it is the replacement rate which adjusts itself.

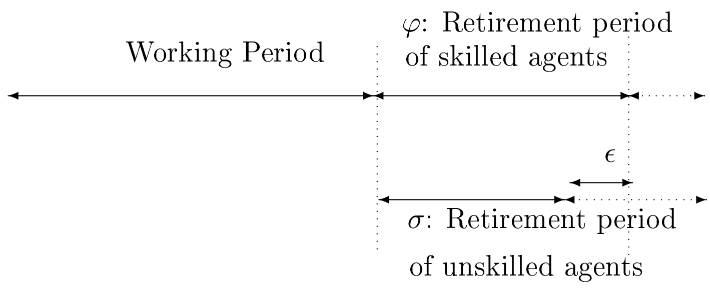

Figure 5.1: Life of agents

5.2.1 Consumers

Each consumer of the economy belongs either to the group of educated agents or to the group of uneducated agents. All agents have the same preferences. These preferences are intertemporally separable and they have the Constant-Intertemporal-Elasticity-of-Substitution form[9]:

$$U_t^i = U(c_t^i) + \beta T^i U\left(\frac{d_{t+1}^i}{T^i}\right) \tag{5.1}$$

It is such that:

$$U(x) = \begin{cases} \frac{x^{1-\eta}}{1-\eta} & \text{if } \eta > 0 \text{ but } \eta \neq 1 \\ \ln(x) & \text{if } \eta = 1 \end{cases}$$

β and η denote the pure time preference factor and the inverse of the elasticity of substitution between c_t and d_{t+1} respectively. U_t^i denotes the utility level of an agent of type i (with $i \in \{s, u\}$) born at the beginning of period t. The superscript s means that an agent is skilled, whereas the superscript u means that an agent is unskilled. c_t^i (d_t^i) denotes the consumption of a young (old) agent. T^i denotes the length of life of an agent of type i. It increases the weight that an agent attaches to his future utility. But at the same time,

[9]Andersen (2008) uses the same utility function.

it decreases the consumption flow of the second period of life[10][11]. We have:

$$T^i = \begin{cases} \varphi & \text{if } i = s \\ \sigma & \text{if } i = u \end{cases}$$

The budget constraints of an agent of type i are:

$$c_t^i = w_t^i(1 - \tau_t) - S_t^i \tag{5.2}$$

$$d_{t+1}^i = R_{t+1} S_t^i + p_{t+1}^i \tag{5.3}$$

with R_{t+1} the interest factor, τ_t the tax rate used to finance a PAYG pension system. S_t^i denotes the savings of a young agent of type i at period t. p_{t+1}^i denotes the pension that an agent of type i receives when he is retired. Every agent maximizes (5.1) with respect to S_t^i given the budget constraints (5.2) and (5.3). We obtain:

$$S_t^i = \frac{T^i(\beta R_{t+1})^{1/\eta} w_t^i (1 - \tau_t)}{T^i(\beta R_{t+1})^{1/\eta} + R_{t+1}} - \frac{p_{t+1}^i}{T^i(\beta R_{t+1})^{1/\eta} + R_{t+1}} \tag{5.4}$$

Without a pension system, i.e. for $\tau = p^i = 0$, saving is an increasing function of the length of life. The longer the length of life is, the more agents value their future consumption, and thus, the more they save. Moreover, if τ or if p^i increases, then saving decreases. Firstly, because it decreases the net wage of the first period of life of agents. Secondly, because it increases the revenues from the second period of life of agents.

5.2.2 Firms

We assume that the technology of firms has the following form:

$$Y_t = F(L_t^u, G(K_t, L_t^s)) \tag{5.5}$$

with $F(.,.)$ and $G(.,.)$ two homogenous functions of degree 1. L_t^s (L_t^u) denote the quantity of skilled (unskilled) labor used in the production function. $F()$ and $G()$ are such that[12]:

[10] See d'Autume (2003) and the appendix of chapter 1.

[11] Using the budget constraints and the market conditions defined below, it is straightforward to show that T^i can also denote the probability of dying of an agent of type i. In that case, it is sufficient to assume that the educated and the uneducated population are sufficiently large and that there is a perfectly competitive annuity market for each group of agents.

[12] g_i denotes the derivative of $g()$ with respect to its *ith* argument. g_{ij} denotes the derivative of g_i with respect to its *jth* argument.

$$F_i > 0 \text{ and } G_i > 0$$
$$F_{ii} < 0 \text{ and } G_{ii} < 0$$
$$F_{ij} > 0 \text{ and } G_{ij} > 0 \text{ with } i \neq j$$

We assume a perfect competition on the final good market and on inputs markets. It implies that, at equilibrium, wages and capital return are:

$$w_t^u = F_1((1-e), G(k_t, e)) \equiv w^u(k_t) \tag{5.6}$$

$$w_t^s = F_2((1-e), G(k_t, e))G_2(k_t, e) \equiv w^s(k_t) \tag{5.7}$$

$$R_t = F_2((1-e), G(k_t, e))G_1(k_t, e) \equiv R(k_t) \tag{5.8}$$

with $k_t = K_t/N_t$. Given the properties of the functions $F()$ and $G()$, it can be shown that:

$$\frac{\partial R(k_t)}{\partial k_t} < 0$$
$$\frac{\partial w^u(k_t)}{\partial k_t} > 0$$

A priori, we cannot determine the sign of the derivative of $w^s(k_t)$ with respect to k_t. As in Duffy *et al.* (2004), in the rest of this chapter we make the following assumption:

Assumption 1: $\dfrac{\partial w^s(k_t)/\partial k_t}{w^s(k_t)} k_t > \dfrac{\partial w^u(k_t)/\partial k_t}{w^u(k_t)} k_t.$

This assumption necessarily implies that $\partial w^s(k_t)/\partial k_t > 0$. It means that the elasticity of the wages of skilled agents with respect to capital per capita is higher than the one concerning the wages of unskilled agents. Appendix 2 shows that this assumption is correlated with the capital-skill complementarity technology. In this chapter, Iw denotes the wage inequality ratio. It is such that:

$$I_{w,t} = \frac{w_t^s}{w_t^u} \tag{5.9}$$

Given assumption 1, this wage inequality ratio is an increasing function of k_t.

Furthermore, we make the following assumption[13]:

Assumption 2: There exists a threshold value $\tilde{k}$ sufficiently small, such that $w^u(k_t) < w^s(k_t)$, $\forall k_t \geq \tilde{k}$ with $\tilde{k} \geq 0$. We have $\tilde{k} = 0$ if $w^s(0) \geq w^u(0)$, and $\tilde{k} > 0$ if $w^s(0) < w^u(0)$.

[13]This assumption is reasonable as long as unskilled labor is not scarce.

Even if this assumption seems obvious, the general form we use does not necessarily implies that $w^u(k_t) < w^s(k_t)$, $\forall k_t \geq 0$. However, given assumption 1, we can reasonably assume that there exists a small threshold value $\tilde{k}$ such that the wage level of skilled agents is higher than the one of unskilled agents.

5.2.3 Government

We assume a PAYG pension system. The revenues of this system come from a proportional tax on wages: τ_t. These revenues are used to provide a pension for elderly people. Their pension depends on the wages of young agents having the same productivity as them, and on the average wage of the economy. Their respective weighting is λ and $(1-\lambda)$. The first part of this pension represents the Bismarckian component, whereas the second part represents the Beveridgian component of this system (Casamatta *et al.*, 2000). λ measures the indexation of pensions on the wages of agents. If $\lambda = 0$, each agent receives the same pension and the pension system is completely Beveridgian. Conversely, if $\lambda = 1$ the level of pensions only depends on the wage of agents and the pension system is purely Bismarckian. **The smaller λ is, the more this pension system is redistributive.**[14].

Consumers receive only a fraction ν (with $0 < \nu \leq 1$) of this weighted average, and only during their second period of life T^i. ν denotes the average replacement rate of the pension system.

The pension level of an agent of type i is:

$$p^i_{t+1} = \nu\left(\lambda w^i_{t+1} + (1-\lambda)\bar{w}_{t+1}\right)T^i \tag{5.10}$$

The budget constraint of the Government can be written:

$$\tau_t N_t \overline{w}_t = N_{t-1}\left[e\nu(\lambda w^s_t + (1-\lambda)\overline{w}_t)\varphi + (1-e)\nu(\lambda w^u_t + (1-\lambda)\overline{w}_t)\sigma\right] \tag{5.11}$$

with $\overline{w}_t$ the average wage of the economy. It is obtained by:

$$\overline{w}_t = e w^s_t + (1-e) w^u_t \tag{5.12}$$

[14]In this chapter the term "redistributivity" only concerns the instantaneous redistribution of pension systems and not the long run redistribution of pension systems. The long run redistribution, which is the discounted difference between tax paid and amount received, can be very different because of the life expectancy differential.

Some simple manipulations imply that the tax rate can be expressed as:

$$\tau_t = \frac{\nu}{1+n}\left[\varphi - \epsilon(1-e)\left(\lambda\frac{w^u(k_t)}{\overline{w}(k_t)} + 1 - \lambda\right)\right] \tag{5.13}$$

If $\epsilon = 0$, the tax rate is simply equal to the product between the replacement rate and the old-age dependency ratio of the economy (d'Autume, 2003). The second component between brackets is the ratio between the pensions not paid to unskilled agents because of their lower length of life, and the average wage of the economy. *Ceteris paribus*, under the reasonable assumption that $w_t^u < w_t^s$, $\forall t$, the tax rate is an increasing function of λ. This result is very intuitive. Indeed, educated agents have the longer length of life. Therefore, an increase in λ (i.e. a decrease in the redistributivity of the pension system) increases the indexation of pensions on wages. It implies that the pension of skilled agents increases. Moreover, they benefit from these pensions for a longer period of time than other agents. Consequently, the tax rate has to increase to finance these additional expenditures.

Another interesting point is that the tax rate τ_t is an increasing function of k_t. Indeed, under assumption 1, it is straightforward to show that $w^u(k_t)/\overline{w}(k_t)$ is a decreasing function of k_t. It means that the relative cost not paid to unskilled agents, because of their lower length of life, decreases with the level of capital per capita.

5.3 The Equilibrium and its Properties

All markets clear at each period t if:

$$L_t^s = eN_t \tag{5.14}$$

$$L_t^u = (1-e)N_t \tag{5.15}$$

$$K_{t+1} = eN_tS_t^s + (1-e)N_tS_t^u \tag{5.16}$$

The dynamics of the economy is obtained using equations (5.4), (5.6), (5.7), (5.8), (5.10), (5.12), (5.13), (5.14), (5.15) and (5.16).

It is straightforward to show that we obtain:

$$LHS(k_{t+1},\lambda) \equiv$$
$$(1+n)k_{t+1} + e\frac{p^s(k_{t+1},\lambda)}{\varphi(\beta R(k_{t+1}))^{1/\eta} + R(k_{t+1})} + (1-e)\frac{p^u(k_{t+1},\lambda)}{\sigma(\beta R(k_{t+1}))^{1/\eta} + R(k_{t+1})} =$$

$$(1-\tau(k_t,\lambda))[e\frac{\varphi(\beta R(k_{t+1}))^{1/\eta}}{\varphi(\beta R(k_{t+1}))^{1/\eta}+R(k_{t+1})}w^s(k_t)+$$

$$(1-e)\frac{\sigma(\beta R(k_{t+1}))^{1/\eta}}{\sigma(\beta R(k_{t+1}))^{1/\eta}+R(k_{t+1})}w^u(k_t)] \equiv RHS(k_{t+1},k_t,\lambda) \quad (5.17)$$

Since the dynamics of the economy is complicated we make two further assumptions[15]:

Assumption 3: The equilibrium trajectory is unique and increasing. It can be written: $k_{t+1}=\Psi(k_t)$, with $\Psi_1(k_t)>0$.

Assumption 4: There exists at least one non-trivial stable steady state (k^s_{SS}) such that $k^s_{SS}>\tilde{k}$.

These two assumptions are sufficient to establish the following proposition.

Proposition 5.1 *An increase in λ has a negative impact on every stable steady state (k^s_{SS}). Consequently, it also decreases wage inequalities.*

Proof: See appendix 1.□

Proposition 1 shows that a more redistributive pension system (a decrease in λ) increases the level of capital per capita. Indeed, a decrease in λ has two kinds of effects on saving. The first one is to decrease the *tax rate* τ for a given level of capital per capita. Then, the net wage of every consumer increases. It has a positive impact on saving. The second one is a *pension effect*. Unskilled agents decide to decrease their saving because they benefit from a more redistributive pension system. But at the same time, skilled agents increase their saving because pensions are less indexed on wages. The increase in saving of skilled agents overcompensates the decrease of the one of unskilled agents because skilled agents live for a longer period of time.

[15]We show in the next section that the following assumptions are checked for reasonable values of parameters.

Given our technology and under assumption 1, a decrease in λ also implies that wage inequalities (I_w) increase.

Let us now consider the impact of this redistributive policy on the wealth level and on the welfare level of every agent. These analytical results are obtained at steady state to simplify the exposition. Every derivative is thus a comparison between steady states.

The wealth level of an agent of type i, with $i \in \{s, u\}$, can be written:

$$W^i(\lambda) = w^i(k(\lambda))(1 - \tau(k(\lambda), \lambda)) + \frac{p^i(k(\lambda), \lambda)}{R(k(\lambda))} \tag{5.18}$$

with $k'(\lambda) < 0$. The net impact of a decrease in λ on $W^i()$ is:

$$-\frac{dW^i(\lambda)}{d\lambda} = \underbrace{-\frac{dw^i(k)}{dk}\frac{dk}{d\lambda}(1 - \tau(k, \lambda))}_{A>0} + \underbrace{w^i(k)\left(\frac{\partial\tau(k,\lambda)}{\partial k}\frac{dk}{d\lambda} + \frac{\partial\tau(k,\lambda)}{\partial\lambda}\right)}_{B \lesseqgtr 0}$$
$$\underbrace{-\frac{1}{R(k)}\left(\frac{\partial p^i(k,\lambda)}{\partial k}\frac{dk}{d\lambda} + \frac{\partial p^i(k,\lambda)}{\partial\lambda}\right)}_{C>0 \text{ if } i=u,\ C \lesseqgtr 0 \text{ if } i=s} + \underbrace{p^i(k,\lambda)\frac{\frac{dR(k)}{dk}\frac{dk}{d\lambda}}{(R(k))^2}}_{D>0} \tag{5.19}$$

Element A is positive because of the positive impact of the redistributive pension system on the wage level, through capital accumulation. Element B can be positive or negative. Indeed, as mentioned above, a more redistributive pension system has a direct negative impact on the tax rate because the pension system redistributes resources in favor of agents having a short life expectancy. However, as this policy increases capital accumulation, it reduces the relative share of expenditures not spent because of the mortality differential[16]. The net impact on the tax rate is thus ambiguous, but we can reasonably assume that the direct impact is higher than the one going through capital accumulation. It implies that a more redistributive pension system reduces the tax rate, which has a positive impact on the wealth level.

Element C has an ambiguous sign. Indeed, a more redistributive pension system increases the wage level of every agent, which has a positive impact on pensions. However, λ has a direct impact on pensions through the indexation of pensions on wages. For unskilled agents, a more redistributive pension system increases their pensions, and thus C is positive

[16] See the discussion about equation (5.13).

since every effect has the same sign. But for skilled agents, a more redistributive pension system decreases their pension for a given level of capital per worker. Consequently, the net impact on C is ambiguous for skilled agents.

Finally, element D is positive because of the negative impact of a more redistributive pension system on the interest factor.

We can reasonably conclude from this analysis that it is almost sure that unskilled agents highly benefit from a more redistributive pension system. However, the final impact is ambiguous for skilled agents. The indirect impact on capital accumulation has to be large for educated agents to benefit from this policy. It implies that the life expectancy differential has to be high[17].

Let us now consider the impact of this policy on the welfare level of agents.

Using equations (5.1), (5.2), (5.3) and (5.4) we obtain:

$$U^i(\lambda) = \frac{(c^i(\lambda))^{1-\eta}}{1-\eta} + \beta (T^i)^\eta \frac{(d^i(\lambda))^{1-\eta}}{1-\eta} \tag{5.20}$$

with:

$$c^i(\lambda) = \frac{W^i(\lambda)}{1 + \beta^{1/\eta}(R(k(\lambda)))^{(1-\eta)/\eta} T^i} \tag{5.21}$$

and,

$$d^i(\lambda) = \frac{W^i(\lambda)\beta^{1/\eta}T^i}{(R(k(\lambda)))^{-1/\eta} + \beta^{1/\eta}T^i(R(k(\lambda)))^{-1}} \tag{5.22}$$

The impact of a decrease in λ on consumption flows is *a priori* ambiguous. Indeed, a more redistributive pension system has an impact on the wealth level of agents. However, as it also has an impact on the interest factor, it influences the price of the second period consumption (d)[18]. Let us first consider that a decrease in λ has a positive impact on the wealth level of agents (W^i). Then, two cases have to be considered. The first one is such that $\eta < 1$. It implies that the intertemporal elasticity of substitution is high. In that case, a decrease in λ has a positive impact on the first period consumption (c). Indeed, the wealth level and the price of the second period consumption increase. Consequently, agents prefer increasing their first period consumption. The net impact on the second

[17] In the calibration exercise, we emphasize this point.

[18] The price of the second period consumption is the inverse of the interest factor. A decrease in R increases the price of the second period consumption.

period consumption is ambiguous. It depends on the scale of the wealth effect and of the price effect (through the interest factor).

The second case which has to be considered, is the case in which $\eta > 1$, i.e. the case of a small intertemporal elasticity of substitution. In that case, the impact of a more redistributive pension system on consumption levels is ambiguous whatever the period considered.

Finally, if a decrease in λ has a negative impact on the wealth level, then if $\eta > 1$, consumption levels of every period decrease. If $\eta < 1$, then the second period consumption decreases whereas the impact on the first period consumption is ambiguous. However, we can reasonably assume that the consumption level of every period will decrease.

Since we cannot determine *a priori* the impact of a more redistributive pension system on wealth, on welfare, and on wealth inequalities, we calibrate and we solve our model numerically.

5.4 Calibration and Results

Firstly, we specify our production function and we detail our calibration choices. Then, we give the results of the numerical resolution of our model.

5.4.1 Calibration

In this section, we specify the functional form of the production function and we calibrate our model to study the impact of the redistribtutivity of pension systems on macroeconomic variables.

Firstly, we assume that the production function has the following form:

$$F(L_t^u, X_t) = A\left[\alpha(L_t^u)^{\upsilon} + (1-\alpha)X_t^{\upsilon}\right]^{1/\upsilon} \tag{5.23}$$

and,

$$X_t \equiv G(K_t, L_t^s) = \left[bK_t^{\gamma} + (1-b)(L_t^s)^{\gamma}\right]^{1/\gamma} \tag{5.24}$$

with $A > 0$ the level of the technology, $\upsilon, \gamma < 1$ and $b, \alpha \in (0,1)$. Using Sato's (1967) results and the study of Duffy *et al.* (2004), we show that there exists a capital-skill complementarity if and only if $\upsilon > \gamma$. This condition is necessary and sufficient for capital and skilled labor to be more complementary than capital and unskilled labor[19].

It can be shown that the wage inequality ratio: $I_w = w^s(k_t)/w^u(k_t)$ is an increasing function of k_t if and only if $\upsilon > \gamma$.

We calibrate the parameters υ and γ of the production function to match the findings of Fallon and Layard (1975). They find that, for a restricted set of rich countries, the elasticity of substitution between capital and unskilled labor is 1.85, whereas the one between capital and skilled labor is 0.55. Given our technology, it implies that $\upsilon = 0.46$ and $\gamma = -0.81$.

Secondly, we calibrate the basic parameters of the model. The length of each period is 40 years. The growth rate of the population is $n = 0.3$. It corresponds to an annual growth rate of the population of 0.65% (Charpin, 1999). The pure time preference factor is $\beta = 0.6$ (d'Autume 2003), i.e. an annual psychological discount rate of 1.3%. Moreover, we assume that $\eta \geq 1$. It means that the intertemporal elasticity of substitution is low. Agents prefer smoothing their consumption. This assumption is in accordance with Attanasio *et al.* (1999) and with Cooley and Prescott (1995). It is the same as the one used in Casamatta *et al.* (2000). More specifically, we assume that $\eta = 1.5$ as in Sommacal (2006).

Moreover, we first assume that $\varphi = 0.55$, $\sigma = 0.45$ and thus that $\epsilon = 0.1$. It means that the length of life of an educated agent is 82 years[20], and the one of unskilled agents is 78 years. The life expectancy differential is 4 years. The life expectancy gap is smaller than the one found by Mesrine (1999) between the highest and the lowest socio-professional group. Since the indirect effects depend on the value taken by ϵ, we discuss the impact of a redistributive policy for different values taken by this parameter.

The share of the educated population is 0.4 ($e = 0.4$) as in Sommacal (2006) and in Acemoglu (2002). As in their studies, our model will have to match the wage gap I_w found by Acemoglu (2002).

Thirdly, we calibrate our model to match some empirical facts. The average replacement rate is fixed for the tax rate of the pension system to be around 20% as reported in Nyce and Schieber (2005). We obtain $\nu = 0.55$. This value is in accordance with the empirical

[19]See Duffy *et al.* (2004). This condition is necessary and sufficient using either the Allen-Uzawa partial elasticity of substitution, or the Hicks-Allen direct partial elasticity of substitution.

[20]In our model we do not include the very first period of life during which an agent is young. We assume that the length of this period is 20 years. Thus the life expectancy of skilled agents is obtained by: $20 + 40 + \varphi \times 40$.

Parameter	Value	Parameter	Value
n	0.3	η	1.5
β	0.6	A	3
φ	0.55	b	0.12
σ	0.45	α	0.23
ν	0.55	υ	0.46
e	0.4	γ	-0.81

Table 5.1: Calibration of the model

findings of Nyce and Schieber (2005), and this value is the same as the parametrization used in d'Autume (2003).

There only remains to fix α, b and A. We fix them for our model to reproduce three facts. Firstly, the capital share in total output has to be near 0.33. Secondly, the annual interest rate has to be in the reasonable interval $[0.03, 0.05]$. And finally, following Sommacal (2006) and Acemoglu (2002), the wage premium w^s/w^u has to be near 1.7. If $(A, b, \alpha) = (3, 0.12, 0.23)$, then these three facts are observed at the steady state of our economy. With this calibration, we find a tax rate τ for the pension system around 0.21. Table 1 summarizes our calibration.

5.4.2 Steady State Effects

In this section we analyze the impact of a decrease in λ on the economic variables at steady state (see Figures 2-10).

We observe that the capital level per capita is a decreasing function of λ as in our proposition 1. It implies that wage levels and wage inequalities increase with the redistributivity of pension systems. Another important point is that the decrease in λ has a negative impact on the pensions of educated agents. It means that the direct decrease in the indexation of pensions on wages overcompensates the increase in wages implied by a more redistributive pension system. Obviously, the pensions of unskilled agents decrease with λ because all effects go in the same way.

The total impact of a decrease in λ on the wealth of every agent is positive. It means that the decrease in the pensions of skilled agents is overcompensated by the increase in their wages. For unskilled agents the positive impact is more trivial as wages and pensions

Variables	$\epsilon = 0.02$	$\epsilon = 0.1$	$\epsilon = 0.15$
k	+	+	+
I_w	+	+	+
p^s	-	-	-
p^u	+	+	+
W^s	-	+	+
W^u	+	+	+
I_W	-	-	+
U^s	-	+	+
U^u	+	+	+
τ	-	-	-

Table 5.2: Impact of a decrease in λ on macroeconomic variables

increase.

We also find that wealth inequalities decrease with the level of redistribution of pension systems. It implies that the direct redistribution of pension systems overcompensates the increase in wage inequalities.

The tax rate is a decreasing function of the Beveridgian part of pension systems as can be observed empirically. Finally, we find that every agent of the economy benefits from a more redistributive pension system because his utility decreases with λ.

Let us now consider different values which can be taken by ϵ, denoting the mortality differential. It can be numerically shown that for $\epsilon \in (0, 0.15)$, A, b and α can keep the same value without altering the matching properties of our model too much. Let us recall that the larger ϵ is, the higher inequalities of length of life are. In table 2 we test the robustness of our results for different values of ϵ, keeping φ at its initial value ($\varphi = 0.55$). The sign + $(-)$ means that the redistributivity of pension systems has a monotonous positive (negative) impact on the variable.

Firstly, comparing the first column to the second one, we observe that if the life expectancy differential is not sufficiently high, the indirect effects of a more redistributive pension system are small, thus skilled agents do not benefit from such a policy. Their utility level decreases because the decrease in their pension level is higher than the increase in their wage level.

Secondly, comparing the case in which $\epsilon = 0.1$, to the case in which $\epsilon = 0.15$, we observe that each qualitative result remains unchanged for different values of σ, except for the wealth inequality ratio. Indeed, we observe that an increase in the redistributivity of pension systems decreases wealth inequalities as long as ϵ is not too large. However, for ϵ sufficiently high, an increase in the Beveridgian part of pension systems increases wealth inequalities. The main explanation is that if ϵ is sufficiently high, unskilled agents do not benefit from the redistributive properties of pension systems for a long time. Then, the increase in wage inequalities overcompensates the decrease in pension inequalities.

5.4.3 The Transitional Dynamics

Let us now consider the transitional dynamics of our macroeconomic variables if the pension system becomes more redistributive. To study the transitional dynamics, we assume that $\epsilon = 0.1$. Thus, we consider the case in which, at steady state, the welfare level of every agent increases. We try to know if there is a transitional cost for agents.

We assume that an economy is initially (at period 1) at steady state. This steady state is characterized by a given value of the parameter λ. The Government changes the value of this parameter from period 3 on, and every agent expects this change[21]. We study the dynamics of our model if the pension is initially Bismarckian ($\lambda = 0.885$, as in Hairault and Langot (2008) on the French case), and if it becomes more Beveridgian ($\lambda = 0.685$).

Concerning capital accumulation, we have $k_1 = k^i$, with k^i the steady state value of k with the initial value of λ. Moreover, as equation (5.17) determines k_2 with λ in the RHS and in the LHS, we also have: $k_2 = k^i$. However, as k_3 depends on the level of pensions of period 3 received by agents born in period 2, then k_3 differs from k^i. Then, k_t converges towards its new steady state value. Figure 5.11 illustrates this convergent dynamics.

Figure 5.13 shows that the welfare level of unskilled agents continuously increases along the transitional path. However, figure 5.12 shows that the welfare level of skilled agents of

[21] In the previous chapter we made the same exercise with an unexpected change in λ. We have obtained the same qualitative results.

generations 2 and 3 decreases compared to the welfare level of skilled agents of period 1, whereas in the long-run skilled agents benefit from this redistributive policy.

Finally, figure 5.14 shows that, at first, the wealth inequality ratio highly decreases because of the direct redistribution of public resources in favor of unskilled agents. However, from period 3 on, this ratio increases and converges towards its new value because of the increase in the wage inequality ratio (linked to capital accumulation).

5.5 Conclusion

In this chapter we show that an increasing redistribution of pension systems increases capital accumulation. Given our capital-skill complementarity technology, it implies that a more redistributive pension system increases wage inequalities. However, in a life-cycle perspective, this policy redistributes wealth among the population if the mortality differential is not too large. Moreover, it is possible even for rich agents to benefit from this structure because of the increase in capital accumulation.

A future work will introduce an endogenous labor supply and study the distorsive impact of a more redistributive policy. Intuitively, the tax rate should decrease and the capital accumulation should increase. It would imply that labor supply would be an increasing function of the degree of redistribution of pension systems. It would dramatically contrast with the results of Sommacal (2006).

5.6 APPENDIX

Appendix 1: Proof of Proposition 1.

Given assumption 3, equation (5.17) can be rewritten as:

$$LHS(\Psi(k_t),\lambda) - RHS(\Psi(k_t),k_t,\lambda) = 0$$

Differentiating this equation with respect to k_t gives:

$$LHS_1()\Psi_1() - RHS_1()\Psi_1() - RHS_2() = 0$$

with $f_i()$ the derivative of $f()$ with respect to its ith argument. It implies:

$$\Psi_1() = \tfrac{RHS_2()}{LHS_1()-RHS_1()}$$

Under assumptions 3 and 4 there exists at least one stable steady state k^s_{SS}. It implies that: $LHS_1() - RHS_1() - RHS_2() > 0$.

Let us now consider the net impact of an increase in λ on k^s_{SS}. We differentiate equation (5.17) with respect to k^s_{SS} and λ. We obtain:

$$\left[LHS_1() - RHS_1() - RHS_2()\right] dk^s_{SS} = \left[RHS_3() - LHS_2()\right] d\lambda$$

The factor before dk^s_{SS} is strictly positive under the assumption of stability of the equilibrium. To determine the sign of $dk^s_{SS}/d\lambda$ it is sufficient to know the sign of $(RHS_3() - LHS_2())$. We show that it is negative. Indeed as long as $w^s(k) > w^u(k)$, $RHS_3()$ is a decreasing function of λ because of the positive impact of λ on τ. There only remains to know the sign of $LHS_2()$. If it is positive then we prove the proposition.

$LHS()$ can be rewritten as:

$$LHS(k,\lambda) = (1+n)k + \frac{\nu\varphi\overline{w}(k)}{\varphi(\beta R(k))^{1/\eta} + R(k)} + (1-e)\nu(\lambda w^u(k) + (1 - \lambda)\overline{w}(k))\left[\frac{-\epsilon R(k)}{\left[(\varphi-\epsilon)(\beta R(k))^{1/\eta} + R(k)\right]\left[\varphi(\beta R(k))^{1/\eta} + R(k)\right]}\right]$$

It implies that as long as $w^s(k) > w^u(k)$, then $LHS_2() > 0$.□

Appendix 2[22]:

Firstly, let us give a preliminary result. We consider the case of a perfect competition on the final good market and on inputs markets.

Lemma 1: If there exists a production function $Y = F(K, L)$ satisfying the Inada conditions, and which is homogenous of degree 1 then we have:

$$\frac{dw}{w} = \frac{\alpha_K}{\sigma_{K,L}} \frac{dk}{k} = \frac{\alpha_K}{\sigma_{K,L}} \left[\frac{dK}{K} - \frac{dL}{L} \right] \tag{5.25}$$

and

$$\frac{dr}{r} = -\frac{1-\alpha}{\alpha} \frac{dk}{k} = -\frac{1-\alpha}{\alpha} \left[\frac{dK}{K} - \frac{dL}{L} \right] \tag{5.26}$$

with $k = K/L$, $\alpha_K = \frac{KF_K()}{Y}$ and $\sigma_{K,L}$ the elasticity of substitution between K and L defined by:

$$\sigma_{K,L} = \frac{dk/k}{d(\frac{w(k)}{r(k)})/(\frac{w(k)}{r(k)})} \tag{5.27}$$

Proof: We start with the result: $w = f(k) - kf'(k)$. Then it is sufficient to use equation 5.27 knowing that $dr/dk = f''(k) = -dw/dk \times 1/k$.□

There only remains to apply this result to the production function (5.5). Then we have:

$$\frac{dw^u}{w^u} = \frac{\alpha_F}{\sigma_F} \left(\frac{dG}{G} - \frac{dL^u}{L^u} \right) \tag{5.28}$$

or

$$\frac{dw^u}{w^u} = \frac{\alpha_F}{\sigma_F} \left(\alpha_G \frac{dK}{K} + (1-\alpha_G) \frac{dL^s}{L^s} - \frac{dL^u}{L^u} \right) \tag{5.29}$$

α_F denotes the share of $G()$ in function $F()$, α_G denotes the share of K in $G()$, and σ_F denotes the elasticity of substitution between the two arguments of $F()$. In the same way, using the fact that at equilibrium we have:

$$w^s = F'_G . G'_{L^s} \tag{5.30}$$

we have:

$$\frac{dw^s}{w^s} = -\frac{1-\alpha_F}{\sigma_F} \left(\frac{dG}{G} - \frac{dL^u}{L^u} \right) + \frac{\alpha_G}{\sigma_G} \left(\frac{dK}{K} - \frac{dL^s}{L^s} \right) \tag{5.31}$$

[22] I thank A.d'Autume for the proof presented here.

with σ_G the elasticity of substitution between the two inputs of function $G()$. The previous equation implies:

$$\frac{dw^s}{w^s} = \left[-\frac{1-\alpha_F}{\sigma_F}\alpha_G + \frac{\alpha_G}{\sigma_G}\right]\frac{dK}{K} + \left[-\frac{1-\alpha_F}{\sigma_F}(1-\alpha_G) - \frac{\alpha_G}{\sigma_G}\right]\frac{dL^s}{L^s} + \frac{1-\alpha_F}{\sigma_F}\frac{dL^u}{L^u} \quad (5.32)$$

Assumption A1 implies that for $dL^s/L^s = dL^u/L^u = 0$ and if $dK/K > 0$ then:

$$\frac{dw^s}{w^s} > \frac{dw^u}{w^u} \iff -\frac{1-\alpha_F}{\sigma_F}\alpha_G + \frac{\alpha_G}{\sigma_G} > \frac{\alpha_F}{\sigma_F}\alpha_G \iff \frac{1}{\sigma_G} > \frac{1}{\sigma_F} \quad (5.33)$$

It means that wage inequalities increase with capital accumulation iff the elasticity of substitution between the arguments of function $F()$ is higher than the one between the arguments of function $G()$.□

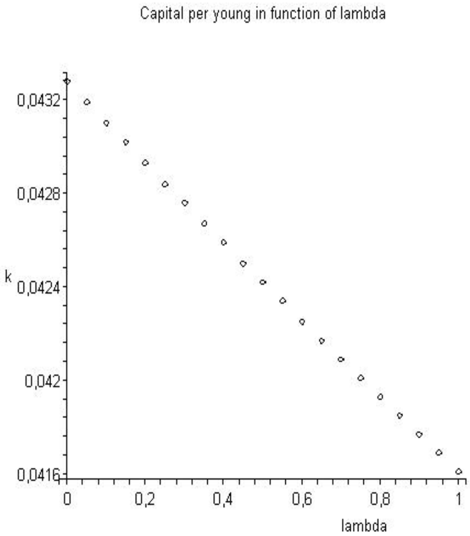

Figure 5.2: Capital per young (k^s) in function of λ

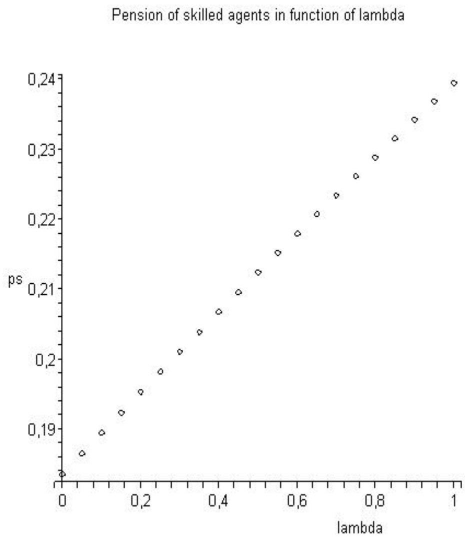

Figure 5.3: $p^s(k)$ in function of λ

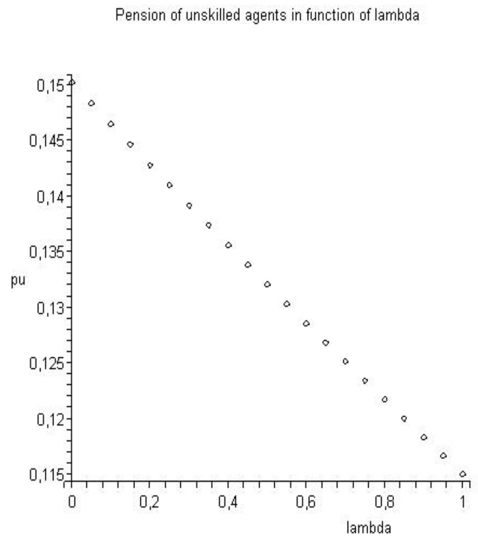

Figure 5.4: $p^u(k)$ in function of λ

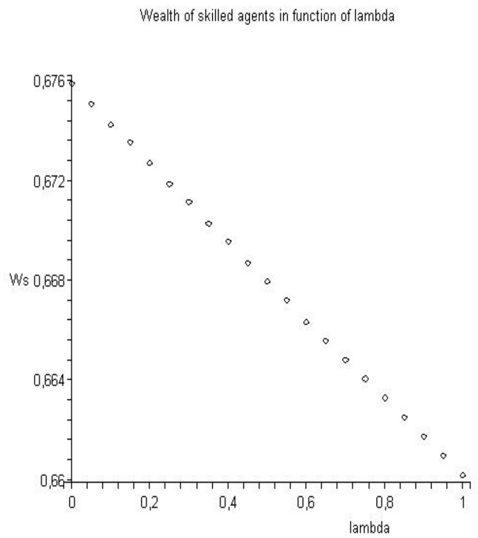

Figure 5.5: $W^s(k)$ in function of λ

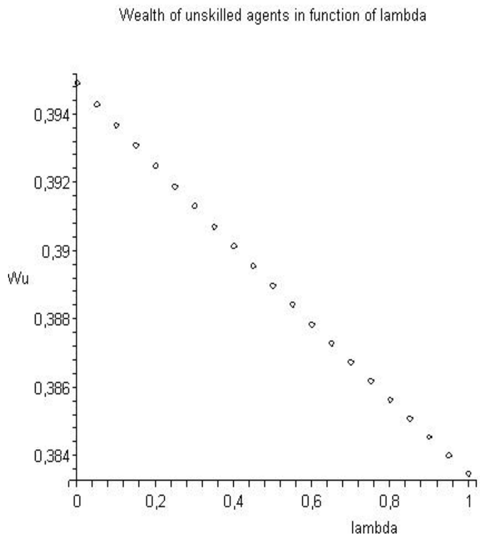

Figure 5.6: $W^u(k)$ in function of λ

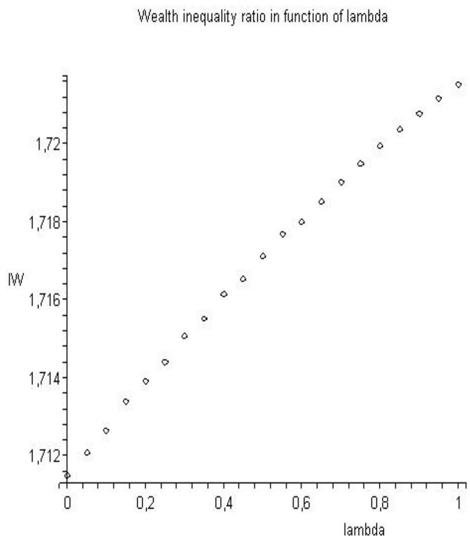

Figure 5.7: $I_W(k)$ in function of λ

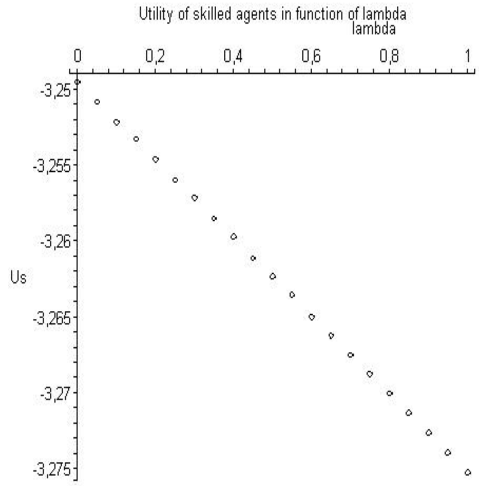

Figure 5.8: $U^s(k)$ in function of λ

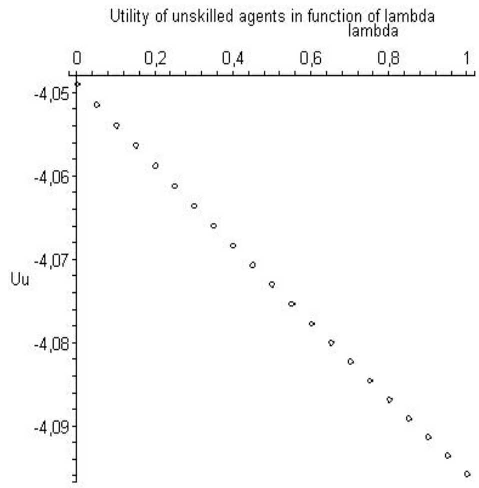

Figure 5.9: $U^u(k)$ in function of λ

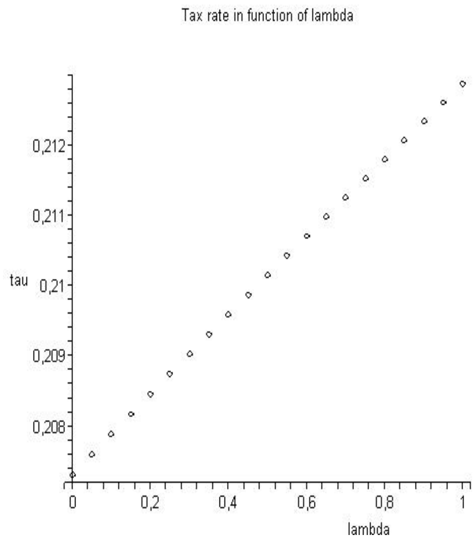

Figure 5.10: $\tau(k, \lambda)$ in function of λ

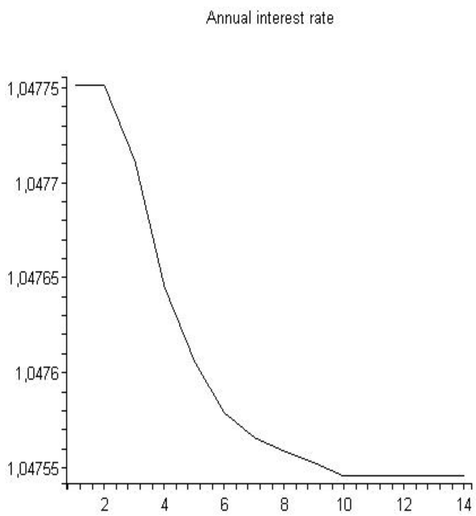

Figure 5.11: Transitional dynamics of the Annual Interest Rate

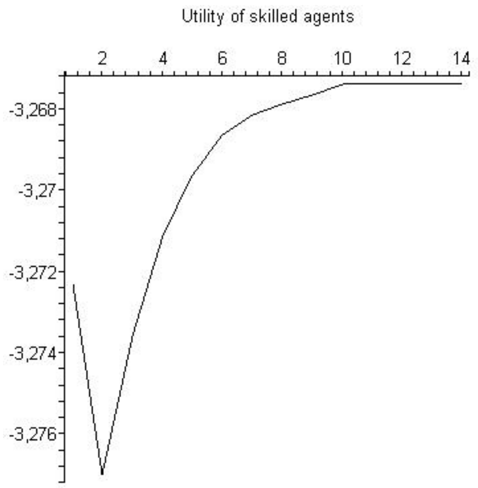

Figure 5.12: Transitional dynamics of U^s

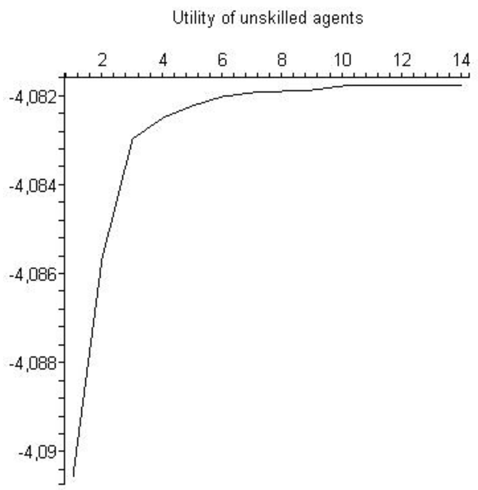

Figure 5.13: Transitional dynamics of U^u

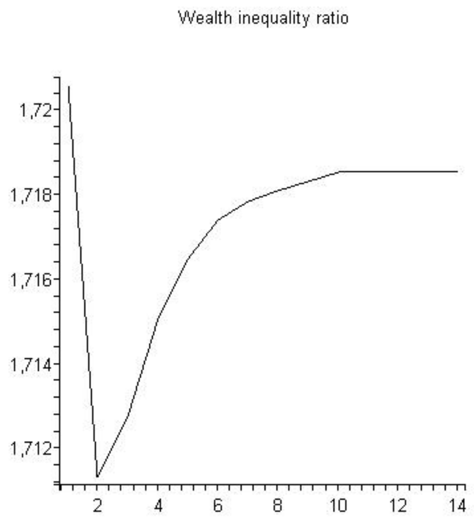

Figure 5.14: Transitional dynamics of I_W

Part III

The Long Run Redistributive Properties of Pension Systems and their Consequences

Chapter 6

Who Really Benefits from Pension Systems? When Life Expectancy Matters

6.1 Introduction

A wide literature studies the macroeconomic impact of the ageing of the population, notably on the sustainability of pension systems[1]. This ageing process implies that reforms have to be adopted by developed countries in order to limit the fiscal burden of pension systems. Another dimension of the length of life has been explored recently. Indeed, agents differ by their life expectancy. More particulary, these differences between socio-professional groups are wide. Mesrine (1999) studies the inequalities of length of life according to socio-professional groups in France[2]. The most striking feature of his paper is that a worker has a probability to die between 35 and 65 almost twice higher than that of an executive manager. Furthermore, their life expectancy at 35 is 38 and 44 respectively. The same qualitative results are observed in the United-States (Panis and Lillard 1995, Deaton and Paxson 2000). Finally, Robert-Bobbée and Cadot (2007) show that this inequality is also observed for elderly people. For agents who are 86, the ones with highest education level can expect to live 20% longer than the ones with lowest education level.

As socio-professional groups are linked to earnings, we can conclude from the previous results that earnings have an impact on the length of life. Some empirical studies deal with this link[3]. This life expectancy differential can have strong implications for the redistributive properties of pension systems. Indeed, as rich agents live and benefit from a pension for a longer period of time, pension systems are not as redistributive as they seem. Consequently, to study these properties, two dimensions have to be analyzed: (1) the distinction between Beveridgian and Bismarckian systems, and (2) the relationship between wages and life expectancy. Some authors have studied the empirical implications of this distinction. Coronado *et al.* (2000) and Liebman (2001) show that the pension system of the United-States is far less progressive than is usually mentioned. Legros (1994) and Bommier *et al.* (2003) find the same qualitative results on french data.

However, theoretical implications of these new results have not been clearly studied. Mitchell and Zeldes (1996) explain that:

[1]See d'Autume (2003), Charpin (1999), Cremer and Pestieau (2000) or Galasso and Profeta (2004) among others. Here and in the rest of this chapter we only consider Pay-As-You-Go pension systems.

[2]These inequalities also depend on other factors such as sex or the geographical localization. For example, in France, the life expectancy of women is 84.1, whereas that of men is only 77.2 (INSEE, 2006). Moreover, Rican and Salem (1999) show that there are strong disparities according to the localization of people in France.

[3]See Attanasio and Emmerson (2001), Bommier *et al.* (2003) or Adams *et al.* (2003) for a survey.

> *"Despite its intent, the system [the pension system] is less progressive than it might seem, because there is a positive correlation between lifetime earnings and length of life."*

They emphasize the main role played by the inequalities of length of life but this sentence raises a question: can a pension system be regressive?

Borck (2007) has exploited this idea. He shows that the size of a pension system can be determined by a coalition of elderly, very poor and very rich agents. Poor agents benefit from the Beveridgian part of the pension system, whereas rich agents benefit from the pension system for the longest period of time. This paper of Borck can be seen as a study on the consequences of the redistributive properties of pension systems when inequalities of length of life are taken into account. Nevertheless, it is not an analysis of the redistributivity itself. Consequently, this chapter is the next step to clarify this last point analytically.

A pension system is purely Beveridgian if every agent receives the same pension. Conversely, a pension system is purely Bismarckian if pensions completely depend on the wages of agents. A pension system is mixed if it has a Beveridgian and a Bismarckian component. The more Beveridgian a pension system is, the higher intra-generational transfers are. Countries highly differ by this intra-generational component. France, Germany and Italy have a Bismarckian structure. Canada, the Netherlands and New-Zeland are essentially Beveridgian. Finally, Japan, the United-Kingdom and the United States have mixed pension systems (Sommacal 2006, Casamatta *et al.* 2000).

In this chapter, we use an overlapping generations model in which agents differ by their wage and by their length of life. The pension system can be Beveridgian, Bismarckian or a mix of the two as in Casamatta *et al.* (2000). To study the redistributivity of pension systems we use the concept of "net contributions". The net contribution of an agent with a given wage, is the actualized difference (at the growth rate of the population) between the tax paid and the benefit from the pension system (Drouhin 2001a). A positive net contribution implies that agents pay more than they receive from the pension system.

In order to understand the qualitative changes induced by the inequalities of length of life we first consider that every agent has the same length of life. In this specific case, it is possible to show that if pension systems are at least partially Beveridgian then it is also progressive. But if the pension system is Bismarckian then the net contribution for each agent is nil. Afterwards, we introduce inequalities of length of life and we consider the case of a Beveridgian, a Bismarckian and of a mixed pension system successively. If the pension system is Beveridgian then it is progressive, but the poorest do not necessarily benefit the most from the pension system. Furthermore, the share of the population who benefits from a negative net contribution changes. If the pension system is Bismarckian, then we show that it is regressive because the poorest have the most positive net contribution and the richest the most negative one. If we have a mixed pension system our analytical results in terms of net contributions can be generalized only if pension systems tend towards either a Beveridgian or a Bismarckian structure. For intermediate cases, a numerical resolution calibrated on French data is used. It does not describe the exact structure of the French pension system but it emphasizes its important qualitative properties. We use different calibrations for the function which links the length of life to the wage level. We show that it is possible that the ends of the distribution of wages benefit from a negative net contribution.

This chapter is organized as follows. In section 2, we present our model. In section 3, the study of the redistributive properties of pension systems without inequalities of length of life is detailed. In the following sections we assume that the length of life is linked to the wage level. In section 4 and 5, we study a Beveridgian and a Bismarckian pension system respectively. In section 6, we emphasize the main properties of mixed pension systems. In section 7 we calibrate our model on French data. Section 8 provides some concluding remarks.

6.2 The Model

We consider a small open economy in which agents live two periods[4]. At each period t, the number of young agents is N_t. The population is assumed to grow at a constant rate n, such that $N_t = (1+n)N_{t-1}$. These agents are heterogenous since each of them has a wage w which belongs to the interval $\Omega_w = [w_-, w_+]$, with $w_- > 0$. Wages are distributed randomly among the population. $f(w)$ denotes the density function of the random variable w. Consequently, it is also the fraction of the population having a wage level w[5]. The average wage of this economy can be written:

$$\bar{w} = \int_{\Omega_w} w f(w) dw \tag{6.1}$$

[4]The length of each period is normalized to 1.

[5]We assume that the size of each generation is sufficiently large to apply the law of large numbers.

Our framework is static because we assume that wages and the distribution of wages are constant over time.

Furthermore, we assume that agents live only a fraction T of their second period of life[6]. This length of life is supposed to be linked to the wage level of agents: $T = T(w)$, and more specifically we assume that $T'(w) > 0$. It represents the inequalities of length of life according to socio-professional groups. The average length of life is:

$$\bar{T} = \int_{\Omega_w} T(w)f(w)dw \tag{6.2}$$

The linkage between the wage level and the length of life is measured by the covariance:

$$COV_{T,w} = \int_{\Omega_w} T(w)wf(w)dw - \bar{T}\bar{w} \tag{6.3}$$

As we assume that $T'(w) > 0$, then we have $COV_{T,w} > 0$[7].

Moreover, we make the following assumption about the function $T(w)$:

Assumption 1: $T'(w) > 0$, $T''(w) < 0$, $E_{T/w} = T'(w)\frac{w}{T(w)} < 1$, and $-T''(w)\frac{w}{T'(w)} < 2$.

The first part of this assumption is standard and represents the decreasing marginal impact of wages on the length of life. The assumption on the elasticity implies that an increase in wages of $x\%$ implies an increase in the length of life of less than $x\%$. The last part of this assumption is only a technical hypothesis which will be used later in this chapter[8].

Each agent works when he is young and retires at the end of his first period of life[9]. It is the same assumption as the one used in Casamatta *et al.* (2000) or in Borck (2007). As

---

[6]There is no uncertainty in our model. Consequently agents are sure to live until the end of the fraction T of their second period of life. However, our model can also be interpreted as a model in which $(1 - T)$ would be the probability of dying at the end of the first period of life as in Drouhin (2001a, 2001b).

[7]See Appendix A.

[8]Some main functions respect this property. For example: $T(w) = aw+b$, or $T(w) = \gamma w^{\xi}$ with a,b,$\gamma > 0$ and $1 > \xi > 0$.

[9]In this chapter we do not consider the length of education and the retirement age even if they have an impact on the redistributive properties of pension systems. Indeed, the length of education and the retirement age are positively correlated with the wage level. Then, a strong link between education and wages reduces the progressivity of pension systems, whereas the positive correlation between the retirement age and the wage level increases the progressivity of pension systems. However, in this chapter we only

long as an agent works, he pays a payroll tax τ. This tax is used to finance a PAYG pension system. When old, an agent receives a pension $p(w)$. Pensions are paid as long as agents are still alive, i.e. during a fraction $T(w)$ of their second period of life (d'Autume 2003). Furthermore, pensions per unit of time are partly indexed on the wage of the first period of the agent and on the average wage of the economy $(\lambda w + (1-\lambda)\bar{w})$. λ measures the size of the Bismarckian part of the pension system, whereas $(1-\lambda)$ measures the Beveridgian part of the pension system. When $\lambda = 1$ then the pension system is completely Bismarckian because pensions are only indexed on the wages of each agent. Conversely, when $\lambda = 0$ every agent receives the same pension. In that case, the pension system is Beveridgian[10].

Finally, agents receive only a fraction (ν) of this weighted average $(\lambda w + (1-\lambda)\bar{w})$ per unit of time. ν denotes the average replacement rate of the pension system. Consequently, the pension $p(w)$ which an agent receives during his second period of life is[11]:

$$p(w) = \nu(\lambda w + (1-\lambda)\bar{w})T(w) \tag{6.4}$$

We assume that there is no debt in this economy. It implies that all pensions have to be financed by a tax on wages. The budget constraint of the Government can be written:

$$N_t \int_{\Omega_w} \tau w f(w) dw = N_{t-1} \int_{\Omega_w} \nu(\lambda w + (1-\lambda)\bar{w})T(w) f(w) dw$$

Some straightforward calculations imply that:

$$\tau = \nu \frac{\bar{T}}{1+n}(1 + \lambda \rho^2) \tag{6.5}$$

with $\rho^2 \equiv \frac{COV_{T,w}}{\bar{T}\bar{w}}$. ρ^2 denotes the covariation coefficient between life expectancy and the wage level.

---

analyse the redistributive properties of unfunded pension systems when the life expectancy differential is taken into account. In this way we emphasize the main role played by the mortality differential, neutralizing other channels.

[10] This formula does not capture the high complexity of real pension systems, but its simplicity is useful (i) to capture the basic redistributive properties of pension systems, and (ii) to simply represent in the same framework different pension systems. Obviously, the usual redistributive properties of pension systems of each country would need more attention, and the use of micro-simulation models such as Liebman (2001).

[11] The aim of this chapter is not to capture the entire complexity of pension systems but just to give some intuitions about their redistributive properties through a theoretical model. That is why we use the simplifying formula (6.4).

Let us first study the case in which every agent has the same length of life, that is: $T(w) = \bar{T}$, $\forall w$. In that case we have $COV_{T,w} = 0$, and the tax rate becomes (d'Autume 2003):

$$\tau = \nu \frac{\bar{T}}{(1+n)}$$

The higher the replacement rate is, the higher the tax rate used to finance the pension system is. Indeed, an increase in the generosity has to be financed by a higher tax on young agents. Furthermore, the higher the old age dependency ratio $\left(\frac{\bar{T}}{(1+n)}\right)$ is, the higher the tax rate has to be. If for each worker there are more old agents, and for a given generosity of the pension system, then the tax rate has to increase to finance these additional pensions.

Let us now consider the case where $COV_{T,w} > 0$, i.e. there are inequalities of length of life. *Ceteris paribus*, the introduction of inequalities of length of life has a positive impact on the tax rate of the pension system. Indeed, the richer agents are, the longer they live and then the longer the period during which they receive a pension is. Conversely, the poorer agents are, the shorter the period during which they receive a pension is. Even if pensions are partially indexed on the wage of agents (by a coefficient λ) then the highest pensions are paid to people who live for the longest period. The tax rate has to increase to finance this additional spending. Now, if the pension system has a Beveridgian structure ($\lambda = 0$), then even if the covariance is large, the decrease in spending for agents with a length of life smaller than $\bar{T}$ exactly compensates the increase in spending for the others[12].

At each period t, a group of agents with a wage w pays: $\tau w f(w) N_t$, whereas at the same time agents with the same productivity receive $p(w)f(w)N_{t-1}$. Our main objective is to know if, at each period, a group of agents with a wage w receive more from the pension system than they pay for it, i.e. if $p(w)f(w)N_{t-1}$ is larger than $\tau w f(w) N_t$. But the model also makes it possible to determine the wage of the group of people who benefit the most from the pension system, i.e. who receive the most given the amount they pay.

At each period t, the net contribution of a group with a wage w is:

$$CN_{w,t} = \tau w f(w) N_t - p(w)f(w)N_{t-1}$$

or,

[12] Because of the linearity of pensions in (6.4).

$$CN_{w,t} = N_t f(w) \left[\tau w - \frac{p(w)}{1+n} \right]$$

The member between brackets represents the net contribution of an agent with a wage w if he uses the growth rate of the population as actualisation rate (Drouhin 2001a). A positive net contribution means that a group pays more for the pension system than he receives from it. Using equations (6.4) and (6.5) we obtain:

$$CN_{w,t} = \nu f(w) N_{t-1} \left[\bar{T}(1+\lambda\rho^2)w - (\lambda w + (1-\lambda)\bar{w})T(w) \right] \tag{6.6}$$

Integrating this function over the interval Ω_w, it is straightforward to show that $\int_{\Omega_w} CN_{w,t}dw = 0$.

But this amount is biased by the size of each group. Indeed, net contributions are correlated with the size of the groups of agents. So it is better to use the individual net contribution for each group. It can be written:

$$CN^i_{w,t} = \nu \left[\bar{T}(1+\lambda\rho^2)w - (\lambda w + (1-\lambda)\bar{w})T(w) \right] \equiv \nu A(w) \tag{6.7}$$

Note that the size of the pension system (ν) has only a quantitative effect because it does not influence the sign of $A(w)$. ν amplifies the net contribution of each agent. In order to well understand the relationship between net contributions and the wage level, we differentiate equation (6.7) with respect to w:

$$\nu A'(w) = \nu \left[\underbrace{\bar{T}(1+\lambda\rho^2)}_{C} \underbrace{-\lambda T(w)}_{D} \underbrace{-(\lambda w + (1-\lambda)\bar{w})T'(w)}_{E} \right]$$

If every agent has the same length of life ($T(w) = \bar{T}$, $\forall w$), then $A'(w) = \bar{T}(1-\lambda) > 0$, i.e. an increase in the wage level has a positive impact on the net contribution. Indeed, the increase in the tax paid is higher than the increase in the amount of pensions received. However, once inequalities of length of life are introduced ($T'(w) > 0$), then the increase in the wage level can have a stronger impact on the amount of pensions received. The element C of the previous equation represents the increase in contributions due to the higher wage level of agents. The element D represents the increase in the pension received during the length of life of the agent because of the indexation of pensions on wages. The element E represents the increase in the pension received due to an increase in life expectancy. Indeed, the wage level has a positive impact on the length of life, which implies that agents endowed with a high wage level, benefit from the pension system for a longer period of time. Thus, D and E can overcompensate C, i.e. an increase in the wage level can have a negative impact on the net contribution of agents.

Consequently, the form and the sign of the function $A(w)$ are not trivial because of the function $T(w)$. To understand the main implications of such a function we first study the case where there are no inequalities of length of life. Then, we introduce inequalities of length of life and we study the case where $\lambda = 0$ and the case where $\lambda = 1$ successively. Finally, we give the main properties of a mixed pension system, i.e. a system with a Beveridgian and a Bismarckian part.

6.3 The Benchmark Case

This section details the results for the case in which each agent in our economy has the same length of life: $T(w) = \bar{T}$, $\forall w$. It is an usual assumption. Agents only differ by their wages. This uni-dimensionality of the heterogeneity often simplifies the analysis but masks a very different reality. Let us first study the conclusions that would be obtained if we had only considered wage inequalities.

If $T(w) = \bar{T}$, $\forall w$, then $COV_{T,w} = 0$ and finally:

$$A_1(w) = (1 - \lambda)\bar{T}(w - \bar{w}) \tag{6.8}$$

$A(w)$ is a strictly increasing function of w as long as $\lambda \in [0, 1)$ (see figure 6.1). Furthermore, the net contribution is negative for agents having wages below the average. It implies that poor agents receive more from the pension system than they pay for it. Conversely, the net contribution is positive for agents having wages above the average wage.

Proposition 6.1 *(i) Every agent with a wage below (above) the average wage has a negative (positive) net contribution. (ii) The poorest benefit the most from the redistributive properties of the pension system. This net benefit is a decreasing function of w.*

Proof: (i) $A(w)$ is negative for $w < \bar{w}$ and positive for $w > \bar{w}$. (ii) $A(w)$ is a strictly increasing function of w.□

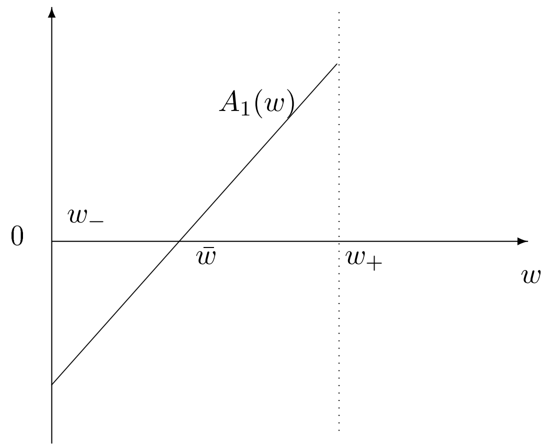

Figure 6.1: $A_1(w)$ for $T(w) = \bar{T}$ and $\lambda \in [0, 1)$

The second part of this proposition implies that the poorest have the highest negative net contribution. This result depends on the assumption that the pension system is at least partly Beveridgian ($\lambda \in [0, 1)$). If $\lambda = 1$ then the pension system is completely Bismarckian because the pension is only indexed on the wage of agents.

Proposition 6.2 *If* $\lambda = 1$ *then the net contribution of each group is nil. The pension system is not redistributive.*

Proof: See equation (6.8).□

This result ensures that a Bismarckian pension system, in an economy in which agents only differ by their wages, is neutral in terms of net contribution. Pensions exactly compensate contributions to the pension system.

These two results are usual in the economic literature (Casamatta *et al.* 2000). The following sections show that these results depend on the assumption that every agent has the same length of life. If it is not the case, then the higher the wage is, the longer agents live, and the more they benefit from the pension system. Consequently, the intuition is that the redistributive properties mentioned above change. Firstly because the poorest do not necessarily benefit the most from a Beveridgian pension system. Secondly because the

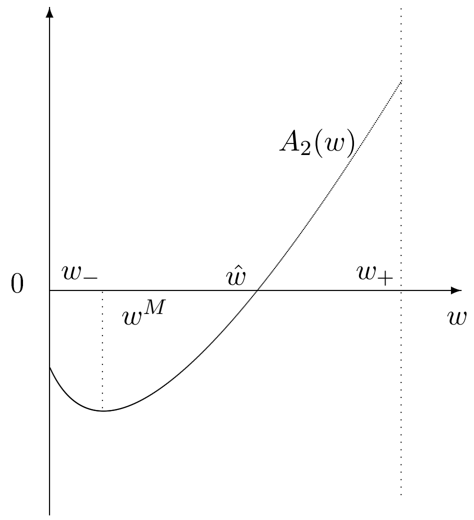

Figure 6.2: $A_2(w)$ with $w^M > w_-$

Bismarckian pension system is not neutral.

6.4 Pure Beveridgian Pension Systems

For pure Beveridgian pension systems we have $\lambda = 0$, i.e. every agent receives the same pension. This pension is indexed on the average wage of the economy. From this section we assume that agents also differ by their length of life: $T(w)$, with $T'(w) > 0$. The expression for $A(w)$ becomes:

$$A_2(w) = w\bar{T} - \bar{w}T(w) \tag{6.9}$$

The properties of this function are such that: $A_2'(w) = \bar{T} - \bar{w}T'(w)$ and $A_2''(w) = -\bar{w}T''(w)$. The sign of $A_2'(w)$ is indeterminate, but $A_2''(w)$ is clearly positive under assumption 1 (see figure 6.2).

Proposition 6.3 *There exists a threshold $\hat{w}$ such that the net contribution is negative (positive) for $w < \hat{w}$ ($w > \hat{w}$). Furthermore $\hat{w} > \bar{w}$ iff $\bar{T} < T(\bar{w})$.*

Proof: Under assumption 1, we know that $T(w)/w$ is a decreasing function of w. Furthermore, $A_2(w) > 0$ iff $T(w)/w < \bar{T}/\bar{w}$. $A_2(w)$ cannot be positive for every w since the sum of net contributions is equal to 0. Finally, as $T(w)/w$ is a decreasing function of w, we conclude that there exists a threshold value $\hat{w}$ such that $A_2(w) < 0$ for $w < \hat{w}$, and $A_2(w) > 0$ for $w > \hat{w}$. $\hat{w} > \bar{w}$ iff $A_2(\bar{w}) < 0$.$\square$

This result ensures that poor agents (with a wage below $\hat{w}$) have a negative net contribution. Moreover, if the average length of life is smaller than the length of life of the average wage, then a larger group benefits from the redistributive effect of the pension system. The more $\bar{T}$ is different from $T(\bar{w})$, the more $\hat{w}$ removes away from $\bar{w}$. This result is intuitive. Indeed, a high average length of life implies that the taxe rate of the pension system is higher. Then the share of the population who benefits from a negative net contribution decreases ($\hat{w} < \bar{w}$).

But we also have to compare the net contribution $A_2(w)$ with $A_1(w)$. It is easy to show that $A_2(w) > A_1(w)$ ($<$) as long as $T(w) < \bar{T}$ ($>$). It implies that poor workers benefit less from the pension system and that rich workers pay less for the pension system.

Because of the convexity of the function $A_2(w)$, if there exists a value w^M such that $A_2'(w^M) = 0$, then it is a minimum. It has two implications for our analysis. (i) The richest contribute the most to the pension system because of the convexity of $A_2(w)$. They have the highest positive net contribution. (ii) The poorest do not necessarily benefit the most from the redistributivity of the pension system.

Lemma 6.1 *$\hat{w} > w^M > w_-$ iff $A_2'(w_-) < 0$. $w^M = w_-$ iff $A_2'(w_-) \geq 0$.*

Considering $A_2'(w_-)$, this expression is negative if $T'(w_-)$ is sufficiently large. It means that the marginal impact of wages on the length of life is large for the low values of wages. Agents with a wage slightly higher than w_- can expect to live much longer than agents with a wage w_-. They receive a pension during this additional time. Consequently, the agents from the group w^M benefit the most from the pension system and w^M can be different from w_-.

The two main implications are: (i) The share of the population who benefits from the

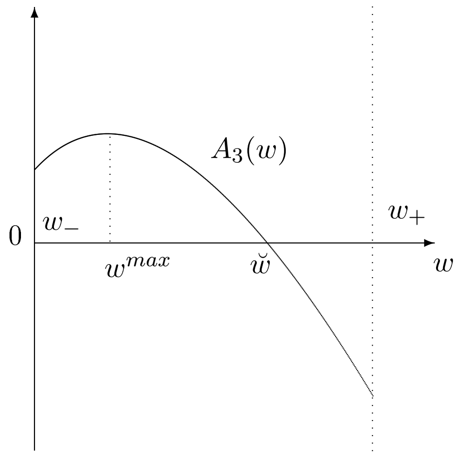

Figure 6.3: $A_3(w)$ for $w^{max} > w_-$

pension system can differ from the interval $[w_-, \bar{w}]$. (ii) The agents who benefit the most from the pension system are not necessarily the poorest. It depends on the properties of the function $T(w)$.

6.5 Pure Bismarckian Pension Systems

For pure Bismarckian pension systems, pensions are only indexed on wages ($\lambda = 1$). In section 3, we obtained that the net contributions are nil for every group if there are no inequalities of length of life. But the introduction of these inequalities changes the qualitative results considerably.

For $\lambda = 1$, $A(w)$ can be written:

$$A_3(w) = w\bar{T}(1+\rho^2) - wT(w) \tag{6.10}$$

The sign of $A_3'(w) = \bar{T}(1+\rho^2) - T(w) - wT'(w)$ is indeterminate, but $A_3''(w) = -2T'(w) - T''(w)w$ is clearly negative under assumption 1. Then $A_3(w)$ is a concave function of w (see figure (6.3)).

Proposition 6.4 *There exists a threshold $\breve{w}$ such that the net contribution is positive (negative) for $w < \breve{w}$ ($w > \breve{w}$). $\breve{w}$ is determined by the following equation:*

$$T(\breve{w}) = \bar{T}(1+\rho^2)$$

Proof: $A_3(w) > 0$ iff $T(w) < \bar{T}(1+\rho^2)$. As $T(w)$ is a strictly increasing function of w and as the sum of the net contributions is equal to 0, then as long as $w < \breve{w}$ we have $T(w) < T(\breve{w})$ and the net contribution is positive. But for $w > \breve{w}$ the net contribution is negative.□

This result is completely different from that of proposition 2. Rich agents benefit from a negative net contribution whereas poor agents have a positive net contribution. But the result of this section is even more surprising. Indeed, the highest negative value of $A_3(w)$ is obtained for $w = w_+$. Furthermore, as $A_3(w)$ is a concave function of w then if there exists a value w^{max} such that $A_3'(w) = 0$, it is a maximum. Given the result of proposition 4 we can write that $w^{max} \in [w_-, \breve{w})$, and more precisely:

Lemma 6.2 *$w^{max} = w_-$ as long as $A_3'(w_-) = \bar{T}(1+\rho^2) - T(w_-) - w_- T'(w_-) \leq 0$. $w^{max} \in (w_-, \breve{w})$ iff $A_3'(w_-) > 0$.*

$A_3'(w_-)$ is positive if $T'(w_-)$ is not too large, i.e. if agents with a wage slightly higher than w_- can expect to have a life expectancy only just higher than that of agents with a wage w_-.

Two main conclusions can be drawn from this analysis. (i) Pure Bismarckian pension systems are regressive because poor agents have a positive net contribution and rich agents have a negative one. (ii) If $T'(w_-)$ is sufficiently large, then the poorest have the highest positive net contribution. The concavity of $A_3(w)$ also implies that the richest have the highest negative net contribution, which reinforces our previous conclusion.

6.6 Mixed Pension Systems

Let us assume from now on that $\lambda \in (0, 1)$. If λ tends towards 1 then the pension system becomes more Bismarckian. Conversely, if λ tends toward 0 the pension system becomes more Beveridgian as the pension depends less on the wage of agents. The function $A(w)$ can be written as:

$$A_4(w,\lambda) = w\bar{T}(1+\lambda\rho^2) - (\lambda w + (1-\lambda)\bar{w})T(w) \tag{6.11}$$

But this function has indeterminate properties. Indeed $\partial A_4(w,\lambda)/\partial w$ can be positive, negative or null. And $\partial^2 A_4(w,\lambda)/\partial w^2$ can also be positive, negative or null.

In order to obtain clear analytical results we specify the function $T(w)$. We assume it has the following form:

$$T(w) = \varrho w^{\xi} \tag{6.12}$$

with $\varrho > 0$ and $1 > \xi > 0$. This function respects each property of assumption 1. Let us first study the concavity and the convexity of the function $A_4(w)$. Note that:

$$\frac{\partial A_4(w,\lambda)}{\partial w} = \bar{T}(1+\lambda\rho^2) - T'(w)(\lambda w + (1-\lambda)\bar{w}) - \lambda T(w) \tag{6.13}$$

and that:

$$\frac{\partial^2 A_4(w,\lambda)}{\partial w^2} = -T''(w)(\lambda w + (1-\lambda)\bar{w}) - 2\lambda T'(w) \tag{6.14}$$

Lemma 6.3 *There exists a threshold value $w_p(\lambda)$ such that if $w < w_p(\lambda)$ ($> w_p(\lambda)$) then $\frac{\partial^2 A_4(w,\lambda)}{\partial w^2} > 0$ (< 0). It is such that $\frac{\partial w_p(\lambda)}{\partial \lambda} < 0$. Furthermore, there exists an interval $(0,\lambda_+)$ such that for $\lambda \in (0,\lambda_+)$ we have $w_p(\lambda) > w_+$; and a second interval $(\lambda_-,1)$ such that for $\lambda \in (\lambda_-,1)$ we have $w_p(\lambda) < w_-$, with $\lambda_+ < \lambda_-$.*

Proof: Using equation (6.12) we obtain that $T''(w)/T'(w) = (\xi - 1)/w$. Then using equation (6.14) it is straightforward to show that $\frac{\partial^2 A_4(w,\lambda)}{\partial w^2} > 0$ if and only if $w < \bar{w}\frac{(1-\xi)(1-\lambda)}{\lambda(1+\xi)} \equiv w_p(\lambda)$, with $w_p'(\lambda) < 0$. $w_p(0) = +\infty$ and $w_p(1) = 0$. Figure 6.4 illustrates this lemma.□

This result gives the properties in terms of convexity and of concavity of the function $A_4(w)$. It implies that for a value of λ sufficiently small (inferior to λ_+) $A_4(w)$ is convex. Conversely, with α sufficiently high (superior to λ_-) $A_4(w)$ is concave.

Proposition 6.5 *If $A(w_-,\lambda_-) > 0$ and if $A(w_-,\lambda_+) < 0$ then for $\lambda \in (0,\lambda_+)$ net contributions behave qualitatively as in the pure Beveridgian case (figure 6.2). And for $\lambda \in (\lambda_-,1)$ net contributions behave qualitatively as in the pure Bismarckian case (figure 6.3).*

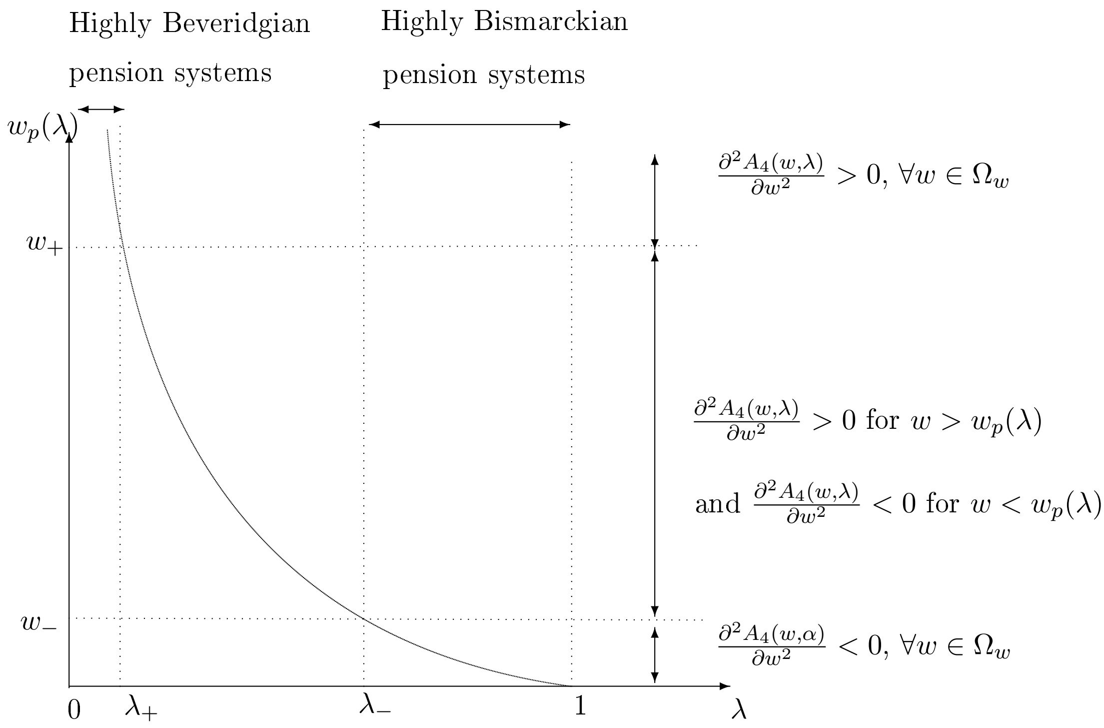

Figure 6.4: Illustration of lemma 3

Proof: The first two conditions ensure that the poorest have a positive (negative) net contribution in the interval $(\lambda_-, 1)$ $((0, \lambda_+))$.$\square$

This result extends the qualitative properties of Bismarckian and Beveridgian pension systems to intervals for the parameter λ. Each pension system with a λ in the interval $(\lambda_-, 1)$ is regressive. Conversely, each pension system with a λ inferior to λ_+ is redistributive, whereas the poorest do not necessarily benefit the most from this pension system.

For $\lambda \in (\lambda_+, \lambda_-)$, we have to calibrate our model to know the form of net contributions.

6.7 Calibration on French Data

In order to calibrate our model we first have to specify wage inequalities. We assume that wages belong to the interval: $\Omega_w = [0.2, 9]$. This interval implies that the wage w_+ is 45 times higher than w_-. Piketty (2002), studying the distribution of wages in France, finds a ratio of 5 between the wages of the first and of the last decile. The gap between this empirical fact and our calibration can be explained by the fact that we use the two extreme values of a *continuum* and as a consequence wage inequalities are greater. We choose this interval for Ω_w because once it is combined with the density function of w, our model matches the Gini coefficient of the wage distribution calculated by Hairault and Langot (2008) on French data.

The density function of the distribution of wages among the population has to respect the essential property: mode<median<mean (Lambert 2001, p.23). This property is a common feature of most industrialized countries. It implies that the wage distribution among the population is asymetric. The most common income level is less than the median wage. And because of strong wage inequalities the median wage is less than the average wage of the economy. Furthermore, the Gini index of wages has to tend towards 0.32 (Hairault and Langot 2008). Lambert (2001) shows that the Gini index can be calculated as:

$$G = -1 + 2\int_{w_-}^{w_+} \frac{wF(w)f(w)}{\bar{w}} dw$$

A useful density function is the density function of Weibull. It is asymmetric and it has the following form:

$$f(w) = \frac{c}{b}\left(\frac{w-a}{b}\right)^{c-1} e^{-\left(\frac{w-a}{b}\right)^c} \tag{6.15}$$

if $w > a$, and $f(w) = 0$ for $w \leq a$, with b and $c > 0$. The only problem with the use of this function is that the $\int_{\Omega_w} f(w)dw$ is not exactly 1. But the following calibration is such that this integral is approximately 1 over our interval Ω_w: $a = 0.2$, $b = 2$ and $c = 1.55$. Furthermore, it implies that $F(\bar{w}) > 0.5$, with $F()$ the cumulative distribution function, $\bar{w} = 2$, $w_{median} = 1.78$ and $G = 0.3245$.

We now have to specify the relationship between wages and the length of life. We assume that it has the following form:

$$T(w) = \varrho w^{\xi} \tag{6.16}$$

with ξ the elasticity of the length of life with respect to wages. ϱ is a scale parameter. We assume that $\varrho = 0.2$. In the following analysis we study the impact of a change in the parameter ξ, knowing that this value is clearly less than 1 (Bommier *et al.* 2006).

Finally, we assume that $\lambda = 0.885$ (Hairault and Langot 2008), but we study what happens if λ varies around this benchmark value.

For the moment, let us consider that $\xi = 0.09$, i.e. an increase in wages of 1% implies an increase in the length of life of 0.09%. Figure 6.5 illustrates this case.

For the value of λ calculated by Hairault and Langot (2008), then it is clear that agents with a wage below 1.6 and above 5.2 benefit from the pension system. The richest have a net contribution more negative than the net contribution of the poorest. Agents with a wage in the interval [1.6,5.2] have a positive net contribution and finance the negative net contributions of the two extremes. If λ increases, i.e. the pension system becomes more Bismarckian, then less poor agents benefit from a negative net contribution whereas the net contributions becomes negative for wages above 2.8. It is almost a regressive pension system. Finally, if $\lambda = 0.7$, then the redistributive properties of the pension system endure. The poorest benefit the most from the pension system and the richest have the highest positive net contribution.

Let us now assume that $\xi = 0.18$. The elasticity of the length of life is twice as high as before. Then the previous qualitative results endure (see figure 6.6), but the pension system

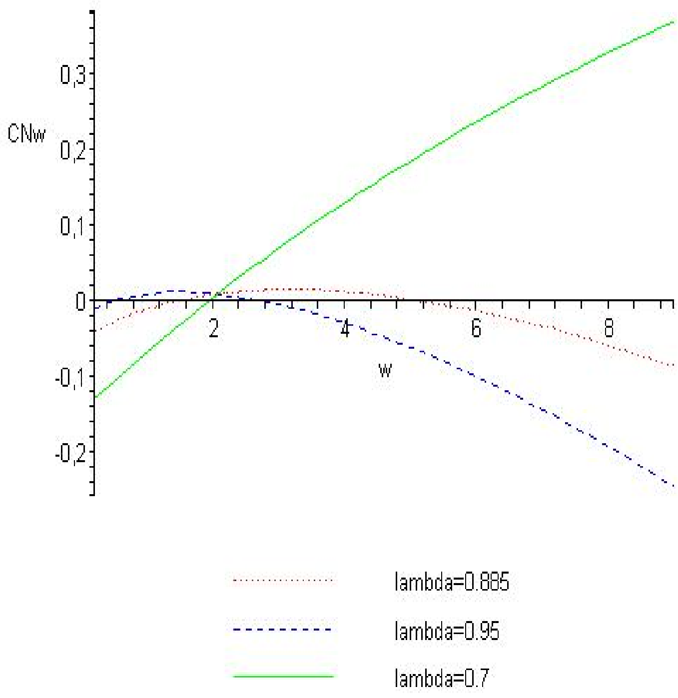

Figure 6.5: Net Contributions for different values of λ and for $\xi = 0.09$

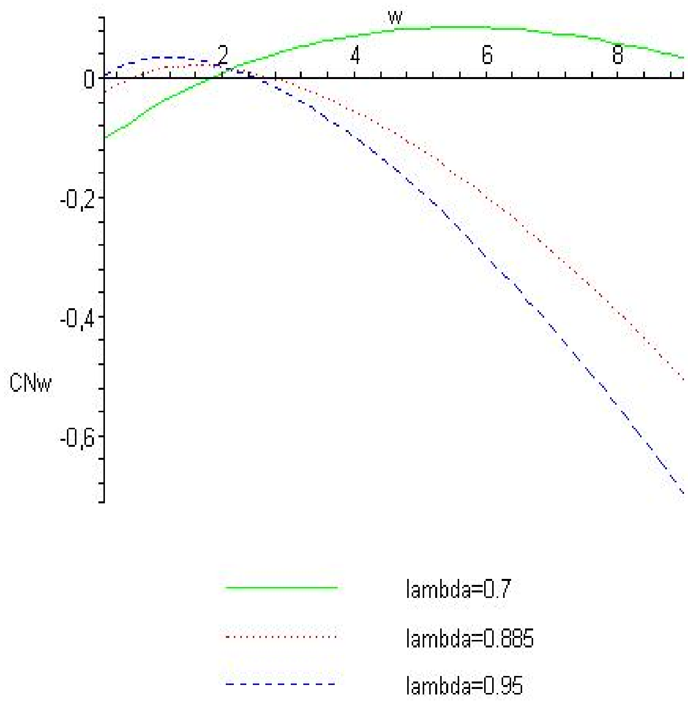

Figure 6.6: Net Contributions for different values of λ and for $\xi = 0.18$

is clearly regressive because only very small wages benefit from a negative net contribution. Agents with a wage above 2.4 have a negative net contribution and the richest benefit the most from the pension system. This result is also truer for $\lambda = 0.95$. Finally, even for $\lambda = 0.7$ the redistributive properties of the pension system are less clear.

6.8 Conclusion

Our contribution clarifies the debate on the redistributive properties of pension systems theoretically, when there are inequalities of length of life. We explain more precisely the sentence from Mitchell and Zeldes (1996) cited above. It is shown that Beveridgian pension systems are less progressive than they seem and that Bismarckian pension systems are regressive. Moreover, it makes it possible to show that for mixed pension systems, there can be a redistribution of resources from the middle to the ends of the distribution of wages. This last point would have to be taken into account for empirical analyses which study the progressivity of pension systems.

Our theoretical analysis is a first attempt to clarify the debate on the progressivity of pension systems once the life expectancy differential is taken into account. The next step on our research agenda would be to use a micro-simulation model, as in Liebman (2001), on French data. It would permit to quantify the impact of the life expectancy differential on the redistributive properties of pension systems.

6.9 APPENDIX

APPENDIX A

Indeed, the covariance can also be written as: $\int_{\Omega_w}(w-\bar{w})(T(w)-\bar{T})f(w)dw$. But as $\int_{\Omega_w}(w-\bar{w})f(w)dw = 0$, we can write that: $\int_{\Omega_w}(w-\bar{w})(T(w)-\bar{T})f(w)dw = \int_{\Omega_w}(w-\bar{w})(T(w)-X)f(w)da$, with X a constant, whatever the value of X. So it is particulary true for $X = T(\bar{w})$. Then we can write that: $\int_{\Omega_w}(w-\bar{w})(T(w)-\bar{T})f(w)dw = \int_{\Omega_w}(w-\bar{w})(T(w)-T(\bar{w}))f(w)dw$. The RHS is positive as it is an integral on a product of terms which have the same sign because $T'(w) > 0$.$\square$

Chapter 7

Education and the Progressivity of Pension Systems: When the Life Expectancy Differential Matters

7.1 Introduction

Since the introduction of unfunded public pension systems, their generosity has highly increased. The replacement rate of pension systems was 0.5 in 1975, and it has become 0.65 in 1995 in France. This trend has been observed in every developed country (Nyce and Schieber 2005). In the same time, the payroll tax rate used to finance this pension system has increased. For example, it was 8.5% in France in 1967, and it has become 19.8% in 1995[1]. This evolution describes a defined-benefit structure, i.e. a pension system in which the tax rate is the adjustment variable.

A growing literature studies the consequences of this evolution on macroeconomic variables. A first part of this literature emphasizes the negative net impact of a higher payroll tax rate on the physical capital accumulation (Feldstein 1974, Saint-Paul 1992). A second part of this literature studies the impact of the generosity of pension systems on educational choices (Lambrecht *et al.* 2005, Docquier and Paddison 2003, Casarico and Devillanova 2007) . In models with a representative agent, an increase in the size of pension systems can increase the time that agents dedicate to education. Indeed, pensions are a significant part of returns on educational investments. However in models with heterogenous agents, two dimensions of education have to be taken into account: the share of the population who decides to educate itself; and the time dedicated to education for those who choose to educate themselves. Docquier and Paddison (2003) show that an increase in the payroll tax rate has a negative impact on the share of the educated population and thus on the growth rate of the economy. Le Garrec (2005) obtains the same qualitative result but he shows that for agents who decide to educate themselves, the impact on the time dedicated to education is widely positive, what overcompensates the first negative impact. The common element of the last two studies is the decrease in the share of the educated population. Le Garrec (2005) explains it rightly by the fact that pension systems are redistributive[2]. Indeed, pension systems are at least partially Beveridgian. Consequently, an increase in the size of pension systems has a negative impact on the share of the educated population.

[1]Some authors emphasized this point:

> *"In the developed nations today, these programs are largely financed through payroll taxes and pay a defined benefit." (Nyce and Schieber 2005)*

[2]In the paper of Docquier and Paddison (2003), there is another channel. Indeed, the increase in the size of pension systems has a positive impact on the interest rate what decreases the discounted value of returns on education.

However, as the previous chapter shows, a pension system can be regressive if we consider inequalities of length of life. The previous conclusions are not robust to this change and the objective of this chapter is to study the implications of an increase in the size of pension systems once inequalities of length of life are taken into account.

A pension system is purely Beveridgian if every agent receives the same pension. Conversely, a pension system is purely Bismarckian if pensions completely depend on the wages of agents. A pension system is mixed if it has a Beveridgian and a Bismarckian component. The more a pension system is Beveridgian, the higher intra-generational transfers are. Countries highly differ by this intra-generational component. France, Germany and Italy have a Bismarckian structure. Canada, the Netherlands and New-Zeland are essentially Beveridgian. Finally, Japan, the United-Kingdom and the United States have mixed pension systems (Sommacal 2006, Casamatta *et al.* 2000).

Some empirical studies, as Coronado *et al.* (2000) and Liebman (2001), show that the pension system of the United-States is far less progressive than is usually mentioned once the life expectancy differential is taken into account. Bommier *et al.* (2003) find the same qualitative results on French data. Borck (2007) shows that pure Beveridgian and Bismarckian pension systems are respectively progressive and regressive. However, if the pension system is a mix of the two, then the poorest and the richest receive more from this system than they pay for it[3], whereas agents with intermediate wages have a negative net benefit. In that case, it is shown that an increase in the size of pension systems can increase wealth inequalities. We argue in this chapter that it can change educational choices. Gorski *et al.* (2007) study the implications on the share of the educated population of a change in the *structure* of pension systems when educated agents live longer than uneducated ones. They show that if the pension system becomes more Bismarckian then less agents decide to educate themselves[4]. We use the same kind of models to study the impact of a change in the *size* of pension systems on the share of the educated population.

The link between educational choices and the expected life expectancy is not widely documented in experimental analysis. Mirowski and Ross (2000) find that each additional year of education increases the subjective life expectancy of 0.7 years. They conclude:

"*Educational attainment clearly influences subjective life expectancy. The per-*

[3]Borck (2007) shows that the size of a pension system can be determined by a coalition of elderly, very poor and very rich agents. Poor agents benefit from the Beveridgian part of the pension system, whereas rich agents benefit the longest time from the pension system.

[4]See the general introduction for a more detailed presentation and for the critics about their model.

> *ceived payoffs to education apparently extend beyond employment and economic resources to the anticipation of a secure, healthy, and long future. (Conversely, people with little education implicitly anticipate a risky future.)*" (Mirowski and Ross 2000, pp.145)

Two remarks can be made concerning this study. Firstly, educational choices do not only concern the financial benefits but also the benefits in terms of health. Secondly, the model we use in this chapter is more restrictive because we assume that every agent rightly expect the distribution of life expectancy among the population. Consequently, if they decide to remain uneducated, they know that they renounce to higher wages and to a higher length of life. This assumption of a common knowledge concerning differences in life expectancies is also used in Borck (2007) and in Gorski *et al.* (2007).

In this chapter we study the impact of an increase in the generosity of pension systems on the share of the educated population should inequalities of length of life exist. In this chapter, the generosity of pension systems is defined as the average replacement rate of the economy. We use an overlapping generations model, with a small open economy, where agents live for three periods. They can be young, adult or old. Every agent lives his whole two first periods of life. Moreover, educated agents stay alive during a fraction φ of their third period of life, whereas uneducated agents live only during a fraction σ of this period, with $\sigma < \varphi$. It implies that there are inequalities of length life according to the socioeconomic status of agents. Each agent bears a psychological cost if he decides to educate himself during his first period of life. He does so if his utility once educated is higher than his utility if he remains uneducated. The cost of education represents the time which cannot be spent on leisure. Every agent works and offers inelastically one unit of labor during his second period of life. Furthermore, there is an unfunded public pension system which has a mixed structure (Beveridgian and Bismarckian). The Government levies a tax on wages of workers to finance this unfunded pension system. What matters for our purpose is the instantaneous redistributivity of this pension system which has a positive influence on its progressivity. Conversely, life expectancy inequalities tend to increase the regressivity of pension systems. We find that the impact of the size of pension systems on the share of the educated population is ambiguous. The main intuition is that the instantaneous redistributivty of pension systems has a negative impact on the share of the educated population, whereas the possible long run regressivity, induced by life expectancy inequalities, has the opposite effect. Consequently, if a pension system is progressive, then it increases the opportunity cost of education and a larger share of agents prefer remaining uneducated. But, if a pension system is regressive, then the Government redistributes resources in favor of educated agents and a larger share of the population decides to educate itself. Then, using a calibrated version of the model we define an area for the inequalities of length of life and for the Bismarckian part of pension systems such that an increase in the generosity of pension systems has a positive impact on the share of the educated population. The dynamics of the educated part of the population is also interesting. Indeed, if we assume an expected change in the generosity of pension systems, there appears a strong increase in the share of the educated population, before decreasing until the new higher steady state. Finally, we find that an increase in the size of pension systems only has a transitory positive impact on the welfare of agents.

This chapter is organized as follows. In section 2 the main elements of our model are described. In section 3, we study the macroeconomic equilibrium and its properties. In section 4, we propose a numerical resolution of our model. Finally, section 5 provides some concluding remarks.

7.2 The Model

We assume a small open economy in which agents live for three periods and differ by their ability to educate themselves. In the first period of their life, agents decide to educate themselves or to remain uneducated. In the second one, agents work. And finally, in the third period of their life, agents go into retirement. Furthermore, we assume a constant growth rate of the population such that: $N_t = (1+n)N_{t-1}$.

When an agent is born, he is randomly endowed with an educational cost denoted by θ, which the agent bears if he decides to educate himself. This educational cost is related to the time necessary for education, i.e. to the time which cannot be spent on leisure. Consequently, in our model, agents differ by the time which they should invest if they wanted to be educated. The density function and the cumulative distribution function of θ are respectively denoted by $f(\theta)$ and $F(\theta)$. Because of the law of large numbers, $f(\theta)$ also represents the share of the population endowed with a cost θ. θ takes its values in the interval $\Omega_\theta = [\underline{\theta}, \bar{\theta}]$, with $\underline{\theta} > 0$ and $\underline{\theta} \to 0$. Thus, we have $F(\underline{\theta}) = 0$ and $F(\bar{\theta}) = 1$. A small θ implies that the agent has a large learning ability. Conversely, a high θ implies that the agent has a large learning cost. Given their cost endowment, agents decide to educate themselves or not during their first period of life. We assume that it is the only economic decision which can be taken by young agents. The educational decision is a binary choice. If an agent decides not to educate himself, then he spends the time of his first period of life on leisure[5]. The utility level of leisure is normalized to 0. Conversely, if an agent decides to educate himself, then he has to bear a psychological cost, denoted by $(-v(\theta))$, related to the learning process, and thus to the time which cannot be spent on leisure. In order to simplify the formal analysis, we assume that this cost is the only opportunity cost of education. In this way, we do not consider that uneducated agents can begin to work earlier[6].

We assume that $v(\theta)$ has the following properties:

Assumption 1: $v'(\theta) > 0 \ \forall \theta \in \Omega_{\theta}$, $v(\underline{\theta}) \leq 0$ and $v(\bar{\theta}) > 0$.

The first part of this assumption implies that a longer time spent on the learning process reduces leisure time, and thus the utility level. As for the second part of this assumption, it implies that for a fraction of the population there could exist a psychological benefit of being educated. This benefit is all the more large as θ tends toward $\underline{\theta}$. For these people, education is a way to blossom out. $v(\underline{\theta}) \leq 0$ also ensures that there necessarily exists an educated population at equilibrium. The last part of this assumption means that there exists a fraction of the population for which the time spent on education is a cost. Otherwise every agent would choose to educate himself.

During his second period of life, an agent supplies his work inelastically. He retires at the very beginning of his third period of life. Educated agents can expect to live a fraction φ of their third period of life, whereas uneducated agents live only during a fraction σ of theirs, with $1 > \varphi \geq \sigma > 0$. Let ϵ denote the differential between φ and σ, i.e. $\varphi > \epsilon = \varphi - \sigma \geq 0$. If $\epsilon = 0$, then educated and uneducated agents have the same length of life. Figure 7.1 sums up this point.

[5]The length of each period is normalized to 1.

[6]However, we suspect that it would reinforce our results concerning the regressivity of pension systems and that the dark region of figure 7.2 would be larger.

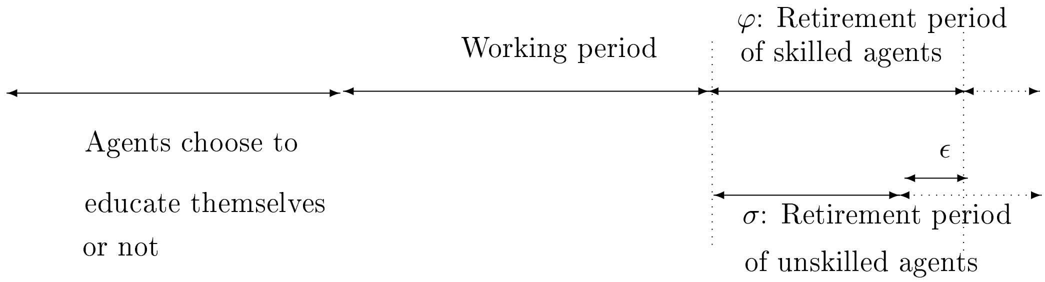

Figure 7.1: Life of agents

7.2.1 Consumers

Preferences of a representative agent born at period $t-1$ are the following[7][8] (d'Autume 2003)[9]:

$$U_{t-1}^i(c) = \ln(c_t^i) + \beta T^i \ln\left(\frac{d_{t+1}^i}{T^i}\right) - (1-I)v(\theta) \tag{7.1}$$

with $I = 1$ if the agent decides to remain uneducated ($i = ne$), and $I = 0$ if he educates himself ($i = e$). c_t and d_{t+1} denote the consumption levels of the second and the third periods of life respectively. Furthermore, we have:

$$T^i = \begin{cases} \varphi & \text{if i=e} \\ \sigma & \text{if i=ne} \end{cases}$$

β is the pure psychological discount factor. The first two parts are standard and imply that an increase in the length of life (T^i) increases the weight of the third period utility but reduces the third period consumption flow. The third term in the utility function denotes the psychological cost of education.

Let us define R_{t+1} as the interest factor. Moreover, W_t^e and W_t^{ne} will denote the wealth of an educated and of an uneducated agent born at period $t-1$ respectively, assessed at

[7]Neither Borck (2007), nor Gorski *et al.* (2007) introduce the inequalities of length of life in the utility function whereas agents in their models take into account the net impact of inequalities of length of life on the budget constraint.

[8]We assume logarithmic preferences in order to emphasize the results induced by the introduction of inequalities of length of life on educational choices. This is the simplest way we found to exhibit the properties of this new channel. It is clear that results of the following section, notably for $\epsilon = 0$, depend on this assumption. Taking a more general utility function would have complicated the analysis and it would have added other effects that could have masked ours.

[9]See the appendix of chapter 1 for a justification.

period t. Consequently, the budget constraint in period t of an agent born in $t-1$ can be written:

$$c_t^i + \frac{d_{t+1}^i}{R_{t+1}} = W_t^i \tag{7.2}$$

The wealth level of agents depends on the wages obtained on the labor market, and on the social transfers. The wealth level can be written:

$$W_t^i = w_t^i(1-\tau_t) + \frac{p_{t+1}^i}{R_{t+1}} \tag{7.3}$$

w_t^i denotes the wage of an agent of type $i \in \{e, ne\}$ obtained on the labor market. τ_t denotes the tax rate used to finance the pension system. Finally, p_{t+1}^i denotes the pension of an agent of type i at period $t+1$.

The sequence of decisions is: (1) the agent decides to educate himself or not; (2) he saves a share of his second period of life net income for his third period consumption. Using a backward resolution, we first have to determine the indirect utility function of an agent for each scenario (educated or not). Then the agent compares the utility level obtained in each case and he chooses the best option.

Maximizing (7.1) subject to the budget constraint (7.2) we obtain:

$$c_t^i = \frac{W_t^i}{1+\beta T^i} \tag{7.4}$$

$$d_{t+1}^i = \beta T^i R_{t+1} \frac{W_t^i}{1+\beta T^i} \tag{7.5}$$

Finally, denoting by U_t^e (U_t^{ne}) the utility of an educated (uneducated) agent, we obtain with equations (7.1), (7.4) and (7.5):

$$U_t^{ne} = \ln\left(\frac{W_t^{ne}}{1+\beta\sigma}\right) + \beta\sigma \ln\left(\beta R_{t+1} \frac{W_t^{ne}}{1+\beta\sigma}\right) \tag{7.6}$$

and,

$$U_t^e = \ln\left(\frac{W_t^e}{1+\beta\varphi}\right) + \beta\varphi \ln\left(\beta R_{t+1} \frac{W_t^e}{1+\beta\varphi}\right) - v(\theta) \tag{7.7}$$

An agent decides to educate himself iff $U_t^e > U_t^{ne}$, i.e. if:

$$(1+\beta\varphi)\ln\left(\frac{W_t^e}{W_t^{ne}}\right) + \epsilon\beta \ln(W_t^{ne}) + \chi_{t+1} > v(\theta) \tag{7.8}$$

with $\chi_{t+1} = \epsilon\beta \ln(\beta R_{t+1}) + (1+\beta\sigma)\ln(1+\beta\sigma) - (1+\beta\varphi)\ln(1+\beta\varphi)$. This equation is used by consumers to decide if they educate themselves or not. We see that they prefer being educated if θ is not too large, i.e. if their leisure cost is low. But it is very intuitive that this equation determines implicitly a threshold value $\hat{\theta}_t$ such that each agent for whom $\theta < \hat{\theta}_t$ ($\theta > \hat{\theta}_t$) chooses to educate himself (remain uneducated). This threshold value will be determined later in this chapter. We assume that agents have perfect foresight and know the equilibrium value $\hat{\theta}_t$ when they make their decisions.

The first term of the inequality (7.8) is the wealth differential if the agent decides to educate himself. The wealth level increases the consumption level of both periods. The second term is the loss of consumption because of the shorter length of life if the agent decides to remain uneducated. The higher ϵ is, the larger this loss is. The RHS (Right-Hand-Side) is the leisure cost of education.

7.2.2 Firms

A continuum of mass one of identical firms produces the final good using the constant returns to scale Cobb-Douglas technology:

$$Y_t = AK_t^{\alpha} L_t^{1-\alpha} \tag{7.9}$$

with $1 > \alpha > 0$, $A > 0$ the level of the technology, K_t the physical capital and L_t the labor. We normalize to 1 the productivity of uneducated agents and to q, with $q > 1$, the productivity of educated agents. The labor demand at each period t is:

$$L_t = qL_t^e + L_t^{ne} \tag{7.10}$$

with L_t^e (L_t^{ne}) the demand for educated (uneducated) labor.

We assume a perfect competitive economy on the final good market and on the inputs markets. Consequently, interest factor and wages are fixed at the marginal productivity of capital and of labor respectively. Finally, as mentioned above, we assume a small open economy in which the interest factor R_t is fixed exogenously and is constant over time ($\bar{R}$). The first order condition for capital implies:

$$K_t = \left(\frac{A\alpha}{\bar{R}}\right)^{1/1-\alpha} L_t$$

Then, using this equation with the first order condition of L_t implies:

$$w_t = A(1-\alpha)\left(\frac{A\alpha}{\bar{R}}\right)^{\alpha/1-\alpha}$$

w_t is constant over time ($w_t = w$, $\forall t$) and it denotes the wage per efficiency unit of labor. Calling w^e (w^{ne}) the wage of educated (uneducated) agents, we finally obtain:

$$w^e = qw \tag{7.11}$$

and

$$w^{ne} = w \tag{7.12}$$

Note that w is completely independent of $\hat{\theta}_t$ because of the perfect substituability of the different kinds of labor. In the rest of this chapter $\bar{w}_t$ will denote the average wage of the economy at time t. It is obtained by:

$$\bar{w}_t = \int_{\underline{\theta}}^{\hat{\theta}_t} w^e f(\theta)d\theta + \int_{\hat{\theta}_t}^{\bar{\theta}} f(\theta)w^{ne}d\theta = w\left[qF(\hat{\theta}_t) + (1 - F(\hat{\theta}_t))\right] \tag{7.13}$$

with $F(\hat{\theta}_t)$ the share of the educated population. A higher $\hat{\theta}_t$ has a positive impact on $\bar{w}_t$ because of the increase in the share of the educated population which has the higher productivity level.

7.2.3 Government

As mentioned above, we assume that agents pay a tax τ_t on their wage, and receive a pension during their last period of life as long as they are alive. As in Casamatta *et al.* (2000), agents receive a pension indexed on their wage and on the average wage of the economy. The pension per unit of time of an agent born at period $t-1$ is: $\nu\left(\lambda w_t^i + (1-\lambda)\bar{w}_t\right)$, with $i \in \{e, ne\}$. ν is the average replacement rate of the economy. λ is a measure of the instantaneous redistributivity of the pension system[10]. If $\lambda = 0$, then the pension system is Beveridgian because every agent receives the same amount from the Government. If $\lambda = 1$, each agent receives a pension which is proportional to his wage. This pension system is called Bismarckian. Finally, the agents receive this pension only during a period T^i, i.e. as long as they are alive. The Government budget constraint can be written:

$$\tau_t\left[w^e F(\hat{\theta}_t) + w^{ne}(1 - F(\hat{\theta}_t))\right] N_t =$$

$$\nu\left[\varphi(\lambda w^e + (1-\lambda)\bar{w}_{t-1})F(\hat{\theta}_{t-1}) + \sigma(\lambda w^{ne} + (1-\lambda)\bar{w}_{t-1})(1 - F(\hat{\theta}_{t-1}))\right] N_{t-1} \tag{7.14}$$

[10]See the previous chapter for a more specific discussion on the distinction between the instantaneous and the long run redistributitvity of pension systems.

Furthermore, we assume a defined-benefit structure of the pension system. It means that it is τ which adjusts itself to economic and demographic changes. This assumption is realistic given the evolutions of the tax rate and of the replacement rate since the end of the second world war [11].

7.3 The Macroeconomic Equilibrium

At each period t, the labor supply is equal to the labor demand.

$$L_t^e = N_t F(\hat{\theta}_t) \tag{7.15}$$

and

$$L_t^{ne} = N_t(1 - F(\hat{\theta}_t)) \tag{7.16}$$

The dynamics of this economy is completely characterized by the dynamics of the threshold cost, i.e. by the sequence $\{\hat{\theta}_t\}_{t=0..\infty}$. A higher $\hat{\theta}_t$ implies that a larger share of the population is educated ($F(\hat{\theta}_t)$ increases). $\hat{\theta}_t$ is determined by the following equation:

$$(1 + \beta\varphi)\ln\left(\frac{W_t^e}{W_t^{ne}}\right) + \epsilon\beta\ln(W_t^{ne}) + \chi_{t+1} = v(\hat{\theta}_t) \tag{7.17}$$

Given that wages verify the first order conditions of firms (7.11) and (7.12), that the Government budget constraint (7.14) has to be satisfied, and that at each period the labor markets of educated and uneducated agents (7.15) and (7.16) have to be at equilibrium.

In order to correctly understand the implications of the inequalities of length of life we first study the case with $\epsilon = 0$. Then we detail the results obtained for $\epsilon > 0$. But before making such an analysis, we define the steady state in this economy.

Definition: The economy is at steady state if $\hat{\theta}_t = \hat{\theta}$.

Because of the complexity of the study of the dynamics of $\hat{\theta}_t$, we only consider steady states for the moment. We show in the numerical exercice that there exists some specifications for functions such that the dynamics of $\hat{\theta}_t$ is unique and stable.

[11] Feldstein and Liebman (2002) and Nyce and Schieber (2005) provide an empirical evidence concerning this assumption.

7.3.1 The case with $\epsilon = 0$ ($\varphi = \sigma$)

In that case we consider that there is no life expectancy differential. At steady state, the budget constraint of the Government can be written:

$$\tau = \nu \frac{\varphi}{1+n} \equiv \tau(\nu) \tag{7.18}$$

It is usually the case, the tax rate is the product between the generosity of the pension system (ν) and the old-age dependency ratio[12] ($\varphi/(1+n)$). To find the critical value $\hat{\theta}$, there only remains to use equation (7.17) but with $\epsilon = 0$. It is straightforward to show that it becomes:

$$(1+\beta\varphi)\ln\left(\frac{W^e(\hat{\theta})}{W^{ne}(\hat{\theta})}\right) = v(\hat{\theta}) \tag{7.19}$$

with $W^e(\hat{\theta}) = qw(1-\tau(\nu)) + \frac{\nu}{R}(\lambda qw + (1-\lambda)\bar{w}(\hat{\theta}))\varphi$, $W^{ne}(\hat{\theta}) = w(1-\tau(\nu)) + \frac{\nu}{R}(\lambda w + (1-\lambda)\bar{w}(\hat{\theta}))\varphi$, and $\bar{w} = w\left[qF(\hat{\theta}) + (1-F(\hat{\theta}))\right]$. It can easily be shown that the LHS (Left-Hand-Side) is a decreasing function of $\hat{\theta}$. Moreover, the RHS is an increasing function of $\hat{\theta}$. Consequently, there exists at most one steady state value $\hat{\theta}$ which satisfies equation (7.19).

Lemma 7.1 *There exists at most one steady state threshold cost $\hat{\theta}$. It is unique and it belongs to $(\underline{\theta}, \bar{\theta})$ iff $LHS(\bar{\theta}) < v(\bar{\theta})$.*

Proof: It is obtained using the properties mentioned above and assumption 1.□.

This lemma shows that, at steady state, the share of the educated and of the uneducated population are not null ($F(\hat{\theta}) > 0$ and $(1 - F(\hat{\theta})) > 0$). More specifically, a fraction of the population decides to remain uneducated if the maximum leisure cost is high.

We also assume that if $\hat{\theta}$ is an interior solution, then it is stable. Let us now study the properties of $\hat{\theta}$ when the replacement rate changes. We say that an increase in ν corresponds to an increase in the size of the pension system.

Proposition 7.1 *An increase in ν leads to a decrease in $\hat{\theta}$ for $\lambda \in [0,1)$, i.e. an increase in the generosity of the pension system leads to a decrease in the share of the educated population. But if the pension system is purely Bismarckian ($\lambda = 1$), then an increase in ν has no impact on $\hat{\theta}$.*

[12] See d'Autume (2003).

Proof: It is obtained with equation (7.19) and differentiating this equation with respect to $\hat{\theta}$ and ν.$\square$

Because there are no inequalities of length of life, the pension system is redistributive [13]. Consequently, an increase in the generosity of the pension system implies that agents prefer remaining uneducated. For the special case of a purely Bismarckian pension system, wealth is proportional to wages. It implies that ν has only a multiplicative effect on wealth, and this factor is the same for educated and uneducated agents. In that special case, an increase in ν has no impact on the size of the educated population. ν is said to be neutral for educational choices.

The increase in ν has an ambiguous impact on the welfare of agents of our economy. Thanks to equations (7.6) and (7.7) we see that the utility level depends on ν only through wealth. That is why when we study the impact of ν on welfare we only need to consider its impact on the wealth of agents.

$$\frac{dW^i(\hat{\theta},\nu)}{d\nu} = \underbrace{-w^i\frac{d\tau(\nu)}{d\nu}}_{1} + \underbrace{\frac{T^i}{\bar{R}}(\lambda w^i + (1-\lambda)\bar{w}(\hat{\theta}))}_{2} + \underbrace{T^i\frac{\nu}{\bar{R}}(1-\lambda)\frac{d\bar{w}(\hat{\theta})}{d\hat{\theta}}\frac{d\hat{\theta}}{d\nu}}_{3} \tag{7.20}$$

An increase in ν has three effects on wealth. Firstly, it increases the tax rate used to finance the pension system (element 1<0). Secondly, it increases the generosity of the pension system (element 2>0). And finally, it decreases the average wage because of the decrease in $\hat{\theta}$ (element 3<0). Whereas the first and the last effects are negative, the second one is positive. Consequently the net impact on U_t^i is ambiguous, for $i \in \{e, ne\}$.

7.3.2 The case with $\epsilon \in (0,\varphi)$ $(\varphi > \sigma)$

If there is a life expectancy differential between educated and uneducated agents, then, at steady state, the budget constraint of the Government can be written:

$$\tau(1+n)\bar{w} = \nu\left[\varphi\bar{w} - \epsilon(\lambda w^{ne} + (1-\lambda)\bar{w})(1-F(\hat{\theta}))\right]$$

The first part of the RHS is standard and corresponds to the previous case when $\epsilon = 0$. But the second term of the RHS denotes the pensions which are not paid to uneducated

[13] See the previous chapter and Borck (2007).

workers because their life expectancy is shorter than that of educated agents. This amount is all the more significant as ϵ is high and as $\hat{\theta}$ is small.

After some straightforward calculations we obtain:

$$\tau = \frac{\nu}{1+n}\left[\varphi - \epsilon\left(\frac{\lambda}{\bar{w}(\hat{\theta})} + 1 - \lambda\right)(1 - F(\hat{\theta}))\right] \equiv \tau(\nu, \hat{\theta}) \tag{7.21}$$

Consequently, τ is an increasing function of ν and of $\hat{\theta}$. Whereas the first relation is obvious, the second one is less so. Indeed, an increase in $\hat{\theta}$, first of all, reduces the share of uneducated agents and therefore the amount that is not paid during a period ϵ. Secondly, an increase in $\hat{\theta}$ has a positive impact on $\bar{w}$, which reduces the ratio between the pension which is not paid and the average wage of the economy.

Finally, *ceteris paribus*, τ is an increasing function of λ which represents the Bismarckian part of pension systems. Indeed, if pensions are more indexed on wages, the pensions which are paid to agents having the shorter length of life are smaller. Conversely, educated agents, who have the longer length of life, benefit from a higher pension per unit of time. Consequently, for a given $\hat{\theta}$, an increase in λ has a positive impact on the tax rate τ.

We assume that there exists a unique $\hat{\theta}$ and that it is stable at least locally. Moreover, given our assumption of a small open economy, χ_t is constant over time. Then equation (7.17) which defines $\hat{\theta}$ becomes:

$$g(\hat{\theta}, \nu) \equiv (1 + \beta\varphi)\ln\left(\frac{W^e(\hat{\theta}, \nu)}{W^{ne}(\hat{\theta}, \nu)}\right) + \epsilon\beta\ln(W^{ne}(\hat{\theta}, \nu)) + \chi = v(\hat{\theta}) \tag{7.22}$$

with $W^e(\hat{\theta}, \nu) = qw(1-\tau(\nu, \hat{\theta})) + \frac{\nu}{R}\varphi(\lambda qw + (1-\lambda)\bar{w}(\hat{\theta}))$, $W^{ne}(\hat{\theta}, \nu) = w(1-\tau(\nu, \hat{\theta})) + \frac{\nu}{R}\sigma(\lambda w + (1-\lambda)\bar{w}(\hat{\theta}))$, and $\bar{w}(\hat{\theta}) = w\left[qF(\hat{\theta}) + (1 - F(\hat{\theta}))\right]$. We cannot analytically prove that the value of $\hat{\theta}$ is unique, but let us assume that it is for the moment. We define the LHS of equation (7.22) as $g(\hat{\theta}, \nu)$. Then, the following result can be established:

Lemma 7.2 *If $g_2(\hat{\theta}, \nu) > 0$ and if $\hat{\theta}$ is unique and belongs to $(\underline{\theta}, \bar{\theta})$, then an increase in ν has a positive impact on $\hat{\theta}$.*

Proof: First, note that $g_i(.,.)$ denotes the derivative of $g(.,.)$ with respect to its *ith* argument. Then, because the RHS of equation (7.22) is a strictly increasing function of $\hat{\theta}$ and because it does not depend on ν, we only have to study the impact of ν on $g(.,.)$. A simple graphical illustration of the condition of the lemma proves the result. Moreover, using assumption 1 and the fact that $w^e > w^u$, we prove that $g(\underline{\theta}, \nu) > v(\underline{\theta})$.□

This result is important and describes a new channel through which the generosity of pension systems influences educational choices.

Indeed, if there are inequalities of length of life, pension systems can be regressive despite the indexation of pensions on the average wage of the economy. It means that pension systems can increase the wealth of rich agents more than that of poor agents [14].

Secondly (second effect), the second term of equation represents the impact on the opportunity cost in terms of consumption (7.22). Indeed, if an agent decides to remain uneducated he looses consumption because of his shorter length of life. This opportunity cost increases iff:

$$\frac{\partial W^{ne}(\hat{\theta}, \nu)}{\partial \nu} = -w \frac{\partial \tau(\nu, \hat{\theta})}{\partial \nu} + \frac{\sigma}{\bar{R}}(\lambda w + (1-\lambda)\bar{w}(\hat{\theta})) > 0 \tag{7.23}$$

This condition can be satisfied if $\bar{w}$ is sufficiently high and if $\bar{R}$ is sufficiently small. Consequently, these two effects are such that increasing the generosity of pension systems can imply that more agents decide to educate themselves.

Because the condition of lemma 2 is not always satisfied, the general impact of ν on $\hat{\theta}$ cannot be determined *a priori*.

Proposition 7.2 *The net impact of the generosity of the pension system (ν) on the share of the educated population ($F(\hat{\theta})$) cannot be determined a priori because the system can be regressive.*

As we cannot conclude as for the net impact of ν on $\hat{\theta}$, we numerically solve our model and we study its properties in the next section.

Let us now consider the impact of an increase in ν on the welfare level of agents. As mentioned above, ν influences the welfare of agents only through the wealth level. The wealth of agents increases iff:

$$\frac{dW^i(\hat{\theta}, \nu)}{d\nu} = \underbrace{-w^i \left(\frac{\partial \tau(\nu, \hat{\theta})}{\partial \nu} + \frac{\partial \tau(\nu, \hat{\theta})}{\partial \hat{\theta}} \frac{d\hat{\theta}}{d\nu} \right)}_{B}$$

[14] See the previous chapter or Borck (2007).

$$+\underbrace{\frac{T^i}{\bar{R}}(\lambda w^i + (1-\lambda)\bar{w}(\hat{\theta}))}_{C} + \underbrace{T^i\frac{\nu}{\bar{R}}(1-\lambda)\frac{\partial \bar{w}(\hat{\theta})}{\partial \hat{\theta}}\frac{d\hat{\theta}}{d\nu}}_{D} \quad (7.24)$$

Under the conditions of lemma 5.2, we have: $B < 0$, $C > 0$ and $D > 0$. Element B denotes the impact of an increase in the generosity of pension systems on net wages. Element C denotes the direct impact of ν on pensions. Finally, element D denotes the impact on the pensions received by agents through the average wage of the economy. Note that in this new framework there are two positive effects and only one negative. Conversely, in the case in which $\epsilon = 0$, we had two negative effects and only one positive.

7.4 The Numerical Resolution of the Model and its Properties

Two dimensions have to be analyzed. The first one is the behavior of $\hat{\theta}$ at steady state, i.e. its comparative statics. The second one concerns the dynamic behavior of the economy. This last point also enables us to show that the dynamics can be stable. Right now we assume that $v(\theta) = \ln(\theta/\mu)$[15].

7.4.1 The steady state

The main objective of this subsection is to exhibit the behavior of $\hat{\theta}$ when ν increases. We have emphasized above that it essentially depends on two parameters: inequalities of length of life (ϵ) and the Bismarckian part of pension systems (λ). Consequently, we calibrate every parameter of the model, including φ, except ϵ (and thus σ) and λ. Then we determine which couples of parameters (ϵ,λ) imply that ν has a positive impact on $\hat{\theta}$.

Let us first calibrate our model. We assume that the length of each period is 40 years. Our model implies that educated agents live for $(1+\varphi) \times 40$ years, whereas uneducated agents live only during $(1+\sigma) \times 40$ years.

The distribution of educational costs ($f(\theta)$) is assumed to be uniform over the interval $\Omega_\theta = [0.000001, 10]$. Furthermore, we assume that $\bar{R} = 4.8$ which corresponds to an annual

[15] We did not make such an assumption right from the beginning of the paper, because it does not change our main result which is that ν does not have a clear impact on $\hat{\theta}$. More specifically, we cannot provide a clear figure which determines the value of parameters such that ν has a positive impact on $\hat{\theta}$.

Parameter	Meaning	Value	Source(s)
$\bar{R}$	Interest Factor	4.8	Standard, AIR[a]=4%
α	$R_t K_t / Y_t$[b]	0.33	Sommacal (2006) among others
A	The technology level	$\left(\frac{\bar{R}}{\alpha}\right)^{\alpha}\left(\frac{1}{1-\alpha}\right)^{1-\alpha}$	Such that $w = 1$
q	Productivity of educated agents	1.7	Acemoglu (2002)
μ	Parameter in the educational cost	2	Such that $\hat{\theta} \simeq 4$ (Acemoglu 2002)
φ	$(1+\varphi)$=LEEA[c]	0.55	LEEA is 62 years
β	Psychological discount factor	0.6	d'Autume (2003)
n	Population's growth rate	0.3	AGR=0.65%[d], Charpin (1999)

[a]AIR=Annual Interest Rate
[b]The share of income spent on capital.
[c]LEEA=Life Expectancy of Educated Agents
[d]AGR=Annual Growth Rate.

Table 7.1: Basic Calibration of the model

interest rate of 4%. Following Charpin (1999), $n = 0.3$, i.e. the annual population growth rate is 0.65%. We assume that $\beta = 0.6$ (d'Autume 2003), which corresponds to an annual psychologic discount factor of approximately 0.983. Concerning the production function, we assume that $\alpha = 0.33$ and that A is such that $w = 1$.

q is equal to 1.7, which implies that the wages of educated agents are 1.7 times higher than the wages of uneducated agents. In this way we calibrate q in order to match Acemoglu's findings (2002). In Acemoglu (2002) the corresponding share of the educated population is 0.4. We calibrate μ for $\hat{\theta} \simeq 4$. We obtain $\mu = 2$. We normalize φ to 0.55, which corresponds to a life expectancy of 62 years[16].

Table 7.1 sums up our calibration.

Let us now consider the net impact of an increase in the replacement rate (ν) from 0.5 to 0.6. We use this evolution as Nyce and Schieber (2005) report that the replacement rate has increased from 0.5 in 1975 to 0.6 in 1995 in most European countries (France, Germany, Belgium). The parameters λ and ϵ play a significant role to determine the sign

[16]Remember that in our model we do not take into account the period when an agent is young. Assuming that the length of this period is 20 years we obtain a life expectancy for educated agents of 82 years.

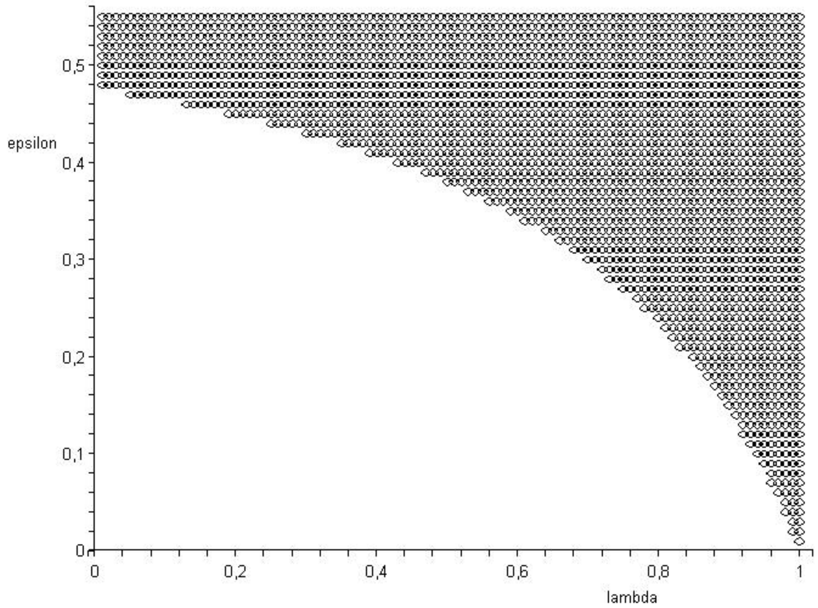

Figure 7.2: Couples (ϵ, λ) which imply that $d\hat{\theta}/d\nu > 0$.

of $d\hat{\theta}/d\nu$. The dark part of figure 7.2 represents the couples of λ and ϵ which are such that ν has a positive impact on $\hat{\theta}$[17].

Given our remarks of the previous section, the results are intuitive . Indeed, as in the previous chapter and as in Borck (2007), a pension system is regressive if inequalities of length of life are sufficiently high and if the pension system is highly Bismarckian. In our case, if inequalities of length of life are sufficiently small, then the Bismarckian part of the pension system has to be very high for the pension system to be regressive. Higher inequalities of length of life imply that the pension system is regressive even if the Bismarckian part of the pension system is smaller. Finally, we see that if $\epsilon = 0.33$, i.e. uneducated agents live only 22% of their last period of life, or equivalently that inequalities of length of life are very high; then whatever the structure of the pension system, an increase in ν has a positive impact on $\hat{\theta}$.

Now, the question is to know if an economy can belong to this dark area. Let us consider the French case. Hairault and Langot (2008) and Disney and Johnson (2001)

[17]This figure is obtained using a simple algorithm available upon request.

show that $\lambda = 0.885$. Thanks to the previous figure, we obtain that if $\epsilon > 0.07$ then an increase in ν can have a positive impact on $\hat{\theta}$. Robert-Bobée and Cadot (2007) show that an educated agent can expect to have a life expectancy at 86 years old 20% higher than the life expectancy of uneducated agents. But they also show that inequalities of length of life decrease with age. It implies that at 60, ϵ will be strictly higher than 0.11[18]. Consequently, for the French economy, an increase in the generosity of the pension system (ν) can have a positive impact on the share of the educated population ($F(\hat{\theta})$).

7.4.2 The dynamics

Let us now consider an increase in ν and its impact on the dynamics of $\hat{\theta}_t$. To do so, we use the same calibration as in the previous subsection, but we precise that $\epsilon = 0.1$ and $\lambda = 1$.

We consider an economy which is initially at steady state with $\nu = 0.5$. Then we consider an expected increase in ν from 0.5 to 0.6 from period 3 on. The adult generation of period 1 is not affected by the change in period 3 of the replacement rate. This generation is used as a reference. The adult generation of period 2 is taxed with the initial replacement rate but benefits from the larger replacement rate when old. From period 3 on (included), the replacement rate is 0.6 for each generation.

The dynamics is obtained using equation (7.17). It can be written:

$$(1+\beta\varphi)\ln\left(\frac{W_t^e}{W_t^{ne}}\right) + \epsilon\beta\ln(W_t^{ne}) + \chi = \ln(\hat{\theta}_t/\mu) \tag{7.25}$$

The budget constraint of the Government describes a relation of the following form: $\tau_t = f(\hat{\theta}_t, \hat{\theta}_{t-1})$. $\hat{\theta}_{t-1}$ is a state variable at period t. Every other variable only depends on $\hat{\theta}_t$. Consequently, the dynamic behavior of $\hat{\theta}_t$ is obtained using the previous equation.

For $\nu = 0.5$, we have at steady state $\hat{\theta} = 4.0278$. But, for $\nu = 0.6$, we obtain $\hat{\theta} = 4.035$. This positive impact is explained by the fact that the couple (ϵ, λ)=$(0.1, 1)$ belongs to the dark region of the previous subsection. The dynamics of $\hat{\theta}_t$ between these two steady states is represented by figure 7.3[19].

There is an overshooting reaction of $\hat{\theta}$ before stabilizing to its new steady state value[20]. Indeed, the adult generation of period 2 pays only for the past replacement rate but benefit

[18]We find ϵ such that $\varphi - \epsilon = 0.8 \times \varphi$.
[19]The algorithm is available upon request. The abscissa is a time index for adult generations.
[20]We also observe a negative overshooting reaction if ν decreases.

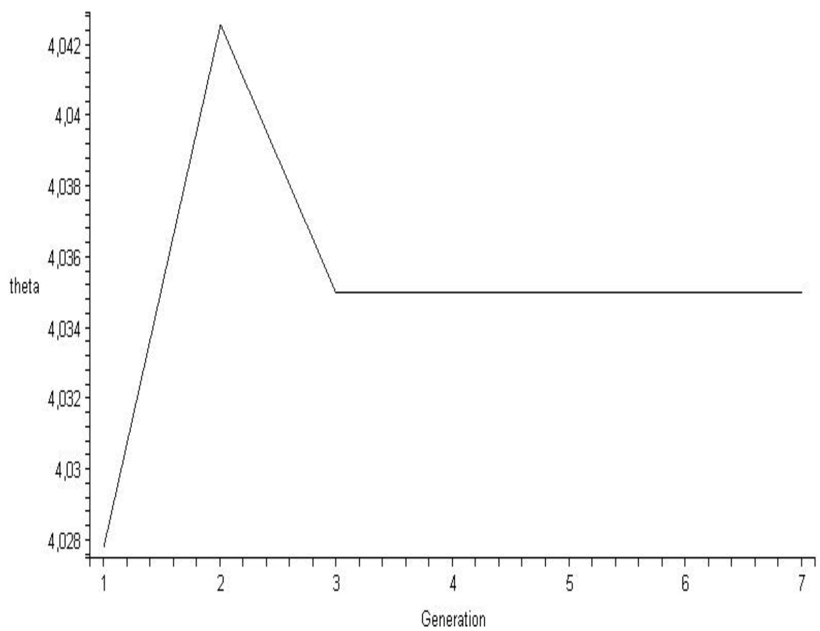

Figure 7.3: Dynamics of $\hat{\theta}_t$ if ν increases.

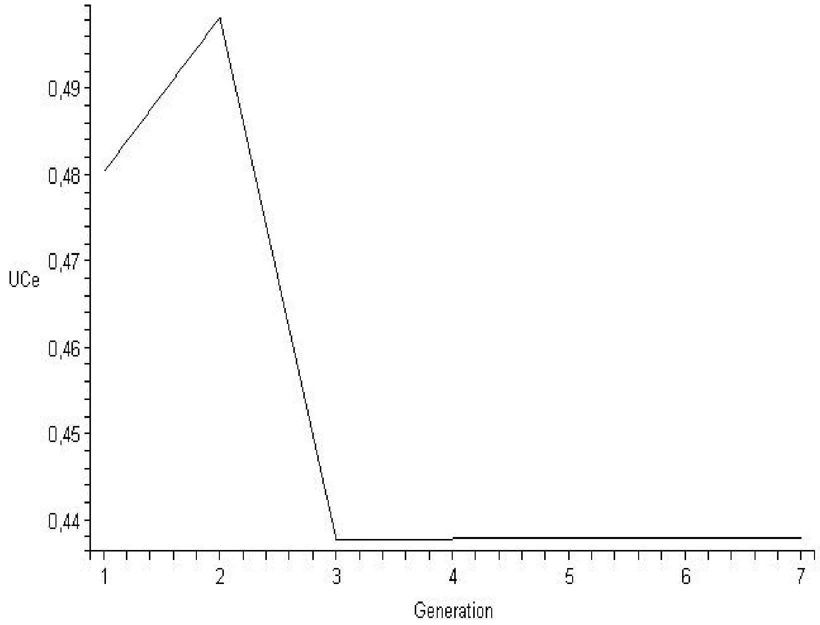

Figure 7.4: Dynamic of UC_t^e if ν increases.

from a high replacement rate in the second period of their life. Consequently, more agents want to educate themselves. For the following periods, $\hat{\theta}$ is near its new steady state value.

Figures 7.4 and 7.5 illustrate the welfare that each generation gets from their consumption, without carrying about educational costs. The time index refers to the period at which a generation is adult. It means that $U_{t-1}^i = UC_t^i - (1 - I)v(\theta)$.

The increase in the replacement rate only has a positive impact on the welfare of the adult generation of period 2. For every following generation the welfare is lower. This result comes from the fact that the social security wealth of every agent (educated or not) is negative, because of the higher return of savings compared to that of the pension system.

7.5 Concluding Remarks

Usually models assume that agents only consider the positive impact of educational choices on financial variables. In that case, with an heterogenous population, an increase in the

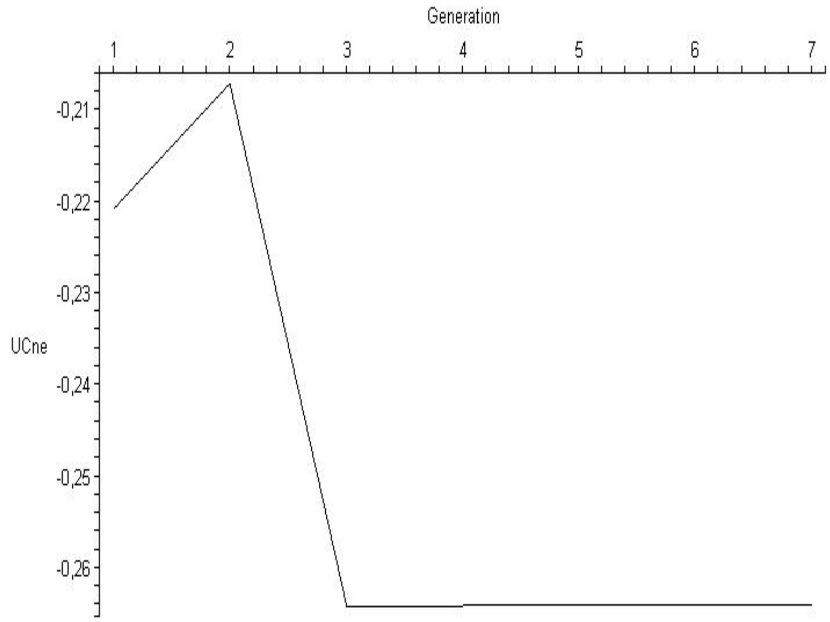

Figure 7.5: Dynamic of UC_t^{ne} if ν increases.

size of the pension system has a negative impact on the share of the educated population as in Docquier and Paddison (2003) and as in Le Garrec (2005). However, we have shown that if they also consider the benefit in terms of length of life, as in Gorski *et al.* (2007) or as in Borck (2007), we can obtain the opposite result. In this chapter, we define every situation in which such a result can appear. We show that the life expectancy differential has to be significant and that the system has to be sufficiently Bismackian. Finally, we study the dynamics of the share of the educated population for an expected increase in the generosity of the pension system. We show that an overshooting reaction can appear.

The French economy has all the characteristics to belong to the special case of this chapter. Inequalities of length of life are not only high, but they also tends to increase (Monteil and Robert-Bobée 2005). Consequently, the increase in the generosity of the pension system in France can have a positive impact on the share of the educated population, and this phenomenon is reinforced by the increased inequalities of length of life. Therefore, further empirical results have to be obtained to conclude as for the net impact of the generosity of pension systems on educational choices.

General Conclusion

Cette thèse a pour objectif de compléter la littérature économique portant sur le caractère redistributif des systèmes de retraite par répartition. Ces systèmes opèrent à la fois une redistribution inter-générationnelle, mais aussi une redistribution verticale des ressources. Notre réflexion s'articule autour (1) de l'importance de cette redistribution pour la détermination endogène des systèmes de transferts; (2) de l'impact de la redistributivité instantanée des systèmes de retraite par répartition sur l'accumulation du capital, la richesse et le bien-être des agents; (3) des différences entre la progressivité instantanée et celle de long-terme dues aux inégalités face à la mort.

Les apports de cette thèse

Dans la première partie de cette thèse, nous nous sommes attachés à étudier la détermination endogène des systèmes de transferts assurant à la fois une redistribution inter-générationnelle, mais aussi une redistribution verticale des ressources.

Le premier chapitre étudie la détermination endogène d'un système de retraite par répartition assurant uniquement une redistribution inter-générationnelle des ressources. Cette étude s'appuie à la fois sur des arguments positifs, mais aussi sur des arguments normatifs. Nous montrons ainsi que le vieillissement de la population a un impact ambigu sur le taux de taxe et sur le taux de remplacement du système de retraite.

Le deuxième chapitre cherche à analyser la dynamique des systèmes de transferts assurant une redistribution verticale pure des ressources, lorsque la structure de la population est déterminée de façon endogène. Nous montrons alors que les éléments subjectifs tels que les anticipations et les croyances, jouent un rôle aussi important que les éléments objectifs tels que les salaires, pour la détermination et la dynamique des politiques redistributives.

Le troisième chapitre montre que si l'on conçoit les systèmes de retraite par répartition comme un élément d'une politique redistributive plus globale, alors il est possible d'obtenir une dynamique en trois temps pour les politiques redistributives avec tout d'abord des transferts de faible montant; puis le montant des transferts augmente avant de se réduire. Nous montrons que les inégalités de salaire ainsi que la structure de la population sont à l'origine de cette dynamique.

Dans la seconde partie de cette thèse, nous étudions l'impact de la redistributivité instantanée sur l'accumulation du capital, la richesse et le bien-être des agents.

Dans le quatrième chapitre, nous montrons qu'un accroissement de la redistributivité instantanée peut bénéficier à tous les agents de l'économie à long-terme, si le taux de taxe est la variable d'ajustement. Cependant, il semble qu'il existe un coût de la transition vers un tel système pour la population la plus éduquée.

Le cinquième chapitre généralise les résultats du chapitre précédent et montre en plus que la redistributivité instantanée, via son impact sur l'épargne agrégée, modifie les inégalités de salaire.

Dans la troisième partie, nous étudions les différences entre la redistributivité instantanée et celle de long-terme induites par les inégalités face à la mort.

Le sixième chapitre étudie plus précisément cette distinction. Nous montrons alors que la progressivité instantanée n'assure pas nécessairement une progressivité de long-terme des ressources du fait de l'impact régressif que jouent les inégalités face à la mort.

Enfin, *le septième chapitre* exploite les résultats du chapitre précédent pour montrer que l'accroissement de la taille d'un système progressif réduit la part de la population éduquée. En revanche, si les inégalités face à la mort sont telles que le système de retraite devient régressif, alors en augmenter la générosité revient à subventionner la population éduquée, ce qui a un effet positif sur la part de cette population.

Les pistes de recherche

Les pistes de recherche sont très nombreuses c'est pourquoi nous n'en mentionnerons ici que quelques unes :

i. Concernant la première partie, d'autres éléments déterminent la demande de redistribution et notamment l'accessibilité des marchés financiers. Ainsi, la baisse des coûts d'accès aux marchés peut être un facteur explicatif de la remise en question récente des systèmes de sécurité sociale. Un autre point important concerne l'accessibilité au système de sécurité sociale lui-même. En effet, puisque les entrepreneurs n'ont pas la possibilité de se couvrir contre le risque de défaillance, alors la disponibilité d'un système de sécurité sociale va avoir un impact sur l'activité entrepreneuriale. Les agents les plus averses au risque vont alors préférer rester salarié.

ii. Concernant la deuxième partie, une extension serait d'endogénéiser d'autres variables sur lesquelles la structure des systèmes de retraite aurait un impact distorsif (offre de travail ou choix d'éducation).

iii. La principale extension de la troisième partie serait de procéder à des microsimulations sur le cas français pour savoir à quel point les inégalités face à la mort peuvent rendre le système de retraite moins progressif qu'il n'y paraît.

Bibliography

Aaron, H., 1966. "The Social Insurance Paradox". *The Canadian Journal of Economics and Political Science*, 32(3), 371-374.

Acemoglu, D., 2002. "Technical Change, Inequality, and the Labor Market". *Journal of Economic Literature*, 40, 7-72.

Adams, P., Hurd, M.D., McFadden, D., Merrill, A., Ribeiro, T., 2003. "Healthy, Wealthy and Wise ? Tests for Direct Causal Paths between Health and Socio-Economic Status". *Journal of Econometrics*, 112, 3-56.

Alesina, A., Angeletos, G-M., 2005. "Fairness and Redistribution". *American Economic Review*, 95(4), 960-980.

Alesina, A., Rodrik, D.,, 1994. "Distributive Politics and Economic Growth". *Quarterly Journal of Economics*, 109, 465-490.

Andersen, T.M., 2005. "Social Security and Longevity". *Cesifo Working Paper* 1577.

Andersen, T.M., 2008. "Increasing longevity and social security reforms-A legislative procedure approach". *Journal of Public Economics*, 92, 633-646.

Atkinson, A.B., 2000. "The Welfare State, Budgetary Pressure and Labour Market Shifts". *The Scandinavian Journal of Economics*, 102(3), 445-462.

Attanasio, O.P., Banks, J., Meghir, C., Weber, G., 1999. "Humps and Bumps in Lifetime Consumption". *Journal of Business and Economic Statistics*, 17.

Attanasio, O., Emmerson, C., 2001. "Differential Mortality in the UK". *IFS Working Paper* 01/16.

d'Autume, A., 2003. "Vieillissement et choix de l'âge de la retraite. Que peut nous dire le modèle à générations ?". *Revue Economique*, 54 (3), 561-571.

Belan, P., Pestieau, P., 1998. "Privatizing Social Security: A Critical Assessment". *The Genova Papers on Risk and Insurance*, 24(1), 114-130.

Bisin, A., Verdier, T., 2000. "A model of cultural transmission, voting and political ideology". *European Journal of Political Economy*, 16, 5-29.

Bisin, A., Verdier, T., 2002. "Public Policies and the Dynamics of Cultural Values in the Welfare State". *Annales d'économie et de statistique*, 63-64, 215-233.

Bommier, A., 2006. "Uncertain Lifetime And Intertemporal Choice: Risk Aversion As A Rationale For Time Discounting". *International Economic Review*, 47(4), 1223-1246.

Bommier, A., Magnac, T., Rapoport, B., Roger, M., 2006. "Droit à la Retraite et Mortalité Différentielle". *Economie et Prévision*, 168, 1-16.

Borck, R., 2007. "On the Choice of Public Pensions when Income and Life Expectancy are Correlated". *Journal of Public Economic Theory*, 9, 711-725.

Breyer, F., Straub, M., 1993. "Welfare effects of unfunded pension systems when labor supply is endogenous". *Journal of Public Economics*, 50(1), 77-91.

Browning, E.K., 1975. "Why the social insurance budget is too large in a democracy". *Economic Inquiry*, 13, 373-388.

Casamatta, G., Cremer, H., Pestieau, P., 2000. "The political economy of social security". *Scandinavian Journal of Economics*, 102, 503-522.

Casarico, A., Devillanova, C., 2007 (forthcoming). "Capital Skill complementarity and the Redistributive Effects of Social Security Reforms". *Journal of Public Economics*, doi:1016/j-jpubeco.2007.06.007.

Chakraborty, S., 2004. "Endogenous Lifetime and Economic Growth". *Journal of Economic Theory*, 116(1), 119-137.

Charpin, J.M., 1999. "L'avenir de nos retraites". *La documentation française.*

Cooley, T.F., Prescott, E.D., 1995. "Economic Growth and Business Cycles", in T.F. Cooley, editor, *Frontier of Business Cycle Research.* Princeton University Press.

Coronado, J.L., Fullerton, D., Glass, T., 2000. "The Progressivity of Social Security". *NBER working papers* 7520.

Cremer, H., De Donder, P., Maldonado, D., Pestieau, P., 2007 (forthcoming). "Voting over Type and Generosity of a Pension System when some Individuals are Myopic". *Journal of Public Economics*, doi:10.1016/j-jpubeco.2007.06.001.

Cremer, H., Pestieau, P., 2000. "Reforming our pension system: Is it a demographic, financial or political problem?". *European Economic Review*, 44(4-6), pp.974-983.

de la Croix, D., and Michel, P., 2002. "A theory of Economic Growth". *Cambridge University Press.*

Deaton, A. and Paxson, C., 1999. "Mortality, education, Income, and Inequality among American Cohorts". *NBER Working Paper* No. 7140.

Disney, R., Johnson, P. (eds), 2001. "Pension Systems and Retirement Incomes across OECD Countries". *Edward Elgar*, Cheltenham.

Docquier, F., Paddison, O., 2003. "Social Security Benefit Rules, Growth and Inequality". *Journal of Macroeconomics*, 25, 47-71.

Drouhin, N., 1997. "Système de retraite et accumulation du capital: un modèle à générations imbriquées avec duré de vie incertaine". *Recherches Economiques de Louvain*, 63(2), 133-151.

Drouhin, N., 2001a. "Inégalités face à la mort et systèmes de retraite". *Revue d'Economie Politique*, Numéro spécial "Epargne et Retraite".

Drouhin, N., 2001b. "Life Uncertainty and Time Preference". *Theory and Decision*, 51, 145-172.

Drouhin, N., 2006. "Uncertain lifetime, prospect theory and temporal consistency". *Working Paper.*

Duffy, J., Papageorgiou, C., Perez-Sebastian, F., 2004. "Capital-Skill Complementarity? Evidence from a panel of Countries". *Review of Economics and Statistics*, 86(1), 327-344.

Dutta, J., Kapur, S., Orszag, J.M., 2000. "A portfolio approach to the optimal funding of pensions". *Economics Letters*, 69(2), 201-206.

Esping-Andersen, G., 2007. "Les Trois Mondes de l'Etat-Providence". Paris, *PUF*, 2ème éd. française.

Fallon P.R., Layard, R., 1975. "Capital-Skill Complementarity, Income Distribution, and Output Accounting". *Journal of Political Economy*, 83, 279-302.

Feldstein, M., 1974. "Social Security, Induced Retirement, and Aggregate Capital Accumulation". *The Journal of Political Economy*, 82(5), 905-926.

Feldstein, M., 1985. "The Optimal Level of Social Security Benefits". *The Quarterly Journal of Economics*, 100(2), 303-320.

Feldstein, M., Liebman, J.B., 2002. "Social Security", in *Handbook of Public Economics*, volume 4, Edited by Auerbach, A.J. and Feldstein, M..

Galasso, V., Profeta, P., 2002. "The Political Economy of Social Security: a Survey". *European Journal of Political Economy*, 18, 1-29.

Galasso, V., Profeta, P., 2004. "Lessons for an ageing society: the political sustainability of social security systems". *Economic Policy*, 19(38), pp.63-115.

Galasso, V., Profeta, P., 2007. "How does ageing affect the welfare state?". *European Journal of Political Economy*, 23(2), 554-563.

Galor, O., Zeira, J., 1993. "Income Distribution and Macroeconomics". *The Review of Economic Studies*, 60(1), 35-52.

Gorski, M., Krieger, T., Lange, T., 2007. "Pensions, Education and Life Expectancy". *Working Paper*, ZEW DP Nr. O7/04.

Griliches, Z., 1969. "Capital-Skill Complementarity". *The review of Economics and Statistics*, 51(4), 465-468.

Groezen, B.v., Meijdam, L., Verbon, H.A.A., 2007. "Increased Pension Savings: Blessing or Curse ? Social Security Reform in a Two Sector Growth Model". *Economica*, 74(296), 736-755.

Hairault, J.O., Langot, F., 2008. "Inequality and Social Security Reforms". *Journal of Economic Dynamics and Control*, 32 (2), 386-410.

Heer, B., Maussner, A., 2005. "Dynamic General Equilibrium Modelling: Computational Methods and Applications". *Springer Berlin-Heidelberg.*

Homburg, S., 1990. "The efficiency of unfunded pension schemes". *Journal of Institutional and Theoretical Economics*, 146, 640-649.

INSEE, 1999. "Revenu et patrimoine des ménages". 28.

Juillard, M., 1996. "A Program for the Resolution and Simulation of Dynamic Models with Forward Variables through the Use of a relaxation Algorithm". *CEPREMAP Working Paper 9602.*

Krusell, P., Rios-Rull, J-V, 1996. "Vested Interests in a Positive Theory of Stagnation and Growth". *Review of Economic Studies*, 63, 301-329.

Krusell, P., Quadrini, V., Rios-Rull, J-V, 1997. "Politico-economic Equilibrium and economic growth". *Journal of Economic Dynamics and Control*, 21, 243-272.

Krusell, P., Ohanian, L.E., Rios-Rull, J-V, Violante, G.L., 2000. "Capital-Skill Complementarity and Inequality: a Macroeconomic Analysis". *Econometrica*, 68(5), 1029-1053.

Lambert, P.J., 2001. "The Distribution and Redistribution of Income". *Manchester University Press.*

Lambrecht, S., Michel, P., Vidal, J-P, 2005. "Public Pensions and Growth". *European Economic Review*, 49, 1261-1281.

Le Garrec, G., 2005. "Social Security, Inequality and Growth". OFCE, *working paper* 2005-22.

Legros, F., 1994. "Caractère Redistributif des Systèmes de Retraite". *Revue Economique*, 4, 805-817.

Liebman, J.B., 2001. "Redistribution in the Current U.S. Social Security System". *NBER working paper 8625.*

Lindert, P.H., 2004. "Growing Public: Social Spending and Economic Growth Since the Eighteenth Century". *Cambridge University Press.*

Meltzer, A., Richard, S., 1981. "A Rational Theory of the Size of Government".

Journal of Political Economy, 89, 914-927.

Merrien, F.X., 2007. "L'Etat-Providence". *Que sais-je ?*, PUF, 2007.

Mesrine, A., 1999. "Les différences de mortalité par milieu social restent fortes". *Données Sociales*, INSEE, 228-235.

Mirowski, J., Ross, C.E., 2000. "Socioeconomic Status and Subjective Life Expectancy". *Social Psychology Quarterly*, 63, 133-151.

Mitchell, O.S., Zeldes, S.P., 1996. "Social Security Privatization : A Structure for Analysis". *American Economic Review*, 88 (2), 363-367.

Monteil, C., Robert-Bobée, I., 2005. "Les différences sociales de mortalité : en augmentation chez les hommes, stables chez les femmes". *INSEE première*, No.1025.

Nelissen, J.H.M., 1999. "Mortality Differences Related to Socioeconomic Status and the Progressivity of Old-Age Pensions and Health Insurance: The Netherlands". *European Journal of Population*, 15, 77-97.

Nyce, S.A., Schieber, S.J., 2005. "The Economic Implications of Aging Societies: The Cost of Living Happily Ever After". *Cambridge University Press.*

Pannis, C., Lillard, L., 1995. "Socioeconomic Differentials in the Returns to Social Security". *RAND Working Paper*, Santa Monica, CA.

Perotti, R., 1993. "Political Equilibrium, Income Distribution and Growth". *Review of Economic Studies*, 60, 755-776.

Persson, T., Tabellini, G., 1994. "Is Inequality Harmful for Growth ?". *American Economic Review*, 84, 600-621.

Persson, T., Tabellini, G., 2000. "Political Economics: Explaining Economic Theory". *The MIT Press.*

Piketty, T., 2002. "L'économie des Inégalités". *Repères: La découverte.*

Ponthière, G., 2006. "Growth, Longevity and Public Policy". *Cesifo Working Paper* 1780.

Rawls, J., 1971. "A theory of Justice". *Harvard University Press.*

Razin, A., Sadka, E., Swagel, P., 2002a. "The Aging Population and the Size of the Welfare State". *Journal of Political Economy*, 110(4), 900-918.

Razin, A., Sadka, E., Swagel, P., 2002b. "The Wage Gap and Social Security: Theory and Evidence". *American Economic Review: Papers and Proceedings*, 92(2), 390-395.

Razin, A., Sadka E., 2004. "Privatizing Social Security: A Political Economy Approach". *IMES DP 2004-E-16.*

Rican, S. and Salem, G., 1999. "Les disparités spatiales de mortalité des adultes en âge d'activité". *Données Sociales*, INSEE, 236-242.

Robert-Bobée, I., Cadot, O., 2007. "Mortalité aux Grands Ages : encore des écarts selon le diplôme et la catégorie sociale". *INSEE première*, No.1122.

Saint-Paul, G., 1992. "Fiscal policy in an endogenous growth model". *Quarterly Journal of Economics*, 107, 1243-1260.

Sato, K., 1967. "A Two-Level Constant-Elasticity-of-Substitution Production Function". *Review of Economic Studies*, 34, 201-218.

Snower, D.J., 1993. "The Future of the Welfare State". *The Economic Journal*, 103(418), 700-717.

Sommacal, A., 2006. "Pension systems and intragenerational redistribution when labor supply is endogenous". *Oxford Economic Papers*, 58, 379-406.

Uebelmesser, S., 2004. "Political Feasibility of Pension Reforms". *Topics in Economic Analysis and Policy*, 4(1).

Villa, P., 2004. "Typologie et Equivalence des Systèmes de Retraites". *CEPII WP 2004-09.*

Yaari, M.E., 1965. "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer". *Review of Economic Studies*, 32, 137-150.